

A user collaboration model for urban passenger transport

by

António Júlio Albuquerque Diogo Nunes

Thesis submitted to the Faculty of Engineering, University of Porto for the
Doctoral Program in Transport Systems

Supervisors

Maria Teresa Galvão Dias

João Bernardo de Sena Esteves Falcão e Cunha

June 2016

This thesis was supported by Fundação para a Ciência e Tecnologia (FCT) through the MIT Portugal Program under grant SFRH/BD/52356/2013. The research was initially supported by TICE – OneStopTransport project 13843 and TICE – MOBIPAG project 13847, with funding provided by the COMPETE QREN Programme, managed by AdI in the context of European Union FEDER.

Abstract

This thesis is concerned with the development and evaluation of a model of collaboration between users of an urban passenger transport system. This model explores an opportunity created by the widespread adoption of advanced personal mobile devices with ubiquitous access to wireless communication networks, to harness information that is distributed throughout urban passenger transport systems. Its ultimate goal is to facilitate collaboration between users of those systems, using personal mobile devices and sensor-based data to ease their journeys.

This thesis firstly evaluates the potential of user collaboration on the improvement of resource utilisation, and enhancement of experience associated with more sustainable travel alternatives. This thesis subsequently defines the conceptual foundations of the user collaboration model, which hinges upon the creation of dynamic social networks, and requires the active participation of users as providers, consumers, and validators of spatiotemporally structured information. This thesis then describes the development and evaluation of two tools for the user collaboration model. The first relates to a methodology for estimating the destination of passenger journeys from Automated Fare Collection (AFC) data to produce an Origin-Destination (O-D) dataset; it applies to entry-only AFC systems with distance-based fare structures. The second tool relates to a novel concept of Temporary User-Centred Networks (TUNs) for passenger transport, which link users based on circumstantial collaboration potential. Lastly, this thesis describes the design and evaluation of a prototype service *Journata* based on the user collaboration model.

Three main contributions are made by this thesis. The first is the intersection of theory topics to introduce a service science perspective to transport. The second is arguably the most significant, and consists of the user collaboration model more generally, and specifically of the tools to implement it. The third consists of the design and evaluation of the prototype service *Journata* as an instance of the user collaboration model.

Resumo

A presente tese descreve o desenvolvimento e a avaliação de um modelo de colaboração para utilizadores de sistemas de transportes urbanos de passageiros. O modelo tira partido da crescente utilização de dispositivos móveis pessoais sofisticados, com acesso às redes de comunicação sem fios, para recolher informação distribuída em sistemas de transportes urbanos de passageiros. O objectivo é facilitar a colaboração entre os utilizadores e auxiliar as suas viagens, através do uso dos dispositivos móveis pessoais e de dados provenientes de sensores.

A presente tese avalia inicialmente o potencial da colaboração entre utilizadores na maximização de recursos, e na melhoria da experiência associada a alternativas de mobilidade mais sustentável. Esta tese posteriormente define os fundamentos teóricos do modelo de colaboração, centrados na criação de redes sociais dinâmicas, e na participação ativa dos utilizadores na criação, consumo, e validação de informação espaço-temporal. Esta tese descreve ainda o desenvolvimento e avaliação de duas ferramentas para o modelo de colaboração. A primeira destas trata-se de uma metodologia que estima o destino de viagens registadas num sistema de bilhética de transporte público, gerando um conjunto de dados origem-destino; esta aplica-se a sistemas de bilhética que combinam tarifários com base em distância e leitura de bilhetes apenas na origem. A segunda ferramenta trata-se de um novo conceito de redes temporárias centradas no utilizador, que estabelecem ligações entre utilizadores com base numa medida de potencial de colaboração circunstancial. Por último, esta tese descreve o desenho e avaliação de um protótipo de um serviço designado *Journata* que materializa o modelo de colaboração.

A presente tese faz três contribuições para o estado-da-arte. A primeira é a aplicação da ciência dos serviços ao domínio dos transportes. A segunda, que será talvez a mais significativa, trata-se da criação do modelo de colaboração e das ferramentas que o permitem implementar. A terceira centra-se no desenho e avaliação do protótipo *Journata* como uma instância real do modelo de colaboração.

Acknowledgements

It has been a long, winding journey. I wish to thank the kind and talented people who have challenged and supported me, without whom the journey would have got nowhere. Fate had it that a common interest in urban passenger transport would lead me towards my supervisors Prof. Teresa Galvão and Prof. João Falcão e Cunha. I thank Prof. Teresa Galvão for the guidance and encouragement throughout the way, and time and patience for reading my ideas over and over again. I thank Prof. João Falcão e Cunha for the trust and support, and for having a solution for nearly every problem I found as a PhD candidate. I thank Prof. Chris Zegras for welcoming me to his research group during my visit to the MIT – Massachusetts Institute of Technology, and for challenging my models and assumptions. I thank Prof. Jorge Pinho de Sousa for the trust and enthusiasm in my research, and for encouraging me to join the Doctoral Program in Transport Systems. I thank Prof. José Luís Borges for sharing an interest in knowledge extraction, which turned out to steer my research focus.

The present research was enriched by working in collaboration with three Masters students in Informatics and Computing Engineering. I thank Tiago Gonçalves, Luís Dias, and Marco Amador for their interest in my research proposals. I learned a lot advising their studies, and really enjoyed their company in the research group. The shared office space has been a fun and intellectually stimulating environment. I thank my research project colleagues Marta Ferreira, André Cunha, Pedro Costa, Tânia Fontes, and Vera Costa, for the companionship, support, and fruitful contributions to my research. I also thank other talented researchers whom I shared the office with, for their willingness to share knowledge, with a special mention to Carlos Gomes, Fabrício Sperandio, Miguel Oliveira, and Eduardo Pinto. I thank Inês Frade, my fellow doctoral student and expat buddy at the MIT, for the friendship and advice. And I thank all participants in various experiments for supporting the development of *Journata*.

I thank STCP and TIP for kindly providing the valuable data used in this research. I thank the people at INEGI for making me feel welcome at their offices, for the day-to-day admin support, and for the financial contribution towards my tuition fees. I also thank other organisations for the involvement of their staff advising on this research, namely OPT, AMTP, and Metro do Porto.

I am deeply grateful to my parents, sisters, niece, and nephews, for their emotional support throughout the way, and for patiently listening to my concerns. I wish I keep making you proud, so that it is all worth it in the end. I thank my parents-in-law for helping Marta and I, often sacrificing their own plans to let us enjoy our scarce spare times. Last but not least, I owe my deepest gratitude to my small but warm household. To Pixel for constantly reminding me there is a life to be lived outdoors, in any weather. And to my dearest wife Marta, who, eleven years, two thesis, and three international relocations later, still is by my side everyday. I love you!

Publications and presentations

The following articles, abstracts, and posters were created in the course of the research described in this thesis.

Articles in scientific journals, first author

Nunes, António A., Teresa Galvão Dias, Chris Zegras, and João Falcão e Cunha. 2016. “Temporary User-Centred Networks for Transport Systems.” *Transportation Research Part C: Emerging Technologies* 62: 55-69. doi:10.1016/j.trc.2015.11.006.

Nunes, António A., Teresa Galvão Dias, and João Falcão e Cunha. 2016. “Passenger Journey Destination Estimation from Automated Fare Collection System Data Using Spatial Validation.” *IEEE Transactions on Intelligent Transportation Systems* 17 (1): 133-42. doi:10.1109/TITS.2015.2464335.

Articles in international conference proceedings, first author

Nunes, António A., Teresa Galvão, and João Falcão e Cunha. 2014. “Urban Public Transport Service Co-Creation: Leveraging Passenger’s Knowledge to Enhance Travel Experience.” *Procedia - Social and Behavioral Sciences* 111: 577–85. doi:10.1016/j.sbspro.2014.01.091.

Nunes, António A., Teresa Galvão, João Falcão e Cunha, and Jeremy V. Pitt. 2011. “Using Social Networks for Exchanging Valuable Real Time Public Transport Information among Travellers.” In *IEEE Conference on Commerce and Enterprise Computing*, 365–70. doi:10.1109/CEC.2011.60.

Articles presented at international conferences, first author

- Nunes, António A., Tiago Gonçalves, and Teresa Galvão. 2013. “A Prototype for Public Transport Service Co-Creation Using Social Media: Results from Usability Testing.” In *International Conference on Exploring Service Science 1.3*.
- Nunes, António A., Teresa Galvão, and João Falcão e Cunha. 2012. “Public Transport Service Co-Creation: A Social Media Based Framework.” In *The Art & Science of Service Meeting*.

Abstracts and posters presented at national conferences, first author

- Nunes, António A. 2015. “Linking public transport passengers in real-time.” In *IEMS '15 | 6th Industrial Engineering and Management Symposium*.
- Nunes, António A. 2014. “Reimagining the social network for urban public transport.” In *IEMS '14 | 5th Industrial Engineering and Management Symposium*. Winner of best poster award.
- Nunes, António A. 2013. “The use of pervasive mobile computing based services to improve urban public transport experience.” In *IEMS '13 | 4th Industrial Engineering and Management Symposium*.

Articles in international conference proceedings, co-author

- Costa, Pedro Maurício, Tânia Fontes, António A. Nunes, Marta C. Ferreira, Vera Costa, Teresa Galvão Dias, José Luís Borges, João Falcão e Cunha. 2016 (accepted for presentation). “Application of collaborative information exchange in urban public transport: the Seamless Mobility solution”. In *Proceedings of 6th Transport Research Arena*.
- Costa, Pedro Maurício, Tânia Fontes, António A. Nunes, Marta C. Ferreira, Vera Costa, Teresa Galvão Dias, José Borges, João Falcão e Cunha. 2015. “Seamless Mobility: a disruptive solution for public urban transport”. In *22nd ITS World Congress, Bordeaux*.

Table of contents

1. Introduction	1
1.1. Motivation.....	1
1.2. Problem and objectives	5
1.3. Methodological overview	8
1.4. Contributions	10
1.5. Thesis structure	11
2. Literature review	13
2.1. Customer experience.....	13
Experience quality.....	15
Travel experience.....	16
User experience.....	24
2.2. Value co-creation	24
The notion of value	26
Who are the co-creators?	27
Co-creation in the public sector	29
Co-creation in passenger transport	31
2.3. ICT-based user collaboration	33
Crowdsourcing spatiotemporal data	35
Crowdsourcing in passenger transport.....	36
Passenger information requirements.....	39
Evaluation of user collaboration potential	42
2.4. Travel pattern extraction and estimation.....	43
Mobile phone network data	45
Traffic sensor data.....	46
Public transport AFC system data.....	47
2.5. Journey path analysis	51
Similarity metrics	51
Clustering methods	53
Carpool service problem	54

Relevance in information retrieval	55
2.6. Review summary	55
3. Foundations of the collaboration model.....	59
3.1. Leveraging collective intelligence	59
3.2. Information value	62
Information value for travellers	63
Information value for transport providers	65
3.3. User participation	67
3.4. Reimagining the social network	69
3.5. Validation and incentive mechanisms	70
3.6. Model summary.....	72
4. Passenger journey destination estimation.....	75
4.1. Methodology scope.....	76
4.2. Dataset description	77
4.3. Data preparation	81
4.4. Destination estimation methodology	83
4.5. Algorithm structure and implementation	89
4.6. Discussion of results	93
4.7. Multi-modal upgrade.....	97
4.8. Chapter conclusions	98
5. Temporary User-Centred Networks (TUNs)	101
5.1. Concept	102
5.2. Mathematical formulation.....	104
Journey similarity (SIM)	105
Journey substitutability (SUB).....	107
Relevance	111
5.3. TUNs algorithm structure	113
Sourcing data.....	114
Calculating SIM	116
Calculating SUB.....	118
Compiling relevance.....	120
5.4. Data subset description.....	121

5.5.	Implementation	125
5.6.	Discussion of results.....	126
5.7.	Indexing method	133
5.8.	Chapter conclusions	135
6.	Case study: <i>Journata</i>	137
6.1.	Non-functional prototype: the concept	138
6.2.	Functional prototype, first iteration	141
	Implementation	144
	Evaluation	145
	Results	149
6.3.	Functional prototype, second iteration	152
	Evaluation	156
	Results	157
6.4.	<i>Journata</i> in <i>Seamless Mobility</i>	158
6.5.	Chapter conclusions	160
7.	Conclusions	163
7.1.	Findings and contributions.....	163
7.2.	Future research directions.....	167
	References	169

List of figures

Figure 1.1 Research map	7
Figure 1.2 Methodology	8
Figure 2.1 Literature review map.....	56
Figure 3.1 Public transport user entries, bus and light-rail, 15 minute interval, Porto	61
Figure 3.2 Fundamentals of the user collaboration model.....	73
Figure 4.1 Geographic spread of boarding transactions in the subset of data and structure of travel zones	80
Figure 4.2 Application of key assumptions exemplified	85
Figure 4.3 Methodology flowchart.....	90
Figure 4.4 Daily average journey transactions throughout the day	91
Figure 4.5 Single daily journeys with multiple stages (three stage example).....	92
Figure 5.1 Journey similarity (SIM) and journey substitutability (SUB)	106
Figure 5.2 Journey with SIM and SUB scores	108
Figure 5.3 Journey paths with maximum SUB at shorter substitute portion.....	109
Figure 5.4 SUB according to distance between location points within SPODs for various levels of substitution potential.....	111
Figure 5.5 Flowchart of main activities	113
Figure 5.6 Time chart	114
Figure 5.7 Flowchart of activity ‘sourcing data’	115
Figure 5.8 Flowchart of activity ‘calculating SIM’	117
Figure 5.9 Flowchart of activity ‘calculating SUB’	119
Figure 5.10 Flowchart of activity ‘compiling relevance’	120
Figure 5.11 Similar portion within substitute portion.....	121
Figure 5.12 Journey stage transaction records throughout the day	122
Figure 5.13 Time chart of ongoing journeys.....	123
Figure 5.14 Ongoing journeys at 11:00	124
Figure 5.15 Aggregate results at 11:00.....	127
Figure 5.16 TUN evolution for selected passenger	129

Figure 5.17 Journey paths of selected passengers and highly relevant peers at 11:00	130
Figure 5.18 Journeys hourly across simulation period	134
Figure 5.19 New distinct journeys hourly across simulation period.....	134
Figure 6.1 <i>Journata</i> prototype iteration evolution.....	137
Figure 6.2 <i>Journata</i> logo in alternative shade combinations.....	137
Figure 6.3 Non-functional prototype interface screen mock-ups	139
Figure 6.4 Interface screen mock-ups created in Photoshop.....	140
Figure 6.5 First iteration functional prototype interface screens	143
Figure 6.6 Architecture of the <i>Journata</i> prototype.....	145
Figure 6.7 Experiment environment	147
Figure 6.8 Post-task questionnaire results.....	149
Figure 6.9 Post-test questionnaire results.....	151
Figure 6.10 Mock-up samples of preliminary interface designs	153
Figure 6.11 Second iteration functional prototype interface screens	155
Figure 6.12 Post-task questionnaire results.....	157
Figure 6.13 Social component interface screens of <i>One Ride</i>	160
Figure 7.1 Methodology (reprinted)	165

List of tables

Table 2.1 Summary of travel experience factors in the literature.....	22
Table 2.2 Overview of passenger information requirements identified in the literature	41
Table 4.1 Data summary	79
Table 4.2 Network summary (dataset used).....	81
Table 4.3 Irregular transaction records after pre-processing.....	82
Table 4.4 Results of the methodology applied to the Andante dataset	95
Table 4.5 Sensitivity analysis of the cut-off Euclidean distance parameter	96
Table 5.1 Data subset summary.....	125
Table 5.2 Relevant peers to selected passenger	128
Table 5.3 Analysis of TUNs across time intervals at various relevance threshold levels	132
Table 6.1 Participant characteristics	146
Table 6.2 Post-task questionnaire	148
Table 6.3 Post-test questionnaire-based interview	148
Table 6.4 Post-task questionnaire	156

List of acronyms

AFC	Automated Fare Collection
APC	Automatic Passenger Count
AVI	Automatic Vehicle Identification
AVL	Automatic Vehicle Location
C2C	Customer-to-Customer
DRT	Demand Responsive Transport
EC	Experience Component
EF	Experience Factor
ETC	Electronic Toll Collection
FEUP	Faculty of Engineering, University of Porto
HCI	Human-Computer Interaction
ICE	In-Car Entertainment
ICT	Information and Communication Technology
ITS	Intelligent Transport System
JPRI	Journey Pair Relevance Index
O-D	Origin-Destination
POI	Point of Interest
PTAL	Public Transport Accessibility Level
RAM	Random Access Memory
RFID	Radio-Frequency Identification
SIM	Journey similarity
SPOD	Shared Potential Origins or Destination
STCP	Sociedade de Transportes Colectivos do Porto, SA
SUB	Journey substitutability
TfL	Transport for London
TUNs	Temporary User-Centred Networks
UX	User Experience
VD	Vehicle Detector

1. Introduction

This thesis is concerned with the development and evaluation of a model of collaboration between users of an urban passenger transport system. The thesis sets out the theoretical framing, conceptual foundations, and tools for the model, and describes the evolution of a case study that emerged from it. The model explores an unprecedented opportunity created by the widespread adoption of advanced personal mobile devices with ubiquitous access to wireless communication networks, to harness information that is distributed throughout urban passenger transport systems. Its ultimate goal is to improve utilisation of available resources to enhance travel experience and to promote sustainable urban transport. This chapter describes the motivation for the research, its objectives and methodological approach, and how it contributes to the state-of-the-art.

1.1. Motivation

The motivation for this research consists of two aspects, a need and an opportunity. The need is high in the political agenda, and relates to the reduction of adverse environmental and social impacts of urban passenger transport. The opportunity came in recent years with advances in Information and Communication Technologies (ICT), particularly with the evolution of personal mobile devices and their rampant popularity, combined with the proliferation of wireless networks. Starting from the need, and in order to fully grasp its complexity, one may reflect on the historical role of transport in the shaping of contemporary cities and as an enabler of modern urban lifestyle. The city has been described metaphorically throughout history as a human body. This dates at least as far back as Classical Greece in the writing of Plato who argued that cities exhibit the same virtues of human souls (Plato and Bloom 1991). The analogy has been perpetuated in the vocabulary that is commonly used to describe urban form. The understanding of the city as a human body has allowed the mapping of centre as its heart and of transport routes as arteries expanding outwards from it (Shibatani and Thompson 1999).

Figuratively speaking, transport is the blood of the city; literally speaking, it keeps the city alive.

Yet this romanticised vision has been challenged in most cities around the world. Technological advances in transport have granted the general public in developed countries the freedom to travel faster and across greater distances. However, the utopian gift of personal mobility was largely awarded at the expense of the environment and society as a whole (Low 2003, 1). Increasing reliance on private cars and fossil fuels for personal travel resulted in ‘auto sprawl syndrome’, which consists of two main spiralling symptoms, suburbanisation and private car dependency (SceneSusTech 1998, 100).

The first of these symptoms is widespread car ownership since World War II, initially and foremost in North America but also to a great extent in Western Europe. It acted as a catalyst for the suburbanisation of cities and consequently to the weakening of their hearts (Lucy and Phillips 1997, 260; Batty, Besussi, and Chin 2003, 9; Antrop 2004, 13; Glaeser and Kahn 2004, 2483; Richardson and Bae 2004, 1). As for the second symptom, suburbanisation led to greater travel distances and to a sustained increase in car use, since disperse forms of sprawling urbanisation reduce the viability of public forms of transport (Jansen 1993, 9). Rising car use in turn overwhelms road network capacity producing traffic congestion and heightening the environmental impact of personal travel. Meanwhile, this problem also became a reality in developing countries such as China and Brazil where car ownership levels began rising dramatically and are forecast to grow steadily (Han and Hayashi 2008; Motta, Silva, and Brasil 2012).

The adverse social impacts of car dominance have been exposed from a social geography perspective of inclusion (Boschmann and Kwan 2008), and from the urban planning and design perspectives of city liveability (Gehl 2010). The social inclusion aspect relates to the provision of transport that support equitable access to employment, leisure, and services, preventing isolation of lower income groups. The city liveability aspect applies to the preservation of cities as places for walking and cycling that encourage street activity and democratic use of public space, whilst safely

catering for travel demand. Today, many city dwellers travel on a daily basis for a variety of purposes, and rely on the transport systems that are available to them. Urban populations keep growing globally, and so will their travel needs and the strain on urban sprawl. In the year 2007 the global urban population equalled the global rural population, and is expected to become two-thirds of the world population by the year 2050 (United Nations 2014).

