Visualizing Flixbus Travel Accessibility in Europe

Botond Papp
Anatol Botond Vass
University of Bremen
Bremen, Germany
pappbo@uni-bremen.de
anatolb1@uni-bremen.de

ABSTRACT

Understanding the accessibility and connectivity of intercity bus services is really important to improve European public transportation. In this project, we develop a visualisation to understand the accessibility of Flixbus travel in Europe using the General Transit Feed Specification (GTFS) dataset. We focus on developing visualisations showing available Flixbus routes, travel times between major cities, frequency of connections, and areas with low accessibility. From analyzing the GTFS data, it shows that service coverage is highest in Germany, home to Flixbus. The most under-provided regions are those of the Balkans and Iberia—in other words, Spain and Portugal. The coverage is excellent around Stockholm. Precomputed data and fast algorithms are necessary for large-scale GTFS visualizations; precomputed data is more effective. This will be a valuable tool for travellers, transportation planners, and policymakers to better understand and improve bus connectivity throughout Europe. There is also related work on urban transit in the literature, for example, Prommaharaj et al.'s "Visualizing Transit Systems" (2020) [1] and Ma et al.'s "Visualization of Spatio-temporal Data of Bus Trips" (2012) [2]. However, what sets our project apart is its continental scale.

KEYWORDS

Data Visualisation, Public Transportation, GTFS, Flixbus, Accessibility Analysis

ACM Reference Format:

Botond Papp and Anatol Botond Vass. 2024. Visualizing Flixbus Travel Accessibility in Europe. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Understanding accessibility to travel is essential in the quest to improve public transportation systems. Flixbus, one of the largest companies operating in the domain of intercity bus travel throughout Europe, has extensive route networks connecting a significant number of cities.

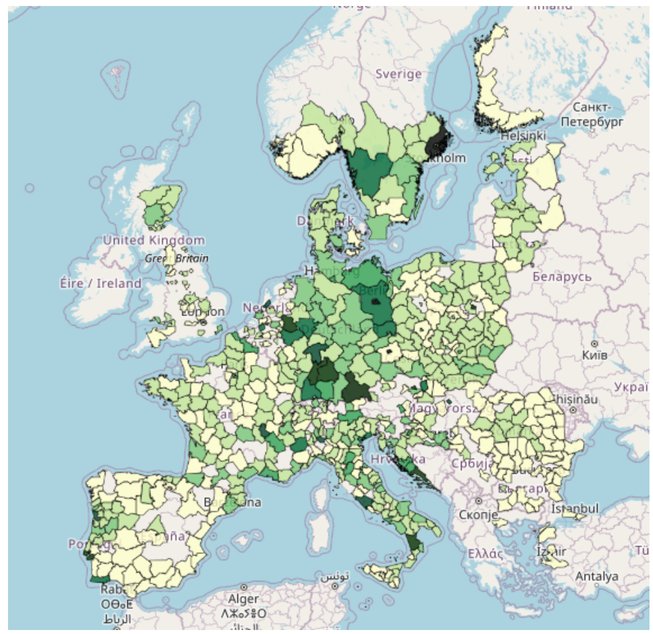

Figure 1: *Heatmap of FlixBus availability in different regions of Europe*

Public transport is one of the most significant elements in urban and regional development, as it provides a way to reduce mobility by moving people. Consequently, visualization of these links and identification of areas that have not been well serviced will aid in better travel planning and policy formulation, hence supporting infrastructure improvements and service planning. However, even though Flixbus has such a well-established network, coverage across various European regions is highly uneven. In some areas, the connectivity provided is good. In contrast, in others, it is not as good, decreasing accessibility and making the overall bus transit experience cumbersome for both locals and travellers. More importantly, the current Flixbus map interface lacks critical filters, and on the map itself, comprehensive travel data do not appear.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
Conference'17, July 2017, Washington, DC, USA
© 2024 Association for Computing Machinery.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

It only portrays destinations and prices, but not the interchange stations and the time it could take to travel. Most of the existing visualisation tools and research work tend to operate in and elaborate regional or local public transportation in cities, where the scope of analysis in large dimensions of intercity connectivity as a continent is neglected. We have designed a project using Flixbus GTFS data to demonstrate travel accessibility through Flixbus travels across Europe. Therefore, our primary attention was focused on travel times between major cities, the frequency of connecting lines, and the location of low-accessible areas. We hope to make this tool very informative and helpful for many users, from everyday travellers to transportation planners and policymakers, using advanced data processing and visualisation techniques.

The main findings of our project point to the existence of pronounced regional disparities in the coverage of the Flixbus service. The country with the most significant number of stops is Germany. It is the land where Flixbus has its headquarters. The worst serviced areas are the Balkans, Spain and Portugal. Stockholm presents a very high number of services. In addition, our results indicate that for the representation of large-scale GTFS visualisations to work efficiently, quick algorithms or precomputed data are needed, the latter more successful in the analyses that we carried out. The key contributions of this work are (1) a comprehensive visualisation tool showing the accessibility of FlixBus travel across Europe, including visibility of regional disparities and connectivity; (2) a methodological framework for the use of GTFS data for the analysis and visualization of intercity bus networks on a continental scale; and (3) empirical findings into the most and least served regions in Europe, presenting essential data for transportation planners and policymakers in their efforts to improve public transport services. In addition, our tool overcomes the weaknesses of the current Flixbus map interface with the following: It adds filters and complete travel information—such as how many changes are needed in a travel plan, to enhance usability and usefulness for all kinds of users.

