

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Visualization of service reliability of public transportation

Tiago José Grosso Pacheco

DISSERTATION REPORT



Mestrado Integrado em Engenharia Informática e Computação

Supervisor: Teresa Galvão Dias

Second Supervisor: Thiago Sobral

June 28th, 2019

Visualization of service reliability of public transportation

Tiago José Grosso Pacheco

Mestrado Integrado em Engenharia Informática e Computação

June 28th, 2019

Resumo

Melhorar a fiabilidade dos transportes públicos é importante não apenas para aumentar a atratividade destes serviços para a população em geral, como também para minimizar os custos de operação das redes de transportes por via do aumento da eficiência na alocação de recursos. *Frameworks* de visualização podem ser extremamente úteis como ferramentas de apoio à decisão para que operadores da rede de transportes públicos identifiquem problemas, tendências e padrões acerca da fiabilidade do serviço.

O número de passageiros e o cumprimento de horário são as duas principais dimensões para se avaliar a qualidade do serviço; no entanto, a visualização destas duas dimensões pode tornar-se bastante dispersa se a informação for apenas mostrada ao nível de cada linha, daí que tal visualização deverá ser interativa e capaz de alterar a granularidade dos dados, tendo, simultaneamente, a capacidade de ajustar rapidamente o intervalo de tempo desejado. Determinar o cumprimento de horário poderá ser conseguido correlacionando dados de AVL (Localização Automática de Veículos) com o horário da linha. Transformar estes dados em informação útil que se relaciona com a consistência do serviço continua a ser tópico de discussão, com alguns autores a optar por usar o tempo médio de espera para os passageiros como um indicador, outros estudam as vantagens de usar uma abordagem baseada em *buffer time* e outros escolhem estudar o cumprimento de horário como um todo.

Uma das causas mais significativas da redução da fiabilidade do serviço é o *vehicle bunching*, que pode ser difícil de visualizar quando a abstração dos dados é demasiado elevada. Outro problema com *frameworks* existentes é a falta de filtros de procura para localizar mais facilmente as secções pretendidas do sistema de transportes. Por último, como as redes de transportes públicos podem ser extremamente extensas, há um vácuo por preencher de funcionalidades que apontem o utilizador para potenciais problemas e que permitam que esse utilizador navegue os dados de forma eficiente e eficaz.

Esta dissertação propõe uma *framework* de visualização, denominada TransViz, orientada para a análise da fiabilidade de transportes públicos, adotando uma abordagem centrada no utilizador que segue os princípios de Interação Pessoa-Computador (HCI). Como caso de estudo, serão usados dados de transportes públicos da área de Grande Boston, obtidos através da Massachusetts Bay Transportation Agency.

A avaliação da *framework* desenvolvida foi realizada com um grupo selecionado de *domain users* de operadoras de transportes públicos e por investigadores da área dos transportes. O processo de design encontra-se descrito de início ao fim e os resultados são discutidos de forma a apresentarem conclusões relativamente ao trabalho de dissertação e à iteração atual da *framework* TransViz.

Abstract

Improving the reliability of public transportation is important, not only to increase the attractiveness of these services to the general population, but to minimize the transportation network operation costs by increasing its resource allocation efficiency. Visualization frameworks can be very useful as decision support tools for transportation domain users to identify issues, tendencies and patterns regarding reliability and quality of service. Ridership and schedule adherence are the two main dimensions for evaluating the quality of service; however, the visualization of these two dimensions can become quite disperse if the information is only shown on a route level. Hence such visualization should be interactive and enable the change in the granularity of the data while having the capability to rapidly adjust the desired time frame. Determining schedule adherence can be done by correlating AVL (Automatic Vehicle Location) data with the route's schedule. Transforming that data into usable information that relates to service reliability remains a topic of discussion, with some authors opting to use the passengers' average wait time as an indicator while others study the advantages of using a buffer time approach and others foregoing those measures and evaluating schedule adherence as a whole. One of the most significant causes of undesirable service reliability is vehicle bunching which can be cumbersome to visualize at the higher levels of abstraction. Another problem with existing frameworks is the lack of search filters to more easily locate desired sections of the transportation system. Lastly, since the public transportation network can be overwhelmingly extensive, there is an unfilled void for features that direct the focus of the user to potential problems and allow them to effectively and efficiently navigate the data. This dissertation proposes a visualization framework, entitled TransViz, oriented towards the analysis of the reliability of public transportation adopting a user-centred approach that follows the principles of Human-Computer Interaction (HCI). As a case study, the Greater Boston region public transportation data, provided by the Massachusetts Bay Transportation Agency will be used. The evaluation of the developed framework was carried out with a group of selected domain users from public transportation operators and by researchers in the transportation area. The design process is described from beginning to end and the results discussed in order to provide conclusions regarding the dissertation work and the current state of the TransViz framework.+

Acknowledgements

I would like to express my gratitude to both my supervisor, Prof. Teresa Galvão and my co-supervisor, Thiago Sobral, for both the opportunity to work on this topic and the tremendous support and advice given throughout this dissertation.

I would also like to give thanks to everyone who gave up a substantial amount of their time to help me evaluate the dissertation work.

Finally, I would like to thank my family and friends. They made this whole endeavour much easier than it would have been otherwise.

Tiago Grosso

This work is partially financed by the ERDF - European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme and by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e Tecnologia within project POCI-010145-FEDER-032053

“What is great in man is that he is a bridge and not an end.”

Friedrich Nietzsche, *Thus Spoke Zarathustra*

Contents

1	Introduction	1
1.1	Overview	1
1.2	Context	1
1.3	Motivation and Objectives	2
1.4	Dissertation Report Structure	3
1.5	Summary	4
2	State of the Art	5
2.1	Overview	5
2.2	Interaction Design	5
2.2.1	Requirements	6
2.2.2	Human-Computer Interaction	6
2.2.3	Information Visualization	7
2.3	Key Performance Indicators	8
2.3.1	Schedule Adherence	9
2.3.2	Headway Regularity	9
2.3.3	Travel Time	10
2.3.4	Wait Time	10
2.3.5	Buffer Time	10
2.3.6	Stop Accessibility	11
2.3.7	Ridership	11
2.4	KPI Visualization Techniques	11
2.4.1	Table-Based Visualization	11
2.4.2	Graph-Based Visualization	12
2.4.3	Calendar-Based Visualization	13
2.5	Influential Factors	14
2.6	Developed Tools and Frameworks	14
2.7	Summary	17
3	Methodology	19
3.1	Overview	19
3.2	Research	19
3.3	Development	20
3.4	Evaluation	21
3.4.1	Case Study and Data Set	21
3.4.2	Testing Methodology	23
3.5	Public Transportation Reliability	24
3.6	Summary	26

CONTENTS

4	TransViz	27
4.1	Overview	27
4.2	Objectives	27
4.3	Initial Requirements	28
4.3.1	Other Functionalities	29
4.4	Non-Functional Prototype	29
4.4.1	Focus-group evaluation	30
4.5	Requirements Revision	32
4.5.1	VIZ01 - Stacked Columns Chart	34
4.5.2	VIZ02 - 24 Hour Clock	36
4.5.3	VIZ03 - Vehicle Location Chart	37
4.5.4	VIZ04 - Map	38
4.5.5	VIZ05 - Space-time Diagram	40
4.5.6	VIZ06 - Colour Calendar	40
4.6	General Evaluation	41
4.7	Proposed Guidelines	42
4.8	Summary	43
5	Conclusions	45
5.1	Overview	45
5.2	General Conclusions	45
5.3	Real World Applications	46
5.4	Objectives Satisfaction	46
5.5	Future Work	47
A	Usability Tests for feedback collection regarding the TransViz prototype	53
A.0.1	Introduction	53
A.1	Test 1 – Stacked Columns Chart	54
A.2	Test 2 – 24 Hour Clock	54
A.3	Test 3 – Vehicle Location Chart	55
A.4	Test 4 – Map	56
A.5	General Questions	56

List of Figures

2.1	An example of the Waterfall process lifecycle [20]	7
2.2	Timeboxes in the RAD model. (Adapted from [10])	8
2.3	Average passenger waiting time spatial distribution for route 15 westbound am peak hours [18]	13
2.4	Space-Time Diagram of the Dublin Bus route 46A, outbound, No. 8th, 2012. [5]	14
2.5	An example of a calendar for ridership values visualization. Values closer to red are considered undesirable [16]	15
2.6	An overview of the MetroViz tool	16
3.1	Downtown map of the MBTA public transportation system [2]	22
4.1	The main page of the non-functional prototype. Notice the hamburger button on the top left	30
4.2	The main page of the non-functional prototype with the expanded drawer menu .	31
4.3	The data visualization page of the non-functional prototype	32
4.4	The data visualization page of the non-functional prototype with the "new visualization" overlay	33
4.5	The Stacked Columns Chart visualization in "lines" mode	34
4.6	The Stacked Columns Chart visualization in "stops" mode	35
4.7	The 24 Hour Clock visualization	36
4.8	Vehicle Location Chart	38
4.9	Map Visualization	39
4.10	Example space-time diagram [5]	40
4.11	Colour Calendar Visualization	41

LIST OF FIGURES

List of Tables

2.1	Table-Based KPI Visualization for some indicators [13]	12
3.1	AVL Data Example	22
3.2	Schedule Adherence Data Example	23

LIST OF TABLES

Abbreviations

AVL	Automatic Vehicle Location
APC	Automatic Passenger Counting
API	Application Programming Interface
DIS	Deviation index based on stops
EIS	Evenness index based on stops
HCI	Human-Computer Interaction
KPI	Key Performance Indicator
MBTA	Massachusetts Bay Transportation Authority
PIR	Punctuality index based on routes
RAD	Rapid Application Development
UX	User Experience

Chapter 1

Introduction

1.1 Overview

This chapter provides context for the dissertation work. It explains the public transportation environment and the problems in determining and visualizing the reliability of transportation networks. It clarifies how work in this area can help transportation domain users and researchers identify patterns in these networks using key performance indicators for service reliability and explains the critical role of visualizations in the decision making process of public transportation domain users, which makes clear the motivation behind this dissertation.

The structure of this report is presented at the end of this chapter with a brief description of each chapter.

1.2 Context

Public transportation is a complex topic with multiple branches of study. One of these branches is the study of service reliability, which generally refers to the probability that a system or service will perform its intended function properly during a certain period of time. In the context of public transportation, the concept of service reliability is not limited to deviations from schedule and advertised services: a reasonable distribution of passengers and between vehicles and the number of people which can use the service are also components of service reliability of public transportation. Therefore, service reliability of public transportation is a measure of the capability of public transportation networks to consistently provide the scheduled services with quality in regards to aspects such as passenger load and time [6].

Knowing what service reliability is is still not enough because there still needs to be an answer to the question: how can the reliability of a public transportation network be calculated? There are many attempts at answering this question over the course of the last few decades with many indicators being proposed, such as schedule adherence¹, vehicle-bunching² and buffer time³ [24].

¹How well a vehicle keeps up with its schedule

²How close together and evenly spaced the vehicles in a route are

³Extra travel time required to allow passengers to arrive on time to their destination

Determining the service reliability of a public transportation network is often done by developing indexes suited for the time of evaluation being done. Such indexes aim to reduce the complexity of the vast amounts of data into a meaningful number. For example, regarding vehicle-bunching, an index can be used to get an idea of the headway regularity of vehicles in a route. [13].

The study of the service reliability of public transportation networks plays a major role in determining the efficiency at which public transportation networks run in terms of resource allocation and management as well as how attractive the service is for the population. For example, the travel time reliability has been associated with the satisfaction reported by public transportation users [7], not just because unreliable travel times can generate delays for the user but also because it harms the perception the users have on the service they are using. Travel time, however, is not the only indicator to look at. Headway regularity, i.e., how evenly spaced vehicles are in high-frequency routes; the waiting time for the users; and transfer times also impact the reliability of public transportation services [24].

Another important factor to take into account when analysing the reliability of the service is the scale at which the data analysis is being performed. Since a public transportation network is comprised of many elements of different types, there is the possibility to look at the service reliability at the stop, route and network level. Passengers might be more sensitive to issues at the stop level since it is where their perception is focused on, where the resource allocation for a transportation company might be underperforming due to issues at the route and networks levels.

1.3 Motivation and Objectives

Service reliability of a public transportation network is related to user satisfaction and the resource allocation efficiency of the service providers. As such, improving reliability can be extremely advantageous in reducing costs associated with inefficiencies in the network and increasing profits via the attraction of new users and the retention of existing ones. Moreover, there is an interest in cities, especially larger ones, to have a network of public transportation that serves its population, industry and services, providing an extensive coverage while compromising as little as possible in reliability and efficiency.

Creating an efficient public transportation network requires an evaluation of service reliability. However, visualizing service reliability is not a trivial task. Even with the use of indexes, the data is still quite dispersed and it is vital to analyse and correlate data from different parts of the network. Furthermore, poor performance in a stop or route is not enough information for the decision making process. There needs to be a myriad of other data that the user can easily access in order to establish the causes for a certain problem. For example, route length, distance from a stop to the bus terminal and the use of exclusive bus lanes are factors which have been identified as potentially influential of service reliability [13] and transportation domain users might be interested in assessing how these indicators are affecting a certain network.

Lastly, the time factor is also to be considered. The times of the day, week, month and year influence the usage and performance of a public transportation network and, as such, need to be

taken into account when analysing the data. An issue in the network might only appear on rush hours or on holidays, so the ability of a transportation domain user to discriminate data via a rather complete manipulation of the time frame is also an important requirement.

The objective of this dissertation is to investigate and develop a framework for the visualization of service reliability in public transportation. First, it requires a definition of the KPIs to be used for measuring service reliability. Secondly, research into the factors that might influence such KPIs, such as route length and the number of buses in a route will be performed. Lastly, the transportation domain users' requirements will be assessed and used to guide the development of the framework.

During the development of the framework, a data set will be used to test the intended functionalities and ascertain if the proposed requirements are being fulfilled. This dissertation will make use of data acquired from the MBTA V3 API as a case study for that effect.

1.4 Dissertation Report Structure

This Dissertation Report follows a structure that aims to provide a basic understanding of the concepts discussed and the current state of the work on this area of research before moving into a more detailed discussion of a solution to the problems presented. The report will also explain the needs that the developed framework needs to address and the requirements it must fulfil, as well as how the evaluation of that framework was performed. It will then describe the followed design process and finally, it will present conclusions on the work done so far and its potential impact, and will lay the road ahead for all the future steps of this dissertation.

As such, Chapter 2 describes the state of the art of the research and tools developed on the topic of service reliability in public transportation with an emphasis on the study of KPIs and influential factors for public transportation reliability. A critical revision of some of the existent tools and frameworks is presented with a focus on ways that they could be improved during the course of this dissertation work.

Chapter 3 describes the methodology that will guide this dissertation, with insights into the research process and the design process for the framework to be developed. It goes on to describe the evaluation methodology that was used to validate the dissertation work and the case study was used to give real-world meaning to the developed prototype. Finally, it describes the selected KPIs to be integrated into the developed framework and the additional information that should be integrated into the framework. It establishes the requirements that the framework must fulfil as well as the functionalities that it should implement.

A description of the design process along with its results and the evaluation of the prototypes will be discussed in Chapter 4 along with an explanation of how those results have shifted the focus and requirements of the framework.

Lastly, the conclusions are presented in Chapter 5 with a report on the objectives satisfaction, the potential applications of this dissertation's work and the future work that could be carried out on top of the developed framework.

1.5 Summary

Public transportation networks are extremely complex systems that require specialized tools to be properly analysed. There are a number of influential factors and performance indicators that can be used to measure service reliability of public transportation and its causes, but its effect are quite clear. Service reliability impacts the networks' resource allocation efficiency and the appeal it has for potential passengers, making it imperative to create tools that can be utilized to analyse public transportation networks and support the decision making process of public transportation domain users. This dissertation work followed the principles of HCI to create a visualization framework for that purpose.