Provided that transport keeps cities alive, restraining mobility cannot be seen as an option if urban vitality is to be reinforced or even to remain unchanged. Hence it is widely acknowledged that urban mobility patterns must be readapted to meet increasing needs and desires to travel albeit in a more sustainable way (European Commission 2011). This raises the quintessential question on the meaning of sustainable transport. Several definitions exist, but one that attempted to harmonise them captures most successfully how sustainable transport is understood here:

A sustainable transport system is one that provides transport and mobility with renewable fuels while minimizing emissions detrimental to the local and global environment, and preventing needless fatalities, injuries, and congestion. (Black 2010, 10, emphasis in the original)

The definition is entirely aligned with the identified need for mitigating the socially and environmentally adverse impacts of transport, as long as it is understood that the social inclusion aspect is implicit in the notion of providing transport and mobility. One way to decrease those impacts is to encourage a shift from private to public modes of transport for personal travel. Public transport has potential to reduce emissions and congestion, to lessen the probability of traffic related injuries, to provide equitable accessibility for being affordable, and to release public space taken up by traffic and parking towards social activities and ‘softer’ modes of transport (Litman 2015, 2). Car use may also become more sustainable with the adoption of lower emission vehicles using renewable fuels (Granovskii, Dincer, and Rosen 2006; Faria et al. 2012), and with the introduction of new technologies designed to improve safety (e.g. Kusano and Gabler 2012). However, cars will hardly mitigate social inclusion issues, and their scope for reducing congestion and releasing public space is

constrained. Achievable improvements among these topics may largely relate to the increase of average vehicular occupancy, to the growth of car sharing and ride sharing schemes, and to the reduction of travel time via routing intelligence.

The promotion of sustainable urban transport is already happening in many different ways, and the most successful approaches will likely combine several complementary measures (Sloman et al. 2010). Hence this thesis is not intended to overcome singlehandedly that challenge, but to contribute towards the reduction of the adverse impacts of urban passenger transport from a novel perspective of collaboration between its users. This guides the present argument to the second aspect providing the motivation for this research, mentioned earlier as the opportunity.

In recent years, the rapid evolution and widespread adoption of personal mobile devices with advanced computing and sensing abilities (e.g. the ‘smartphone’), combined with the ubiquity of wireless communication networks, has brought us closer to a ‘pervasive computing’ paradigm (Saha and Mukherjee 2003). That has profoundly changed how people access and share information with each other without the need for fixed computing terminals or wired communication networks. It also acted as a catalyst for the development and growth of location-based services, which tailor information according to the context of the user (Rao and Minakakis 2003). The ability to communicate permanently and on the move increased significantly the timeliness in which information can be produced and consumed, and how useful it may be to others (Manovich 2009). Combined with the scattering of advanced mobile devices (Smith et al. 2015), a holistic view of an urban transport system can be obtained at any time by harnessing information that is distributed throughout it.

Besides mobile devices, advances in ICT have also materialised into the proliferation of embedded sensors across transport system infrastructures. Whilst designed and implemented with very specific purposes, a variety of sensors for collecting fares, locating vehicles, and counting passengers, generate huge volumes of data able to provide insight into the dynamics of a transport system and users within it (Rosado 2014). All things considered, transport systems of today are mobile-saturated, data-

intensive environments that facilitate collaboration between its users. Examples of potential collaboration include the sharing of transport-related information in real-time, and the identification of ride sharing options. The aforesaid environment affords an opportunity to aim for a reduction on the adverse impacts of urban passenger transport, by improving utilisation of available resources and enhance the experience associated with more sustainable travel alternatives. Perhaps the greatest motivation for this research stems from a belief that this opportunity has largely been neglected to the present day. This thesis wishes to trigger a change to that.

1.2. Problem and objectives

The objectives for this research were driven by an overarching goal of contributing towards the sustainability of urban passenger transport. That is effectively the need identified in the previous section. Despite it being widely acknowledged, urban local governments have been struggling for years to stimulate behavioural shifts towards more sustainable transport. This thesis takes on that endeavour from a user collaboration perspective that leverages recent advances in ICT. Hence the overarching goal translates into the following research problem:

How to facilitate collaboration between users of an urban passenger transport system using personal mobile devices and sensor-based data to ease their journeys?

This thesis has three main research objectives (RO), which were specified to divide the abovementioned research problem into a well-defined set of steps:

RO.1. Assess the potential of user collaboration on the improvement of resource utilisation and enhancement of experience associated with more sustainable travel alternatives;

RO.2. Define the conceptual foundations and develop the tools for a user collaboration model in urban passenger transport; and

RO.3. Design and evaluate a prototype service for mobile devices, based on the user collaboration model, in the urban public transport domain.

The first objective (RO.1) provides the theoretical framing and sets the scene for the remainder of the thesis. Its purpose is to understand how collaboration helps to make better use of existing transport resources, and to what extent that can improve travel experience. Achieving this objective requires an in-depth theoretical understanding of the concept of customer experience and the identification of factors that may influence that experience in the context of passenger transport. Furthermore, it involves the identification of passenger information requirements that can be obtained through user collaboration, in order to mitigate travel experience issues or generate new travel alternatives. Lastly, it involves a review of collaboration mechanisms from various service domains and an understanding of their applicability to urban passenger transport.

The second objective (RO.2) is considered to be the main theoretical contribution of this thesis. It provides the conceptual foundations for the user collaboration model, inasmuch as the mathematical formulation and algorithmic tools to implement it across urban passenger transport domains. The third objective (RO.3) is related to the empirical evaluation of the feasibility of the model through its application to the urban public transport domain. The design of a prototype service for mobile devices called *Journata*, which was iteratively developed in collaboration with other researchers initially, and more recently as part of a project with a wider institutional consortium called *Seamless Mobility*, provided a case study for evaluating the model. The majority of time dedicated to this research consisted of interchanging between RO.2 and RO.3, which allowed a dialectical refining of theoretical aspects from practice, and practicing the theory on the prototype service. Figure 1.1 summarises this section as a research map.

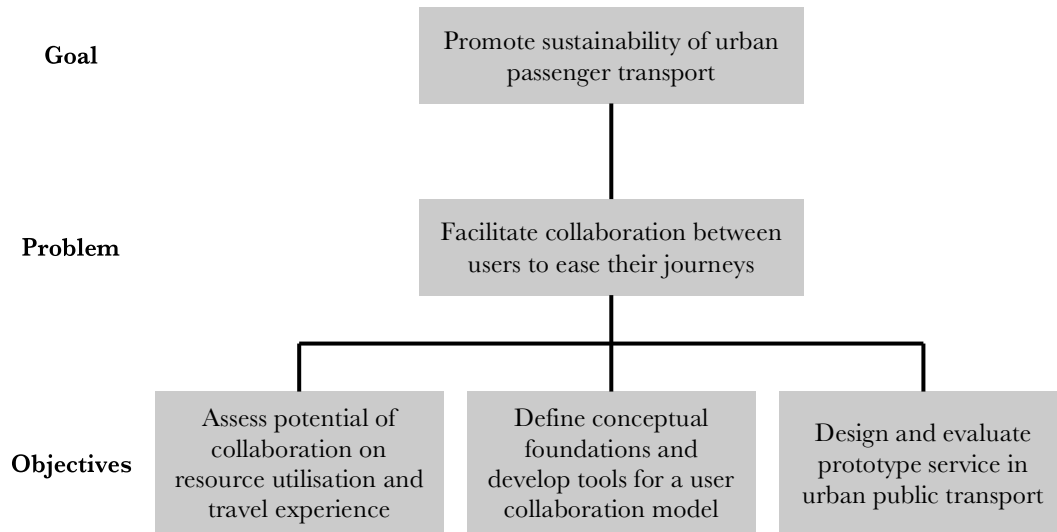


Figure 1.1 Research map

This thesis was originally intended to address exclusively the urban public transport domain, and focus on how collaborative exchanges of information between passengers could improve travel experience. However, as the model evolved, its future potential for application across other urban passenger transport domains became evident, and additional forms of collaboration for improving the utilisation of resources were identified. Hence the problem was broadened in its scope. Private car use, urban cycling, and Demand Responsive Transport (DRT) are among those other domains, in which user collaboration may potentially yield diverse benefits including the identification of ride sharing opportunities, and the reduction of travel time. The fundamentals of the model, including its conceptual foundations and mathematical formulation, are generalised across domains. The algorithmic tools are structurally similar in most cases, but their implementation details must accommodate the specifics of the domain and actual system of interest. For the research described in this thesis, the specifics reflect the initial and primary interest in the urban public transport domain and the characteristics of the system where data came from.

As a result, for the purpose of disambiguation, the expression ‘urban passenger transport system’ in the research problem is used throughout the thesis as a generalisation. It can either relate to a specific mono-modal system or network (e.g. bus services or road network), integrated multi-modal system (e.g. various services

represented by a transport agency), or to the entire set of passenger transport alternatives available to the public within a specific urban area, including cycling.

1.3. Methodological overview

The research typology adopted combined applied research and experimental development. The first two objectives (RO.1 and RO.2) are characteristic of applied research, being concerned with the theoretical framing and development of a new concept that has materialised into the user collaboration model for urban passenger transport. The last objective (RO.3), relating to the design and evaluation of *Journata*, falls within experimental development because it drew from the model to deliver a new service that is expected to be publicly available in the near future following its ongoing field trial. Figure 1.2 outlines the adopted methodology, which is described throughout this section.

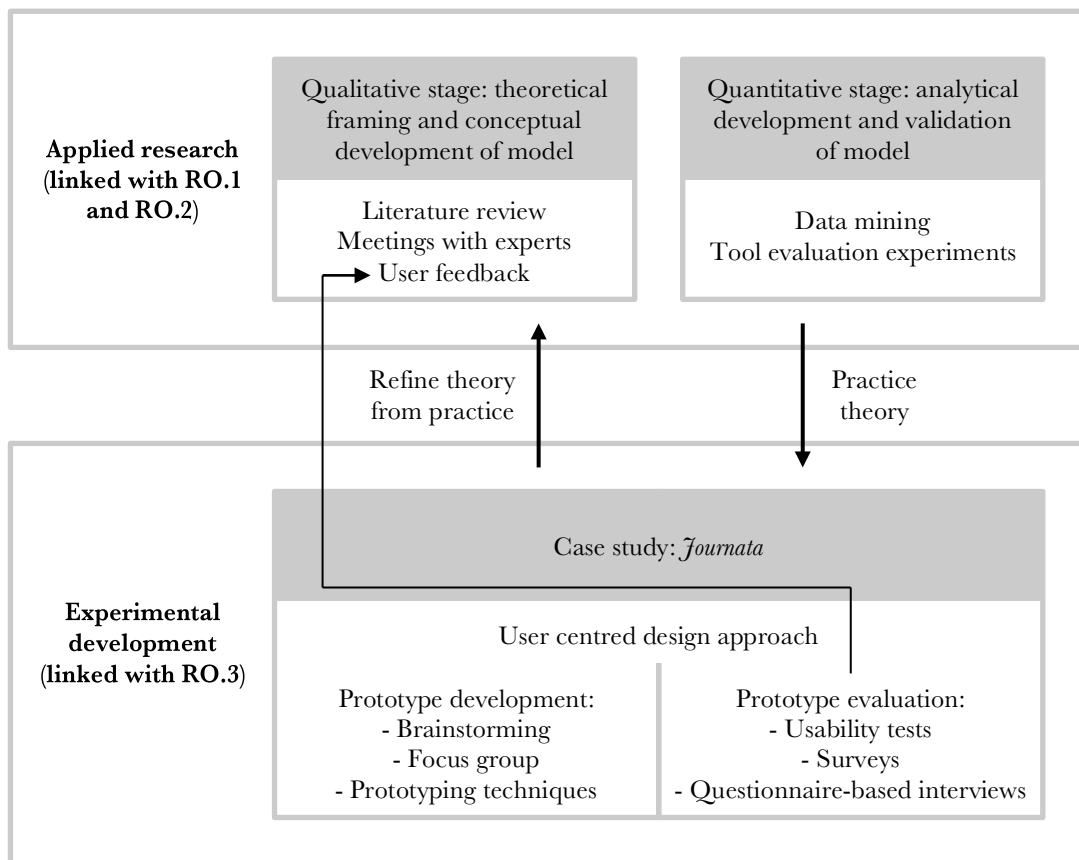


Figure 1.2 Methodology

The applied research part followed a mixed methods approach, qualitative in the outset but having gained a quantitative bias upon mathematical formulation of the model. The literature review was the starting point and reflects the multidisciplinary nature of the problem by covering several topics within transport research, service science, information science, and network science. The literature review provided the theoretical framing for the research. In addition, it contributed towards the conceptual foundation of the user collaboration model, alongside opinions gathered from meetings with experts, and feedback obtained from users of the first iteration of *Journata* as detailed later in this section. The shift from concept to object came with the development of a theoretical notion of spatiotemporal relevance between users of a passenger transport system. This key outcome allowed the mathematical formulation of the model as a measure of circumstantial collaboration potential between users. This marked the turning point to the predominantly quantitative stage.

Two algorithms were developed as data mining tools for the model, and were applied to a large dataset in the urban public transport domain for evaluation. The first is concerned with estimating the destination of passenger journeys to produce an Origin-Destination (O-D) dataset. The second builds on that dataset to implement the mathematical formulation of the model and produce dynamic estimates of collaboration potential between users. The effectiveness and efficiency of these algorithms were evaluated using spatial validation rules and mathematical concepts relating to random graph theory.

The experimental development part wrapped several research methods into a case study. A distinctive feature of this case study is that it emerged from the iterative development of *Journata*. Therefore, the case study itself is an outcome of the thesis, having evolved alongside the user collaboration model to enable its empirical evaluation. The development of *Journata* followed a user centred design approach. Research methods included a brainstorming session and a focus group to elicit requirements, prototyping techniques to define the design, and evaluation with users. The latter consisted of usability tests accompanied by surveys and questionnaire-

based interviews. Each design iteration of *Journata* incorporated new advances from the model, and its evaluation provided feedback to close the loop.

1.4. Contributions

This thesis provides three main contributions. The first is to the state-of-the-art, by bringing a service science perspective that is scarce in the transport literature. This is a reaction to the critique that the public sector, which many urban passenger transport providers are part of or controlled by, is lagging behind in the adoption of new value creation practices and tools (Cassia and Magno 2009). This contribution draws from the literature in other service domains and highlights the potential role of user collaboration in the mobile-saturated, data-intensive setting of urban passenger transport to improve resource utilisation, enhance travel experience, and reduce adverse environmental and social impacts.

The second main contribution is also to existing theory and consists of the user collaboration model. The concept is believed to be entirely new, and replicable across urban passenger transport domains and systems around the world. It provides tools to explore collaboration opportunities for the benefit of travellers and transport providers with improved experience and travel decision support, and for the benefit of society and the environment with the promotion of more sustainable transport. At a detailed level, each of the algorithms developed as tools for the model also contribute to the state-of-the-art.

The first algorithm for estimating the destination of passenger journeys refines assumptions from previous works, and is the first in the literature dedicated to entry-only Automated Fare Collection (AFC) systems with distance-based fare structures. The meaning of these terms will be explained in Chapter 4, but essentially entry-only refers to systems where only passenger departures are recorded, and distance-based means that journey fares depend on distance travelled. The output of the algorithm can produce reliable O-D matrices at any level of aggregation and geographic coverage, which are also useful for service performance monitoring and transport planning support. The second algorithm for producing dynamic estimates of

collaboration potential between users is also believed to be innovative. It streamlines computation of path similarity for the case of passenger transport where those paths have to follow a predefined network grid, it introduces a measure of path substitutability as the degree to which two paths constitute alternatives to each other, and it presents an aggregate measure of spatiotemporal relevance between users. This algorithm is key to materialise the model.

The third main contribution is the prototype service *Journata*. Not only it served as a case study for the empirical testing and validation of the user collaboration model, it is also expected to be generally available to the public in the near future if it succeeds in its ongoing field trial. *Journata* is the corollary of this thesis as a predominantly applied research effort, being a tangible outcome that is value brought back to the society.

1.5. Thesis structure

This thesis is organised in the following way. Chapter 2 provides a review of the literature and other background material to this research. Due to the multidisciplinary nature of the problem, the review covers several topics. It reviews the notion of customer experience, its relation to passenger transport, and impact on travel behaviour. It reviews recent value creation trends in the context of services, namely the rise of the Web 2.0 as a form of value creation through collaboration. It reviews the state-of-the-art of ICT-based user collaboration in urban passenger transport, including its potential in light of graph theory, and relevant mobile platforms and how they overcome information needs of travellers. In regards to the development of algorithmic tools, it surveys existing types of data sources in urban passenger transport domains and methodologies for extracting and estimating travel patterns. It also reviews trajectory similarity metrics and clustering methods and their suitability to the measurement of collaboration potential between users.

Chapter 3 describes the conceptual foundations of the user collaboration model. It begins by describing *what* the opportunity and main objective of this model consist of. Then it discusses *why* travellers and transport providers may benefit from the model.

Lastly, it defines *how* the model works. This includes the notion of users as information providers, reimagining the social network concept to account for the dynamics of urban passenger transport, and the role of validation and incentive mechanisms that are applicable to some forms of collaboration, for sustaining the reliability and breadth of information obtained from users.

Chapter 4 presents the first tool for the model and its application to a dataset in the urban public transport domain. Firstly it details the characteristics of the transport system and the attributes of the data available. This is followed by the formal definition of the methodology to estimate the destination of passenger journeys, and by a description of the structure of the algorithm to implement it. Lastly, it presents and discusses the results, and provides a sensitivity analysis of the trade-off between extensiveness and accuracy of estimates.

Chapter 5 presents the second tool for the model and its application to an enriched dataset obtained from the algorithm described in Chapter 4. It defines the proposed notion of spatiotemporal relevance between users of a passenger transport system. Then it provides essential definitions, the mathematical formulation of the user collaboration model, and the structure of the algorithm to perform the required calculations. This is followed by the implementation of the algorithm, and the presentation and discussion of initial results. Lastly, some efficiency improvements to the algorithm are described.

Chapter 6 describes the evolution of the case study *Journata* from its inception as non-functional prototype through to its implementation as a commercial prototype service. This includes the details of two intermediate versions of functional prototypes, and the research methods adopted for eliciting requirements, designing, and evaluating them. The results obtained from the evaluation at different stages of development are presented and discussed.

Chapter 7 concludes this thesis. It summarises the main findings and contributions of the research, and presents future research directions.

2. Literature review

This literature review covers several topics within transport research, service science, information science, and network science, reflecting the multidisciplinary nature of the research problem. It attempts to establish the necessary relationships between different theoretical fields to support the conceptual foundation of the user collaboration model. Section 2.1 reviews the notion of *customer experience*, how it relates to passenger transport, and its potential role influencing travel behaviour. Section 2.2 reviews *value creation* trends in services to frame the potential benefits of active engagement of transport users. Section 2.3 reviews the state-of-the-art of *ICT-based user collaboration* as a form of value creation in services in general, but particularly in the urban passenger transport domains. This section also links collaboration to graph theory, as a way to evaluate its potential to overcome some information needs of travellers. Section 2.4 reviews existing types of data sources and methodologies for *estimating travel patterns*, as prerequisites for implementing the user collaboration model. Section 2.5 reviews *trajectory similarity metrics* and *clustering methods*, and the notion of *relevance in information retrieval*. This provides the background for the measurement of collaboration potential between users, and so to the materialisation of the model. Section 2.6 summarises the literature review.

2.1. Customer experience

Over the past two decades, the marketing literature has given attention to the creation of value for customers in the shape of experiences. Pine and Gilmore (1998) coined the term ‘Experience Economy’ as the latest phase in the historical progress of economic value, after the chronological sequence of the agrarian, industrial, and service economies. The authors argued that services were becoming increasingly commoditised, as did products in a previous phase, and that the new competitive battleground lied on staging experiences. The argument drew from practices that were common in the entertainment business, yet it became influential by expressing that the experience economy was spreading and becoming a differentiating factor

across many business sectors. Neither their seminal article nor subsequent book (Pine and Gilmore 1999) explicitly defined experience. Instead the authors provided economic distinctions and characterised experience as occurring “when a company intentionally uses services as the stage, and goods as props, to engage individual customers in a way that creates a memorable event” (Pine and Gilmore 1998, 98), adding that it is “inherently personal, existing only in the mind of an individual who has been engaged on an emotional, physical, intellectual, and even spiritual level” (99). The authors gave various examples, one of which from passenger air transport. This example mapped the ‘commodity mind-set’ to the functional view of moving people from A to B cheaply and on time, and mapped experience to providing a respite for travellers away from their hectic lives, using the travel service merely as a stage (Pine and Gilmore 1998).