2 BACKGROUND

Understanding and improving public transportation systems requires practical tools for visualising accessibility and connectivity of travel. Various existing projects and tools, each focussing on different aspects of transit data visualisation and trip planning, give valuable insight into this field.

OpenTripPlanner (OTP) [4] is a popular multi-modal trip planning tool using GTFS data for transfer optimisation. OTP is written in Java and is flexible with complex routing queries, integrating different modes of transport such as bike sharing, car sharing, and subways. Its advanced algorithm is very efficient in handling multiple transfers, so this will be a good reference point for similar projects. The project used Python to make it easily accessible and user-friendly.

The other important visualisation tool is **ChronoTrains** [3], which gives the user an idea about the travel distances in Europe. It will provide an intuitive view of how far you can go by train within a predetermined period from any station. The information will help highlight the accessibility of trains and their connectivity while pointing out some areas in which public transport planning could be improved.

In the bus transit area, visible works are on visualizing spatio-temporal bus trip data. Some of such projects have provided comprehensive visualizations to support the understanding of dynamics in bus travel patterns and consequently to enable better transit planning and operations (Visualization of Spatio-temporal Data of Bus Trips [2]).

An example of this would be the work by Prommaharaj et al [1]. on the visualisation of transit systems, where they graphically described the transit data for analysing public transportation efficiency and accessibility. This work only emphasizes how critical visual tools are in garnering insights about transit systems that might not be obvious solely from raw data (Visualizing Transit Systems).

2.1 Route Planning Algorithms

Practical route planning algorithms are critical in optimising public transport systems and achieving a better user experience. These algorithms have to process large data sets to determine the most efficient paths between the origin and destination points, considering several things like time of travel, options for changeovers or transfers, and service schedules.

OTP is a powerful Java-based tool for route planning. It links different means of transport such as buses, trains, bike sharing, and car sharing for more integrated trip planning. The basic functioning of the OTP is based on data retrieved from the General Transit Feed Specification (GTFS), which includes rich data on transit routes, stops, schedules, and locations. The OTP algorithm uses graph traversal techniques with the representation of a transit network using stops as nodes and routes as edges.

An extension of Dijkstra's algorithm optimized with heuristics, known as the A* search algorithm, is used to find the shortest path through the graph, which brings an improvement in search efficiency.

Heuristics and Optimizations: Heuristics used in the A* algorithm at OTP focus on some paths, reducing the computational overhead and increasing speed in the route calculation process. This makes OTP ideal for real-time trip planning applications where the quickness of response is a must.

Transfers and Intermodal Integration: OTP would be capable of handling multiple transfers and integrating different modes of transport with the help of detailed scheduling information within the GTFS data. This enables synchronization of connections of varying transit services, a property rendering it versatile in handling many complex routing queries common in urban and regional transit planning.

API and Extensibility: OTP comes packed with a rich API that allows developers to include the functionality of planning trips in other applications. The API thus supports a wide range of queries, from straightforward point-to-point routes to very complex itineraries involving many transfers. Being extensible, OTP can lend enhanced utility across different applications and platforms.

Our work builds upon the methodologies used within OTP, but the scope is limited: intercity bus travel with GTFS data from Flixbus. Differently from OTP's multi-modal idea, our approach towards route planning provides a way to maintain the uniqueness of this type of long-distance bus travel—where it concerns lesser, more crucial transfers and central city-to-city travel time.

Efficiency and Precomputed Data: Our algorithm uses precomputed travel times and routes for between-city travel. That contrasts the real-time route computation in OTP, which allows compelling visualizations and easy access. This remarkably lowers the computational load at user interaction and makes our tool much more responsive, hence user-friendly, for long-distance bus travel.

Adapting OTP's methods: Concretely, our project uses techniques for graph traversal and the concept of using GTFS data for route planning as seen in OTP. In adopting these approaches, we ensure that our route calculations become accurate and efficient. The heuristics used by the A* search algorithm, optimized in OTP, have inspired us on our path to reducing computational overhead while doing better on the efficiency of our search. Thus, whereas OpenTripPlanner is rich and general in its ability to find and display routes in urban and regional travel, our algorithm carves a specialized tool for visualizing intercity bus travel. With a focus on the user having that one particular use, the Flixbus traveler, our project will take the precomputed data to come up with an efficient, cost-effective, and excellent solution to look at.

2.2 GTFS Data Visualisation

Adequate visualization of GTFS data will be vital to understanding public transport systems and their respective improvements. Using raw transit information, visualisation tools provide insight while giving visibility into the patterns that help make informed decisions.

