Technische Universität München
TUM School of Computation, Information and Technology

The Rationality of Irrationality: A Software Architecture for Multi-Dimensional Navigation

Jens Klinker

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz:
Prof. Dr. Jörg Ott

Prüfende der Dissertation:
1. Prof. Dr. Stephan Jonas
2. Prof. Dr. Holger Regenbrech
3. apl. Prof. Dr. Georg Groh

Die Dissertation wurde am 14.03.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 23.07.2024 angenommen.

Preface

My greatest appreciation goes to my academic colleagues and personal contacts. This dissertation journey would have been insurmountable without the support from numerous individuals in both spheres. I wish to take a moment to extend my sincere thanks to all who have assisted me in reaching this milestone.

First, I would like to thank my Ph.D. advisor, Prof. Dr Stephan Jonas, who provided me with the freedom, support, guidance, and environment to conduct my research. Your approach to leading our research group could not have fit my needs more closely. I also want to express my personal gratitude to my Mentor, Dr. Mariana Avezum Mercer. Be it as a Co-Author or in personal feedback sessions. She would always help me find and take the next step to achieve my goals. I also want to thank Prof. Dr. Jörg Ott, Christine Lissner, Helma Schneider, and Ruth Demmel for their support in any kind of bureaucratic matter.

My special thanks goes to the students who participated in the development of TrafSim, either as student employees or as part of their Theses. To mention a few, I thank Nikolai Madlener, Benjamin Pabst, Berke Esmer, Olena Galitska, Joe Yu, Kaethe Vrettos, Mohamed Hechem Selmi, Julian Ostarek, and Bernhard Kreminski. Though not involved in the development of TrafSim, I also want to thank Felix Moser for his support as a student employee, helping me to refine game theoretical questions surrounding my research. Similarly, I thank my colleagues Maximilian Kapsecker, Lara Reimer, Marko Jovanovic, Jonas Schulte-Coerne, Leon Nissen, Florian Schweizer, and Elene Mille for the feedback and the gained insights from our discussions at our offices in Munich.

Last, I want to thank my family and wife, Sophia, for their unconditional support and honest feedback throughout my entire Ph.D. Your help was truly appreciated.

Abstract

Problem: In the context of traffic networks, determining the perfect path for a single traveler is a daunting task. Navigating a human necessitates an understanding of the individual's short-term and long-term aspirations intertwined with the current state of the world around them. This translates into a multi-dimensional utility function composed of an unpredictable number of relevant variables. Financial costs, travel comfort, route scenery, stress, or required physical activity are just a few variables influencing our travel behavior. To complicate this issue, travelers weigh these variables subjectively, further individualizing the equation. This means that each traveler's understanding of a "rational" travel behavior is, in part, catered to their personal preferences. Depending on the extent to which travelers are able to empathize with another traveler's preferences, they will find rationality in that behavior or not. When the results of all these individual travel preferences then collide in a network with limited capacities, the limited supply to meet the compound traveler demands results in traffic congestion.

Description: To efficiently solve this issue, most navigation systems simplify the problem by focusing on universally significant variables. Navigation algorithms like Google Maps or Apple Maps mostly optimize routes for travel duration and distance. However, pursuing low-dimensional optimization assumes that human routing preferences follow similarly low-dimensional utility functions. In reality, human utility functions are high-dimensional, including the traveler's current situation, general personality, and physicality. This causes low-dimensional routing algorithms to underfit the multi-dimensional potential of human travel preferences. Underfitting human travel preferences causes conflicts that could have been avoided in higher dimensionality. Similar to how adding lanes to roads (2nd dimension) or building bridges instead of crossroads (3rd dimension) reduces conflicts, adding more optimization dimensions to routing algorithms also reduces conflicts.

Results: Delving into this issue, we propose a novel concept demonstrating how including human "irrationality" in routing benefits all network participants. The dissertation details the development and implementation of a software architecture that supports the transition from Navigation 2.0 (current) to Navigation 3.0 (future). A primary feature of this work is the construction of a simulation environment that compares the behaviors of Navigation 2.0 and Navigation 3.0. The development of an algorithm for Navigation 3.0, capable of integrating user-specific needs into the routing process and manipulating network supply, along with adjustments in physical, psychological, and situational optimization variables, is a key contribution.

The simulation demonstrates how individual travel needs affect travel behavior within a network. Utilizing an Origin-Destination Matrix (ODM), the simulation provides a detailed representation of travel behavior for both individual travelers and network-wide patterns. The research highlights the effects of psychological, physiological, and situational changes on travel patterns, as well as the influence of structural or regulatory modifications in traffic flow.

Another focus of the work is on the effects of implementing navigation algorithms based on the newly developed principles for Navigation 3.0. This includes the impact of network adjustments (edges/nodes) and their influence on various travel needs. The research aims to ensure that the enhancements from these customizations are greater than their associated costs. The dissertation offers insights into balancing the maximization of benefits against the interaction and computational expenses, aiming to prevent a shift from underfitting to overfitting in navigation algorithms.

Zusammenfassung

Problem: Im Kontext von Verkehrsnetzen ist die Bestimmung des perfekten Weges für einen einzelnen Reisenden eine gewaltige Aufgabe. Die Navigation eines Menschen erfordert ein Verständnis der kurz- und langfristigen Ziele des Einzelnen, die mit dem aktuellen Zustand der Welt um ihn herum verwoben sind. Daraus ergibt sich eine mehrdimensionale Nutzenfunktion, die sich aus einer unvorhersehbaren Anzahl von relevanten Variablen zusammensetzt. Finanzielle Kosten, Reisekomfort, landschaftliche Gegebenheiten, Stress oder erforderliche körperliche Aktivität sind nur einige Beispiele für Variablen, die unser Reiseverhalten beeinflussen. Hinzu kommt, dass Reisende diese Variablen subjektiv gewichten, was die Gleichung weiter individualisiert. Das bedeutet, dass das Verständnis eines jeden Reisenden von einem "rationalen" Reiseverhalten zum Teil auf seine persönlichen Präferenzen abgestimmt ist. Je nachdem, inwieweit ein Reisender in der Lage ist, sich in die Präferenzen eines anderen Reisenden hineinzuversetzen, wird er dieses Verhalten als rational empfinden - oder eben nicht. Wenn die Ergebnisse all dieser individuellen Reisepräferenzen in einem Netz mit begrenzten Kapazitäten aufeinandertreffen, führt das begrenzte Angebot zu Verzögerungen.

Beschreibung: Um dieses komplexe Problem efficient zu lösen, vereinfachen die meisten Navigationssysteme das Problem, indem sie sich auf die allgemeingültigsten Variablen zu konzentrieren. Navigationsalgorithmen wie Google Maps oder Apple Maps optimieren Routen meist nach Reisedauer und Entfernung. Eine niedrigdimensionale Optimierung setzt jedoch voraus, dass die menschlichen Routingpräferenzen ähnlich niedrigdimensionalen Nutzenfunktionen folgen. In Wirklichkeit sind die menschlichen Nutzenfunktionen hochdimensional, einschließlich der aktuellen Situation des Reisenden, seiner allgemeinen Persönlichkeit und seiner körperlichen Verfassung. Dies führt dazu, dass niedrigdimensionale Routing- Algorithmen das mehrdimensionale Potenzial der menschlichen Reisepräferenzen nicht ausreichend berücksichtigen. Die unzureichende Berücksichtigung dieser Präferenzen führt zu Konflikten, die bei höherer Dimensionalität hätten vermieden werden können. Ähnlich wie das Hinzufügen von Fahrspuren zu Straßen (2. Dimension) oder der Bau von Brücken anstelle von Kreuzungen (3. Dimension) Konflikte reduziert, werden Konflikte ebenfalls durch das Hinzufügen zusätzlicher Optimierungsdimensionen in Routenfindungsalgorithmen reduziert.

Ergebnisse: Um dieses Thema zu vertiefen, schlagen wir ein neuartiges Konzept vor, das zeigt, wie die Einbeziehung menschlicher "Irrationalität" in das Routing für alle Netzwerkteilnehmer von Vorteil ist. Die Dissertation beschreibt die Entwicklung und Implementierung einer Softwarearchitektur, die den Übergang von Navigation 2.0 (aktuell) zu Navigation 3.0 (zukünftig) unterstützt. Ein wesentliches Merkmal dieser Arbeit

ist der Aufbau einer Simulationsumgebung, die das Verhalten von Navigation 2.0 und Navigation 3.0 vergleicht. Die Entwicklung eines Algorithmus für die Navigation 3.0, der in der Lage ist, benutzerspezifische Bedürfnisse in den Routingprozess zu integrieren und das Netzangebot zu manipulieren, sowie Anpassungen der physischen, psychologischen und situativen Optimierungsvariablen, ist ein wesentlicher Beitrag.

Die Simulation zeigt, wie sich individuelle Reisebedürfnisse auf das Reiseverhalten in einem Netzwerkt auswirken. Unter Verwendung einer Start-Ziel-Matrix (ODM) bietet die Simulation eine detaillierte Darstellung des Reiseverhaltens sowohl für individuelle Reisende als auch für netzweite Muster. Die Untersuchung zeigt die Auswirkungen psychologischer, physiologischer und situativer Veränderungen auf das Reiseverhalten sowie den Einfluss struktureller oder regulatorischer Veränderungen im Verkehrsfluss.

Ein weiterer Schwerpunkt der Arbeit liegt auf den Auswirkungen der Einführung von Navigationsalgorithmen, welche auf den neu entwickelten Prinzipien für Navigation 3.0 aufbauen. Insbesondere auf den Auswirkungen von Netzwerkanpassungen (Kanten/-Knoten) und deren einfluss auf unterschiedliche Reisebedürfnisse. Die Forschung zielt darauf ab, sicherzustellen, dass die Verbesserungen durch diese Anpassungen größer sind als die damit verbundenen Kosten. Die Dissertation bietet Einblicke in die Abwägung zwischen Nutzenmaximierung und Interaktions- und Rechenaufwand, um eine Verschiebung von Underfitting zu Overfitting bei Navigationsalgorithmen zu verhindern.

Contents

Preface iii

Abstract iv

Zusammenfassung vi

Contents viii

List of Figures x

List of Tables xii

1 Introduction 1
- 1.1 Societal Interpretation of Behavioral Variety . . . 1
- 1.2 Exploration of Behavioral Variety . . . 3
- 1.3 Behavioral Variety in Traffic Networks . . . 4

2 Principles of Navigation 6
- 2.1 Determinants of Human Behavior: Past Actions, Emotions, Norms, and Personality . . . 6
- 2.2 Supply, Demand, and Situation . . . 7
- 2.3 The Evolution of Navigation . . . 10
  - 2.3.1 Game Theoretical Concepts . . . 10
  - 2.3.2 Navigation 1.0 . . . 11
  - 2.3.3 Navigation 2.0 . . . 13
- 2.4 Navigation 2.0: An Underfitting Solution . . . 15
- 2.5 Homogeneity of Navigation 2.0 . . . 20
  - 2.5.1 From Water's Behavior to "The Law of the Fastest Route" . . . 21
  - 2.5.2 Game Theoretical Issues . . . 22
  - 2.5.3 Supply- and Demand Distortion . . . 22
  - 2.5.4 Evolution of Navigation Strategies . . . 23
  - 2.5.5 Traffic Planning Methods . . . 26
- 2.6 Research Approach . . . 27
- 2.7 Research Questions . . . 29

3 Methodology 30
- 3.1 The Traffic Simulation Tool "TrafSim" . . . 31
  - 3.1.1 A Model for Navigation Theory . . . 32

3.1.2 An Architecture for a Preference-Oriented Navigation Simulation . 35
3.1.3 The Network Layer 36
3.1.4 The O-D Layer 40
3.1.5 The Constraint Layer 42
3.1.6 The Routing Layer 43
3.1.7 The Traffic Simulation WebApp 46

4 Content 49
4.1 Publications 49
4.1.1 Publication 1, None-Core-Publication: A Practical Prototype . . . 49
4.1.2 Publication 2, Core-Publication: An Improvement in the Realism of the Data 51
4.1.3 Publication 3, Core-Publication: An Approach for Further Validation 54

5 Summary of Results 57
5.1 Concise Answers 61

6 Discussion 62
6.1 Contribution to Research 62
6.2 Limitations, Problems, and Assumptions 64

7 Outlook 65
7.1 Technical Improvements 65
7.2 Digital Representation 66
7.3 Closing Thoughts 68

Bibliography 69

8 Appendix: Full-Text Publications 74
8.1 Full-Text Version of Publication 1 74
8.2 Full-Text Version of Publication 2 85
8.3 Full-Text Version of Publication 3 96

List of Figures

1.1 Visual depiction of "delusional sensing" and "shifted", using W.E. Hill's cartoon (My Wife and My Mother-In-Law, 1915, right) to demonstrate delusional sensing and an extended version of a normal distribution[1] (left). The cartoon displays both an old and a young woman, causing observers to question each other's sanity as they see opposing versions of reality. On the left, our extension of the normal distribution showcases how the likelihood of the majority of the population (μ) to consider a person or group to be "shifted", or in other words crazy, increases with their distance ($+\sigma$ or $-\sigma$) from the population's mean (red curve). 3

1.2 Symbiotic relationships in nature, showcasing advantages of diverting strategies. Water Buffalo and Oxpecker Bird (parasite control/food source), Bee and Flower (food source/reproduction), Clown Fish and Sea Anemone (protection/cleaning), Shark and Remora Fish (food source/cleaning). [2345] 4

2.1 Left: An example for supply information in Munich[6], graphically visualized as a 2 dimensional map. Right: Demand information in Munich[7], visualized as expected travel speeds (Dark Green: 80 km/h, Bright Green: 61-80 km/h, Yellow: 41-60 km/h, Orange: 21-40 km/h, Red: 0-20 km/h) 8

2.2 An example for situational weather influences on cyclists' travel behavior, published by Ahmed et al. [1]. 9

2.3 Navigation 1.0: Sextants[8], maps[9], and individual knowledge[10] as typical representatives of navigation prior to the invention of satellite-connected devices (GPS and mobile internet) 12

2.4 Apple Maps[11] as a typical representation of modern navigation applications representing Navigation 2.0. 14

2.5 Study data from the Master's thesis of Bayram Ahmadov, showcasing the distribution of planned in buffer time depending on the travel purpose [2]. 17

2.6 Study data from the Master's thesis of Bayram Ahmadov, showcasing the distribution of travel behavior considering more or less affecting factors [2]. 18

2.7 Overfit (Navigation 1.0), underfit (Navigation 2.0), and optimal fit (Navigation 3.0) as an exemplary parallel to flaws in Navigation 1.0 and 2.0 and setting a goal for navigation 3.0.[12] 19

3.1 The iterative process of the scientific method, visualized by Vestin [3]. . . 30

3.2 Generalized concept of navigation, including human, automated, and distributed navigation. 33

3.3 Overview of the different subsystems and support layers used by the Traffic Simulation WebApp. 35
3.4 Graphical Model visualizing data sources and data formats used to build the graph file in the Network Layer. 37
3.5 Graphical user interface for selecting a GTFS and OSM source. In combination with a router configuration, a network graph is built. 38
3.6 Example snippet of the GTFS-Visualizer, along with the edit functionality. Due to the large number of entries, this section of WebApp is difficult to display in the format constraints of this dissertation. 39
3.7 Model visualizing the concept behind the O-D Layer, describing the used data sources and format. 40
3.8 GUI for selecting an O-D Matrix, and a source graph, together with individual routing parameters called "constraints" (e.g. walking speed, biking speed, maximum walking distance, ...). 41
3.9 Snippet from the Constraint Manipulation Tool user interface, showcasing editable constraints in a policy termed "Client Application". 42
3.10 Model illustrating the concept behind the Constraint Layer, outlining the user interface and data storage. 43
3.11 Model illustrating the architecture of the Routing Layer, inclusive of the utilized routing library and its integration with data from other layers. . . 44
3.12 Screenshot of the Comparison View provided in the WebApp. 45
3.13 Screenshot of the Traffic Visualizer View provided in the WebApp. 46
3.14 Generalized concept of navigation, encompassing human, automated, and distributed navigation. 47

5.1 Commuter satisfaction model by St-Louis et al. [4] 57
5.2 Compares the results of traveling with avoidance feature (right) and without avoidance feature (left). The above figures use isochrones to visualize the impact on travel time, while the below figures compare heat maps to visualize the impact on station utilization. 58
5.3 Documentation and onboarding page for setting up the Navigation 3.0 oriented Traffic Simulator TrafSim. 59
5.4 Confusion matrix showing the prediction precision in the diagonal (sum of column = 100). The respective recall values are, Car: 73%, Walk: 100%, Bike: 95%, Train: 85%, Bus: 96%, Tram: 94%. 60

List of Tables

4.1 Publication 1: Published. 49
4.2 Publication 2: Published. 51
4.3 Publication 3: Published. 54

1 Introduction

We often assume that others think rationally. Especially when traveling, we assume that everyone's goal is to reach their destination as quickly as possible. But each person's reasoning can take a unique path. Financial costs, travel comfort, route scenery, stress, or required physical activity are just a few examples of other influences on our travel behavior. This means that our understanding of rational travel behaviors is, in part, catered to our preferences. Depending on the extent to which we are able to empathize with another traveler's preferences, we will find rationality in that behavior or not. This means that we are left uncertain of the motives and reasons for somebody's travel behavior unless they match ours or we are able to empathize with their preferences.

But how do we humans cope with people whose behaviors do not fit into our own range of rational behaviors? How do we classify a person who prefers to ride their bike 30 kilometers to work every day instead of using public transport or a car?

1.1 Societal Interpretation of Behavioral Variety

Once somebody else behaves so differently from our own expectations that we can no longer identify the reasoning behind it, we tend to label their actions as irrational, crazy, or erratic. In German, terms for crazy, like "verrückt" and "wahnsinning" are used to describe such behaviors, framing them as irrational. However, the literal translation of these words provides insight into how craziness was initially depicted in the German language. "Verrückt" literally means shifted, explaining a typification of irrationality in which a shown behavior is "shifted" outside the normative range of behaviors.

But how is a normative behavior defined? There are two perspectives: the individual's and the population's. For any individual, the range of rational behaviors is defined from their own perspective. When we, as individuals, classify behaviors as rational or irrational, we consider our own experiences, knowledge, physical ability, and values in this process. The more closely the exerted behavior aligns with our own perspective, the more likely it is that we see it as rational. If a chosen strategy diverts from our personal perspective, we use a skill called empathy to switch to the other person's perspective. Using this skill, we can extend our range of acceptable behaviors. Here, the rule is: The further another person's experiences, knowledge, physical ability, and values differ from our own, the more empathy we must utilize to take their perspective. In return, the more closely we are aligned with people's perspectives around us, the less empathy is required.

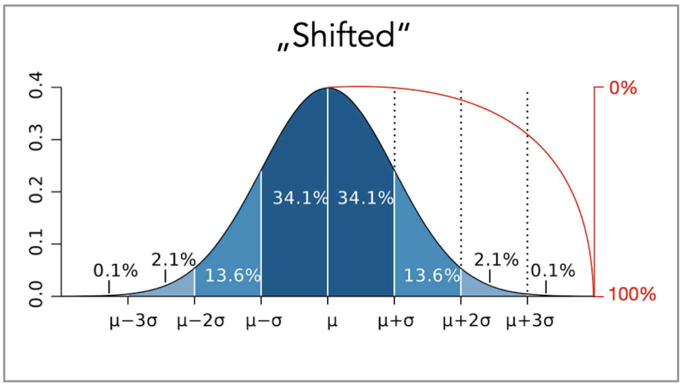
From the perspective of an entire population, these individual perspectives are summarized and normally distributed. In such normal distributions, the peak describes the most commonly held perspective of a population. Further strafing to the left and right, we find less and less popular opinions on rational behavior. This means that from the perspective of an entire population, the ends of the distribution are where irrationality is found. Similarly to before, more and more empathy is required for the center of the distribution to find rationality in the behavior of the distribution's ends.

This relation is depicted in Figure 1.1 (left), symbolically showing how the likelihood for the majority in the center (μ) to classify an individual as irrational increases with the individual's distance ($+\sigma$ or $-\sigma$) to the center (μ) of the normal distribution, as more and more empathy is required. What defines the previously mentioned normative perspective can be described as the areas close to the center (μ) of the normally distributed perspectives of the relevant population. Similarly, an individual placed at the edge of a normal distribution (based on experiences, knowledge, physical ability, and values) will require much empathy to understand the perspective of the seemingly shifted center.

The term "wahnsinning" provides a different interpretation, assuming false information instead of false interpretation. When translated literally, "wahnsinning" can be expressed as "delusional sensing". Here, the German language takes a different route, implying that it's not the person's knowledge, background, or values that are causing irrational behavior but their false perception of facts. This perspective on irrationality does not target a difference in interpretation but rather suggests a difference in the initial recording of reality using our biological sensors of vision, hearing, smell, touch, or taste. The assumption here is that the individual making a judgment is operating on false data and, therefore, never had the chance to make a rational choice. It thereby assumes that the person is not generally shifted but they are merely misinformed. People operating under this assumption of irrationality will try to inform the other person, relying on their own perception of reality. However, there is potential for both perceptions to be correct.

To further illustrate the possibility of both individuals sensing correctly, consider the image created by W.E. Hill in 1915 (Figure 1.1, right). Depending on one's perspective, the image could depict either an old or a young woman. The picture hosts two similar truths that are difficult to perceive simultaneously, forcing us to take one perspective at a time. This demonstrates how our senses, here vision, can be deceived, reporting different versions of truth depending on our perspectives.

[1] https://en.wikipedia.org/wiki/Standard_deviation Accessed: 15.11.2023



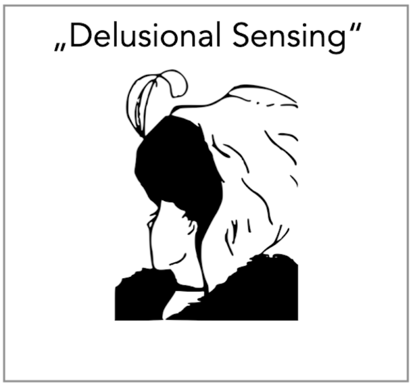


Figure 1.1: Visual depiction of "delusional sensing" and "shifted", using W.E. Hill's cartoon (My Wife and My Mother-In-Law, 1915, right) to demonstrate delusional sensing and an extended version of a normal distribution[1] (left). The cartoon displays both an old and a young woman, causing observers to question each other's sanity as they see opposing versions of reality. On the left, our extension of the normal distribution showcases how the likelihood of the majority of the population (μ) to consider a person or group to be "shifted", or in other words crazy, increases with their distance (+σ or -σ) from the population's mean (red curve).

1.2 Exploration of Behavioral Variety

While differences in truth generally cause conflicts between individuals, they can also cause the precise opposite and prevent conflicts. In eco-systems, different species can benefit from shifted strategies, either by directly helping them, a concept called symbiosis [5, 6] (Figure 1.2) or by targeting a different food source, called coexistence. Symbiotic relationships mark a trade between two entities in which both entities have a positive return from engaging in the trade. So, while a Water Buffalo sees flies as parasites, it does not mind the Oxpecker Bird eating them off its back. The Oxpecker Bird, in return, sees them as a food source. While an Oxpecker eating flies might appear irrational to a Water Buffalo, it actively benefits from the bird's behavior. The second concept is coexistence, in which two entities do not interfere with the other's strategy, causing no conflicts between them [7]. Both concepts advertise a heterogeneity of strategies to spread the demand evenly across a provided supply and maximize the total number of animals occupying the ecosystem.

Both symbiosis and coexistence allow for a wider range of different entities to exist without causing additional conflicts [8, 9, 10]. Bascompte et al. even describes that when habitats are initially homogeneous, the homogeneous population will eventually become scattered over different strategies in a heterogeneous pattern, causing patchy distributions, each focused on their own supply niches [10]. In their book Animal Conflict, Huntingtonford describes how conflicts between animals only revolve around either resources (e.g., food) or outcomes (e.g., killing or survival of prey) [11]. One way to alleviate conflict is, for example, achieved through the utilization of uncontested food sources, leading to an increase in ecosystem heterogeneity. A logical consequence of such heterogeneity is a normal distribution of preferences, demonstrating the rationality of irrationality.

Figure 1.2: Symbiotic relationships in nature, showcasing advantages of diverting strategies. Water Buffalo and Oxpecker Bird (parasite control/food source), Bee and Flower (food source/reproduction), Clown Fish and Sea Anemone (protection/cleaning), Shark and Remora Fish (food source/cleaning). [2][3][4][5]

1.3 Behavioral Variety in Traffic Networks

From a game theoretical perspective, traffic networks show similarities to ecosystems. Both are seen as non-cooperative n-person games. Such games provide a range of supplies (e.g., network paths or food sources) that are fought over by a range of different demands (e.g., travelers or animals). When traveling from an origin to a destination, travelers fight over the existing supply resources, trying to find a subjectively optimal path. In homogeneous traffic systems, all travelers share the same definition of optimal, similar to an ecosystem consisting of just a singular animal species, forcing them to fight over the same path or food source, leading to a supply shortage and unused

[2]https://www.scuba.com/blog/5-marine-symbiotic-relationships/ Accessed: 15.11.2023
[3]https://animalsymbiosis.weebly.com/cattle-egrets.html Accessed: 15.11.2023
[4]https://sharktourshawaii.com/blog/shark-remora-fish-unique-relationship/ Accessed: 15.11.2023
[5]https://www.earth.com/news/how-do-symbiotic-relationships-evolve-between-species/ Accessed: 15.11.2023

supply sources. In traffic networks, we call these supply shortages congestion. Heterogeneous systems contain a variety of traveler types or animal species, each with different preferences for different resources. Here, resource preferences are distributed between travelers or species, allowing for larger populations before resource limits are met. In traffic networks, this means that a larger distribution of traveler preferences results in a larger total amount of travelers traveling from origin to destination before congestions emerge.

In this dissertation, we argue that reasons for unusual travel choices that might seem irrational to others are actually beneficial as long as they utilize otherwise vacant supplies. Similar to the Water Buffalo not minding the Oxpecker Birds' "irrational" hunger for the flies on its back, a car commuter using the highway will not mind his neighbor's "irrational" preference to walk to work, consequently leaving more space on the highway for the commuter.

Yet, technological advances in traveler navigation have caused a growing homogeneity of traveler strategies, mostly optimized for the fastest or shortest routes. Instead of optimizing an individual's routing preferences, modern navigation solutions like Google Maps or Apple Maps focus on singular optimization variables for routing, causing an undesirable centralization.

Building upon these principles, this dissertation will demonstrate how heterogeneously distributed strategies for public transportation may be perceived as irrational on an individual level but are rational on a systemic level, spreading the demand more evenly over the available network supply. A new approach to navigation will be introduced through a simulation tool using a new type of navigation algorithm optimized for individualized routing. This, in turn, possesses the potential to alleviate congestion on the more conventionally logical and efficient routes (fastest and shortest) by allowing travelers to opt for personalized routes.

2 Principles of Navigation

2.1 Determinants of Human Behavior: Past Actions, Emotions, Norms, and Personality

In the previous chapter, we mentioned how technological advances in traveler navigation contribute to an increase in traveler homogeneity, thereby fostering resource conflicts. Before discussing how technology caused this effect, we first introduce the basic principles constituting routing decisions. Aarts et al. have delved into the prediction of human behavior based on past actions, emphasizing whether these predictions result from repeated decision-making or if they stem from established habits [12]. This work aligns closely with the research conducted by Ouellette and Wood, where they explored the multi-faceted processes through which past behavior influences future actions, highlighting the significant role that habits play [13]. Furthermore, Ronis et al. stress the importance of attitudes, decisions, and habits in determining recurring behavior [14].

Parker et al. expanded the theory of planned behavior by introducing the role of personal norms. Their study adds depth to our understanding of how personal norms play an integral role in determining behavior [15]. Similarly, Biddle et al. discuss the influence of norms, preferences, and identities on retention decisions, emphasizing how these aspects impact decision-making processes [16]. Charng et al. delve deeper into the role of identity, specifically role identity, and its importance in predicting repeated behavior [17]. Richard et al. looked into the influence of anticipated emotional reactions on the prevention of AIDS, shedding light on how emotions play a pivotal role in decision-making, especially when it comes to vital decisions concerning health [18]. This is somewhat connected to Tessler and Schwartz's exploration of help-seeking behaviors in relation to self-esteem and achievement motivation, discussing how the anticipation of societal judgments can influence an individual's actions [19].

Roccas et al. investigated the interplay between the big five personality factors and personal values. Their research provides insights into individual personalities and how these can influence and be influenced by personal values, thus playing a crucial role in shaping behavior [20]. In a related context, Cobb-Clark and Schurer delved into the stability of the big five personality traits, offering an understanding of the consistency and changes in these traits over time [21].

2.2 Supply, Demand, and Situation

In any given network, travelers can choose from many routes to reach their destination. Each person ranks these paths using their subjective "utility function" that considers a potentially endless list of factors such as travel time, comfort, or emissions. To pick the best path, we have identified three clusters or types of information that travelers base their decisions on:

Concept 1: Supply Information
Supply Information provides insights into the routes, roads, and public transportation options available within the transportation network, representing the theoretical availability. Additionally, weights on edges and nodes are provided, indicating distances and expected travel times without demand. To do so, speed limits, traffic light sequences, and public transport operation schedules are used to assume an "unobstructed" travel time for actions performed (e.g., traveling an edge, waiting at a node) on the network.

Background: Supply information, as the backbone of understanding transportation infrastructure, has been widely studied in transportation research. Ortuzar and Willumsen (2011) offer comprehensive insights into the planning and modeling of transportation systems. Their work showcases the importance of understanding the routes, roads, and available public transportation options within any given network, underscoring the essence of supply information [22]. Similarly, Rodrigue et al. (2016) delve deeper into the spatial structures and patterns of transportation systems, discussing the various components that form the transportation supply [23]. Hall and Tewdwr-Jones (2019) emphasize how the supply of transportation infrastructure, for example, the rail networks or ports, directly influences the socio-economic trajectory of regions, emphasizing its critical role in urban planning and development [24]. Winters et al. [25] highlight how built environment influences can sway the route selection for both bicycles and cars, indicating the complexity of choices. Borst et al. [26] emphasized how environmental street characteristics can significantly impact the walking routes chosen by the elderly. Meanwhile, Moran et al. [27] investigated the role of trip destinations and neighborhood attributes in shaping the route choices of children, emphasizing the importance of understanding diverse environmental influences.

Concept 2: Demand Information
In addition to this theoretical availability of supply information, demand information offers details about the current traffic situation. This information is based on data reflecting the current distribution of travelers across the network. From this, one can infer the anticipated travel speed on all network sections based on the current balance of the traffic network's supply and demand.

Background: Winters et al. also addressed demand information, stating that traffic demand is the cumulative outcome of countless intersecting paths chosen by travelers, navigating the networks that connect their current location to their preferenced endpoint

[25]. Mahmassani elaborates that transportation system analysis, especially in urban areas, needs accurate demand data to anticipate travel speeds and network congestions [28]. Similarly, Ben-Akiva and Lerman highlight the significance of capturing real-time demand information for traveler's route choice behavior modeling [29]. They contend that the accuracy of travel demand models improves with the incorporation of real-time traffic information. Furthermore, the work by Nagurney and Dong supports Concept 2 by emphasizing the importance of understanding the current distribution of travelers (or demand information) and how this can influence traffic dynamics and route choice behaviors in a network [30].

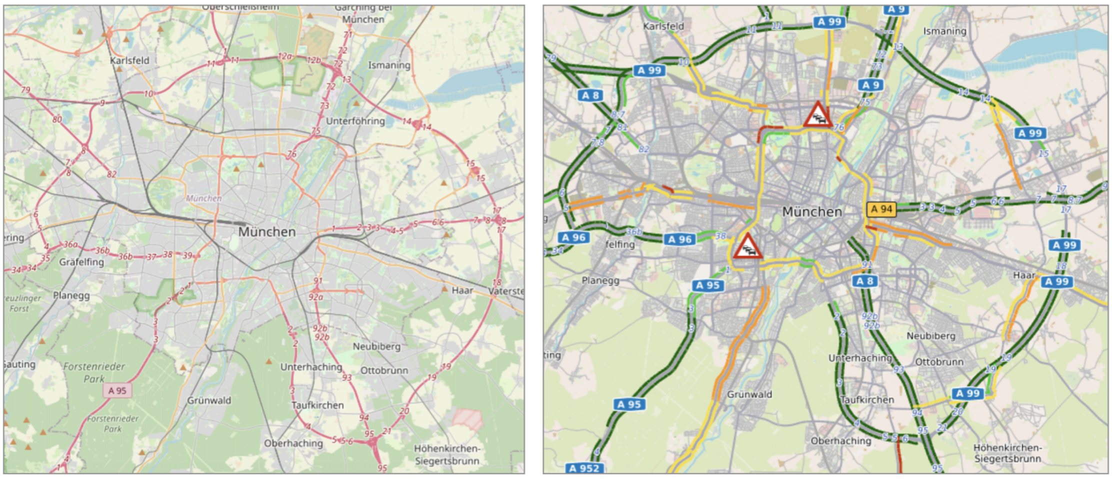


Figure 2.1: Left: An example for supply information in Munich[6], graphically visualized as a 2 dimensional map. Right: Demand information in Munich[7], visualized as expected travel speeds (Dark Green: 80 km/h, Bright Green: 61-80 km/h, Yellow: 41-60 km/h, Orange: 21-40 km/h, Red: 0-20 km/h)

Concept 3: Situational Information
Situational information contains additional meta-data that may or may not influence a traveler's route choice, such as the current weather conditions, the financial price of a route, the expected emissions for a transport connection, or the proximity of fast food restaurants near the selected route. Situational information is usually required by the traveler's current circumstances, such as hunger, lack of money, or unfitting clothing for heavy rain.

Background: Situational information is an often underappreciated component that significantly impacts travel behaviors and choices. Liu et al. describe the influence of weather information on travel behavior [31]. Cools and Creemers investigate the impact of weather forecasts on travel behavior, identifying a significant effect. Consequently,

[6] https://www.openstreetmap.org/#map=11/48.1420/11.547 Accessed: 15.11.20233
[7] https://www.br.de/nachrichten/verkehr/index.html (Data: 13.10.2023, 08:30:00) Accessed: 13.10.2023

they suggest the implementation of a road weather information system that is directly linked to the weather forecasts. However, the impact of situational information on routing behavior contains more variables than the weather conditions. Franken and Lenz describe that in the everyday context, people often act as they did before in the same or similar situations, meaning that without a situational change, travelers are likely to stick to established behavior patterns. Yet, they (re)consider how they act only if situations are completely new or unknown so that previous behavioral patterns do not fit [32]. However, a new situation may only be identified and incorporated if information about the situation changes reaches the traveler.

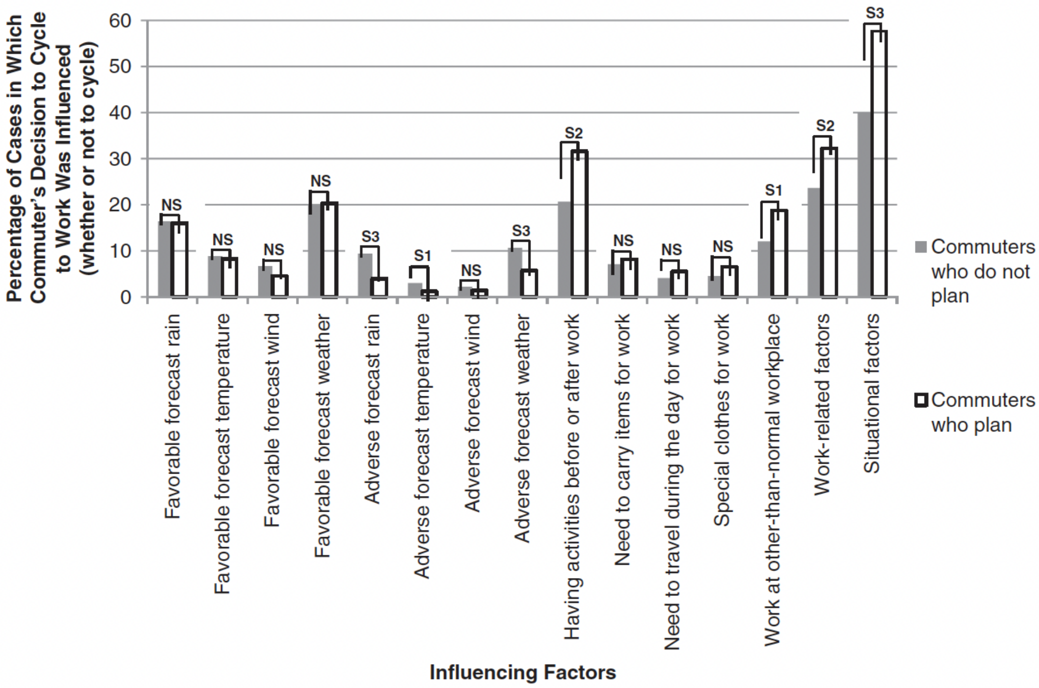


Figure 2.2: An example for situational weather influences on cyclists' travel behavior, published by Ahmed et al. [1].

An example of both supply- (left) and demand information (right) is provided in Figure 2.1. An example of situational information influences is provided in Figure 2.2, demonstrating how factors such as the weather forecast, work-related factors, activities before and after work, and other situational factors each impact the decision to cycle to work. Based on this information, either the travelers themselves, a person they have put in charge, or a navigation device they are using calculate various routes to their destination. Optimally, the calculated route will pay full respect to the information described in concepts 1-3. From the available options, the traveler then selects the route that will maximally satisfy the traveler's current routing preferences. Whether the selected route will actually result in the highest satisfaction remains uncertain, as the network, its demand, the current situation, and the traveler's needs constantly change with time passing. Yet, at the moment of decision-making, the traveler will always opt for the route they currently perceive as subjectively optimal to them.

2.3 The Evolution of Navigation

But how does the traveler select the route with the greatest satisfaction? Before we answer this question, we first need to look at the previously mentioned types of routers: The traveler themselves, a human representative (e.g., taxi drivers, tour guides, parents, etc.), and a digital representative (e.g., Google Maps, Apple Maps, built-in car navigation systems, etc.).