A set of objectives that this dissertation aimed to achieve have been defined. The work involved the definition of the relevant key performance indicators that were to be used during the development of the framework and research on how to visualize those metrics in the most comprehensive way was then followed by a design process which produced a functional prototype of the framework.

Chapter 2

State of the Art

2.1 Overview

This chapter describes the previous work in regards to service reliability of public transportation and its visualization. Since visualization is one of the pillars of this dissertation work, this chapter will develop the subject of visualization and HCI before going through the more theoretical research on the topic of service reliability of public transportation, which includes the Key Performance Indicators to measure service reliability at various levels and in different circumstances and it explores how these KPIs have been visualized.

Lastly, it goes through some of the tools and frameworks that have been developed to try and visualize service reliability.

2.2 Interaction Design

Interaction Design is based on creating user experiences with the aim of enhancing the way people work, communicate and interact with systems[28]. It can also be explained as designing around the why and the how of users' daily interaction with computers.

A lot of components make up interaction design since it takes into account the user's cognitive processes and their limitations, as well as the limitations of the systems for which something is being designed. It can be said that the user experience (UX) is the central pillar of Interaction Design. UX encompasses all aspects of the user's interaction with all parts of a product, system, service or company [26] which means that every physical product or piece of software is subject to the scrutiny of a UX evaluation.

One important aspect to take into account is that user experience cannot be designed but one can design for user experience [28]. An illustrative example of this characteristic is the cellphone, which can be designed to be light, sturdy, fast and bright and, if designed correctly, will invoke the user experiences of comfort, safety, responsiveness, among others. UX design has the goal of creating positive sensual, cognitive and emotional user experiences.

2.2.1 Requirements

The first and perhaps most important task in the design process is the definition of the requirements that will guide the project, which requires an understanding and discussion about the users, their capabilities, tasks and goals and the constraints and conditions under which the product/service will be used. [28]. In software engineering, requirements can be divided into two types: functional and non-functional requirements. Functional requirements specify the capabilities of the system, such as business rules, certification and authentication functionalities, among others. Non-functional requirements describe the constraints there are on the system and its development [28].

2.2.2 Human-Computer Interaction

Human-Computer Interaction differs from Interaction Design in matters of scope, with the latter being much wider. HCI narrows the focus of Interaction Design to "the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" [31] and, as such, it relates to creating positive and powerful user experiences in computer systems.

2.2.2.1 UX Design Methodologies

There are many methods that can be followed when developing a system such as the Waterfall process, The Rapid Application Development model and Agile development approaches.

The Waterfall process is one of the earliest and simplest forms of methodology for software development [20], and is so named due to its linear sequence of lifecycle activities, each of which cascades into the next one, resembling a waterfall, as illustrated in Figure 2.1.

The Waterfall process benefits from its simplicity: not only it is easy to understand and implement with easily identifiable milestones, it places an emphasis on documentation for each phase and source code, which means that new team members have an easier task when familiarising themselves with the project [19]. However, this methodology is not suited for changing requirements that can come from evaluations and unexpected difficulties, leading to increased costs from modifying the problem deep into the development phase [27].

The Rapid Application Development (RAD) model puts more emphasis on an adaptive process rather than planning. RAD can be characterized by small development teams of both developers and users who can make design decisions; timeboxes (see Figure 2.2), which are delivery deadlines and should be met even at the cost of cutting requirements; incremental prototyping and phased deliveries; the use of rapid development tools and highly interactive, low complexity projects [10].

The RAD model is equipped and even expects the change in requirements of the course of the design process. It involves the user in the whole process and is inherently iterative which, by means of rapid prototyping, can increase create creativity through quicker user feedback. However, early

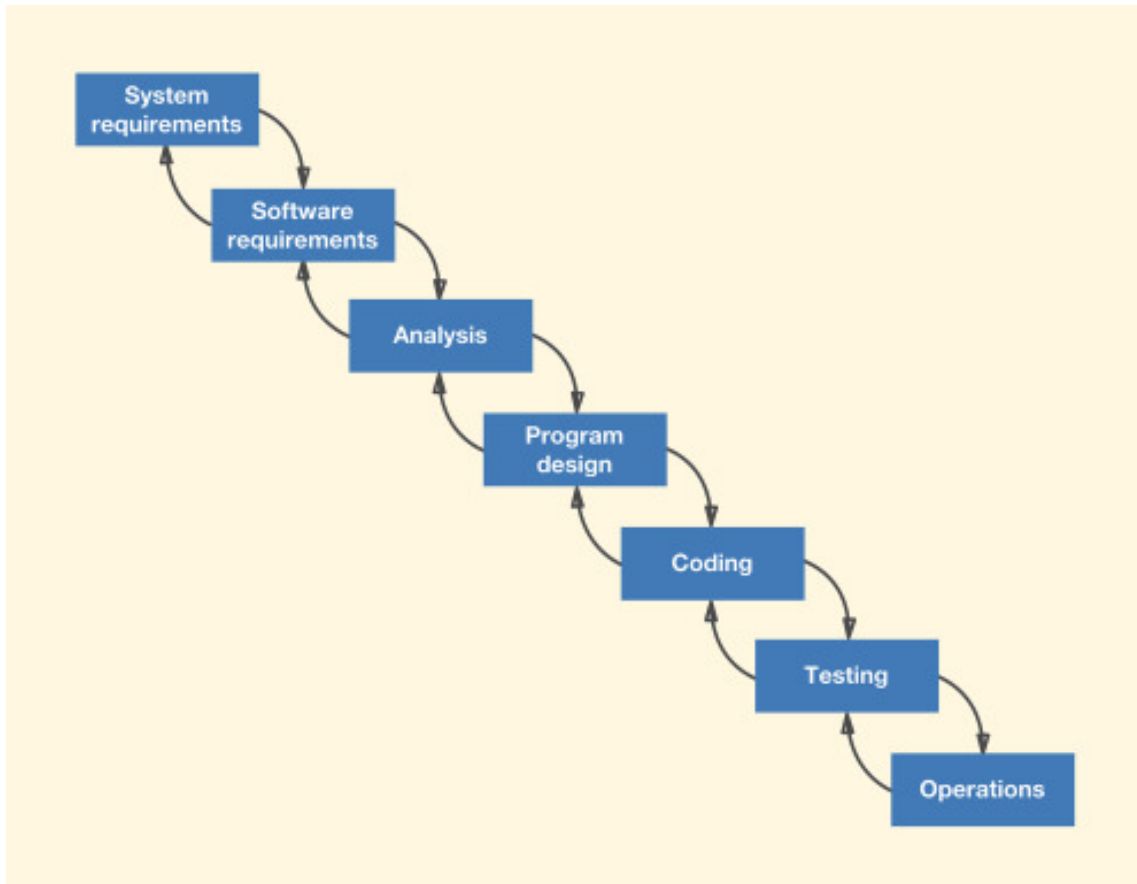


Figure 2.1: An example of the Waterfall process lifecycle [20]

prototypes can lead to a premature commitment to a design and to feature creeping which can inflate the design to an unmanageable scale [33].

The Agile model is typically an iterative approach to development where the requirements and features evolve through the effort of cross-functional teams alongside the system's end user [14]. There are a number of agile development methods, such as Extreme Programming, Scrum and Feature Driven Development, among others [4].

As with the RAD model, Agile processes respond well to change and uncertainty. This methodology brings the end user, potentially a customer, closer and more involved in the project due to frequent deliveries. However, its heavy reliance on functional tests and its short iterations can negatively impact the usability of the system [15].

2.2.3 Information Visualization

Interaction design plays a key role in visualization: a data set can be powerful and a tool can be feature complete but if the visualization of the relevant aspects of the data is not done properly, the user experience will be poor and the actual effectiveness of the tool will be significantly reduced. Information visualization techniques are computer-generated graphics that represent complex data, while typically being both interactive and dynamic, with the goal of amplifying human cognition

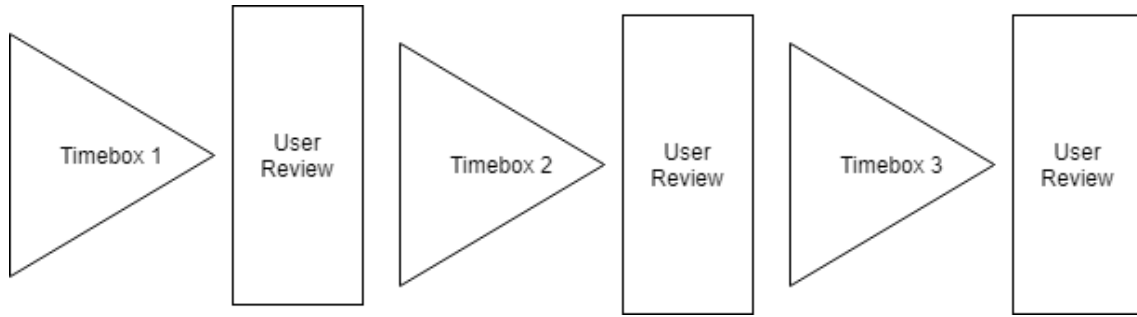


Figure 2.2: Timeboxes in the RAD model. (Adapted from [10])

and enabled users to make otherwise difficult or impossible inferences such as recognizing patterns, trends and anomalies in the data [11].

Information visualization techniques can reduce the time and effort necessary to draw conclusions and inferences about a certain topic or data set and, as such, they allow users to perceive things they couldn't easily perceive otherwise [28]. The principles of interaction design apply here: the intent and mindset of the users is important in deciding how to construct a visualization. Some of the factors that influence the development of visualizations are the data characteristics (dimensions, granularity, continuity, etc.); the visualization objectives (comparison, trend over time, distribution, etc.); and the reasons for visualization (discover, summarize, present, identify, etc) [30]. A visualization can be evaluated on its effectiveness, its expressiveness, readability and interactivity.

2.3 Key Performance Indicators

A Key Performance Indicator(KPI) is a measurable value that reports on well a system, service or company is performing. In the context of this dissertation, KPIs point to the reliability of the public transportation network.

Over the years, research has been made on many (potential) KPIs for evaluating the reliability of public transportation networks. The most prevalent ones in research are Schedule Adherence, Headway Regularity, Wait Time and Travel Time. There are other KPIs that have also been investigated, although to a lesser extent, such as Buffer Times, Transfer Times, among others. Measuring these KPIs requires access to certain types of data that are not collected by the infrastructure of all public transportation operators. In particular, some require access to Automatic Vehicle Location (AVL) data, others to Automatic Passengers Counting (APC) data and others to both (mostly to be measured to a higher degree of precision).

KPIs can be divided into two groups: physical indicators and psychometric indicators [29]. Physical indicators describe the system as it is while psychometric indicators describe it as it appears to be. Psychometric indicators are much harder to calculate and require access to a large number of inputs from public transportation passengers if one does not wish to rely on algorithmic estimations of psychometric indicators derived from vehicle and schedule data alone.

2.3.1 Schedule Adherence

Schedule Adherence, often referred to (or measured by) On-Time Performance, is a measure of how well a network performs at accomplishing its schedules: if a route suffers from consistent delays, then its Schedule Adherence will be low. It is important to take into account that vehicles arriving earlier than scheduled also contributes to poor Schedule Adherence and early arrivals do now counterbalance later arrivals; on the contrary, they should be added up.

The exact definition of the limits of Schedule Adherence is a case for debate. A survey based on 146 answers from bus operators showed that most operators use the definition of no more than 1 minute earlier and no more than 5 minutes later than scheduled [7], with an almost complete agreement that this is an important indicator for service quality and reliability in the context of public transportation.

The impact of Schedule Adherence is greater on low-frequency routes[24] since passengers tend to plan their arrival to stops in a way that minimizes their waiting time. high-frequency routes are defined as those where the frequency of vehicles is smaller than a reasonably threshold for Schedule Adherence, which means that passengers will not feel a significant impact on early or late arrivals of vehicles.

Measuring Schedule Adherence requires data regarding the schedules of the vehicles and their arrival times at each stop. Schedule adherence can be measured as a recurrence of values beyond certain thresholds [18], as the average difference between arrival and scheduled time at stops [16], or visualized with a visualization tool [9].

2.3.2 Headway Regularity

Headway Regularity refers to the uniformity of distance between vehicles performing service in a route or line. While many other indicators related to headway can be measured, such as the comparison between actual and scheduled headway, a Headway Regularity index can be used to measured vehicle-bunching situations.

We say that vehicle-bunching occurs between two (or more) vehicles if the distance between them remains below a certain threshold for a significant amount of time (or stops). Vehicle-bunching has been associated with many reliability and efficiency issues in public transportation networks, including uneven wait times and passenger counts as well as overcrowding [18]. It is also a self-feeding pattern: a late vehicle will encounter more passengers which will increase the boarding time and the delay of that vehicle. The next vehicle on the line will run faster due to a decrease in boarding time caused by the higher numbers of passengers in the previous vehicle. In fact, there is a tendency for buses, for example, to pair together over the course of their service in a route [25].

On routes with high-frequency services, where Schedule Adherence is no longer an appropriate indicator, Headway based measures, such as Headway Regularity, play an important role [12] since passengers tend to arrive at the stops in random intervals of time. As such, Headway Regularity can be used as an indicator of service quality and reliability[8].

Measuring Headway Regularity requires data regarding the location of the vehicles at each moment. That location could be reasonably estimated by their arrival times at each stop for the effects of calculating a precise enough measure of Headway Regularity. Headway regularity can be visualized on a space-time diagram or calculated using a regularity index[13] which measures how evenly distributed the vehicles are at the stop or route level.

2.3.3 Travel Time

Travel Time refers to the time elapsed between the arrival of a vehicle at two stops. Most Travel Time indicators use Travel Time distributions [22].

A range of physical indicators can be measured only from data on the vehicle arrival times, from standard deviations of the scheduled time to the percentage of late trips and threshold-based tardiness indicators[24]. Travel time has been studied with the use of an index defined as the difference between an upper percentile for the travel time during the selected time interval and the median travel time across some days[34].

2.3.4 Wait Time

Wait Time is a measure of the time passengers wait on a stop for the arrival of a vehicle performing the service they are seeking and represents one of the most important components of service reliability perception for public transportation passengers.

Wait Time Indicators can be separated into two categories [35]. Mean-Variance Indicators measure an Excess Wait Time which is the difference between the Average Wait Time and the Scheduled Wait Time. Scheduled Wait Time is defined as the average wait time for passengers if the service was operating as scheduled. The other category, Extreme-Value Indicators is used with the assumption that passengers are more sensitive to extreme values in their Wait Time and attempt to measure the probability of passengers waiting for more than a certain amount of time for their vehicle to arrive[24].

Measuring Wait Time indicators requires data from the arrival time of vehicles at stops. For high-frequency routes, wait times can be estimated as half the headway between vehicles based on the assumptions that passengers arrive randomly at stops and catch the first vehicle[17].

2.3.5 Buffer Time

Buffer Time indicators are related to the extra time a passenger should reserve for the expected completion of a trip[24]. Buffer Time is usually defined as the difference between a certain percentile and the average travel time. Buffer Time is used as a service reliability indicator because they are indicative of other problems in the network, from Headway Irregularity to poor scheduling or inconsistent travel times. As such, Buffer Time indicators can be used on a first stage analysis to identify issues in the network which would then be followed by a more detailed analysis into the specific problems that are occurring. It is also extremely relevant to the passenger perception and experience of the public transportation service.

Buffer Time can be measured using data from the arrival time of vehicles at stops. The difference between the sum of the actual travel and wait times and the scheduled travel and wait times results in the buffer time. Buffer time indicators can be determined by the recurrence of extreme values or by means of an average of the calculated buffer times [24].

2.3.6 Stop Accessibility

Stop Accessibility refers to how many people are at a reasonable distance from a stop. In general, the more people can reach a stop, the better for a public transportation operator since it means more potential passengers. However, extreme unevenness in the accessibility to stops can be harmful to the reliability of the network, as it could create points of overcrowding and points of under crowding. Stop Accessibility could also be expanded to how easily a person can reach a stop by walking or using public transportation which would grant another layer of analysis regarding the connectivity of the network.