The business world has since been charmed with the notion of customer experience as a competitive differentiator (Bitner, Ostrom, and Morgan 2008, 69). And indeed the recipes of the experience economy accelerated the development of practices of ‘customer experience management’, which recognise the significance of emotional bonds between companies and customers, and how these are nurtured in every interaction across multiple channels (Berry, Carbone, and Haeckel 2002; Smith and Wheeler 2002; Schmitt 2003; Verhoef et al. 2009). This continued focus on customer experience triggered attempts to produce definitions for it (e.g. Shaw and Ivens 2002, 6; Gentile, Spiller, and Noci 2007, 397; Meyer and Schwager 2007, 2) but a struggle to find consensus remains today. This thesis is not intended to dwell into the debate and embraces the last revision of an evolving definition that corroborates the characterisation above. It highlights the sensorial and emotional elements of experience, inasmuch as the role of the level of consciousness of the customer:

A Customer Experience is an interaction between an organization and a customer as perceived through a customer’s conscious and subconscious mind. It is a blend of an organization’s rational performance, the senses stimulated and emotions evoked, and intuitively measured against customer expectations across all moments of contact. (Shaw, Dibeehi, and Valden 2010, 3)

The focus on customer perceptions in the above definition implies a subjective nature to the construct. In their early work on customer experience engineering, Carbone and Haeckel (1994, 9) noted that customers could not help but have an experience with a product or a service. Yet a *good* experience is not guaranteed (Berry and Carbone 2007, 27). Their standpoint may seem cautious if compared with the dichotomy of service delivery and staging of experiences inherent to the experience economy, but Berry and Carbone (2007) were equally adamant of benefits deriving from the new focus. The authors have argued that an organisation cannot control the subjective emotional responses from customers but can systematically manage the clues that trigger them. Those clues are embedded in every way customers experience the organisation, and if managed effectively they can collectively meet or exceed their expectations (Haeckel, Carbone, and Berry 2003, 20). What is vital about their argument is that those emotions will drive behaviour, hence are the crux of building loyalty. “Emotional connection extends far beyond customer satisfaction” (Berry and Carbone 2007, 32). People who love a brand or identify themselves with the values of an organisation are bound to be loyal customers.

Experience quality

Notions of customer experience management kept attracting scholarly interest, alongside the continued effort from organisations to adopt them into their value propositions to raise loyalty (e.g. Frow and Payne 2007; Verhoef et al. 2009). Therefore, the design of methodologies to measure experience quality and assess its impacts on customer behaviour has been a concern, in spite of conceptual difficulties arising from the degree of subjectivity involved, and practical obstacles in developing and implementing scales that are able to capture its affective components (Palmer 2010). Those methodologies are generally rooted in previous notions of service quality and customer satisfaction, and in quality management frameworks such as SERVQUAL (Parasuraman, Zeithaml, and Berry 1988), yet fundamental distinctions exist.

Service quality can be objectively measured from a specific set of functional attributes, whereas experience quality is subjectively and holistically evaluated by the

customer (Otto and Ritchie 1996). Adding to the cognitive nature of service quality assessment, the holistic evaluation of experience quality comprises an affective component that is subjective and emotional (Otto and Ritchie 1996). Shaw, Dibeehi, and Valden (2010, 3) argued that more than half of an experience is about the emotions and feelings it evokes for the customer, hence its assessment extends beyond the evaluation of service quality. The aforementioned SERVQUAL framework has indeed been argued not to consider the affective component contributing to the overall customer experience (Fick and Brent Ritchie 1991).

Customer satisfaction in turn relates to “the perceived discrepancy between prior expectation and perceived performance after consumption” (Chen and Chen 2010, 30). This implies less depth as a mainly cognitive construct, but greater breadth as a judgement that exceeds the actual experience alone. In other words, it suggests a post-consumption, more pondered assessment of the gap between expected and delivered performance. Another perspective considers customer satisfaction to be the net outcome of a series of customer experiences benchmarked against expectations (Meyer and Schwager 2007, 2). According to this, understanding how to improve satisfaction requires its deconstruction into the constituent experiences.

Existing methodologies to measure experience focus on the identification of relevant factors in a specific service domain of interest, for subsequent qualitative analysis (e.g. Rowley 1999) or quantitative evaluation combining factor analysis and structural equation modelling (e.g. Chen and Chen 2010; Maklan and Klaus 2011). The latter type enabled for an empirical demonstration that experience quality has greater explanatory power than customer satisfaction on loyalty (Maklan and Klaus 2011), supporting theoretical assumptions on the importance of customer experience for organisations.

Travel experience

Turning back to the aims of this thesis, the review of customer experience management practices for building loyalty will be directed towards the passenger transport context. This initially requires a finer understanding of the loyalty concept.

Customer loyalty is said to consist of ‘three Rs’, which stand for retention, repeat business, and referrals (Heskett, Sasser Jr., and Schlesinger 1997). These relate respectively to keeping relationships with customers, to repeating business from the same or other services, and to raising new customers via endorsements. With respect to the promotion of urban passenger transport, this conception can arguably be interpreted to map customer loyalty into retaining passengers, increasing patronage of specific modes of transport and modal alternatives, and encouraging modal shift through positive testimonials. Nonetheless, despite the long-standing interest on customer experience across several service domains, and potential benefits for transport providers and society in general, the existing literature on travel experience is scarce.

This reflects particularly on public transport operation, where traditional customer satisfaction surveys remain a key instrument to identify service quality factors and prioritise improvements (Eboli and Mazzulla 2011; Eboli and Mazzulla 2012a). In fact, several previous studies focused on service quality and travel satisfaction rather than on the actual travel experience. Some have examined factors that influence satisfaction in the daily commute (Cantwell, Caulfield, and O’Mahony 2009; Ettema et al. 2012; Eriksson, Friman, and Gärling 2013) and public transport travel (Stuart, Mednick, and Bockman 2000). Others have identified factors to benchmark satisfaction across travel modes (Stradling, Anable, and Carreno 2007), to analyse the effect of satisfaction on modal choice (Habib, Kattan, and Islam 2011; Abou-Zeid and Ben-Akiva 2012; Abou-Zeid et al. 2012), and to assess the relationship between travel satisfaction and subjective well-being (Ettema et al. 2011). Others have emphasised the identification of key service quality factors influencing satisfaction, both in public transport (Friman 2004; Geetika and Nandan 2010; Eboli and Mazzulla 2011; Eboli and Mazzulla 2012b; de Oña et al. 2013) and air travel (Babbar and Koufteros 2008).

Some studies have mostly focussed on service quality factors themselves. These include understanding passenger perceptions of service quality (dell’Olio, Ibeas, and Cecin 2010; Cirillo, Eboli, and Mazzulla 2011) and desired levels of service (dell’Olio, Ibeas, and Cecin 2011). These also include evaluating the effect of service

quality on ridership (Paulley et al. 2006; Litman 2011) and on attitudes driving modal choice (Beirão and Sarsfield Cabral 2007), and analysing customer satisfaction surveys results to derive service quality assessment methods (Nathanail 2008; Tyrinopoulos and Aifadopoulou 2008; de Oña, de Oña, and Calvo 2012). Amongst the fewer studies that addressed travel experience, some focus on ‘experience-centric’ services such long-haul air travel, cruise trips, and leisure motorcycling experiences that are outside the scope of this thesis (Le Bel 2005; Zomerdijk and Voss 2010; Zomerdijk and Voss 2011).

Li (2003) is the earliest study found in the literature that deals with the quality of urban travel experience. His study presented an alternative to the vast body of existing research approaches for explaining travel behaviour based on conventional decision theory, and particularly using random utility models. Li (2003) has criticised those approaches on the grounds of the Nobel winning ‘prospect theory’ (Kahneman and Tversky 1979; Kahneman and Tversky 1984), for ignoring the important role of experience value and the hedonic motivations of decision-makers. The author drew from research in psychology to propose a model that considered the subjective perception of travel time, as opposed to clock time, to be fundamental for the passenger evaluation of travel experience. Based on a literature review, Li (2003) hypothesised that a range of factors within four categories influence perceived travel time and subsequently, from his logic, the urban travel experience.

The first is ‘commute characteristics’ and includes objective duration and number of journey stages; the second is ‘journey episodes’ that subdivide into the ride, wait, and access and transfer travel states; the third is ‘travel environment’ and includes comfort, and availability and quality of entertainment provided; and the fourth is ‘expectancy’ that subdivides into the expected travel time and service reliability (Li 2003, 56). Furthermore, the author recognised that other factors could have a moderating influence. Among these were ‘goal attainment’ representing the fulfilment of the journey objective, ‘economic values’ representing the monetary payoff, and ‘time urgency’ that may either be a personality trait, or a personality state stemming from time demands of the external environment (Li 2003, 57-58). The transport provider does not control some factors, notably the mediating factors,

questioning the appropriateness of studies that focus on service quality alone. However, it is debatable whether the fourth category on expectancy relates to experience or satisfaction because it already suggests a pondered assessment of prior expectation against measurable service delivery performance. Lastly, it is noted that the relationships of the model were tentative, being unclear if it has since been empirically evaluated.

Two years later, Anable and Gatersleben (2005) examined and compared the relative impact of various functional (often worded as ‘instrumental’) and affective journey factors on travel experience, by carrying out two studies relating to work travel and leisure day trips. Overall the authors found that affective factors are considered more important for leisure travel, but also play a role in the work travel experience. The studies were based on responses to travel survey questionnaires in respect to journeys by car and by public transport in both cases, plus walking and cycling in the work travel one. The choice of factors appears to be informed by their literature review, but the rationale behind it is vaguely described. Anable and Gatersleben (2005) considered flexibility, convenience, cost, predictability, environmental friendliness, and health as functional factors, and lack of stress, sense of control, sense of freedom, relaxation, and excitement as affective factors. One of the main conclusions drawn from this study is that the affective journey attributes were measurable and relevant constructs that should be considered alongside functional factors in order to understand travel experience and influence modal choice. It may however be argued that stress, excitement, and relaxation are not factors but affective responses, as differentiated in other studies.

Around the same time, Ory and Mokhtarian (2005) have published research on travel liking that studied what sorts of people liked to travel, and under what circumstances was travel enjoyed. Whilst the study did not identify a set of factors that are relevant for this thesis, its results suggest that travel liking is an intrinsic human characteristic, which inevitably bears a subjective influence on travel experience. Two years later, Stradling et al. (2007) surveyed Edinburgh residents living close to a designated ‘Quality Bus Corridor’ to understand what they disliked about the urban bus travel experience. The study was based on a self-completion

questionnaire containing items obtained from the literature and from discussions with bus users. The authors conducted factor analysis that revealed eight overarching factors of dislike, revealing not only functional but also social and affective concerns. Those were feeling unsafe, preference for walking or cycling, service provision issues, unwanted arousal, preference for car use, cost, disability and discomfort, and self-image. Whilst most of these are largely self-explanatory, unwanted arousal deserves further clarification. This factor contained questionnaire items relating to issues such as overcrowding, uncleanliness, harsh driving, and impoliteness of other passengers.

Meanwhile, two studies applied the seminal Circumplex Model of Affect (Russell 1980) to travel experience. The model represents affective concepts in an orthogonal spatial model of two dimensions, pleasure-displeasure in the horizontal axis, and degree of arousal in the vertical axis (Russell 1980). The first study examined affective appraisals of the daily commute experience across various modes of travel (Gatersleben and Uzzell 2007). It suggested that different travel modes elicited different affective responses. Notably, the use of private cars and public transport were generally found to be associated respectively with excess and a lack of arousal, hence the former to be stressful and the latter to be boring. Furthermore, attitudes towards public transport were found to be more negative than for other modes. This led to the insightful conclusion that strategies for promoting public transport use should consider the main sources of pleasure with its travel experience, reportedly reading, listening to music, interacting with people, and enjoying the scenery. The second study used confirmatory factor analysis to validate that public transport travel experience is multidimensional (Olsson et al. 2012). The authors found one cognitive dimension relating to judgement of service quality factors plus two affective dimensions named positive activation and positive deactivation. These range from enthusiasm to boredom and from relaxation to stress, and derive from combinations of the dimensions originally proposed by Russell (1980).

Carreira et al. (2013; 2014) authored the latest research on travel experience reviewed in this thesis. The authors initially carried out a qualitative study with bus passengers regarding mid-distance experience-centric and utilitarian journey types, and confirmed travel experience to depend both on service quality and aspects not

controlled by the transport provider (Carreira et al. 2013). Their starting point was to identify experience factors in the literature, which was scarce and required the addition of service quality and satisfaction studies. The authors built on those studies that focussed on the cognitive assessment of service aspects controlled by the transport provider, to study more comprehensively the multi-channel customer experience. After reviewing the literature, the analysis of interviews with passengers and other stakeholders yielded relevant experience factors. The authors made a clear distinction between ‘Experience Factors (EFs)’ and ‘Experience Components (ECs)’, shedding some light on the erratic use of vocabulary in previous studies. EFs were defined as service provision aspects that drive travel experience, and ECs as internal customer responses to the service provided. Their comprehensive work significantly improved the breadth and depth of existing research on travel experience.

From the first qualitative study (Carreira et al. 2013), eleven identified EFs were grouped into two main categories; the first, trip conditions, consists of cleanliness, comfort, easy accessibility, safety, scenery visibility, and waiting time; the second, supplementary services, consists of information provision, on-board entertainment, and off-board services; two additional factors, social environment and staff skills were identified outside those categories. ECs were divided into cognitive, sensorial, and emotional. The latter included positive emotions such as excitement and joy, and negative emotions such as annoyance, discontentment, nervousness, and fear.

A second quantitative study based on a questionnaire survey regarding mid-distance bus journeys (Carreira et al. 2014) refined the definition of EFs into a shorter list of seven using exploratory factor analysis, and evaluated their impacts on ECs and loyalty behaviours. It consisted of individual space, information provision, staff skills, social environment, vehicle maintenance, off-board facilities, and ticket line service. The results confirm the hypothesis that travel experience influences passenger loyalty, directly and indirectly, and emphasise its holistic nature involving both cognitive and emotional components, and factors not controlled by the transport provider. On-board aspects were found to be the core of experience, followed by information provision and staff skills. However, the cognitive experience component was said to involve not only quality assessment but also customer satisfaction (Carreira et al. 2014,

37). This appears to go against aforementioned distinctions between customer experience and customer satisfaction. These studies by Carreira et al. (2013; 2014) have also explicitly split the travel experience into its constituent before, during, and after moments. Although other studies (e.g. Stradling et al. 2007) had already considered factors outside the duration of the journey, such as the walking distance to access transport services and the quality of waiting facilities, their distinction between those moments provides further depth to the understanding of travel experience.

Table 2.1 Summary of travel experience factors in the literature

		Li (2003)	Anable and Gatersleben (2005)	Stradling et al. (2007)	Gatersleben and Uzzell (2007)	Carreira et al. (2013)	Carreira et al. (2014)
Mainly controlled	Travel environment						
	Comfort (includes crowding)	✓		✓ (d)		✓	✓ (k)
	Entertainment	✓			✓	✓	
	Cleanliness / maintenance	✓ (a)		✓ (d)		✓	✓
	Accessibility			✓ (e)		✓	
	Driver skills (public transport only)			✓ (d)		✓ (i)	✓ (i)
	Journey characteristics						
	Journey time	✓	✓ (b)	✓ (f)	✓ (g)	✓ (j)	
	Journey stages & episodes	✓					
	Cost		✓	✓			
Other	Information provision			✓ (f)		✓	✓
	Safety			✓	✓	✓	
	Off-board services					✓	✓
	Travel setting						
	Social environment			✓ (d)	✓	✓	✓
Mainly uncontrolled	Scenery				✓	✓	
	Subjective appraisal						
	Convenience / flexibility		✓	✓ (f)	✓		
	Sense of freedom & control		✓				
	Modal preference		✓ (c)	✓	✓ (h)		
Self-image			✓				
Notes:							
(a) Included in the comfort category; (b) alluded to in the predictability factor; (c) alluded to in the environment and health and fitness factors; (d) included in the unwanted arousal factor; (e) included in the disability and discomfort factor; (f) included in the service provision factor; (g) alluded to in terms of delays; (h) alluded to in terms of enjoyment; (i) included in the staff skills factor; (j) partially considered as part of the waiting time factor; (k) included in the individual space factor							

Table 2.1 summarises the factors that influence travel experience and were identified in studies reviewed in this section. An attempt was made to harmonise concepts and organise factors regarding the degree of control from the transport provider, as understood by the author of this thesis. The aim is ultimately to set out the factors that may be directly or indirectly influenced by the collaboration model proposed throughout the following chapters. That influence may be direct, on factors such as information provision and flexibility of specific travel alternatives, or indirect through the mitigation of factors via collaboratively generated information. The cognitive, sensorial, and affective responses, or experience components according to the terminology of Carreira et al. (2013; 2014), are not depicted in Table 2.1 but are constituent of travel experience.

Aside the marketing and transport literatures reviewed in this section, studies in the realms of sociology and psychology have focused specifically on the car driving experience. In spite of dissimilar theoretical framing, these studies also highlight the existence of an emotional response that undermines transport policy-making decisions based on conventional decision theory (Sheller 2004). Aware of the complexity behind modal choice, (Sheller 2004, 222) conceptualised car use as a culture that has “social, material and, above all, affective dimensions that are overlooked in current strategies to influence car-driving decisions”. Steg (2005) obtained empirical evidence aligned with that, and suggested that transport policy makers should not only focus on the functional incentives of car use but also on various social and affective motivations. This stresses the value of experience management in transport, which car manufacturer have known well how to exploit (Sheller 2004). Whilst flexibility, convenience, and freedom are the common arguments for private car use, one could argue that peace of mind, relaxation, and the ability to engage in other activities would offset those advantages of cars in favour of public transport. This section showed that it is not as simple; ultimately, travel behaviour is also about emotions and the holistic travel experience.

User experience

Prior to closing this section on customer experience, it is worth clarifying how it differs and relates to the also popular concept of User Experience (UX). While these concepts are sometimes used interchangeably, the first obvious difference is that the customer may or may not be the user. For example, a firm may be the customer of an office supplies merchant, but the actual users are its employees. A second difference is that a user, such as a visitor to an e-commerce website, may not be nor even become a customer. Yet, the third and most important difference that is not evident from the outset, is that UX has mostly been used in the context of Human-Computer Interaction (HCI). An early definition related UX to “all the aspects of how people use an interactive product” (Alben 1996). Consequently, UX is intimately associated with the notion of usability as an ease-of-use assessment of a user interface (Nielsen 1993).

Since UX relates specifically to the use of a product or service, it may be considered a subset of customer experience. According to Meyer and Schwager (2007, 1) the scope of customer experience “encompasses every aspect of a company’s offering – the quality of customer care, of course, but also advertising, packaging, product and service features, ease of use, and reliability”. Hence customer experience includes UX but is broader. Nonetheless, the notion of UX is also relevant for this thesis in regards to the iterative development of the prototype service *Journata*, which followed a user-centred interaction design approach (Preece, Rogers, and Sharp 2002). This approach involves users throughout the design and development processes, and in the evaluation of usability.

2.2. Value co-creation

The customer experience construct introduced in the previous section puts interaction under the spotlight. As a result, it fosters the active involvement of customers in shaping their own experiences. Prahalad and Ramaswamy (2000) popularised the term ‘co-creation’ (initially spelt as ‘cocreation’), used to describe an observed evolution of the role of customers from a traditionally passive audience to

active players creating and competing for value. According to their argument, customers became a source of competence that firms could leverage to achieve competitive advantages. The phenomenon was largely ascribed to the Internet, which let customers engage themselves in dialogue with product manufacturers and service providers, and even in discussions with fellow consumers aside these firms and their sphere of control. Four years later, Prahalad and Ramaswamy (2004) matured their argument in another influential paper titled ‘Co-creation experiences: The next practice in value creation’ that is considered to mark the official debut of the expression ‘value co-creation’.

As the title suggests, Prahalad and Ramaswamy (2004) reinforced the notion that the process of value creation was shifting towards personalised customer experiences, and clarified the differentiating features of co-creation in relation to previously established practices of customer involvement. They have argued that value co-creation was neither about raising productivity by transferring activities to customers as in self-services or self-checkouts, nor about involving a subset of customers to co-design products or services, because in either case *“the firm is still in charge of the overall orchestration of the experience”* (Prahalad and Ramaswamy 2004, 8, emphasis in the original). Instead, it was about emerging *“communities of connected, informed, empowered, and active consumers”* (Prahalad and Ramaswamy 2004, 8) challenging the traditional firm-centric perspective by demanding personalised interactions. Hence, co-creation was claimed to focus on the *“consumer-company interaction as the locus of value creation”* (Prahalad and Ramaswamy 2004, 10, emphasis in the original). Ebay and Amazon were mentioned as leading examples, fitting into the value co-creation philosophy by facilitating the personalisation of experiences, involving communities, and easing dialogue. Despite the authors having extensively used the term value, which meaning was claimed to be changing, it was loosely defined in the paper.

Around the same time, Vargo and Lusch (2004) published their influential notion of service-dominant (S-D) logic based on eight foundational premises. The sixth one called ‘FP₆: The Customer Is Always a Coproducer’ dealt with the process of value creation (Vargo and Lusch 2004, 10). It described an opposition between the traditional goods-dominant (G-D) logic, in which production is notionally separated

from consumption to maximise its efficiency in terms of output, and the new S-D logic. The latter contends that “the consumer is always involved in the production of value. Even with tangible goods, production does not end with the manufacturing process; production is an intermediary process” (Vargo and Lusch 2004, 11). Their new logic endorsed that customers were also able to produce value continuously through consumption. These authors have since updated the foundations of the S-D logic (Vargo and Lusch 2006; Vargo and Lusch 2008) and rephrased the sixth premise as the “customer is always a co-creator of value” (Vargo and Lusch 2006, 44). This was a response to a critique (Prahalad 2004) that they dismissed on the basis of having been unintentionally stuck in G-D logic vocabulary (Vargo and Lusch 2008, 7).