2.3.1 Game Theoretical Concepts

As we move from the individual to the collective, we transition to the realm of game theory, where these individual choices intersect with broader societal patterns. Navigation 3.0, encapsulated in Concept 9, advocates for a balanced diversity in routing, reminiscent of the stability found in ecosystems abundant with biodiversity. This concept resonates with game-theoretical principles, especially when considering the multi-dimensional Nash Equilibrium. In this equilibrium, individuals' pursuit of personal utility contributes to an overarching system balance, much like the bio-diverse routes contribute to a resilient navigation system.
The forthcoming section on Game Theoretical Concepts seeks to unravel the strategic interplay that underpins the traffic market theory. Here, we are not just observers of traffic as a flow of vehicles or pedestrians but as a strategic game played by rational actors. These actors, or travelers, though often modeled simplistically as uniformly rational, are anything but homogenous. They bring with them the complexities of their physical, psychological, and situational realities into a strategic environment where not just the fastest route but the most balanced and resilient system is the ultimate goal.
In the realm of traffic market theory, travelers behave in a manner that's analogous to participants in financial markets. These travelers, often perceived as rational, are characterized by a model built on a few assumptions [33]. However, this simplicity often leads to the neglect of essential variables. Although this makes it straightforward to model, the end results tend to diverge from reality.

When examining the supply, we primarily consider the underlying networks [34]. In contrast, the demand side encapsulates all travelers as a collective unit. Within this framework, univariate optimization stands out. It focuses solely on travel time, providing calculations that are easier to compute. Nevertheless, this simplicity offers no mechanism to bypass potential bottlenecks.

Moving on to traffic psychology, there's a distinct parallel between financial markets and the behaviors of what we term the "irrational traveler". This category of traveler presents a more complex modeling challenge but results in outcomes closer to real-world behaviors [35]. One key aspect is the motivation system. An analogy can be drawn from lab experiments where rats pull on springs. These experiments yield a singular value as a result, even when influenced by numerous inputs. Similarly, human route choices are motivated by multiple factors, which become nearly impossible to deduce solely from the chosen route.

In the context of information arbitrage, the absence of complete information often paves the way for decisions that aren't optimal. It's akin to a scenario where an incomplete understanding of better alternatives leads individuals to settle for the "best known" choices contentedly. Further complicating the picture is the concept of multivariate optimization, which requires the understanding of theories like the Nash Equilibrium [36, 37]. The Nash Equilibrium describes, in part, how the introduction of subjective utility functions contributes to alleviating bottlenecks by creating a wider range of subjective optima.

Addressing complexity reveals that an increase in available options also demands an increase in the volume of subjective user information that needs to be input into the algorithm. These input variables need to be computational and comparable to optimize effectively. Here, the challenge lies in translating human preferences into feasible, calculable solutions.

2.3.2 Navigation 1.0

Before the advent of computers, route determination rested with individual travelers or the designated navigator within a traveling group, a human representative. This direct involvement made it easier for travelers to customize their routes based on their needs. Either the travelers were in charge themselves, only requiring them to be aware of their own needs, or they would have to communicate their needs to their human representative. Either way, the interfaces for communicating personal routing preferences were close to ideal. The issue with such navigation resides in the responsible human's informedness and ability to calculate routes based on traveler preferences and their knowledge of the network. Before the advent of automated and interconnected navigation devices, acquiring the necessary information was a challenge. Relying primarily on static information sources like maps or personal memories, these sources were often incomplete. In a similar inefficient fashion, information about roadblocks, congestion, or delays was drawn from radio broadcasts (if already available) or other travelers, forcing the travelers themselves to connect this information with their planned path.

> In the 1800s, John, departing from Harrington to Crestwood, unfolded his weathered map, tracing the route with his finger. Recognizing key landmarks, occasionally asking passersby to confirm his direction, and retrieving information about the current road quality, he navigated the maze of dirt roads. At crossroads, he'd gauge the sun's position or study the moss on trees to ascertain his heading. Despite the absence of clear signposts, John's keen observational skills and prior knowledge of the terrain ensured he stayed on course, leading him to Crestwood by evening's fall.

If all this information is indeed available, humans run into another issue: their calculation capabilities. The human capacity to calculate the optimal route from available information is limited. Experienced travelers within a particular network traditionally have an advantage due to their prior knowledge. Local taxi drivers are an example of such local experts who utilize their extensive training and resulting knowledge of the traffic network to calculate routes quickly. Only new information has to be processed, while static knowledge is automatically included using self-established patterns.

Figure 2.3: Navigation 1.0: Sextants[8], maps[9], and individual knowledge[10] as typical representatives of navigation prior to the invention of satellite-connected devices (GPS and mobile internet)

[8]https://de.wikipedia.org/wiki/Datei:US_Navy_031025-N-8955H-005_Quartermaster_2nd_Class_Martineau,_from_Ft._Lauderdale,_Fla.,_uses_a_sextant_to_shoot_the_sun_line_from_the_port_bridge_wing_of_USS_Blue_Ridge_(LCC_19).jpg Accessed: 15.11.2023

[9]https://thewalkingmermaid.com/blog/things-to-do-before-going-on-a-road-trip Accessed: 15.11.2023

[10]https://www.universaltaxidispatch.com/blog/2015/08/taxi-etiquette-is-it-okay-to-strike-a-conversation-with-the-cab-driver/ Accessed: 15.11.2023

In contrast to these local experts, tourists and other types of travelers who are unfamiliar with local static knowledge lacked this advantage, resulting in long calculation times and sub-optimal routes. Researchers of this era, like Vetta [38], Roughgarden and Tardos [39], or Golledge [40] argued that this method of route determination was sub-optimal due to travelers' inability to optimize their routes. They, too, identify a lack of information or ability to process said information as a source for this issue. They proposed central control mechanisms as a solution to manage traffic more intelligently and prevent such inefficiencies. We call this type of Navigation, which does not utilize automated digital support systems **"Navigation 1.0"**.

In the 1970s, John, departing from Harrington to Crestwood, pulled out his well-folded paper map from the glove compartment of his classic sedan. As he drove along the asphalt roads, he'd tune in to the local radio stations, listening for any traffic updates or reports of road closures. Occasionally, he'd pull over at gas stations, not just to refuel but to ask attendants about the best routes or any recent changes in the roadways. The highway signs, now more prevalent, guided him at major junctions. Although there were no GPS devices or digital maps to rely on, John's trust in his paper map, his sense of direction, and occasional interactions with locals ensured a smooth journey, leading him to Crestwood by evening's fall.

2.3.3 Navigation 2.0

With the digital age and the rise of the internet, we also witnessed a surge in digitized information, computational power, and, eventually, semi-autonomous navigation systems. While initially overshadowed by local experts, the competence of these systems has increased with digitization.

In the year 2000, John, setting out from Harrington to Crestwood, powered on his car's early-model GPS navigation system. The device, with its pixelated screen, plotted a clear route for him. As he drove along the modern highways, John occasionally glanced at the GPS, appreciating the satellite-assisted precision it offered. However, for real-time traffic updates, he still relied on the radio, tuning into local stations broadcasting any traffic jams or roadwork ahead. Sometimes, he'd have to cross-reference the radio updates with his GPS, making manual detours when necessary.

In places with frequent traffic data capture, local experts have little to no advantage over these systems. Unlike these experts, digital navigation systems can access multiple data sources simultaneously. Additionally, local experts have to rely on their senses, extended by radio broadcasts, past experiences, and their own assumptions. Applications like Google Maps or Apple Maps instead utilize real-time data drawn from all active users' current locations and movements. The more users participate, the more accurate the representation of the traffic situation becomes. So, with the continuous digitization of real-world data, we can also see a shift in the most efficient tactic available for navigation. As this shift describes a substantial change to the navigational process, away from local knowledge experts and Navigation 1.0 toward computerized navigation systems, we frame this type of Navigation: **"Navigation 2.0"**.

In 2020, John, embarking on his journey from Harrington to Crestwood, connected his smartphone to his car's infotainment system. As he input the destination into his preferred navigation app, it instantly calculated the quickest route, factoring in current traffic conditions sourced from the real-time data of millions of users. The high-resolution display showed turn-by-turn directions, with voice assistance guiding him at every junction. Along the way, the system alerted him of any upcoming congestion, suggesting alternate routes when beneficial. Additionally, if there was an accident or unexpected road closure, the app would reroute him automatically. On the highway, digital signs overhead complemented his app's info, providing speed limit changes and traffic updates

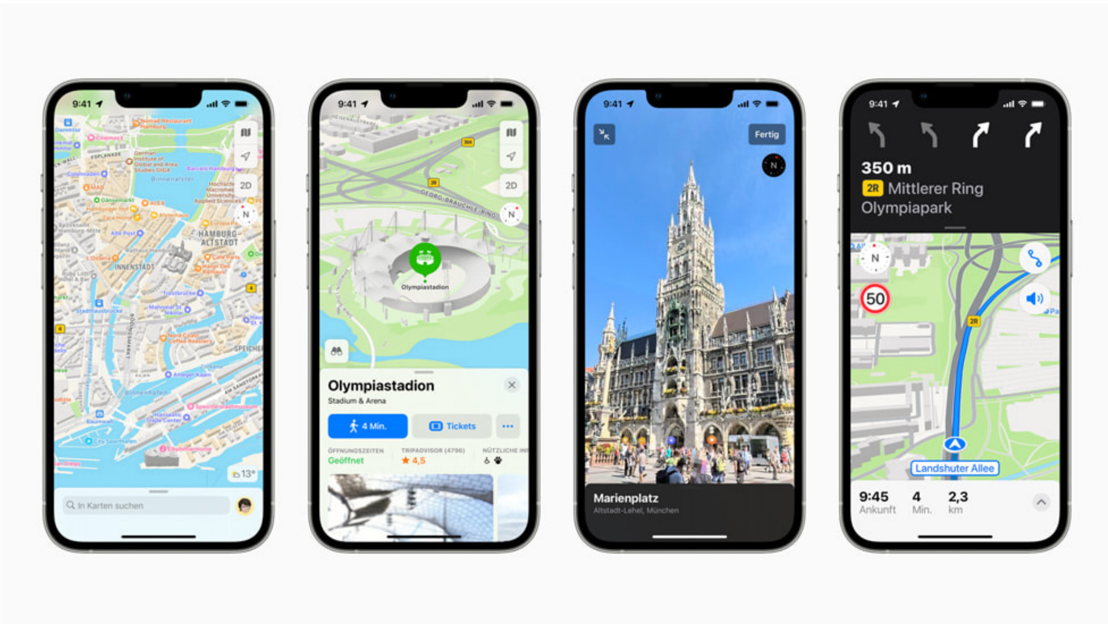


Figure 2.4: Apple Maps[11] as a typical representation of modern navigation applications representing Navigation 2.0.

As the change from 1.0 to 2.0 is continuous and yet substantially different, we chose to take the approach used for labeling versions of software increments in computer science. By showing that versions 1.x, and 2.x of navigation in between these hard version clusters

[11]https://www.apple.com/newsroom/de/images/product/apps/standard/Apple-Maps-update-2022-DE-hero_big.jpg.large.jpg Accessed: 15.11.2023

do exist, we intend to highlight the gradual change in the dominance of each system using this labeling approach. While computers today can rapidly process vast amounts of data to continuously calculate an optimal route for bike rides, walks, public transport, or car rides, this was not the case at the start of this transition process. Instead, their ability to outperform local knowledge experts gradually evolved over time, growing from support systems that were handy yet regularly overruled by local experts to the now dominant method for navigation in sufficiently digitized areas of the world.

In a potential future, John, setting off from Harrington to Crestwood, activated his car's Navigation 3.0 system. Inputting his destination, the system, familiar with John's habits, crafted a personalized route. Prioritizing scenic views, as it had learned he preferred. It avoided crowded highways, guiding him along tranquil countryside roads instead, as John did not mind slight delays if the scenery was compensating. Recognizing the current light autumn rain and John's unusual fear of slipping on wet leaves, the system chose to avoid paths along wooded areas. When traffic issues arose, it rerouted, factoring in John's dislikes and likes. If fitting, the voice guidance offered suggestions: "John, a viewpoint is nearby. Interested?" To John, the journey was as much about the experience as his timely arrival at his destination. While others might not like the route John chose, finding the suggestion of a viewpoint unnecessary and invasive and regretting an untimely arrival, to John, it felt just right.

2.4 Navigation 2.0: An Underfitting Solution

Navigation 2.0 brings about an issue that was less prominent before. While it greatly improves the ability of travelers to calculate optimal routes using a vast amount of data, it falls short in integrating the traveler's personal routing preferences into the process. We previously mentioned that travelers will always select a route they think will leave them most satisfied. The following concepts 4, 5, and 6 extend the initial concepts and illustrate in which categories travelers' routing preferences may be clustered.

Concept 4: Physical Routing Preferences
Physical Routing Preferences represent a traveler's individual physical capabilities. For example, stairs or long distances can be challenging for some travelers, depending on their physical condition. Injuries, young age, old age, and disabilities are typical examples of such physical conditions. Travelers with unusual physical traits may also have unusual routing preferences, making them a relevant factor for their navigation.

Background: Physical Routing Preferences address the tangible constraints a traveler may experience due to personal physical limitations. Musselwhite and Shergold delved into the transportation needs of older people, emphasizing how physical impairments cause these individuals to eventually give up driving cars. Their findings suggest that although a similar pattern was found between the trigger and life post-car, not all older people go through the stages of giving up driving in the same way [41]. Hakamies-Blomqvist and Wahlstrom, and Marottoli et al. make similar observations, explaining that declining health is often a predictor for a reduction in driving, ending in giving up driving entirely [42, 43]. Tsai discusses the environmental factors that impact the mobility of people with disabilities, providing insight into how physical aspects influence routing choices [44]. Similarly, Card et al. highlight the mobility challenges faced by disabled travelers, recommending travel providers consider the travel intentions and health conditions of the physically disabled [45].

Concept 5: Psychological Routing Preferences
Psychological aspects, like a traveler's personality, also influence route choice and are rather consistent. Measures, such as the Big Five personality traits, capture these consistent personality traits in categories, suggesting that travelers who differ in these personality traits value the quality of a route differently. For example, a strongly neurotic traveler will perceive travel risks differently than their hardly neurotic counterpart by preferring safer routes and planning longer time buffers. A hardly neurotic traveler, in return, is likely to accept higher risks in their routes.

Background: Travelers' psychological makeup undeniably impacts their route selection. Costa and McCrae's work on the NEO Personality Inventory, often referred to as Big Five Factors of Personality, offers a detailed analysis of their measurement and reliability [46]. Jani explores the relationship between the Big Five Factors of personality and travel personality [47, 48]. Their findings suggested that travelers with different personality traits exhibit distinct preferences and aversions, shaping their route decisions. Dahlen et al. look into correlations between Big Five Factors, trait driving anger, and sensation seeking, finding consistent correlations to unsafe and aggressive driving [49]. Morar et al. investigated changes in travel behavior during the COVID-19 pandemic and the role of tourists' personalities. The results of their study suggest that specific information about COVID-19, coping mechanisms, fear of travel, and neuropsychological personality traits may affect travel behavior in the pandemic period [50]. Terzić et al. also utilized the COVID-19 pandemic to investigate human psychology in relation to travel behavior. Using a random sample of roughly 50.000 participants from 29 European countries, they conclude that risk perceptions and behavioral responses of travelers differ significantly between individuals, groups, and even nations [51]. Prevendouros also investigates potential associations between personality and individual travel behavior characteristics. Prevendouros concludes that personality characteristics tend to correlate well with residence location selection, automobile ownership, and travel characteristics [52].

Concept 6: Situational Routing Preferences
Situational factors include things like the traveler's current clothing, their current financial situation, their current stress levels, or their current preference for scenic views. Major events, like the SARS-CoV-2 pandemic in 2020, also fall under situational influences that can change how travelers choose their routes. While the physical and

psychological factors come from the travelers themselves, situational ones are external.

Background: Situational factors, whether predictable like weather or unforeseeable like pandemics, sway route choices. A recent event, the COVID-19 pandemic, made this quite clear. Various researchers described the influence of the prevailing situation on travel behaviors and routing preferences. Bratic et al. focused on observed changes in travel risk perception and vacation travel behavior [53]. Thomas et al. and Harrigton et al. looked into changes in commuting behaviors during and after the pandemic [54, 55]. Lancée et al. revealed that the effect of different ways of commuting differs across situations, especially with regard to commuting mode, concluding that there is no one way of commuting that is optimal for everybody [56]. So and Lehto investigate situational differences between Japanese travelers by comparing travel with friends, traveling with family members, and traveling alone. Their results indicate unique situational characteristics during family travel but similarities in travel behaviors for friend groups and solo travelers [57]. Endsley's works on situational awareness in aviation and other domains illustrate how the Traveler's current situation impacts decisions [58, 59].

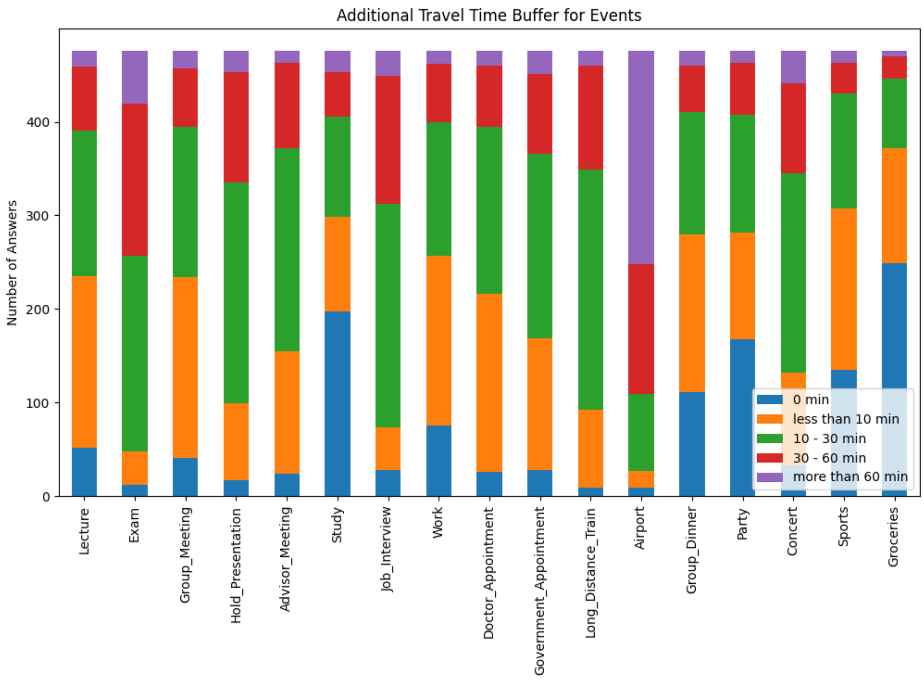


Figure 2.5: Study data from the Master's thesis of Bayram Ahmadov, showcasing the distribution of planned in buffer time depending on the travel purpose [2].

In the context of an ongoing Master's thesis by Bayram Ahmadov (publication in progress), we asked students of the Technical University and the Ludwig Maximilian University in Munich to estimate how much additional time they would plan as a safety measure when traveling [2]. We then faced these students with various events of different severity that they know from their daily lives. A trip to the airport, a university lecture, a university exam, a study group meeting, a leisure activity, and some more. Provided five answer options, the student answers generally resulted in normal distributions. However, the center of each normal distribution would shift based on the event type, causing students to include more additional time when traveling to the airport or for an exam and much less for leisure activities or lectures. The results are visualized in Figure 2.5. Similarly, we also asked students how other factors would affect their travel choices (Figure 2.6), such as travel time constraints, developing body odor, their own physical conditions, or route safety. Here again, we found that students' answers are normally distributed with shifts to the left or right based on the event type. The idea of the study is to showcase how physical, psychological, and situational preferences impact travelers (in this case, students) differently, showing how they each follow their own multi-varied equations when evaluating the quality of a route. With 476 participating students for a population of 100,000 students studying at TUM and LMU worldwide, the study is representative at an error span of 5% and a confidence level of 95%.

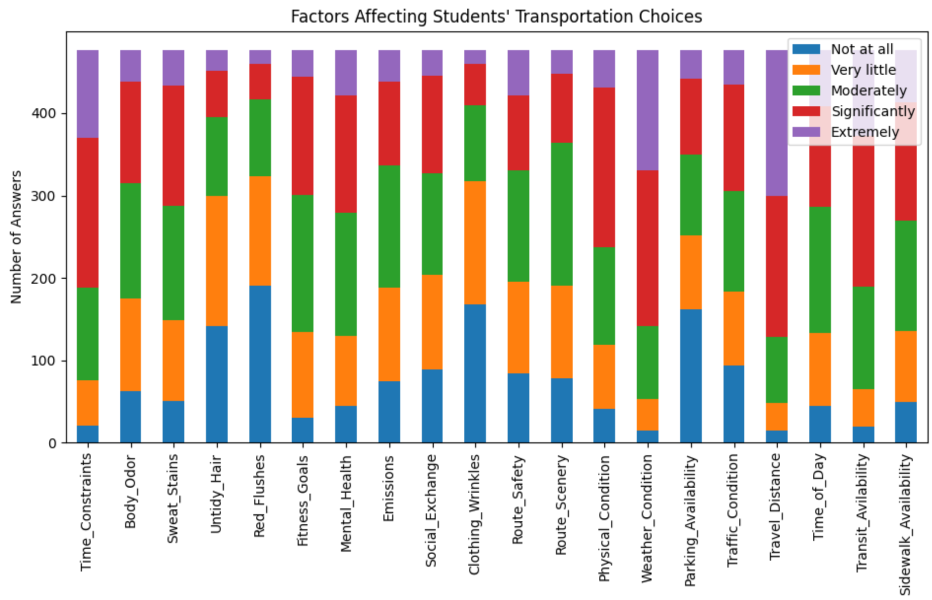


Figure 2.6: Study data from the Master's thesis of Bayram Ahmadov, showcasing the distribution of travel behavior considering more or less affecting factors [2].

All these physical, psychological, and situational factors add complexity to routing. To simplify things, current routing algorithms tend to focus on singular optimization goals: Finding the fastest, shortest, or cheapest route. What these optimization variables share is their situational nature. None of them take the traveler's psychological or physiological biases into account. So even though Navigation 2.0 finds faster, cheaper, and shorter routes than most travelers using Navigation 1.0 can, it brings up a new issue: Navigation 2.0 hardly accounts for personal routing preferences. To better understand this issue, we draw parallels to the concepts of overfitting and underfitting that are prominent in statistics and machine learning.

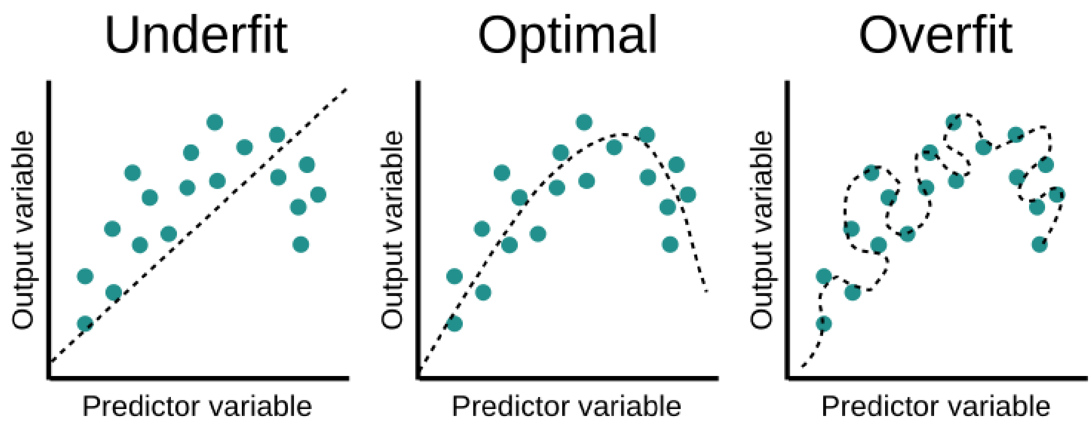


Figure 2.7: Overfit (Navigation 1.0), underfit (Navigation 2.0), and optimal fit (Navigation 3.0) as an exemplary parallel to flaws in Navigation 1.0 and 2.0 and setting a goal for navigation 3.0.[12]

Overfitting occurs when a model (representation of reality) studies the training data too well, reaching a point where it becomes overly complex. It captures noise and nuances in the data that aren't necessarily representative of the broader data set, leading to poor generalization of new data. Similarly, Navigation 1.0 can be seen as an "overfitted" system. The route determination, based largely on personal memories and static sources, was intimately tailored to the individual's experiences and biases. Local taxi drivers and seasoned travelers, with their intricate knowledge of specific areas, would navigate the terrain with a depth of detail akin to an overfitted model. Their decisions were, however, heavily based on the small details of their experiences, which might not have been optimal or relevant when new or unexpected travel challenges arose.

The term underfitting is used when a model is too simple, failing to capture underlying patterns in the data. This results in a model that neither performs well on the training data nor on new data. Navigation 2.0, with its reliance on broad computational data, can be thought of as "underfitting" the navigational challenge. While these systems

[12] https://www.fastaireference.com/overfitting Accessed: 15.11.2023

excel at aggregating vast amounts of data to quickly calculate and suggest routes, they often overlook the nuances and personal preferences that individual travelers might have, be they physical, psychological, or situational.

In Figure 2.7, the differences between these two states of fitting are visually represented, with the suggestion that a potential **Navigation 3.0** might find the optimal balance. Just as the ideal model in machine learning strikes a balance between complexity and simplicity, the ideal navigation system balances the broad computational prowess of Navigation 2.0 with the intricate personalization of Navigation 1.0.

The journey from overfitting to underfitting, or from Navigation 1.0 to Navigation 2.0, has been one of refining the balance between personal preferences and generalized computational power. The challenge that remains is finding the "just right" fit - a Navigation 3.0, perhaps - that combines the depth of personal preferences with the informational breadth and computational power of digitized navigation systems.

2.5 Homogeneity of Navigation 2.0

The previous section discussed how current navigation systems tend to oversimplify and, therefore, underfit human complexity in navigation. Navigation algorithms show a tendency to prefer rational and easily measurable variables of a situational nature over psychological and physiological aspects. When travelers interact with a Navigation 2.0 device, they are usually just required to provide the starting (or current location) and destination, along with their preferred mode of transportation and departure time. The navigation device will then calculate routes for all typically available modes of transport connecting start and destination while focusing primarily on the traveler's initially preferred mode of transport. Options for other routes are not displayed on the map, but in most cases displayed by indicating the travel time with other modes of transport. For such routing solutions, the minimal travel time is generally equated with their traveler's satisfaction. Due to the success of solutions like Google Maps, Apple Maps, and other representatives of Navigation 2.0, equating travel time and travel satisfaction seems to be a sufficient solution for many travelers. But it is also an efficient solution for the routing algorithms as calculating travel times for routes is straightforward to compute if data about network supply and demand are available. But, in the process of calculating routes, these algorithms must make various assumptions about the traveler, for example, their travel speed using each transportation mode. These assumptions cause centralization of routes as each traveler, independent of their situation, physiology, or psychology, will receive the same routes, solely depending on their routing origin, destination, mode of transport, and departure time.

2.5.1 From Water's Behavior to "The Law of the Fastest Route"

As every participant in the traffic network using Navigation 2.0 is forced to compete for the same resource, the fastest route connecting their origin and destination, more conflicts are caused as mostly one variable is optimized: The travel time. While conflicts are hardly avoidable in traffic systems, causing travelers to interfere with other travelers' paths, they are reduced by spreading travelers evenly across the entire traffic network. However, if every traveler optimizes their routes for the fastest connection between their origin and destination, then they are likely to choose similar routes. Chosen routes are then filled until another route becomes more time-efficient. Navigation 2.0 devices then identify this congestion as travelers' travel speeds start to decline on certain edges in the network graph. This change is then included in route calculations for following travelers by increasing the expected travel time on a congested route. Similar to prices in markets that increase when a good is strongly requested but supply remains limited, the expected travel time of a route will increase if it is strongly requested.

In this sense, obeying their navigation devices makes traveler behavior more predictable. With closer adherence to navigator instructions, rising informedness of the devices, and uni-dimensionality of optimization variables, traffic shows traits of fluids like water. The behavior of water, always seeking the path of least resistance, also follows this singular optimization method. As every water particle adheres to the same laws of physics without any deviations in psychology, physical composition, or situational preferences, it does not show any deviations along these dimensions. So, while simulating them is complex, it mostly follows rationally observable rules. So the better informed Navigation 2.0 devices are, and the more reliably travelers adhere to the device's instructions, the more the traveler neglects their individual preferences, replacing them with the commonly shared preference for the quickest route. In this regard, traffic flow in Navigation 2.0 is comparable to the way water fills a river with multiple channels. Just like water follows gravity, filling the lowest channels first, travelers, increasingly relying on navigation devices, follow what might be termed as "The Law of the Fastest Route." However, this focus on speed might not always equate to traveler satisfaction, ignoring the potential for distributing travelers along other optimization variables. As we have already shown, in Figure 2.5 and 2.6, many other influences exist and are relevant to travelers. So, in reality, a traveler's genuine optimization goal is not their travel time. It's actually their individual travel satisfaction, which involves a multi-variate equation, of which one variable is travel time. This equation uses a multitude of variables (x_i), each with an individual weight (w_i), to compute the highest possible satisfaction value ($y_{satisfaction}$) for the traveler. Here, travel time is a variable among others that is to be optimized according to the weight placed on it by each traveler.

$$y_{satisfaction} = w_1 * x_1 + w_2 * x_2 + w_3 * x_3 + \dots w_n * x_n \tag{2.1}$$

Yet, current navigation systems, emphasizing minimal interaction, are not tailored for this holistic approach. The challenge lies in crafting a navigation system that integrates individual needs without adding significant interaction burdens. Subsequent sections will explore potential solutions for a more inclusive and efficient navigation system. Publication 1 discusses solutions for minimizing costs for the traveler that come along with the benefits of customized routing. Alongside a theoretical analysis, approaches for the cost-efficient inclusion of travelers' situations, physical conditions, and psychological constitutions are presented.

2.5.2 Game Theoretical Issues

Critics like Vetta [38] correctly explain that travelers make suboptimal path choices due to a lack of knowledge or limited computational abilities. Critics of Navigation 1.0 also describe that the overall travel time for all travelers is minimized if all act to minimize their own travel time in synchronization with current traffic information. It is, however, not correct to assume that a minimization of travel time results in a maximization of overall traveler satisfaction. As previously explained, the actual goal of a traveler is not to minimize their time spent traveling but their individual satisfaction with the route. While travel time is an important element of this satisfaction function, it's just one of a potentially endless list of variables that the traveler wishes to optimize and weigh subjectively. This centralizes previously distributed travel interests, causing an unnecessary reduction in travel flow. As human navigation was flawed and highly individualized, this individuality led to a more even distribution of travelers across various routing options, thereby reducing the stress on the theoretically fastest route. In the era of Navigation 2.0, knowledge and calculation errors are eliminated, but the multi-dimensionality of travel needs is also lost in the process. In relation to the concept known as "the tragedy of the commons", there were many different types of commons in the times of Navigation 1.0 that were each operated inefficiently [60]. While a distribution across many commons is desirable, their inefficient operation is undesirable. Navigation 2.0 forces all travelers to use the same, single common, operating it most efficiently.

2.5.3 Supply- and Demand Distortion

While under Navigation 1.0 traffic jams may arise because travelers don't know how to bypass them, under Navigation 2.0, traffic jams arise because travelers' subjective preferences are neglected. Looking at a traffic jam, both systems can cause them. In Navigation 1.0, because travelers don't know how to get around them, and in 2.0, all travelers take the same route until a faster route is found. It's unclear whether a traffic jam that occurred under the reign of the one system would also have occurred under the other.

So, paradoxically, the apparent solution to the supply distortion problem of Navigation 1.0 leads to the emergence of a new problem in the form of demand uncertainty in Navigation 2.0. If we view road traffic in the age of Navigation 1.0 or 2.0 from an economic perspective, we would have to describe it as an "inefficient" market in which either large arbitrage opportunities (Navigation 1.0) remain untapped or the market does not reflect the actual demand (Navigation 2.0). A frequently cited concept in this context is the "Efficient Market Hypothesis," which states that in an efficient market, prices reflect all available information at all times [61]. This information includes both the currently available knowledge as well as the prevailing needs of all market participants. If this is not the case, there is a market inefficiency.

Consequently, the individualization of navigation algorithms contributes to a more realistic representation of supply and demand in the transportation market. While it is expected that this approach will increase the average travel time, it is conversely anticipated to enhance the overall travel satisfaction of all travelers.

2.5.4 Evolution of Navigation Strategies

However, this gain in utility comes at a cost, as greater individualization requires the identification and consideration of more variables from the traveler during route calculation. From a univariate optimization equation, a multivariate optimization equation emerges. Imagine a traveler asking a taxi driver to get them to their destination as quickly as possible but to avoid traffic jams and choose routes with as many green spaces along the roadside as possible. As these objectives can contradict each other, the taxi driver is forced to calculate various routes and have the traveler evaluate them to get a "feel" for their preferences. This process is laborious for both the taxi driver and the traveler. The same problem exists in the interaction with the navigation system. The key here is to primarily place the costs of enhancing individualization on the device and the algorithm so that the user is disproportionately less affected by the cost increase. To be more concrete, we formulate the following concept descriptions for Navigation 1.0, 2.0, and 3.0.

Concept 7: Navigation 1.0 - Unnecessary Conflicts
Drawing from the ecological concept of competitive exclusion[55], two species cannot coexist in a habitat if they are competing for exactly the same resources. In the era of Navigation 1.0, individuals operated based on personal knowledge and individualized strategies. Just like species that haven't found their specific niches, travelers often overlapped in their choices of routes. This overlap created unnecessary congestion and conflicts. Even if satisfying solutions existed, the travelers lacking knowledge of their existence left these options vacant, creating room for arbitrage. Just as two similar species might fight over the same territory or food source without a clear division of ecological roles, travelers jostled over the same popular routes without a systematic distribution.

Background: Just as species might clash for resources when their ecological roles aren't clearly defined, travelers often grapple over the same popular routes due to a lack of information dissemination. Helbing et al. [62] discussed traffic dynamics and related them to concepts from social science. They found that individual strategies based on limited knowledge can lead to sub-optimal route choices, echoing the conflicts seen in competitive exclusion.

Concept 8: Navigation 2.0 - Homogeneity
Moving to Navigation 2.0, the scenario can be liked to a monoculture in ecology. When a single species dominates a habitat, it leads to a lack of ecological diversity. With Navigation 2.0, introducing a universally "optimal" path determined by algorithms forced everyone onto the same route, just as a dominant species in a monoculture outcompetes or replaces other species. While this might seem efficient, it creates vulnerabilities. In an ecological monoculture, if a pest or disease targets the dominant species, it can devastate the entire habitat. Similarly, if an issue arises on the "optimal" route (like a roadblock or an accident), it can cause disproportionate disruptions as everyone is funneled through that path.

Background: Navigation 2.0's framework can be related to ecological monocultures, where a single species or approach dominates the landscape, leading to vulnerabilities. Barabasi and Albert [63], in their pioneering work on network dynamics, emphasized the fragility of networks dominated by highly connected nodes. This is analogous to the idea that when everyone follows the same "optimal" route in Navigation 2.0, disruptions to that route can have cascading effects. Research by Goodchild on spatial cognition and GIS sheds light on how geospatial technologies influence (and are influenced by) human cognition [64].

Concept 9: Navigation 3.0 - Balance
In the concept of Navigation 3.0, we look towards the principles of ecological stability and resilience derived from biodiversity. In a diverse ecosystem, various species have carved out their niche, coexisting without direct competition and ensuring the system's overall health. If one path gets disrupted, only a subset of travelers is affected, just as if one species in a diverse ecosystem faces a threat, the entire system doesn't collapse.

Background: Drawing inspiration from the stability offered by biodiversity, Navigation 3.0 aims for a balance in route choices among travelers. Page [65] highlighted the strengths of diversity in problem-solving and decision-making. In terms of navigation, a diverse set of routes taken by different travelers ensures better resilience against disruptions. This diversity also aligns with the concept of a multi-dimensional Nash Equilibrium, where travelers make decisions based on personal optimizations rather than a singular META-Strategy, as discussed by Nash himself [36, 37].

From the perspective of the Nash Equilibrium [36, 37], Navigation 1.0 describes a non-equilibrium state in which "players" have room for optimization, as information arbitrage between travelers exists. Navigation 2.0 mostly eliminates information arbitrage between "players" by providing a META-Strategy (Most Efficient Tactic Available). This causes all players to optimize for the same strategy, achieving a one-dimensional Nash Equilibrium. Navigation 3.0 adds subjectivity to the definition of a META-Strategy, thereby allowing for a multi-dimensional Nash Equilibrium.

A Nash equilibrium refers to a situation in game theory where each player's strategy is optimal given the strategies chosen by the other players [36, 37]. No player can benefit by changing their strategy, while the other players keep their strategies unchanged. The concept can be applied to games with different dimensions in terms of strategy space. Here's the distinction between a unidimensional and multidimensional Nash equilibrium:

1. **Unidimensional Nash Equilibrium:**
The strategy space of the game is one-dimensional for each player. This means a single scalar value can describe each player's strategy. An example might be a game where players decide how much to invest, with the investment amount ranging from \$0 to \$100. The equilibrium would be a certain amount where neither player has the incentive to change their investment given the choice of the other player. Graphically, this can often be represented on a single line or curve.

2. **Multidimensional Nash Equilibrium:**
The strategy space for at least one player is multi-dimensional. A vector rather than a scalar describes a player's strategy in such games. Consider a game where a company decides on both the price and the quantity of a product it will produce. Both the price and quantity constitute the strategy, making it two-dimensional. Finding a Nash equilibrium in multidimensional spaces can be more complex because one has to consider changes in multiple dimensions and how they interact. Graphically, the strategy space might be represented by a surface or a region in higher-dimensional space. In such games, the equilibrium is a point (or set of points) in this multi-dimensional space where no player is incentivized to change their strategy in any dimension, given the strategies chosen by the other players. In essence, the difference between unidimensional and multidimensional Nash equilibria boils down to the complexity of the strategy space and the number of dimensions in which players can make decisions.

Navigation 3.0, therefore, aims to embrace the individualized nuances of Navigation 1.0 but with the informed structure of Navigation 2.0. By allowing for diverse travel routes that cater to individual preferences yet are optimized to prevent overlapping and congestion, it finds a balance. It's analogous to a mature ecosystem where each species, while following its evolutionary strategy, coexists in balance with others, leading to a stable and resilient environment.