Accessibility to the stops has also been proposed as indicative of service reliability and some accessibility maps have been created to attempt to study that correlation [32].

Stop Accessibility requires location data for the stops in each route to be determined.

2.3.7 Ridership

Ridership refers to the number of passengers using public transportation. It is a measure of service reliability[16] not only because it speaks to the core business part of public transportation operators but also because extreme values of ridership contribute to a decrease in efficiency and perceived reliability by the passengers.

Ridership can be measured using APC data. Ridership indicators can be based on the average passengers[16] or simply the total number of passengers in a certain section of a network[23].

2.4 KPI Visualization Techniques

Although KPIs are measurable values, the sheer amount of data on public transportation networks hinders the meaningful visualization of those values due to the rapid changing nature of the data and, therefore, of the KPI values themselves. Because of that, methods have been developed and studied to facilitate such visualization, ranging from a detailed calendar navigation that displays the selected KPI at various degrees of granularity to dispersion graphs which better illustrate the fluctuation of values beyond reasonable thresholds.

2.4.1 Table-Based Visualization

Table-Based Visualization techniques reduce KPIs to a number or series of numbers. In Table 3.1, several indicators are presented in the form of indexes: Punctuality index based on routes (PIR), which measures the probability of an on-time arrival at the terminals; Deviation index based on stops (DIS), which defines the probability that a bus will maintain the headway between

Route Number	Reliability		
	PIR	DIS	EIS
Route 1	0.795	0.378	0.443
Route 34	0.891	0.605	0.526
Route 39	0.617	0.530	0.466
Route 44	0.430	0.476	0.244
Route 45	0.538	0.566	0.122
Route 57	0.663	0.442	0.263
Route 101	0.756	0.702	0.452
Route 108	0.671	0.451	0.494
Route 125	0.569	0.719	0.315

Table 2.1: Table-Based KPI Visualization for some indicators[13]

successive buses at each stop; and Evenness index based on stops (EIS), which describes how even the headway between vehicles is along a route.

As the table illustrates, this type of visualization can be used to easily compare the indexes between routes. However, it does not help the user identify exactly where problems or patterns are occurring since each indicator is reduced to a value for the whole route. If the raw values used to calculate the indexes were used instead, the density of the data would increase and more dimensions would be added (such as the stop and the temporal dimension), which would hinder the ability to compare values from different sections of the network.

This approach also suffers from a lack of scalability. If a transportation domain user intends to analyse the entirety of the network, without restricting it to a small number of routes, the data density would make it so that it would take an additional effort to extract meaningful information from the table. Identifying issues would require an increased amount of work for the user.

2.4.2 Graph-Based Visualization

Graph-Based Visualizations can be compelling due to the plethora of conclusions they allow a user to reach. One of their main advantages is that they allow the rapid comparison between different sections of the network, making it much easier to find over and underperforming routes. They also make it possible to visualize the evolution of the system over time with bombarding the user with numbers.

Figure 2.3 shows an example of a graph that offers the ability to compare the wait time for passengers in each of the buses presented.

Space-Time Diagrams are a specific type of graph that is often used in the context of public transportation for visualizing the headway of vehicles and identifying vehicle-bunching problems and trip irregularities[5]. Figure 2.4 shows an example of such a diagram where a transportation domain user would immediately identify the occurrence of some headway issues while also having the capability of easily analysing and comparing bus speeds and detecting problematic times of the day for the route.

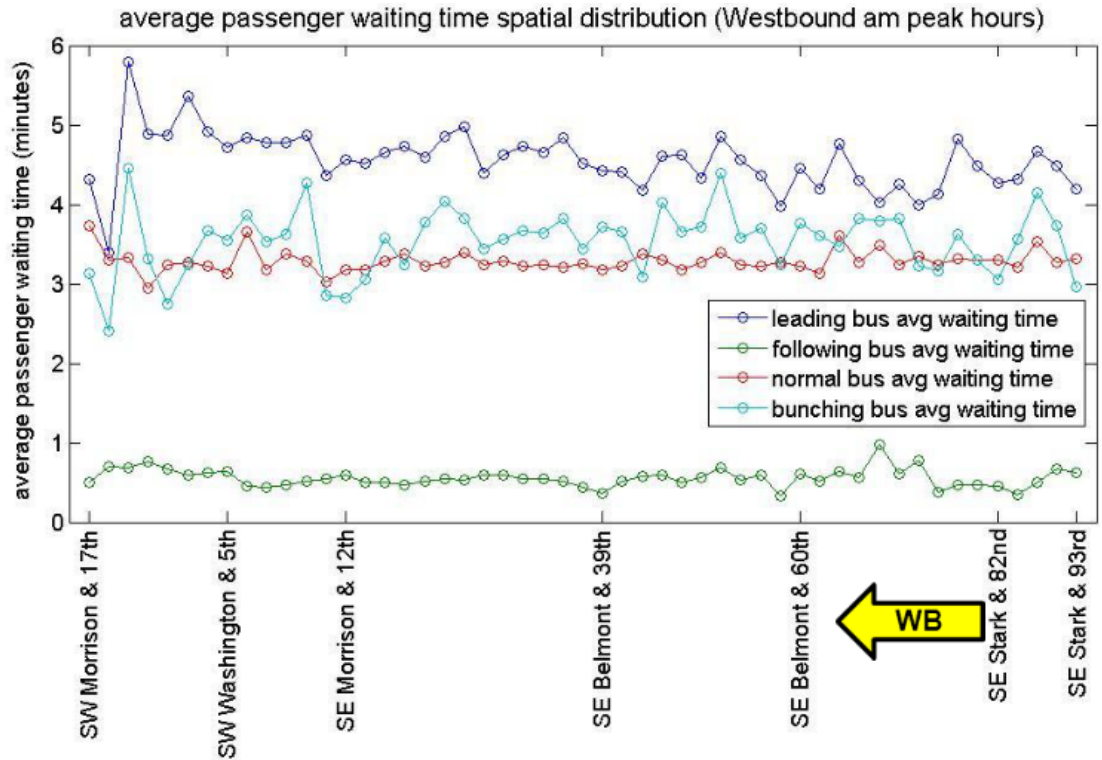


Figure 2.3: Average passenger waiting time spatial distribution for route 15 westbound am peak hours [18]

Although graphs allow for simple comparisons, intuitive interpretation of the data being presented and the extraction of a vast amount of information, they also suffer from the scalability issues that hinder the use of Table-Based Visualizations: packing the information of several routes into a graph requires the time-frame to be the same for all routes; visualizing more than one indicator in several routes increases the effort necessary to make inferences. Nevertheless, graphs, much like tables, are quite versatile and can be used in just about every scenario with a reasonable degree of usability.

2.4.3 Calendar-Based Visualization

Calendar-Based Visualizations, as the name implies, present an interactive calendar that can be used to navigate large amounts of temporally separated data. By specifying a time range, the user can be presented with the data that he intends to see, be that in a table, graph or other forms.

As such, one might assume that a calendar would only be used to navigate data and not exactly visualize it. However, calendars can employ colour to easily draw the user's attention to potential problems in the network. Figure 2.5 illustrates just that: as the user looks to the calendar, he immediately sees that there was an undesirable ridership value on November 8th, 2011.

Thus, a calendar can be paired with other visualization methods to provide the user with the high-level status of the network's reliability as well as a grainier sense of the data being presented

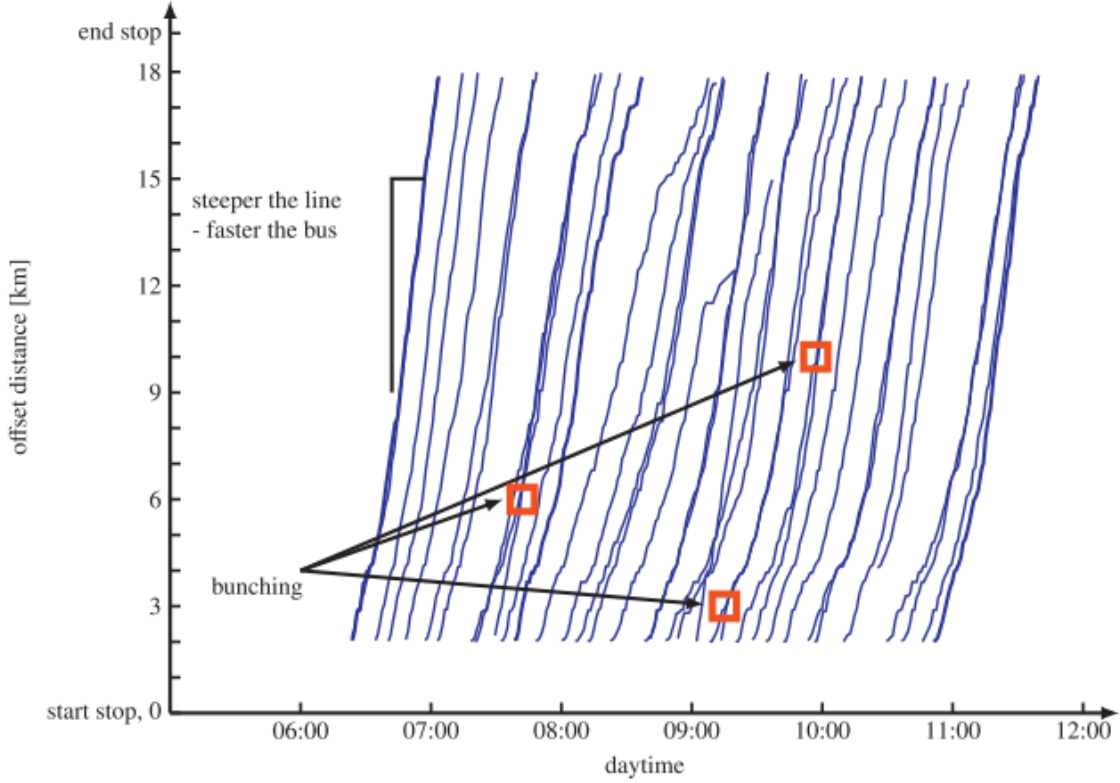


Figure 2.4: Space-Time Diagram of the Dublin Bus route 46A, outbound, No. 8th, 2012. [5]

to him in order to extract valuable information from a disperse data set.

2.5 Influential Factors

KPIs are extremely useful to detect the performance of a system or, in this case, a transportation network. However, there needs to be an understanding of the factors that influence those KPIs or there will be no progress made towards the intended goals. Those factors might be related to time, weather, location, vehicles and passengers themselves.

Correlations have been found between service reliability and distance from the stops to the origin terminal; route length; scheduled headway and the use of exclusive bus lanes[13].

There are other factors which are very likely to influence service reliability but are difficult to evaluate in such a way that the findings could be generalized. Examples of those are the driver attitude, the state and facilities of the vehicles and stops, information at the bus stop regarding schedules and destination and bus fares and discounts [21].

2.6 Developed Tools and Frameworks

Research on service reliability of public transportation using AVL and APC data has mostly intended to define service reliability or develop algorithms to predict travel time or optimize certain

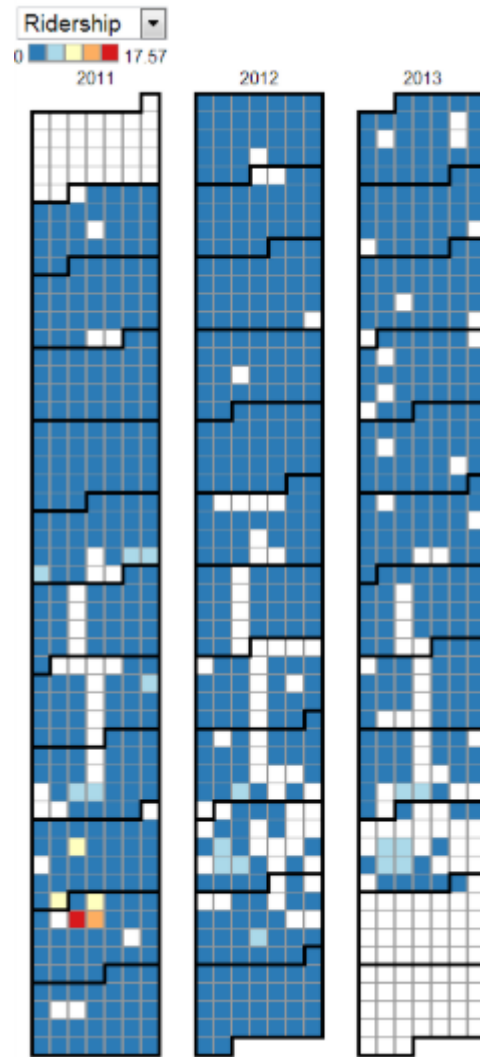


Figure 2.5: An example of a calendar for ridership values visualization. Values closer to red are considered undesirable [16]

aspects of the network. Nevertheless, research has also led to the development of some tools and frameworks aimed at allowing a transportation domain user to easily assess the performance of each section of the network. However, many gaps in those tools and frameworks still exist.

MetroViz is a tool for visual analysis of public transportation data [16]. MetroViz is composed of three levels, the stop/station level, route level and trip level; and three views: the map view, route view and calendar view.

The idea of this tool is to present the user with an overview of the network and the ability to navigate to the desired section using the map to select a route or stop. On the right side, the user can use the calendar to adjust the time frame and granularity of the data being presented. On top of the calendar, the user can select the type of data to visualize (ridership and adherence).

MetroViz makes extensive use of colour to display information and status: the selected route and/or stop is highlighted on the map, the calendar uses colour to give a high-level, the type of

fares are colour coded, and so on. However, once the time frame and section are selected, MetroViz presents data using several single colour bar graphs.

This tool succeeds in creating efficient data navigation for a large data set and in displaying system status to the user. Its capabilities as an effective decision support agent remain to be evaluated as the authors only evaluated its usability.

MetroViz, as the authors conclude, suffers from long load times, an excessive amount of scrolling and the inability to sort routes by adherence and ridership. Other voids not referred by the authors include the lack of filters for the search results, the lack of system-wide alert that directs the user's attention to potential issues and no configuration whatsoever. Most importantly, though, MetroViz does not allow for the correlation of any type of influential factor with the values of the ridership and adherence indicators (or, as a matter of fact, for the visualization of any of those factors besides the fare type), nor does it allow for the comparison between different routes.

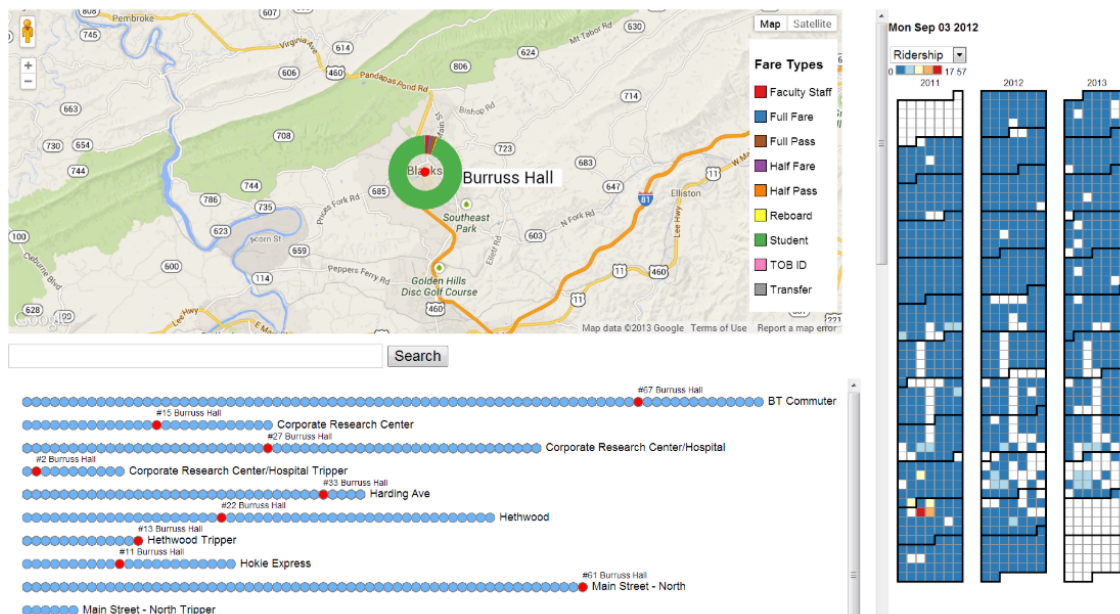


Figure 2.6: An overview of the MetroViz tool

Another framework based on a Buffer Time indicator makes use of AVL data to create a service reliability visualization [24]. The framework aims to be a first step in studying the use of AVL data for measuring service reliability and it far from complete or robust, being limited to presenting several graphs and charts that measure Buffer Time indicators. It is, by no means a visualization tool and it only presents static data.