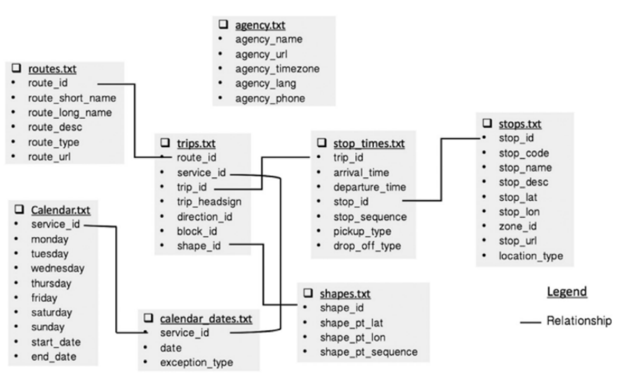

Figure 2: *GTFS Data Structure from Visualizing public transit system operation with GTFS data: A case study of Calgary, Canada [1]*

The Visualisation of Spatiotemporal Data of Bus Trips, by Kwan-Liu Ma et al. [2], provides users with much more detailed visualisations to represent dynamism in travel patterns over a network using buses. The research provides a way in which spatio-temporal data should be analyzed and presented; in this case, how bus services vary both temporally and spatially. The project applied interactive maps to illustrate how the bus routes and timetables were changing throughout the day to pinpoint times for peak travel and areas with poor service. The methods in that work informed our approach to processing the GTFS data—primarily to identify salient patterns and service frequencies essential for building a coherent dataset for visualization.

According to Prommaharaj et al., this study, Visualising Transit Systems, focusses on the use of different techniques. It approaches visualisation of transit data for the analysis of public transport efficiency and accessibility. The most important contribution this study makes is the emphasis put on the need for various visualization tools to provide insights concerning transit system performance. Developed visual tools have enabled the user to realize complex transit networks by explicitly presenting routes, stops, and service frequencies. This study will influence the way we display Flixbus routes and stops by using visual tools such as heat maps or interactive maps that show interconnections and travel times between different cities. Additionally, it informed our approach to presenting data on service frequency and regional connectivity through statistical charts. Using a combination of the techniques, our project presents a pool of strengths from previous research to generate robust and user-friendly visualisations of the accessibility of Flixbus travel. This improves the detailed understanding of the features that express inter-city bus connectivity, providing strong support for making decisions in the field of travel for the traveller and transport planner.

2.3 Visualising Time-Related Data on Maps

Visualising time-related data on maps is crucial to understanding temporal patterns in public transport. This type of visualisation helps identify trends and variations in service availability and travel times, providing valuable information to plan and optimise transit systems.

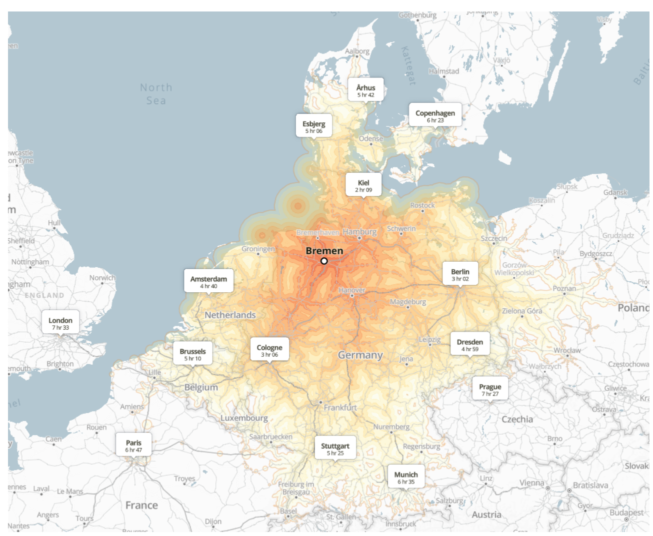

Figure 3: *Map from Chronotrains website showing reachable destinations from Bremen within 8 hours [3]*

ChronoTrains is a notable tool that focuses on visualizing train travel times across Europe. It provides an intuitive map-based interface showing how far one can travel by train within a specified time from any given station. ChronoTrains uses a heatmap-like visualisation where areas reachable within certain time frames are shaded differently, allowing users to see the extent of the train network's reach over time. This method effectively highlights accessibility and connectivity, helping users understand the temporal dynamics of train travel.

Inspired by ChronoTrains, our project visualizes destinations reachable from a starting city within specified travel time limits. However, instead of using a heatmap, we employ colour-coded markers to represent different travel times. Each marker on our interactive map is colour-coded based on the travel time required to reach the destination, providing a clear visual distinction between closer and further destinations.

This approach allows users to quickly grasp the reachability of various destinations according to travel time. Colour-coded markers make it easy to identify which areas are accessible within the user's specified parameters, enhancing the overall understanding of intercity bus connectivity. This visualization technique provides valuable insights for both travelers and transportation planners, facilitating better travel planning and service optimization.

By integrating these visualization techniques, our project leverages the strengths of previous research to create a robust and user-friendly tool for visualizing Flixbus travel accessibility. This approach enhances the understanding of intercity bus connectivity and supports better decision-making for travelers and transportation planners.

3 METHODOLOGY

The methodology of this project outlines the systematic approach taken to develop an efficient and user-friendly tool for visualizing and planning intercity bus travel across Europe. This section details the processes involved in data collection, preprocessing, analysis, and the implementation of algorithms and visualizations. The code of the final interface is available here: https://gitlab.informatik.uni-bremen.de/anatolb1/ds_2024

3.1 Goals and Objectives

The project includes two primary components: a heatmap for visualizing the connectivity and reachability of European cities and an interactive trip planning interface. The heatmap analyses and highlights underserved regions, offering critical insights to improve service coverage and accessibility. This visualization tool helps transportation planners and policymakers identify and address network gaps. The interactive interface allows users to select a starting city, a time interval, the number of allowable transfers, and the time of day they wish to travel. Based on these parameters, the system dynamically displays all reachable destinations, improving the trip planning experience and helping users understand their travel options comprehensively.