2.5.5 Traffic Planning Methods

While different forms of navigation propagate a demand distortion, different forms of traffic planning may result in a supply distortion. A common yet controversial process of traffic planning is opinion-based decision-making. Here, individuals or groups discuss plans for structural network changes, either through construction or ruling. While decision-making based on a representative sample of the population is scientifically sound, it is hardly met in practice. The process has a tendency to be driven by minorities of deeply affected individuals or by politicians championing personal endeavors and not by a representative sample of the population. Relying on strongly biased and, therefore, non-representative minorities or on a single politician's vision is problematic, likely resulting in adaptions of the traffic network that do not properly reflect the population's needs. The result is a supply distortion. However, a realistic representation based on statistical concepts is ideally rooted in comprehensive data samples that correctly represent an underlying population.

An additional oversight in current traffic planning is focusing on vehicular flow rather than the individual travelers causing the flow. At best, planners base decisions on observed traffic patterns. While this is scientifically sound, this method does not provide insights into travelers' actual motivations, origins, and destinations, as exerted behaviors are influenced by the available options provided by the network supply. This symptom-based approach, while quantifiable, is akin to treating the symptoms of a disease while possibly not addressing the underlying cause.

Current simulation systems for traffic flow fail to cater to the vast differences in traveler profiles. Instead, they focus on fluid mechanical flow optimization, which is an important aspect of traffic analysis but caters to a different problem. If fluid-based simulations are used to determine traffic flow, then the resulting optimization will, in return, be optimized for fluid-like behaviors. While Navigation 2.0 does promote fluid-like behaviors, Navigation 3.0 does not. For a successful transit from Navigation 2.0 to Navigation 3.0, the network supply in the form of streets, sidewalks, bike lanes, and transit options requires aid from a new type of simulation, including traveler preferences.

2.6 Research Approach

- **Goal: Transition from Navigation 2.0 to Navigation 3.0**
  - Develop a software architecture for Navigation 3.0.
  - Implement the architecture in a functional prototype.
  - Identify and explore traveler preferences
- **Goal: Build a Simulation Environment**
  - Compare behaviors of Navigation 2.0 and Navigation 3.0.
  - Develop an algorithm for Navigation 3.0 that integrates user-specific needs.
  - Ensure the algorithm can manipulate network supply (e.g., avoidance, blocking, prioritizing nodes/edges).
  - Include physical optimization variables (e.g., walking speed).
  - Incorporate psychological variables (e.g., buffer times during transfers).
  - Factor in situational variables (e.g., weather conditions).
- **Goal: Simulate Realistic Traveler Behaviors and Networks**
  - Utilize an Origin-Destination Matrix (ODM) for a detailed representation of daily travel behavior.
  - Simulate behaviors of travelers using various navigation algorithms.
  - Demonstrate interaction of different need groups (Demand) with transport network (Supply).
  - Highlight differences from the standard algorithm (Navigation 2.0).
- **Goal: Analyze Impact of Changes in Navigation Algorithms**
  - Simulate effects of structural or regulatory changes in traffic flow.
  - Analyze and compare impacts according to Navigation 2.0 and 3.0.
  - Evaluate effects on different interest groups and individual travelers.
  - Focus on depicting consequences arising from changes towards Navigation 3.0.
- **Goal: Optimize Simulation Customization**
  - Allow adjustments to the network (edges/nodes) and to demand (optimization variables of the algorithm).
  - Ensure gains from customization exceed costs to prevent overfitting.
  - Aim for optimizing benefit increase while minimizing interaction and calculation costs.

Description:

The objective of this dissertation is to support a transition from Navigation 2.0 to Navigation 3.0. To achieve this, a software architecture for Navigation 3.0 will be developed. To prove that this architecture actually works, it needs to be implemented in a functional prototype. The goal is to build a simulation environment capable of comparing behaviors from Navigation 2.0 and Navigation 3.0 against each other. For this purpose, an algorithm must be developed that is capable of Navigation 3.0, i.e., capable of integrating user-specific needs into the routing process. To do this, the algorithm must be able to manipulate the network supply (avoidance, blocking, or prioritizing certain network nodes/edges). Additionally, physical optimization variables, such as walking speed, must be adjustable. The same applies to psychological variables, e.g., longer buffer times during transfers and situational variables like the weather.

At the same time, the simulation should also be able to simulate not only individual travelers but entire networks. For this, the simulation must realistically simulate travelers. To approach realism, a detailed representation of daily travel behavior in the form of an Origin-Destination Matrix (ODM) is needed. Based on this ODM, various forms of navigation algorithms should then simulate the behavior of the travelers. The aim of simulating a traveling network is to demonstrate how travelers (Demand) from different need groups (Navigation 3.0) interact with an existing transport network (Supply), thereby highlighting differences to the standard algorithm (Navigation 2.0). These differences describe the effect of psychological, physiological, and situational changes previously made to the algorithm. In addition, the simulation should also allow the simulation of the effect of structural or regulatory changes in traffic flow. These can then be analyzed and compared both according to Navigation 2.0, i.e., without individual adaptation of the algorithm to the traveler, and with Navigation 3.0. This makes it possible to simulate the effect of such a change in the transport network for different interest groups and, if necessary, even for the individual traveler.

The aim of the simulation is to depict general behaviors and travel needs. The goal is not the fluid representation of travel flows. Instead, the focus is on depicting the consequences that arise from a change in navigation algorithms towards Navigation 3.0. The completed simulation should allow adjustments to the network (edges/nodes) and adjustments to demand (manipulation of the optimization variables of the algorithm). Any gains achieved through this customization must exceed the costs to avoid a switch from under- to overfitting. Therefore, the aim is to optimize the benefit increase while minimizing the interaction and calculation costs that arise simultaneously.

2.7 Research Questions

RQ1: How can a navigation algorithm be individualized?

RQ2: How can costs associated with improved individualization be minimized for travelers?

RQ3: How can the impact of such individualization on traffic planning and travel behavior be communicated?

3 Methodology

The iterative process of the scientific method suggests a repetition of creating new hypotheses, designing experimental environments, conducting experiments, and analyzing the results as visualized in Figure 3.1 [3]. Based on the analysis step, two actions may be invoked. If the experiment was conducted without any issues, then hypotheses are rejected, approved, or devised. However, if the experiment shows signs of contamination through unintended influences, the experiment's design or its implementation is revised. This cycle may be repeated indefinitely.

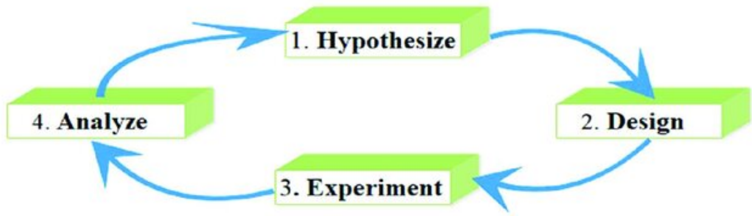


Figure 3.1: The iterative process of the scientific method, visualized by Vestin [3].

In this dissertation, we initiated this process by formulating new hypotheses and theories about traffic systems formalizing them as concepts. These concepts are accumulated in the idea of a more balanced type of navigation called Navigation 3.0. We derive the idea of Navigation 3.0 by picturing traveling as a non-cooperative n-person game. By drawing parallels to other types of n-person games like biological ecosystems and extending principles such as the Nash-Equilibrium, we cluster the currently dominant navigation solutions as Navigation 2.0 solutions, leaving room for optimization. Adhering to the principles of non-cooperative n-person games, we describe a path for improving Navigation 2.0 solutions to approach Navigation 3.0. The core idea is to improve the fitness of routing algorithms to traveler preferences yet with a disproportionately smaller increase in costs for the traveler (e.g., interaction time, interaction complexity, or route calculation time).

After formulating hypotheses, design concepts for a traffic simulation are developed to test said hypotheses. In Section 3.1.1, a first design concept for navigation theory (Figure 3.2 is introduced, which is then further formalized into a software architecture for an aligned simulation tool in Figure 3.3. The alignment between theory and software architecture is highlighted in Figure 3.14.

Within this architecture, the first implementation of the traffic simulation is approached by a customizable navigation algorithm. This algorithm considers traditional navigation factors but also supports dimensions drawn from Navigation 3.0, such as the traveler's psychology, physical condition, and current situation. A practical prototype is implemented and published, simulating the impact of cautious travelers avoiding crowded public transport stations in the city of Munich (Publication 1). To assess the potential impact of this algorithm at the market level, it is embedded in the traffic simulation architecture. By extending the first concept from Publication 1, other types of traveler groups are simulated next. Special emphasis is placed on certain groups – such as children, elderly, and physically impaired individuals. While the manipulation of routing algorithms is successfully demonstrated in Publication 1, the origin and destination pairs used in the demonstration are selected randomly.

To address the issue of unrealistic origin-destination pairs, Publication 2 introduces a statistical approach for correctly transforming standardized German traffic surveys into Origin-Destination-Matrices (ODMs).

To conduct the analysis phase, additional information outside the simulation results is required. Publication 3 presents an algorithm that allows travelers to automatically collect and label their travel behavior, requiring minimal user input. This behavioral data can then be compared to simulated behaviors, allowing us to analyze differences and either repair the experimental setup or update the initial hypotheses.

3.1 The Traffic Simulation Tool "TrafSim"

Congruent to these publications, a Traffic Simulation WebApp was implemented, allowing operators (like us) to simulate travel behaviors based on Navigation 3.0 principles. "TrafSim" offers a Graphical User Interface (GUI), including the following functionalities:

1. **GUI for the manipulation of traffic network supply:** Allows users of TrafSim to manipulate uploaded GTFS files freely. This includes adding new public transportation connections, including existing lanes and transportation types, but also any imaginable new types of public transport (hyperloop, air taxi, cable car, ...). Deleting any existing public transportation connections (Bus, Tram, Train, Subway, ...). Editing existing connections, for example, by changing departure and arrival times of transport connections or changing transportation frequencies, with both effectively increasing or reducing the travel time.
2. **GUI to manipulate traveler demand:** This functionality is met by allowing simulation operators to manipulate traveler's routing preference-related variables such as their walking speed, biking speed, allowed modes of transport, avoided

stations and lanes, a maximum of changes between transport vehicles, perceived costs for changing lane or mode of transport, maximum and minimum walking distance, maximum and minimum biking distance, and many more variables.

3. **GUI for simulation comparisons:** Allows TrafSim users to compare multiple simulations against each other and download the results as CSV files or compare them directly in the GUI of the WebApp (Map plotting and tabular format). Comparisons are done both on a collective level by comparing changes in average routing and also on an individual level, comparing differences between the same individual under different supply, demand, or both at the same time. Used comparison metrics are differences in travel time by mode of transport, travel distance by mode of transport, lane changes, and more.

4. **GUI for data import:** TrafSim users may either use pre-installed default data in the form of existing Origin-Destination-Matrices (ODMs) of Munich, network supply data of Bavaria from OpenStreetMap (Roads, sidewalks, bike lanes, etc.), public transportation data (GTFS) from Munich, or upload their custom ODMs, OpenStreetMap network data, or GTFS public transportation data.

5. **GUI for change management:** To save and manage changes to the transportation graph, the traveler demand, and the underlying data, sessions have been introduced. Any alteration of the just mentioned factors may be saved and loaded at any time. This allows TrafSim operators to compare saved changes to one another.

Following, we will first explain the underlying theory and software architectural implications of that theory before diving deeper into the actually implemented architecture. Finally, we will compare the actual implementation to the initially described theory to verify how well the implementation meets the theoretical requirements.

3.1.1 A Model for Navigation Theory

To build a simulation architecture that meets the navigation behavior of every type of traveler, we must first understand and define relevant elements of navigation. To do so, we provide a general concept of navigation. Figure 3.2 illustrates our theoretical understanding of human navigation, which was inspired by existing concepts [66, 67, 68]. We expand these concepts by dividing the overarching idea of navigation into four interacting subsystems that we call **Traffic Simulations**, **Navigation Solutions**, **Network Graphs**, and **Traffic Participants**.

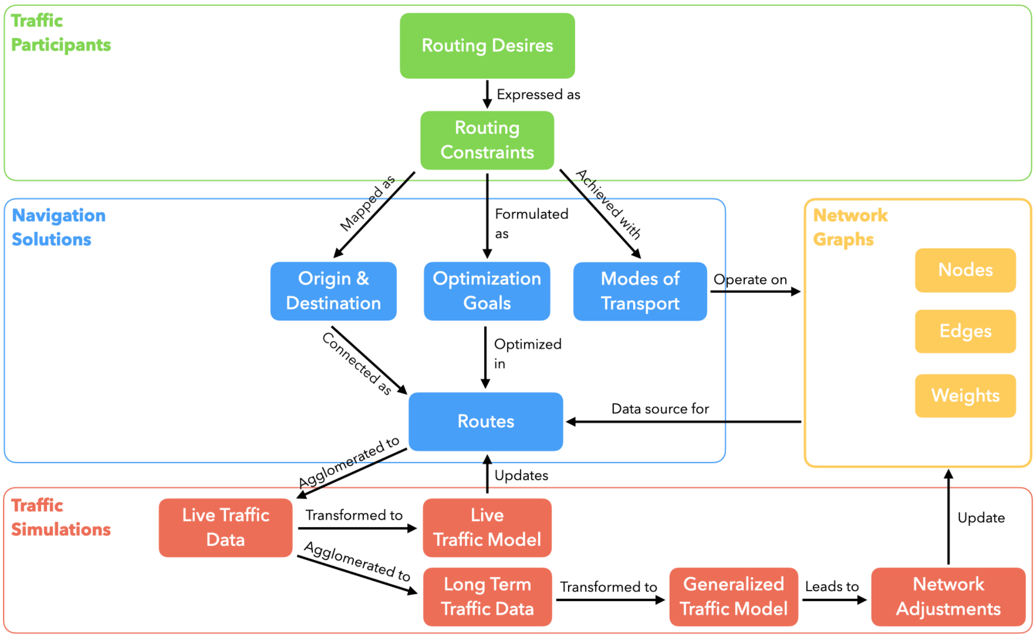


Figure 3.2: Generalized concept of navigation, including human, automated, and distributed navigation.

These four subsystems constitute the basis of navigation that we intend to include in a simulation. Each subsystem consists of elements that describe the processes inside the subsystem, as shown in Figure 3.2. Some of these elements may be prioritized or neglected in individual navigation processes, but each presented element remains fundamental to navigation. In this context, Figure 3.2 serves as a universal blueprint, showcasing all relevant factors affecting navigation. This blueprint is designed to facilitate the creation of technology-based navigation solutions.

To accurately describe as many types of navigation as possible, the model is kept correspondingly general. To demonstrate its range, we showcase how two opposing extremes of navigation are depicted using this model. The first extreme, which we call "Fully Isolated Navigation", is a process that involves no external tools, solely depending on the human navigator's senses. Essentially, this type of navigation relies on the traveler's innate sense of orientation and prior knowledge, devoid of other navigational aids. On the contrary, the other extreme could be called "Fully Distributed Navigation", representing the utmost integration of available distributed knowledge sources and computing aids into the navigation process. In the following two examples, we will briefly demonstrate how Figure 3.2 incorporates both navigation styles. However, it is applicable to any other kind of navigation between these two extremes.

Applying the model presented in Figure 3.2, a traveler adhering to the "Isolated Intuitive Navigation" approach would possess **Routing Preferences** that require a clearer expression through constraints. Let's consider such a traveler currently residing in a remote area in Canada. The preference to travel to a destination, starting at their present location (**Origin and Destination**), possibly keeping certain **Optimization Goals** in mind (e.g., avoid wildlife, avoid hills, travel in daylight, use the shortest route, ...). They next proceed to select a mode of transport (e.g., ski-doo, walking, dog-sled, ...), thereby also selecting an available set of paths that form a **Network Graph**. As this traveler doesn't utilize external knowledge, they depend on their own understanding of the local environment to develop said **Network Graph**. This graph, comprising **Nodes**, **Edges**, and **Weights on the Edges**, represents their personal estimation of possible routes and travel costs for each **Edge**. Using this graph, the **Origin and Destination**, and the **Optimization Goals**, the traveler formulates a preliminary set of potential **Routes**. Considering potential wildlife encounters, the traveler would benefit from live knowledge of the routes of other traffic participants. But, as the traveler is fully isolated, they cannot aggregate such knowledge into a **Live Traffic Data** set. Instead, they solely rely on their own senses to adapt a route, utilizing their **Live Traffic Model** to respond quickly if wildlife is encountered. The traveler can, however, use their aggregated knowledge of past encounters (**Live Traffic Data**), predicting wildlife movements during the current season (**Long Term Traffic Data**). Experiences that can be generalized into a **Generalized Traffic Data** facilitate **Network Adjustments**, ultimately updating their **Network Graph**. **Live Traffic Data**, on the other hand, is utilized to directly modify a route based on current network knowledge, such as a bear obstructing the intended path. Similarly, the model accommodates the other extreme — "Fully Distributed Navigation". Let's consider a commuter in San Francisco who needs to navigate to their workplace. Their requirements are conveyed as **Routing Constraints**, which may include needing a coffee stop, avoiding construction areas, and a preference for bike lanes. This commuter utilizes an automated **Navigation Solution** - a mobile app - that translates these constraints into a route. This requires an **Origin and Destination** (their home and workplace), an **Optimization Goal** (such as the fastest or least congested route), and a chosen **Mode of Transport** (in this case, a bicycle). The selected transport mode dictates the **Network Graph**, which in this scenario, is collated from multiple distributed data sources - city biking maps, current roadwork data, and local business locations - and merged into one. The outcome - a set of **Routes** - is influenced by traffic flow hypotheses formulated in a **Live Traffic model**. The model incorporates **Live Traffic Data**, sourced from real-time traffic reports or user data. This data is also archived as **Long Term Traffic Data**. The data then aids in identifying patterns, like increased congestion on specific routes at certain times, that are expressed in a **Generalized Traffic Model**. Based on these insights, **Network Adjustments** such as traffic rules or construction changes to the infrastructure are made. These adjustments are then reflected in the updated **Network Graphs** and influence future route suggestions.

As previously mentioned, any other type of navigation can be expressed along the modeled relations in Figure 3.2, similar to these two extremes. Building upon this model, we next introduce our traffic simulation tool.

3.1.2 An Architecture for a Preference-Oriented Navigation Simulation

To communicate the concept illustrated in Figure 3.2, we have developed the TrafSim simulation tool that assists individuals, researchers, and decision-makers by visualizing the **Individual Routing Preferences** of travelers within a network.

To accomplish this, every aspect that leads from the **Individual Routing Preferences** to an update of the **Network Graph** must be formally implemented. In the subsequent chapters, we introduce the underlying software architecture capable of fulfilling the requirements set by Figure 3.2. We split the subsequent software architecture into the subsystems described in Figure 3.2 (highlighted in different colors).

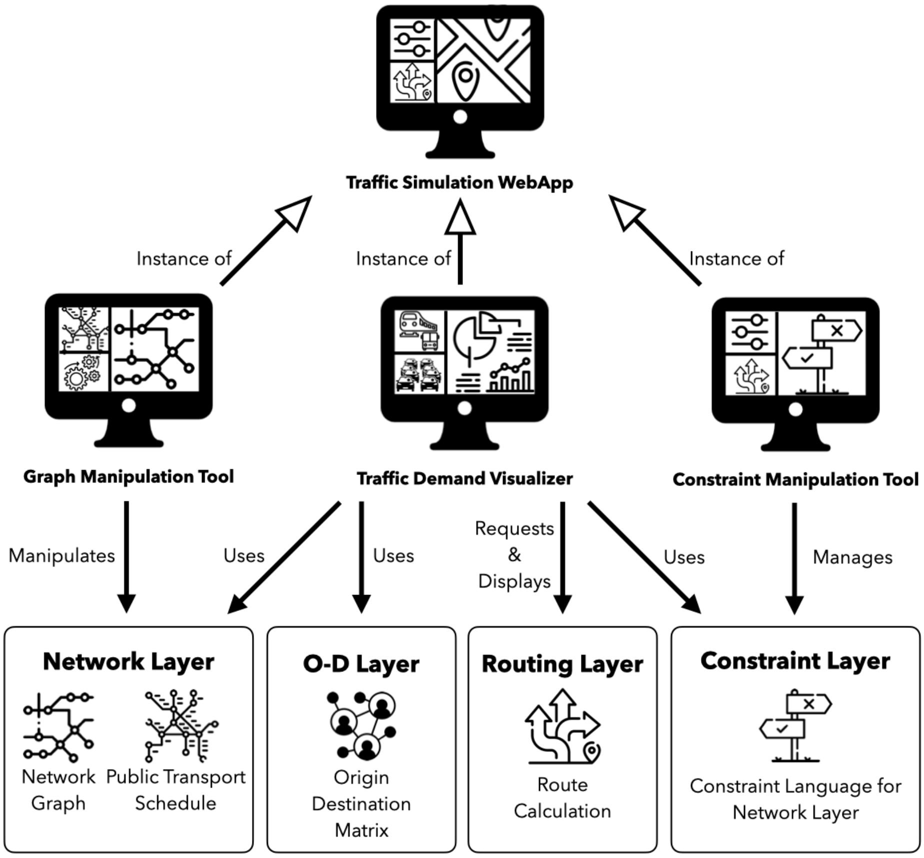


Figure 3.3: Overview of the different subsystems and support layers used by the Traffic Simulation WebApp.

In Figure 3.3, we demonstrate how these subsystems are connected to the three "sub-interfaces" that, when combined, constitute the simulation's GUI: the **Graph Manipulation Tool**, the **Traffic Demand Visualizer**, and the **Constraint Manipulation**

Tool. These, in turn, function based on information sourced from four distinct layers, namely the **Network Layer**, the **O-D Layer**, the **Routing Layer**, and the **Constraint Layer**. Everything is combined into a single user interface, which we term the **Traffic Simulation WebApp**. This interface enables devices equipped with web browsers to interact with and control the simulation tool.

To comprehend how the three user sub-interfaces operate, we will first address the four underlying layers.

3.1.3 The Network Layer

The **Network Layer** can be considered as the storage for available knowledge about the **Network Graph** (refer to Figure 3.2). If a Network Graph is constructed for a public transportation mode, then the departure times and the routes operated are additionally stored in the form of a **Public Transportation Schedule**. As illustrated in Figure 3.2, the Network Graph itself comprises nodes, edges, and weights. Nodes are connected via edges, and weights offer supplementary information about nodes and edges. A typical example of weight is the expected travel time.

To the navigation algorithm utilized by the traveler, whether it is in their heads or in an external device, edges and nodes constitute a geographic network that is defined as "known". As long as the traveler is geographically on this network, their routing algorithm remains fully operational. If the traveler departs from this network, they enter "unknown" territory. In these scenarios, any type of navigation will resort to estimating the general direction that the traveler should follow, hoping they will sooner or later encounter a part of the "known" network. This implies that the quality of a **Network Graph** depends on its information-completeness, indicating that every piece of available information that exists in reality is included in this modeled graph of reality. We choose the term information-completeness to create a clearer distinction from game theoretical completeness, which pertains to a different concept. Achieving information-completeness, however, is unattainable and will always be so, as acquiring this knowledge would necessitate an instantaneous connection to reality, which is unfeasible due to communication delays. Even if we were able to recognize all changes occurring in a network, we would still be hindered by the time the information needs from its occurrence to our observation and then to our publication of this change into the **Network Graph**. Even in the absence of processing delays, the flow of information is still restricted by the speed of light. This extreme example demonstrates that graph networks can never be an information-complete representation of reality. However, they can approximate information-completeness. This proximity to information-completeness is the quality measure by which the value of a graph network is determined.

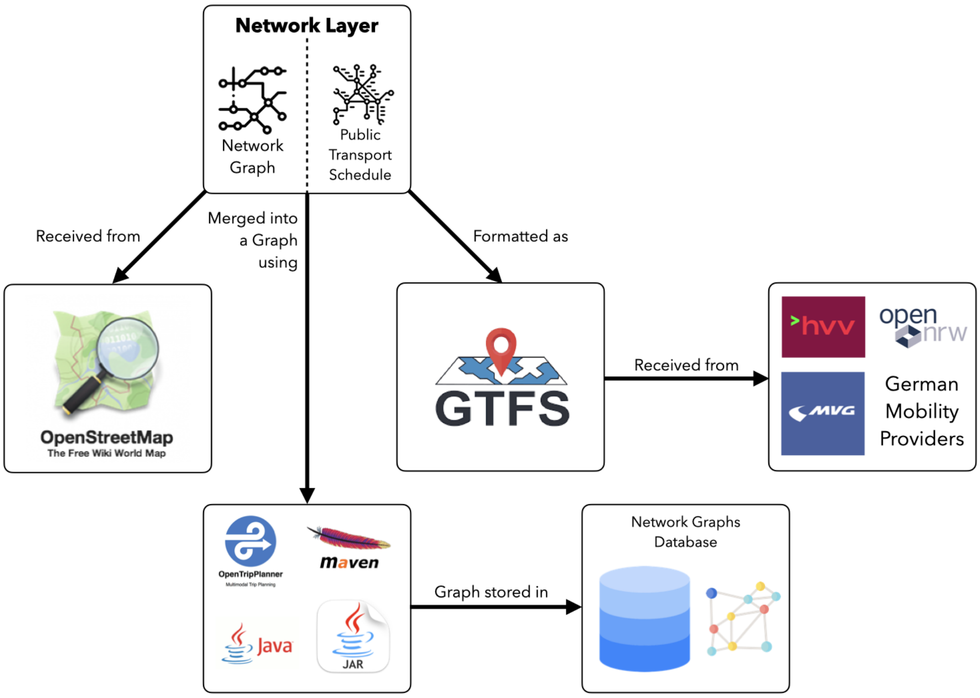


Figure 3.4: Graphical Model visualizing data sources and data formats used to build the graph file in the Network Layer.

Unless researchers have access to unique additional information sources, they are solely reliant on the best, publicly available information sources to construct a network graph. Therefore, to create our network graph, we utilize OpenStreetMap (Figure 3.4), as it is arguably the most information-complete publicly available graph. A Network Layer using the OpenStreetMap graph could, however, be improved if knowledge from local experts or data from companies such as Google or Apple were incorporated. In addition to OpenStreetMap, we also employ publicly available transportation schedules in the GTFS format published by German travel agencies. These schedules are utilized to generate public transportation network maps for major German cities, including departure and arrival times (Figure 3.4).

To construct a graph file from this data, we employ a combination of Maven[1] and Java[2], together with OpenTripPlanner[3]. The said Graph file is then provided to the Graph Manipulation Tool and the Traffic Demand Visualizer in the form of a .jar file[4]. Using the Graph Manipulation Tool also allows simulation operators to create modified versions of the graph file. These altered versions are stored, along with the original, unaltered

[1]https://central.sonatype.com Accessed: 15.11.2023
[2]www.java.com Accessed: 15.11.2023
[3]www.opentripplanner.org Accessed: 15.11.2023
[4]https://docs.oracle.com/ Accessed: 15.11.2023

graph file, in the Network Graphs Database. This database is subsequently used by the operator in the Traffic Demand Visualizer to simulate how a selected graph influences traffic behavior. The resulting user interface for operators of the WebApp is shown in Figure 3.5.

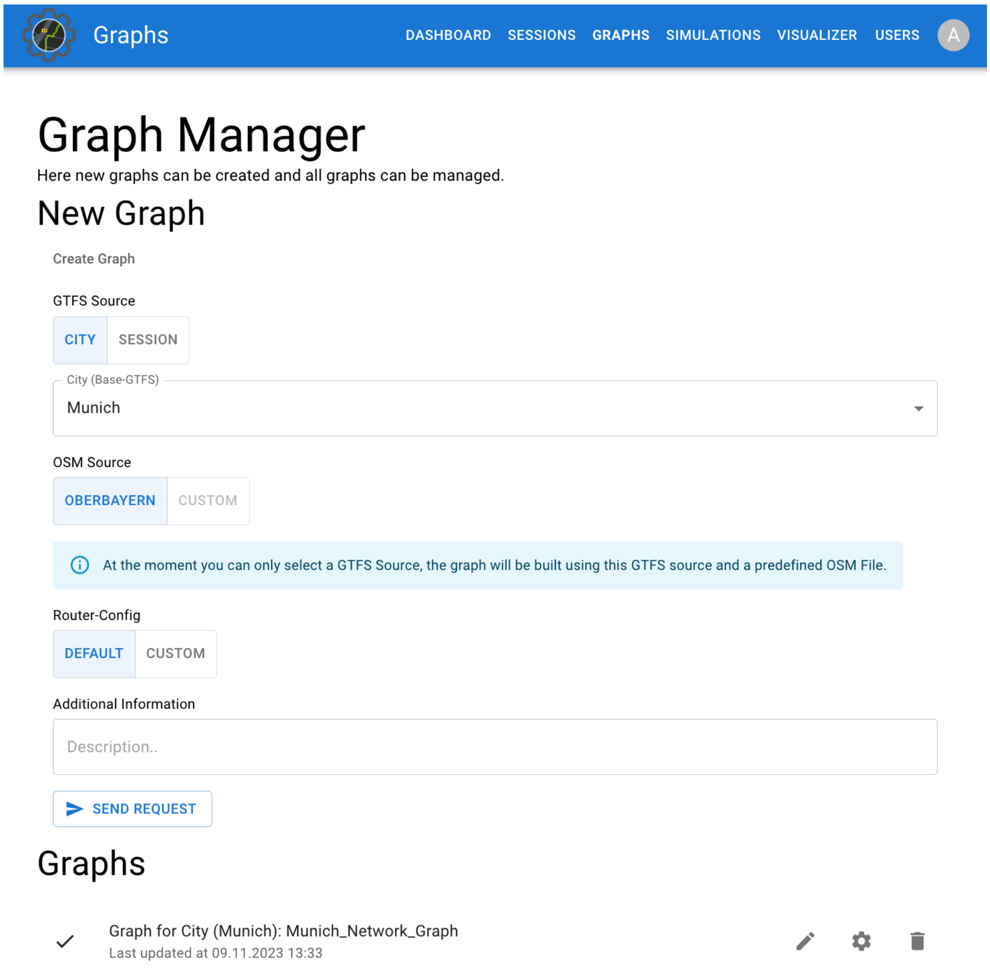


Figure 3.5: Graphical user interface for selecting a GTFS and OSM source. In combination with a router configuration, a network graph is built.

In addition to loading and using existing network layers, TrafSim allows operators to manipulate any uploaded GTFS files and add, edit, or delete public transportation connections. As GTFS files can be quite complex to understand, TrafSim provides a GUI, visualizing the existing content in a provided GTFS file. Along with the visualization, TrafSim also provides the previously mentioned functionalities. A small example snippet of this visualization is presented in Figure 3.6. While Figure 3.6 only shows the edit function for trips operating on Route "U6" in Munich, similar functionality is provided on the "Route" and "Stop" levels, giving operators the freedom to change and delete or extend any entries in the File. Due to the large number of entries, the GTFS-Visualizer section of WebApp is generally difficult to display in the format constraints of this dissertation.

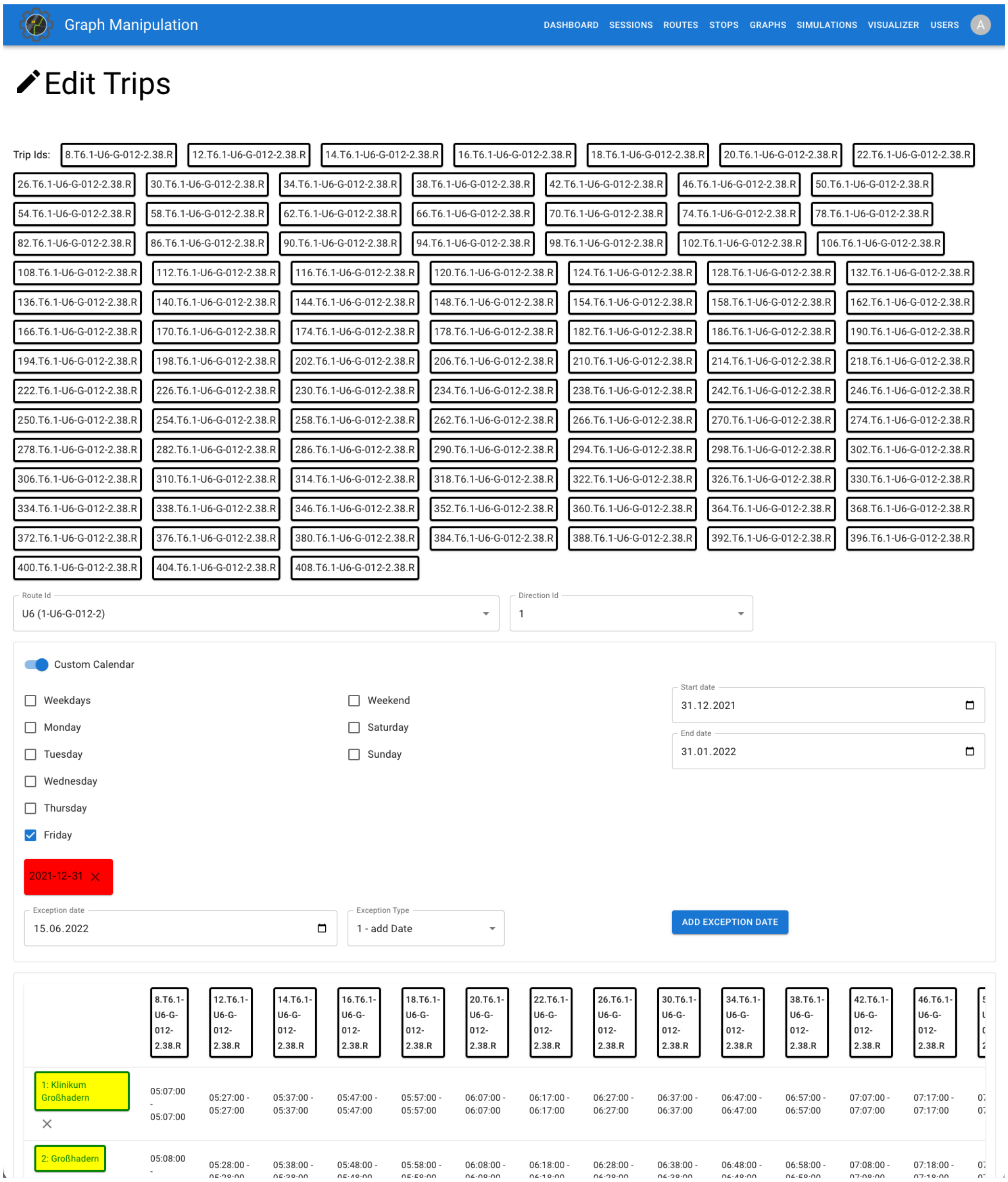


Figure 3.6: Example snippet of the GTFS-Visualizer, along with the edit functionality. Due to the large number of entries, this section of WebApp is difficult to display in the format constraints of this dissertation.

3.1.4 The O-D Layer

The foundation for any traffic simulation is a realistic representation of demand. This demand stems from travelers who traverse paths from an origin to a destination (O-D) at a specific time. However, our simulation tool is not designed to showcase conflicts causing congestion and delays. Instead, we simulate how travelers would utilize the traffic system if they could move freely without causing conflicts with other travelers. This is achieved by calculating routes determined by the travelers' **Routing Preferences**, introduced in Figure 3.2. The goal is to generate a representative data set of travelers that contains realistic travel origins and destinations, including their travel time and selected modes of transport. Additional meta-information about the travelers' optimization goals will be incorporated in the Constraint Layer.

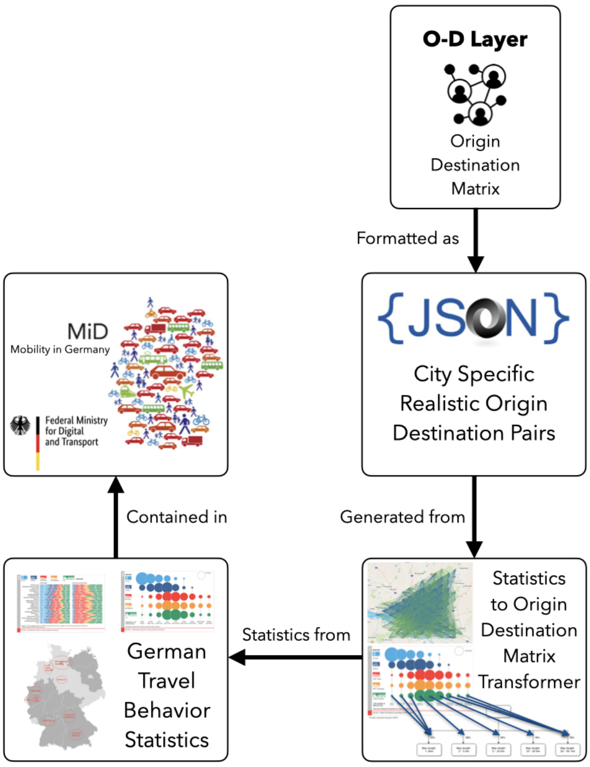


Figure 3.7: Model visualizing the concept behind the O-D Layer, describing the used data sources and format.

To construct the O-D Layer, we produce large origin-destination matrices (ODMs) for each city, commencing with realistic origin-destination pairs for a standardized day. Since our aim is not to simulate traffic but instead the absence of traffic, the O-D Layer does not delve into further detail. The core of these ODMs is a set of traffic data collated by the German Federal Ministry of Transport and Digital Infrastructure in a study known as the MiD-Study. For instance, in Munich, the study gathered 48,627 travel routes from 15,693 participants in the city alone, as well as 41,404 additional routes from 13,660 participants in areas surrounding Munich. The study is not a one-time effort; it is replicated periodically across various German cities [69, 70]. As the data collected in each city and in each cycle of the study is similarly structured, the methods utilized here could effortlessly be applied to other cities in Germany.

The deliberate absence of a comprehensive representation of travelers' mindset, physical condition, and current situation serves a purpose. The intention is to provide researchers the opportunity to test hypotheses by adjusting these variables in the Constraint Layer, which will be discussed in the following chapter. Publication 1 in Section 4.1.2 describes this process in greater detail.