A more complete approach, still with no meaningful data navigation, has been made following a "snapshot" approach [18]. This framework is superior to the one previously referred in almost every way since it is able to display a wider range of indicators and can not only make use of graphs and charts but it also displays the information dynamically on top of a map, creating a very compelling visualization of the information regarding a certain route. The "snapshot" part of the framework comes in the form of time controls that allow the user to move forward and backwards

in time to see the data from different time periods, while also providing a Play feature that makes the data go forward in time automatically.

2.7 Summary

The research on the topic of service reliability for public transportation is extensive, yet it is not completely solidified. Much of the research is focused on studying certain indicators of reliability which has left a void for connections between different factors and indicators.

The visualization of service reliability can be invaluable when it comes to the decision-making process but there is also another aspect of visualization tools that can be helpful for the future of research in this field and that is the identification of correlations and patterns that might warrant further investigation into what service reliability means and how it should be measured.

The use of software tools is beneficial for visualizing the vast amounts of data that exists for public transportation. However, there is a lack of such tools and a lack of visualization frameworks which provide insights on how those visualizations should be built and what they should achieve.

On the next chapters of this report, the definition of service reliability in the context of this dissertation will be explained out and the problems presented will be met with a set of proposed solutions based on interaction design principles for visualization.

State of the Art

Chapter 3

Methodology

3.1 Overview

This chapter describes the methodology that guided the realization of the dissertation work. It describes the approach made into researching and how the development of a prototype followed the design process. It also presents the case study that was used for evaluation purposes and how the data was put together to fit the needs of the dissertation.

3.2 Research

Research into public transportation is vast and disperse and when it comes to its visualization of service reliability, it's hard to find all the topics connected in a single place.

As such, research was divided into different parts that aimed to establish the foundations on top of which this dissertation work would be built. The first part consisted of understanding the importance of an efficient and well performing public transportation network so that the main objectives for the dissertation could be contextualized. The second pertained to the definition of a meaning for service reliability in the public transportation environment, as well as how that reliability could potentially be measured or analysed. These two parts lead to the creation of a specific vocabulary for the context which aimed to standardize, in the context of this dissertation, the many different ways that have been used to describe the same issues and factors over the years (E.g.: "low-frequency routes" and "high headway routes" refer to the same thing) .The third part of research had to do with investigating the visualization aspect of the dissertation: how the network might be visualized and navigated; how certain indicators might be presented in an easy to understand and meaningful fashion; among others.

The research aspect of this dissertation stretched itself over the course of most of the work, although it was taken more and more into the background as the work progressed and the development phase takes priority. Nevertheless, studying the design process, interaction guidelines and visualization techniques was a regular activity throughout the whole dissertation.

3.3 Development

The development of the framework followed a user-centred design process based on the principles of Interaction Design for Visualization in order to develop a framework that took into account the user's needs while also having a high degree of usability and usefulness.

For the development of the TransViz prototype, the RAD model was followed. The choice of following the RAD model was based on its iterative nature, which corresponded well with building somewhat independent functionalities, with the delivery of each visualization and refinement of the previous one corresponding to each timebox. This way, the feedback about the strengths and shortcomings of each visualization could be used to shift the requirements of the next visualizations in order to create visualizations that complement each other.

The Waterfall process was ruled out due to both the iterative nature of the development of the TransViz and the need for a somewhat high involvement by the end users which would provide expert feedback on what a service reliability visualization needs to achieve. The use of Agile models was not justified by the scale of the development project which did not require the use of a multi-disciplinary team to be completed.

Since the framework was to be used by transportation domain users, an analysis of such users was made to create the first set of requirements. These were divided into Functional Requirements, which specify the capabilities of the system, such as business rules, certification and authentication functionalities, among others, and are vital to the system, taking priority during the development phase; and Non-Functional Requirements, which are linked to the user experience, such as how aesthetically pleasing the interface is, how closely the system matches the user's mental model, how responsive it is, how it retains the users attention, how clearly it displays information, and so on. Defining the requirements has to be done with the user's needs and goals in mind.

The next step was the development of non-functional prototypes for discussion in a focus group setting. This prototype was invaluable as it was the basis for the collection of a large amount of early feedback which shifted the initial requirements. These prototypes broadly illustrated how the requirements would be fulfilled and how the user would interact with the final product and were evaluated in questionnaires and focus groups scenarios.

The development followed an iterative approach with each iteration producing a new visualization and refining the usability of the previous ones by means of feedback collection and implementation.

The next phase saw a more thorough evaluation of the functional prototypes either through usability tests. The feedback collected from this phase was registered and the changes/ideas proposed have either been prototypes to be evaluated again or documented for future work.

The evaluations played a crucial role in understanding the prototypes strengths and shortcomings, as well as providing ideas for future work on top of the developed framework.

3.4 Evaluation

3.4.1 Case Study and Data Set

The evaluation of the developed framework required the application of a real life scenario. For that effect, public transportation data was collected from the Greater Boston Area through the V3 API, provided by the MBTA. The information from the API was filtered and compiled in order to obtain a robust data set for a subset of the vast network which was then inserted into the framework for testing and evaluation purposes.

The creation of the data set aimed to provide relevant information for service reliability measurements. As such, not only did the data set contain AVL data for vehicles on the selected routes, it also contained information that would allow calculating deviations from schedule. The data set was complemented with the static part of the network's data, i.e., data regarding the location of stops, among others.

The Greater Boston area was selected as a case study for developing and evaluating TransViz because of the accessibility to its data: MBTA¹ provides the V3 API² for free which can be used to obtain a plethora of information regarding real-time schedules, vehicle location, routes, trips, stops, among others.

The MBTA public transportation network encompasses Subway Lines, Bus Routes, Commuter Rail Lines, Ferry Routes and The RIDE - a door-to-door service for users who cannot easily use or access the rest of the system, totalling over 200 routes and lines [1]. Ridership values for MBTA services are very high, totally 1,297,650 average passengers per weekday across all services as of April 2019 [3]. Figure 3.1 shows a map of all the lines and routes of the MBTA public transportation services in the downtown area of Boston. A full map of those services beyond the downtown area can be found at the [MBTA Website](https://www.mbta.com).

Since the selected case study encompasses hundreds of lines and multiple types of vehicles, its scope was shrunk to encompass a few of the major lines and routes of the Greater Boston area. In particular, Route 1 and Route 747 were selected for buses; for subways, the Red and Green line were selected, which totalled almost half a million passengers per weekday as of April 2019 [3]. The Green line is subdivided into Green-B, Green-C, Green-D and Green-E lines.

The data set had the purpose of allowing for measuring the proposed KPIs on the selected routes and lines. As such, it required data on the schedule of the vehicles, their location and their arrival at stops.

AVL data was stored in a CSV file with each line following the structure:

```
Vehicle ID; Update Time; Latitude, Longitude, Route ID, Direction ID,
Next Stop ID
```

With this structure, it becomes easy to track each vehicle and it also becomes trivial to aggregate vehicles by route.

¹Massachusetts Bay Transportation Authority - <https://www.mbta.com>

²<https://api-v3.mbta.com/>

Methodology

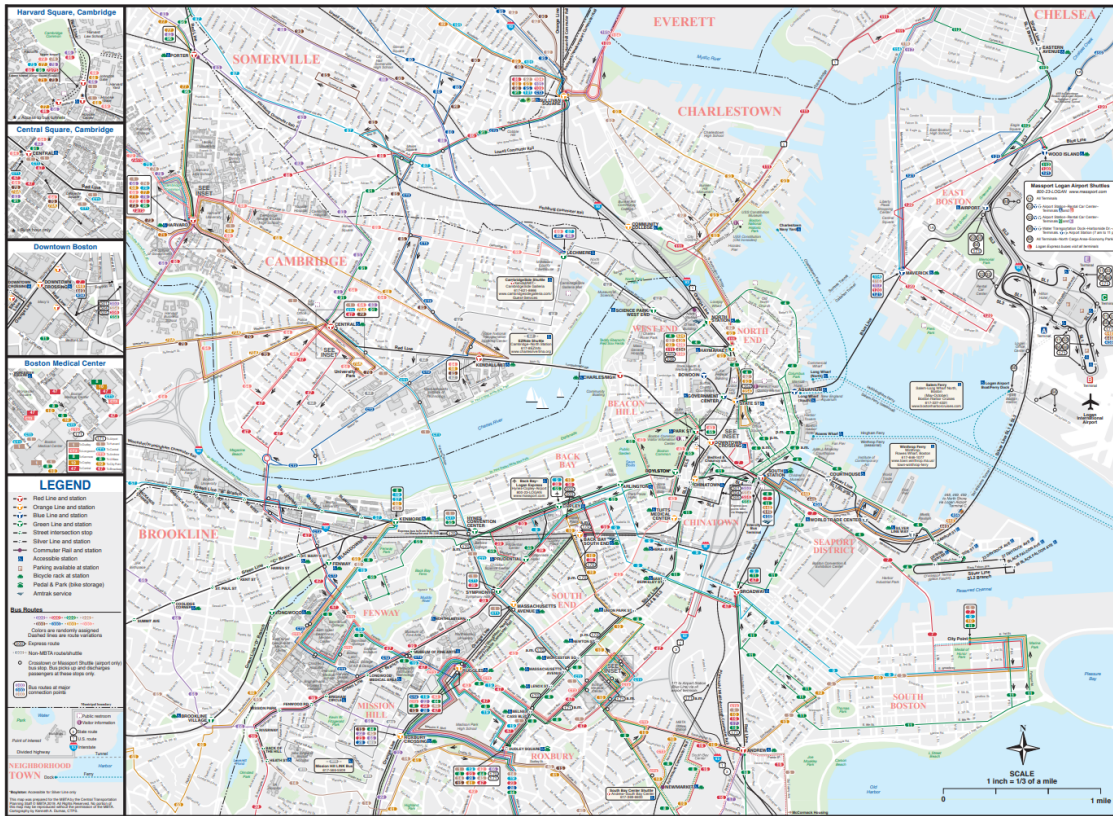


Figure 3.1: Downtown map of the MBTA public transportation system [2]

Vehicle ID	Last Updated At	Latitude	Longitude	Route ID	Direction ID	Next Stop ID
G-10120	2019-01-09T09:37:49-05:00	42.34838104248047	-71.13526916503906	Green-B	1	70128
R-545A8DC2	2019-01-09T09:37:16-05:00	42.32057189941406	-71.0525894165039	Red	0	70085
y1900	2019-01-09T09:37:46-05:00	42.337711334228516	-71.07845306396484	1	0	87

Table 3.1: AVL Data Example

Data regarding the schedule and arrival of vehicles at stops was more disperse, requiring multiple different API calls and the aggregation of the retrieved data from each call. As such, results for each of the selected routes were stored in different CSV files, each with the structure:

```
Trip ID; Next Stop ID; Update Time; Scheduled Arrival At Next Stop;
Predicted Arrival at Next Stop
```

It's easy to notice the unexpected use of Predicted Arrival data. This was done because the API does not provide information on the actual arrival times of each vehicle at stops but it does provide a prediction based on an MBTA algorithm. That prediction is regularly updated, which means that, at most, the difference between the real value and the collected value is one minute, which is the refresh time of the API call.

Methodology

Trip ID	Next Stop ID	Last Updated At	Scheduled Arrival	Predicted Arrival
39366150	83	2019-01-09T09:37:27-05:00	2019-01-09T09:38:00-05:00	2019-01-09T09:37:26-05:00
39366162	87	2019-01-09T09:37:48-05:00	2019-01-09T09:32:00-05:00	2019-01-09T09:39:33-05:00
39366265	77	2019-01-09T09:37:46-05:00	2019-01-09T09:40:00-05:00	2019-01-09T09:38:07-05:00

Table 3.2: Schedule Adherence Data Example

Since the V3 API does not give historical data, a JavaScript program was created to extract information periodically. Every minute from January 9th to January 23rd, 2019, that program was called via the Windows Task Scheduler and the data extracted was appended to the different files. V3 API provides a substantial amount of information in relatively small bundles but it does so by extensive use of IDs to connect the various levels and elements of the network. At any given moment, vehicles have associated with them IDs for their trip and their next stop and trips have an ID for the route to which they belong. As such, the JavaScript program made several consecutive calls to the API each time it was invoked to obtain all the necessary information and joined it together in the specified files.

This data was complemented with GTFS data to also encompass the names and locations of stops. Three major problems appeared during the creation of the data set:

1. For some stops, their ID existed in the V3 API but not on the GTFS data which made it so their geographical coordinates were not obtained;
2. A significant amount of lines of data collected using the V3 API came with 'null' or 'undefined' values. Those values were disregarded from the data set;
3. There was no information on which stops were part of each route. This problem was somewhat overcome by making a list of all the stops that appeared in the file for each route's arrivals.

The next step for this data set was to clean the data. For example, for a reasonably accurate estimation of the arrival of a vehicle at a stop, there is no need to store all the predictions for that vehicle and stop, only the last one. The data set contains 2.089.413 lines of data over 8 files.

3.4.2 Testing Methodology

Evaluating the various stages of this dissertation work is crucial to validate its results, assess the decisions made and implemented in the framework and prepare the next iterations of the prototype. For testing purposes, this dissertation used usability tests and focus groups. Before the tests, the users heard an explanation of the goals and scope of the TransViz framework. After each test, a few questions were asked to classify the user experience in interacting with each visualization and the system as a whole.

After the creation of a non-functional prototype, a focus group setting was carried out with five researchers in the area of transportation and mobility with the aim of reevaluating the requirements and obtaining more detailed insights into the positive and negative aspects of the current

iteration of the TransViz framework. This focus group was characterized by a discussion of the actions being performed by the users and what the framework displays after each action in order to ascertain the mental model users are creating for the system and ways to potentially enhance the user experience in the next iteration of TransViz. The users were showed each part of the prototype before engaging in a discussion regarding the screen in front of them. This focus group significantly reshaped the initial requirements and the scope of the TransViz framework and was an essential step before advancing to the development of a functional prototype.

Usability tests were conducted with researchers and transportation domain users to evaluate the usability of the functional prototype and to ascertain how useful such a tool could be for public transportation operators. These tests took place in the final phase of the design process where not many changes could be implemented. Nevertheless, the evaluation results were documented and included in the discussion of results in the next chapters of this report. The script for the usability tests is included in Annex A. During the usability tests, the users were asked to perform a series of tasks representative of the system's functionalities with the end goal being the identification of tasks that are problematic and warrant a different design approach. These tests were conducted individually with each user which allowed for a final discussion with the user on the various aspects of the framework and future work that could be done in this subject.

3.5 Public Transportation Reliability

Including every researched KPI would increase the complexity of this dissertation to a point where time would only permit the development of a broad and shallow system. Since access to information on APC data is very limited, the developed framework will not make use of indicators that depend on APC data. It is important, however, to choose indicators that serve for both low and high-frequency routes and that convey a large amount of usefulness.

3.5.0.1 Schedule Adherence

In the context of this dissertation, Schedule Adherence will be measured as the percentage of vehicles beyond a certain threshold of tardiness or earliness. This approach makes it so that small discrepancies from the schedule, which are not significant, are not taken into account and it also makes it so that the KPI is easy to visualize. Schedule Adherence (SA) can be measured at the stop level by:

$$SA = \frac{Ne + Nl}{N} * 100 \quad (3.1)$$

where Ne is the number of vehicles that arrived earlier than a certain Earliness Threshold, Te ; Nl is the number of vehicles that arrived later than a certain Lateness Threshold, Tl , and N is the total number of vehicles

Similarly, the same formula can be applied to the route and even network level by simply considering Ne , Nl and N to refer to all arrivals at the stops that comprise the route or network.