In that critique, Prahalad (2004) attempted to clarify the meaning of co-production by describing conventional customer engagement approaches that he considered to be informed by a traditional firm-centric perspective. Chathoth et al. (2013) attempted to harmonise concepts and clarify the distinction between co-production and co-creation further based on the opposition between G-D logic and S-D logic. They associated co-production with the traditional firm-centric perspective of customer involvement articulated by Prahalad and Ramaswamy (2004), and argued that it is informed by the G-D logic. Co-production is seen as production-oriented, to regard customers as passive, and to deem service as secondary. The authors find co-creation to be informed instead by the S-D logic. Co-creation emphasises a joint, collaborative, interactive, and continued effort, in which engagement in dialogue and learning from customers are the underpinnings of mutually beneficial relationships. Whilst this contribution goes a long way to unify concepts, the actual meaning of value, which so often accompanies co-creation, remains ambiguous in this review.

The notion of value

As Gupta and Lehmann (2005) have noted, two sides of value exist, value for the customer and value for the firm. Their logic mapped the first part to the investment, and the second part to the financial return on that investment. Whilst the two are interrelated, the aforementioned literature on value co-creation implicitly

focuses on value for the customer, presumably under the assumption that it will also be a driver of value for the firm. Woodall (2003) argued that value for the customer became a hot business topic, but came to find that the expression had been used with a range of different meanings, which were grouped into five main notions. These correspond to ‘intrinsic value’ determined in terms of product attributes (Fronzini 1971); ‘value-in-exchange’ as a purely economic interpretation; ‘value-in-use’ as the Aristotelian notion of benefits derived from consumption experience (e.g. Holbrook 1999); ‘utilitarian choice’ as an evaluation of benefits against sacrifices (e.g. Heskett, Sasser Jr., and Schlesinger 1997, 40; Kotler 2000, 6); and a blend of the ‘value-in-exchange’ and ‘intrinsic value’ notions, relating to the deviation from a pre-established benchmark price (Woodall 2003).

Whilst finding a definition that is both precise and unanimous is impossible, value for customers in the broadest sense means that customers become ‘better off’ (Kotler 2000, 7; Grönroos 2008, 303) or increase their ‘well-being’ (Vargo, Maglio, and Akaka 2008, 150) from service. The S-D logic hinges upon two of the notions of value identified by Woodall (2003). Vargo and Lusch (2004) argued that the meaning of value shifts from value-in-exchange defined by the producer in the G-D logic, to value-in-use determined by the customer in the S-D logic. In the latter, a customer who accepts a value proposition from a provider trusts that value-in-use will meet or exceed value-in-exchange (Lusch, Vargo, and O’Brien 2007, 13).

Who are the co-creators?

The notion of value in the S-D logic prompted critique of the premise that the customer is always a co-creator of value. Grönroos (2011) has argued that if value creation is characterised by value-in-use, logically the customer is always a value creator, not a co-creator. And because value-in-use implies that it is realised during consumption, the development, design, manufacturing, and delivery processes of the provider merely facilitate value creation (Grönroos 2008). As a result, Grönroos (2011, 289, emphasis in the original) specified that “*the customer creates value, and the firm facilitates value creation*”. The author still maintains that co-creation exists, but only through the interaction between customer and the firm. Customers are in charge of

value creation because it is them who determine whether value emerges or not, but the direct interaction lends an opportunity for the firm to engage with them, influence their value creation processes, and rise from a value facilitator into a value co-creator (Grönroos 2011; Grönroos and Voima 2013).

This view that customers are always value creators falls short of considering that value may also be co-created amongst customers themselves. And indeed, the literature on S-D logic generally does not give attention to the interaction between customers (Rihova et al. 2013). However, a number of recent studies have been focussing on the understanding of Customer-to-Customer (C2C) interactions (Libai et al. 2010). Nicholls (2010) noted that C2C interactions are a common phenomenon across a range of service industries, namely passenger transport. Rihova et al. (2013) have proposed a conceptual framework to guide service providers in facilitating C2C value co-creation in socially dense, shared and collective consumption contexts. The authors argue that service providers can benefit from creating environments that soften social barriers and stimulate communing practices among customers, who co-create value by generating a sense of unity, shared identity, and ultimately of belonging. Furthermore, when those communing practices exceed the realms of service provision, encouraging reciprocity and solidarity between customers, value emerges as social capital (Ostrom 2000). Communing practices raise collective awareness of service environment, which is a critical element to enable collaborative efforts between customers (Pitt et al. 2013).

Whilst the framework that Rihova et al. (2013) proposed requires further validation, a previous empirical study had already demonstrated that C2C interactions have an effect on customer experience and are positively associated with loyalty towards the firm (Moore, Moore, and Capella 2005). Moreover, case study based research has also shown how a firm can benefit from fostering a sense of community and easing communication amongst customers, as a catalyst of experience co-creation (Rowley, Kupiec-Teahan, and Leeming 2007). In passenger transport specifically, Harris and Baron (2004) conducted an ethnographic study that shows the interaction between railway service users led to a decrease in travel anxiety and reduced dissatisfaction with the service.

Joining the aforementioned perspectives, it can be concluded that not only the firm can co-create value by interacting with customers, customers can also co-create value by interacting themselves. Either way, value co-creation invariably emerges from interaction. Amid the diversity of examples found in the literature, the involvement of customers in beta software testing (e.g. Banks and Humphreys 2008; Roser et al. 2009) is a prime example of value co-creation emerging from the interaction between customers and the firm, whereas collaborative content creation using Web 2.0 technologies is arguably the most representative and widespread form of value co-creation centred in interactions among customers (or users more generally). Web 2.0 relates to “*a set of economic, social, and technology trends that collectively form the basis for the next generation of the Internet – a more mature, distinctive medium characterized by user participation, openness, and network effects*” (Musser and O’Reilly 2007, 4, emphasis in the original). It is geared towards value co-creation by harnessing collective intelligence (O’Reilly 2005), which is defined as “the capacity of human collectives to engage in intellectual cooperation in order to create, innovate and invent” (Lévy 2010, 71). In Web 2.0, the role of providers is one of facilitators of the creation of value by delivering platforms that unite individual contributions, whilst real value is co-created by, and for, its users (O’Reilly and Battelle 2009).

In a study of the tourism industry, Neuhofer, Buhalis, and Ladkin (2012) brought together notions of customer experience, value co-creation, and customer-to-firm and C2C interactions mediated by Web 2.0 technologies. The authors highlighted that Web 2.0 technologies are increasingly mediating experiences, helping customers getting involved in co-creation to enhance those experiences.

Co-creation in the public sector

Urban passenger transport services around the world have traditionally been owned or subsidised, and mostly remain regulated, by the public sector (Pucher and Markstedt 1983; Erdmenger and Führ 2005; Sreenivas and Sant 2008). Hence the provision of urban passenger transport is largely considered to be a public service that caters for the mobility needs of populations. Despite a recent privatisation trend across various countries, through franchises and concessions, the public sector retains

control by defining service terms and awarding contracts. Therefore, value co-creation examples in the public sector, and specifically in passenger transport, are scarce but therefore of particular interest to this thesis. It has been observed that the public sector is lagging behind the adoption of value co-creation practices and tools due to a lack of customer orientation from public officials (Cassia and Magno 2009), and for being generally characterised as averse to change (Alves 2012). This is hardly surprising if taking into account that public services are often not as pressured to focus on the customer as private services are in order to be successful in highly competitive environments; some public services are even natural monopolies (e.g. Evans 1991). Nevertheless, public services in many developed countries are currently facing tighter spending restrictions whilst pushed to maintain levels of service (Pollitt 2010; Evans, Hills, and Orme 2012). Alves (2012) suggests that co-creation has potential to drive innovation in public services, drawing on the capacities and knowledge of citizens to overcome spending restrictions.

Alford (2009a; 2009b) identified several activities in which citizens play a role in the co-production of public services. These include health services that require patients to behave in specific ways such as taking medicines or resting properly to ease their conditions; waste recycling services that require households to separate their rubbish in categories (Alford 2009a); job centres that require employment seekers to prepare for recruitment processes; and projects for the reduction of anti-social behaviour in public housing estates that require tenants to adopt a sense of ownership over those spaces (Alford 2009b). However, as reviewed earlier in this section, co-production largely relates to the traditional G-D logic. The recycling and surveillance examples are largely focused on the transfer of labour, and the public service provider still orchestrates the process in all four examples without incorporating the competence of customer citizens to generate value interactively.

Therefore, Wise, Paton, and Gegenhuber (2012) are possibly the first to have published a review of value co-creation initiatives in the public sector. It focuses on initiatives in the US and in Europe to leverage collaborative Internet media to engage broader communities in the co-creation of value, and promote collective intelligence in the spirit of Web 2.0. The authors reviewed several US initiatives

within the open government effort by The Administration of President Barack Obama, and emphasised two cases. One is a collaborative web-based platform linked to an Idea Management system to leverage the ideas of individuals working with veterans in order to develop innovative solutions to help them. The other called innovation.ed.gov (relocated to the challenge.gov platform at the time of writing), was aimed at sourcing innovations to improve the standard and quantity of students in education system, allows stakeholders to contribute with ideas and opinions to shape future education initiatives.

The authors also reviewed two European cases targeting contentious topics. One was 'Your Country Your Call' (yourcountryyourcall.com), which was introduced in Ireland following the recent financial crisis, to leverage social capital in developing ideas for reconstructing the economy. It not only allowed individuals to contribute with ideas but also to vote and decide on the best one, which would be adopted by the government. The other is a portal called 'Beteiligungshaushalt Freiburg', (www.beteiligungshaushalt-freiburg.de) that allows the citizens of Freiburg to participate in the government budget decisions, by deciding the most important issues to be addressed, and how much should be spent on them (Wise, Paton, and Gegenhuber 2012). All four cases are interesting examples of how public services can facilitate collective action as a driver of value co-creation with their customers, i.e. the public they serve or represent. The focus is indeed to integrate the competence of customers through provider-to-customer, and also C2C, interactions.

Co-creation in passenger transport

Gebauer, Johnson, and Enquist (2010) are amongst the few authors who attempted to analyse value co-creation practices in passenger transport, with a qualitative case study of the Swiss Federal Railway operator SBB. Their study was based on content analysis of communications, reports, and publications from SBB and its railway association, and on the analysis of interviews with its senior executives. The authors found that SBB had historically been driven by G-D logic but had been evolving towards becoming a value co-creator according to the S-D logic. From studying that shift, the authors demonstrated the strategic role of value

co-creation in public transport, both in terms of performance improvements and support of environmental sustainability. Gebauer, Johnson, and Enquist (2010, 527) conclude that the pursuit of these benefits “should not only seek to increase transport capacity, but should also encourage value co-creation by engaging customers in marketing activities, offering self-servicing opportunities, creating customer experiences, solving customer problems, and co-designing services in collaboration with customers”. In another study, the authors research these opportunities across four case studies involving the railway operators DB from Germany, SJ from Sweden, and SBB, plus the multi-modal transport agency ZVV from Zurich (Johnson, Gebauer, and Enquist 2010).

In spite of their significant contribution to the literature, both in terms of theoretical and managerial implications, it appears that the value co-creation activities endorsed in these studies are at odds with the aforementioned notion introduced by Prahalad and Ramaswamy (2004). These activities seem aligned with examples that Prahalad and Ramaswamy (2004) related to co-production, which were new to public transport but short of a true S-D logic. Gebauer, Johnson, and Enquist (2010) may have interpreted some approaches described in the invited commentary to the S-D logic by Prahalad (2004) as co-creation examples, while in fact they were supposedly meant as examples of co-production.

Alexander and Jaakkola (2011) also examined value co-creation in public transport using the ‘Adopt a Station’ scheme in Scotland as a case study. The scheme allows local communities to occupy vacant facilities within stations for providing services such as bookshops and cafes, or carrying out improvements like gardening. It provides a complex setting whereby value co-creation occurs in interactions involving businesses, consumers, and the community. The authors found that the success of such scheme hinged upon each actor receiving adequate value-in-use to stimulate engagement. Each actor experienced value-in-use differently. For example, adopters benefit from the feeling of owning an important building and recognition from the community, passengers benefit from an improved station environment, and the rail operator benefits from increased safety and number of passenger journeys. The results led the author to conclude that firms that are prepared to engage with a range

of actors beyond a traditional provider-customer relationship may derive benefits from opening up access to their value network. This appears to apply particularly well to public transport services that generally treat customers as passive (Gebauer, Johnson, and Enquist 2010; Johnson, Gebauer, and Enquist 2010). Jaakkola and Alexander (2014) have later built on this case study and suggested that firms should consider the involvement of customers further and let them operate as proactive collaborators.

The study above (Alexander and Jaakkola 2011) already suggests that the co-creation in passenger transport may exceed the relationship between firm and customer. In that case, passengers also received additional value-in-use from the station improvements, and so does the society in general with the improvement of travel choices that reduce the socially and environmentally adverse impacts of urban passenger transport.

This review has already focused on three important recommendations from the literature that apply to passenger transport providers. The first is to *focus on staging holistic travel experiences to increase passenger loyalty*. The second is to *engage customers in co-production activities for performance improvements*. The third is to *encourage customer collaboration as a form of value co-creation*, which is the focus of the next section.

2.3. ICT-based user collaboration

The collaborative creation of content using Web 2.0 technologies has been argued to be an important form of value co-creation. It is particularly relevant for this thesis, which is concerned with a model of collaboration that leverages recent advances in ICT. Since an essential aspect of Web 2.0 is harnessing collective intelligence, its technologies are characterised by the active participation of users, implicitly or explicitly (Musser and O'Reilly 2007). Implicit user participation consists of algorithmic techniques that continuously improve applications as people use them, such as those behind the recommendation engine from Amazon or Google translate. Explicit user participation generally relates to content creation and other forms of collaboration. It is associated with the notion of 'crowdsourcing', a term coined by

Jeff Howe (2006) to describe a process that “revolves around large groups of people or a community handling tasks that have traditionally been associated with a specialist or small group of experts” (Greengard 2011, 20).

Geiger, Rosemann, and Fielt (2011) identified four types of crowdsourcing, which were named crowd processing, crowd rating, crowd solving, and crowd creation. Crowd processing combines individual contributions toward solving a problem, for example, the identification of objects from satellite images with Tomnod. Crowd rating is the collective assessment of a specific item, for example, media reviews on Amazon and hospitality reviews on TripAdvisor.com, user reputation on eBay and Uber, and film scores on IMDb. Crowd solving sources different individual contributions to find the best solution for a problem, for example, solving medical cases with CrowdMed. Lastly, crowd creation relates to all kinds of user-generated content, for example, articles on Wikipedia, videos on YouTube, maps on OpenStreetMap, and posts on social networking services such as Facebook, LinkedIn, or Twitter (Geiger, Rosemann, and Fielt 2011). The users involved co-create value in several ways through explicit collaboration, including problem-solving know-how and decision-making support. That collaboration allows for various kinds of interpersonal affinity to materialise into actual ties between users. Friendship and professional relationships are tangible social structures replicated in ICT-based social networking services. Yet affinity may also be a common interest revealed in a crowd rating system, or knowledge about a specific activity shared in a crowd solving system (Boyd and Ellison 2008). Whatever affinity consists of, the resultant ties between users have enabled unprecedented collaboration-based value co-creation opportunities by harnessing collective intelligence (O’Reilly and Battelle 2009).

Upon describing the motivation for this research (Section 1.1), it was argued that transport systems of today present opportunities for user collaboration that largely remain to be explored. However, drawing greater value from interpersonal affinity in a transport setting most likely has to account for the inherent spatiotemporal dynamics. Therefore, this review shall focus next on the crowdsourcing of spatiotemporal data.

Crowdsourcing spatiotemporal data

Greengard (2011) highlighted the organisation Ushahidi as a leader in spatiotemporal crowdsourcing. Ushahidi created a platform based on the simple idea of volunteers reporting events as they unfold through a web browser or mobile phone, which are then mapped by time and location. Despite its simplicity, the platform had already been used, for example, to provide disaster relief following the recent earthquakes in Haiti and Japan, to identify medical shortages in the Philippines, and to monitor local elections in Afghanistan, India, and Mexico (Gao, Barbier, and Goolsby 2011; Greengard 2011). The organisation kept growing to date with an ethos of empowering ordinary citizens to generate local information, which is crowdsourced for the benefit of society. Crowdsourcing of spatiotemporal data has also been implemented for scientific purposes. The project eBird is an example, having been developed to collect data on bird occurrence, referenced to a specific time and location. Bird watchers around the world provide the information, which is stored in a publicly accessible database (Sullivan et al. 2009). Scientists and environmental conservationists use that data to develop a better understanding of avian biological patterns, movements, and distributions towards the development of conservation strategies (e.g. Fink, Damoulas, and Dave 2013; Fink et al. 2014).

Wolfson and Xu (2010) have described potential uses and research issues of spatiotemporal data in urban passenger transport. The authors addressed various topics, including the potential use of ICT-based social networking services to crowdsource spatiotemporal data of interest to travellers, traffic managers, and transport planners, such as real-time traffic congestion and availability of parking spaces, and ride sharing opportunities. Three research issues were identified as the need for providing incentives for users to participate in information sharing, the need for validating mechanisms to support the reliability of data provided, and the need to maintain user privacy.

Crowdsourcing in passenger transport

Tiramisu (Steinfeld et al. 2011; Steinfeld et al. 2013), Moovit, Trafi (O’Hear 2015), and Waze (Olson 2014), are interesting examples of mobile applications that leverage crowdsourcing of spatiotemporal data from public transport users, in the first three cases, and from private car users in the latter. Tiramisu, Moovit, and Trafi all combine journey planning features with the ability to crowdsource route and vehicle-specific information provided by passengers, including crowding levels and subjective service quality appraisals. Waze collects traffic-related information by crowdsourcing reports of traffic incidents and other hazards provided actively by users, together with data captured by sensors in their mobile devices while the application is running (this is often referred to as ‘crowdsensing’). The application combines both types of information for helping users make informed routing decisions based on real-time traffic conditions. Xuan, Sengupta, and Fallah (2010) claimed that Waze successfully tackled the aforementioned issue of incentivising users to participate. It features a cumulative user scoreboard that gives a gaming feel to the service, to the extent that it can be considered to be a ‘serious game’ (Susi, Johannesson, and Backlund 2007). Waze user reports have also fed research on the relationship between incident hotspots and level of police coverage to reduce accident prevention inefficiencies and improve road safety (Fire et al. 2012).

Whilst the above applications deal with crowdsourcing of spatiotemporal data, they do so without linking users in a social network fashion. On the other hand, public transport operators are ever more present in social networking services such as Twitter and Facebook (Austin 2010), which are poorly suited to cater for the spatial element of real-time transport-related information. A recent study identified five reasons for that involvement, which relate to providing timely updates, to distributing public information, to engaging with citizens, to recognising employees, and to a form of entertainment (Bregman 2012). Another study then suggested that transport operators can leverage social networking services as a channel for maintaining longer-term relationships with customers and providing real-time travel disruption updates (Gault et al. 2014). Hence, these studies suggest that transport operators have mainly been using social networking services through a top-down approach to reach

out to customers, without exploiting a bottom-up crowd creation approach to harness collective intelligence that is scattered across their transport systems.

Earlier in this section it has been argued that user collaboration through crowdsourcing materialises several types of interpersonal affinity in actual ties between users. But if social networking services connect users based on tangible interpersonal relationships, can they generate crowd creation opportunities between travellers, and particularly in real-time? Some recent studies have focussed on the relationship between human mobility and social ties through the evaluation of mobile phone network data, which contains trajectory and communication records (Wang et al. 2011; Phithakkitnukoon, Smoreda, and Olivier 2012; Toole et al. 2015). These studies assume that the frequency and reciprocity of caller connections are an indicator of de facto social ties, even to a greater extent than ICT-based social networking services where people have hundreds of distant contacts (Phithakkitnukoon, Smoreda, and Olivier 2012). Whilst the findings are unanimous indicating that human mobility and social ties are correlated, given the coarse location accuracy of mobile phone trajectories and the temporal granularity of communication records, they mainly identify the impact of friendship on the likelihood of visiting a specific place instead of providing a finer assessment of the equivalence between daily travel patterns.

Location-based social networking services, in which users check-in in specific locations (e.g. Foursquare) have also been used to study the relationship between human mobility and social ties. Whilst user check-ins are generally less frequent than caller connections, they have the advantage of providing accurate locations. From the analysis of such data, Cho, Myers, and Leskovec (2011) argued that social relationships somewhat influence long-distance travel, but short-range travel is not affected by them. Hence, as one may expect, social networking services may not be well suited for stimulating crowd creation opportunities between urban travellers in real-time. Aguiléra, Guillot, and Rallet (2012, 669) have argued that “the possibilities of interaction offered by the combination of the transportation system and ICTs enables new configurations of social networks that in turn may transform travel

demand”. Those new configurations arguably imply the need for reimagining affinity to account for the spatiotemporally dynamics of transport.