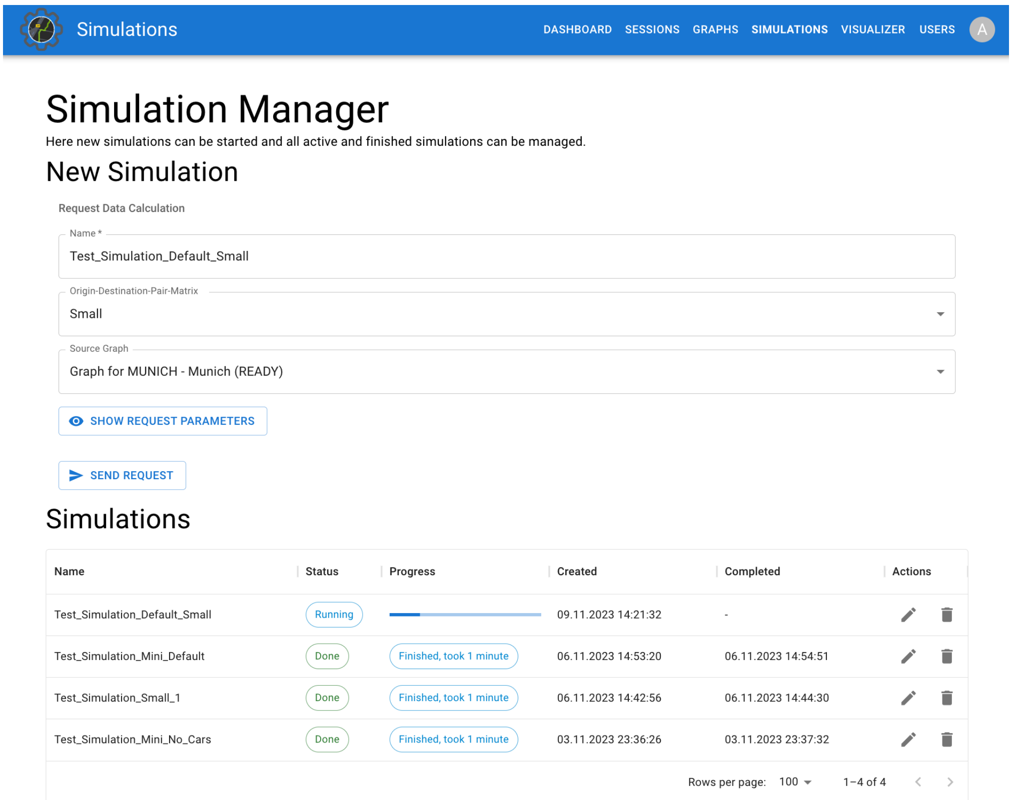


Figure 3.8: GUI for selecting an O-D Matrix, and a source graph, together with individual routing parameters called "constraints" (e.g. walking speed, biking speed, maximum walking distance, ...).

3.1.5 The Constraint Layer

In the preceding chapter, we noted that supplemental meta-information about the travelers' optimization objectives would be incorporated into the Constraint Layer. Ideally, this meta-information would encompass a comprehensive snapshot of the traveler's mental and physical condition, along with their current situation. Given that such detailed information is generally unavailable, we adopt a more streamlined approach. The Constraint Layer confers the responsibility of including this information to the researcher or traffic planner utilizing the simulation tool.

The sole aim of the Constraint Layer is to enable the simulation operator to inject hypotheses about the simulated travelers' condition (mental, physical, situational) into the routing algorithm, facilitating the examination of implications on the network. We term these settings as routing constraints, and their assembly in a set of constraints is referred to as a policy.

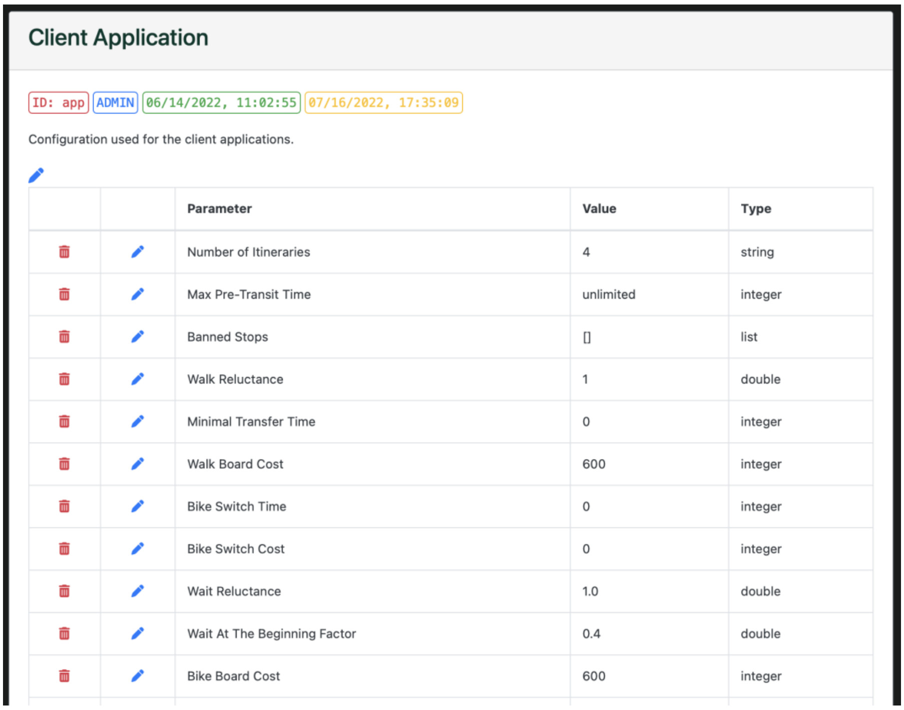


Figure 3.9: Snippet from the Constraint Manipulation Tool user interface, showcasing editable constraints in a policy termed "Client Application".

What exactly constitutes a policy? In our system, a policy entails constraints with customized settings tailored to suit a researcher's hypothesis concerning a group's travel behavior. Within this framework, a policy can also be perceived as a persona - a prototype character symbolizing typical behavior for a cluster of individuals. This enables simulation operators to assimilate the routing preferences of a particular group of individuals and convert them into a set of constraints. Figure 3.9 exhibits a snippet from the Constraint Manipulation Interface, presenting some of the adjustable routing constraints. The constraint selection is toggled by pressing the "Show Request Parameter" button in Figure 3.8. The duty lies with the person currently steering the simulation to bridge the gap between the routing preferences of a group and the presumed routing constraints they might utilize for navigation. An instance of a constraint is the walking speed assumed by the simulation's routing algorithm. Altering this variable permits operators to simulate the travel behavior of travelers with lower walking speeds, such as children or the elderly. The operators can subsequently scrutinize the simulation results to observe how the simulated travelers utilizing this policy interact with the transport network and how the network caters to their requirements. The purpose of the Constraint Layer is to facilitate the process of adjusting policies for the simulation operator.

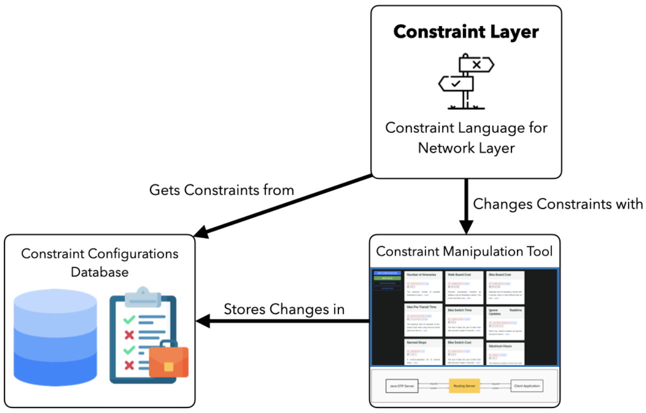


Figure 3.10: Model illustrating the concept behind the Constraint Layer, outlining the user interface and data storage.

Figure 3.10 offers a concise graphical synopsis of the Constraint Layer. The Layer delivers a graphical user interface designated as the Constraint Manipulation Tool. This interface enables operators to construct policies. These policies are stored in a central database to be employed by the other layers to buttress the Traffic Demand Visualizer (Figure 3.3), which allows simulation operators to simulate, display, and compare the impact of a policy on travel behavior.

3.1.6 The Routing Layer

While we have previously mentioned the Traffic Demand Visualizer (Figure 3.3) in the preceding chapter, we must take a moment to briefly describe the last key functional layer before diving into the intricacies of the Visualizer: The Routing Layer.

The Routing Layer, though relatively simple in design, plays a crucial role within the simulation tool. It calculates routes for each Origin-Destination (O-D) pair provided by the O-D Layer using OpenTripPlanner, effectively simulating the likely path a traveler might choose. In the route computation process, OpenTripPlanner considers the Network Graph and Public Transport Schedule from the Network Layer. Additionally, it integrates a chosen policy set, outlined by the Constraint Layer, to tailor the routes in line with these optimization goals. This architectural framework is graphically depicted in Figure 3.11. As the routing itself is a process run in the system backend, this process is not graphically visualized but instead triggered by pressing the "Send Request" button in Figure 3.8. The resulting paths are then stored in the "Simulations" table below the button.

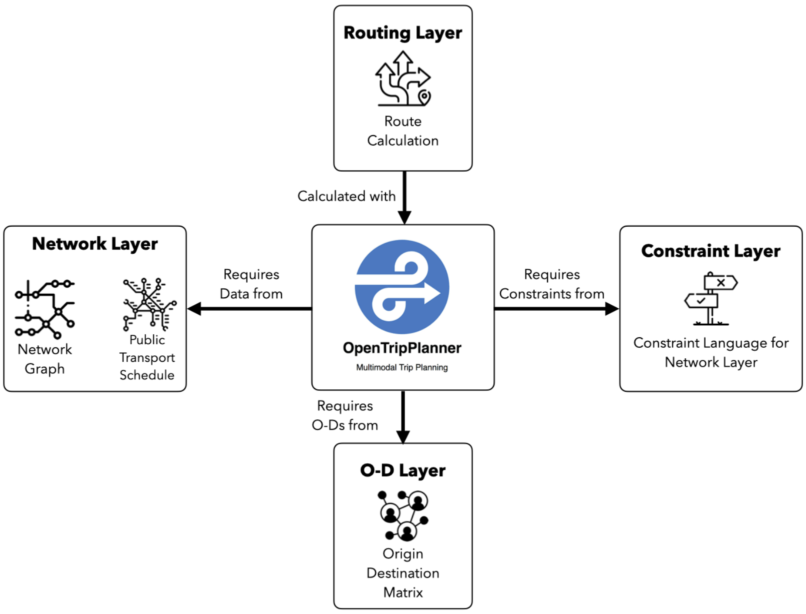


Figure 3.11: Model illustrating the architecture of the Routing Layer, inclusive of the utilized routing library and its integration with data from other layers.

Each simulation produces a list containing a summary of routing results. These may be compared to other simulations using TrafSim's compare function. The comparison can either be done in the tools web interface or by downloading a simulation CSV file to conduct comparisons manually. An example of a comparison using TrafSim's comparison interface is shown in Figure 3.12. Here again, we just show an example snippet of the entire functionality, as the comparison data is quite extensive and, therefore, visually separated into multiple tabs.

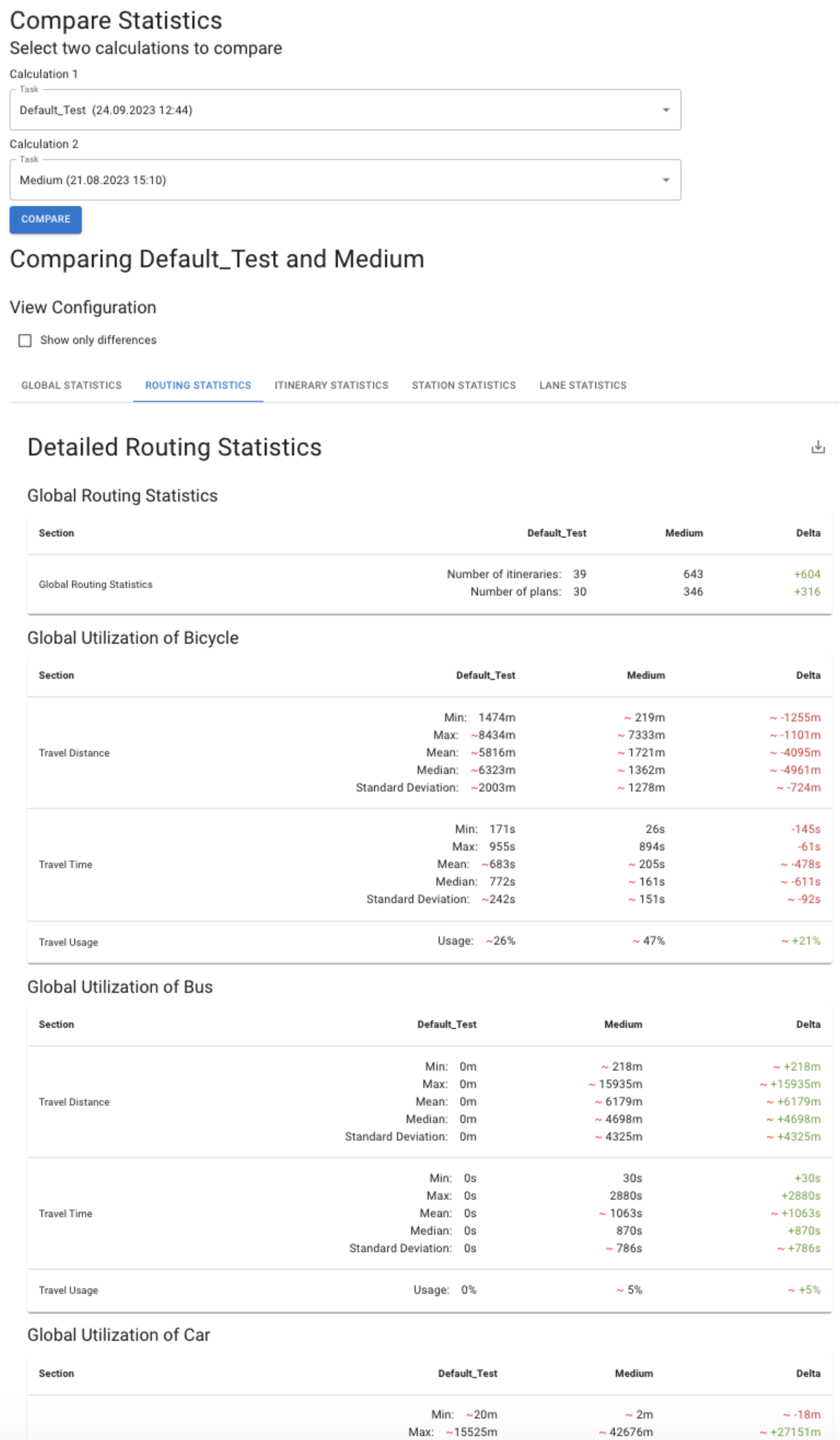


Figure 3.12: Screenshot of the Comparison View provided in the WebApp.

The comparison tool provides the functionality to compare the overall travel behavior between both simulations (”Routing Statistics” tab), as well as differences in singular travelers (”Itinerary Statistics” tab), for example, identical O-D pairs that were simulated in both simulations. If different Graphs were used in the compared simulations, these differences are also highlighted in the ”Global Statistics” tab. Differences in traveler utilization of lanes and stations between both simulations are displayed in the tabs ”Station Statistics” and ”Lane Statistics”.

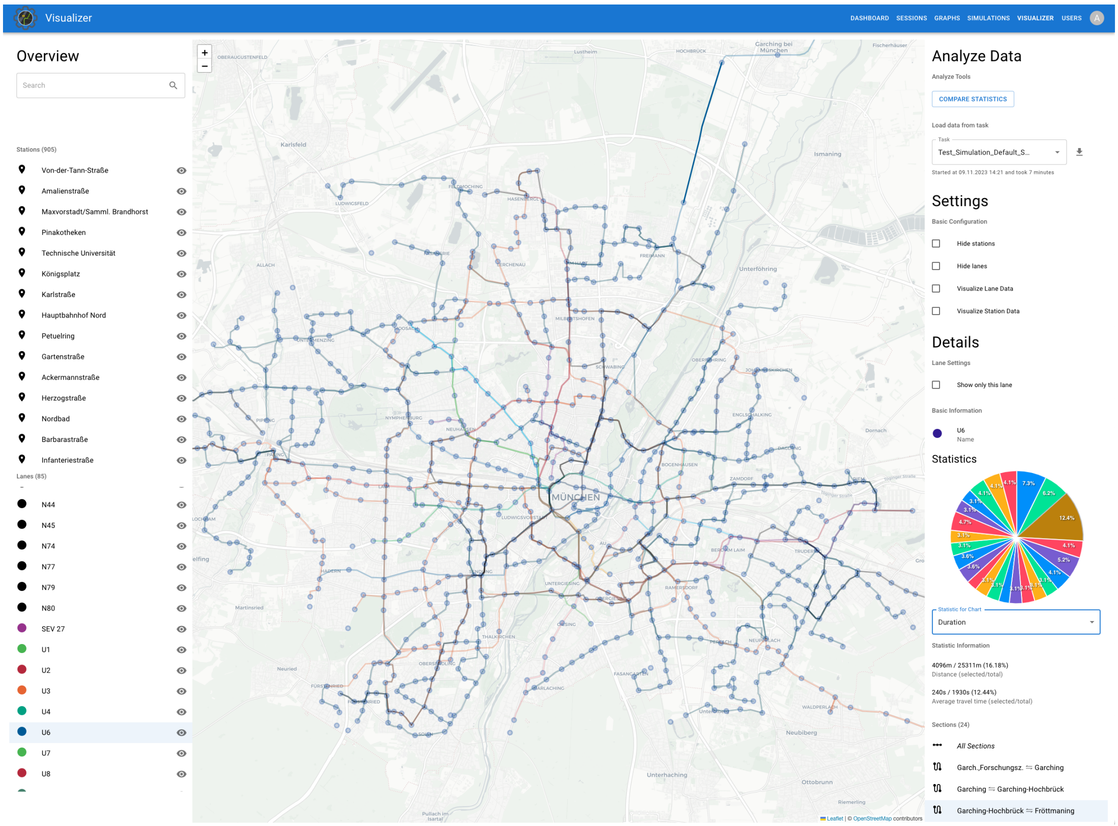


Figure 3.13: Screenshot of the Traffic Visualizer View provided in the WebApp.

In addition to a tabular comparison, a graphical visualization is provided, allowing simulation operators to visually compare the utilization of various public transportation routes in the network. For visual discrimination, operators may toggle lane, station, and utilization data on or off as needed. The GUI of this Visualizer is displayed in Figure 3.13.

3.1.7 The Traffic Simulation WebApp

In the architectural design of the Traffic Simulation WebApp, we adhere to the generalized concept of navigation, introduced in Figure 3.2. Figure 3.14 displays an updated version of Figure 3.2, delineating how the architecture's layers and subsystems align with the concept. For consistency, we've maintained the visual elements introduced in Figure 3.3. Nevertheless, we introduce a new element, the "Operator Hypothesis". This element indicates that the individual or group operating the simulation may introduce a hypothesis in this section of the model.

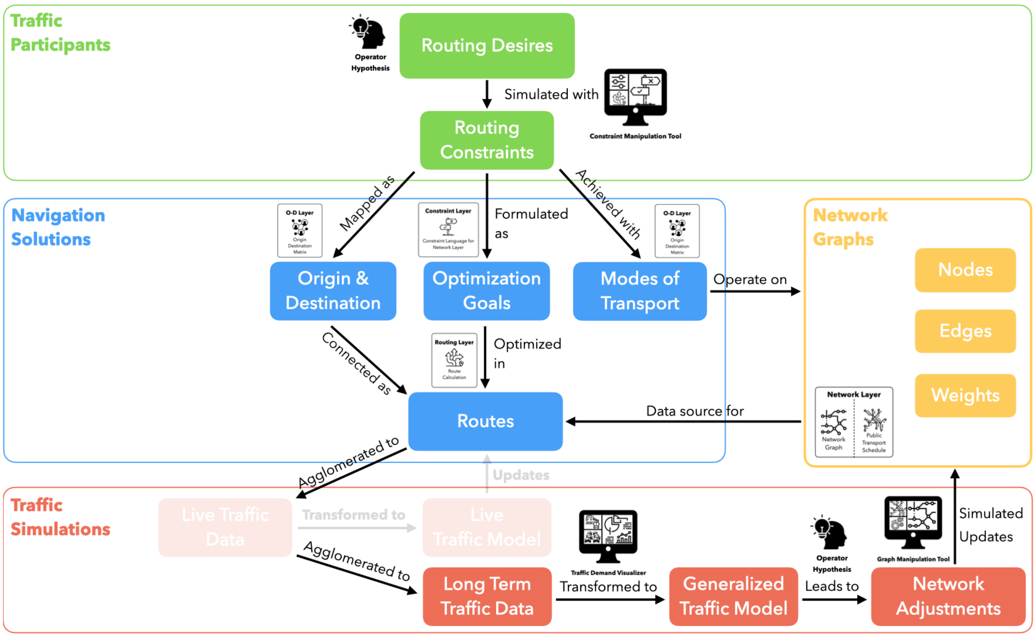


Figure 3.14: Generalized concept of navigation, encompassing human, automated, and distributed navigation.

This implies that commencing with the **Routing Preferences**, the simulation operator may introduce a hypothesis about these preferences into the simulation. Utilizing the **Constraint Manipulation Tool**'s user interface, this hypothesis is then emulated as a set of **Routing Constraints** named policy. The constraints contained within this policy are employed as **Optimization Goals** for route calculation on the **Routing Layer**. The **Origins and Destinations** are supplied by the **O-D Layer**, along with the **Modes of Transport** used for each O-D Pair in the ODM. To simulate the routes, a **Network Graph** is selected, which is supplied by the **Network Layer**. Subsequently, the **Routes** connecting each O-D Pair are simulated.

Given that we don't work with real-time observational traffic data but rather with data derived from travel behavior questionnaires, the Traffic Simulation WebApp does not offer **Live Traffic Data** or a **Live Traffic Model** that updates traveler routes. Since the MiD-Study only provides information about an "average day" with no further insights into travel times, we lack reliable data about travel times throughout the day. We also don't possess information about the capacity limits of edges and nodes in the network, which means we can't model when a capacity limit is reached and how this would affect traffic flow. As previously stated, the goal of this traffic simulation tool isn't to analyze live observational traffic data. Traffic simulations based on observational data excel at identifying and describing actual behavior in response to a network's design and traffic demand.

Our simulation tool bypasses the distinction between **Live** and **Long Term Traffic Data**. Hence, all generated **Routes** are directly conglomerated to form **Long Term Traffic Data**. This data is then presented to the simulation operator using the **Traffic Demand Visualizer**, who may use this information to incorporate it explicitly or implicitly into a **Generalized Traffic Model**. If done explicitly, then a formalized physical object (e.g., map, simulation, or other) is updated. If done implicitly, only the operator's own mental model of the network is updated.

In addition to, or instead of, hypotheses about travelers' **Routing Preferences** and **Routing Constraints**, the operator may also have hypotheses about the **Network Graph**. This signifies the second possibility for the simulation operator to introduce hypotheses into the simulation tool. The **Graph Manipulation Tool** provides an interface to the operator, allowing them to implement their hypotheses into the **Network Graphs**. The adjusted, or any other **Network Graph**, may be selected for subsequent simulation iterations, enabling the operator to test and compare their hypotheses against the simulated results.

4 Content

4.1 Publications

4.1.1 Publication 1, None-Core-Publication: A Practical Prototype

Table 4.1: Publication 1: Published.

Title	Introducing a Navigation Algorithm for Reducing the Spread of Diseases in Public Transport Networks
Authors	Jens Klinker (jens.klinker@tum.de) Mohamed Hechem Selmi (selmi@in.tum.de) Mariana Avezum (m.avezum@tum.de) Stephan Jonas (stephan.jonas@ukbonn.de)
Full Text	The full text of this publication is included in the Appendix in Section 8.1
Publication	dHealth 2021, Vienna, Austria
Status	Published
First Author Contribution	Problem definition, literature analysis, interpretation, management, result validation

Abstract: Reducing passenger flow through highly frequented bottlenecks in public transportation networks is a well-known urban planning problem. This issue has become even more relevant since the outbreak of the SARS-CoV-2 pandemic and the necessity of retaining safe distances between passengers. We propose an approach that allows to dynamically navigate passengers around crowded stations to better distribute the passenger load across an entire urban public transport network. New constraints are introduced into routing requests that enable the avoidance of specific nodes in a network. These requests consider walks, bikes, metros, subways, trams, and buses as possible modes of transportation. An implementation of the approach is provided in cooperation with the Munich Travel Corporation (MVG) for the city of Munich to simulate the effects on a real city's urban traffic flow. The impact on travel time was simulated, given that the two major exchange points in the network were to be avoided. With an increase from 26.5 to 26.8 minutes in average travel time, the simulation suggests that the time penalty might be worth the safety benefits.

Summary

Introduction: This paper was written during the Sars-Cov-2 pandemic. It uses this event to illustrate how travel behavior changed during this period. Due to the increasing importance of social distancing, there was a need to avoid certain stations, as they were prone to overcrowding and thus presented a higher risk of infection. In this context, a new navigation algorithm was developed and integrated into a simulation that allows for the simulation of avoiding (Avoidance) one or more specific stations. In collaboration with the Munich Transport Corporation (MVG), an implementation for Munich is presented, showing that the average travel time increased only slightly, while a second optimization variable - in this case, the risk of infection – was improved.

- **Case Study of Avoidance:** A case study is presented to demonstrate how the algorithm works. A trip in the Munich network is compared with and without the avoidance function during peak times. The route from Nordfriedhof to Theresienstraße is examined. The route calculated without the avoidance function uses the subway to Sendlinger Tor, and a line change is made there. The avoiding route, on the other hand, uses a bus. The travel time is extended by one minute.
- **Network-Wide Effect of Using Avoidance:** Avoiding a station affects both travel time and passenger distribution. The aim is to ensure that other stations are not overloaded. The analysis is based on 1,000 random routes in Munich. It is demonstrated that by avoiding Sendlinger Tor and Hauptbahnhof, the passenger load is distributed more evenly.
- **Impact of Avoidance on Travel Time:** Isochrones are used to visualize how avoidance affects travel time in the Munich network. Accessibility within 20 minutes remains almost the same. However, an increase in travel time for destinations 30 to 40 minutes away is observed.

Concluding Remarks: However, uncertainties remain, as the actual willingness of people to ignore safer routes and the actual increase in safety can only be estimated at present. Moreover, the results are based solely on data from Munich, so it is unclear whether they can be transferred to other cities. The biggest limitation is the generation of 1,000 random routes to illustrate a potential network effect of this behavior. Since travelers do not travel randomly but follow specific patterns, it is advisable to improve the realism of these routes.

4.1.2 Publication 2, Core-Publication: An Improvement in the Realism of the Data

Table 4.2: Publication 2: Published.

Title	Presenting a Statistical Approach for Transforming Standardized German Traffic Surveys into Origin-Destination Matrices
Authors	Jens Klinker (jens.klinker@tum.de) Joe Yu (joe.yu@tum.de) Mariana Avezum-Mercer (m.avezum@tum.de) Stephan Jonas (stephan.jonas@ukbonn.de)
Full Text	The full text of this publication is included in the Appendix in Section 8.2
Publication	2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST), Detroit, USA
Status	Published
First Author Contribution	Problem definition, literature research and analysis, interpretation, writing

Abstract: This paper presents a method for generating Origin-Destination Matrices (ODMs) for the city of Munich using traffic count data from a Germany-wide study conducted by the German Federal Ministry of Transport and Digital Infrastructure (MiD-Study). The results show that the data provided by the MiD-Study was correctly translated into an ODM, thereby providing an interpretable demand format for traffic simulations. Due to the consistent design of the MiD-Study, the approach is also applied to Hamburg and is extensible to 18 further cities and one city-state (Bremen) covered in the MiD-Study. The produced ODMs for Munich and Hamburg are accessible for researchers at: https://nextcloud.in.tum.de/index.php/s/gT48xDzT88YJGQK

Summary

Introduction: Changes in transport networks are costly. They bring financial burdens, worsen travel times, cause environmental pollution, and other inconveniences. At the same time, adjustments are necessary to meet the requirements of the network. To be able to represent the consequences of a change without causing the aforementioned costs, traffic simulations can be used. This paper supports traffic simulations by providing them with more realistic datasets. To interpret traffic data, such simulations require so-called Origin-Destination Matrices (ODMs). This paper deals with the generation of ODMs from Munich traffic count data. A dataset collected by the German Federal Ministry of Transport and Digital Infrastructure (MiD study) is used for this.

- **Data Basis:** The MiD study contains 48,627 travel routes from 15,693 participants in Munich and 41,404 travel routes from 13,660 participants from the surrounding suburbs.

- **Assumptions:** Since the study does not provide exact origin and destination information for individual journeys, assumptions were made. It is assumed that all journeys start and end at a residential address in Munich, with travelers eventually returning to their places of residence. This excludes potential intermediate destinations. Each journey must either start directly at a residential address or return to one.

- **Population Distribution:** The probability of a journey starting or ending at a residence is estimated based on the distribution of apartments in Munich's districts. The total number of daily trips in Munich is estimated at 4.8 million.

- **Forming Origins:** The number of journeys in the population (4.8 million) is halved (2.4 million) and then allocated proportionally to the city districts in Munich. The MiD study provides additional information about travel purposes, which are included in the analysis.

- **Adding Destinations:** To complete the first 240,000 journeys, only one destination is needed. Since the MiD study does not provide destinations, these are estimated based on the information stored in the incomplete journeys. For this, a journey distance for each trip is first required.

- **Calculating Distances:** Distance groups were reduced and adapted to travel purposes.

- **Determining Destinations:** For each journey, a search for suitable destinations is conducted. This is based on a mixed dataset that matches the travel purpose of the journey. The distance is calculated using an API provided by the OSRM project.

- **Creating Return Journeys:** After destinations have been assigned to all journeys, their reversals are created by swapping the origin and the destination to represent the return journeys, representing the other half of the total dataset.

Concluding Remarks: Overall, this approach provides a systematic way to create ODMs for Munich using specific assumptions and data from the MiD study. The method can also be applied to other cities in Germany where MiD studies have been conducted.

4.1.3 Publication 3, Core-Publication: An Approach for Further Validation

Table 4.3: Publication 3: Published.

Title	Improving GPS-Based Mode of Transport Detection in Multi-Modal Trips using Stop Analysis
Authors	Jens Klinker (jens.klinker@tum.de) Mariana Avezum-Mercer (m.avezum@tum.de) Stephan Jonas (stephan.jonas@ukbonn.de)
Full Text	The full text of this publication is included in the Appendix in Section 8.3
Publication	2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST), Detroit, USA
Status	Published
First Author Contribution	Problem definition, data collection, literature research and analysis interpretation, implementation, writing

Abstract: This paper presents an extension to existing GPS-based approaches for tracking modes of transportation in multimodal trips. The extension focuses on analyzing stops and mapping them to surrounding public transport stations in order to improve the accuracy of the mode of transport detection. The proposed method is evaluated using data from the city of Munich, resulting in a 17% improvement of the F1-Score, from 73% to 90%. It is applicable to any GPS-based mode of transport detection system to improve their accuracy potentially.

Summary

Introduction: This paper introduces an algorithm for the automatic detection of modes of transportation. This approach has been tailored specifically for the city of Munich but is expandable to other cities. The algorithm aims to capture GPS traces in the form of routes and subsequently segment these into various transportation modalities. This allows researchers to automatically track and categorize the chosen routes of travelers. Potential applications include the validation of ODMs (Publication 1) and primarily the validation of simulated routes.

- **Concept:** The paper extends existing transportation classification algorithms with the so-called "Stop Analysis." Here, data segments in GPS traces, in which little or no movement occurred are closely examined to derive information about the used mode of transport.

- **New Stop Features for Proposed Segments:** Five new characteristics are introduced, including the number of stops, duration of stops, and the number of stops identified as bus, tram, or train stops.

- **Assignment of Stops to Stations:** The paper initially clarifies that public transport stops are timeless objects possessing a geolocation, name, and type instead of a GPS position. To learn more about stops in a GPS trace, nearby stations are sought. These stations are retrieved from an OpenStreetMap database. A bounding box demarcates a geographic area within which stations are fetched. For Munich, this results in a list of 6,966 stations categorized into tram, train, and bus. This list is advised to be updated weekly, given the infrequency of station list changes.

- **Impact of Avoidance on Travel Time:** To efficiently filter this list and select only those stations proximate to a stop, a technique was developed that employs decimal degrees of latitude and longitude. This method employs a grid, where each cell varies in size due to its earthly position. The algorithm rounds both the list of stations and the stops to 3 decimal places. As a result, similar values are rounded to the same value. A challenge arises when stations in close proximity could be overlooked due to rounding. To counter this, all rounded cells immediately surrounding the stop cell are also considered. This ensures that all relevant candidates are considered and the number of candidates for which the distance needs to be calculated is reduced. Every station that passes the distance test conveys its type to the stop, determining its transportation modality.

- **Merging Data Windows into Segments:** This section demonstrates how adjacent data windows are fused into segments. The approach relies on labeling data windows as "Walking" or "Not-Walking" since a walking segment typically separates transitions between modes of transport.

Results and Concluding Remarks: The paper presents a new approach to recognizing modes of transport in multi-modal routes by utilizing stops and public transportation stations. The average F1-Score across all modes of transport was 90%, marking a 17% improvement over a previous approach by Avezum. A comprehensible decision tree has been developed, which could be applied in other cities. However, it is recommended to adapt the tree to account for the idiosyncrasies of other transport networks, especially in countries differing significantly from German infrastructure. The automatic labeling of GPS data by modes of transport has various applications, such as the automatic calculation of CO2-equivalents or for traffic simulations. Manual labeling by travelers is prone to errors, making the presented automated approach, which offers high accuracy, especially valuable. Nevertheless, it's recommended to tailor the algorithm to the specific region, ensuring high-quality and effortless data acquisition for traffic research and subsequent labeling of collected GPS data.

5 Summary of Results

The objective of this dissertation was to improve the individualization of travel routes, maximizing the benefits for travelers while minimizing the interaction costs. Additionally, it aimed to illustrate how individualization and changes in transportation offerings influence network effects and various travel needs.

Publication 1 outlines the theoretical foundation upon which this objective is based and addresses issues arising from our ability to navigate. It delves into a fundamental trade-off between computational ability, data availability, and their individual importance to travelers. It concludes that navigation before the age of navigation systems (Navigation 1.0) and navigation with the assistance of digitally connected devices (Navigation 2.0) are both insufficient in their own ways. The solution proposed is Navigation 3.0, which combines the data quality and computational capability of Navigation 2.0 with the incorporation of individual preferences in the route calculation from Navigation 1.0. This concept extends prior research, such as the commuter satisfaction model introduced by St.Louis et al. shown in Figure 5.1

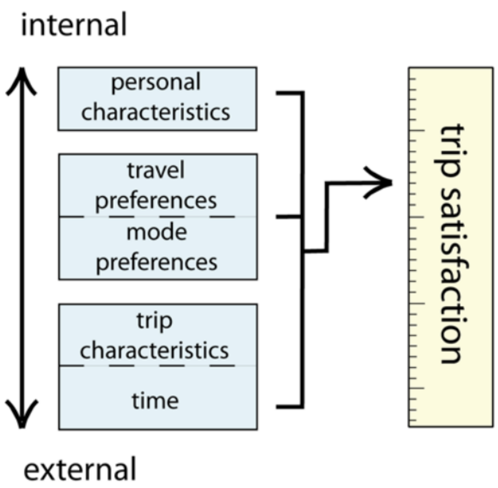


Figure 5.1: Commuter satisfaction model by St-Louis et al. [4]

Publication 1 demonstrates an initial approach to Navigation 3.0 by developing a Constraint Language for the open-source routing tool OpenTripPlanner. Using this Constraint Language, the study simulates how people's travel behavior changes during the Sars-Cov-19 pandemic when avoiding certain transportation hubs in Munich. 1000 routes for randomly selected O-D pairs were calculated, comparing the effects of avoidance on travel behavior to unchanged route calculations. The results are visualized in Figure 5.2.

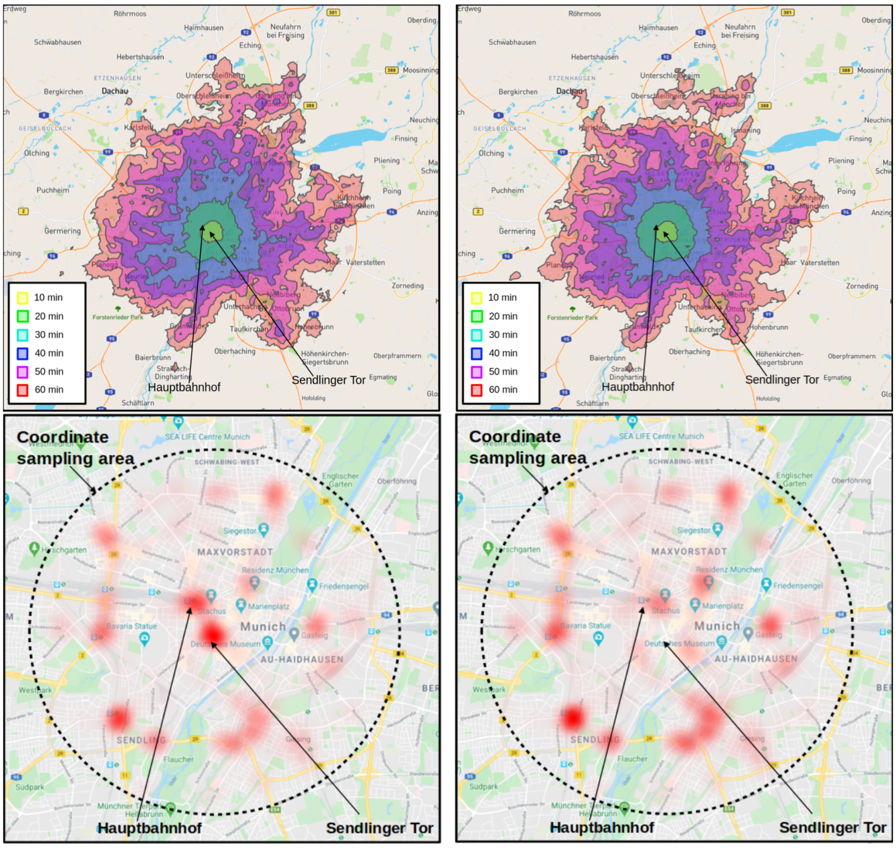


Figure 5.2: Compares the results of traveling with avoidance feature (right) and without avoidance feature (left). The above figures use isochrones to visualize the impact on travel time, while the below figures compare heat maps to visualize the impact on station utilization.