It is important to note how these thresholds vary according to the network in question: population density, quality of infrastructure and even company policies make it so that different transportation domain users might have different values for what should be deemed as "late" and "early".

While understanding how often vehicles arrive at a stop beyond earliness and lateness thresholds, the presence of values beyond the defined thresholds can be utilised as a measure in itself since schedule adherence can be generally high but sporadically poor. If the day is divided into a finite yet reasonably high number of consecutive number of parts, the recurring presence of a value beyond the thresholds at one part of the day is indicative of a problem in the network.

From the passenger's perspective, a high-frequency route's schedule is rather insignificant as they know that, arriving at the stop at any given time, they will not have to wait long for another vehicle. Thus, this indicator is more suited for low-frequency routes, where an out of schedule vehicle could mean a delay getting to the destination or a missed transfer to another route.

3.5.0.2 Headway Regularity

As previously discussed, Headway Regularity is important in detecting vehicle-bunching and large gaps between vehicles. As such, visualizing this KPI should not only mean giving a meaningful value to the transportation domain user but also allowing for the visualization of the location of vehicles in relation to each other by means of, for example, a space-time diagram. The use of a regularity index would allow for the immediate identification of irregular routes, directing the user's to a more detailed analysis of the situation which would be made with the help of a graphical visualization, but the usage of a regularity index could lead to misleading conclusions since it might aggregate the data undesirably due to the fast changing values of headway between vehicles in a route.

Headway regularity is important for avoiding under and overcrowding of vehicles but it is also important for maintaining the desired route frequency. As the vehicles in high-frequency routes do not follow a precise schedule so long as they maintain their announced frequency, this indicator is more suited towards those types of route.

3.5.0.3 Usage of Other Indicators

The developed framework does not encompass Travel Time, Wait Time or Buffer Time indicators due to the focus on creating meaningful visualizations for what were deemed as the most important indicators. Nevertheless, buffer time should be included in a prototype that is not meant to be a proof of concept of visualizations for service reliability and, as such, their inclusion will be discussed in this report when addressing the future work that could be done to improve upon this framework's ideas and guidelines. Travel Time and Wait Time play a much less significant role, as the visualization of Buffer Time and Schedule Adherence should be more than enough to infer the same conclusions from the data.

3.6 Summary

The methodology for this dissertation work mainly consisted on following the Interaction Design Principles, supported by solid research on the subject of service reliability in public transportation, to create a framework that provides transportation domain users with a tool to analyse issues, tendencies and patterns in public transportation networks. The design process followed a RAD methodology and was guided by an early evaluation of a non-functional prototype, important milestones and usability tests for further evaluation. Before the development process could begin, a definition for service reliability of public transportation in the context of this dissertation was given which bases it on Schedule Adherence and Headway Regularity. The use of those metrics guided the creation of the TransViz visualizations.

Chapter 4

TransViz

4.1 Overview

This chapter describes the creation of a visualization framework for service reliability in public transportation which was entitled TransViz. Such a framework follows the established definition of service reliability in the context of this dissertation. This chapter starts by going through the initial requirements and functionalities of TransViz goes on to describe the entire design process and the results of the various evaluations to arrive at the current state of the framework.

4.2 Objectives

Being a visualization framework, TransViz has two main objectives:

- The creation of visualizations that suit the needs of transportation domain users;
- The creation of a set of guidelines for the development of applications that suit the needs of transportation domain users.

For the fulfilment of the first objective, a prototype was created using C# in Windows Forms¹ and using the VTK - Visualization Toolkit². The prototype was comprised of four visualizations which data navigation and selection and relayed on the data set created using data from the MBTA v3 API.

The second objective was pursued in two ways. Firstly, by documenting the interesting ideas that could not be implemented either due to data restrictions or time limitations. Secondly, by collecting feedback on the usability of the prototype and on the needs of transportation domain users to understand how small things like colour contribute to a more complete and useful tool.

¹<https://docs.microsoft.com/en-us/dotnet/framework/winforms/>

²<https://vtk.org/>

4.3 Initial Requirements

When talking about requirements it is important to distinguish between functional requirements, the ones that specify what the system should do, and non-functional requirements, those which have to do with how the user interacts with the system and how it works.

TransViz, being a visualization framework, has a particularly high emphasis on the user experience and on the usability of the data navigation, comparison and visualization functionalities. Nevertheless, the system that supports that user experience should be robust and versatile enough to allow TransViz to be a generalizable framework into the majority of public transportation networks.

During the design process, the current state of TransViz was evaluated several times and from the data gathered from those evaluations, the requirements were revisited in order to assess how relevant they still were and how close TransViz is to achieving them.

The first set of requirements was created and evaluated in a focus group setting via a non-functional prototype developed with Figma³.

Functional Requirements

- FR01 Select/Deselect one or more routes/lines/stops by clicking them on the map;
- FR02 Select/Deselect one or more routes/lines/stops by clicking them on a list;
- FR03 Display basic information regarding the selected routes/lines/stops;
- FR04 Search a section of the network by name, id or area;
- FR05 Adjusting the thresholds for Extreme-Value Based indicators;
- FR06 Select the start and end points of the data in a calendar;
- FR07 Change the granularity of the calendar view;
- FR08 Change the type(s) of data being visualized;
- FR09 Direct the users towards situations that might require their attention, via an alert;
- FR10 Display stop accessibility for the selected section of the network with a click.

Non-Functional requirements

- NFR01 Scalability - The system should support increasing amounts of data typical of a full public transportation network;
- NFR02 Performance - Virtually all users should be able to complete most tasks;
- NFR03 User-friendly - TransViz should be easy to use as a hard time navigating the vast amounts of data would be a major hit into the user experience;

³<https://www.figma.com>

- NFR04 Information Scent - Information should be where the user expects it to be. Users should not waste more than a second or two looking for the information they want;
- NFR05 Flow - Users should intuitively follow a flow for their work which would be something akin to selecting the section(s) they want, selecting the time frame and analysing the data;
- NFR06 Productivity - The user should feel that using TransViz increases their work productivity;
- NFR07 Accessibility - The system should take into account the special needs of some users and implement features to address them, such as colour schemes for colourblindness;
- NFR08 Quality and Reliability - The system should provide a reliable experience for the user during extended periods of use.

4.3.1 Other Functionalities

Non-key functionalities are not integral to the system or its goals but enhance it in ways that often shape user perception. These other functionalities can be very minor, such as allowing for the change of the colour scheme, but summed up they have an important contribution to the refinement of the non-functional requirements and the user experience.

- OF01 Visually represent on the map segments of the network that have exclusive bus lanes;
- OF02 Index creation menu where the user can create a custom measure to be displayed;
- OF03 Filters for the search results;
- OF04 Play feature which makes historical data be displayed in a sequence of snapshots;
- OF05 Use scrolling to change the granularity of the calendar.
- OF06 Preserve the state of the system when the user closes it.

4.4 Non-Functional Prototype

The non-functional, low-fidelity prototype was built using Figma, an online tool for prototyping. This prototype focused heavily on a customizable experience for the user, providing the ability to create new visualizations based on tables or graphs. It also aimed to have various data selection capabilities.

The prototype's main page featured an interactable map for the user to view the state of the network and select which routes and stops they wanted to analyse. Such a selection could also be made from a list (which would be inside the box on the right). Both the list and the map would be updated to highlight any search made at the top (Figure 4.1). There were also buttons for saving the current state of the program, load the state from file and advancing to the next screen.

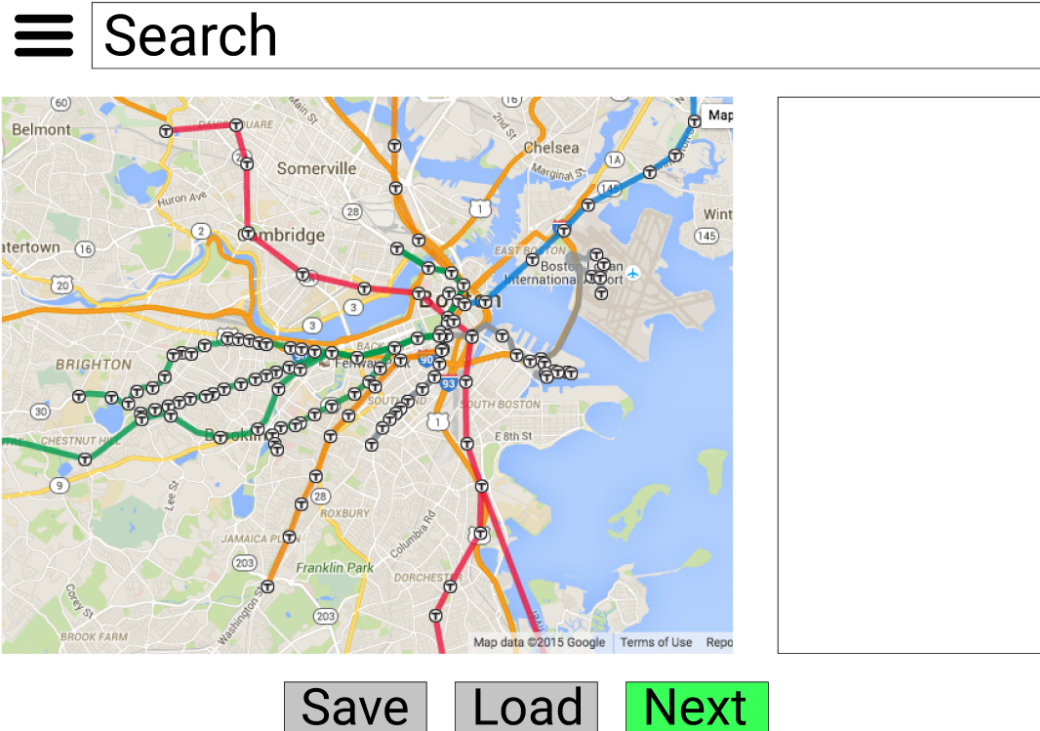


Figure 4.1: The main page of the non-functional prototype. Notice the hamburger button on the top left

Lastly, the main page contained a drawer menu which could be enabled by pressing the hamburger button to push the contents of the page to the right, hiding the list. This menu was to contain a number of filters which the user could use to refine their search (Figure 4.2). Some ideas for filters included: route length; stop distance to the terminal; the existence of bus-exclusive lanes; among others.

After making a selection, the user could advance to the next page which was responsible for the visualization of the data regarding the selected segments of the network (Figure 4.3). On that screen the user could view and create visualizations based on tables or graphs, selecting which axis or columns and lines they preferred. This page also allowed for a calendar visualization much similar to the MetroViz calendar (see Figure 2.5) with the added feature of zooming in or out to change its granularity. The user could replace the calendar view with others that would give additional information on each of the lines and stops selected, line distance to terminal or scheduled headway.

The user could press the "+" button to create a new visualization (Figure 4.4).

4.4.1 Focus-group evaluation

This non-functional prototype was evaluated in a focus-group setting with researchers in the various areas of public transportation so that feedback could be collected early on the design process.

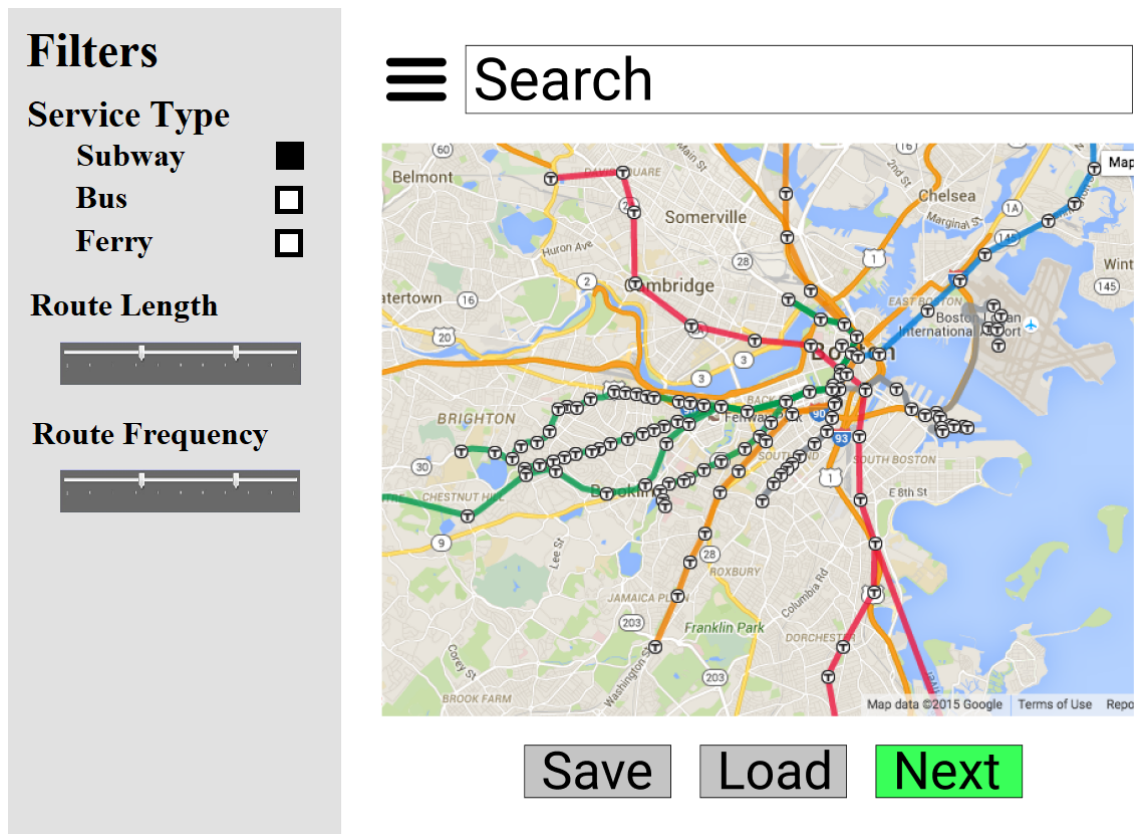


Figure 4.2: The main page of the non-functional prototype with the expanded drawer menu

The objectives of the prototype, and the dissertation work, were explained and contextualized and then the prototype was presented in a guided tour with every comment on it being recorded for analysis later on.

The most notable criticism of the prototype was its large scope: it aimed to do too much which provoked a sense that nothing was very useful. The transition between pages was also criticized as being unintuitive: the fact that one screen was dependent on the other was not clear and the individual purpose of each screen was shadowed in complexity.

The calendar on the second screen received positive feedback because it allowed for simple visualization of the temporal distribution of problems on the network. Other information on the routes seemed to be presented in a way that was not very useful.

The creation of other visualizations seemed to be a solution looking for a problem, with the feedback from the group pointing to the creation of a few very meaningful visualizations instead of a broad brush of graphs and tables.

Overall, the current state of the framework seemed to tailor more to someone who was interested in viewing data on the network rather than a transportation domain user.