3.2 Principles Underlying the Design

Our project design emphasizes user-friendliness and operational efficiency. We achieved user-friendliness by prioritising intuitive navigation, minimising the learning curve, and ensuring accessibility for users with varying levels of technical proficiency. Clear visual cues, consistent layouts, and interactive elements that respond promptly to user inputs enhance the overall usability of the tool.

Operational efficiency ensures the system performs complex computations quickly and accurately while maintaining a responsive user interface. Given the large dataset and the need for real-time updates, we optimized algorithms for speed and accuracy and implemented effective data caching mechanisms to reduce redundant computations.

3.3 Data Collection

Flixbus authorities provided us with access to the official Flixbus GTFS dataset, specifically the dataset available on TransitLand [5]. This dataset contains data from June 3, 2024, to December 2, 2024. We obtained this dataset by directly contacting Flixbus and requesting the data specifically for this project. In addition to the GTFS dataset, we used the Geonames All Cities with a Population > 1000 [6] dataset, which contains population data for European cities with populations more than 1,000 inhabitants. This dataset is publicly available and can be accessed for detailed demographic information.

3.4 Dataset Structure

The GTFS dataset comprises several files, all in .txt format, including: agency.txt, calendar.txt, calendar_dates.txt, feed_info.txt, routes.txt, shapes.txt, stops.txt, stop_times.txt, transfers.txt, and trips.txt. For our project, we primarily used the following files: stops.txt, stop_times.txt, routes.txt, and trips.txt.

The agency.txt file contains information about the transit agencies that provide the data. The calendar.txt file defines the service dates and days of the week that the transit service is available. The calendar_dates.txt file lists exceptions to the regular service defined in the calendar.txt, such as holidays or special events. The feed_info.txt file provides additional information about the feed itself, such as the publisher, version, and expiration date. The routes.txt file describes the routes that the transit agencies operate. The shapes.txt file defines the shape of the transit routes in terms of sequences of latitude and longitude coordinates.

The stops.txt file lists all the stops where transit vehicles pick up or drop off passengers, providing essential details such as stop ID, stop name, latitude, and longitude. The stop_times.txt file lists the times that a vehicle arrives at and departs from individual stops for each trip, including fields such as trip ID, arrival time, departure time, stop ID, and stop sequence. The transfers.txt file defines the rules for transferring between routes. The trips.txt file defines the trips for each route, where a trip is a sequence of two or more stops, and includes fields such as route ID, service ID, trip ID, trip headsign, and direction ID.

trip_id	stop_id	arrival_time	departure_time	timepoint	stop_sequence	stop_headsign	route_short_name	pickup_type	drop_off_type	shape_dist_traveled
N504-2-2005092024-TO#PRG-00	dcbfe731-9603-11e6-9066-549f350fcb0c	08:05:00 PM	08:05:00 PM		1			0	0	
N504-2-2005092024-TO#PRG-00	dcc02ea8-9603-11e6-9066-549f350fcb0c	10:35:00 PM	10:40:00 PM		2			0	0	
N504-2-2005092024-TO#PRG-00	dcc03216-9603-11e6-9066-549f350fcb0c	26:15:00	26:20:00		3			0	0	
N504-2-2005092024-TO#PRG-00	f4a4521a-4808-45fa-a382-4acfbc27bc4a	27:55:00	28:00:00		4			0	0	

Figure 4: *Structure of stop_times.txt*

stop_id	stop_name	stop_lat	stop_lon	stop_code	stop_desc	zone_id	stop_url	location_type	parent_station	wheelchair_boarding	stop_timezone	platform_code
00308c11-2e6c-44e5-916f-f6286ebbb339	Kamenmost	43.429554	17.172177	WKM							Europe/Zagreb	
0038c230-7c73-4b50-a567-aa447870b6ac	Wolpertshausen (Hohenlohe)	49.170225	9.860071	WOL							Europe/Berlin	
00b5a645-29bd-4aa4-b4dc-99c606ec31a8	Val Vomano	42.607785	13.754997	VVM							Europe/Rome	
014f2224-3982-4329-b188-6f1a50f9d5fe	Ostrava (central bus station)	49.830387	18.280333	OSA							Europe/Prague	

Figure 5: *Structure of stops.txt*

route_id	trip_id	service_id	trip_short_name	trip_headsign	route_short_name	direction_id	block_id	shape_id	wheelchair_accessible	trip_bikes_allowed	bikes_allowed	ticketing_trip_id	ticketing_type
1	001-18-0850062024-M#FR-00	001-18-0850062024-M#FR				0		b68d58fc818a218eb1dc2ecdc3be9044					1
1	001-18-0850072024-M#FR-00	001-18-0850072024-M#FR				0		b68d58fc818a218eb1dc2ecdc3be9044					1
1	001-18-0850082024-M#FR-00	001-18-0850082024-M#FR				0		b68d58fc818a218eb1dc2ecdc3be9044					1
1	001-18-0850092024-M#FR-00	001-18-0850092024-M#FR				0		b68d58fc818a218eb1dc2ecdc3be9044					1