Publication 2 addresses the quality of the O-D pairs, which were randomly generated in Publication 1. Since travelers exhibit certain patterns due to factors like living and working locations, these aren't random. To improve this data and make it more realistic, Publication 2 uses a representative study for larger German cities. By statistically transforming the results of this study into an Origin-Destination matrix, the quality of traffic simulations can be made more realistic, but only for cities covered in the study. A visualization of this transformation process is summarized in 3.7.

Our traffic simulation prototype: TrafSim (Section 3.1), integrates concepts introduced in Publication 1 (customizable routing algorithm) and Publication 2 (more realistic ODMs) as components in a complete traffic simulation. It also introduces enhancements to these components, visualization systems, and a user interface. The foundation for this architecture is derived from a newly established navigation theory that is shown in Figure 3.2. It defines four basic systems required to represent navigation entirely.

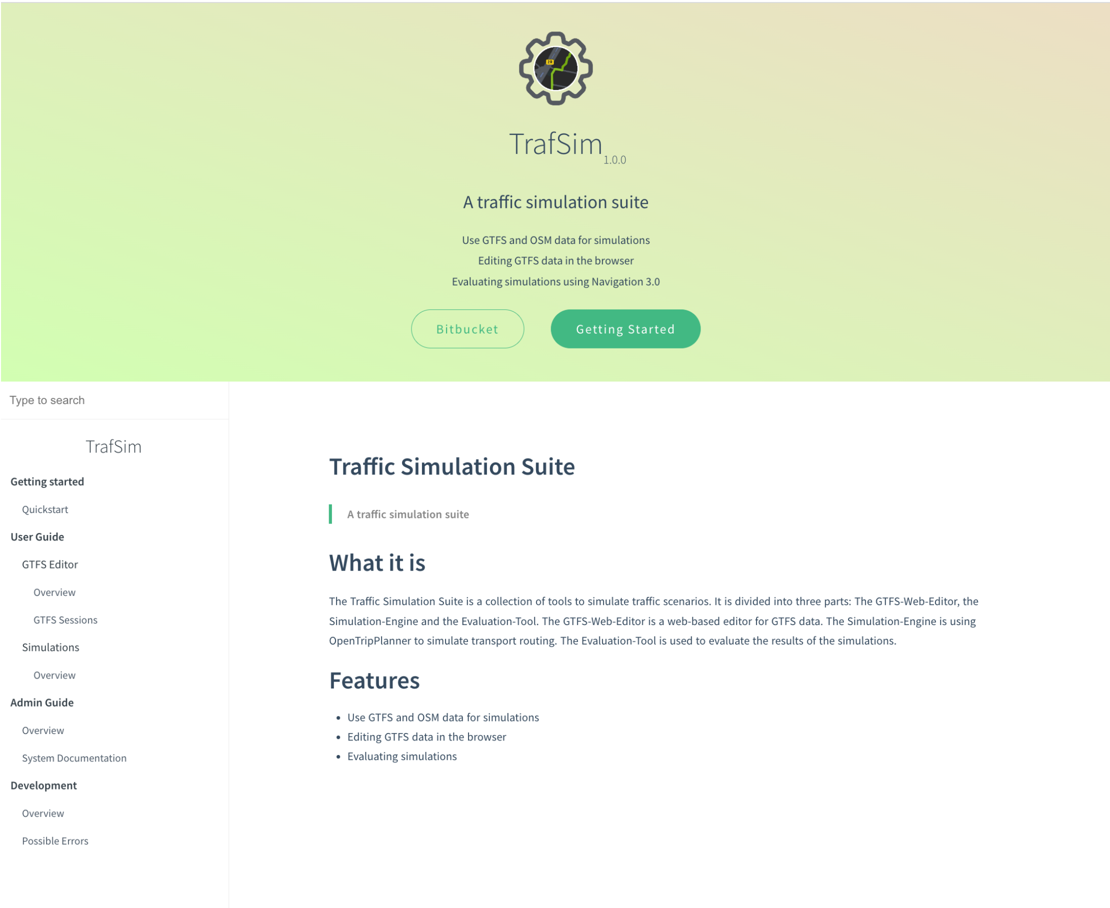


Figure 5.3: Documentation and onboarding page for setting up the Navigation 3.0 oriented Traffic Simulator TrafSim.

These basic systems are implemented as layers in the coherent traffic simulation. The traffic simulation's architecture is depicted in Figure 3.3. The lowest level consists of the fundamental subsystems that are described in Figure 3.2. The middle level provides user interfaces to interact with the below subsystems. Each middle layer provides a functional benefit to the user by itself. The top layer combines all user interfaces into one comprehensive Traffic Simulation WebApp, allowing users to manipulate the supply and demand of a traffic system and visualize the results of said manipulation.

Publication 3 provides an algorithm for validating simulation results within user studies. User studies capturing travel behavior involve significant effort from participants since they must track their routes via GPS and label them. Consequently, participants make labeling mistakes or quit the study due to its demands. The algorithm introduced in this paper achieves approximately 90% labeling reliability, surpassing the data quality produced by an inexperienced test group. Moreover, the algorithm is implemented in Swift and Python, making it compatible with both iOS and Android applications. The resulting precision values for detecting each mode of transport in GPS traces are shown in Figure 5.4.

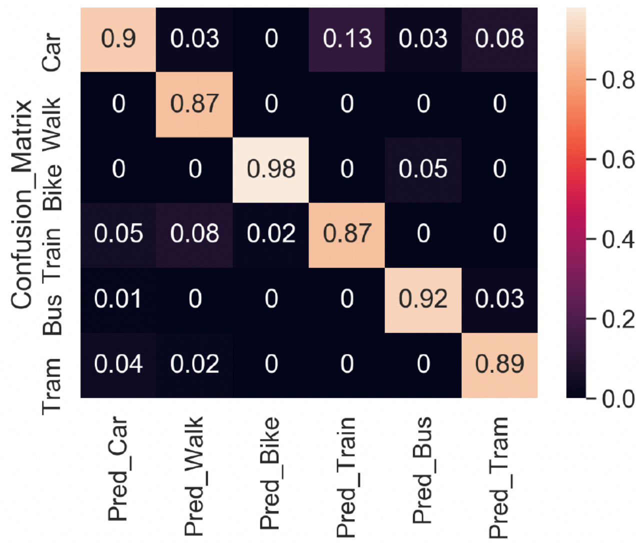


Figure 5.4: Confusion matrix showing the prediction precision in the diagonal (sum of column = 100). The respective recall values are, Car: 73%, Walk: 100%, Bike: 95%, Train: 85%, Bus: 96%, Tram: 94%.

5.1 Concise Answers

RQ1: How can a navigation algorithm be individualized?

A1: By developing a Constraint Language interacting with routing algorithms like OpenTripPlanner, route requests can be tailored more personally. Some functions are used as provided by OpenTripPlanner itself, while others are added to the algorithm either before or after the main processing to filter routes based on specific parameters. An example is incorporating a minimum or maximum walking distance as a routing parameter, enabling individuals with health goals to efficiently integrate them into their travel or those with mobility restrictions to minimize these distances.

RQ2: How can costs associated with improved individualization be minimized for travelers?

A2: Actively tracking route selection, travel speeds, and clustering situational behavior profiles as route options can reduce costs for users. To also protect the traveler's privacy, anonymizing techniques like Federated Learning are suggested, which perform this optimization locally on users' devices and only communicate the resulting model, not the original user data.

RQ3: How can the impact of such individualization on traffic planning and travel behavior be communicated?

A3: Traffic simulation is especially suitable for this approach. In Publication 1, we refer to this type of data communication as a "Level 3" solution, which is associated with the most development effort but achieves the maximum knowledge transfer. Knowledge transfer costs are borne by the simulation developer, not the information consumer or emitter.

6 Discussion

This dissertation demonstrates the core principles of scientific rigor. It begins by formulating hypotheses grounded in reasoned chains of arguments, a technique prevalent in humanities and theoretical research. These hypotheses are then tested through experiments designed to be repeated until enough data is collected. This data, assessed using statistical measures, helps determine the probability of observed patterns occurring by chance. The process involves creating experimental setups with tools for empirical validation, a practice typical in engineering sciences. These tools aim to reduce extraneous variables, enhancing the impact of the variables under study. The dissertation introduces the simulation tool TrafSim, designed to aid researchers in conducting experiments that validate both their hypotheses and the tool's effectiveness.

TrafSim is introduced as an initial prototype for simulating concepts of Navigation 3.0, including a foundational software architecture. The development process was primarily focused on internal testing and refinement, leading to iterative improvements in the system. External testing with operators has recently commenced, recognizing that this is an essential step in the evolution of any scientific tool. At this stage, TrafSim is not presented as a comprehensive and finished solution for individual travel navigation but rather as a prototypical framework subject to further improvements.

In line with common practices in research and development, TrafSim is in an ongoing process of optimization. The current limitations, including the presence of errors in simulation results, are primarily attributed to the system's recent inception and the extensive range of functionalities integrated at an early stage. This approach facilitated the rapid incorporation of diverse concepts, albeit with a trade-off in the granularity of detail. Future versions of TrafSim will focus on enhancing robustness and the refinement of these functionalities based on feedback from both internal and external testing.

6.1 Contribution to Research

This dissertation contributes to the scientific discourse by providing a simulation tool based on theoretical concepts described as Navigation 3.0. Providing a technical solution for conducting experiments is a common practice for advancing the scientific iterative process. For instance, the proof of gravitational waves, which Albert Einstein theoretically described in 1916, was only achieved 99 years later in 2015, as the development of a corresponding experimental setup to control for confounding variables was complex. In recognition of overcoming this complexity, Rainer Weiss, Kip Thorne, and Barry Barish received the Nobel Prize in Physics for this achievement. While this is an extreme case, it demonstrates that scientific contributions are rarely finished results but rather processes that are constantly refined and improved.

This dissertation provides two theoretical problem descriptions and pairs those with a testing environment for empirical validation. The first theoretical contribution encourages viewing traffic as a market-economic optimization problem, where travelers combine the benefits of individual actions with the advantages of machine calculation efficiency and data access.

The second theoretical contribution lies in the development of a "Generalized Concept of Navigation" (Figure 3.2), which systematically describes navigation behavior and thus provides a formal reference point for its implementation as a software system. Both principles are practically implemented, aiming to verify their correctness through experiments. Both concepts are reflected in the implemented traffic simulation. The market-economic theory forms the basis for adding the sub-levels "Graph Manipulation Tool" and "Constraint Manipulation Tool". This allows researchers to modify and simulate both the traffic system and travelers' needs. Additionally, this perspective led to the visualization of the simulation not focusing on fluid mechanics concepts of traffic flow but instead highlighting utilization. The Generalized Concept of Navigation was the basis for constructing the system architecture. It clarifies which aspects of a navigation process are represented by which subsystems. It also points out when an aspect is not or inadequately represented. Thus, it serves as an objective for the quality of navigation solutions since they can be compared against the ideal state described by the model. Lastly, we devoted ourselves to measuring actual travel behavior data to improve the associated data collection process. The goal was to facilitate the collection of travel data by automating the labeling process of collected GPS traces. One reason is the effort associated with collecting data for test subjects, and another is the reduction of labeling errors. The quality of the algorithm was experimentally confirmed for the city of Munich.

The fit of the traffic simulation to the requirements defined in the Generalized Concept of Navigation (Figure 3.2) is visualized in Figure 3.14. Even though most requirements set by the model were addressed, Live Traffic Data and Live Traffic models were excluded, as no live traffic data could be obtained. Although the addition of live data offers interesting potential, it also was of a lower priority for this traffic simulation. The reason is that it is not supposed to visualize current traffic behavior but rather general trends to aid in structural decision-making. While the simulation's code is not yet free of engineering mistakes, most have been addressed over many such iterative refinement cycles.

Currently, TrafSim is being tested by an external research group from the Chair of Automotive Technology at the TUM School of Engineering and Design. In collaboration with this research group, we are currently also initiating the first external validation of TrafSim by comparing simulation results to travel data in Munich. To optimize the testing process, different-sized ODMs for the city of Munich were provided along the simulation, containing 96 (small), 960 (medium), 9.600 (large), and 96.000 (extreme) representative origin-destination pairs for travelers in Munich. Depending on the complexity of a route and the number of selected routing constraints, it is able to calculate between 10 and 100 routes per minute using a Linux server with 16 GB of available RAM. This means that running a simulation roughly ranges between (format: dd.hh.mm) 00.00.01 - 00.00.10 (small), 00.00.10 - 00.01.40 (medium), 00.01.40 - 00.16.00 (large), 00.16.00 - 06.16.00 (extreme).

6.2 Limitations, Problems, and Assumptions

TrafSim still contains some unresolved issues and assumptions. For instance, the encoding of the start and end station in a GTFS file. Even though the GTFS file structure is standardized, the data contained inside this standardized structure is not. This causes hard-coded decoders to function on one provider's GTFS files but might fail on others simply because the data is decoded incorrectly. Currently, the start and end locations of certain public transport lines are retrieved manually from GTFS data files. These endpoints are required for visualizing the GTFS files in TrafSim. Our current solution extracts the start and end stations directly from the routes.txt file present in any GTFS file. However, the solution is generic, yet. It is not able to adapt to alterations in the data formatting inside a routes.txt file. To be able to work with custom formats from different GTFS data providers, we are still looking for a more generic and resilient solution that identifies the data formatting in a GTFS file and adapts to it.

Another limitation is that the default settings and data sets of TrafSim are currently focused on Munich and Bavaria. The used OpenStreetMap data (roads, intersections, sidewalks, ...) currently only contains information about the region of Oberbayern. To simulate other regions, they must be added to the database and selected during Graph building. However, simulation operators adding new data should be aware of the region's size. The more nodes and edges a region contains, the longer the graph-building process is going to take. Similarly, the GTFS data contained in TrafSim is from the municipality of Munich.

We are also aware of some technical issues that are still present in the current version of TrafSim. For example, docker images are not deleted when a graph is deleted by the operator, causing the hard drive to gradually fill up until a manual wipe of all docker images is administered. We are also experiencing a bug causing newly built docker images to initially generate 50 requests to the internal docker-utils server, unintentionally slowing down the graph-building process. Due to the complexity of the graph-building process and its many dependencies, we are unable to dash out precise error messages to the operator in the GUI of TrafSim, forcing us to display a rather generic error message like "Error in Graph building, please try again". If the issue persists, the operator is, therefore, forced to investigate the log files for debugging.

7 Outlook

The goal of this dissertation was to aid both individuals and collectives by providing a more comprehensive and less biased view of travel behavior. This issue arises as the constructional design and the policies that govern traffic are not based on representative statistics but on biased sampling. Even if a traffic decision is based on the structural load on the network, both the existing supply and the reduction of demand on singular variables have an impact on exerted behaviors. To address this issue, we implemented a layered architecture that allows individuals to describe their preferred travel behavior more closely. Simultaneously, we provide tools to test hypotheses about the traffic network to simulate their effect on different travel behaviors. This allows individuals and groups to showcase and compare the effect of a policy or constructional change to the network to their own exerted behavior, allowing them to express realistic and measurable concerns with said changes.

The simulation provides a first prototype along with a theoretical structure and a blueprint for designing this type of traffic simulation. To further improve the simulation, the next steps could include the possibility of altering the road network graph on top of the existing solution for GTFS-graph manipulations. While the addition of live traffic data is intriguing, it would most likely be reduced to validating the ODMs, as its live component has little practical use for this kind of traffic simulation. However, traffic simulations focusing on traffic flow would profit severely from such data. Adding more routing constraints to the existing list of constraints is also recommended to allow for further individualization of traveler preferences. This would be done by extending the existing constraint language with additional variables.

7.1 Technical Improvements

To improve the simulation's predictive power, the next steps could also revolve around improving the data quality. Currently, the ODMs only provide information about daily travel routines without discriminating by the time of the day. However, to showcase potential risks for congestion and other types of influences, it is also required to know at which times travelers are expected to travel. While we did find assumptions and estimates for travel densities in various cities, we could not determine which paths would be invoked at which times. As we could not pinpoint if adding these estimates would improve or harm the realism of the simulation, we had no ground for adding them to the ODMs. We generally advise sticking to this design concept and refraining from adding any additional discriminators if their positive impact on the predictive power of the simulation is not calculable. Economists have shown again and again that adding such "guesses" leads to more harm than gain. Instead, we advise simply treating such elements as what they are: unpredictable.

The routing layer is functional yet slow. While this can be solved with hardware, it is likely to be attributed to the increases in routing complexity due to the addition of constraints. An additional issue in this regard is the usage of OpenTripPlanner 1.x. While it does not pose any issues at the current point in time, it will likely cause them in the future. It is advisable to either use a custom-created routing algorithm to decouple ourselves from this dependency entirely or update the latest version of OpenTripPlanner 2.x. Both solutions pose similar threats. While decoupling creates independence, it also neglects the benefits of an open-source solution with multiple independent contributors from around the world. On the other hand, relying on an OpenTripPlanner creates a permanent dependency on an external piece of code that is placed in the heart of the traffic simulation.

Further validations against reality are required for all system components to improve their predictive power and test their resilience. Currently, the traffic simulation is biased towards German cities, particularly Munich. While Germany is a great example of the described problems of Navigation 2.0, we recommend extending the traffic simulation and its algorithms to other countries as well. To do so, the most important dependency to obtain are ODMs of other cities, as both GTFS-Data and traffic network data are easier to obtain. This is mostly due to the success of OpenStreetMap, providing open access to network data, as well as the worldwide dominance of the GTFS-Formatting standard for public transport schedules.

Finally, the quality of the constraint language is to be validated. This validation could be approached in various ways, either by interviewing experts, representatives of certain behavior groups (e.g., wheelchair users), or randomly selected individuals. The goal of these interviews is to identify ideal routing constraints for individuals and verify if the constraint language meets these requirements.

7.2 Digital Representation

One novel concept introduced in this dissertation is a trainable system to represent the user digitally. While most humans are being represented daily by other humans, such as politicians, friends, family members, or colleagues, digital representatives have hardly been explored. Interestingly, adjusting or training a digital representative, as demonstrated for routing algorithms in this dissertation, offers an unusual potential.

We discussed how even taxi drivers, who function as human champions of a network's various supply and demand streams, have resorted to utilizing automated support in the form of navigation devices. We also explained that even though these algorithms only offer limited customization, they are mostly preferred over the traveler's routing ability. This is due to the algorithm's ability to incorporate a much larger mass of information in their decision processes, leading to better-informed routing suggestions. This demonstrates how most humans prefer a well-informed decision over a highly customized one.

But what about a highly informed and highly customized decision? Here, our theory of Navigation 3.0 suggests that an optimal representative is both maximally informed yet also maximally customized to the traveler. As rising customization further reduces the number of individuals fitting said customization, maximizing both values will lead to a digital representative for each individual. At this point, the advantages of such a representative become quite clear. If the representative is fully informed and fully incorporates the preferred behavior of an individual, then the representative would be a perfect digitized copy of said individual, but with access to all advantages of digitized routing algorithms.

However, a truly perfect representation of an individual is impossible to reach. In Navigation 3.0 we describe how a training process to increase the adaption of said digital representative from Navigation 2.0 to the needs of the traveler further improves its value to them. However, it also translates cost structures from exponential to linear. Using TrafSim, travelers may participate in any number of traffic votes or decisions just by keeping their travel profile up-to-date. In TrafSim, their participation is automized simply by running simulations based on the traveler profiles that have been provided without adding the effort of voting. TrafSim also does not require travelers to inform themselves, as their behavior in accordance with the status of the network is similar to the traveler's actual behavior. This way, travelers do not have to compute the potential impacts of a traffic network change on them. They only need to describe their own behavior. So, while consuming all available information becomes an evermore complex task, as over time more and more information is accumulated, understanding our own desires may be a challenge, yet a linear one.

Hence, it is easier for humans to teach their digital representatives who they are, what they want, and how they behave than trying to compete with the digital representative's ability to consume and process information to form a decision. We thereby imply that for humans to navigate through an ever-growing landscape of information, they will have to rely more and more on digital representatives. Here, we draw a connection to any other type of navigation or representation currently existing. Geographical navigation through traffic systems is, in essence, no different from any other form of navigation through information. Search engines like Google or large language models like chatGPT are also navigators, using different forms of models for decision-making. While human representatives may hold an emotional importance, we see greater potential for digital representatives. We therefore deem it likely that the principles of Navigation 3.0 are applicable to other systems of representation, too.

TrafSim is an example of such a solution. It enables travelers of a transport network to create a digital representative of themselves. This representative can then be integrated into traffic simulations to more realistically determine which travelers would be affected in which way by changes in the transport network. This way, travelers can participate in any number of traffic decisions simultaneously simply by keeping their travel profile up-to-date. Compared to a system based on human representatives, systems like TrafSim are more informed and more representative while retaining a low level of effort for the individual traveler.

7.3 Closing Thoughts

This leads us to conclude that from a purely rational perspective, digital representatives outperform human representatives with further progressing digitization. However, the moral implications of this conclusion remain unresolved. We do not and can not answer if other aspects outside geographical navigation should be surrendered to digitized efficiency. But in the realm of geographical navigation, the transformation to digital representation seems irreversible. Bans on the use of existing navigation algorithms would likely meet resistance from the public, as they have become accustomed to their benefits. An attack on data storage in the form of bans is conceivable, but here, too, a contrary trend seems to be observable. Therefore, if both the data storage and the algorithms remain intact, a return to "Navigation 1.0" seems unlikely. Thus, the choice remains between Navigation 2.0 - efficient but impersonal - and Navigation 3.0 - efficient and personal. In this dissertation, we advocate for Navigation 3.0, which caters more closely to individual preferences and follows statistical principles. We show that including the supposedly irrational preferences of individuals using the principles of Navigation 3.0 may lead to an overall increase in travel satisfaction. We thereby demonstrate how multi-dimensional navigation can be used to utilize supposed irrationality to alleviate stress on the network.

Bibliography

[1] F. Ahmed, G. Rose, C. Jakob, and M. R. Hoque. Examination of clusters for better understanding commuter cyclists' travel behavior: insights from a longitudinal panel survey. *Transportation research record*, 2665(1):40–50, 2017.

[2] B. Ahmadov. Active commuting feedback: Evaluating the potential of sports-comparable metrics in promoting active lifestyles, 2024.

[3] J. Vestin. *SDN-Enabled Resiliency, Monitoring and Control in Computer Networks*. PhD thesis, Karlstads universitet, 2020.

[4] E. St-Louis, K. Manaugh, D. van Lierop, and A. El-Geneidy. The happy commuter: A comparison of commuter satisfaction across modes. *Transportation research part F: traffic psychology and behaviour*, 26:160–170, 2014.

[5] A. De Bary. *Die erscheinung der symbiose: Vortrag gehalten auf der versammlung deutscher naturforscher und aerzte zu cassel*. Trübner, 1879.

[6] E. Parmentier and L. Michel. Boundary lines in symbiosis forms. *Symbiosis*, 60:1–5, 2013.

[7] P. Chesson. Chesson's coexistence theory. *Ecology*, 101(11):1–7, 2020.

[8] D. Tilman, C. L. Lehman, and K. T. Thomson. Plant diversity and ecosystem productivity: theoretical considerations. *Proceedings of the national academy of sciences*, 94(5):1857–1861, 1997.

[9] I. A. Hanski, M. E. Gilpin, and D. E. McCauley. *Metapopulation biology*, volume 454. Elsevier, 1997.

[10] J. Bascompte and R. V. Solé. Spatiotemporal patterns in nature. 1998.

[11] F. A. Huntingford. *Animal conflict*. Springer Science & Business Media, 2013.

[12] H. Aarts, B. Verplanken, and A. Van Knippenberg. Predicting behavior from actions in the past: Repeated decision making or a matter of habit? *Journal of applied social psychology*, 28(15):1355–1374, 1998.

[13] J. A. Ouellette and W. Wood. Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological bulletin*, 124(1):54, 1998.

[14] D. L. Ronis, J. F. Yates, and J. P. Kirscht. Attitudes, decisions, and habits as determinants of repeated behavior. *Attitude structure and function*, 213:39, 1989.

[15] D. Parker, A. S. Manstead, and S. G. Stradling. Extending the theory of planned behaviour: The role of personal norm. *British journal of social psychology*, 34(2):127–138, 1995.

[16] B. J. Biddle, B. J. Bank, and R. L. Slavings. Norms, preferences, identities and retention decisions. *Social psychology quarterly*, pages 322–337, 1987.

[17] H.-W. Charng, J. A. Piliavin, and P. L. Callero. Role identity and reasoned action in the prediction of repeated behavior. *Social psychology quarterly*, pages 303–317, 1988.

[18] R. Richard, J. van der Pligt, and N. de Vries. Anticipated affective reactions and prevention of aids. *British Journal of Social Psychology*, 34(1):9–21, 1995.

[19] R. C. Tessler and S. H. Schwartz. Help seeking, self-esteem, and achievement motivation: an attributional analysis. *Journal of personality and social psychology*, 21(3):318, 1972.

[20] S. Roccas, L. Sagiv, S. H. Schwartz, and A. Knafo. The big five personality factors and personal values. *Personality and social psychology bulletin*, 28(6):789–801, 2002.

[21] D. A. Cobb-Clark and S. Schurer. The stability of big-five personality traits. *Economics Letters*, 115(1):11–15, 2012.

[22] J. d. D. Ortuzar and L. G. Willumsen. *Modelling transport*. John Wiley & Sons, 2011.

[23] J.-P. Rodrigue, C. Comtois, and B. Slack. *The Geography of Transport Systems*. Routledge, 2016.

[24] P. Hall and M. Tewdwr-Jones. *Urban and regional planning*. Routledge, 2019.

[25] M. Winters et al. How far out of the way will we travel? built environment influences on route selection for bicycle and car travel. *Transportation Research Record*, 2190(1):1–10, 2010.

[26] H. C. Borst et al. Influence of environmental street characteristics on walking route choice of elderly people. *Journal of Environmental Psychology*, 29(4):477–484, 2009.

[27] M. R. Moran, D. A. Rodríguez, and J. Corburn. Examining the role of trip destination and neighborhood attributes in shaping environmental influences on children's route choice. *Transportation Research Part D: Transport and Environment*, 65:63–81, 2018.

[28] H. S. Mahmassani. Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and spatial economics*, 1:267–292, 2001.

[29] M. Ben-Akiva and S. R. Lerman. *Discrete choice analysis: theory and application to travel demand*. MIT press, 1985.

[30] A. Nagurney and J. Dong. A multiclass, multicriteria traffic network equilibrium model with elastic demand. *Transportation Research Part B: Methodological*, 36(5):445–469, 2002.

[31] C. Liu, Y. O. Susilo, and A. Karlström. The influence of weather characteristics variability on individual's travel mode choice in different seasons and regions in sweden. *Transport Policy*, 41:147–158, 2015.

[32] V. Franken and B. Lenz. Influence of mobility information services on travel behavior. In *Societies and Cities in the Age of Instant Access*, pages 167–178. Springer, 2007.

[33] J. G. Wardrop. On the statistical distribution of traffic speeds. *Journal of the Royal Statistical Society: Series A (General)*, 115(2):202–208, 1952.

[34] M. G. Bell and Y. Iida. *Transportation network analysis*. John Wiley & Sons, 1997.

[35] T. Richter and J. Schlaich. Traffic psychology today: Review and future directions. *Transportation research part F: traffic psychology and behaviour*, 15(5):551–566, 2012.

[36] J. F. Nash. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.

[37] J. Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[38] A. Vetta. Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 416–425. IEEE, 2002.

[39] T. Roughgarden and É. Tardos. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259, 2002.

[40] R. G. Golledge. Path selection and route preference in human navigation: A progress report. In *Spatial Information Theory A Theoretical Basis for GIS: International Conference COSIT'95 Semmering, Austria, September 21–23, 1995 Proceedings 2*, pages 207–222. Springer, 1995.

[41] C. B. Musselwhite and I. Shergold. Examining the process of driving cessation in later life. *European Journal of Ageing*, 10:89–100, 2013.

[42] L. Hakamies-Blomqvist and B. Wahlström. Why do older drivers give up driving? *Accident Analysis & Prevention*, 30(3):305–312, 1998.

[43] R. A. Marottoli, C. F. M. De Leon, T. A. Glass, C. S. Williams, L. M. Cooney Jr, L. F. Berkman, and M. E. Tinetti. Driving cessation and increased depressive symptoms: Prospective evidence from the new haven epese. *Journal of the American Geriatrics Society*, 45(2):202–206, 1997.

[44] C.-Y. Tsai. The physical disabilities' travel behaviors. *African Journal of Business Management*, 4(4):512, 2010.

[45] J. A. Card, S. T. Cole, and A. H. Humphrey. A comparison of the accessibility and attitudinal barriers model: Travel providers and travelers with physical disabilities. *Asia Pacific Journal of Tourism Research*, 11(2):161–175, 2006.

[46] P. T. Costa Jr and R. R. McCrae. The five-factor model of personality and its relevance to personality disorders. *Journal of personality disorders*, 6(4):343–359, 1992.

[47] D. Jani. Big five personality factors and travel curiosity: are they related? *Anatolia*, 25(3):444–456, 2014.

[48] D. Jani. Relating travel personality to big five factors of personality. *Tourism: An International Interdisciplinary Journal*, 62(4):347–359, 2014.

[49] E. R. Dahlen and R. P. White. The big five factors, sensation seeking, and driving anger in the prediction of unsafe driving. *Personality and individual differences*, 41(5):903–915, 2006.

[50] C. Morar, A. Tiba, B. Basarin, M. Vujičić, A. Valjarević, L. Niemets, A. Gessert, T. Jovanovic, M. Drugas, V. Grama, et al. Predictors of changes in travel behavior during the covid-19 pandemic: The role of tourists' personalities. *International journal of environmental research and public health*, 18(21):11169, 2021.

[51] A. Terzić, B. Petrevska, and D. Demirović Bajrami. Personalities shaping travel behaviors: post-covid scenario. *Journal of Tourism Futures*, 2022.

[52] P. D. Prevedouros. Associations of personality characteristics with transport behavior and residence location decisions. *Transportation Research Part A: Policy and Practice*, 26(5):381–391, 1992.

[53] M. Bratić, A. Radivojević, N. Stojiljković, O. Simović, E. Juvan, M. Lesjak, and E. Podovšovnik. Should i stay or should i go? tourists' covid-19 risk perception and vacation behavior shift. *Sustainability*, 13(6):3573, 2021.

[54] F. M. Thomas, S. G. Charlton, I. Lewis, and S. Nandavar. Commuting before and after covid-19. *Transportation Research Interdisciplinary Perspectives*, 11:100423, 2021.

[55] G. Hardin. The competitive exclusion principle: an idea that took a century to be born has implications in ecology, economics, and genetics. *science*, 131(3409):1292–1297, 1960.

[56] S. Lancée, R. Veenhoven, and M. Burger. Mood during commute in the netherlands: What way of travel feels best for what kind of people? *Transportation Research Part A: Policy and Practice*, 104:195–208, 2017.

[57] S.-I. So and X. Y. Lehto. The situational influence of travel group composition: Contrasting japanese family travelers with other travel parties. *Journal of Travel & Tourism Marketing*, 20(3-4):79–91, 2007.

[58] M. R. Endsley. A taxonomy of situation awareness errors. *Human factors in aviation operations*, 3(2):287–292, 1995.

[59] M. R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human factors*, 37(1):32–64, 1995.

[60] G. Hardin. The tragedy of the commons. *Science*, 162(3859):1243–1248, 1968. URL: `https://www.science.org/doi/abs/10.1126/science.162.3859.1243`, `arXiv:https://www.science.org/doi/pdf/10.1126/science.162.3859.1243`, `doi:10.1126/science.162.3859.1243`.

[61] B. G. Malkiel. *Efficient Market Hypothesis*, pages 127–134. Palgrave Macmillan UK, London, 1989. URL: `https://doi.org/10.1007/978-1-349-20213-3_13`, `doi:10.1007/978-1-349-20213-3_13`.

[62] D. Helbing. Traffic and related self-driven many-particle systems. *Reviews of modern physics*, 73(4):1067, 2001.

[63] A.-L. Barabasi and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[64] M. F. Goodchild. Gis and transportation: status and challenges. *GeoInformatica*, 4:127–139, 2000.

[65] S. E. Page. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies - New Edition*. Princeton University Press, 2008.

[66] D. R. Montello. *Navigation*, page 257–294. Cambridge Handbooks in Psychology. Cambridge University Press, 2005. `doi:10.1017/CBO9780511610448.008`.

[67] R. G. Golledge. *Wayfinding behavior: Cognitive mapping and other spatial processes*. JHU press, 1999.

[68] J. M. Wiener, S. J. Büchner, and C. Hölscher. Taxonomy of human wayfinding tasks: A knowledge-based approach. *Spatial Cognition & Computation*, 9(2):152–165, 2009.

[69] B. Lenz, C. Nobis, K. Köhler, M. Mehlin, R. Follmer, D. Gruschwitz, B. Jeske, and S. Quandt. Mobilität in deutschland 2008. 2010.

[70] C. Nobis and T. Kuhnimhof. Mobilität in deutschland- mid: Ergebnisbericht. 2018.

8 Appendix: Full-Text Publications

8.1 Full-Text Version of Publication 1

Suggested Citation: Klinker, J., Selmi, M. H., Avezum, M., Jonas, S. M. (2021, May). Introducing a Navigation Algorithm for Reducing the Spread of Diseases in Public Transport Networks. In dHealth (pp. 113-121).

Status: Published.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210097

Introducing a Navigation Algorithm for Reducing the Spread of Diseases in Public Transport Networks

Jens KLINKER[a,1], Mohamed Hechem SELMI[a], Mariana AVEZUM[a] and Stephan JONAS[a]
[a] *Professorship for Digital Health, Technical University Munich, Munich, Germany*

Abstract. Reducing passenger flow through highly frequented bottlenecks in public transportation networks is a well-known urban planning problem. This issue has become even more relevant since the outbreak of the SARS-CoV-2 pandemic and the necessity for minimum distances between passengers. We propose an approach that allows to dynamically navigate passengers around dangerously crowded stations to better distribute the passenger load across an entire urban public transport network. This is achieved through the introduction of new constraints into routing requests, that enable the avoidance of specific nodes in a network. These requests consider walks, bikes, metros, subways, trams and buses as possible modes of transportation. An implementation of the approach is provided in cooperation with the Munich Travel Corporation (MVG) for the city of Munich, to simulate the effects on a real city's urban traffic flow. Among other factors, the impact on the travel time was simulated given that the two major exchange points in the network were to be avoided. With an increase from 26.5 to 26.8 minutes on the average travel time, the simulation suggests that the time penalty might be worth the safety benefits.

Keywords. Physical Distance, Social Control, Social Distancing, COVID-19 Pandemic, Risky Health Behavior, Outbreaks, Prevention and Control

1. Introduction

Conventional navigation systems like Google Maps offer a number of fastest or shortest routes to get to a destination. While these options were sufficient in the past, the outbreak of the SARS-CoV-2 pandemic led to a new requirement. Since then, travelers have also been looking for safe routes through cities, that could lower the risk of an infection with the virus. In the case of SARS-CoV-2, one effective measure to reduce this risk is to keep a safe distance to other humans [1, 2].
In Germany, the minimal required safety distance in public places was set to 1.5 meters by the Federal Ministry of Health [3]. Any distance below that threshold is considered unsafe from a social distancing perspective.

Fulfilling these social distancing requirements turns out to be a challenging task for passengers in urban public transportation networks. This is mostly due to overcrowded transportation vehicles and exchange stations during rush-hour times that force people to break social distancing rules. Relatively in the beginning of the outbreak, public

[1] Corresponding Author: Jens Klinker, Technical University Munich, Germany, E-Mail: jens.klinker@tum.de

authorities in Germany took action to prohibit any events with more than 100 people. However, when considering that the average Munich subway has capacity for over 900 people, any public transportation ride operating at over 11% of its total capacity can already be considered a large event [4]. This issue is even more critical at the central exchange points in Munich's public transportation network.

In exchanges with city planners form the Munich Transport Corporation (MVG), it was explained that congestions rarely occur at most public transport stations. Instead, there is a small number of stations in Munich’s public transportation network that get overcrowded reliably and predictably on a daily basis, due to their centrality in the network. Such stations, like “Sendlinger Tor” or “Hauptbahnhof”, that are mainly used to change between public transport lines, pose a constant infection risk. At “Sendlinger Tor” for example, an average of 250.000 passengers board, alight or change public transport every day [5]. While this is an unusual threat during a pandemic, it is a constant threat to people in high-risk groups.

To solve this issue, new solutions have to be implemented that reduce the passenger load on central traffic nodes, so that safety distances can be established again [6]. The goal of such an optimization problem is to fulfill the safety requirement while causing the minimal necessary impact on other travel goals such as minimizing overall travel time to a destination.

2. Methods

To tackle this issue, we introduce a new routing algorithm that offers a safer route to a traveler’s destination by avoiding highly congested stations in the public transport network. As the congestion level of a node is not constant, but depends on a large number of factors, a collaboration with the Munich Transport Corporation (MVG) was established. Using insights provided by the MVG, public transport stations can be dynamically toggled on and off before they reach a dangerously high congestion level in which safe distances can no longer be maintained. Currently, the assessment whether a station is considered safe or overcrowded is delegated to the MVG. To make this assessment as transparent and scientific as possible, it is advisable to take a closer look at parameters influencing the spreading risk in public transport stations. Due to its complexity, we haven’t included this additional step in the scope of this publication.

The avoidance algorithm then uses the information about the available and unavailable stations to calculate a safer route with less congested exchange points, giving the control for social distancing back to the travelers. To simulate the impact on the travel time and other important features, the MVG provided a dataset containing the scheduled departure and arrival times of all public transport lines in Munich. Both the data and the congestion knowledge are used to simulate the results produced by an avoidance algorithm for safer routes.

To evaluate the performance the avoidance algorithm is compared to the results of the same algorithm without the avoidance feature. Core features like the change in travel time are then used to measure how the convenience of a route is impacted by the feature. As the convenience of a route is highly subjective, no final measure for the convenience could be produced. Instead, some underlying factors that partially correlate with a route’s perceived convenience are presented.