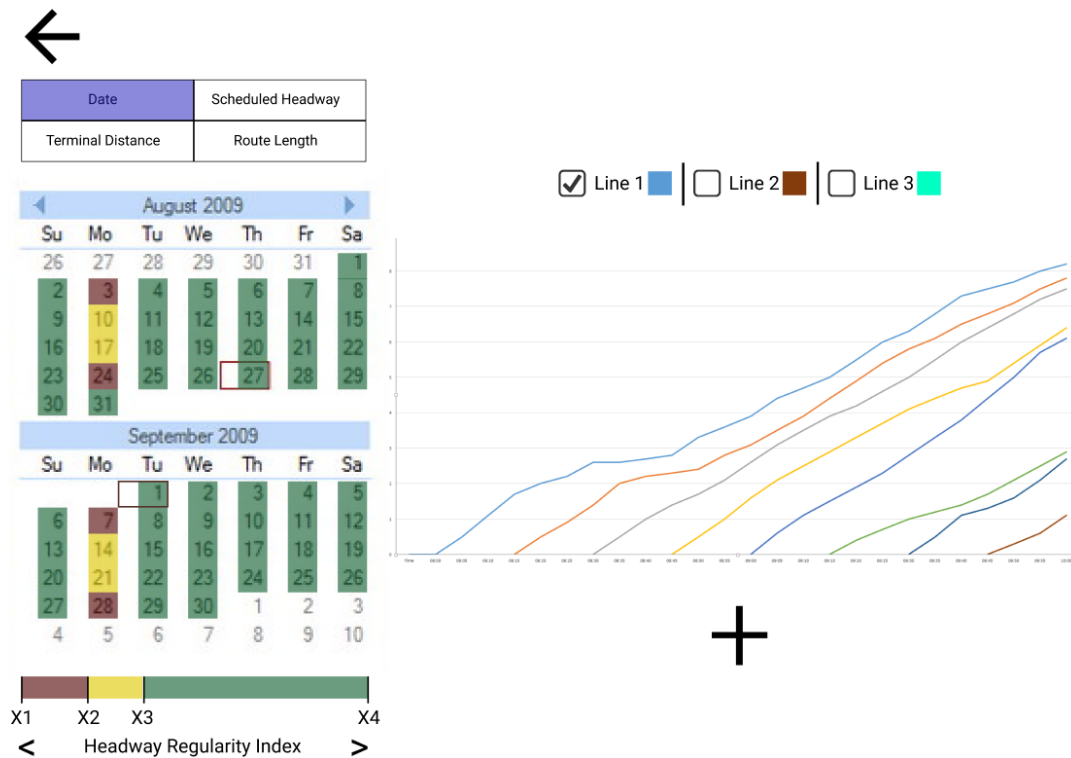


Figure 4.3: The data visualization page of the non-functional prototype

4.5 Requirements Revision

From the feedback collected on the non-functional prototype in the focus group setting, the initial set of requirements was trimmed down for the dissertation work and the scope narrowed to encompass only the creation of a few visualizations that would give public transportation domain users the ability to identify and analyse problems and patterns within a network. Those visualizations would have to be tailored to represent the KPIs that were selected in a suitable and intuitive way.

The following visualizations were pondered:

- VIZ01 A stacked columns chart for visualizing Schedule Adherence via the percentages of early, late and on-time arrivals for a line/stop based on changeable earliness and lateness thresholds.
- VIZ02 A 24-hour circular clock with the work days individually represented for visualizing Schedule Adherence at each part of the day via the presence of values beyond the defined earliness and lateness thresholds.
- VIZ03 A representation of the distance between vehicles in a line at any given moment.
- VIZ04 A 3D map view of the Schedule Adherence at each stop, calculated by the presence of values beyond the earliness and lateness thresholds.



Figure 4.4: The data visualization page of the non-functional prototype with the "new visualization" overlay

VIZ05 A space-time diagram with the location of vehicles.

VIZ06 A colour-based calendar view that conveys the days where there are problems on the network.

Each of these visualization proposals had its requirements refined and their feasibility and utility evaluated. However, as a whole, the TransViz prototype would have to follow some requirements as well:

FR01 Navigate between the various visualizations;

FR02 Select the time frame for the data;

FR03 View the selected time/date.

UR01 Provide options for colourblind users;

UR02 Provide hints to the functioning of the prototype;

UR03 Achieve each requirement with as few actions as possible (preferably fewer than 3).

Each of these requirements was achieved in a simple way: navigation is made through tabs, each tab corresponding to a visualization; both selecting the time frame and viewing the currently selected time/date are made through a calendar which, for some visualizations, is complemented by additional controls; colourblind users can select their preferred colour palette, even though all three of the colour palettes implemented are suited for most or all types of colourblindness; hints are provided via a tooltip on hover over virtually every element; limiting the amount of actions needed to perform a task is done by implementing simple yet robust controls.

Visualizations 1-4 were implemented and evaluated in usability tests with public transportation operators and researchers. Visualizations 5 and 6 were not implemented but will nonetheless be discussed as they could play an important role in a framework such as this one.

4.5.1 VIZ01 - Stacked Columns Chart

This visualization (Figure 4.5) is aimed towards the analysis of Schedule Adherence at each line and each stop. The user can select the start date on a calendar and the number of days that should be included in the calculation. The user can also change the earliness and thresholds that are used to calculate the percentage of early and late arrivals and see the changes happen instantly. By default, the user is presented with the data for each of the lines in the network.

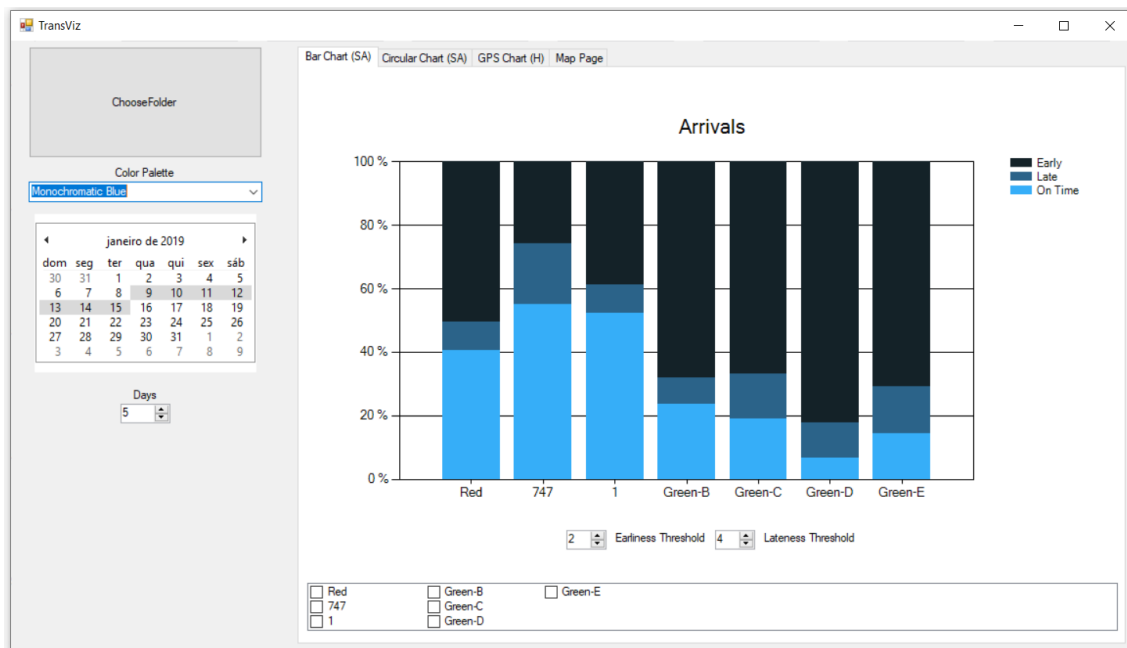


Figure 4.5: The Stacked Columns Chart visualization in "lines" mode

When the user selects a line from the options at the bottom of the screen, the chart columns will be replaced by the columns corresponding to the arrivals at each of the stops of the selected line, displayed in geographical order (Figure 4.6). The user still retains the same control over the data selection and the thresholds as they do when no lines were selected.

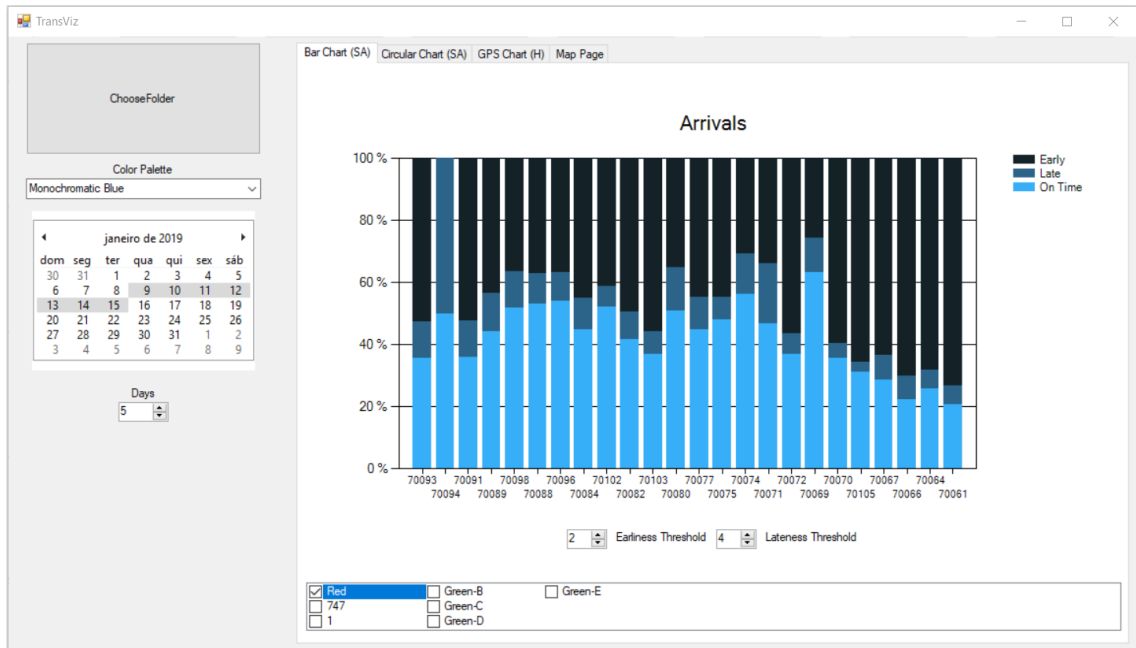


Figure 4.6: The Stacked Columns Chart visualization in "stops" mode

Early versions of the prototype included a slider instead of a numeric up-down for selecting the thresholds. However, the slider limits were harder to define while maintaining a good scale and were opted out.

Since this is a visualization for Schedule Adherence, it is more suited towards low-frequency routes. This visualization allows the user to identify problematic lines and stops in the network. Allow the information provided by this visualization can be constrained by the low granularity of the data presented, it is an important first step in understanding what problems are occurring and where and this visualization is best used when in conjunction with the one presented next which provides a higher granularity for visualizing schedule adherence.

4.5.1.1 Evaluation

The feedback for this visualization was positive. It was stated that it was simple to read and understand. The use of the variable thresholds was remarked as a positive feature as the needs of each network and line could differ and the definition of "late" and "early" could vary from city to city, line to line and even company to company.

The fact that all stops inside a line are given equal highlight was noted as a potential flaw since there are stops that, for the public transportation operators and the decision-making process are more important. These stops are the start and end terminals and checkpoints (or time-control points) which are defined by the transportation domain users taking into account the infrastructure of the stop/station. These are the points in the network where there is a systematic collection of the vehicle data (is it late?; is it overcrowded?) and it is here that vehicles can be held in order to deal with problems that are occurring in real time, like vehicle-bunching. It was noted

that this framework has importance in evaluating the network at all stops but it is critical at these checkpoints and emphasis should be given to those stops whenever possible.

A suggestion was made to explore the idea of using this visualization in a map of the selected line: instead of providing the values for each stop, the values for each segment between the stops would be shown. This suggestion was not explored as the benefits of such a visualization would be minimal in comparison to the current one and time constraints deemed such small improvements as unfeasible.

Lastly, some suggestions were made regarding the white background colour which could be changed to something more neutral like light grey and the use of a more self-explanatory title than "Arrivals", which is something that should also be taken into account in the other visualizations.

4.5.2 VIZ02 - 24 Hour Clock

This visualization (Figure 4.7) is aimed towards the analysis of patterns in Schedule Adherence at each line. Each disk represents a day of the week, going from Monday, the inner disk, to Friday, the outer disk. The darker colours identify times of day where there are early or late arrivals for the selected line, based on the selected thresholds. For the purposes of this prototype, each day is divided into 250 equal parts, a number that was selected on an attempt to balance the high rendering time with the precision: more parts result in more precision but also on poorer rendering performance.

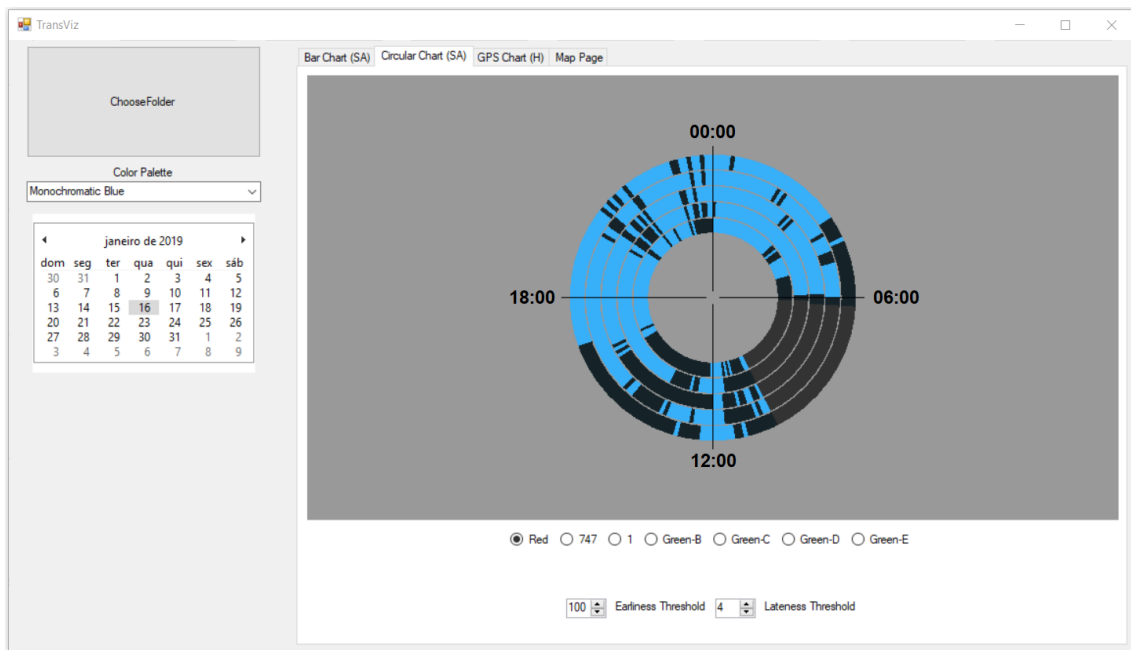


Figure 4.7: The 24 Hour Clock visualization

When the user selects a date from the calendar, the visualization is updated to start on the Monday of the selected week. The user can select the earliness and lateness thresholds but experience showed that, when both are set too low, the amount of issues that the visualization detects

overshadows any information that the user could withdraw otherwise. Lastly, the user can navigate the 2D space of the visualization using the mouse.

As before, this is a visualization for Schedule Adherence and so it is more suited towards low-frequency routes. This visualization allows the user to identify parts of the day where problems occur. The concentric disks allow for faster analysis of the entire work week and, thus, for easier identification of patterns that occur throughout the week. As previously mentioned, a user could identify a poor Schedule Adherence in the first visualization and move to this one to further analyse the reasons for that poor performance.

Ideally, this visualization would allow for a "Play" feature which would illustrate one week after the other in order to understand if the patterns identified in one week persist during the month or even the year. However, since the data set created only contained data for two weeks and the render time for this visualization is in the order of the few seconds, the implementation of such a feature was not included in the work plan.

4.5.2.1 Evaluation

This visualization was well received and it was noted that it is useful in detecting problematic sections of the day and patterns that occur during the week.

A comment was made regarding the alignment of the data in that the beginning of the data for each day (for lines without a 24-hour schedule) was particularly aligned with any line of the clock. As such, it was suggested that, for each day, the clock was not aligned by the hours of the day but by the start of the schedule for that line. Such a suggestion merited discussion but ultimately was not implemented as there was not a significant enough reason to do so.

Another suggestion was made to divide the disks by sectors - day, evening and night. This would make it easier to locate problems specific to each time of the day.