Recent studies on harvesting transport-related information using text-mining techniques attempted to overcome the aforementioned inadequacies of social networking services and harness collective intelligence from a bottom-up crowd creation perspective. Carvalho, Sarmiento, and Rossetti (2010) have proposed a text classification approach to perform real-time sensing of traffic-related information from posts on Twitter. However, the authors estimated those to account for less than 0.05% of all posts in their sample, requiring a very large and continuous data stream for the method to be worthwhile. Gal-Tzur, Grant-Muller, Minkov, et al. (2014) and Gal-Tzur, Grant-Muller, Kuflik, et al. (2014) explored the combination of two sides of engagement via social networking services: the potential of a top-down approach from transport providers, and the potential value of bottom-up user generated information for policy development. From one study, the authors concluded that the involvement in social networking services can initiate public discussion, and mining that data afterwards may reveal public needs and perspectives on a range of transport policy and strategy issues (Gal-Tzur, Grant-Muller, Minkov, et al. 2014). In the other study, focused specifically on extracting information from transport-related posts, the authors developed a hierarchical approach to categorise data and applied it to posts on Twitter, and concluded that social networking services are a valuable source of information for developing and delivering transport policy goals (Gal-Tzur, Grant-Muller, Kuflik, et al. 2014). Their bottom-up approach is therefore geared towards supporting policy making, unlike Carvalho, Sarmiento, and Rossetti (2010), whose goal was to sense information in real-time for the immediate benefit of travellers.

A few studies have also focused on the integration of transport-specific social networking features with other platforms. Alves, Chaves, and Steinmacher (2011) have proposed an application for Facebook users to create, collaboratively edit, watch, and share journey routes. The aim has been to promote collaboration among urban public transport users and harness their collective intelligence for the identification of interesting places and pleasant routes. This proposal appears to be best suited for leisure travel than for the daily commute. Lüke et al. (2009) have

proposed the architecture of an integrated mobile service that combines personal preferences obtained through social networking services with journey planning, smart ticketing, and entertainment features. The authors have described how an integrated solution may fulfil some travel requirements regarding flexibility and convenience of public transport journeys.

Apart from social networking services, crowdsourcing has also been argued to potentially improve journey planning applications (Cotfas, Croicu, and Cotfas 2009), and to stimulate the involvement of larger groups of stakeholders in transport planning processes via technology-mediated forms of participation (Misra et al. 2014). Whilst this section has tried to provide an overview of crowdsourcing processes in passenger transport, the information needs that they fulfil must also be understood. This review will move on to passenger information requirements and how they can be met through ICT-based collaboration.

Passenger information requirements

Adler and Blue (1998) had an early vision of future traveller information systems. Years before the Web 2.0 revolution, the authors forecast the new generation of those systems to be capable of personalising assistance by learning and adapting to traveller preferences and behaviours. And indeed, Chorus, Molin, and Van Wee (2006) later suggested that travellers wish to have access to travel information beyond 'basic' travel times and costs, to include softer characteristics to match their preferences, such as levels of convenience, comfort, and privacy associated with their travel alternatives. Chorus et al. (2007) maintained that in addition to the 'basic', other types of information to ease travel were needed, including early warning functions to inform about travel disruptions, and personalised information to account for preferences on modes, routes, and departure times. Caulfield and O'Mahony (2007) also highlighted a general desire from public transport passengers to have access to real-time information, particularly news on travel disruptions. Fonzone (2015, 2) has gone further to conclude from a study of bus passengers that real-time information "can have remarkable consequences on travel choices and eventually on system performances". Ferris, Watkins, and Borning

(2010a; 2010b; 2010c) evaluated the impact of providing real-time arrival information and service alerts to public transport passengers through OneBusAway and found various positive outcomes in terms of increased satisfaction, patronage, and feelings of safety, and reductions in waiting times.

In one of their studies, Ferris, Watkins, and Borning (2010b) did note that crowdsourcing might be a good alternative for providing service alerts on disruptions, either when transport providers are not able to provide that information, or in order to increase its timeliness. And indeed, an urban passenger transport system is typically populated by a large number of users spread across it at any time of operation. Each user has limited visibility of the system as a whole, but a single, location-specific perspective of it may be harnessed to form collective intelligence. That certainly includes, but may not be limited to, the identification of service disruptions. Filippi, Fusco, and Nanni (2013) showed that public transport service improvements might be achieved by empowering users in order to foster bottom-up development in a setting of decreasing public funding, namely through ICT-based crowdsourcing. The authors highlighted the potential of mobile devices as two-way communication media, allowing users to share first-hand information on different aspects of service performance, reducing delays in supplying that information to others. Hence, their research advocated the development of applications to distribute advanced mobility information generated from past user behaviour, and collective intelligence from exchanges between users. The information would be personalised and include comparison of costs, time, reliability, and impacts of travel alternatives.

It has been argued that such applications that harness collective intelligence from collaborative exchanges of information amongst transport users can mitigate passenger information requirements, in terms of service delays, vehicle breakdowns, indication of best and worst route alternatives in real-time, assessments of service and route quality, and assessments of safety across geographical areas (Chaves, Steinmacher, and Vieira 2011). At the same time, increasing the availability of richer and personalised information has potential to allow travellers to adjust their travel choices to their preferences (Costa et al. 2012). Some studies have therefore aimed to evaluate the value of smarter information on travel choice from a variety of

perspectives, and found positive impacts that include improved perceived utility of alternatives (Chorus, Walker, and Ben-Akiva 2010), reduced uncertainty with scheduling (Line, Jain, and Lyons 2011), and improved efficiency of choices (Iryo, Yamabe, and Asakura 2012). Other studies have implicitly addressed passenger information requirements by associating the provision of real-time information via urban screens (Foth and Schroeter 2010) and mobile devices (Windmiller, Hennessy, and Watkins 2014) with the improvement of travel experience and satisfaction.

Table 2.2 Overview of passenger information requirements identified in the literature

	Adler and Blue (1998)	Chorus, Molin, and Van Wee (2006)	Chorus et al. (2007)	Caulfield and O'Mahony (2007)	Chorus, Walker, and Ben-Akiva (2010)	Ferris, Watkins, and Borning (2010b)	Chaves, Steinmacher, and Vieira (2011)	Filippi, Fusco, and Nanni (2013)	Fonzzone (2015)
Basic information									
Travel routes /route planning	✓ (a)	✓ (b)	✓ (c)	✓	✓ (d)	✓	✓ (e)	✓ (c)	✓
Travel times		✓	✓		✓	✓	✓ (f)	✓	✓
Travel cost		✓			✓			✓	
Advanced information									
Personalised information (*)	✓		✓					✓	
Travel assistance (**)	✓ (a)		✓						
Comfort		✓					✓ (g)		
Disruption warnings / alerts			✓	✓	✓	✓			
Location-specific information			✓						
Real-time vehicle location (***)				✓		✓	✓ (f)		✓ (i)
Safety							✓ (h)		
Reliability							✓ (g)	✓	
Notes: (*) Personalised information to reflect personal preferences on modes, routes, departure times, etc. (**) Travel assistance taking account of real-time travel conditions. (***) Real-time vehicle location is specific to public transport modes. (a) Included in route guidance based on real-time conditions; (b) mentioned as convenience; (c) included in multi-modal information; (d) mentioned as alternative assessment and generation; (e) alluded to as best or worse routes; (f) alluded to in terms of delays and vehicle problems; (g) alluded to as best or worse bus routes and companies; (h) included in terms of danger areas; (i) included in real-time arrival information.									

Table 2.2 is a summary of passenger information requirements that are mentioned in many of the aforementioned studies. An attempt was made to harmonise and aggregate information topics across studies, as understood by the author of this thesis.

The aim is to scope the potential for mitigation of these requirements through ICT-based user collaboration.

Evaluation of user collaboration potential

This section has reviewed so far several mechanisms of collaboration and their potential to mitigate information needs of travellers. But prior to moving on to the next section, some attention should be given to the potential for harnessing collective intelligence through user collaboration in urban passenger transport. This can be illustrated in light of graph theory. Random graphs consist of a set of N nodes and E edges, in which a pair of nodes has a probability p of being connected by an edge. Let affinity in urban passenger transport be defined in the context of this thesis as the degree of equivalence between the travel patterns of users in the system (the rationale and details of this definition will be addressed in Chapter 3). Let the users of a transport system at a specific instant be the nodes, and let affinity between them be the edges of a random graph.

Paul Erdős and Alfréd Rényi (1960) found the interesting property that when the average degree $\langle k \rangle$ (average number of edges connected to a node) reaches 1, it is almost certain that a great cluster will appear in the graph. Barabási (2003) wittily illustrates this property with an example of a gossip spreading at a party. If each person knows at least one other guest, then everybody will get to know it rapidly. Unlike gossip at a party, it may not be expected that many users of urban passenger transport willingly engage in some sort of ICT-based collaboration. Yet, as long as some do, and the average degree $\langle k \rangle$ of the resulting graph reaches 1, any information shared amongst them is likely to spread and produce collective intelligence. Section 5.6 of this thesis will demonstrate that the likelihood of this is high, even when conservative assumptions are applied.

At a greater level of detail, percolation theory suggests that a graph will likely collapse when a specific fraction of nodes or edges fail (Stauffer and Aharony 2003). This raises the question of how many nodes (N) or edges (E) are necessary for maintaining a cohesive graph. Generally, if a small fraction of nodes is removed, only small

clusters break from the main cluster, known as giant component G . However, at a critical threshold p_c (the percolation threshold), the giant component G gets fragmented into small clusters. For large random graphs, it is suggested that the threshold p_c is the multiplicative inverse of the average degree $\langle k \rangle$ (Newman 2010). Graphs with a large $\langle k \rangle$ can withstand the loss of many of its nodes while keeping their main connectivity. A graph loses its connectivity when the number of failed nodes reaches $\mathcal{N} - \lfloor \mathcal{N} * p_c \rfloor$. Hence, in order to assess the potential for harnessing collective intelligence, the expected minimum number of users engaging in collaboration should exceed the threshold value $\lfloor \mathcal{N} * p_c \rfloor + 1$ (Daqing Li et al. 2015).

2.4. Travel pattern extraction and estimation

The previous sections of this literature review provide the theoretical background for developing the conceptual foundations of the user collaboration model. The present and next sections focus instead on existing data sources and methodologies for extracting and estimating travel patterns (Section 2.4), and to measure collaboration potential (Section 2.5) to support the development of algorithmic tools for the model. In the context of this thesis, collaboration potential will be estimated in terms of the spatiotemporal equivalence of travel patterns (Chapter 5). Therefore, capturing those patterns is an enabling step for implementing the model, and that is the focus of this section.

Today, a plethora of sensors are used throughout urban passenger transport systems, generating large volumes of data that can be used to reveal travel patterns. Urban transport users carry mobile devices such as smartphones, tablets, and In-Car Entertainment (ICE) units, which often have embedded GPS receivers and accelerometers that track and sense their movements, and leave traces by getting signal from radio antennas and joining wireless local area networks. Besides that, sensors for collecting fares, locating vehicles, and counting passengers, increasingly populate transport system infrastructures (Rosado 2014). Whilst these are designed and implemented for other purposes, the wealth of data they generate, often via interaction with other sensors carried by users themselves, such as Radio-Frequency

Identification (RFID) tags, is also useful for tracking and predicting travel behaviour (Zhao, Rahbee, and Wilson 2007; Wang, Attanucci, and Wilson 2011). And so is the use of ICT-based journey planning and ticketing applications, which capture travel intentions that may translate into future journeys (Falcão e Cunha and Galvão 2014).

This diffusion of sensors has driven data-intensive research methods and practices in urban passenger transport. In recent years, the widespread adoption of advanced personal mobile devices with advanced sensing abilities triggered a relatively new field called ‘participatory sensing’ (Kanhere 2011). It leverages the ubiquity of mobile devices in urban areas to monitor phenomena of interest. Participatory sensing has found several applications in passenger transport domains, including real-time monitoring of traffic conditions (Work and Bayen 2008; Herring et al. 2010) and road conditions (Mohan, Padmanabhan, and Ramjee 2008; Ndoye et al. 2011). At a lower level of data aggregation, participatory sensing has been used to infer the travel mode of individual passengers (Zheng et al. 2008; Reddy et al. 2010; Hemminki, Nurmi, and Tarkoma 2013) and for personal walking behaviour recognition (Bujari, Licar, and Palazzi 2012) from GPS receiver and accelerometer readings. In addition, several mobile applications exist that track and record cycling and jogging workouts (e.g. Mapmyride). These have proved hugely popular and are an additional source of journey data. But whilst data obtained from mobile sensors is detailed and locationally accurate, it relies on users having to download and run specific applications in their mobile devices.

That issue does not apply to mobile phone network data, but that comes at the price of location accuracy, which is much coarser in this case. Still it has been used as an alternative for the monitoring of traffic conditions (Bolla and Davoli 2000; Lovell 2001; Yim and Cayford 2001; White, Quick, and Philippou 2004; Bar-Gera 2007; Calabrese, Colonna, et al. 2011) and to analyse personal travel behaviour (Yuan, Raubal, and Liu 2012; Jiang et al. 2013). But even more importantly for this thesis, mobile phone network data enabled studies on the estimation of travel patterns and on the understanding of human mobility behaviours, which are subsequently reviewed.

Mobile phone network data

The availability of reliable O-D information is vital for the planning and monitoring of passenger transport systems. An O-D matrix contains, for a set of geographic zones defined, the number of trips going from each origin to each destination during a specific time period. O-D matrices depict passenger travel demand but are generally challenging to obtain. Their estimation traditionally relies on extensive travel surveys carried out in a periodic basis, which are expensive to conduct, time-consuming, and prone to response bias (Barry et al. 2002). In addition, those travel surveys depict a single snapshot at a certain time (Caceres, Wideberg, and Benitez 2007). Therefore, several studies have focused on the estimation of dynamic O-D matrices and traffic counts from anonymised mobile phone network data. But this strategy requires collaboration from the mobile network operators, who own the data, and their willingness to provide it. Some authors have worked around this difficulty by testing their proposed methods against simulated phone probe trajectory information, and obtained promising results (Caceres, Wideberg, and Benitez 2007; Sohn and Kim 2008; Zhang et al. 2010; Chen, Bian, and Ma 2014).

Other authors who have proposed methods for estimating dynamic O-D matrices have indeed had access to real mobile phone network data of different kinds. Friedrich et al. (2010) used location-area updates data, which is recorded from any mobile phones in standby mode, combined with a transport network model, and recorded traffic and passenger counts. That comprehensive set of data allowed them not only to generate aggregate O-D matrices but also separate O-D matrix estimates for passenger cars, lorries, and rail passengers. Calabrese, Di Lorenzo, et al. (2011) used network connection data, in which a record is generated every time a device connects to the mobile phone network, either making or receiving a call, sending or receiving a text message, or connecting to the Internet. The authors have proposed an algorithm that allowed for the estimation of weekday and weekend aggregate O-D matrices over a large metropolitan area, and the results showed a high degree of accuracy evaluated against travel to work census data. Despite the positive outcome, the authors did highlight some limitations of the above methods. These include difficulty in extrapolating to the entire population given the market share of the

mobile phone network operator and specific customer base, the error associated with determining location from triangulation of radio signals, and people carrying multiple devices (Calabrese, Di Lorenzo, et al. 2011).

The availability of mobile phone network data has also given researchers the possibility to enlighten our understanding of human mobility patterns. González, Hidalgo, and Barabási (2008) studied the trajectories of anonymised mobile phone users for an extended period of time, and found that human mobility shows a high degree of temporal and spatial regularity when compared with random walk models. The authors argued that an individual has a time-independent characteristic travel distance, and shows significant probability to return to a small number of highly frequented locations. Other studies have related human mobility with urban morphology. Isaacman et al. (2010) demonstrated different human mobility patterns in the cities of Los Angeles and New York, for example that travel distance in Los Angeles is nearly two times greater than in New York. Kang et al. (2012) looked at the impact of city compactness and size on travel distance with a comparative study of intra-urban travel across eight cities in northeastern China. As may be expected, the authors found larger or less compact cities to be associated with greater travel distances. Further research usages of mobile phone network data include testing of a model of migration patterns (Simini et al. 2012) and studies on the uniqueness, predictability and variability of human mobility and activity patterns (Song et al. 2010; de Montjoye et al. 2013; Järv, Ahas, and Witlox 2014).

The studies reviewed in the previous paragraph illustrate the range of additional applications of mobile phone network data, which include epidemic prevention, emergency response, urban planning and modelling, and agent-based modelling (González, Hidalgo, and Barabási 2008; Isaacman et al. 2010; Calabrese et al. 2013).

Traffic sensor data

Another stream of research has focussed on estimating dynamic O-D traffic matrices from Automatic Vehicle Identification (AVI) data or Vehicle Detector (VD) based counts. Zhou and Mahmassani (2006) have proposed an O-D matrix

estimation approach to leverage the growing popularity of AVI technologies that provide point-to-point traffic flows. Their approach relies on a combination of AVI counts, link counts, and historical information into a multi-objective optimisation framework, which proved effective upon testing with simulated data. Other studies have addressed the problem of estimating dynamic O-D matrices by mapping O-D flows to VD based link traffic counts (Kattan and Abdulhai 2012; Toledo and Kolehkina 2013). Prior to these, Hu and Wang (2008) had proposed a mathematical programming framework for determining the most desirable VD locations for O-D matrix estimation purposes.

Public transport AFC system data

This sub-section is largely based on Nunes, Galvão Dias, and Falcão e Cunha (2016), “Passenger Journey Destination Estimation from Automated Fare Collection System Data Using Spatial Validation”, IEEE Transactions on Intelligent Transportation Systems © 2016 IEEE.

AFC systems are used in urban public transport systems throughout the world. Whilst collecting fares automatically is their primary design function, AFC systems continuously generate data that is also of interest for the extraction of travel patterns (Pelletier, Trépanier, and Morency 2011). In some cases, that task requires additional logic for estimating attributes that are not present in raw data. Such is the case of entry-only public transport AFC systems, which only record passenger entry time and, in some cases, entry time plus location. In this configuration, passenger fare media (e.g. smartcards, more recently mobile devices) are only read at the beginning of a journey, creating a transaction corresponding to a single data record. The large dataset that was available for this research is from an entry-only AFC system, which is integrated with an Automatic Vehicle Location (AVL) system. For that reason, this review will cover in greater depth previous methodologies for the estimation of O-D matrices from entry-only AFC systems. These informed the methodology proposed in Chapter 4, which ultimately provided an enriched journey dataset to evaluate the user collaboration model. The methodology itself is a contribution of this research to the literature.

The estimation of an O-D matrix from entry-only AFC system data relies on estimating journey destinations. Destination estimation and O-D matrix estimation are different issues, but the designations are often used interchangeably. The difference falls upon the level of data aggregation. An O-D matrix contains aggregate journey counts between O-D pairs, whereas destination estimation looks at each journey individually. But most O-D matrix estimation works rely on a destination estimation algorithm, which outputs are subsequently aggregated. Hence the topics are intertwined. Barry et al. (2002) pioneered this topic with a methodology applied to entry-only AFC system data from the New York subway. The authors introduced two seminal assumptions that have been applied in several studies afterwards, and which are based on the fact that passenger journey origins are known from entry-only AFC data, but destinations are not. The first of these assumptions is that the most likely destination of a passenger journey is the origin of the next journey (hereafter referred to as continuity of daily travel). The second is that the most likely final daily destination of a passenger is the first daily origin (hereafter referred to as circularity of daily journey chains). From these assumptions they were able to estimate the destination for 83% of boarding transactions in a single day sample.

Barry, Freimer, and Slavin (2009) expanded this approach to include both New York City subway and bus data. Adding bus data required a slight modification to the abovementioned assumptions, to consider that an estimated destination may not be the same but the nearest stop to the related origin. Evidently, these assumptions do not always hold. A passenger may carry out an intermediate journey on foot or by car for instance, which will probably break the assumption of the destination being nearest to the next origin. Similarly, if a public transport passenger stays in different places overnight the second assumption of the final daily destination being nearest to the first origin will likely be broken. Therefore, the validity of results obtained from O-D estimation methodologies should be verified. Barry et al. (2002) validated their methodology using travel diary information and found that the assumptions held for 90% of passengers surveyed. These authors compared their estimated destination totals with exit counts at the stations, and estimated peak load point passenger volumes to increase the robustness of their study. This sort of validations has been

called exogenous because they rely on external datasets instead of the actual AFC system data which the assumptions were applied to (Munizaga et al. 2014).

The aforementioned assumptions (Barry et al. 2002) have been used in later research work on the same topic. Trépanier, Tranchant, and Chapleau (2007) applied a comparable methodology to bus service AFC system data from Gatineau, Québec, but introduced an endogenous validation step into their methodology. It required the assumed destination to be within a 2000 m Euclidean distance from the related origin, otherwise it was assumed that an intermediate journey in an unrecorded mode of transport was likely to have taken place. Another difference was how boarding transactions that were single in a day for a specific passenger were dealt with; instead of not estimating the destination for those records, the authors carried out journey regularity analysis for the passenger and estimated the destination on that basis whenever enough data was available. The authors were able to estimate the destination for 66% of journeys in their sample (Trépanier, Tranchant, and Chapleau 2007). Zhao, Rahbee, and Wilson (2007) have proposed a comparable methodology and used Chicago Transit Authority rail AFC system data, but with a maximum transfer Euclidean distance of only 400 m. The authors did not attempt to estimate destinations of single day journeys, but were still able to do it for 71% of journeys, and their method was partially validated at aggregate level using O-D survey data.