As the proposed solution was developed with data from- and in cooperation with public authorities in Munich, the researched results and examples use cases are limited to the city of Munich.

2.1. Foundations

Public transportation has long been suggested as an alternative for road bottlenecks. During rush-hour, however, these are already overcrowded, and some cities thus prefer to limit passenger flow in bottleneck stations. In order to motivate passengers to consider going around an overcrowded station, two main reviews are presented: (1) Providing passengers with personalized multi-modal routing options and (2) calculating routes which have further advantages than a fastest travel time.

Review (1) is based on the work of Bucher *et al.* who describe how to integrate personal constraints on calculation of multi-modal routes [7]. To incorporate more travelling options into the multi-modal route calculation, the authors save a series of constraints in the form of a user profile.

As far as multi-modal routing goes, the state-of-the-art implementation is provided by OpenTripPlanner [8]. The open-source project uses public transportation schedules formatted as General Transit Feed Specification (GTFS) to calculate multi-modal routes that, among others, can combine cycling and public transportation segments. Since the consideration of multi-modal routes can be seen as an interesting approach to reduce the load of overcrowded public transportation intersection, OpenTripPlanner contains a "bannedStop" parameter that can be passed through a route request.

While the calculation of routes is a technical challenge, actually ensuring that the passengers use that route is a social one. Review (2) mentioned above relates to providing route advantages other than simply travel time. The basis for this is the work of Bunds *et al.* who analyzed how different route attributes were perceived by travelers and how these affected the choice of route [9]. Bunds et al. showed how air pollution, traffic, and noise level are the determining factors when deciding which route to walk through [9]. This allows us to infer that if a traveler can control for these variables during a route calculation, they can be more likely to accept longer walking segments. Furthermore, the fact that traffic and noise levels are directly correlated with overcrowded regions (even in public transportation), means that presenting this information to the traveler can serve as an important motivation for them to avoid these regions.

2.2. Constraint-based route personalization

OpenTripPlanner calculates multi-modal routes in the network based on a routing request that consists of a list of query parameters. In order to define the context of the route search, the request must specify the following information:

- **fromPlace:** Latitude and longitude of the start location.
- **toPlace:** Latitude and longitude of the end location
- **date:** Date on which the trip should depart.
- **time:** Time when the trip should depart.
- **mode:** Set of modes that a commuter is willing to use. The main modes supported by the system are walk, bike, car, and transit (buses, trains, trams).

In addition, the existing system supports multiple optional parameters that can be used to further manipulate the results. The most useful parameters for the problem at hand are:

- **bannedRoutes:** A comma-separated list of banned public transportation lines.
- **bannedStops:** Banned stations cannot be used to board or alight from a public transportation mode, but it is still possible to travel through them. This is achieved by blocking the pre-board and pre-alight edges that connect the transit network to the street network.
- **bannedStopsHard:** Stations that are removed from the network. It is no longer possible to board, alight or travel through these stations.

While the bannedStops parameter realizes the avoidance that the system aims to achieve for crowded stations, the time aware usage of this parameter to automatically avoid the stations during rush hour periods is still to be implemented.

2.3. Constraint integration in multi-modal routing

Similar to most state-of-the-art route planning services, OpenTripPlanner uses the A* algorithm to search for routes in the transportation network [10]. This algorithm keeps track of an ordered list of tentative routes and during each iteration the one with the smallest weight is extended. In order to achieve multi-modal routes, the algorithm is modified to loop over the available transportation modes during each iteration. For all outgoing edges of the last node, each mode that matches the type of the edge is used to traverse it. For instance, edges from the street network can be used for walking, biking, or driving, whereas edges of the transit network are restricted to a specific public transportation mode.

3. Results

3.1. Example use case

The following use case exemplifies how the avoidance algorithm calculates routes. We consider a trip in the Munich transit network and compare the route generated by the system with and without the automatic avoidance during rush-hour periods. For this demonstration, the experts from the Munich Transport Corporation (MVG) recommended to use the Sendlinger Tor, as it is one of the most congested stations in the network. In this case the routing request sent to the trip planner could, for example, use the following parameters:

- **fromPlace:** Nordfriedhof station with latitude 48.17312 and longitude 1.59686.
- **toPlace:** Theresienstraße station with latitude 48.15139 and longitude 11.56444.
- **date:** May 5th 2020.
- **time:** 08:00 am.
- **mode:** Walk, transit (buses, trains, trams)

This trip starts near a dorm for students and ends at a station used to access the technical university in the city, which makes it a realistic trip that students take on a daily basis. The time of the request is within the morning rush-hour period from 07:00 am to 09:00 am.

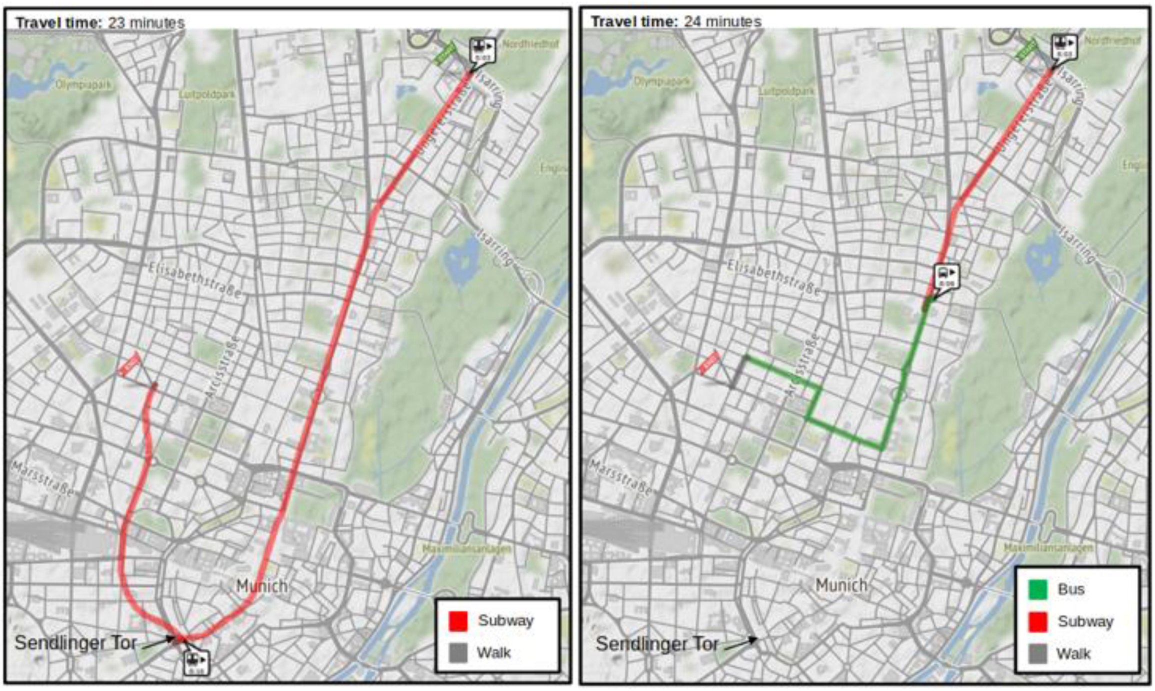


Figure 1. Route from Nordfriedhof to Theresienstraße. Left: Without automatic rush-hour avoidance. Right: With automatic rush-hour avoidance.

The route generated without avoidance can be seen in Figure 1 (left). It uses the subway from the origin, marked with a green flag, to get to Sendlinger Tor station. From there, a different subway line is used to get to the destination marked with a red flag. Including the time-aware avoidance results in a route that successfully avoids changing lines at Sendlinger Tor station as shown in Figure 1 (right). The new route cuts the subway part short before reaching banned stations and instead uses a bus to get to the destination. With activated avoidance, the travel time increases by one minute from 23 to 24 minutes.

3.2. General effect of avoidance on line changes

In this section we analyze how avoiding one station would affect the travel time and passenger distribution. The goals are to avoid overcrowding other stations in an attempt to scatter passengers to other stations and to maintain a reasonable travel time. To do so, we analyze the number of routes that use these stations to board, alight, or change lines. This can be considered as an estimation of the number of passengers at the stations. The analysis is based on a set of 1000 random routes located in the city of Munich. For each route, coordinates for the origin and the destination are sampled from an area centered at Sendlinger Tor with a radius of 4 kilometers. In addition, the following parameters are used for all routes:

- **date:** May 5th 2020
- **time:** 08:00 am.
- **mode:** Bicycle, walk, transit (buses, trains, trams)

The coordinates of the stops used to **board**, **alight**, or **change** lines are then extracted from each route and used to generate two heat maps. The first heat map shown in Figure 2 (left) summarizes the routes where no stops were avoided. The heat map in Figure 2 (right) considers the routes where Sendlinger Tor and Hauptbahnhof were automatically avoided during rush hour periods.

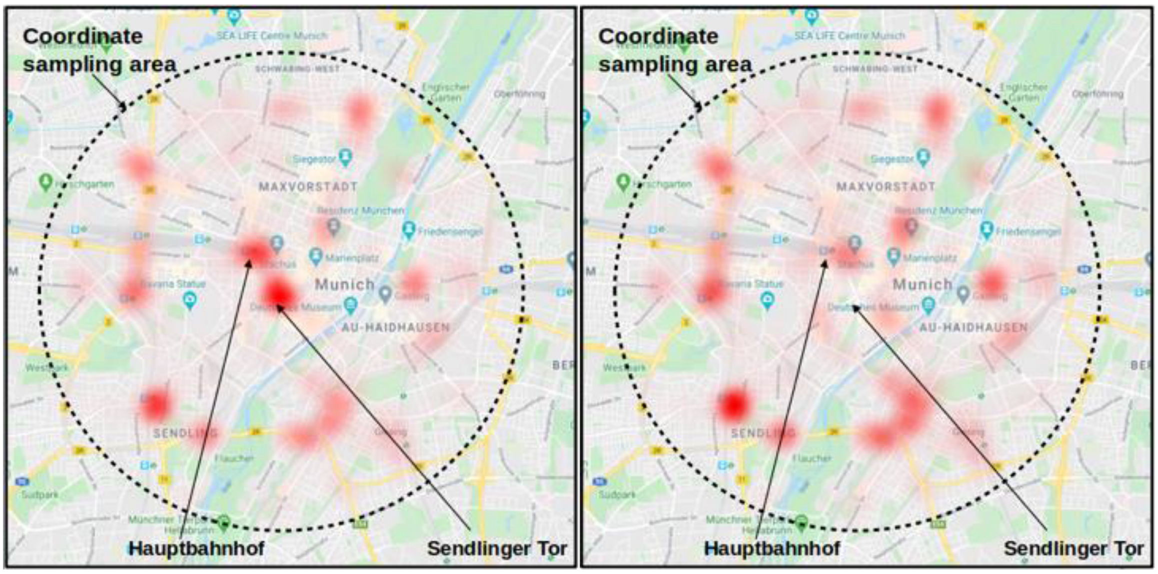


Figure 2. Heat maps showing how often transit stations are used to board or alight public transportation vehicles based on 1000 random routes. Left: Without activated avoidance feature. Right: With activated avoidance feature.

The left image in Figure 2 visualizes how Hauptbahnhof and Sendlinger Tor are originally the most used stations in the network, indicated through the dark red color of theses hot spots. While some areas on the outskirt of the network have similar size and darkness, they represent stations where commuters mostly board or alight a transit line but do not change them. With the automatic avoidance feature activated, the heat signature at the Sendlinger Tor and Hauptbahnhof both brighten up, while the direct areas around them darken slightly. This effect visualizes the distribution of the passenger load from these two stations to other nearby stations in the network. Instead of overcrowding a new station in the network, the load was distributed rather evenly over multiple surrounding stations located in close vicinity of the banned ones.

Aside the distribution of line transfers, the avoidance did not have notable effects on other route characteristics. The mean biking distance per route rose from 959 to 976 meters and the average walking distance slightly shrank from 469 meters to 468 meters. The mean waiting time at public transportation stations also remained unchanged at around 3.8 minutes per route. The mean runtime required to calculate a route decreased from 4.7 seconds to 3.7 seconds when the avoidance was activated. Finally, we consider the travel time of the routes used for the creation of the heat maps. The mean travel time increased from 26.5 minutes without avoidance to 26.8 minutes with avoidance.

3.3. Effect of avoidance on travel time

In this section we use isochrones to visualize the effect of the avoidance on the travel time in Munich's traffic network. Isochrones are graphs that measure location reachability from a specific origin. They consist of curves with equal travel time (Figure

3). OpenTripPlanner provides a service for generating isochrones out of the box by sending a request to the system similar to how routes are generated. In Figure 3 Sendlinger Tor was picked as the origin parameter for calculating the isochrones.

The generation of an isochrone starts by calculating a shortest path tree. OpenTripPlanner then builds a regular grid of samples covering the whole shortest path tree area. Finally, the sample points are connected based on their travel time to form the curves of the isochrones. The isolines are computed with the help of the Delaunay Triangulation Algorithm [11].

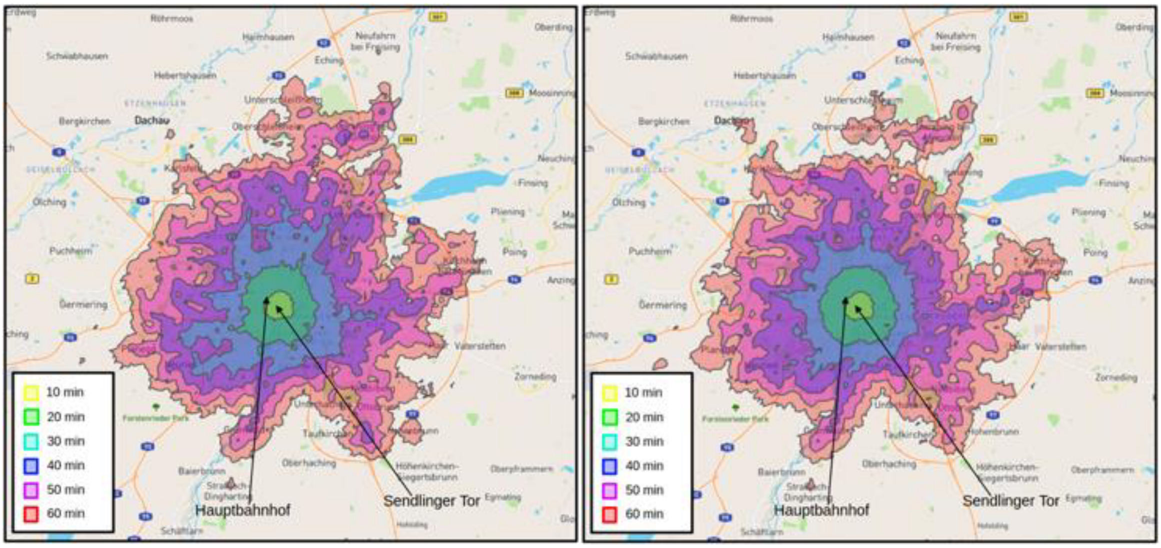


Figure 3. Isochrone of the travel time. Left: Without avoidance of Sendlinger Tor and Hauptbahnhof. Right: With avoidance of Sendlinger Tor and Hauptbahnhof.

Comparing the yellow and green areas in Figure 3 (left) and Figure 3 (right), shows that the avoidance of Hauptbahnhof and Sendlinger Tor did not have a noticeable effect on the area reachable within 20 minutes of the origin Sendlinger Tor. However, the curves that were affected the most are the ones from 30 to 40 minutes (blue and purple areas). When the crowded intersections are avoided, these areas became noticeably smaller. This reflects an increase in travel time for the destinations located within these areas. The last two curves that represent 50 and 60 minutes were also affected by the avoidance. In general, the areas reachable within 50 and 60 minutes are very similar in both isochrones, with a small decrease in reachability when avoidance is included.

4. Discussion

4.1. Result interpretation

Deleting two critical stations in the transportation network results in an overall increase in travel time, particularly for medium long routes. However, this increase is to be expected since the deleted nodes represent important connections in the transit network. Also, for most routes, the increase was so small that it is neglectable. This is specifically true for the random test set presented in Figure 2, where the average travel time for 1000 randomly generated routes only increased by 0.3 minutes. For longer routes, avoiding the central exchange points in Munich's traffic network usually results in less direct routes with more line changes and an increase in waiting time. Regarding the travel time, the increase is rather small, when compared to the potential crowd reduction benefits presented in Figure 2. In contrast to our own expectations, passengers were evenly distributed by the algorithm, thereby preventing neighboring stations from overcrowding. However, this method could still pose a risk, if those neighboring stations are much smaller than the avoided station and therefore could be overloaded even by a comparably small number of passengers. Assuming that this is not the case, the changes to the overall travel time, as well as the stable physical activity level (walk/bike distance) required for a route, would be small enough to compensate for the potential health benefits. This is specifically the case for high-risk groups with a higher need for a safe passage.

4.2. Limitations

The greatest threat lies in the theoretical conception of the motivational aspects. Without real data from travelers, the true threshold for individuals to ignore a safe route can only be estimated. Also, a safety increase through the navigation around congested bottlenecks, can only be assumed but not measured yet. Hence, we do not know how effective such a measure could be unless it has been tested. Another bias lies in the selection of one single city to serve as a prove of concept. Even though the simulation produces promising results in Munich, there is currently no foundation on which the results could be compared to other cities.

4.3. Outlook

We suggest repeating the same evaluation with other cities in Germany, or even Europe, to understand how different transportation networks respond to the avoidance feature. With this measure, it could be determined whether the simulated results of this research are generalizable or whether they merely occurred due to the specific layout of Munich's traffic network. To reproduce these results in other cities, the public transport schedule data and position of traffic hot spots would be required. The format and accessibility of the public transport schedule data may vary between cities. During first investigations, we were able to obtain similar data sets for Nuremburg, Berlin and Duesseldorf. The second requirement for the extension of this work is to obtain the city specific knowledge about the most frequented nodes. For this knowledge a cooperation with local traffic experts would have to be established.

The results presented in the previous sections leave room for a variety of subsequent research areas. As hinted before, our research so far focused on simulations with planned trips. Consequently, the next step could be a comparison of these results to the real travel behavior of Munich's citizens. Currently such a data set does not exist, as there is no technological solution yet that can reliably track the lines, change points and modes of transport a traveler used in a route. Given such a solution was implemented, it could be used to compare the schedule data to the real travel behavior of passengers who are given the chance to test the avoidance router.

Finally, this router could not only be used to prevent infections with contagious diseases, but also to individualize routes to fit to the needs of physically impaired passengers. Firsts test with our algorithm have shown that it is possible to tailor the amount of physical activity through steps in a route to the settings of a passenger. Given that a passenger struggles with walking, it would be possible to generate routes that minimize the number of steps. Vice versa it is also possible to create routes that contain a minimum number of steps for passengers who need or want to include more physical activity in their daily routines.

5. Conclusion

The presented approach simulated how safer routes that avoid overcrowded nodes in the traffic network of Munich could be generated, without causing the travel time to increase significantly. Even though it could not be determined how much safer such routes are, our simulations suggest that it could dissolve existing congestions without causing new ones at neighboring stations. This would allow passengers to maintain a safe distance to other passengers while they are in a public transport station. Determining when a node is at risk of being too congested was established through a collaboration with city specific transportation flow experts. In the here presented use case for the city of Munich, this knowledge was provided by the MVG who already monitor the public transportation network but required additional tools to steer the passenger flow and prevent congestions.

References

[1] C. Courtemanche, J. Garuccio, A. Le, J. Pinkston, A. Yelowitz, Strong Social Distancing Measures In The United States Reduced The COVID-19 Growth Rate: Study evaluates the impact of social distancing measures on the growth rate of confirmed COVID-19 cases across the United States, *Health Affairs* **39**(7) (2020), 1237-1246.

[2] K. Prather, C. Wang, R.T. Schooley, Reducing transmission of SARS-CoV-2, *Science* **368**(6498) (2020), 1422-1424.

[3] German Federal Ministry of Health, Novel Coronavirus SARS-CoV-2: Information and practical advice, https://assets.ctfassets.net/eaae45wp4t29/29lyMngJAvYypiWvqOTtVr/5707e9a6f5ea1681c4d0d1d5e09019f4/BMG_Corona_Information_englisch.pdf, last access: 31.01.2021

[4] Munich Transport Corporation, Aktuelle Zahlen und Fakten, https://www.mvg.de/dam/mvg/ueber/unternehmensprofil/mvg_in_zahlen_s, last access: 02.01.2021

[5] Munich Transport Corporation, Modernisierung U-Bahnhof Sendlinger Tor, https://www.mvg.de/ueber/mvg-projekte/bauprojekte/modernisierung-sendlinger-tor.html, last access: 13.03.2021

[6] T.R. Frieden, C.T. Lee, Identifying and interrupting superspreading events-implications for control of severe acute respiratory syndrome coronavirus 2, *Emerging infectious diseases* **26**(6) (2020), 1059.

[7] D. Bucher, D. Jonietz, M. Raubal, *A heuristic for multi-modal route planning*, in: Progress in Location-Based Services 2016, Springer, Germany, 2017. Pp. 211-229.

[8] C. Narboneta, K. Teknomo, OpenTripPlanner, OpenStreetMap, General Transit Feed Specification: Tools for Disaster Relief and Recovery, *CoRR*, arXiv:1609.01472 (2016).

[9] K.S. Bunds, J.M. Casper, J.A. Jonathan, J. Koenigstorfer, Recreational walking decisions in urban away-from-home environments: The relevance of air quality, noise, traffic, and the natural environment, *Transportation research part F: traffic psychology and behaviour* **65** (2019), 363-375.

[10] P.E. Hart, N.J. Nilsson, B. Raphael, A Formal Basis for the Heuristic Determination of Minimum Cost Paths, *IEEE Transactions on Systems Science and Cybernetics* **4**(2) (1968), 100-107.

[11] M. Duckham, L. Kulik, M. Worboys, A. Galton, Efficient Generation of Simple Polygons for Characterizing the Shape of a Set of Points in the Plane, *Pattern Recogn.* **41**(10) (2008), 3224-3236.

8.2 Full-Text Version of Publication 2

Suggested Citation: J. Klinker, J. Yu, M. Avezum-Mercer and S. Jonas, ”Presenting a Statistical Approach for Transforming Standardized German Traffic Surveys into Origin-Destination Matrices,” 2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST), Detroit, MI, USA, 2023, pp. 177-186, doi: 10.1109/MOST57249.2023.00026.

Status: Published.

2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST)

2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST) | 979-8-3503-1995-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/MOST57249.2023.00026

Presenting a Statistical Approach for Transforming Standardized German Traffic Surveys into Origin-Destination Matrices

1st Jens Klinker
School of CIT
TU-Munich
Munich, Germany
jens.klinker@tum.de

2nd Joe Yu
School of CIT
TU-Munich
Munich, Germany
joe.yu@tum.de

3rd Mariana Avezum-Mercer
Mobility Department
Vay Technology GmbH
Munich, Germany
m.avezum@tum.de

4th Stephan Jonas
Institute for Digital Medicine
University Hospital Bonn
Bonn, Germany
ORCID: 0000-0002-3687-6165

***Abstract*—This paper presents a method for generating Origin-Destination Matrices (ODMs) for the city of Munich using traffic count data from a Germany-wide study conducted by the German Federal Ministry of Transport and Digital Infrastructure (MiD-Study). The results show that the data provided by the MiD-Study was correctly translated into an ODM, thereby providing an interpretable demand format for traffic simulations. Due to the consistent design of the MiD-Study, the approach is also applied to Hamburg and is extensible to 18 further cities and one city-state (Bremen) covered in the MiD-Study. The produced ODMs for Munich and Hamburg are accessible for researchers at: https://nextcloud.in.tum.de/index.php/s/gT48xDzT88YJGQK**

***Index Terms*—traffic intensity, ODM, data analysis, data processing, traffic data**

I. Introduction

Structural modifications to traffic networks are costly. Amongst financial costs, they include deteriorations in travel time, pollution, noise, and other inconveniences to travelers and citizens [1] [2]. On the other hand, structural modifications to traffic networks a required to accommodate changes in the demand on the network. City planners and politicians try to negotiate between both the required network changes and the costs linked to their implementation. They thereby try to estimate a break-even point between the long-term benefits of the network improvement and the costs for the affected part of the public.

Sander summarizes that the "currently existing measures to inform the public and collect potential disagreement are based on classical media and require to lookup plans and project descriptions during the opening hours of municipal offices where those materials are at display [3]." They further describe that "a promising approach to overcome these problems is to involve inhabitants in public processes actively [3]."

This paper will contribute to the active involvement approach Sander describes by aiding traffic simulations with origin-destination matrices (ODMs) for cities in Germany. These ODMs are required to simulate how a traffic network is currently used by travelers, thus providing the so-called demand on the network. ODMs also help identify which parties might be affected to which extent by the gains and costs of a structural change.

II. Fundamentals

If an active involvement of stakeholders is a promising approach, why is it not the go-to solution for city planning? The reason becomes apparent once the effort for creating a traffic simulation is considered. When traffic simulations are compared to other options by their communication effectiveness for the public and implementation effort for the city, interactive solutions rank the highest in effort.

A. Data Communication Levels

For an easier understanding, we cluster approaches into three "Levels" of increasing implementation effort.

- Level 1 Solutions - Raw Data:
  Properties: (Low effort, low effectiveness)
  Visualization: (Lists of values, usually in tabular formats)
- Level 2 Solutions - Visualized Charts:
  Properties: (Low effort, mediocre effectiveness)
  Visualization: (Charts, figures, images, videos)
- Level 3 Solutions - Interactive Simulations:
  Properties: (High effort, high effectiveness)
  ("Games" showing the impact of a stakeholder's decision in the simulated environment)

Level 1 solutions are the easiest to generate, but also the hardest to process for humans. They are comparable to displaying the source code of a program to a user instead of the graphical user interface (GUI). Only trained experts will be able to understand this level of communication.

Level 2 solutions provide a decent level of abstraction compared to level 1, but reduce the raw data to a selection of snapshots. This could be compared to showing screenshots of a program to a user. Researchers such as Rosling developed tools like Gapminder to increase the frame rate of these snapshots

979-8-3503-1995-8/23/$31.00 ©2023 IEEE
DOI 10.1109/MOST57249.2023.00026

to maximize their communicative power, bringing them to the very edge of a Level 2 solution. [4].

But even if the frame rate is increased to a level where the screenshots turn into a video, a lack of interactability for the involved individual remains. Users might grasp the nature of the presented data but can not directly interact with the raw data. It may be compared to watching videos on how to fly a plane. Even though a theoretical concept is communicated, an interactive exercise would provide new insights that viewers are not able to derive from a video. For similar reasons, pilots train in flight simulators and as co-pilots before flying an actual plane by themselves. For this purpose Level 3 solutions were established. They offer an interactive simulation to the users, allowing them to select and interact with the raw data, visually showing them the effect of their actions. Referring back to traffic networks, this would mean releasing an interactive traffic network simulation to decision-makers.

Level 3 solutions are the most effective form of communication, but they are also significantly harder to build. Due to the involved effort, most communicators seem to prefer level 2 solutions. Nonetheless, if a simulation is available, it is the best tool for the active involvement of all stakeholders in the decision process.

B. Traffic Networks

To support city planners and decision-makers in the process of finding an optimal solution for structural changes to a traffic network, this paper contributes to the implementation of a realistic traffic simulation. While a visual simulation is nothing more than a computer game, it is the word "realistic" that carries complexity and can never be fully achieved. To approach an acceptable level of realism in the simulation of a city's traffic network, four different building blocks are required:

1) **Traffic Network Data**: Expressed through nodes, edges, weights, and the modes of transport that operate on each graph element. This builds the "supply" that is provided by the network.

2) **Travel Behavior Data**: Expressed by the travelers' origin, destination, departure times, and chosen modes of transport. This results in the "demand" on the network, generated by its users.

3) **Route Planning**: An algorithm to calculate routes that are most likely taken by an individual, given an origin, destination, and choice of transportation mode. The algorithm effectively finds a balance between the supply offered by the network and the individual demand of a traveler.

4) **Traffic Load Balancing**: If multiple individuals travel a network, the Route Planning is affected, as the demand influences travel time (weights) on the edges of the network. Point 3) expresses the optimal route in the traffic network's default state, Traffic Load Balancing adapts 3) to the current state of the network based on how the travelers occupy its capacities.

Traffic simulations are mathematical models that are used to represent and analyze traffic patterns, behaviors, and their impact on a given transportation system. Like any other optimization model, traffic simulations are based on assumptions and approximations, and they may not accurately reflect all of the complex and dynamic factors that influence traffic in the real world. As a result, traffic simulations are not perfect and may not always provide accurate or reliable predictions.

However, traffic simulations can be useful tools for understanding and analyzing traffic patterns, but should not be relied upon exclusively or treated as a substitute for real-world data and observations. Instead, the goal of optimization is to constantly reduce differences between real and modeled behaviors. The accuracies of both 3) Route Planning and 4) Traffic Load Balancing depend on the underlying accuracy of the used data. While 1) Traffic Network Data is already openly and quite accurately available through platforms like OpenStreetMap. But 2) Travel Behavior Data is still scarce, decentralized, low resolution, or not obtainable at all. One reason for this discrepancy is the difference in privacy invasion in the generation of the two datasets. 1) Traffic Network Data is an objective fact, that may be generated without interference with personal privacy. In comparison, 2) Travel Behavior Data is highly subjective data that could potentially be used to derive delicate information about a person's private life.

The objective of this paper is to improve 2) Travel Behavior Data by providing a system for generating data sets. These data sets also called "origin-destination matrices" (ODMs), are generated based on available research that is published in a Level 1 or Level 2 format. These studies are then transformed into a statistically representative ODM to build Level 3 formats. These ODMs consist of an origin, a destination, a departure time, and a selected mode of transport. The approach is demonstrated for the greater areas of Munich, Hamburg, and Bonn. As the approach is kept generic, it is extensible to most other cities in Germany. It is based on a Germany-wide study, which is addressed in the Methods section.

III. Related Work

Traffic networks unite different research areas. Graph theory, game theory, and psychology are required to understand and describe the network, the travelers, and path optimization methods.

A. Game Theory: Supply

Probabilistic risk analysis and game theory, such as the work of Hausken, have shown that "introducing the behavior dimension into Probabilistic Risk Analysis (PRA) is important because players in a dispersed system expend valuable resources trying to increase system reliability interpreted as a public good. Risk is affected by behavioral, technological, and natural factors, controllable to different extents. Individual strategies at the component or unit level do not always match collective desires at the system level" [5].

Transferred to traffic networks, the set of nodes, edges, and weights that form the network are the grounds for a game's theoretical optimization. If Hausken's analysis is correct, a game's theoretically optimal solution is nonetheless influenced by the behavior of the individuals in the game. In traffic networks, this means that traveler goals, such as their travel path's origin, destination, and departure time add subjectivity to the question of which route is optimal. Additionally, the travelers' personal preferences like the selected mode of transport further influence a rationally optimal solution. This leads to the observation that an optimal travel route is influenced by objective factors (lowest travel time between origin and destination), as well as subjective factors (individual preferences).

The basis for any game theoretical optimization is 1) Traffic Network Data, which is nowadays obtainable using publicly available open-source solutions such as OpenStreetMap. Due to this data's high level of detail and obtainability, it is comparably easy to find rationally optimal routes for any provided set Origin-Destination Pairs (O-D-Pairs) [1].

B. Psychology: Demand

On the other hand, obtaining these sets of O-D-Pairs is significantly harder. Battigalli describes how motivations for a game's theoretical action can be influenced by a multitude of "belief-dependant motivations" [6]. They argue that traditional game theory, which assumes that individuals make decisions based solely on their own preferences and beliefs about the payoffs they will receive, is incomplete. Instead, they propose an extension of game theory that takes into account belief-dependent motivations, which are motivations that are influenced by beliefs about others' motivations.

This concept of belief-dependent motivations can be applied to the behavior of individuals in traffic networks and to designing policies and interventions that take into account the role of beliefs and motivations in transportation decisions.

C. ODMs: Measuring and Modeling Demand

Due to the complexity added by these psychological influence factors, researchers building data sets containing realistic ODMs, resort to observation instead of intelligent design. An approach for observation-based ODMs that has been used for over 4 decades now is traffic counts. Erlander used traffic counts in 1979 to build unique optimal solutions based on the assumption that traffic counts are available for all edges in the network.

To generally estimate an ODM based on traffic counts, researchers first collect traffic count data at a number of locations in the network. These data are used to estimate the flow of traffic between pairs of locations in the network. To then build an ODM for the entire network, approaches start differing slightly. Yang used the priorly described information, along with assumptions about the distribution of traffic within the network, to estimate the ODM [7]. Erlander used traffic counts in 1979 to build unique optimal solutions, based on the assumption that traffic counts are available for all edges in the network [8]. Van Zuylen (1980) and Fisk (1988) built models taking congestions in these networks into consideration. Like Erlander, they based their models on traffic counts [9] [10]. Spiess (1987) used traffic counts to create a maximum-likelihood model. They then used the model to reproduce the originally observed traffic count data and derive an optimal path solution [11]. Kawakami also used traffic counts to construct ODMs while taking choices for different modes of transportation into account, but neglecting congestions [12].

Bera describes a shift in ODMs, now splitting them into two types: Static and dynamic. Bera explains that "in static methods, the traffic flows are considered as time-independent and an average O-D demand is determined for long-time transportation planning and design purpose. Whereas from last two decades, different dynamic approaches are proposed which are meant for short-term strategies like route guidance, traffic control on freeways, intersections, etc" [13].

Today, differences are more distinct, as most traffic simulations include departure times in their ODM making them mostly dynamic ODMs. The differences are rather found in the way 2) Travel Behavior Data is used to build an ODM, and the choice of methods to perform 3) Route Planning and 4) Traffic Load Balancing.

Yadav (2020) designed a traffic estimation framework generated from open traffic cameras using deep learning technology. As a basis for their traffic network, they relied on OpenStreetMap while exposing their results through visualizations back to OpenStreetMap [14].

The German Aerospace Center (DLR) spent the last 20 years developing an open-source traffic simulation package called SUMO, thereby making it one of the longest research-based simulation tools. Behrisch provided a status update in 2011, describing how it is a purely microscopic traffic simulation, as each vehicle is uniquely represented in the system. Like Yadav, they mostly use OpenStreetMap as a source for their network's underlying graph of nodes and edges. The system's traffic behavior models were evaluated against footage of actual traffic cameras at certain intersections. At this stage, they have also been working on the integration of other modes of transport, such as walking, biking, and buses. In 2018 Lopez published an update on the SUMO system. They describe how SUMO is still using OpenStreetMap to build network graphs. To add further convenience to users of the system, they are featuring an additional download and import functionality for OpenStreetMap data. Also, SUMO tools such as OD2TRIPS have been added to support the integration of network demand data in the form of ODMs. Additionally, SUMO further includes rail traffic and its intersections with the road network in its calculations [15] [16].

[1] www.openstreetmap.org

IV. Methods

This paper introduces an approach for generating ODMs from traffic count data for the city of Munich. The basis for the ODMs is provided by a traffic data set collected in a study conducted by the German Federal Ministry of Transport and Digital Infrastructure (MiD-Study). The study consists of 48.627 travel paths collected from 15.693 participants in Munich and 41.404 travel paths from 13.660 participants from the outskirts surrounding Munich. The study is reproduced in regular time intervals for multiple cities in Germany [17] [18]. As the studies produce data in a similar structure in each iteration and for each city, the approach is generally extensible to other cities in Germany.

The study generally provides statistics and insights about travel behaviors, including mode of transport choice, travel purpose, and travel distance. It does however not provide specific information about the origins or destinations of individual trips.

In the following sections, both the approach and the underlying study data are introduced in greater detail.

A. Scope and Assumptions

Since the study does not provide an actual ODM, certain assumptions had to be made to translate the data into an ODM format. The goal is to translate the provided statistics from the MiD report into a realistic set of Origin-Destination-Pairs (O-D-Pairs), which are then agglomerated to an Origin-Destination-Matrix (ODM) representing the demand on Munich's traffic network.

The study was generated by conducting interviews with citizens of the respective city, meaning that it does not include tourism or routes starting outside the city boundaries. There was also no clear data source obtainable describing the percentage of touristic trips in Munich. Hence, the influence of tourism is neglected in the ODM.

The scope for this paper was also narrowed down to trips originating in the city of Munich, excluding the greater Munich municipality.

As no knowledge about actual origins or destinations was provided in the study, the assumption was made that all trips start and end at a residential address in the city of Munich, thereby assuming that travelers eventually return to their homes. This approach however neglects potential intermediate travel destinations. Obeying this assumption, any taken trip starts from a citizen's home to a purpose-driven destination. The consecutive trip will have the inverse direction, starting from the purpose-driven destination, and ending at the citizen's home again. This means that, for example, a trip to work, then to the supermarket, and finally, home is not considered. Instead, the path from work must directly go to the person's home before a following trip to the supermarket can be conducted.

Allowed trip structure:

- Home to Point of interest
- Point of interest to Home

This leads to the rule that a trip must always start at a residential address unless it is a direct return trip.

B. Specifying the Population

Given this scope, the likelihood of a trip starting or ending at a home location based on the distribution of homes in Munich's districts is estimated. As precise data about Munich's population spread across districts was obtainable, as displayed in Figure 1, this proportion of citizens was utilized to attribute a similar proportion of trips to each district[2]. If further data with an even higher level of detail is found, it could be used to determine the travel origin more precisely.

City district	Population	Inhabitants in % of the total population in Munich
1 Ramersdorf - Perlach	108.244	7,39
2 Neuhausen - Nymphenburg	95.906	6,55
3 Thalkirchen - Obersendling - Forstenried - Fürstenried - Solln	90.790	6,20
4 Bogenhausen	82.138	5,61
5 Milbertshofen - Am Hart	73.617	5,03
6 Pasing - Obermenzing	70.783	4,83
7 Schwabing - Freimann	69.676	4,76
8 Trudering - Riem	67.009	4,57
9 Schwabing West	65.892	4,50
10 Au - Haidhausen	59.752	4,08
11 Feldmoching - Hasenbergl	59.391	4,05
12 Sendling - Westpark	55.405	3,78
13 Laim	54.030	3,69
14 Untergiesing - Harlaching	51.937	3,55
15 Maxvorstadt	51.642	3,53
16 Moosach	51.537	3,52
17 Obergiesing - Fasanengarten	51.499	3,52
18 Ludwigsvorstadt - Isarvorstadt	50.620	3,46
19 Hadern	48.945	3,34
20 Berg am Laim	43.068	2,94
21 Aubing - Lochhausen - Langwied	42.305	2,89
22 Sendling	39.953	2,73
23 Allach - Untermenzing	30.737	2,10
24 Schwanthalerhöhe	29.663	2,02
25 Altstadt - Lehel	20.422	1,39
Munich total	**1.464.961**	**100,00**

Fig. 1. Considered city districts in Munich, by total population and percentage of the total population.