Figure 6: *Structure of trips.txt*

agency_id	route_id	route_short_name	route_long_name	route_type	route_desc	route_url	route_color	route_text_color	bikes_allowed	route_sort_order
FLIXBUS-eu	N862	FlixBus N862	Prague - Pula	3			73D700	FFFFFF		
FLIXBUS-eu	858	FlixBus 858	Prague - Chomutov - Most	3			73D700	FFFFFF		
FLIXBUS-eu	507	FlixBus 507	Geneva - Milan - Bologna	3			73D700	FFFFFF		
FLIXBUS-eu	563	FlixBus 563	Rome - Sorrento	3			73D700	FFFFFF		

Figure 7: *Structure of routes.txt*

3.5 Data Preprocessing

Initially, we eliminated rows with missing and inconsistent data from the stop_times.txt and stops.txt files by filtering for NaN values. In the stop_times.txt file, we normalized arrival and departure times using a function that converts the time into days, hours, minutes, and seconds to handle instances where times exceeded the 24-hour format. The following code snippet is used for this normalization:

```
def normalize_time(t):
    if pd.isna(t):
        return t
    h, m, s = map(int, t.split(':'))
  return pd.Timedelta(hours=h % 24, minutes=m, seconds=s)
        + pd.Timedelta(days=h // 24)
```

To reduce the number of stops, we first eliminated multiple stops within the same city. We extracted the city name using regex functions to remove words that are not part of the city name, such as 'Bus Station', 'Airport', 'FlixBus stop', 'Central Station', 'Station', 'East', 'West', 'North', 'South', etc. We then merged the extracted city names with the "Geonames All Cities with a Population > 1000" dataset to assign population figures to each city. Cities with populations less than 50,000 were filtered out, ensuring that our map remained clear without an excessive number of stops.

We consistently handled missing values appearing in critical columns of the stop_times.txt and stops.txt files by continuously filtering out NaN values. For data processing, we used the pandas and numpy libraries available in Python.

3.6 Data Analysis

Our data analysis and implementation followed a systematic approach to ensure comprehensive and accurate results, including data cleaning, normalization, and the implementation of a Dijkstra-like shortest path algorithm to optimize route planning and visualization.

We began by extracting stop information from the stops.txt and stop_times.txt files. These files provided the necessary details such as stop ID, stop name, latitude, longitude, arrival time, and departure time. Data integration involved merging the stop_times dataframe with the stops dataframe using the pandas merge function. This integration facilitated easier data access during analysis by combining relevant information from different GTFS files.

To implement the Dijkstra-like shortest path algorithm, we used a priority queue (min-heap) with the heapq library. The algorithm steps are as follows: First, we identified the stop IDs for the starting city and initialized a priority queue with these stops, each having zero travel time and zero transfers. The core data structures included dictionaries to track the shortest travel times and the number of transfers for each stop. We then loaded the necessary GTFS files (stops.txt, stop_times.txt, routes.txt, and trips.txt) into pandas DataFrames, extracting and normalizing the time data to ensure consistency, particularly for times exceeding 24 hours.

Using the priority queue, we managed the processing of stops. The algorithm iteratively processed each stop, checking conditions for transfers and travel times. For each stop, it calculated the travel time to the next stop and updated the priority queue if a shorter path was found. The priority queue ensured that the most promising (shortest time) paths were prioritized, reducing unnecessary computations. We calculated transfer times by adding waiting times between transfers to the cumulative travel time, providing realistic travel time estimates, including both transit and layover times. Filters were applied based on specified time intervals to restrict trips to desired time frames, allowing users to query trips within specific periods and enhancing the tool's flexibility and usability. Finally, we aggregated all reachable stops into a dataframe for the final output, providing a consolidated view of all destinations that could be reached within the specified constraints.

To determine the frequency of bus connections, we loaded the stop times data from the stop_times.txt file. We calculated the number of trips servicing each stop by counting occurrences of each stop ID in the stop times DataFrame and merged this information with the stops DataFrame to associate each stop with its corresponding frequency.

For spatial analysis, we performed a spatial join between the stops with their frequencies and geographic regions using the NUTS regions dataset. After calculating the trip frequencies and merging with the stops DataFrame, we converted this data to a GeoDataFrame using geopandas. We then performed a spatial join between the NUTS regions and the converted stops GeoDataFrame to aggregate trip frequencies by geographic regions. This process allowed us to analyze the spatial coverage and visualize the frequency of bus connections in different regions.

To enhance efficiency and ensure fast response times, especially for repeated queries, we implemented a caching mechanism. This mechanism stores previously computed results, allowing quick retrieval without re-computation. The precomputing and caching were run in Google Colab due to the significant computational resources required, which would have taken too much time on our hardware.

We utilized several advanced visualization techniques and tools to present data and route planning results effectively. Folium enabled us to create interactive maps, allowing users to explore bus connections and their spatial distribution. Streamlit facilitated the rapid development of data-driven web interfaces, integrating input fields, buttons, and visual elements with minimal effort.

3.7 Visualization

We used several key tools to create interactive and informative visual representations for the project. Folium was crucial in generating both the heatmap and the interactive map used in the trip planner. Its capability to handle geographic data and interactive features such as zooming and panning made it an ideal choice. The heatmap visualizes the frequency of bus connections in different regions, providing insight into how well various areas are serviced by FlixBus. The interactive trip planner map allows users to input their starting city, time limit, maximum number of transfers, and optional time intervals. It displays all reachable destinations based on the selected parameters, with markers indicating reachable stops and different colours representing travel times.