Similar logic was applied to other studies. Farzin (2008) has proposed an O-D estimation methodology for bus AFC system data from São Paulo and tried to validate the results with an O-D household survey. Her conclusions were hampered by the scarcity of buses equipped with AVL technology and a time gap between data sources. A methodology proposed by Li et al. (2011) using bus AFC data from Jinan, China, is claimed to have estimated the destination of 75% of journeys, but appears to rely exclusively on a transfer distance based endogenous validation. Wang, Attanucci, and Wilson (2011) have applied their methodology to bus AFC system data from Transport for London (TfL) combining two types of validations. The first was a maximum transfer Euclidean distance, whereas the second was a validation of the results obtained from applying their methodology against an extensive bus O-D

survey. The authors did not attempt to estimate destinations of single day journeys, and were able to estimate the destination of 57% of journeys.

Drawing from several previous methodologies, Gordon (2012) has proposed a sophisticated algorithm to estimate both destination and time of arrival of bus passenger journeys, which was similarly tested with TfL data. Adding the time aspect allowed checking if the passenger had enough time to transfer on foot or not. Destinations were estimated for 74% of journeys. Munizaga and Palma (2012) have proposed a methodology with a variation. In addition to distance, it considered generalised time, consisting of a combination of walking time and vehicle travel time, to determine the potential destination of a journey. It was applied to bus and metro AFC system data in Santiago and yielded an 80% estimation rate with a maximum transfer Euclidean distance of 1000 m. Munizaga et al. (2014) have built on that methodology by proposing more robust endogenous validations methods. Exogenous validation at a disaggregate level from an O-D survey and an experiment with volunteers provided evidence of good performance. Lastly, it is noted that in some of the aforementioned studies, not only the destination but also the origin of journeys has to be estimated from fusing AFC and AVL datasets (Zhao, Rahbee, and Wilson 2007; Farzin 2008; Barry, Freimer, and Slavin 2009; Wang, Attanucci, and Wilson 2011; Ma et al. 2012). That does not apply to the data available for the present research.

The aforementioned studies have focused on the estimation of O-D matrices from entry-only AFC systems across various transport systems around the world. All of these are specific in terms of travel behaviour, dataset availability, and degree of integration between systems. And so are the various methodologies proposed, namely in terms of the variety and strictness of validation rules that are aligned with their specific goals. Hence the outcomes are varied and not directly comparable, but there is a general belief that the main assumptions are valid in the majority of instances. However, all of these works have been applied to entry-only systems with flat fare structures, hence no previous attempts have been made to use additional data resulting from the operation of distance-based fares to increase the accuracy of the destination estimation results. The methodology proposed in this thesis (Chapter 4)

covers this aspect. Lastly, it is noted that other researchers have used AFC system data for passenger travel studies beyond the estimation of O-D matrices, for example, to understand travel behaviour (Lathia et al. 2013) and route choices (van der Hurk et al. 2015), to estimate ridership levels (Reddy et al. 2009) and journey interchange locations (Seaborn, Attanucci, and Wilson 2009; Gordon 2012), and to model public transport demand (Chu and Chapleau 2008).

2.5. Journey path analysis

Collaboration potential in the context of this thesis is measured in terms of spatiotemporal equivalence of travel patterns. This section reviews trajectory similarity metrics, and clustering and optimisation methods that have been used to evaluate that spatiotemporal equivalence and may apply specifically to journey paths. Several studies across a range of domains have proposed quantitative methods for the analysis of movement data (Long and Nelson 2013). The ‘movement’ designation is deliberately used in this context, because some of those methods apply to physical or abstract forms of movement beyond human travel.

Similarity metrics

Spatiotemporal trajectory similarity metrics have been developed in various domains, which not only include transport (e.g. Atev, Masoud, and Papanikolopoulos 2006), but also environmental health (e.g. Sinha and Mark 2005), pattern recognition (e.g. Zhang, Huang, and Tan 2006), and wildlife motion (e.g. Shirabe 2006). The simplest trajectory similarity metrics are based on Euclidean distance (Long and Nelson 2013). Whilst the original metric is purely spatial, movement similarity is often determined in relation to its temporal dimension too. Sinha and Mark (2005) proposed a time-weighted distance metric that incorporates both spatial proximity, measured in terms of Euclidean distance, and temporal duration of that proximity. Yanagisawa, Akahani, and Satoh (2003) proposed a similarity metric based on Euclidean distance between directed discrete trajectories. These authors dealt with the temporal dimension by calculating similarity between trajectories when occurring within the same fixed time intervals.

Other studies have proposed trajectory similarity metrics based on adaptations of the Hausdorff distance, which is a shape comparison metric (Atev, Masoud, and Papanikolopoulos 2006; Shao, Cai, and Gu 2010). Those adaptations are designed to overcome limitations of the Hausdorff distance for the analysis of movement data, because it disregards the ordering of points, and is sensitive to outliers (Long and Nelson 2013). Atev, Masoud, and Papanikolopoulos (2006), for example, addressed those limitations by augmenting the metric to reflect the chronological order of points of a trajectory, and ignoring the worst matching points between two sets. However, the original Hausdorff distance is purely spatial, overlooking the temporal similarity of trajectories. Therefore, the Fréchet distance may provide a better starting point for a trajectory similarity metric. The Fréchet distance measures similarity between curves, taking into account the ordering of points within those curves. It is often described symbolically as the minimum length of the leash connecting a dog and its walker as they move in separate paths without turning backwards. Eiter and Mannila (1994) proposed a discrete variation of the Fréchet distance that approximates the continuous metric to ease its computation. Yet, in most movement data sources relating to passenger transport (see Section 2.4) the locations of objects are recorded periodically but not at the exact same instant, requiring further adaptations to allow measurement of trajectory similarity (Long and Nelson 2013).

Since this thesis addresses specifically passenger transport domains, trajectory similarity metrics that are specific to objects moving along networks are most relevant, yet scarcely covered in the literature. Hwang, Kang, and Li (2005) studied the properties of similar trajectories in road network space and proposed a measurement method based on the Point of Interest (POI) notion. The authors considered two trajectories to be spatially similar when passing by the same POIs, such as road junctions or popular places. They also considered temporal similarity, proposed as the inverse of temporal distance, which is the time difference between two objects passing by the same POI.

Dodge, Laube, and Weibel (2012) introduced a different approach for determining trajectory similarity. The authors proposed the segmentation of trajectories according

to a set of movement parameters such as speed, acceleration, and direction. In their method, a trajectory is transformed into a symbolic representation, and common patterns are identified based on the variation of the chosen parameters over time. Their approach is flexible, allowing a choice of parameters based on the attributes that are available in a specific movement dataset. The authors illustrate this flexibility by testing the approach in two dissimilar case studies, respectively on hurricane trajectories and GPS tracks of city couriers. Shirabe (2006) proposed another alternative take on trajectory similarity measurement, based on the correlation analysis of moving points whose locations are recorded at equal time intervals. The approach was developed for the identification of coordinated motions, and consists of decomposing trajectories into vector time series to compute the correlation coefficient between pairs of movement trajectories. Yet, the requirement for data to be recorded at equal time intervals presents a limitation in terms of applying the method to passenger transport journey data sources.

The paragraphs above reviewed several trajectory similarity metrics without dwelling into results, because they can hardly be compared. There are no best or worse metrics, ultimately it all depends on the way researchers conceptualise movement similarity according to their specific research objectives, and to the type and to the attributes of the data that is available to them.

Clustering methods

Another stream of research has approached movement similarity using methods that identify trajectory patterns and clusters (Gudmundsson, Laube, and Wolle 2008). The interest in their development has grown in recent years with the increasing availability of movement databases, which accumulate location data captured by mobile devices over time. Those methods include k -nearest neighbour queries on moving object trajectories, which are able to identify movement patterns such as convoys travelling together or similar trajectories to a given one (e.g. Gao et al. 2009; Güting, Behr, and Xu 2010). Other studies have focused the adaptation of density-based clustering algorithms, for the same specific purpose of grouping similar moving object trajectories (e.g. Nanni and Pedreschi 2006; Kalnis, Mamoulis, and

Bakiras 2005). Another example is found in Shoshany, Even-Paz, and Bekhor (2007), who have developed a data clustering method based on linear programming. It tracks the dynamic evolution of clusters and provides information regarding the spatial behaviour of entities such as people or animals.

Carpool service problem

In contrast with the variety of trajectory similarity metrics and clustering methods reviewed in this section, no previous studies have been found to address the identification of alternative movement trajectories. Yet, this type of movement trajectories is highly relevant to this thesis, since the definition of spatiotemporal equivalence combines similarity and compatibility of alternative trajectories to achieve identical travel goals, i.e. between common points (see Chapter 5). The recent research interest on the carpool service problem implicitly closed the abovementioned literature gap by looking at the spatiotemporal similarity of travel itineraries without predefined trajectory constraints. In order to illustrate this, the carpool service problem is briefly described. In broad terms, it aims to find the best matches between a driver and passengers travelling in the same direction at the same time, so that they can share a single vehicle. Each of the interested parties provides an origin and a destination, and an algorithm designed to solve the problem determines a common trajectory that fulfils all of their travel needs. So instead of looking at movement similarity, these algorithms only look at the compatibility between sets of itineraries depicted by an origin and a destination at a specific time.

Existing approaches have used genetic algorithms to solve the carpool service problem, which is formally defined as a multi-objective optimisation problem (e.g. Jiau, Huang, and Lin 2013; Huang, Jiau, and Lin 2014a; Huang, Jiau, and Lin 2014b). The various objectives in the problem consist of maximising the total number of passengers matched with drivers and their reputation scores if applicable, and minimizing the travel distances and waiting times, subject to a myriad of restrictions (Huang, Jiau, and Lin 2014a). The complexity of the problem makes it computationally demanding.

Relevance in information retrieval

The connection between information retrieval and journey path analysis is not obvious. But in fact, the evaluation of spatiotemporal equivalence between journey paths may be argued to share some basic characteristics to the search for information on the Internet. Evaluating the degree to which journey paths are equivalent may be seen as a problem of determining the relevance of one journey path to the other. And whilst this does not draw from web search engine algorithms (e.g. Brin and Page 1998), the spatiotemporal equivalence between journey paths is quantified via a proposed composite relevance measure. Kleinberg (1999) stressed that the notion of relevance, which is used to determine the ranking of search results for a given query, is inherently subjective. That subjectivity drives the challenge of improving the quality of search results. The proposed spatiotemporal equivalence measure hinges upon a similarly crucial notion of relevance. However, by virtue of timeliness, the focus is shifted from the actual information to the ranking of users who generate that information. Subjectivity however remains and so does the challenge.

This section has reviewed several existing quantitative methods for the analysis of movement data. The proposed spatiotemporal equivalence measure draws from various notions covered here, namely similarity, compatibility of alternative trajectories, and relevance, in order to evaluate collaboration potential between passenger transport users.

2.6. Review summary

This chapter framed the topic of this thesis (Figure 2.1). It started with a review of theoretical aspects relating to notions of *customer experience* and *value co-creation* that informed the conceptual foundation of the user collaboration model. A key outcome of the first section on customer experience was the identification and summarisation of factors that were identified in the literature to have an influence on travel experience. This is fundamental to inform how to improve that experience through user collaboration. The main outcome from the second section on value co-creation was a confirmation of potential benefits from the active engagement of

transport users, and from interactions amongst their peers and transport providers. This was achieved with a review of studies that illustrate successful co-creation practices in the public domain, and particularly in passenger transport. The next part of this chapter was concerned with practical examples of *ICT-based user collaboration practices*, with a specific focus on passenger transport. An important outcome of this section was a listing of passenger information requirements that can be met with ICT-based user collaboration. The availability of information is as a factor that influences travel experience.

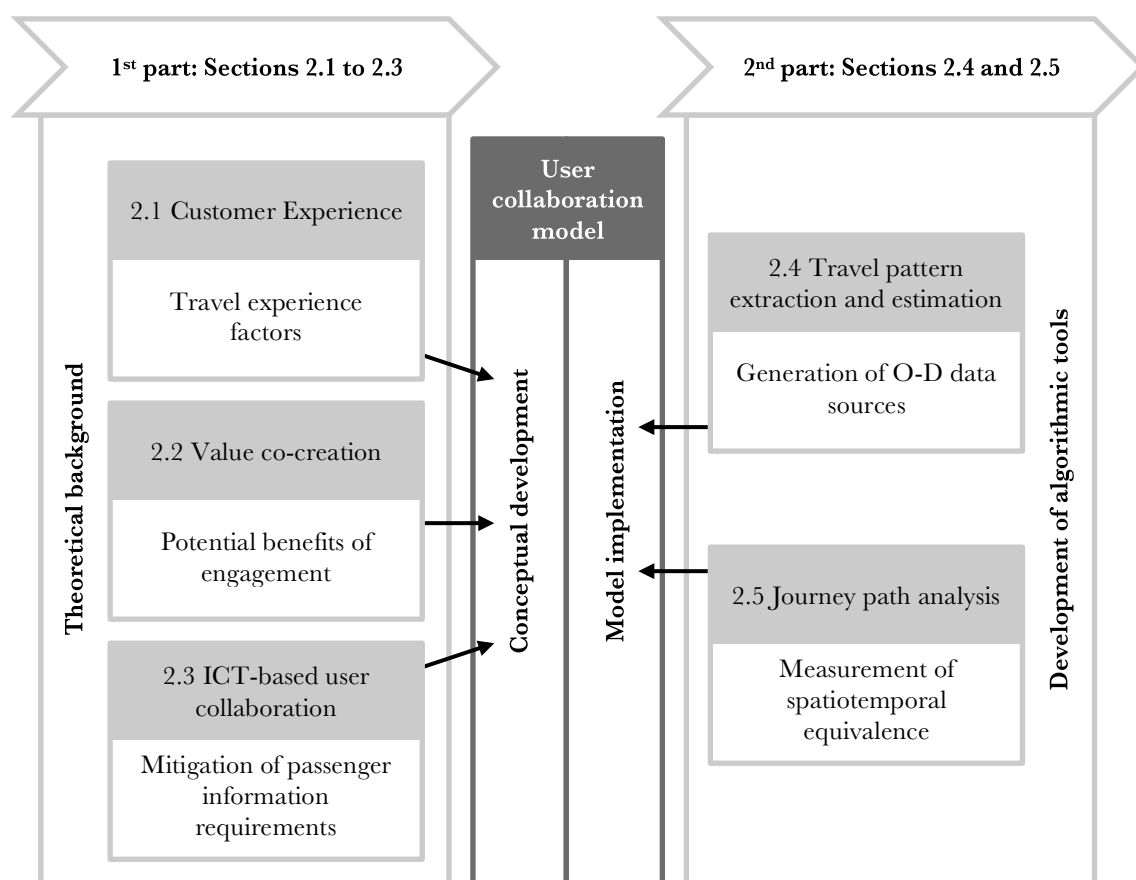


Figure 2.1 Literature review map

Having reviewed the theoretical aspects and state-of-the-art informing the conceptual development of the user collaboration model, the remainder of this chapter focused on sources of data and methodologies relating to its implementation. The fourth section focused on the *extraction and estimation of passenger travel patterns*, giving particular emphasis to methodologies for estimating the destination of individual passenger

journeys from entry-only AFC system data. The aim has been to identify the main characteristics and limitations of those studies in order to frame the contribution and innovative features of the methodology proposed in Chapter 4. The last section reviewed several quantitative methods for the *analysis of movement data*. The purpose, again, has been to frame the innovative aspects of the proposed composite measure of spatiotemporal equivalence described in Chapter 5 and its contribution to the literature.

3. Foundations of the collaboration model

This chapter describes the conceptual foundations of the user collaboration model. It begins by describing *what* the opportunity and main objective are. Section 3.1 frames the opportunity as the widespread adoption of advanced personal mobile devices with ubiquitous access to wireless communication networks, and the objective in terms of harnessing collective intelligence. This chapter continues with the reasoning *why* travellers and transport providers will benefit from the model. Section 3.2 discusses the benefits of leveraging information that is distributed throughout an urban passenger transport system. Finally, this chapter defines *how* the model works. Section 3.3 describes the role of users of an urban passenger transport system as information providers and consumers. Section 3.4 reasons the need for reimagining affinity between these users based on their location and travel patterns, to account for the spatiotemporally dynamics of urban passenger transport. Section 3.5 describes the requirement for validation and incentive mechanisms to support collaboration, and to sustain the reliability and breadth of information obtained from users. Section 3.6 provides a summary of the conceptual foundations of the model.

Few excerpts in this chapter are based on Nunes, Galvão, and Falcão e Cunha (2014), “Urban Public Transport Service Co-Creation: Leveraging Passenger’s Knowledge to Enhance Travel Experience”, Procedia - Social and Behavioral Sciences.

3.1. Leveraging collective intelligence

The motivation for this research was presented in the introductory Chapter 1, and has been said to consist of a need and an opportunity. That opportunity was described as the widespread adoption of advanced mobile devices combined with the ubiquity of wireless communication networks, which led us towards a pervasive computing paradigm. Then, the literature review (Chapter 2) highlighted the role of those advances in ICT in facilitating value co-creation practices based on user collaboration across various service domains. And lastly, the literature review

(Chapter 2) also provided evidence that the development of such practices in urban passenger transport remains short of fulfilling their entire potential in terms of leveraging the collective intelligence of users. Despite indication that some transport providers have willingly started to embrace value co-creation practices, it was noted that monopolistic institutional settings, a general aversion to change, and the inherent spatiotemporal dynamics have created barriers to the maturing of those practices.

In addition to the argument above, Chapter 2 covered recent studies that argue in favour of empowering users with tools to foster bottom-up development, namely taking into account the increasingly common institutional setting of diminishing public funding. Each and every user of urban passenger transport has limited visibility of the system as a whole, but a single, location-specific perspective of its operation as it unfolds. Whilst their knowledge may appear to be of limited value in isolation, taking all users into consideration there is a mass of observers across the system at any given time (e.g. Figure 3.1, showing the entry location of public transport users of light-rail and of the main bus operator in Porto, Portugal, within a selected AM peak 15 minute interval). Hence, there is potential for harnessing that collective intelligence, understood in light of the definition by Lévy (2010) previously cited (see Section 2.2). But that requires the development of mechanisms to facilitate the active involvement of users and incentives to the collaborative co-creation of value. Furthermore, it is argued in this thesis that it requires a new configuration of social networks that reimagines affinity to account for the abovementioned spatiotemporal dynamics of transport.



Figure 3.1 Public transport user entries, bus and light-rail, 15 minute interval, Porto

The main objective of the proposed user collaboration model is to leverage advances in ICT to harness collective intelligence in an urban passenger transport system. As implied in this statement, the focus of collaboration within the model is on the generation of information through ICT. Nonetheless, there may be specific instances where that information also enables collaboration in terms of sharing resources. That would be the case, for example, of exchanges of information between users that lead to ride sharing.

The main objective divides into three tangible aims, the first of which is the *better utilisation of available resources*. From a transport provider perspective, this may translate into better monitoring of the services provided without a significant investment in additional human resources or Intelligent Transport System (ITS) technologies. A better utilisation of resources may also derive from a richer understanding of

customer preferences and travel demand for optimising capacity allocation, for example, in terms of planning routes and frequencies of public transport services. From a traveller perspective, user collaboration has the potential to augment existing travel information resources. Examples include the timely identification of unplanned changes to services, improved insight into travel alternatives based on reviews provided by others, greater ease to tailor a journey to personal preferences, and the discovery of ride sharing opportunities. The concepts of collective awareness and social capital mentioned in section 2.2 are well aligned with this first aim (Ostrom 2000; Pitt et al. 2013).

All the examples above tie in with the second tangible aim of the user collaboration model, which is the *improvement of the travel experience*, not only in terms of information provision but also by acting upon experience-related factors that were summarised in Chapter 2. In other words, whilst the availability of travel information is a factor that influences the travel experience by itself, it has also potential to mitigate issues relating to the convenience and flexibility of specific modes of transport, and to let passenger transport users plan and adjust their journeys to suit their individual preferences in terms of travel characteristics and environment.

If a better utilisation of resources and an improvement of the travel experience are realised in relation to more sustainable travel modes, their use is encouraged and patronage will likely increase. In public transport modes specifically, these objectives may lead to the desirable Mohring effect (Mohring 1972). This designation relates to a virtuous cycle of increased demand justifying higher frequency of services, which in turn reduces waiting times and stimulates even greater levels of demand. This brings clear benefits for the society in terms of the *promotion of more sustainable travel*, which is the third tangible aim of the user collaboration model.

3.2. Information value

The user collaboration model is intended to generate information of value both for travellers and for transport providers. Whilst the information requirements of travellers are documented in literature reviewed in Chapter 2, no previous works

on the information needs of transport providers, which could potentially be met through user collaboration, were known to the author of this thesis at the time of writing. This section will look at information value from both perspectives, starting with travellers and ending with transport providers. As far as the former are concerned, information value is discussed in the following paragraphs in light of the reviewed travel experience factors set out in Table 2.1, and information requirements set out in Table 2.2.