According to the MiD-Study, an estimated total of 4.8 million daily trips are taken in Munich every day. As mentioned in the previous subsection only the city of Munich was considered, excluding the extended municipal public transport area "MVV" (an additional 4.6 million trips) that describes travel behavior in the suburban districts surrounding Munich. Instead, the scope was set to the city, using it as the statistical trip population for our ODM. The system could however be extended to also include these areas in the calculations.

The previous subsection mentioned how the ODM is built on the assumption that every trip must start at a home address

[2]https://suedbayerische-immobilien.de/Einwohner-Muenchen-Stadtteile

unless it is a direct return trip. This means that for every trip away from someone's home, there must be exactly one return trip back home. This means that every trip in the resulting ODM must either start directly at a residential address or return to one, creating a 1:1 relationship. Following this principle, the chosen population of 4.8 million trips, must consist of 2.4 million trips starting at a residential address. Following this logic, the other half of the trips must have one of these residential addresses as its return destination. The initial assumption also dictates that these two partitions can not overlap one another.

Extending this thought leads to the conclusion that only the origin-destination pairs for half the population have to be determined, as the other half must be the inverted version of these pairs.

C. Building Origins

Based on the previous section's logic, the number of trips in the population (4.8 million) is split in half (2.4 million). Next, they are fractionally assigned to **city districts** in Munich, based on the percentages shown in Figure 1. This splitting process is visualized in Figure 2.

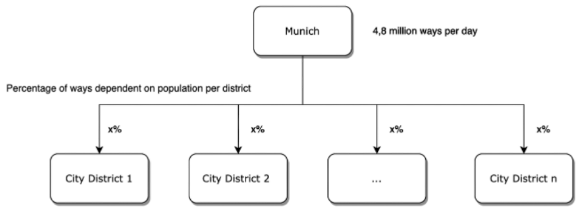


Fig. 2. Visualization of how trips (ways) are assigned to city districts

The MiD-Study offers additional information about **travel purposes**. These purposes are split into work, business, education, shopping, errands, leisure, and escorting, as displayed in the lower half of Table 1. For this analysis step, we are mostly interested in the "Munich City" column, nested in the "Trips in MM per day" column, as it provides us with the number of trips that are daily taken in Munich's city and how many of these trips are linked to the each row's travel purpose.

To show how these travel purposes are taken into consideration, we next focus on one district at a time. Figure 3 shows how percentages are used to build a relationship between **travel purposes** and the **chosen mode of transport**. To build an example, we are initially interested in the route percentages for public transportation, displayed in green on the left half of the Figure.

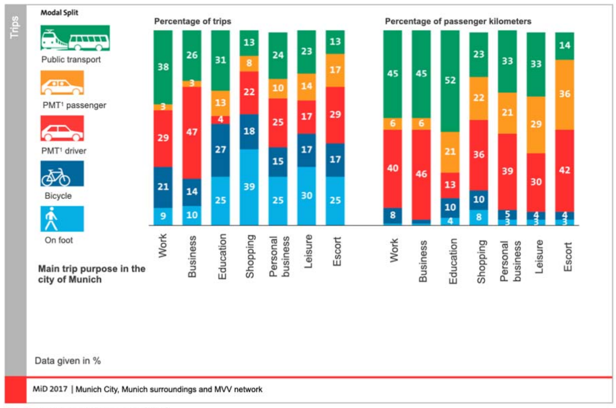
We then take these distributions into consideration to calculate a precise percentage of trips starting in a specific district, with a specific purpose, taken by public transport. For example, we could take the city district of Ramersdorf-Perlach which makes up 7.39% of the households in Munich.

For Ramerstdorf-Perlach, we multiply the remaining 2.4 million trips of our population by the percentage of total households in this district (7.39%). We then apply a travel

	Trips in MM per day: Germany	Trips: MVV³ network	Trips: Munich city	Trips: Munich surroundings	Passenger kilometers in MM per day: Germany	Pkm: MVV³ network	Pkm: Munich city	Pkm: Munich surroundings	Trip length in kilometers 2017 – Germany: Average value	Germany: Median	MVV³ network: Average value	MVV³ network: Median	Munich city: Average value	Munich city: Median	Munich surroundings: Average value	Munich surroundings: Median
Main mode of transport																
On foot	56	2,0	1,2	0,8	93	3,1	1,7	1,4	1,7	1	1,5	1,0	1,4	0,8	1,7	1,0
Bicycle	28	1,4	0,9	0,6	112	5,4	3,1	2,3	3,9	2	3,7	2,0	3,6	2,0	4,0	2,0
PMT¹ driver	111	3,1	1,1	2,0	1.754	56,3	21,4	34,9	15,8	6,7	18,0	5,7	18,9	5,7	17,6	5,7
PMT¹ passenger	36	1,1	0,5	0,7	650	24,4	12,0	12,4	18	5,7	21,8	7,6	25,8	6,7	18,9	7,6
Public transport	26	1,7	1,2	0,5	605	34,4	21,9	12,5	23,1	8,1	20,4	7,6	18,6	6,3	24,6	14,4
Total	**260²**	**9,4**	**4,8**	**4,6**	**3.200²**	**123,5**	**60,0**	**63,5**	**12,5**	**3,8**	**13,2**	**3,8**	**12,5**	**3,6**	**13,9**	**4,8**
Main trip purpose																
Work	42	1,6	0,9	0,7	674	25,6	11,1	14,5	16	8,1	15,6	8,1	12,4	6,3	19,5	14,3
Business	28	0,9	0,4	0,5	539	20,3	10,2	10,1	19	5,7	22,4	6,9	22,7	4,8	22,1	9,5
Education	18	0,7	0,3	0,4	131	4,5	1,9	2,6	7,3	2,9	6,7	2,4	6,3	2,0	7,0	2,7
Shopping	41	1,5	0,8	0,7	217	6,5	2,8	3,6	5,3	2	4,4	1,9	3,7	1,5	5,2	2,3
Personal business	37	1,2	0,6	0,6	376	11,8	5,5	6,3	10,2	3,6	9,7	3,3	9,2	2,9	10,1	3,8
Leisure	71	2,7	1,4	1,3	1.098	48,3	25,4	22,9	15,5	3,9	17,8	3,9	17,6	3,9	18,0	4,8
Escort	21	0,7	0,3	0,4	179	6,2	2,8	3,4	8,6	2,9	8,7	2,9	8,4	2,4	8,9	2,9
Total	**260²**	**9,4**	**4,8**	**4,6**	**3.200²**	**123,5**	**60,0**	**63,5**	**12,5**	**3,8**	**13,2**	**3,8**	**12,5**	**3,6**	**13,9**	**4,8**

¹ Private motorized transport (PMT)
² rounded values
³ Munich Transport and Tariff Association

TABLE I
TABLE SHOWING TRAVEL PURPOSES (LOWER HALF, ROWS) AND TRIPS IN MUNICH (LEFT THIRD, COLUMNS) [17].



¹ Private motorized transport (PMT)

Fig. 3. Visualization mapping travel purpose to used modes of transport for the given purpose [17].

purpose, for example, "work", which represents 18.75% of trips, and multiply it with the previous values. Finally, we multiply everything by the number of people who are using public transport. For the travel purpose "work" this is 38%. As a result, we end up with roughly 0.28% (= 13440 trips) of all trips that have their origin in Ramersdorf-Perlach, the purpose "work", and the travel mode "public transport". This procedure is then repeated for the remaining travel purposes and modes of transport. A visualization of this process for all purposes is provided in Figure 4.

This rather simple formula allows us to estimate a trip's origin on the district level, for any city in which a MiD-Study has been conducted. If information should become available describing residence distributions at a higher resolution, it may be included by replacing the districts. The same rule may be applied to the other abstraction layers such as the travel purpose or the mode of transportation.

As we currently do not have data available at higher res-

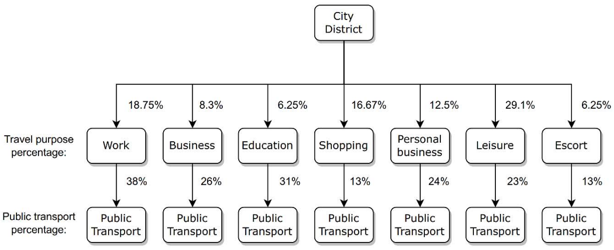


Fig. 4. Explains how trips in city districts are fractionally broken down by travel purpose and mode of transport.

olutions than the district level we randomly select residential addresses that are located in the requested district. To find candidate nodes we use a combination of tools working with data from OpenStreetMap.

First, we select all residential addresses in the respective district and save them in a randomly shuffled list. Next, we draw addresses from this list and add them as origins to the ODM. This process can be continued until the required number of addresses has been distributed. For the district of "Ramersdorf-Perlach", the travel purpose "work", and the travel mode "public transport" resulted in 13440 trips (0.28% of 4.8 million).

While this method allows us to create realistic origins based on the MiD-Study, it also introduces biases. One consequence, for example, is the underrepresentation of multi-family homes in comparison to single-family homes as both are only represented with a single address in the shuffled list but have large differences in the number of inhabitants.

The same procedure is next carried out for the other modes of transport, travel purposes, and city districts. We additionally decided to work with just 10% of the actual number of trips in Munich, shrinking down the final ODM to 480.000 trips instead of 4.8 million. Based on this reduction we eventually finish this step with roughly 240,000 trips with assigned origins. Thereby, the desired 50% of trips for the ODM is reached. Each of these 240,000 trips now contains a designated origin, meta-information about its assigned travel purpose, and the chosen mode of transportation that will be required in the following destination estimation.

D. Adding Destinations

To complete the first 240,000 trips, only a destination is required. As destinations are not provided in the MiD-Study, they are approximated based on the information stored in the uncompleted trips. To estimate a destination, first, a travel distance is required for each trip. This distance is approximated using the values provided in Figure 5.

Figure 5 matches a mode of transport to a distribution of travel distances covered with it. Previously we already matched travel purposes to a distribution of modes of transport used to fulfill the given travel purpose. Therefore, the distance distribution for modes of transportation also includes the travel purpose for our set of 240,000 trips. As no exact percentages

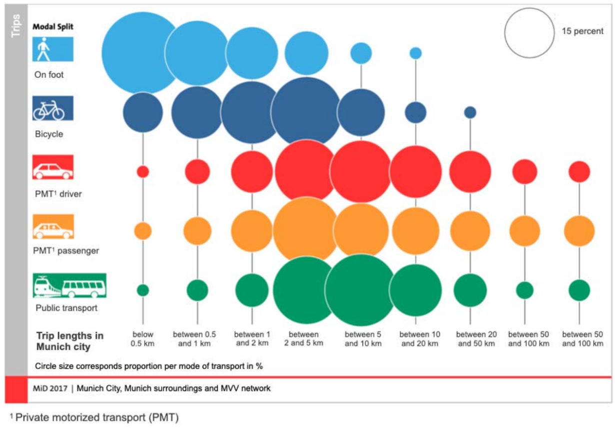


Fig. 5. Depiction describing which mode of transport is used to which degree for altering travel distances [17].

were provided in Figure 5, we developed a simple algorithm to reverse engineer the percentages based on the area of the circles. We also reduced the distance classes to "below 2km", "2 – 5km", "5 – 10km", "10 – 20km", and "20 – 50km". Figure 6 displays this process visually for the transportation mode "Public Transport".

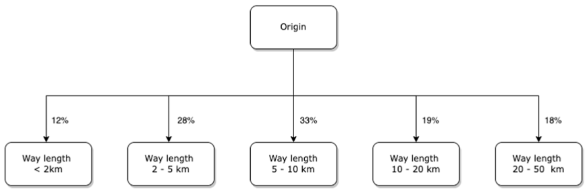


Fig. 6. Provides a visual example for the adapted distance classes.

Once every trip has been assigned a range group we search for fitting destinations. This search is done on a shuffled data set fitting to the travel purpose of the trip. For example, for a trip with the travel purpose of "shopping" (referring to groceries shopping), we collect all nodes in Munich marked as supermarkets in the set for potential destinations. Next, we randomly pick a supermarket from the list and calculate whether the distance from the trip's origin to this destination falls into the travel distance interval assigned to the trip. If it does, it is saved as the trip's destination. If it does not, a new trip is selected and put through the same process. To measure the distance, we use an API provided by the OSRM project, which calculates the driving distance between two nodes[3] [18]. Even though this solution is slightly inaccurate for transportation modes aside from cars, we considered it to be a better approximation than, for example, a distance calculation as the crow flies using the haversine distance.

Once all trips have been assigned a destination, we generate their inverses by swapping the origin and destination to build

[3]http://project-osrm.org

the return trips representing the other half of the total dataset.

V. Results

The goal of this paper was to produce an ODM that correctly represents the findings of the MiD-Study. Figures 7, 8, and 9 compare the distributions described (target value) in the MiD-Study to ODM produced with the method introduced in the previous chapter (actual value). A closer look reveals that the values for the "actual value" always exceed those of the "target value". The reason for this is the way the translation method was implemented. Since this process is supposed to act completely random, to avoid introducing undesired biases, trips with their specific properties are produced until they met or exceeded the target value. Since this is an expensive computational process, queries are batched to calculate multiple routes at once. Thus, there is only a small probability that the target value is hit precisely. In fact, it is much more likely that it will be slightly exceeded, as can be observed in the figures. In Figure 7, combining all errors results in an overall deviation of 2.08% from the target values set by the MiD-Study.

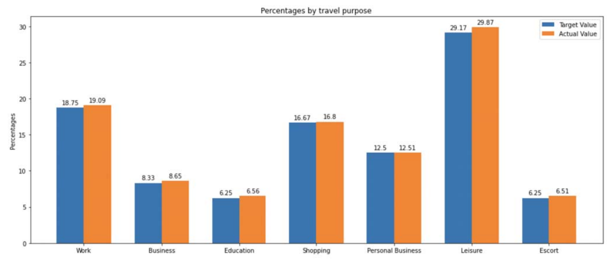


Fig. 7. Matches the targeted percentage of trips with a travel purpose to the actual number of trips with that travel purpose in the newly created ODM.

The relative distribution for transport modes was calculated by multiplying the percentage of the individual travel purposes (above) with the respective shares of the travel mode in modal split statistics below. For example, for Public Transport, the calculation looks like this: $23.92\% = (18.75\% * 0.38\%) + (8.33\% * 0.26\%) + (6.25\% * 31\%) + (16.67\% * 13\%) + (12.5\% * 24\%) + (29.17\% * 23\%) + (6.25\% * 13\%)$ Figure 8 shows the resulting percentages for all modes of transport.

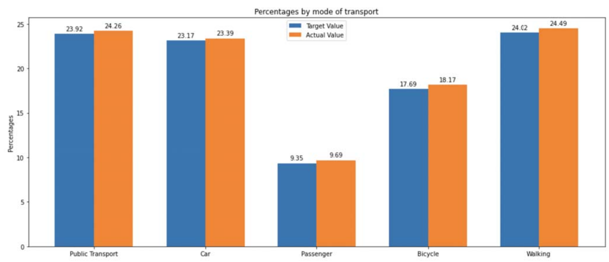


Fig. 8. Matches the targeted percentage of trips with a mode of transport to the actual number of trips with that mode of transport in the newly created ODM.

Figure 9 visualizes the deviation of trips in a defined travel distance cluster. Again the target value is slightly exceeded but by a neglectable amount.

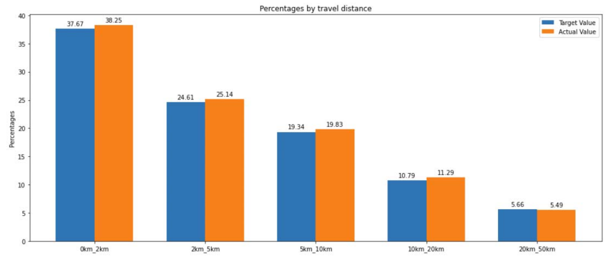


Fig. 9. Matches the targeted percentage of trips within the travel distance cluster to the actual number of trips within that travel distance cluster in the newly created ODM.

Figures 7 to 9 show that the insights of the MiD-Study were correctly transformed into an ODM. Next, we validate this ODM against another ODM. To do so, we used an existing data set obtainable through TomTom move, to extract an ODM recorded traffic data in Munich[4]. TomTom itself is a company that specialized in building navigation systems for cars. So, the GPS data used to generate the TomTom ODM was collected from cars using the TomTom navigation system[5]. To now generate a comparable dataset, we selected all trips that were performed in the same geographical boundaries as in our data set. The selected boundary areas are presented in Figure 10 and match precisely the boundaries used in our approach. Next, a data sample was created as a basis for the TomTom ODM by agglomerating all trips in one week of January 2021 resulting in 1,018,950 trips. The reason for this setting is to have a similar basis to the data provided by the MiD-Study which also based its statistics on a full week that was then combined to express travel behaviors on an average travel day.

To compare both data sets, we generated a list of possible O-D combinations between the 25 different city districts, resulting in 625 possible combinations. Even though a usual handshake problem with 25 participants would result in 300 possibilities, in an O-D scenario two additional alterations are to be considered. First the "direction of the handshake" matters meaning that switching the origin and the destination of an existing possible connection results in a new O-D-Pair, adding 300 more O-D-Pairs to the existing 300. Second, reflexive O-D-Pairs are also relevant meaning that a traveler's origin and destination may lie in the same district (e.g from "Laim" to "Laim"). To include these options, 25 additional O-D-Pairs are added, resulting in 625 possible district O-D-Pairs. We next use a reverse geocoding API, to assign each trip stored in both data sets to one of these 625 O-D-Pairs. The result is a side-by-side comparison of absolute trips of our data set to the TomTom data set as displayed in Figure 11.

As the sum of the total trips differs between both data sets, we converted the absolute trip values to relative percentages.

[4]https://move.tomtom.com
[5]https://support.move.tomtom.com/od-analysis-introduction

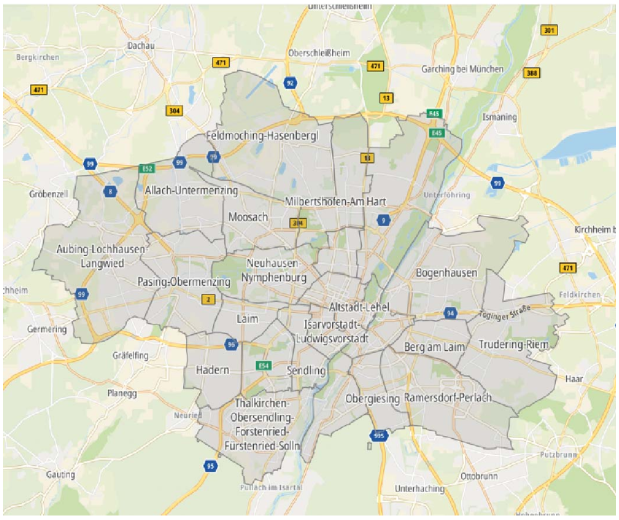


Fig. 10. Snapshot showing the selected regions for both the TomTom ODM and our ODM based on the Munich MiD-Study.

	Origin	Destination	Trips	Trips Simulation	Deviation
0	Feldmoching-Hasenbergl	Feldmoching-Hasenbergl	0.012573	0.017272	1.373737
1	Feldmoching-Hasenbergl	Milbertshofen-Am Hart	0.003366	0.003588	1.065954
2	Feldmoching-Hasenbergl	Schwabing-Freimann	0.004638	0.00157	-2.95414
3	Feldmoching-Hasenbergl	Allach-Untermenzing	0.003293	0.000646	-5.097523
4	Feldmoching-Hasenbergl	Aubing-Lochhausen-Langwied	0.012028	0.000378	-31.820106
5	Feldmoching-Hasenbergl	Pasing-Obermenzing	0.001485	8E-05	-18.5625
6	Feldmoching-Hasenbergl	Hadern	0.000613	5E-05	-12.26
7	Feldmoching-Hasenbergl	Neuhausen-Nymphenburg	0.001346	0.000775	-1.736774

Fig. 11. Snapshot of the 625 entry long list of possible O-D-Pairs, with Trips (TomTom), Trips Simulation (MiD based), and the Deviation between these two columns.

The column "Trips" in Figure 11 shows the percentage of trips for this O-D connection in the **TomTom data set**. The values in the column are generated by dividing the number of trips for that connection by the total number of trips in the TomTom data set. The column "Trip Simulation" refers to **our data set** that was generated from the MiD-Study data. The "Deviation" column shows the deviation between the "Trips" and "Trips Simulation" columns for each row. A value of 1 means that no deviation was observed. A value larger than one means that more trips were recorded in our data set for the given O-D-Pair than in the TomTom data set. Negative deviation values mean that fewer trips were observed by comparison. For example, a positive deviation of 1.3 means that the smaller value (for positive values the "trips" value) deviates by a multiplicator of 1.3 from the larger value. Negative values are used to indicate a change in the direction of the deviation, but function similarly (here: "Trips Simulation" multiplied by "Deviation" results in "Trips"). The resulting spread for deviations is visualized in Figure 12.

For Figure 11 and Figure 12, all modes of transport were used from our MiD-based data set, while the TomTom data

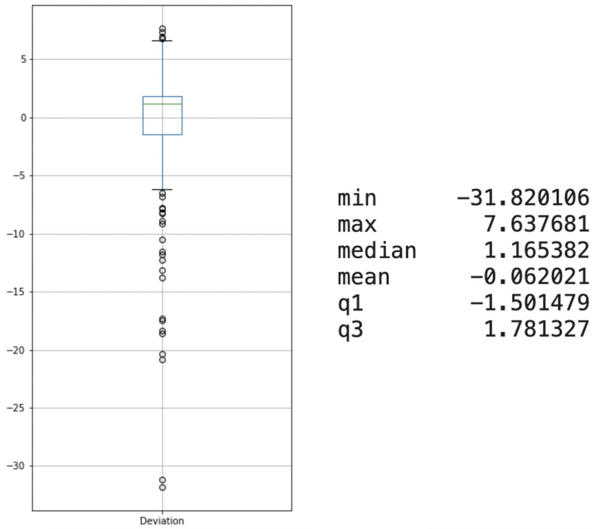


Fig. 12. Boxplot for the deviations between TomTom Trips and Simulation trips for **all modes of transport** in the newly created ODM.

set only provides car data. To also show how the two data sets compare solely on car data, an additional comparison was created in which all modes of transport except for cars were excluded. The resulting deviations are shown in Figure 13.

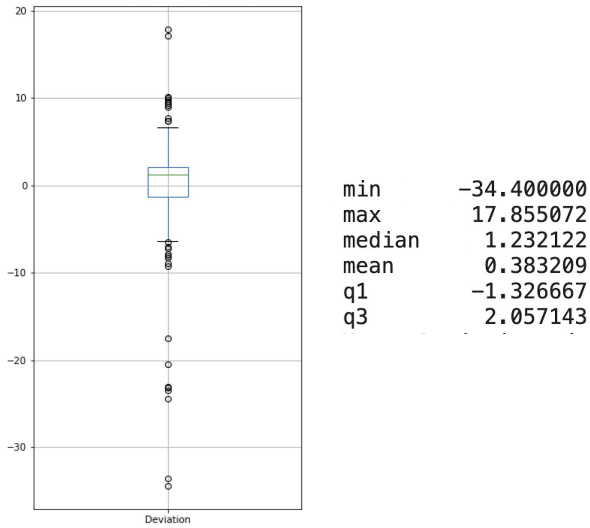


Fig. 13. Boxplot for the deviations between TomTom Trips and Simulation Trips (MiD based) for **cars** in the newly created ODM.

A. Interpretations

The median value for the deviation shown in Figure 12 (TomTom car set's deviation to the MiD set's with all modes of transport included) is roughly 1.165, meaning that a median deviation of 16.5% was observed. However, when comparing the TomTom car set's deviation to the MiD car set's deviation a slightly higher value of 23.2% is observed. This observation is surprising, as we expected to see a reduction in the deviation when noise modes of transport are excluded. We also measured an increase in the variance from 13.37 to 16.63 (std: 3.66 and 4.08) indicating an increase in volatility by the exclusion of other modes of transport. Our best explanation for this increase in variance and median deviation is that generally, a reduction in sample size can lead to greater variances. This, however, is rather speculative. The variance also shows that both data sets are quite different.

To improve our understanding of the differences between the TomTom based and the MiD-based data sets, we used precisely the same methodology to produce two additional data sets for the city of Hamburg. The first is TomTom based and the second is MiD-based. This was done by consulting an external company (vay.io) that had priorly analyzed TomTom data in Hamburg. They explained that TomTom clustering data origins and destinations based on hexagonal grids before publishing it, as shown in Figure 14.

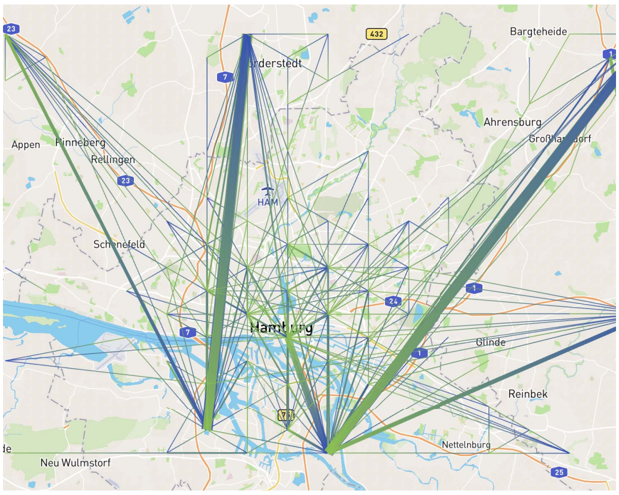


Fig. 14. Map of Hamburg showing **the raw data provided by the TomTom data set**. Lines are O-D-Pairs. The line thickness represents the number of underlying trips. Origins are in blue, and destinations are in green. **Trips run from hexagon to hexagon**.

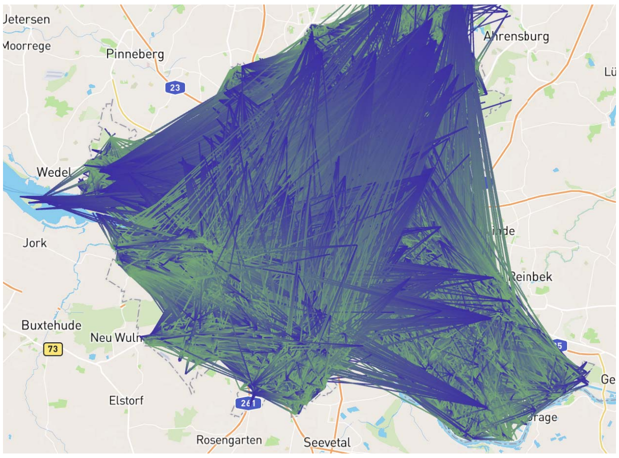


Fig. 15. Map of Hamburg showing **the raw data provided by the MiD-based data set**. Lines are O-D-Pairs. The line thickness represents the number of underlying trips. Origins are in blue, and destinations are in green. **Trips run from address to address**.

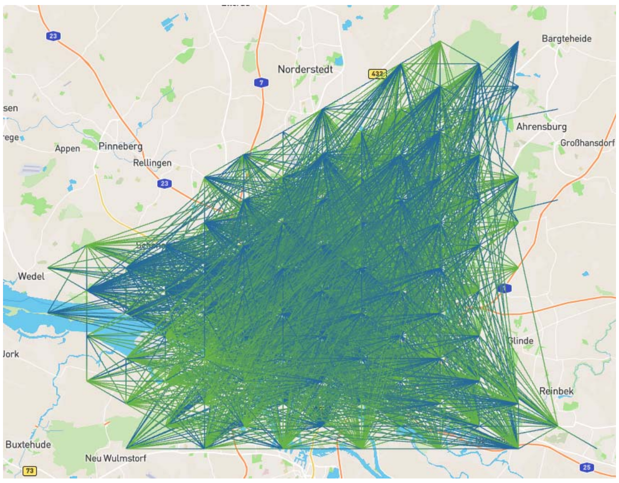


Fig. 16. Map of Hamburg showing **the MiD-based data transformed into a hexagonal representation**. Lines are O-D-Pairs. The line thickness represents the number of underlying trips. Origins are in blue, and destinations are in green. **Trips run from address to address**.

Each line in Figure 14 stands for an O-D connection from one hexagon to another. The blue end of the line represents the origin and the green end the destination. The thickness visually indicates the number of trips that are attributed to each connection. So, while TomTom expresses raw data as trips from hexagons to hexagons, we instead provide individual trips from building address to building address. The result of our approach for the city of Hamburg is shown in Figure 15.

As these approaches are not comparable due to the different resolutions (hexagons vs address to address), we next clustered the raw data from our MiD ODM into similar hexagons, used by the TomTom data set. The results are shown in Figure 16.

When comparing Figure 14 to Figure 16, two things can be observed. First, the TomTom data seems to be more biased towards highways. Second, the MiD data set seems to be more evenly distributed. In addition to that, it becomes apparent that clustering hexagons into city districts might lead to serious biases when compared to an approach that is clustering addresses to city districts. The reason is that by pre-clustering data into hexagons and then into districts, the number of trips in a district could be over- or underrepresented due to the prior reduction in resolution. Vay.io came to the same conclusion as us, recommending rather make use of the MiD-based ODM than the TomTom ODM, as it seemed to be more detailed and less biased. We, therefore, conclude that the comparison to TomTom data should be treated carefully as both the population and data format differ.

VI. Conclusion

In conclusion, this paper presents an approach for generating origin-destination matrices (ODMs) from traffic count data for the city of Munich. The approach is based on data from a study conducted by the German Federal Ministry of Transport and Digital Infrastructure, which includes data on travel behavior in Munich and the surrounding suburbs (MiD-Study). The paper explains how the data from this study was used to create realistic O-D-Pairs and how these pairs were then aggregated to create an ODM representing the demand on Munich's traffic network. The approach is limited to trips originating in the city of Munich and excludes tourism, as well as trips starting outside the city boundaries. The paper also discusses how the likelihood of a trip starting or ending at a home location was estimated based on the distribution of homes in different districts of Munich, and how the ODM was generated based on this information. While the approach was developed specifically for the city of Munich, it is generally extensible to other cities in Germany.

The results show that the insights generated by the MiD-Study were successfully and correctly translated into an ODM, thereby making the data interpretable for traffic simulations using the Sumopy simulation suite for SUMO as demonstrated by Schweizer [20]. A comparison to an external ODM obtained from TomTom showed significant deviations between both datasets. The differences are attributed to different populations and vastly altering collection methods. In addition, a third party (vay.io) was asked to evaluate two new data sets for the city of Hamburg. To create both data sets the same methodology was used. Vay.io conclude that the data set produced by this paper's approach seems to lead to more reliable results. They explain that compared to the ODMs generated from the TomTom data, the results are more detailed and less biased towards highways.

References

[1] Gillen, David W. "Transportation infrastructure and economic development: A review of recent literature." Logistics and Transportation Review 32.1 (1996): 39.

[2] King, W. W. "PORTO RICO. Report from Ponce—Last Case of Smallpox Discharged." Public Health Reports (1896-1970), vol. 16, no. 25, 1901, pp. 1431–1431. JSTOR, http://www.jstor.org/stable/41461819. Accessed 16 Dec. 2022.

[3] Münster, Sander, et al. "How to involve inhabitants in urban design planning by using digital tools? An overview on a state of the art, key challenges and promising approaches." Procedia Computer Science 112 (2017): 2391-2405.

[4] Rosling, Hans, Rönnlund A. Rosling, and Ola Rosling. "New software brings statistics beyond the eye." Statistics, knowledge and policy: Key indicators to inform decision making (2005): 522-530.

[5] Hausken, Kjell. "Probabilistic risk analysis and game theory." Risk Analysis 22.1 (2002): 17-27.

[6] Battigalli, Pierpaolo, and Martin Dufwenberg. "Belief-dependent motivations and psychological game theory." Journal of Economic Literature 60.3 (2022): 833-82.

[7] Yang, Hai, et al. "Estimation of origin-destination matrices from link traffic counts on congested networks." Transportation Research Part B: Methodological 26.6 (1992): 417-434.

[8] Erlander, Sven, Sang Nguyen, and Neil Frederick Stewart. "On the calibration of the combined distribution-assignment model." Transportation Research Part B: Methodological 13.3 (1979): 259-267.

[9] Van Zuylen, Henk J., and Luis G. Willumsen. "The most likely trip matrix estimated from traffic counts." Transportation Research Part B: Methodological 14.3 (1980): 281-293.

[10] Fisk, C. S. "On combining maximum entropy trip matrix estimation with user optimal assignment." Transportation Research Part B: Methodological 22.1 (1988): 69-73.

[11] Spiess, Heinz. "A maximum likelihood model for estimating origin-destination matrices." Transportation Research Part B: Methodological 21.5 (1987): 395-412.

[12] Kawakami, Shogo, Huapu Lu, and Yasuhiro Hirobata. "Estimation of origin-destination matrices from link traffic counts considering the interaction of the traffic modes." Papers in Regional Science 71.2 (1992): 139-151.

[13] Bera, Sharminda, and K. V. Rao. "Estimation of origin-destination matrix from traffic counts: the state of the art." (2011).

[14] Yadav, Piyush, et al. "Traffic prediction framework for OpenStreetMap using deep learning based complex event processing and open traffic cameras." arXiv preprint arXiv:2008.00928 (2020).

[15] Behrisch, Michael, et al. "SUMO–simulation of urban mobility: an overview." Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation. ThinkMind, 2011.

[16] Lopez, P. A., et al. "Microscopic Traffic Simulation using SUMO, 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA." (2018): 2575-2582.

[17] Nobis, Claudia, and Tobias Kuhnimhof. "Mobilitaet in Deutschland-MiD: Ergebnisbericht." (2018).

[18] Huber, Stephan, and Christoph Rust. "Calculate travel time and distance with OpenStreetMap data using the Open Source Routing Machine (OSRM)." The Stata Journal 16.2 (2016): 416-423.

[19] Lenz, Barbara, et al. "Mobilitaet in deutschland 2008." (2010).

[20] Schweizer, Joerg. "Sumopy: an advanced simulation suite for sumo." Simulation of Urban MObility User Conference. Springer, Berlin, Heidelberg, 2014.

8.3 Full-Text Version of Publication 3

Suggested Citation: J. Klinker, M. Avezum-Mercer and S. Jonas, "Improving GPS-Based Mode of Transport Detection in Multi-Modal Trips using Stop Analysis," 2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST), Detroit, MI, USA, 2023, pp. 142-151, doi: 10.1109/MOST57249.2023.00023.

Status: Published.

2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST) | 979-8-3503-1995-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/MOST57249.2023.00023

Improving GPS-Based Mode of Transport Detection in Multi-Modal Trips using Stop Analysis

1st Jens Klinker
School of CIT
TU-Munich
Munich, Germany
jens.klinker@tum.de

2nd Mariana Avezum-Mercer
School of CIT
TU-Munich
Munich, Germany
m.avezum@tum.de

3rd Stephan Jonas
Institute for Digital Medicine
University Hospital Bonn
Bonn, Germany
ORCID: 0000-0002-3687-6165

***Abstract*—This paper presents an extension to existing GPS-based approaches for tracking modes of transportation in multi-modal trips. The extension focuses on analyzing stops and mapping them to surrounding public transport stations in order to improve the accuracy of the mode of transport detection. The proposed method is evaluated using data from the city of Munich, resulting in a 17% improvement of the F1-Score, from 73% to 90%. It is applicable to any GPS-based mode of transport detection system to potentially improve their accuracy.**

***Index Terms*—GPS, Labeling, Data Processing**

I. Introduction

Balancing a group and its individual benefits is a delicate act. Things that are best for the group may hurt some individuals, and vice versa. One scenario to which this applies is public and private transport in urban areas, expressed through traffic jams and overcrowded public transport vehicles. John Nash's "Equilibrium points in n-person games" offers a game theoretical description, explaining the strategic interactions between individuals and groups [1]. In the context of public transportation, this branch of game theory is used to analyze the demand on traffic networks and adapt the network's infrastructure to the demand. These theories need to be based on and validated against the real travel behavior of agents in the system to prove their correctness. This agent behavior is often expressed in the form of trips from an origin to a destination using several modes of transport. To analyze such multi-modal trips, methods for tracking and observing them in the real world are required.

One method utilizes wearable devices. For this purpose researchers developed techniques to passively track individuals' travel behavior using the sensors embedded in these devices. Even though travelers could be asked to manually note and share their travel behavior, they are unlikely to perform this task reliably. For this reason, automatic solutions were developed that identify used modes of transport from device-specific data.

Some of the first publications in this field reach back over two decades. Wolf, for example, looked into inferring modes of transport from data traces in 2000, followed by others such as Chung in 2005 and Tsui in 2006, all focusing on GPS traces [2] [3] [4]. However, the field gained most of its attraction with the introduction of smartphones, as these offer a more user-convenient way to track movement behaviors through embedded sensors. Since then, different smartphone sensors have been tested to predict modes of transport with mixed results, due to differences in the used data sets, used sensors, and used detection algorithms. One of the most promising approaches is achieved through sensor fusion as different sensors have different strengths and weaknesses depending on the circumstances of the traveler [5] [6]. Accelerometers, for example, are specifically potent for tracking physically active modes of transport like walking or riding a bike, but lack accuracy when noise is generated through the user (e.g holding the phone instead of placing it in a pocket). GPS on the other hand is not susceptible to how the traveler handles the phone but requires an unobstructed satellite connection. This connection is easily hindered by tall buildings, trees, or tunnels, causing gaps in the collected data [7] [8]. While most sensors are already well-researched, GPS data offers additional perks. In contrast to accelerometers or gyroscopes, GPS values are represented on a latitude and longitude grid with a time stamp, making this data comparable to context data stored on the latitude and longitude grid. From this connection, a constantly growing amount of information may be derived depending on the amount of available context. Nowadays, websites like google maps or OpenStreetMap constantly add new context to the latitude and longitude network giving researchers the opportunity to add additional context to GPS data (e.g. nearby stores, houses, stations, schools, ...).