Streamlit was used to build the interactive web application interface. This framework facilitated rapid development and deployment, allowing us to create a user-friendly interface where users can input their travel preferences and receive immediate visual feedback. The interface integrates input fields, buttons, and visual elements, providing an efficient and intuitive user experience.

The user interface features several components to enhance usability and functionality. At the top, there are input fields to enter the starting city, selecting the maximum travel hours (ranging from 1 to 8), the maximum number of changes (ranging from 0 to 3), and the desired travel time interval. These intervals include all day, early morning (00:00 - 06:00), morning (06:00 - 10:00), midday (10:00 - 14:00), afternoon (14:00 - 18:00), late afternoon (18:00 - 21:00), and night (21:00 - 00:00). Once the user inputs these parameters and clicks the "Find Trips" button, the interface updates to display the results.

The interactive map, generated with Folium, displays all reachable destinations from the selected starting city, colour-coded based on the travel time. The color gradient ranges from yellow to red, representing increasing travel times. Clicking on a marker reveals detailed trip information, including the destination, travel time, number of transfers, and start and arrival times. Additionally, a legend is included to help users interpret the color-coding of travel times.

Next to the map, the "List of Trips" section provides a searchable list of all reachable destinations. Users can type a search term to filter the list of trips, making it easy to find specific destinations. Clicking on a destination in the list highlights the corresponding marker on the map and displays the trip details in the "Selected Trip Details" box on the right.

The "Selected Trip Details" box shows detailed information about the selected trip, such as the stop name, departure time, arrival time, travel time, and number of changes. This section is updated as the user interacts with the map or the list of trips, ensuring that the most relevant information is always displayed.

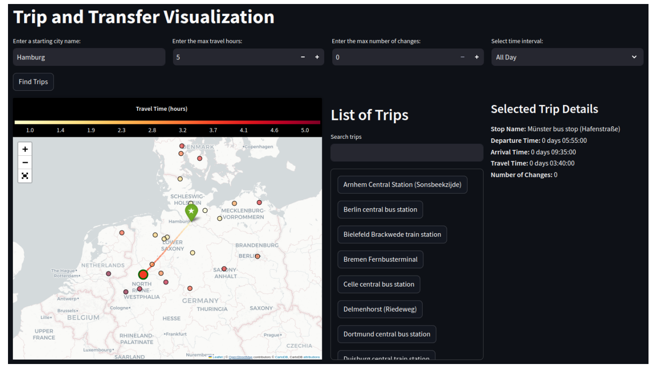

Figure 8: *Interactive interface for trip planning. [7]*

3.8 Key Components

The interface includes a page selection feature in the top left corner, allowing users to switch between the heatmap and the trip planner. On the trip planner page, users initially see the input fields to select their starting city, maximum travel time, number of allowable transfers, and desired time of day. Upon entering these parameters and initiating a search, the interface updates to display the interactive map on the left and the list of reachable destinations in the middle. Detailed trip information appears only after a user selects a destination, either from the map or the list. This progressive disclosure keeps the interface clean and focused, presenting only the most relevant information at each stage of user interaction.

3.9 Appearance and Function

We designed the tool with a focus on simplicity and user-friendliness, avoiding unnecessary complexities. The interface facilitates easy navigation through several features. Clear visual cues guide users through the process, from selecting input parameters to viewing results. The layout and intuitive design elements, such as drop-down menus and symbols to increase and decrease travel parameters, ensure that users can easily understand and interact with the tool.

3.10 Usage Scenarios

The heatmap serves primarily Flixbus to identify underserved regions and optimize service coverage. However, other transportation companies can also use it to expand their services to these underserved areas. By highlighting regions with limited connectivity, the tool provides actionable insights that help these companies make informed decisions about where to introduce new routes or improve existing services.

In addition, the destination finder targets everyday travelers or groups planning trips, helping them discover easily reachable destinations based on their preferences and input parameters. This feature is particularly useful for users looking to optimize their travel plans by considering factors such as travel time, number of transfers, and preferred travel times of the day.

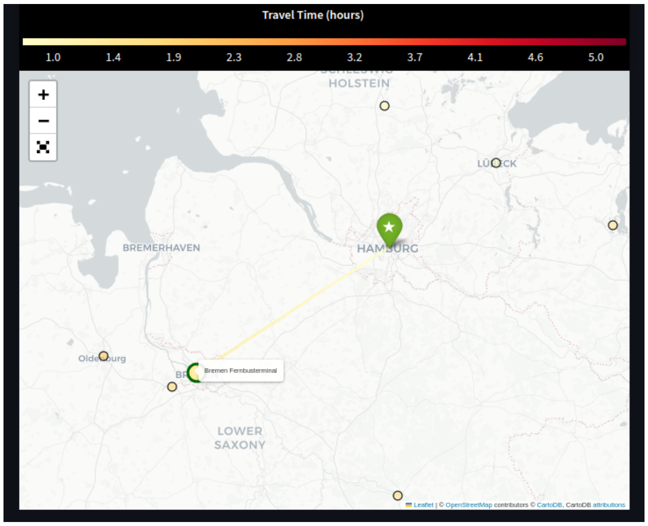

Figure 9: *Map showing a connection from Hamburg to Bremen.*

3.11 Limitations

Despite robust implementation, several limitations were identified throughout the project. One primary limitation is the potential inaccuracy of travel time estimates due to variations in bus schedules, traffic conditions, and seasonal changes, which were not accounted for in the current model. The data used for this project were static and did not include real-time updates, potentially affecting the relevance and precision of the results over time.