Information value for travellers

The travel experience factors relating to travel environment include comfort and crowding, cleanliness and maintenance, and driver skills. Whilst collaboratively generated information may not mitigate any of these issues directly, it provides an additional layer of information to help travellers make and adjust their travel plans according to their individual preferences. For example, a traveller who is sensitive to crowding of public transport vehicles may be willing to choose the travel alternative that is most likely to have spare capacity, based on reported experiences from others. In other words, that information may help the passenger find a pleasant travel environment. A similar logic applies to travel experience factors relating to journey characteristics, namely journey time. This already happens in terms of private vehicle use, whereby crowdsensed traffic conditions (e.g. Waze, see Section 2.3) help drivers plan and adjust their routes in real-time to speed up their journeys. Still, collaboration may potentially yield comparable insights for public transport modes, letting passengers decide on routes based on real-time travel conditions, or at least manage their expectations of travel time. Lastly, this logic also applies to travel experience factors such as safety, social environment, and scenery. The generation of information regarding these aspects may also facilitate personalised travel planning.

The availability and quality of information itself was reviewed to be a factor influencing the travel experience. This is not surprising if the holistic travel experience is considered, particularly in relation to non-recurrent public transport journeys. These often start with a passenger planning a route before entering a station or vehicle, and deciding based on the information that is available. But the

availability and quality of information may have secondary positive impacts on other travel experience factors reviewed. Having a good understanding of available travel alternatives can enhance perceived convenience and flexibility of non-private transport modes, thus inducing a sense of control. Ultimately that will have longer-term impacts on yet another travel experience factor, which is the modal preference of a traveller. Adding to the arguments above, communication between passenger transport users may also generate new modal alternatives such as car sharing or cycling in groups, associated with entirely different, and potentially better, travel experiences. Lastly, communication between users has potential to be stimulated in an engaging and playful way (e.g. a serious game) and improve the travel experience as a form of entertainment.

Regarding the information requirements reviewed and set out in Table 2.2, the collective intelligence of passenger transport users may be helpful in relation to basic types of information, through the identification of deviations to scheduled services. But more importantly, that collective intelligence can play a role in providing more advanced types of information. Travel disruption warnings are an obvious type, and arguably the most useful too. A significant amount of time usually elapses from the moment an unplanned service disruption occurs until affected travellers are warned of it and of its consequences. A transport provider typically needs to receive some form of notification of the disruption, confirm it, and then act upon it by attempting to relay the information to their customers using the communication channels that are available. However, such information can be timely sourced from users scattered across the transport system themselves, who observe events inasmuch as their evolution and travel implications. This requires their empowerment to report such events. Furthermore, it requires adequate ICT-based tools to distribute relevant information efficiently to selected potentially interested parties, whether it originates from travellers or transport providers. Whilst the crowdsourcing of travel disruptions is already happening for private vehicle use (e.g. Waze, see Section 2.3), some limitations arguably exist and will be highlighted later in this chapter.

Collective intelligence may also cover other advanced types of information, including comfort-related aspects, and appraisals of safety and reliability of services from

unbiased user reviews. E-hailing applications (e.g. Uber) that feature driver and passenger ratings are a current example of a collaboratively generated safety appraisal. These ratings let both consumers and service providers review their experiences of the other to lessen safety risks associated with the transport service. However, this model has yet to extend into other transport modes, namely public transport.

Information value for transport providers

Having argued the value of collaboratively generated information from the perspective of travellers, the argument moves on to that of transport providers. As in any other service domain, transport providers should know their customers well. This is particularly relevant to transport providers such as public transport operators and taxi or private hire companies. Today, that knowledge of customers derives mainly from satisfaction surveys and complaints. These information collection methods are limited in several ways, as discussed in the following paragraph.

First, satisfaction surveys do not occur regularly, so passengers are unlikely to recall past events that fundamentally changed their perceptions and attitudes towards the service. Someone may dislike a specific travel mode but no longer remember the underlying events that triggered dissatisfaction. Second, passengers end up having a passive role on satisfaction surveys, because they only answer what and when something is asked from them. In theory, this may even introduce a positive bias because a satisfaction survey fails to capture customers that gave up on the service altogether before the survey was carried out. And third, the information captured via surveys is limited in its scope. For example, the passengers of a bus service may report a general dislike for waiting facilities, but pinpointing the worst ones, and understanding why some are favoured in relation to others may be trickier because there are many external variables at play. This example additionally highlights that customer satisfaction with some aspects of the service provided is influenced by the surrounding environment that is not controlled by the transport provider. In this example, by the quality of the place where the bus stop is located.

Whilst traditional surveys cannot easily capture finer aspects influencing customer satisfaction, ICT-based tools to enable crowdsourcing of timely and continuous location-based reports from passengers are able to mitigate that issue. The availability of such information would enhance the knowledge that a transport provider has of its customers, their perceptions and attitudes, to address service deficiencies quickly and more consistently from real reports. This would impact positively on service patronage and on the overall travel experience. As illustrated in the literature review (Section 2.3), some transport providers have been actively engaging with their customers using ICT-based social networking services such as Facebook and Twitter, but which are poorly suited to capture location and to enable the aggregation of information for subsequent analysis.

Adding to the above limitations, existing methods for collecting information are costly and time consuming. Hence, information that is valuable for transport providers may not be gathered as often as ideally should be. A thorough understanding of travel patterns and modal decisions is essential for adapting the service offer to existing demand. And so is the monitoring of service delivery indicators such as punctuality or crowding in public transport, or spare capacity in a road system. Today, gathering that sort of information typically requires carrying out travel surveys or investment in monitoring technologies based on sensors and tracking devices, which may either be unaffordable or be low priority investments for transport providers.

However, urban passenger transport users are potential observers distributed across a transport system, who hold knowledge about services as they unfold. These users can become involved in providing information that is valuable not only to their peers, but also to transport providers. Such information can be significantly cheaper, and spatially and temporally more relevant for transport providers, for monitoring service delivery indicators in real-time. Transport providers can potentially leverage the ubiquity of personal mobile devices with communication and sensing abilities to crowdsource structured information on aspects such as traffic incidents and delays, and to continuously monitor travel patterns. In the public transport domain specifically, that information can potentially include additional aspects to be sourced

from passengers, such as crowding levels and rating of drivers. Crowdsourced information in real-time may ultimately be incorporated into the information systems of transport providers to assist with live operational decisions. Examples include allocating more capacity to public transport routes where crowding is being reported, and assisting road traffic operations via management of traffic lanes or signalised junctions. This can potentially lead to the improvement of travel experience through timely action upon reported problems. It is worth mentioning, however, that the privacy of passenger transport users supplying information must be safeguarded. That requires data to be crowdsourced anonymously and to be aggregated for later analysis. Besides safeguarding privacy, aggregation may also facilitate the consumption of reports.

3.3. User participation

The main requirement of the user collaboration model is the participation of users of urban passenger transport, as providers of spatiotemporally structured information, through their personal mobile devices. The expression spatiotemporal structuring is used in this thesis in regards to information that is time-stamped and geolocated, and in some cases spatially contextualised (e.g. associated to a specific public transport route or vehicle). Drawing from principles of Web 2.0, the model aims to stimulate the collaborative generation of information for the co-creation of value within an urban passenger transport system. Moreover, the model intends to ease communication between users to leverage the benefits of C2C interactions reviewed in Section 2.2. The participation may involve active user input, for example, for the crowdsourcing of service delivery ratings and travel information reports. Yet, user input may also be automated, for example, for the crowdsensing of traffic conditions or monitoring of characteristics of the travel environment. Whilst existing personal mobile devices are already capable of automatically monitoring several aspects such as movement and noise levels, the permanent evolution of their sensing capabilities may in the future allow the crowdsensing of environmental variables such as air pollution levels (e.g. Holub 2014).

Besides their participation as providers of spatiotemporally structured information, users of urban passenger transport and transport providers are the consumers who reap the value of that collaboratively generated information. The previous section of this thesis focused on the value of that information. Yet, it is noted that the aforementioned forms of participation have previously been explored. Section 2.3 reviewed mobile applications such as Waze, Trafi, Moovit, and Tiramisu, as examples of crowdsourcing and crowdsensing from passenger transport users. Such examples reveal the potential to harness spatiotemporally structured user input, in a way that facilitates the aggregation of data. Contrariwise, these examples fail to leverage interactions amongst users of urban passenger transport. They represent data-centred approaches instead that barely reflect affinity between passenger transport users, thus providing limited ability to proactively filter and forward information to them according to their travel patterns. Trafi is an exception to a limited extent; whilst not reflecting affinity, the application uses an alternative approach based on machine learning to estimate what users may find important.

The proposed user collaboration model attempts to close this gap through an evaluation of collaboration potential between users in real-time, in order to promote interaction and facilitate the distribution and subsequent consumption of information. In practical terms, this might represent an evolution from a situation where users have to actively seek for information (e.g. by planning a journey in Moovit or Waze), to a scenario where information is proactively relayed to them from matching travel patterns in real-time. Furthermore, the evolution proposed by the model may lead to the identification of collaboration opportunities in terms of modal alternatives such as ride sharing.

The following section describes how the model addresses the evaluation of collaboration potential. It proposes reimagining affinity to account for the spatiotemporal dynamics of urban passenger transport.

3.4. Reimagining the social network

The literature review (in Section 2.3) described the growing tendency from transport providers to be present in ICT-based social networking services to reach out to their customers. Yet, these services tend to emulate tangible social structures such as friendship and professional relationships, and materialise actual ties between users. Whilst these ties have enabled collaboration-based value co-creation in specific contexts, they are unlikely to reveal similar opportunities in a dynamic urban passenger transport setting.

This thesis argues that drawing greater value from interpersonal affinity in urban passenger transport most likely has to account for the inherent spatiotemporal dynamics. Consequently, the model considers that the spatiotemporal equivalence of travel patterns is the best indicator of collaboration potential between users of an urban passenger transport system. That equivalence is spatiotemporal because it constantly changes in space and in time. What that means effectively is that passenger transport users who may have relevant travel related information for each other are likely those in equivalent routes, and roughly at the same time. For example, a transport user travelling in a public transport route or driving along a specific road observes events that may inform the travel decisions of others travelling upstream along an equivalent route. Another example is that of groups of users with recurrently equivalent travel patterns, who have greater collaboration potential in terms of sharing rides. Reimagining affinity to link users based on spatiotemporal equivalence of travel patterns, instead of tangible social structures, leads to the creation of dynamic social networks. These are constantly transforming as its users begin and terminate their journeys, in an attempt to maximise collaboration potential in real-time.

The reimagined social network configuration has potential to overcome observed limitations of existing crowdsourcing services for urban passenger transport, by facilitating the proactive distribution of information to relevant audiences. Under the proposed user collaboration model, crowdsourced information is set to be distributed across users sharing the same network, who approximate the relevant audience. A

better matching between providers and consumers of information will likely promote collaborative interactions, not only for the benefit of passenger transport users, but also with advantages for transport providers able to leverage greater volumes of information (see Section 3.2). This implies a profound change in dealing with user-provided, travel-related data. Instead of being only referenced to a specific location (e.g. traffic incident) or service (e.g. crowding in specific public transport route), it will be proactively centred on users and individually personalised. This can be illustrated using the Facebook news feed analogy, in which the information that is displayed reflects the unique structure of the social network of a user; users are not required to search for information within their networks, it is automatically and conveniently relayed to them. The main difference from this analogy is that the reimagined social network in the model is adaptive and circumstantial.

The main challenge is to establish networks in a context where affinity is circumstantial. In this context, communities are ephemeral and exist temporarily in space and in time. *The reimagined social network must reflect the location and travel patterns of every user of an urban passenger transport system in real-time.* This new configuration provides the building block for the development of ICT-based applications to leverage the collective intelligence of passenger transport users. The following chapters of this thesis (Chapters 4 and 5) describe the algorithms that were developed as tools for the model based on the reimagined social network configuration. Chapter 6 describes the development of such an application based on the model.

3.5. Validation and incentive mechanisms

Two fundamental concerns that apply to crowdsourcing applications in general are the promotion of user participation, and confirming the reliability of user-generated data. In some existing crowdsourcing applications, in which the reliability of information is paramount (e.g. Wikipedia), users not only provide and consume information, they also assume a third role of validators of information provided by others. Likewise, in passenger transport, the reliability of information is fundamental, in a sense that users make travel decisions and transport providers

make operational and managerial decisions based on the information that is available to them. In either case, the use of unreliable information may have detrimental consequences. Therefore, *the development of applications based on the user collaboration model must account for the need to get users involved in validation.* The evolution of the case study *Journata* described in Chapter 6 illustrates an approach to this requirement.

Having previously ascertained that two aims of the user collaboration model are better utilisation of available resources and improvement of the travel experience, there is great potential for transport providers to benefit from it. However, attaining those benefits in some cases may require *transport providers to offer their customers incentives to participate as information suppliers.* That would be the case, for example, if a public transport operator decided to rely on crowdsourced GPS data as a substitute to carrying out O-D surveys. Although the improvement of travel experience may generally work as an incentive by itself, it may not suffice to encourage higher levels of participation. Furthermore, it would be fair to reward customers who participate more actively in co-creating value with the transport provider. Taking the case of a public transport operator again, potential reductions in operational costs resulting from crowdsourcing data required to monitor service delivery indicators, and increased patronage levels from a sustained improvement of travel experience, may potentially offset the costs of rewarding passengers. An obvious type of reward would be a travel card discount that varies according to the level of commitment from a passenger. Another type would be vouchers to spend on purchases with partner retailers (e.g. Ferreira and Galvão Dias 2015).

From a financial point of view, getting users involved in the co-creation of value in passenger transport services should not present significant risks. Even modest levels of investment from transport providers may potentially generate attractive returns. Conversely, some transport providers may perceive a greater exposure deriving from the interaction with and amongst customers to carry greater risks for them. Easing interaction will likely accelerate the diffusion of customer appraisals, namely the impact of complaints, with a detrimental effect on a brand image created throughout the history of their organisation. Whilst this thesis acknowledges these concerns, there are relevant arguments against them set out in the following paragraph.

First, higher levels of exposure may encourage transport providers to excel. Much knowledge can be extracted from listening to customer feedback; loyal customers praise the brand if provided with a good experience (see Section 2.1), potentially outweighing negative reviews. Second, many transport providers who are truly committed to improving their services through customer engagement, have already exposed themselves successfully. Such is case of transport providers that are actively present in ICT-based social networking services (e.g. Bregman 2012, see Section 2.3). Furthermore, public transport operators that provide real-time arrival information either in stations or via the Internet have allowed their passengers to pinpoint delays, but the benefits in terms of improving the holistic travel experience by mitigating information needs, and in terms of managing passenger expectations, are praised. Such practices also convey an image of technological progress that enhances the reputation of their organisation. Third, many groups exist across ICT-based social networking services that are dedicated to specific passenger transport services, where several aspects are discussed without the formal involvement of the transport provider (e.g. www.facebook.com/metrodoporto, www.facebook.com/TheMBTA). Faced with the inability to stop the debate, transport providers may instead join it and leverage a mutually beneficial interaction. All things considered, the advantages of incentivising engagement can potentially outweigh the associated financial commitment and risks for the transport provider.

3.6. Model summary

This chapter described the conceptual foundations of the user collaboration model. The model fundamentally hinges upon the creation of dynamic social networks for users of a passenger transport system, by reimagining affinity as spatiotemporal equivalence of travel patterns. The main objective is to leverage advances in ICT to harness collective intelligence in an urban passenger transport system, leading to a better utilisation of resources for transport providers, to improved travel experience for users, and ultimately to more sustainable travel for the society in general. The model requires the active participation of users as providers, consumers, and validators of spatiotemporally structured information,

through their personal mobile devices. Furthermore, in some cases, the provision of incentives by transport providers for users to participate is endorsed to drive greater benefits. The model, and particularly the dynamic social networks at its core, is merely a concept. It only materialises with the development of ICT-based applications based on its conceptual foundations. Figure 3.2 attempts to capture the fundamental aspects of the user collaboration model.

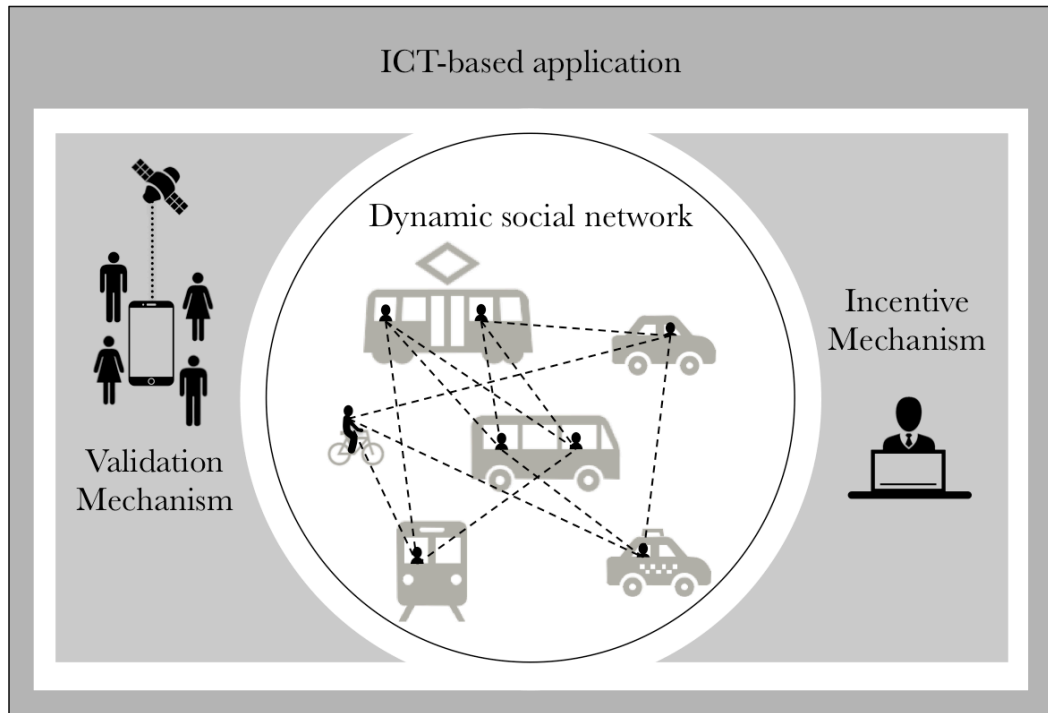


Figure 3.2 Fundamentals of the user collaboration model

4. Passenger journey destination estimation

This chapter presents a methodology for estimating the destination of passenger journeys to produce an O-D dataset. It is the first tool for the user collaboration model. The user collaboration model relies on the estimation of collaboration potential in terms of spatiotemporal equivalence of travel patterns. Hence, the existence of an O-D data source is the enabling step to implement it. Reflecting the initial and primary interest in the urban public transport domain, and the characteristics of the AFC system where the data available for this research came from, the methodology described may require adaptations for application to other entry-only AFC systems, and does not apply to AFC systems where all entries and exits are recorded. Furthermore, in relation to other passenger transport domains, alternative data sources (e.g. mobile phone network data, traffic sensor data) and methodologies for extracting and estimating travel patterns exist, as described in the literature review (Section 2.4); this chapter does not apply to those cases. The ultimate purpose of the methodology described in this chapter is to provide an enriched journey dataset to evaluate the user collaboration model and feed the second tool for the model, described in Chapter 5, which is concerned with estimating collaboration potential. Yet, the methodology itself is a contribution of this research to the literature (Nunes, Galvão Dias, and Falcão e Cunha 2016).

This chapter is organised as follows. Section 4.1 provides an overview of AFC system objectives and configurations to frame the application domain of the methodology. Section 4.2 provides a description of the dataset used in this research study and the characteristics of the AFC system providing the data. Section 4.3 describes a sequence of steps designed to mitigate missing or illogical attributes in the original dataset. Section 4.4 provides the formal definition of the methodology to estimate the destination of passenger journeys to produce the O-D dataset. Section 4.5 describes the structure of an algorithm developed for implementing of the methodology. Section 4.6 discusses the results obtained from the application of the methodology to the dataset, and presents a sensitivity analysis of the trade-off between extensiveness

and accuracy of estimates. Section 4.7 describes additional results from another implementation of the methodology. Lastly, Section 4.8 summarises the conclusions drawn.

This chapter is largely based on Nunes, Galvão Dias, and Falcão e Cunha (2016), “Passenger Journey Destination Estimation from Automated Fare Collection System Data Using Spatial Validation”, IEEE Transactions on Intelligent Transportation Systems © 2016 IEEE.

4.1. Methodology scope

The literature review of this thesis (in Section 2.4) already described that AFC systems have a primary design function of collecting fares, but are also capable of generating data of interest for the extraction of travel patterns. These systems automate the ticketing system, easing public transport use and adding efficiency to revenue collection operations. Additionally, AFC systems enable integrated ticketing across various public transport modes and operators in urban areas. Two main configurations of AFC systems exist. Entry-only AFC systems only record passenger entry time, or entry time and location, since fare media are only read at the beginning of journeys. The alternative configuration requires fare media to be read at the end of journeys too, and records passenger entries and exits. Therefore, entry-only AFC systems require additional logic for estimating the destination of journeys. The entry-only configuration is popular in bus services all over the world, as it avoids alighting delays if the fare media of exiting passengers had to be read upon arrival at a stop, simplifies infrastructure, and reduces maintenance requirements.