Stops in trips are both seen as a nuisance and a carrier of information. On the one hand, they contain no movement behavior and hence no mode of transport-specific movement is detectable. As no movement is recorded, all sensors produce the same data despite the currently used mode of transport. On the other hand, stops are also an indication of forced interaction with the environment surrounding the traveler at the location and time of the stop. For this reason, the exact location and time of a stop are of severe importance when compared to surrounding environmental factors that might be able to explain the stop's occurrence.

This paper introduces an approach that utilizes stops as a source of information, by connecting them to available context to further improve approaches using GPS data to detect modes of transport. This is done by mapping stops in GPS traces to

979-8-3503-1995-8/23/$31.00 ©2023 IEEE
DOI 10.1109/MOST57249.2023.00023

nearby public transport stations. They are thereby used as an indicator for the used mode of transport. The hypothesis is that the more reliably a GPS trace stopped close to a certain type of public transportation station, the more likely it becomes that the GPS trace was generated by that type of transport. As most trips use a combination of modes of transport, the goal is to detect all modes of transport used in a multi-modal trip and to assign them to the correct segment of the GPS trace. In the following sections, an approach for the integration of stops and stations into a mode of transportation prediction algorithm is introduced. It builds upon the approach of Avezum to show how the inclusion of stations as context for stops improves the prediction quality [10]. Similar to Avezum, the approach and data collection conducted in this paper will also focus on the city of Munich.

II. Common Procedures for GPS-based Mode of Transport Prediction

Before engaging into the benefit of a stop to station mapping, some of the existing GPS-based processes for predicting modes of transport are presented.

A. From Uni-Modal to Multi-Modal

Compared to a trip with more than one mode of transport, inferring the used mode of transport from a uni-modal trip is relatively simple. Multi-modal trips can be seen as a sequence of uni-modal trips. But, as the information about a switch from one mode of transport to the next is not provided, an additional dimension for labeling errors is added. This means that the underlying sequence of uni-modal trips inside the multi-modal trip (=: segments) no longer has clear start and end points for each mode of transport. To solve this problem, Zheng proposes a solution that first focuses on identifying the change points from one mode of transport to another before predicting the modes of transport themselves [9]. The reasoning is, that these change points can be used to break up a multi-modal trip into a sequence of uni-modal trips and thereby compressing the complexity back to the uni-modal level.

B. Splitting Multi-Modal Trips into Fixed Segments

To describe multi-modal trips, the term segment is commonly used. A segment is part of a multi-modal trip that was performed using only one mode of transport. Additionally, the term ”real segment” will be used to describe the actual segment as it was labeled by the traveler, while the term ”suggested segment” describes the segment that was calculated by the researchers' algorithms.

To identify the change points in a trip, Bolbol uses a fixed data window of three data points to iterate over the trip data, with a collection frequency of one data point per minute, which leads to a data window duration of roughly three minutes [12]. The window size may vary, as Bolbol uses GPS sensors to generate data points, which require a satellite connection to produce a data point [12]. If no GPS connection can be established, the window size increases until the connection is re-established and a new data point is recorded. Using this method, Bolbol then cut the trip into a list of 3-minute windows. They do not further specify how the last window is handled in the case that only one data point is left in the trip [12].

Each window is then directly treated as a ”suggested segment” and additional features (e.g. average speed) are created for each suggested segment. As a suggested segment is supposed to contain just one mode of transport, it can be treated like a uni-modal trip. As described in the previous section, there are several machine learning models for interpreting uni-modal trips. Bolbol chooses a smart vectoring machine to predict the modes of transport to label each ”suggested segment” [12]. The procedure is repeated until all modes of transport in all suggested segments have been predicted.

A potential weakness of this method may lie in its inflexibility. From a statistical standpoint, each of the real segments in a trip is a population from which samples are drawn. To maximize a sample's significance, an ideal sample would exactly cover the considered population. On a trip level that means having a matching start and end point for the real segment and the suggested segment. Having a perfectly sized suggested segment allows the algorithm to consider the maximal number of data points for predicting the used mode of transport. It also prevents the overlap of two different real segments which could result in misleading data.

Assuming that a real segment of 120 minutes was observed, and the ”suggested segment” drawn from this population was limited to 3 minutes, it would slice the real segment into 40 ”suggested segments”. Of these 40 segments, each has a higher chance of being misclassified as fewer data are contained increasing the impact of outliers.

If instead one window of 120 minutes was created, perfectly matching the ”real segment”, the data amount would increase simultaneously decreasing the likelihood of misclassification. On the other hand, if a ”real segment” of 1 minute is observed using a ”suggested segment” of 3 minutes, two-thirds of the ”suggested segment” now covers data from other modes of transport. These observations can be formalized:

1) The smaller the size of a fixed window, the more likely it is to under-represent the ”real segment” size

2) The larger the size of a fixed window, the more likely it is to exceed the ”real segment” size

This leads to the conclusion that algorithms using a fixed suggested segment size only perform well on data in which the real segment size fits well to the fixed suggested segment size. Therefore, approaches using fixed data windows are likely to overfit if their window size is tailored to the training data and likely to underfit if the window size is not adapted.

C. Dynamic Suggested Segments

As fixed data windows are likely to overfit or underfit, Zheng and Avezum offer a different solution to this problem, based on the high reliability at which walk segments can be predicted with any sensor or model [9] [10]. Both Zheng and Avezum use a fixed data window to iterate over trip data, but instead of differentiating between all modes of transport, they first perform a classification between the classes walk and non-walk in every data window. This means, that they do not automatically generate a ”suggested segment” for every data window [9] [10]. Instead, they add this additional analysis step, that allows them to dynamically adapt a suggested segment’s size to more accurately match the ”real segment” size.

Once a trip (GPS trace) has been divided into data windows with fixed sizes, both approaches start iterating chronologically over these data windows. In the process, each data window is either labeled as walk or non-walk. Once a change from walk to non-walk or vice versa is detected, a change point in modes of transport has supposedly been detected. When a change is detected, the current data window marks the start of a new ”suggested segment”, while the previous ”suggested segment” is closed. Below an example of this process is displayed:

- Walk (w) and non-walk (n-w) labeling:
  (n-w / w / w / n-w / n-w / n-w / w)
- Resulting suggested segments:
  (1st / 2nd / 2nd / 3rd / 3rd / 3rd / 4th)

This approach makes use of a correlation identified by Zheng, in which they describe that most changes from one mode of transport to another are separated by a walk segment [9]. Fortunately for this approach, walking is a mode of transport that behaves quite differently than other modes of transport, making it reliably detectable with comparably low effort. This is due to the low travel speed when walking and the lack of spikes in maximum speeds.

Zheng and Avezum test different window sizes ranging from 30 to 120 seconds [9] [10]. While Zheng got the best results for a window size of 120 seconds, Avezum peaked at a window size of 60 seconds [9] [10]. This can partially be explained by differences in the data sets, as they were recorded in different countries and with different devices. The advantage of this approach is, that it dynamically matches a suggested segment to the real segment sizes, and therefore has the potential to choose the optimal sample for each real segment.

The risk of this method is that it also has the potential to cause large overlaps if two modes of transport aren’t separated by a walk segment, or if walk segments aren’t detected reliably enough. On the other hand, it can also lead to an over-segmentation, if the sensitivity for predicting walks is too high. This can occur in slow traffic situations with many stops, which are typical for traffic jams or densely populated areas with more traffic control measures (e.g. traffic lights).

D. Predicting Modes of Transport

After completing the segmentation process, the multi-modal trip has been broken down into a set of uni-modal trips, the ”suggested segments”. For predicting the mode of transport, Bolbol then use Support Vector Machines (SVMs), while Avezum chooses Decision Trees. Zheng compares four different machine learning models including both SVMs and Decision Trees [9] [10] [12]. The resulting accuracies of Zheng’s comparison are displayed in Figure 1 [9].

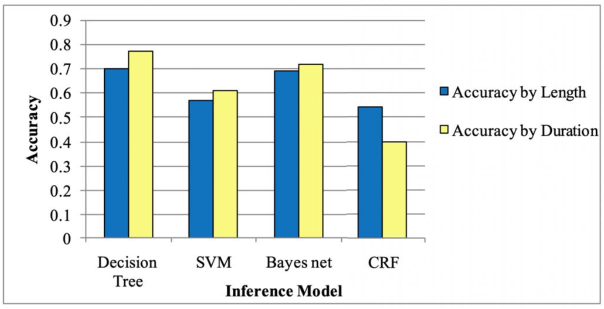


Fig. 1. Comparison of different machine learning models for the prediction of modes of transport in multi-modal trips [9].

According to Zheng, Decision Trees performed best on their data set, combined with their selected segmentation method. As the observed differences were small, they are to be treated with care [9].

E. Likelihood Analysis

Eventually, all segmentation methods result in a list of suggested segments regardless of the priorly applied methods. Additionally, some approaches also perform a secondary analysis based on the conditional likelihood of one suggested segment following the other. In this context, Bolbol, and Zheng analyze whether the sequence of suggested segments is likely to be correct. [9] [12].

III. Stop Analysis

The process of mapping network context to collected GPS traces has already been performed in previous publications. Stenneth was among the first in 2011 to consider public transport as a source of input to improve the mode of transport detection accuracy. They used live location data of buses, trams, and trains in combination with random forests to improve the accuracy of their algorithm from 76% to 93% [14]. This paper will consider static data sources instead, as they are more commonly available than live position data. In the city of Munich for example, live public transport location data is only provided for the S-Bahn, while locations of subways, busses, and trams aren’t published.

A. Creating Basic Features

All GPS-based approaches introduced in the previous section agree on four basic features for inferring modes of transport in any kind of trip. Therefore, in addition to the latitude, longitude, and timestamp, the following features are added to each GPS location in the trips’ GPS log:

1) **Durations** - list of durations (seconds) between GPS

locations

2) **Distances** - list of haversine distances (meters) between GPS locations

3) **Speeds** - list of speeds (km/h) between two GPS locations

4) **Accelerations** - list of speed changes (m/s^2) between three GPS locations.

In preparation for the stop analysis, a fifth and new feature is created:

5) **isStop** - is set to True if the speed of a GPS log element is equal to or smaller than 2 km/h

This new feature indicates if a GPS location is considered to be a stop or not. If the speed value assigned to a GPS location is equal to or smaller than 2 km/h, it is labeled as a stop.

B. Creating Data Window and Suggested Segment Features

Similar to the approaches of Zheng and Avezum dynamic suggested segments are created, by dynamically merging 60-second data windows based on a walk / non-walk discrimination [9] [10]. Sticking to the approach of Avezum, the following features are created to describe a 60-second data window [10]:

1) **TotalDuration** - Sum of the Durations

2) **TotalDistance** - Sum of the Distances

3) **MedianSpeed** - Median of the Speeds

4) **AvgAcceleration** - Mean of the Accelerations

5) **SpeedPercentile85** - 85th quantile of the Speeds

6) **SpeedPercentile99** - 99th quantile of the Speeds

7) **StartLat** - Latitude of first element in data window

8) **StartLong** - Longitude of first element in data window

9) **EndLat** - Latitude of last element in data window

10) **EndLong** - Longitude of last element in data window

When data windows are eventually merged by the walk / non-walk discrimination method these values are agglomerated and used in a similar fashion to describe the resulting suggested segment.

As Avezum reports, the maximum value of a speed carries outliers so the 99th and 85th percentile are used instead [10]. These outliers tend to occur due to connection issues caused by line-of-sight obstruction. When such obstructions occur, GPS signals are displaced in position, causing unnatural jumps in calculated distances and speeds. The impact of such events is lessened by considering the 99th percentile instead of the maximum value. This effect is even stronger for the 85th percentile, which is why it was added as an additional backup.

C. Creating new Stop Features for Suggested Segments

In addition to features used by Avezum, the following five features are newly introduced in preparation for the stop analysis:

11) **NumberOfStops** - Summed number of all stops in the data window

12) **StopDuration** - Summed duration of all stops in the data window

13) **NumberOfbusStops** - Summed number of stops that were labeled as bus stops

14) **NumberOftramStops** - Summed number of stops that were labeled as tram stops

15) **NumberOftrainStops** - Summed number of stops that were labeled as train stops

The NumberOfStops feature is determined by iterating over the part of the GPS trace that was assigned to each data window and counting the number of times that the isStop flag was set to True. The StopDuration feature is calculated by summing up the values stored in the Durations variable, only considering the values that are flagged as a stop. The creation of features 13), 14), and 15) is achieved by mapping the GPS location that is flagged as a stop, to nearby public transport stations of the types bus, tram, and train. As this mapping of stops to stations is a complicated process, it is thoroughly explained in the following section.

D. Mapping Stops to Stations

Stations are timeless objects. They either exist or do not. Therefore, they do not have a GPS location but rather a geolocation. Each station also has a name and a type. Possible station types are train, tram or bus. New stations are created by accessing the open-source traffic network database provided by OpenStreetMap[1], using the Overpass-API[2]. To get all stations, first, a boundary box is defined. The boundary box is a square that is determined by selecting a latitude and longitude for the southwest corner and the northeast corner. For the Greater Munich area, the southwest corner was placed at a latitude of 47.867 and a longitude of 11.224. The northeast corner was placed at a latitude of 48.356 and a longitude of 11.963. The given boundary box is next used

[1] www.OpenStreetMap.org
[2] www.overpass-api.de

to request specific node points that lie inside the box. The following nodes are requested:

a) **["railway"="tram stop"]** - Gets all tram stations from the railway network

b) **["train"="yes"]** - Gets all train stations from the railway network

c) **["subway"="yes"]** - Gets all subway stations from the subway net- work

d) **["highway"="bus stop"]** - Gets all bus stations from the road network

All stations received from a) and d) are directly assigned the types of tram and bus. The stations received in requests b) and c) are both assigned the type train. Each value already contains latitude, longitude, and a station name. Duplicate values of type train are reduced to one value. For the defined boundary box, this results in a list containing 6966 stations and the following distribution:

- Tram: 175
- Train: 769
- Bus: 6022

This list is produced once and can be updated regularly if needed. The update frequency can be set independently, but it is recommended to update the list at most once a week, as the list of stations does not change that regularly.

E. Threshold Determination

Previous sections explain how stops are identified in a data window. Given a stop's latitude and longitude, the next task is to map the stop to nearby stations. The goal of the mapping process is to produce a list of stations that are close to a given stop. Several distance thresholds were first tested to find an optimal discrimination value. The best-performing distance thresholds were 60 meters for buses as well as trains and 70 meters for trams. The difference for trams is likely to be caused by dataset specifics as we were not able to come up with a logical explanation for it. Performance differences for distances between 50 and 70 meters were generally similar, with a slight outperformance for the selected values. We, therefore, recommend considering distance values for stop station mapping that lie between 50 and 70 meters. However, these values might differ for other cities.

F. Efficient Mapping

Previously, a list of stations was created that contains nearly 7000 entries. Different approaches were tested to efficiently filter the stations' list for all stations that are close to a given stop. As the operation for calculating the haversine distance for every entry in the stations' list is not only highly inefficient but also not scalable, a solution was required that reduces the usage of this function by efficiently reducing the number of stations a stop is compared to. To solve this issue, we derived an approach that uses the special traits of the decimal degrees of latitudes and longitudes. The grid that is placed over the earth's globe has cells of different sizes depending on the number of decimal degrees for the latitude and longitude. All stations and all GPS locations in the CEC are created with 5 decimal places, leading to a cell side length of up to 1.1 meters. The grid size varies depending on the latitude position. The closer to the equator, the larger the side length of the grid elements. At a latitude of 0°, the side length of a cell is 1.1 meters. The closer the cell is to the poles, the smaller its side length will become. For this thesis, a minimal side length of 0.43 meters is assumed, which is reached at a latitude of 67°. Our algorithm uses this property to generate a rounded-up version of the stations' list and the stops with 3 decimal places. This means that a cell on the rounded precision level has a side length of 111 to 43 meters. In Munich, where all of our tests were conducted, the side length is roughly 77 meters. With this method, values that have slightly different values are rounded to the same value in the rounded value.

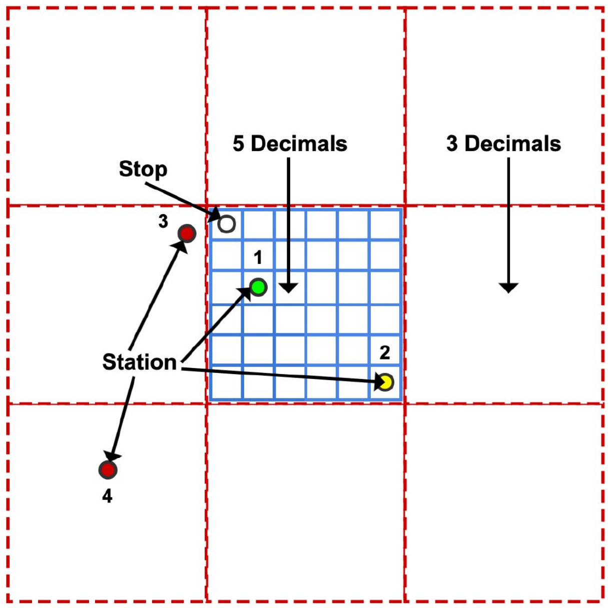


Fig. 2. Cell sizes for a varying number of latitude and longitude decimal places.

This approach makes it possible to simply compare the rounded latitude and longitude of a stop to the rounded latitude and longitude of a station without having to calculate the haversine distance between them. All the indices of the stations found with this method are then used to add the corresponding station from the un-rounded station to a list of candidates. The haversine distance is then calculated for each station in the candidate list. Depending on the station's type a threshold is chosen to determine, if it is close enough to change the stop's "is(tram/bus/train)Stop" feature to True. Figure 2 visualizes the stop-to-station mapping process.

Figure 2 does not only visualize the rounding and mapping process but also unveils a problem with the current approach. In the visualized example, the stop is correctly mapped to station (1) and drops station (2) as it exceeded the threshold. But it did not consider station (3) even though it is actually closer to the stop than station (1). The consequence of the rounding process is a fixed boundary. If a stop and a station are in reality close to each other but were assigned to different cells, then this relation is missed.

To solve this problem, all rounded cells that directly surround the stop's rounded cell are also considered. This is done by adding and subtracting 0.0005 to the 4th decimal of the stop's latitude and longitude before rounding it. With this method 3 to 8 more rounded cells are considered, depending on the stop's position in its rounded cell. Figure 4 shows how this influences the mapping process.

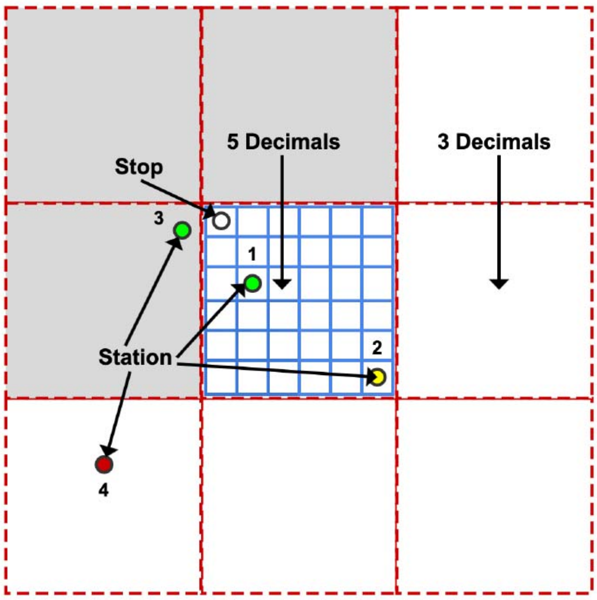


Fig. 3. Cell selection with adapted rounding.

As the stop in Figure 3 is located very close to the top and left rounding boundaries, it is easily pushed into the center-left cell by subtracting 0.0005 from its longitude and adding nothing to its latitude. The top-left tile is reached by again subtracting 0.0005 from the longitude and adding 0.0005 to the latitude before rounding them. The center-top cell is reached by adding nothing to the longitude, and 0.0005 to the latitude before rounding. This procedure is repeated for all of the 9 possible combinations. Only 4 cells were added in Figure 3, as the stop's location inside its rounded cell in this example was too far away from the right and bottom boundaries. Instead of selecting all 9 cells by default, the addition and subtraction ensure to only consider rounded cells that are likely to be inside the distance threshold. Including this method thereby lowers the number of candidates that are on average considered to be close to a stop, while considering all relevant candidates. This then lessens the number of candidates for which the haversine distance needs to be calculated, thereby enhancing the performance of the algorithm and making it scalable to any number of stations in the stations' list.

Every station candidate that also passed the distance test passes its type to the stop. Corresponding to the type the stop's isbusStop, istrainStop or istramStop variable is set to True. This means that the number of stops of the same type nearby does not further influence the result. On the other hand, a stop can be of just one, two, or all of the public transport types. At the same time, a stop can also be of no type.

IV. Merging Data Windows into Segments

This section shows how adjacent data windows are merged into segments. The approach of Avezum labels data windows as a walk or non-walk [10]. This theory is based on insights of both Bolbol and Zheng who observed that transitions from one mode of transport to another are nearly always separated by a walk segment [11] [12]. By detecting these walk segments, the changing point from one mode of transport to another can be identified reliably. According to both Bolbol and Zheng, in GPS-based approaches, walk segments are best detected by looking at the average speed [11] [12]. These results were also reproduced by Avezum [10]. On the other hand, Avezum had issues with over-segmenting their trips due to stops in the data [10]. As any data window with an average speed below 8.8 km/h is predicted to be a walk data window, stops caused the average speed to regularly drop below that threshold. Vice versa, every data window with an average speed above 8.8 km/h is labeled as a non-walk segment. All data windows are iteratively predicted. As long as the predicted class doesn't change, the data windows are added to the same segment. When a change occurs, the current suggested segment is closed and a new suggested segment starts. As this entire process simply distinguishes the two suggested segment types by a speed threshold of 8.8 km/h, the consequence can be the initialization of a new suggested segment whenever a long enough stop occurs that lowers the average speed enough to fall below the threshold.

For this reason, we implemented a new solution that builds up on this concept but adds stop detection to reduce this over-segmentation. The stop detection uses the new data window feature 12) "StopDuration". The StopDuration saves the summed duration of all GPS locations in a given data window that is flagged as stops.

If that duration exceeds one-third of feature 1) "TotalDuration" of the data window, then the window is considered to be a "stop window". As previously described stops are easily mislabeled because a stop looks similar in any mode of transport. If a window is flagged as a "stop window" it is assigned to the same suggested segment as its predecessor without predicting it as walk or non-walk. With this addition, stops no longer influence the segmentation process, thereby reducing noise generated through over-segmentation. Once a suggested segment is closed, the features of its data windows are agglomerated again using the same statistical methods for each feature, as already used in the data windows themselves. An exception is the 5) SpeedPercentile85 and 6) SpeedPercentile99. These two values were previously determined by selecting the 85th and 99th percentile from the data window's Speeds list. On the segment level, the average of the SpeedPercentile85 of all data windows is used to create the segment's SpeedPercentile85. The same is done for the SpeedPercentile99. As both values already select the highest speed values in their data frames, it would be counterproductive to select the highest values yet again. Instead, choosing the mean created more stable results that were less affected by noise outliers in the data.

The values 13) NumberOfbusStops, 14) NumberOftramStops and 15) NumberOftrainStops are also altered for the suggested segment. Previously, they counted the number of stops for which the corresponding label was set to True. In the suggested segment, they are altered to now store the percentage of train, tram, or bus stops in the given segment, by first summing up the values from all data windows and next dividing it by the total 11) NumberOfStops in the suggested segment. The result is a stop percentage for each the public transport types in the suggested segment. Additionally, two new features are introduced that calculate the average distance between stops and the number of stops per kilometer in the suggested segment:

16) **AvgDistBetweenStops** - Average distance between the stops in a suggested segment

17) **NumberOfStopsPerKilometer** - Number of stops per kilometer in a suggested segment

Once all features have been added to the suggested segment, it is labeled using a decision tree.

V. Results

To test the introduced methodology a new data set was collected in the city of Munich. Then two tests were performed on the data set, one using stop analysis and one without stop analysis.

A. Data Collection

The results of the data collection are presented in Figure 4.

The data were collected from 13 different test participants over the course of 6 weeks during the months of June and July 2019. GPS points were generated at a frequency of 1 data point per second, which resembles the maximum frequency for GPS. Even though lower frequencies were tested to be equally accurate, the maximum frequency was chosen, as an artificial reduction of data points can still be performed in later processing steps. Figure 4 shows the distribution of "real segments" in the recorded multi-modal trips by mode of transport. When comparing the different modes of transport it is visible that car data are overrepresented, while bikes are underrepresented. Even though this might appear to be the

Mode of Transport	Data Points	Trips	Avg. Speed [km/h]	Avg. Duration [s]	~ Collected Minutes	~ Collected Hours	~ Share (by Duration)
Walk	8409	42	5.76	2.46	345 Minutes	5.8 Hours	15.8%
Bike	3951	8	16.58	1.14	75 Minutes	1.3 Hours	3.4%
Car	34640	47	50.98	1.34	775 Minutes	12.9 Hours	35.6%
Train	9751	27	67.86	3.23	525 Minutes	8.8 Hours	24.1%
Tram	9223	34	28.35	1.51	230 Minutes	3.8 Hours	10.6%
Bus	9447	27	26.73	1.47	230 Minutes	3.8 Hours	10.6%
All	75421	185	32.71	1.86	2180 Minutes	36.4 Hours	100%

Fig. 4. Results of data collection sorted by real segments (uni-modal) of multi-modal trips as labeled by the study participants.

result of a random collection, this was done deliberately. When looking at possible movement behaviors, cars have the widest range. A car can be as slow as a pedestrian if it is stuck in a traffic jam. Yet, it can also go as fast as a train. Second, a car's movement behavior is far more dependent on the driver than it is on public transport modes. The public transport modes have drivers that were trained to have similar driving behaviors, to be able to stick to specific travel timetables. Third, cars can move freely while public transports have to stick to the road or track network they were assigned to. Finally, cars also differ in their fabrications, leading to further differences in travel speed and acceleration. Overall this leads to a greater variety of behaviors observable in cars, while public transport moves more reliably and thus in clearer patterns. higher speeds or faster accelerations. To take these variations into account, a larger data set was required for cars. In addition to that, different car types were paired with different drivers throughout the data collection period. Among the used car types were a Ford Galaxy, an Audi A4, a BMW Mini Countryman, a Tesla Model 3, a Skoda Oktavia, a BMW 1 and a Mercedes E 220.

Nearly the same arguments that were used for cars could also be applied to bikes. Yet, they are only represented with a 3.4% share in the data set. The reason this collection focuses on cars much more than on bikes is that no CO2 equivalent can be calculated for bikes, while cars are the main emitter in the transportation sector. Also, the city of Munich is not a particularly bike-friendly city. Therefore, it was more challenging to find test participants that regularly use a bike in the city. It was also not an option to artificially increase the number of trips by asking one or two participants that regularly use a bike to focus on collecting more data. Even more than cars, bikes are susceptible to the rider's movement behavior. Consequently, it was not an option to overload the data with bike rides of just two individuals.

Walks were the easiest to record as they occur before and after any mode of transport. To avoid an over-representation of transitioning walks in the data set that only connects two goals, some test participants were randomly selected to also collect independent walk data from aimless shopping trips or strolls in and around Munich.

In contrast to that, it was comparably easy to collect travel data for the public transport, as the individual's behavior, has hardly any impact on the recorded movement behavior of a bus, tram, or train. To increase variance, test participants were asked to even alter their behavior by changing seat positions inside the vehicles between different trips, or to move during a trip.

B. Predicting Modes of Transport using Stop Analysis

After several training iterations on the collection data, the decision tree shown in Figure 5 was produced. Child nodes to the right of a parent node are larger or equal to the value in the parent node, while child nodes to the left are smaller.

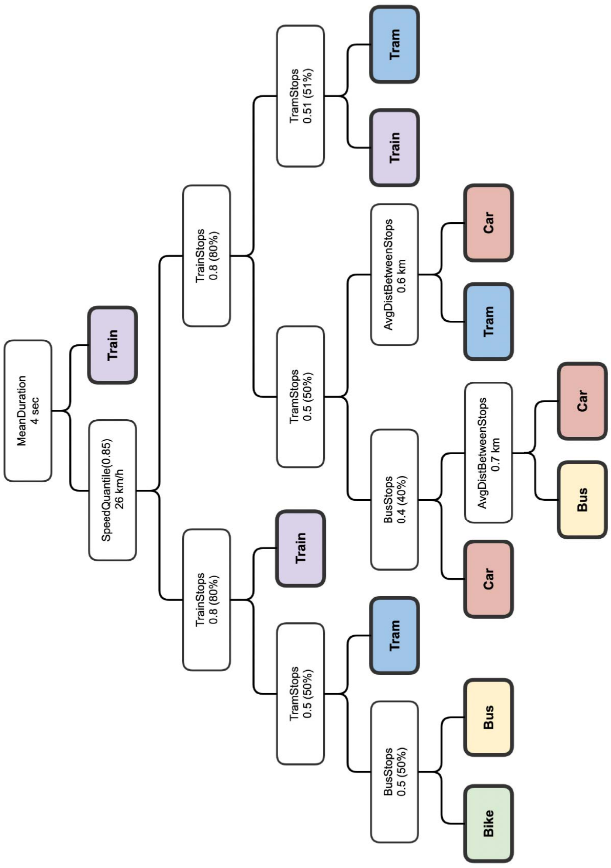


Fig. 5. Decision tree for suggested segments. (Left child smaller than Parent value — Right child larger or equal to Parent value)

An 8-fold cross-validation was used in the training process to avoid overfitting the tree to the data sample. The resulting tree first splits data with respect to MeanDuration. Any suggested segment with a mean duration equal to or larger than 4 seconds is predicted as a train. As the application used for data collection attempts to collect a GPS location every second this indicates severe connection issues to the satellite network. As Munich does not have many long tunnels for cars, buses, trams, or bicycles, this is a telling phenomenon. Neither does Munich have a dense and tall skyline, explaining these connection issues. This means that most modes of transport do not have regularly occurring connection losses. The exception, however, is Munich's train network (Subway and S-Bahn) operates mostly underground in Munich's inner city area. For train trips passing through this area, it is impossible to establish a GPS connection. This is the reason why the mean duration between data point collections is the most reliable discriminator between a train segment and any other mode of transport (except walking). But as the other half of Munich's train network operates above ground this value alone is not sufficient to reliably detect all train segments. Segments by train could also be entirely above ground or a mixture between both above and underground. These hybrids of above and below-ground trips are also the reason for the low mean duration value of 4 seconds.

The 85th speed percentile turned out to be the second node, as it reliably separates a car from a bike trip. Interestingly, the 85th percentile and not the 99th percentile was chosen by the algorithm, indicating that it provides a more reliable discriminator for cars and bicycles. However, for all public transportation modes, this discriminator does not suffice. To separate them from one another and bicycles they were additionally filters for the newly introduced stop type percentages. If 80% of the stops or more were labeled as "trainStops" the suggested segment is labeled as a train ride. Similarly, trams (50%) and then busses (also 50%) are checked. If none of these checks is larger or equal to the checked value, the suggested segment is labeled as Bike.

If the 85th speed percentile was larger than 26 km/h, then the same procedure is conducted for cars. In contrast to the left side of the tree, cars are additionally discriminated by the AvgDistanceBetweenStops. This variable was required as cars may by chance stop close to a public transport station. This regularity increases in the inner city. However, as cars tend to skip some public transport stations while public transport is obliged to stop, an alternation in behaviors was detected. Even though a car may have stopped 3 times in total, each close to a bus, tram, or train station, the distance covered between these stops is rarely comparable due to the previously stated station skipping. This leads to cars having a higher average distance between stops. However, this variable is highly dependent on Munich's transportation network. It is therefore recommended to retrain this value for other cities or neglect it.

C. Prediction Accuracy

The precision values shown in Figure 6 are based on overlapping durations. This means that the duration of a trip is selected as the ground truth for the prediction precision. For example, a value of 0.9 means that out of 100 minutes that were predicted to be car minutes, 90 actually were completed in a car. This value can be derived from summing up the values in the vertical columns. The recall can be observed

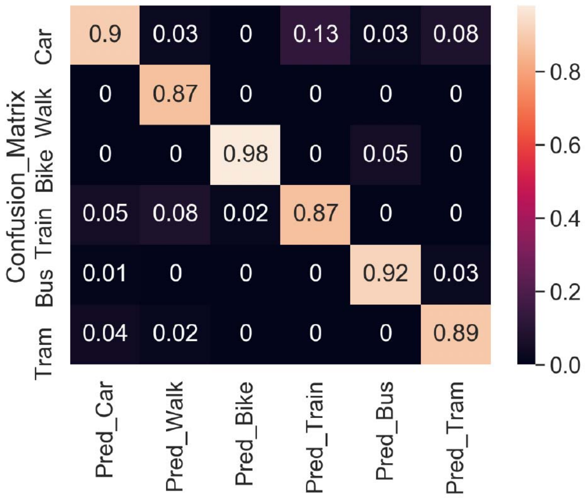


Fig. 6. Confusion matrix showing the precision at which suggested segments labeled correctly.)

by summing the horizontal lines, excluding the cells where the predicted mode of transport and actual mode of transport cross over (values in diagonal running from top left to bottom right). So the recall values were for car 72%, walk 100%, bike 95%, train 85%, bus 96%, tram 94%.

To properly represent the predictive power combining both recall and precision in one variable for each mode of transport we use the F1-Score.

The resulting F1 values are 80% for car, 93% for walk, 96% for bike, 86% for train, 94% for bus and 91% for tram. The resulting average F1-Score is 90%. However, transport modes are not evenly distributed across the dataset (see Figure 4). The following value visualizes the effect of this distribution on the average F1-Score:

*(345 minutes * 93% + 75 minutes * 96% + 775 minutes * 80% + 525 minutes * 86% + 230 minutes * 94% + 230 minutes * 91%) / 2180 minutes = 86.7%*

VI. Conclusion

This paper presents a new approach using stops and public transport stations for the detection of modes of transport in multi-modal routes. The average F1-Score, combining precision and recall for all modes of transport, was 90%. Based on a prior approach from Avezum, the F1-Score was increased by 17% from 73% to 90% [10].

A humanly readable decision tree was produced that may be applied by other researchers in other cities. However, it is recommended to retrain the node values in the tree for the specifics of traffic networks in other cities. Most likely, the values for each decision node will vary for other cities, as some discriminating values are likely to be overfitted to Munich's transportation network. Nonetheless, we do recommend adhering to the tree's general structure and the used features as their discriminatory power are logically sound for other cities, too. However, the applicability in countries that majorly differ from the German infrastructure should be reassessed before implementing the tree structure.

The resulting automatic labeling of GPS traces by mode of transport has various use cases. It could be used to automatically calculate a CO2 equivalent for commuters and travelers, but it could also be used for further research about movement behaviors.

Traffic simulations for example rely on real data to assess their correctness. But obtaining this data is complicated as manual labeling by the traveler is required. From past data collections, we also experienced labeling mistakes caused by careless or unattentive participants. For this reason, we paid special attention to obtaining a small and well-trained participant group for this study, to ensure the correctness of the ground truth. Due to this natural error, using an automated algorithm with an F1-Score close to 90% is likely to be as precise or even more precise than user-based labeling. But, as previously hinted, we do recommend first adapting the algorithm to the specifics of the considered region. This combination of higher data quality and effortless collection from the user's perspective makes automatic detection algorithms quite applicable for other traffic research relying on a ground truth to refer to. It also adds the possibility to label GPS data by mode of transport that was collected in the past. This effectively offers the possibility to transform any existing set of unlabeled GPS traces into a user study that asked travelers to label the mode of transport they were using during their trip.

References

[1] Nash Jr, John F. "Equilibrium points in n-person games." Proceedings of the national academy of sciences 36.1 (1950): 48-49.

[2] Wolf, Jean. Using GPS data loggers to replace travel diaries in the collection of travel data. Georgia Institute of Technology, 2000.

[3] Chung, Eui-Hwan, and Amer Shalaby. "A trip reconstruction tool for GPS-based personal travel surveys." Transportation Planning and Technology 28.5 (2005): 381-401.

[4] Tsui, Sheung Yuen Amy, and Amer S. Shalaby. "Enhanced system for link and mode identification for personal travel surveys based on global positioning systems." Transportation Research Record 1972.1 (2006): 38-45.

[5] Feng, Tao, and Harry JP Timmermans. "Transportation mode recognition using GPS and accelerometer data." Transportation Research Part C: Emerging Technologies 37 (2013): 118-130.

[6] Brondeel, Ruben, Bruno Pannier, and Basile Chaix. "Using GPS, GIS, and Accelerometer Data to Predict Transportation Modes." Medicine and science in sports and exercise 47.12 (2015): 2669-2675.

[7] Soloviev, Andrey, et al. "Assessment of GPS signal quality in urban environments using deeply integrated GPS/IMU." Proceedings of the 2007 National Technical Meeting of the Institute of Navigation. 2007.

[8] Morales, Yoichi, and Takashi Tsubouchi. "GPS moving performance on open sky and forested paths." 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2007.

[9] Zheng, Yu, et al. "Learning transportation mode from raw gps data for geographic applications on the web." Proceedings of the 17th international conference on World Wide Web. (2008): 247-256.

[10] Avezum, Mariana, Jens Klinker, and Bernd Bruegge. "Transportation Mode Recognition on Multi-modal Routes Based on Mobile GPS Data." 2019 4th International Conference on Intelligent Transportation Engineering (ICITE). IEEE, (2019): 141-146.

[11] Zheng, Yu, et al. "Understanding mobility based on GPS data." Proceedings of the 10th international conference on Ubiquitous computing. (2008): 312-321.
[12] Bolbol, Adel, et al. "Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification." Computers, Environment and Urban Systems 36.6 (2012): 526-537.
[13] Prelipcean, Adrian C., Gyözö Gidófalvi, and Yusak O. Susilo. "Transportation mode detection–an in-depth review of applicability and reliability." Transport reviews 37.4 (2017): 442-464.
[14] Stenneth, Leon, et al. "Transportation mode detection using mobile phones and GIS information." Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems. (2011): 54-63.

Total number of pages in the document: 106