Another limitation concerns the granularity of the geographic data. While we focused on major cities and filtered out smaller stops, the representation may still include stops that do not perfectly align with user expectations or needs, especially in less densely populated areas. The filtering threshold of cities with populations over 50,000 might exclude some relevant stops, impacting the comprehensiveness of the coverage.

4 RESULTS

The results we achieved from analysing accessibility to Flixbus travel across Europe showed great regional differences. Germany is highly accessible, while Spain and the Balkans are underserved. Surprisingly, the service frequency in Stockholm was higher, while the number of routes was lower. These findings also pointed to where improvements can be made in the future on both accessibility and connectivity, exactly the conclusions that the other studies came to. The methodology was efficient, with potential for further development both in respect of improving the accuracy and regarding visualization.

4.1 Travel Time Analysis

The travel time analysis revealed huge differences in terms of the number of destinations that can be reached when travelling from major European cities. For example, a large share of these destinations reachable within 8 hours with one transfer was available from Berlin, which is very well connected by Flixbus. Under the same conditions, the number of reachable destinations decreased in Budapest, while in cities such as Barcelona and Bucharest, there were far fewer reachable destinations than from Berlin and Budapest. This underlines the underservicing of Spain and the Balkan regions and specifies that better connectivity is needed in these areas.

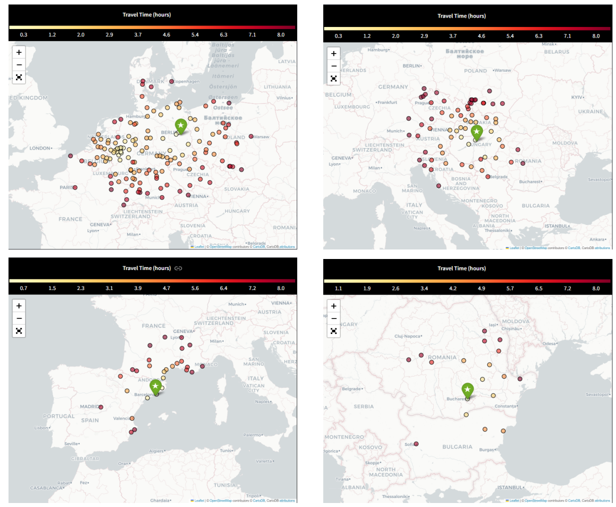

Figure 10: *Destinations reachable from Berlin, Budapest, Barcelona, and Bucharest*

4.2 Spatial Coverage Analysis

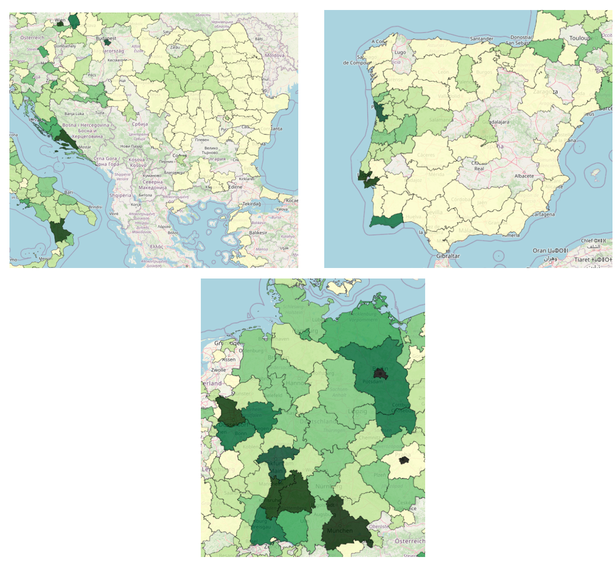

Figure 11: *Heatmap of Balkan region, Spain and Germany*

Region-wise disparities have been discovered through our spatial coverage analysis in the provision of Flixbus services throughout Europe. High-connectivity regions such as Germany, Central Europe (including Czech Republic and Poland), and Italy show high-level spatial coverage, especially considering that Germany is the headquarters of Flixbus. On the contrary, the Balkans, Spain, and Portugal, along with parts of Southern Europe, show a very thin coverage with very few bus stops. This discrepancy highlights a strategic necessity in building improved connectivity for those underserved regions, to make transportation services available in an equitable manner.

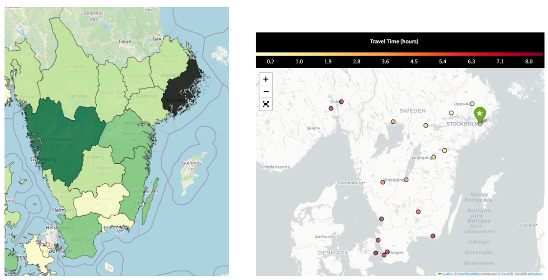

Figure 12: *Heatmap of Stockholm, and the reachable destinations from Stockholm under 8 hours*

In the Stockholm area an anomaly was detected. Whereas the Stockholm area had the best service frequency in whole Europe, an interactive mapping applying the same travel margins (8 hours travel and one transfer) revealed a few accessible destinations. This is not something predictable at all since Stockholm has only 10 routes, while Berlin has 65 and still the latter one has shown a less frequency. This does not have any apparent reason, which calls for reasons once again to dig deep and discover as to why service frequency will be high in Stockholm area.