4.5.3 VIZ03 - Vehicle Location Chart

This visualization (Figure 4.8) presents the user with a chart that displays the location of each vehicle in a line, for both directions, at the selected time and date. The left-most and right-most vertical lines display the start and end station for the line. The other vertical lines each represent a vehicle. The colour of the segments between vehicles is an indicator of the regularity of the spacing between them: vehicles that are too close to each other or too far apart from one another are a problem. The positive and negative height of the segments indicate whether the headway of the vehicles is too large or too small: a positive delta-y refers to a spacing that is too large; a negative delta-y to a spacing that is too small.

The user can use the controls at the bottom to advance or go back in the selected time: each frame is a 2 minutes interval. The "Play" button allows makes the simulation advance automatically, with the time it takes to advance to the next frame being adjustable by the slider. Lastly, the user can navigate the 2D space of the visualization using the mouse.



Figure 4.8: Vehicle Location Chart

This visualization pertains to Headway Regularity and, as such, it is most suited towards high-frequency routes. Using this visualization, a user can observe the location of vehicles at a given moment and observe the persistence of headway irregularities. This allows for the identification of problematic times of day/week/year and sections of the line where issues occur most frequently.

4.5.3.1 Evaluation

This visualization received positive feedback for its ability to analyse vehicle-bunching problems and large gaps between vehicles which is an ever increasing problem as passengers continue to migrate to cities and the number of frequency of routes continues to increase.

A suggestion was made to transpose this chart into a map which would allow for a better analysis of the actual location of the vehicles and stops. To the same effect, another option would be to add the position of a few key stops on the visualizations, so that the user could interpret where the vehicles were located at any given time in relation to these stops and not only to the start and end terminal.

The choice of using the same colour for illustrating both large and small gaps was criticized as the use of different colours for the different extreme values would make it more immediate for the user to identify vehicle-bunching situations.

4.5.4 VIZ04 - Map

This visualization (Figure 4.9) was an experiment on the utility of having a 3D visualization when dealing with data regarding public transportation. It presents the user with a map of Boston with

bars on top of each stop. The height and colour of the bars represent problems in Schedule Adherence for the selected day and hour based on the presence of percentage of early and late arrivals, but it could also be modified to illustrate a regularity index or other measures.

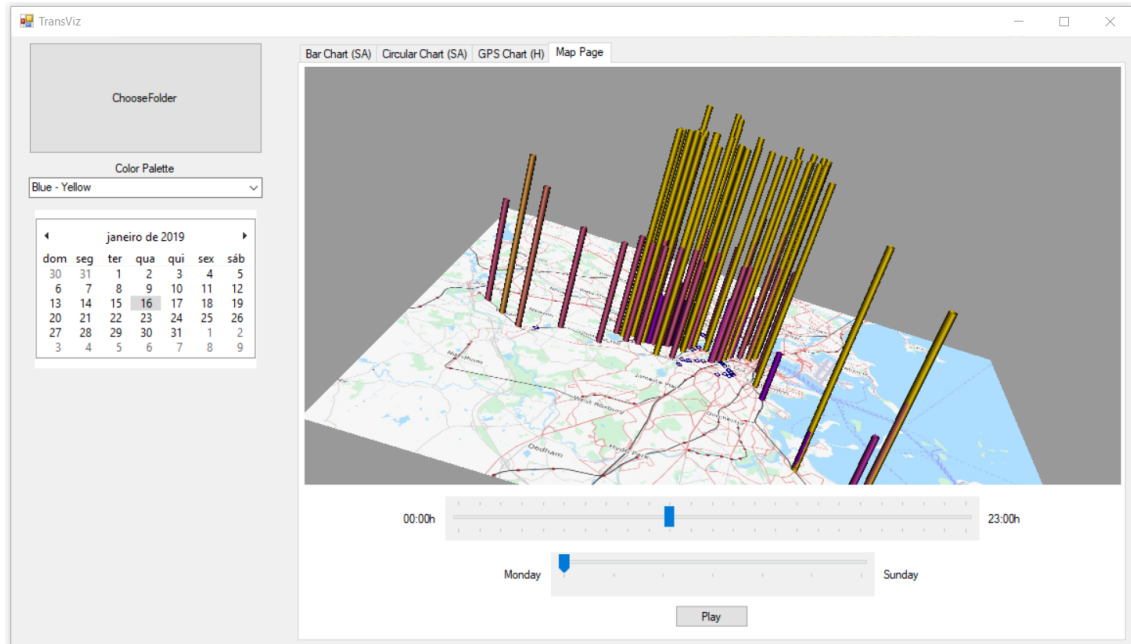


Figure 4.9: Map Visualization

The user can use the "Play" feature to automatically advance the time of the data by intervals of 1 hour or select the desired time and day but moving the sliders. Lastly, the user can navigate the 3D space of the map using the mouse.

4.5.4.1 Evaluation

This visualization received negative feedback due to its low degree of utility.

While the play feature could be used to identify the times of day where problems are more persistent, this visualization is overshadowed by simpler, easier to interpret ones which was something that was promptly mentioned in every usability test conducted.

The feedback for this visualization was unanimous in pointing that this is more of visualization for an enthusiast than for a decision maker since it has very little capabilities for extracting useful information.

Suggestions were made to remake this visualization on a 2D scheme of the network with each sector having a colour that represented its Schedule Adherence value. Such a suggestion was not implemented due to time constraints but it would complement the other schedule adherence based visualizations in identifying issues in specific parts of the network.

4.5.5 VIZ05 - Space-time Diagram

Space-time diagrams (Figure 4.10) are a widely used graph type when analysing the service reliability of public transportation networks for visualizing the headway of vehicles and identifying vehicle-bunching problems and trip irregularities. Such a diagram displays the position of each vehicle (in relation to the start terminal) over time and headway irregularity problems can be easily identified by analysing the distance between the lines that pertain to each vehicle.

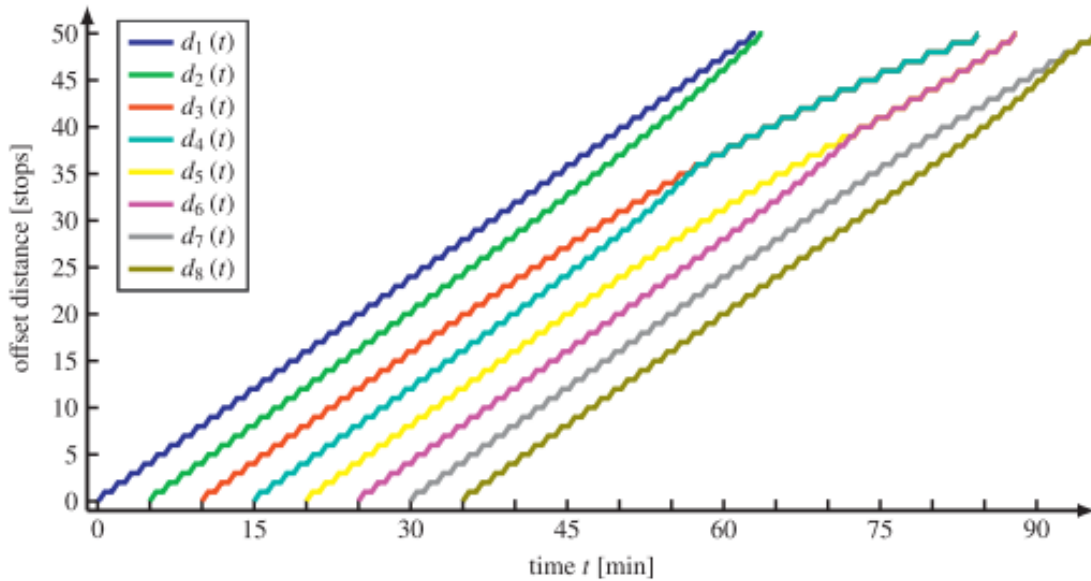


Figure 4.10: Example space-time diagram [5]

This visualization was not implemented in the TransViz prototype due to its already widespread use. It could, however, be used as a precursor to Vehicle Location Chart visualization: the user would use a space-time diagram to identify issues in a day and then use the Vehicle Location Chart to more thoroughly analyse when the issues begin and how they propagate.

4.5.6 VIZ06 - Colour Calendar

This visualization (Figure 4.11) presents the user with a calendar where periods are coloured based on the values of a KPI for the period. The user would be able to select the limits of the gradient that is used to colour the calendar and change the granularity of the calendar by scrolling up and down, moving from a calendar that displays all days in a month or two to a calendar that displays only the weeks or the months of the year.

This visualization was not implemented in the TransViz prototype due to time limitations and complexity of implementation in a Windows Forms application (the native calendar control allows neither the colouring of days or scrolling to change granularity. Nevertheless, the implementation of such a visualization can bring the user the ability to be easily alerted to problems in a network

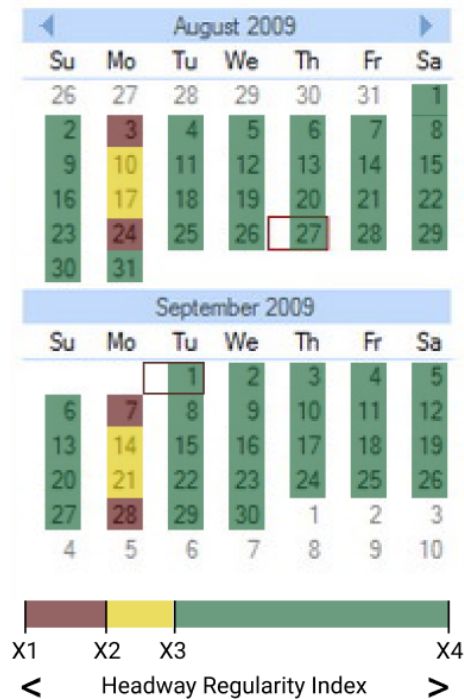


Figure 4.11: Colour Calendar Visualization

and to patterns that are only visible with a high degree of abstraction. For example, this visualization could alert the user to problems that only occur on Fridays, something that would be harder to visualize on the other frameworks which deal with smaller time frames.

4.6 General Evaluation

Besides feedback on each particular visualization, some comments were made regarding the framework as a whole.

The first comment was regarding the use of colourblind adequate palettes which could be replaced by normal Green-Yellow-Red palettes which are more universal and, to address the needs of colourblind users, adopt a Colour Identification System like ColourADD⁴.

Another criticism was made regarding the case study and the data set which lacked the presence of a low-frequency route. While this does not affect the quality of the visualizations, the existence of at least such a route in the prototype could have been interesting in better assessing the potential of the prototype. Also regarding the data set, it was noted that the 1 minute refresh rate of the data might be insufficient in a real world scenario although it was agreed that for testing and proof of concept purposes it was adequate.

Questions were made on the data set used, namely how were arrivals at stops recorded and how was the position of each vehicle in a line determined, which were followed by a description of the shortcomings of the case study and the data set, such as the inability to record the times of

⁴<http://www.colouradd.net/code.asp>

arrivals at each stop and the need to use the "Predicted Arrivals" to get a somewhat accurate arrival time, as well as the assumption that lines were somewhat a straight line in order to position the vehicles correctly on VIZ03 with their geographical coordinates.

Overall, the evaluation of the framework was very positive: users agreed on the importance of a tool like the TransViz prototype to the decision-making process of a public transportation operator, with visualizations addressing the needs of both low-frequency and high-frequency routes. Some discussion on future work also took place with users suggesting the implementation of more visualizations: the use of simple network maps with coloured segments to represent schedule adherence values; the implementation of space-time diagrams into the prototype; and a graph to display the absolute the difference in time between schedule arrivals and actual arrivals for each vehicle in a line.

One of the tests was followed by a discussion of the increasing importance of visualization tools and: with the continued migration from rural locations to urban centres, not only does the complexity of public transportation networks increase, and with it the difficulty in maintaining schedules and frequencies, but also the importance of maintaining good service reliability in very low-frequency routes to assure that people who live on the outskirts of cities can still depend on the public transportation network for their mobility. With this in mind, the importance of the TransViz framework was highlighted since it provides valuable insight on how to create tools that will be even more necessary in the future than they already are today.

4.7 Proposed Guidelines

Besides the creation of a functional prototype, this dissertation work helped aggregate and consolidate information and ideas on how visualization tools for service reliability of public transportation should be built. As such, a list was compiled with guidelines that should be taken into account when developing tools to visualize service reliability.

- Understand the different KPIs associated with service reliability and the different ways they can be measured (extreme-value based, mean-variance, real depictions of the state of the networks, among others);
- Create both schedule adherence and headway regularity visualizations so that both low and high-frequency routes can be analysed with a measure that suits them;
- Provide visualizations that either allow for a change in the granularity of the data (from the stop level all the way to the network level) or that are complemented by visualizations with different levels of abstraction;
- Make use of colour to highlight the most important information, such as problems in the network, and immediately direct the user towards it;

- Provide controls so that the user can adjust the tool to fit the needs of the network they are analysing such as adjustable thresholds or different types of calculation for the selected KPIs;
- Create simple and powerful data navigation functionalities that allow the user to select the time frame of the data as well as the sections of the network he intends to visualize;
- As much as possible, either insert the visualizations onto a map where they can be better contextualized or provide references, such as the location of stops on a line representation;
- Don't focus on 3D visualizations as they are rarely more useful than 2D ones and often just become too dense and confusing;
- Do not reduce the KPIs only to numbers: visualizing service reliability measures can often be done without much abstraction, such as representing the position of vehicles in a map;
- Mind the principles of HCI and interaction design for visualization when developing not only the visualizations themselves but also the tool in each they are inserted.

Following these guidelines and the rest of the evaluation results can be extremely valuable when creating visualization tools for service reliability of public transportation

4.8 Summary

The design process of the TransViz framework began with the creation of the first set of requirements and features that aimed to satisfy the needs of public transportation domain users. Those requirements were evaluated via a non-functional prototype in a focus group setting and were readjusted in nature and scope. TransViz's main focus became the creation of robust and meaningful visualizations of service reliability of public transportation with four visualizations being created and evaluated and two more proposed. General feedback on the TransViz prototype was positive with only one visualization receiving a generally negative evaluation. The entirety of the design process, the research made into the topic and the feedback received from the evaluation of the TransViz prototype allowed for the gathering of insights and the creation of a set of guidelines to follow when creating visualization tools for service reliability of public transportation.

Chapter 5

Conclusions

5.1 Overview

This chapter describes the conclusions arrived at from the research and work. It also presents an overview of the work done and the potential applications of this dissertation work in the academic and public transportation sectors. Finally, it describes the satisfaction regarding the objectives for this dissertation and future work that could be done on top of the current iteration of the TransViz framework.

5.2 General Conclusions

Service reliability in public transportation is not only complex to define but also difficult to visualize. AVL and APC data have previously been used extensively in research into this subject. A large part of that research is geared towards schedule optimization and automatic route planning but there are also extensive studies that analyse the use of data to create indexes and measures of how reliably a network is performing. Service reliability has been measured by means of a number of Key Performance Indicators (KPIs), such as Schedule Adherence, Headway Regularity, Travel and Wait Time, Buffer Time, Ridership and Stop Accessibility.

Putting a value on a service reliability measure is often insufficient or inadequate for ascertain the state of the network which makes it important to create visualizations that represent things more in tandem with the physical world rather than mathematical abstractions. Not many materialisations of the visualization of service reliability have been made in academia under the form of an actual usable tool for transportation domain users. Some frameworks propose studying service reliability by means of certain KPIs but require the intensive study into each and every one of the routes and stops in a network every time any sort of meaningful conclusion wants to be extracted.

The dissertation aimed to create TransViz, a visualization framework of service reliability of public transportation. TransViz's main purpose is to provide guidelines towards the creation of tools that would act as decision support agents by assisting public transportation domain users in the task of identifying problems and tendencies in the network without ordering the analysis of

millions of lines of data. TransViz focuses on KPI visualization as well as data navigation. The framework intends to be used as an identifier of potential problems and patterns that appear in public transportation networks.