4.3 Comparison to Related Work

Figure 13: *Chronotrains map from Germany, Spain and the Balkan region*

When comparing to the ChronoTrains [3] map, we did observe a similar pattern in the service coverage of both the Flixbus and the train networks. Both maps indicated that Germany had the most extensive coverage in 8 hours, while Spain and the Balkan region had significantly less coverage. This clearly cross-validates our results and evidently brings regional disparities as a consistent scene in public transportation under different modes of travel.

4.4 Reflection on Results

The results largely met our expectations based on prior knowledge and hypotheses. The high connectivity in Germany and Central Europe aligns with known patterns of public transportation infrastructure in these regions. However, the high frequency of service in the Stockholm area was surprising and warrants further investigation. The findings underscore the critical need for strategic improvements in bus service coverage, particularly in underserved regions like the Balkans and Southern Europe.

4.5 Usefulness of Methodology

The used methodologies were good for handling large datasets and helped to extract meaningful insights. Precomputed travel times and route frequencies were particularly useful, and visualisation tools allowed for clear, intuitive presentations of data.

4.6 Lessons Learned

If the study was to be repeated, incorporating real-time data feeds would improve the precision of travel time and frequency analyses. Using more advanced computational resources could allow for more granular analysis, including temporal variations and detailed regional studies.

4.7 Future Work

Future research should prioritise the integration of real-time data to enhance the accuracy of travel time and frequency analyses. In addition, improvements are necessary in the visualisation aspect, particularly in the display of routes. Currently, we are unable to visualise the routes on our map due to the complexity of integrating this functionality into the precomputing script. Moreover, the shapes.txt file, which contains the route points, is one of the largest in our dataset at 290 megabytes, posing significant challenges in data handling and processing. Addressing these issues will be crucial for creating a more comprehensive and user-friendly tool.

5 CONCLUSION

This project presents a unique approach in visualizing accessibility to travel with Flixbus across Europe by means of GTFS data, providing essential insight for the improvement of public transportation. Our tool emphasises the regional disparities of high access in Germany and Stockholm and lower service levels in the Balkans, Spain, and Portugal. This visualization supports trip planning by travelers and is useful for transportation planners and policymakers in finding service gaps.

We have therefore noted some key variations in the intercity bus service coverage that requires the precomputed data and efficient algorithms needed to render large-scale visualizations. We expect this tool to serve audiences ranging from travelers interested in route optimization to policymakers looking to enhance access to transportation.

Our contributions include a methodological framework for future research in transportation visualization, promoting a more connected and efficient European transportation network. Ultimately, this work supports better decision-making and aims for a future with improved public transportation accessibility and efficiency.

6 TASK DISTRIBUTION

Botond Papp	Anatol Botond Vass
Load and clean GTFS data	Design and implement the interactive interface using Streamlit
Normalise time fields	Integrate input fields, buttons, and visual elements for trip planning and visualisation
Remove multiple stops within the same city	Develop interface for both destinations and heatmap
Filter small cities based on population	Create interactive maps for visualising routes and destinations using Folium
Implement core logic to calculate travel times with direct connections and transfers	Develop a heat map to reflect connectivity and underserved regions
Optimise travel time calculations	Optimising the precomputing scipt for Colab and cleaning the data set after running the script
Implement filtering for trips based on the time of day and maximum travel hours	Write Background and Conclusion sections
Write Methodology section	Write Results sections
Write Keywords and Task Distribution sections	Write Abstract with Botond
Write Abstract with Anatol	Write Introduction with Botond
Write Introduction with Anatol	

Table 1: Contributions by Team Members

REFERENCES

[1] P. Prommaharaj, K. Becker, M. Seiler, and A. Lamprecht, *Visualizing Transit Systems*, Heliyon, 2020. Available: https://senseable.mit.edu/papers/pdf/20200320_Prommaharaj-etal_VisualizingTransit_Helyon.pdf.

[2] K.-L. Ma, J. Zheng, and B. Ma, *Visualization of Spatio-temporal Data of Bus Trips*, 2012. Available: https://www.researchgate.net/publication/234164026_Visualization_of_Spatio-temporal_Data_of_Bus_Trips.

[3] ChronoTrains, *ChronoTrains: How far can you travel by train in Europe in 8 hours?*, Available: https://www.chronotrains.com/en.

[4] OpenTripPlanner, *OpenTripPlanner*, Available: https://github.com/opentripplanner/OpenTripPlanner.

[5] Flixbus GTFS Feed, Available: https://www.transit.land/feeds/f-u-flixbus/versions/fb9ccd0db123d3915c72200ab34ddd9177988d8c.

[6] Geonames, *All Cities with a Population of 1000 or more*, Available: https://public.opendatasoft.com/explore/dataset/geonames-all-cities-with-a-population-1000/table/?disjunctive.cou_name_en&sort=name.

[7] A. Botond Vass, *Data Science Project Repository*, Available: https://gitlab.informatik.uni-bremen.de/anatolb1/ds_2024.