Visualizations created for service reliability of public transportation have a lot of complexities to answer to, from the intricacies of the network to the variable degree of scrutiny that different sections of the network are under. Understanding the needs of public transportation domain users is vital since those needs are the ones that must be transposed into the requirements of a visualization tool.

5.3 Real World Applications

The most obvious application for this dissertation work is in public transportation operators which could build upon the developed framework to create a system that is capable of helping transportation domain users in planning schedules, routes, stops and making decisions regarding the network.

TransViz's proposed visualizations can be used to identify issues and tendencies in public transportation networks, thus supporting public transportation domain users in their task of evaluating service reliability without requiring an extensive and thorough analysis of the enormous data sets.

Another application would be related to future academic work, as frameworks such as TransViz could help identify and study patterns and tendencies in public transportation network, especially if they are expanded to encompass the registering and automatic detection of recurring patterns.

This dissertation work has led to the conclusion that TransViz could help bridge a gap in the evaluation and measuring of service reliability of public transportation due to the increased accessibility when using large data sets when compared to the most recurrent, thorough and very formal methods of analysing this topic which are resource expensive and time-consuming.

5.4 Objectives Satisfaction

The main objective of this dissertation work has been concluded: to study and create ways to visualize the service reliability of public transportation. This dissertation also produced a functional prototype of some of the proposed visualizations which allowed for their evaluation and produced additional guidelines and ideas on how to visualize service reliability.

The purpose of creating a framework for the visualization of service reliability of public transportation was achieved. Although further evaluation is warranted, especially in applying this framework in a real life use case such as the city of Porto, the TransViz framework has received positive feedback as a decision support tool, as was intended for it.

It can thus be said that the dissertation work satisfies the objectives laid out for it. Although future work can expand the TransViz framework into a complete tool, an important groundwork

Conclusions

has been done during this dissertation in the field of visualization of service reliability of public transportation.

5.5 Future Work

As previously mentioned, TransViz has much expansion potential which could transform it into a tool that is applicable to public transportation networks beginning with the implementation of the two proposed but not implemented visualizations for evaluation.

TransViz has room for many more visualizations. Heat maps of the distribution of vehicles in the network could be used to identify bunching problems, as well as issues with sections of the road which could lead to changes in schedules and route paths. VIZ03 - Vehicle Location Chart could be changed to a disk when the route is circular. Map schemes of the lines in a network could be used with each segment between stops changing colour to represent the state of either schedule adherence or headway regularity. Additional metrics could be added, such as passenger count, for the visualization of more metrics that define service reliability.

A first screen could be added that allowed for the selection of specific parts of the network, like the first screen proposed in the early non-functional prototype, which would be vital in decreasing the density of information and controls when dealing with more than a few lines of the network.

The presence of automatic global alerts would be a benefit as it would direct the transportation domain user to potential problems, reducing wasted time on finding those issues.

The application of some of the TransViz's visualizations, such as VIZ03 - Vehicle Location Chart, to a real-time stream of data, instead of relying on historical data, could prove beneficial to the monitoring of the current state of the network.

Conclusions

Bibliography

- [1] MBTA - Overview.
- [2] MBTA Maps.
- [3] MBTA Performance Dashboard.
- [4] Pekka Abrahamsson, Outi Salo, Jussi Ronkainen, and Juhani Warsta. Agile software development methods Review and analysis. Technical report, 2002.
- [5] Matthias Andres and Rahul Nair. A predictive-control framework to address bus bunching. *Transportation Research Part B: Methodological*, 104:123–148, 10 2017.
- [6] Benedetto Barabino, Massimo Di Francesco, and Sara Mozzoni. Rethinking bus punctuality by integrating Automatic Vehicle Location data and passenger patterns. *Transportation Research Part A: Policy and Practice*, 75:84–95, 5 2015.
- [7] John W. Bates. *Definition of Practices for Bus Transit On-time Performance: Preliminary Study*. Transportation Research Board, National Research Council, 1986.
- [8] Giuseppe Bellei and Konstantinos Gkoumas. Transit vehicles’ headway distribution and service irregularity. *Public Transport*, 2(4):269–289, 11 2010.
- [9] Mathew Berkow, Ahmed M. El-Geneidy, Robert L. Bertini, and David Crout. Beyond Generating Transit Performance Measures. *Transportation Research Record: Journal of the Transportation Research Board*, 2111(1):158–168, 1 2009.
- [10] Paul Beynon-Davies and Hugh Mackay. Rapid application development (RAD): An empirical review. *Article in European Journal of Information Systems*, 1999.
- [11] Stuart K. Card, Jock D. Mackinlay, and Ben. Shneiderman. *Readings in information visualization : using vision to think*. Morgan Kaufmann Publishers, 1999.
- [12] Avishai. Ceder and Avishai. *Public transit planning and operation : theory, modelling and practice*. Elsevier, 2007.
- [13] Xumei Chen, Lei Yu, Yushi Zhang, and Jifu Guo. Analyzing urban bus service reliability at the stop, route, and network levels. *Transportation Research Part A: Policy and Practice*, 43(8):722–734, 10 2009.

BIBLIOGRAPHY

- [14] Ken Collier. *Agile analytics : a value-driven approach to business intelligence and data warehousing*. [Addison-Wesley], 2012.
- [15] Karina Curcio, Rodolfo Santana, Sheila Reinehr, and Andreia Malucelli. Usability in agile software development: A tertiary study. *Computer Standards & Interfaces*, 64:61–77, 5 2019.
- [16] Fan Du, Joshua Brulé, Peter Enns, Varun Manjunatha, and Yoav Segev. MetroViz: Visual Analysis of Public Transportation Data. 2015.
- [17] Wei (David) Fan and Randy B. Machemehl. Do Transit Users Just Wait for Buses or Wait with Strategies? *Transportation Research Record: Journal of the Transportation Research Board*, 2111(1):169–176, 1 2009.
- [18] Wei Feng and Miguel Figliozi. Developing a bus service reliability evaluation and visualization framework using archived AVL / APC data. pages 1–14, 2006.
- [19] Meliha Handzic. Knowledge Management Selection Model for Project Management. pages 157–179. 2017.
- [20] Rex Hartson, Pardha Pyla, Rex Hartson, and Pardha Pyla. Agile Lifecycle Processes and the Funnel Model of Agile UX. *The UX Book*, pages 63–80, 1 2019.
- [21] Erik Jenelius. Public transport experienced service reliability: Integrating travel time and travel conditions. *Transportation Research Part A: Policy and Practice*, 117(August):275–291, 2018.
- [22] Ioannis Kaparias, Michael G.H. Bell, and Heidrun Belzner. A New Measure of Travel Time Reliability for In-Vehicle Navigation Systems. *Journal of Intelligent Transportation Systems*, 12(4):202–211, 11 2008.
- [23] Junlong Li, Xuhong Li, Dawei Chen, and Lucy Godding. Assessment of metro ridership fluctuation caused by weather conditions in Asian context: Using archived weather and ridership data in Nanjing. *Journal of Transport Geography*, 66:356–368, 1 2018.
- [24] Zhenliang Ma, Luis Ferreira, and Mahmoud Mesbah. A Framework for the Development of Bus Service Reliability Measures. *Australasian Transport Research Forum 2013 Proceedings*, (October):1–15, 2013.
- [25] G F Newell and R B Potts. Maintaining a bus schedule. 2(1), 1964.
- [26] Chris Nodder and Jakob Nielsen. *Agile Development that Incorporates User Experience Practices*. Nielsen Norman Group, 2013.
- [27] David L Parnas and Paul C Clements. A Rational Design Process: How And Why To Fake It. Technical report.

BIBLIOGRAPHY

- [28] Jenny Preece, Yvonne Rogers, and Helen Sharp. *Interaction design : beyond human-computer interaction*.
- [29] Cristina Pronello and Cristian Camusso. A Review of Transport Noise Indicators. *Transport Reviews*, 32(5):599–628, 9 2012.
- [30] Gayane Sedrakyan, Erik Mannens, and Katrien Verbert. Guiding the choice of learning dashboard visualizations: Linking dashboard design and data visualization concepts. *Journal of Computer Languages*, 50:19–38, 2 2019.
- [31] Thomas T. SIGCHI (Group : U.S.). Curriculum Development Group., Ronald Baecker, Stuart Card, Tom Carey, Jean Gasen, Marilyn Mantei, Gary Perlman, Gary Strong, and William Verplank. *ACM SIGCHI curricula for human-computer interaction*. Association for Computing Machinery, 1992.
- [32] Fulvio Silvestri. Estimating and visualizing perceived accessibility to transportation and urban facilities. *Transportation Research Procedia*, 31:136–145, 1 2018.
- [33] Eli Steven, D Tripp, Barbara Bichelmeyer, and Steven D Tripp. Rapid Prototyping: An Alternative Instructional Design Strategy. Technical report, 2006.
- [34] David L. Uniman, John Attanucci, Rabi G. Mishalani, and Nigel H. M. Wilson. Service Reliability Measurement Using Automated Fare Card Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2143(1):92–99, 1 2010.
- [35] Niels van Oort and Rob van Nes. Regularity analysis for optimizing urban transit network design. *Public Transport*, 1(2):155–168, 6 2009.

BIBLIOGRAPHY

Appendix A

Usability Tests for feedback collection regarding the TransViz prototype

A.0.1 Introduction

Bom dia/Boa tarde a todos. O meu nome é Tiago Grosso, sou aluno do mestrado do MIEIC e estou a fazer a dissertação em Visualization of Service Reliability of Public Transportation. Como parte do trabalho de dissertação, estou a desenvolver uma framework para visualizar a fiabilidade do serviço de uma rede de transportes públicos, à qual dei o nome de TransViz. Como tal, esta sessão de Testes de Usabilidade tem como objetivos assenta sobre um protótipo da framework e tem como objetivos:

1. Recolha de feedback relativamente às visualizações apresentadas no protótipo;
2. Recolha de feedback relativamente ao design de interação do protótipo;
3. Avaliação da capacidade do protótipo em responder às necessidades que se propõe satisfazer.

Antes de prosseguir com os testes, é importante contextualizar o problema e definir o seu escopo. A fiabilidade de uma rede de transportes públicos tem impacto direto não só no quão eficiente o serviço é no que toca à alocação de recursos, como também na experiência dos passageiros da rede. Assim, a importância de analisar a fiabilidade de uma rede de transportes públicos é evidente. No entanto, esta é uma questão que engloba muitos fatores e que pode ser avaliada através de um elevado número de indicadores. No contexto desta dissertação, foram desenvolvidas visualizações que assentam essencialmente no cumprimento dos horários definidos e na uniformidade do espaçamento entre veículos. Foram desenvolvidas quatro visualizações que pretendem localizar problemas e padrões na rede e progressivamente focar no local e horário exato em que esses problemas ocorrem, de forma a que depois um operador da rede de transportes possa estudar as suas causas e possíveis soluções. Para que as visualizações fossem construídas, foram utilizados dados da rede de transportes públicos de Boston. Devido à extensão dessa rede e ao facto de apenas se tratar de um protótipo, o número de linhas foi reduzido para 7 e os dados foram limitados ao período de tempo entre 9 de Janeiro de 2019 e 24 de Janeiro de 2019. As linhas selecionadas

foram algumas das mais utilizadas na cidade: as que tem um nome correspondente a uma cor são linhas de metro e as restantes são linhas de autocarro.

A.1 Test 1 – Stacked Columns Chart

Esta visualização pretende mostrar até que ponto o horário está a ser cumprido nas várias linhas da rede de Boston. Considera-se que um veículo está atraso ou adiantado quando a sua chegada a uma estação ocorreu com um diferença de tempo superior a um threshold de atraso ou adianto, respetivamente.

Tasks

1. Determinar a percentagem de veículos da linha “747” que chegaram atrasados 5 ou mais minutos entre os dias 14 e 18 de Janeiro, inclusive;
2. Determinar qual a linha com maior percentagem de veículos que chegaram adiantados 3 ou mais minutos no dia 20 de Janeiro;
3. Determinar qual a paragem da linha “Red” com maior percentagem de veículos atrasados 6 ou mais minutos entre os dias 12 e 13 de Janeiro, inclusive.

Post-test questions

1. Quais as maiores dificuldades na seleção do intervalo de tempo dos dados?
2. De que forma a escolha dos thresholds poderia ser melhorada?
3. Que outro tipo de informação deveria estar presente nesta visualização?
4. Comentários gerais.

A.2 Test 2 – 24 Hour Clock

Esta visualização pretende ajudar a identificar padrões no cumprimento dos horários nas várias linhas. Nesta visualização, é apresentado um relógio de 24h e 5 áreas circulares, cada uma correspondente a um dia (de fora para dentro: segunda até sexta) da semana selecionada no calendário. As cores de cada área indicam o intervalo de tempo em que os horários estão ou não a ser cumpridos. O cumprimento do horário continua a ser definido pelos valores para lá dos thresholds selecionados.

Tasks

1. Identificar os intervalos de tempo em que o horário geralmente apresenta atrasos acima de 4 minutos para a linha “Green-E”;
2. Identificar os intervalos de tempo em que a linha “Green-E” não funciona (não tem horário);

3. Identificar os intervalos de tempo em que o horário geralmente apresenta atrasos acima de 7 minutos para a linha “Red”;

Post-test questions

1. Quais as maiores dificuldades na interpretação do significado das várias áreas circulares desta visualização?
2. Quão clara é a seleção da semana que se pretende visualizar?
3. Que outro tipo de informação deveria estar presente nesta visualização?
4. Comentários gerais.

A.3 Test 3 – Vehicle Location Chart

Esta visualização ilustra a distância entre os vários veículos de uma linha, a determinada altura do dia. Vários controlos permitem definir a velocidade a que a animação corre, bem como pausar e resumir a visualização e avançar uma passo de cada vez em ambas as direções do relógio. A animação ilustra os veículos que efetuam serviço em cada uma das direções de forma separada.

Tasks

1. Definir a animação para começar no dia 17 de Janeiro, às 12h, para a linha “Red”;
 - (a) Identificar quais os veículos que estão demasiado próximos uns dos outros;
 - (b) Avançar para a próxima frame;
 - (c) agora quais os veículos que estão demasiado afastados uns dos outros;
 - (d) a animação a correr;
 - (e) o tempo de cada frame para aproximadamente 1.5s;
2. Alterar a animação para a linha “Green-D” e pausar;
 - (a) Determinar quantos veículos estão em serviço nessa linha, em cada direção, à hora selecionada.

Post-test questions

1. Que dificuldades existem em selecionar a hora e o dia da animação?
2. Como poderia ser melhorada a seleção do tempo de cada frame?
3. Quão clara é a informação apresentada na visualização? Quão imediato é identificar distâncias anómalas entre veículos?
4. Quão fácil é controlar o decorrer da animação (pausa, resumo, avançar frames)?

A.4 Test 4 – Map

Esta visualização mostra problemas de cumprimento de horário representados num mapa e de forma animada. Desta forma, operadores da rede de transportes públicos poderão identificar padrões e problemas geográficos da rede. Nesta visualização, o utilizador pode navegar pelo espaço tridimensional utilizando o rato.

Tasks

1. Colocar a animação a começar no dia 14 de Janeiro;
2. Colocar a animação a correr;
3. Pausar a animação quando o dia estiver na terça feira;
4. Colocar a hora nas 14h;
5. Identificar as paragens com menos problemas de cumprimento de horário;
6. Alterar o dia para quinta feira.

Post-test questions

1. Qual é o nível de dificuldade em compreender o significado das barras no mapa?
2. Quais são as dificuldades em compreender os controlos de seleção de dia e hora dos dados?
3. Como classificaria a navegação pelo espaço tridimensional?

A.5 General Questions

Avaliando agora o protótipo como um todo:

1. De que forma considera que este protótipo pode vir a facilitar o trabalho de operadores de redes de transportes públicos?
2. Quão bem o protótipo explica o significado dos seus vários controlos?
3. Quais foram os maiores problemas encontrados no uso do protótipo?
4. Quais foram os controlos mais fáceis de usar/perceber do protótipo?