Entry-only configurations have often been associated with flat fare structures to lessen the need for on-board inspection to control underpaid travel. Yet, exceptions to this are becoming increasingly common with public transport providers being driven to deliver more equitable distance-based pricing. The Andante AFC system in Porto, Portugal, which is the source of data for this research study, is such an exception, as are the Leap Card in Dublin buses and the SL Access in Stockholm buses. In addition, Toronto Transit Commission and Utah Transit Authority have been

considering switching to distance-based fares (Kalinowski 2014; Utah Transit Authority 2014), whereas First buses in the West of England have already done it.

The methodology developed in this research applies to entry-only AFC systems combined with distance-based fare structures. Still, each system will have specificities that must be understood and considered for it to be applied. The methodology contributes to the literature by introducing two spatial validation features that increase the accuracy of destination estimation results and verify the key assumptions found in previous works (Barry et al. 2002). These key assumptions, relating to the continuity of daily travel and to the circularity of daily journey chains, were reviewed in Section 2.4. The methodology further increases the accuracy of destination estimation in comparison with previous works, through a refinement in the approach towards single daily journeys with multiple stages. This will be explained in Section 4.5. The methodology takes advantage of additional data attributes resulting from the operation of distance-based fares. Yet, whilst increasing accuracy raises confidence that destinations are correctly estimated, it decreases the overall estimation rate. The results obtained suggest that the methodology is effective and reliable for estimating the destination of journeys.

4.2. Dataset description

The dataset used in this research study is from the Andante AFC system (www.linhandante.com). Andante is an entry-only AFC system with a distance-based fare structure that covers the metropolitan area of Porto. The fare media are contactless travel cards that can be used across eleven participating public transport operators, which include buses and railways. The distance-based fares are defined by a zoned structure. Andante is divided into geographic travel zones and the journey fare depends on the number of zones travelled between its origin and destination. In addition to having distance-based fares, Andante is a time-based system for pay-per-use passengers. The time limit does not apply to monthly travel card subscriber passengers, who are entitled to unlimited journeys within the travel zones in their subscriptions. Pay-per-use passengers are allowed to make unlimited transfers in a

given time period, ranging between 1h00 and 3h15, which increases with the number of zones that are included in the fare. For example, the time period of a three-zone fare is 1h00, whereas of a five-zone fare is 1h30. A journey relates to a single fare and consists of one or more journey stages in different routes or vehicles.

The Andante AFC system creates a transaction record every time a passenger taps a travel card on a reader. This must happen at the beginning of each journey stage, when changing routes or entering another vehicle, which facilitates the identification of all stages constituent of a journey. Being a time-based system, pay-per-use passengers sometimes tap their cards during a journey to check on the display of the card reader how much time is left for travelling without paying an additional fare. This may also happen when a passenger cannot recall having tapped the card and repeats it for the same journey stage. The dataset used in this research study consists of all types of transaction records described above. Each transaction record contains several data attributes, of which the following are of interest to the methodology:

1. Travel card serial number;
2. Station code or bus stop code where the transaction took place;
3. Route designation (only applicable to bus journeys);
4. Direction of travel (only applicable to bus journeys);
5. Vehicle number (only applicable to bus journeys);
6. Vehicle trip start time (only applicable to bus journeys);
7. Transaction timestamp; and
8. Number of travel zones in the travel card (pay-per-use passengers) or list of travel zones in the travel card (monthly travel card subscriber passengers).

The travel card readers are located within stations in the case of railways modes, and inside the vehicle in the case of buses. Hence, the attributes relating to the vehicle and direction of travel are not recorded in stations; a passenger is able to choose to board whichever train and direction available at that station. Regarding buses, although card readers are located inside moving vehicles, the stop where the transaction took place is recorded because the Andante AFC system is fully integrated with an AVL system. This integration of the AFC and AVL systems spares

the need for inferring journey origins as in some works found in the literature (e.g. Zhao, Rahbee, and Wilson 2007; Farzin 2008; Wang, Attanucci, and Wilson 2011). The AVL system is believed to assign the origin stop of the journey stage with great accuracy, because it is also used to inform passengers of the next bus stop and has been observed to perform that task reliably and consistently. Some data attributes are recorded for the main purpose of allowing on-board inspection, but are useful for estimating the destination of journeys.

The Andante AFC system data used in this research study, which also enabled the evaluation of the methodology, is the set of transaction records of the whole months of September, October, and November 2013 within vehicles of the main bus operator, called Sociedade de Transportes Colectivos do Porto, SA (STCP, www.stcp.pt). STCP runs the vast majority of routes within the city of Porto and into the surrounding metropolitan areas. STCP also runs three classic tramlines with a single vehicle each that account for a small percentage of transaction records, which will hereafter be considered part of the bus routes and fleet for the sake of simplicity. Table 4.1 summarises the number of days and transaction records in each month.

Table 4.1 Data summary

Month	September 2013	October 2013	November 2013	All 3 months
Total days	30	31	30	91
Usable days ^(a)	30	31	29	90
Weekdays	21	23	21	65
Weekend days	9	8	7	24
Public holidays	0	0	1	1
Working days	21	23	20	64
School days	13	23	20	56
STCP staff strike days	0	1	0	1
Total STCP transactions	6441366	7425903	6175182	20040572
Andante STCP transactions	5482486 (85.1% of total)	6420575 (86.5% of total)	5377621 (87.1% of total)	17280682 (86.2% of total)
Weekday average	225166	248824	224810	229886
Weekend average	83777	87203	93802	87843
Andante travel cards	367286	376165	337432	608544
Vehicles	498	496	467	505
Note: ^(a) Usable days refers to the number of complete days of data considering that in terms of transaction records a day starts at 5h00 and terminates at 4h59 of the following day				

All STCP buses were fitted with AVL equipment, which is advantageous for evaluating the methodology. At the time, two types of travel cards existed, the

Andante multi-modal and legacy STCP specific cards. The legacy cards were compatible with the Andante system and also generated boarding transactions, but have gradually been phased out and generated fewer usable data attributes. Hence, only the Andante travel card transactions were considered in this research study, totalling approximately 17 million records over the three months of data considered. The subset of data contains fewer travel card serial numbers (i.e. fewer passengers), but maintains the entire period, which is well suited for illustrating travel patterns. Figure 4.1 depicts the geographic distribution of boarding transactions in the dataset overlaid on the Andante zoned structure.



Figure 4.1 Geographic spread of boarding transactions in the subset of data and structure of travel zones. Adapted from © 2016 IEEE with new data.

The methodology requires the Andante transaction dataset to be fused with three additional data sources depicting the transport network structure, which are publicly available (Table 4.2). The first is the listing of stations and bus stops (for the dataset used only bus stops were required, 2388 in total), along with their code, zone, and location coordinates. The second is the structure of routes (in the dataset used, 70 in total), listing stations and bus stops (for the dataset used only bus stops were required), and their respective sequence in each direction of travel. The third is a matrix with the minimum number of zones crossed travelling between O-D zone pairs (in the dataset used, 18 zones in total). The relevant data from these sources was extracted on the 30th September 2013, and informed the choice of months to be used from the Andante transaction data. The lower the gap between the transaction data and the structure of routes, the lower the likelihood of errors arising from changes to the route structure and stop locations. Three months worth of transaction data were considered plenty for testing and evaluating the methodology.

Table 4.2 Network summary (dataset used)

Data	Records
Bus routes	70
Bus stops	2388
Travel zones	18

4.3. Data preparation

Previous related works reviewed in Section 2.4 have noted that AFC system data often contains transaction records with missing or illogical data attributes (Trépanier, Tranchant, and Chapleau 2007; Barry, Freimer, and Slavin 2009). Hence, a data preparation step is required prior to estimating the destination of journeys to achieve better results. The Andante AFC system data is not an exception. The proportion of transaction records in the dataset missing at least one relevant data attribute was initially 0.7%. Additionally, 2.0% of transaction records were found to have illogical values across two attributes. However, it has been possible to identify and, in 72.0% of those instances, to correct such errors from a thorough analysis that led to the creation of a data pre-processing method.

The most frequent cause of illogical attribute values of the Andante AFC system dataset is the changeover between two consecutive vehicle trips. Upon arrival at the terminus stop, bus drivers need to signal completion of the trip and the beginning of the next one if not returning the bus to the depot. That next trip often is a return in the same route, but in the opposite direction. The dataset show that passengers sometimes board the bus, initiating their journey before that changeover process is completed. This creates an illogical value in terms of the bus stop code attribute of that transaction record, because it fails to make sense that a passenger would board the bus at the terminus stop. These illogical values originating from the changeover process were mitigated by assuming with a high level of confidence that those boarding transactions should instead be assigned to the next trip in that vehicle. The direction of travel can be assigned to these journeys based on boarding transactions records created later in the subsequent bus trip.

The mitigation above reduced the proportion of transaction records in the data sample with missing or illogical values to 0.8%. This breaks down as shown in Table 4.3 into transaction records missing the vehicle trip start time attribute value, and transaction records at unknown bus stop codes or at bus stop codes that are not part of the bus route. The missing bus stop codes are likely the result of a communication failure of the AVL system, whereas the causes for the remaining cases are not entirely clear, but are not a significant share of the dataset. The transaction records that remain with missing or illogical data attributes after pre-processing are not discarded from the dataset otherwise travel patterns would become distorted due to lost journey stages. In other words, it is preferred not to estimate a destination when records are missing attributes, than ignoring those journey stages and risk estimating incorrectly. This approach will improve the overall accuracy of estimates.

Table 4.3 Irregular transaction records after pre-processing

Missing or illogical attribute	% Transaction records
Missing bus stop code	< 0.1%
Missing vehicle trip start time	0.7%
Bus stop code not part of bus route	0.1%
Total	0.8%

Adapted from © 2016 IEEE with new data

4.4. Destination estimation methodology

Despite the dataset described in the previous sections consisting only of transaction records on buses, the methodology described is applicable to multi-modal datasets. The objective of the methodology for estimating the destination of passenger journeys from entry-only AFC system data is to determine the alighting station or stop of each journey stage:

$$\hat{s}_{pjk}^a, \forall p, j, k \quad (4.1)$$

where,

\hat{s}_{pjk}^a estimated alighting station or stop of the j -th journey stage of passenger p on day k .

Indices p , j , and k are used with the same meaning throughout the upcoming equations. The methodology is primarily based on the two key assumptions found in the literature reviewed (Barry et al. 2002; Barry, Freimer, and Slavin 2009), relating to the continuity of daily travel and circularity of daily journey chains. Let a public transport route in a given direction and the set of candidate alighting route stations or stops of a passenger journey stage be respectively defined as:

$$R = \{s_i^R \mid 1 \leq i \leq n^R\} \quad (4.2)$$

$$A_{pjk}^R = \{s_i^R \mid \beta_{pjk} \leq i \leq n^R\} \quad (4.3)$$

where,

i station or stop sequence index;

R route R in a given direction;

s_i^R i -th station or stop of route R ;

n^R number of stations or stops of route R ;

A_{pjk}^R set of candidate alighting stations or stops along route R of the j -th journey stage of passenger p on day k ; and

β_{pjk} boarding station or stop sequence of the j -th journey stage of passenger p on day k .

As indicated in Section 4.2, the Andante transactions only record the route and direction of travel in the case of bus journeys. In rail-based modes the route and direction of travel are usually unknown because card readers are located outside vehicles within stations that may cater for various directions, and possibly various routes. In the case of rail-based modes, the set of candidate alighting stops shown in Equation (4.3) must combine all possible routes and directions R available at the boarding station, hence the reduction of the number of candidate alighting stops is not as effective as in buses. Equation (4.3) takes account that the destination of a journey stage must be downstream from its origin in a given route and direction. The key assumptions can be defined in the following way:

1. The most likely destination of a journey stage is the route station or stop located downstream from its own origin that is nearest to the origin of the next journey stage from that passenger; and
2. The most likely destination of the last journey stage of a day is the route station or stop located downstream from its own origin that is nearest to the origin of the first journey of the day from that passenger.

The above definitions are multi-modal, taking into account that bus stops, unlike stations, are often specific to a direction of travel. Hence, the definitions would be simpler if only related to rail-based modes. Let $d_{(s_x, s_y)}$ be the Euclidean distance between route stations or stops x and y , the key assumptions are respectively formulated as:

$$\hat{s}_{pjk}^a \leftarrow \arg \min_{s_{pjk}^a} \{d(s_{p(j+1)k}^b, s_{pjk}^a), s_{pjk}^a \in A_{pjk}^R\}, j < m_{pk} \quad (4.4)$$

$$\hat{s}_{pjk}^a \leftarrow \arg \min_{s_{pjk}^a} \{d(s_{p1k}^b, s_{pjk}^a), s_{pjk}^a \in A_{pjk}^R\}, j = m_{pk} \quad (4.5)$$

where,

- s_{pjk}^a alighting route station or stop candidate of the j -th journey stage of passenger p on day k ;
- s_{pjk}^b boarding route station or stop of the j -th journey stage of passenger p on day k ; and
- m_{pk} number of daily journey stages of passenger p on day k .

Applying the key assumptions sets out a candidate destination for each boarding transaction record (Figure 4.2), unless there is a single daily journey for a passenger ($m_{pk} = 1$) in which case the candidate of its last stage is not determined (Section 4.5 provides further detail on the reasoning for this). It was decided that the destination of single daily journey would not be estimated from journey regularity analysis as found in a previous work reviewed (Trépanier, Tranchant, and Chapleau 2007). The reasoning for this is if a passenger occasionally changes part of the daily routine, assigning a destination based on past behaviour will return unreliable results. Given that the objective of this methodology is to illustrate travel patterns, the adopted approach avoids adding bias to the travel history of passengers.

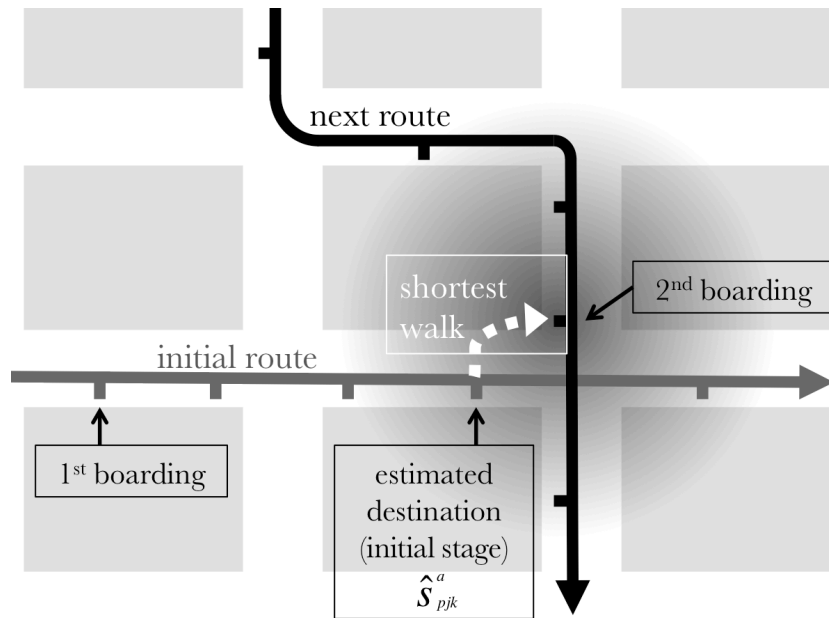


Figure 4.2 Application of key assumptions exemplified. Adapted from © 2016 IEEE.

After establishing candidate destinations, spatial validation rules are used to ascertain if the key assumptions are likely to hold for each individual boarding transaction record. The methodology proposes four endogenous spatial validation rules that can be described through the following questions. The third and fourth are newly introduced (the fourth relates exclusively to bus journeys):

1. Are the origin and candidate destination of a journey stage the same?
2. Is the candidate destination of a journey stage beyond a set Euclidean distance from the next journey origin (or from daily origin if stage is last) for that passenger?
3. Is the number of travel zones exceeded for the passenger to reach the candidate destination?
4. When a travel card is tapped on the reader more than once in the same journey stage, do any transaction records happen downstream from the candidate destination?

The first spatial validation rule has the purpose of verifying, as applicable, if the origins of two consecutive boarding transaction records from a passenger are at the approximate same location, or if the origins of the first and last boarding transaction records of the day from a passenger are at the approximate same location (Munizaga and Palma 2012). Here, an approximate same location is when one boarding origin is also the best candidate alighting station or stop to reach the other boarding origin. This is equivalent to checking the equality:

$$s_{pjk}^b = \hat{s}_{pjk}^a \quad (4.6)$$

If the answer is yes, it is hypothesised that an intermediate journey happened in an unrecorded mode of transport, because it makes no sense to board a vehicle and not go anywhere. Consequently, the destination of that journey is not estimated.

The second spatial validation rule evaluates the likelihood of the candidate destination being the actual destination of a journey stage based on walking distance (Trépanier, Tranchant, and Chapleau 2007). This is equivalent to checking the following inequalities, depending on the passenger journey being the last of the day or not:

$$d(\hat{s}_{pjk}^a, s_{p(j+1)k}^b) > c, j < m_{pk} \quad (4.7)$$

$$d(\hat{s}_{pjk}^a, s_{p1k}^b) > c, j = m_{pk} \quad (4.8)$$

where, c is the cut-off distance. This cut-off distance can be parameterised and represents the maximum transfer Euclidean distance that is considered walkable. Given the purpose of this methodology for this research, which requires a reliable O-D data source for the extraction of travel patterns, the cut-off distance has been set at 640 m. This is intentionally shorter than found in most of the previous literature. It deliberately makes the approach relatively conservative in terms of its endogenous distance-based validation to suit its main purpose, but should be appraised in light of the density of the urban fabric and of the public transport system under analysis. Moreover, the parameter value is not arbitrary. It comes from the Public Transport Accessibility Level (PTAL) methodology, representing an approximate walk catchment area of 8 minutes at 4.8 km/h assumed to be the longest distance a person would normally walk to access a bus service (Transport for London 2010). In relation to evaluating the methodology per se, the sensitivity of the cut-off distance parameter is discussed in Section 4.6 based on experimental results using the Andante AFC system dataset. Going back to the meaning of this rule, if the answer to the second question is yes, it is assumed that an intermediate journey took place in an unrecorded mode of transport and the destination of that journey stage is not estimated.

The third spatial validation rule, introduced by the present methodology, relates specifically to entry-only AFC systems with a distance-based fare structure such as Andante. Whilst Andante is divided into travel zones, other systems exist relying on stages instead, but a similar principle applies. Since the number of travel zones in a travel card is known, and assuming that passengers do not travel beyond the zones they are legally entitled to, it is possible to validate if the candidate destination falls within the allowable travel bounds. This is equivalent to checking the inequality:

$$z_{pjk} \leq z^{min}(\hat{s}_{pjk}^a, s_{pjk}^b) \quad (4.9)$$

where,

z_{pjk} number of zones in travel card of passenger p in the j -th journey stage on day k ; and
 $z^{min}(s_x, s_y)$ minimum number of zones between stations or stops x and y .

The assumption is argued to be reasonable in the context of Andante because fraud levels of any kind are very low totalling 0.49% of inspected passenger journeys (STCP 2014). This is likely a result of frequent on-board inspections and of the penalty applied to underpaid travel being one hundred times the fare. A positive answer to the third question indicates that the key assumptions do not hold for that journey and its destination is not estimated. For passenger journeys made by monthly travel card subscribers, an additional check of whether the candidate destination is within a zone featured in the list of travel zones in travel card can be made.

The fourth spatial validation rule, similarly introduced by the present methodology, owes to the abovementioned Andante characteristic of being a time-based system for pay-per-use passengers. Hence it may or may not apply to other entry-only AFC systems. In Andante, pay-per-use passengers occasionally tap their cards on a reader during a journey because it provides feedback on how much time is left for travelling without paying another fare. Other times, they do so having forgotten whether they tapped the card entering the vehicle or not. Either way, it creates a duplicate transaction record for that journey stage. This only applies to bus journeys, whereby the vehicle number is a recorded attribute in the transaction record that reveals if it is indeed a duplicate. In rail-based modes, the transaction record created by touching the card on a reader in another station is inevitably interpreted as an additional journey stage. This methodology identifies the above duplicate records on bus journeys and uses them to validate if the duplicate transaction happens downstream from the candidate destination in that bus route. This is equivalent to checking the inequality:

$$\rho_{pjk} > \hat{\alpha}_{pjk} \quad (4.10)$$

where,

- ρ_{pjk} highest route stop sequence of duplicate records in the j -th journey stage of passenger p on day k ;
- $\hat{\alpha}_{pjk}$ estimated alighting route stop sequence of the j -th journey stage of passenger p on day k .

If the inequality is verified, the answer to the fourth question is positive and the destination is not estimated because the passenger was confirmed to be travelling in the same bus trip beyond that location.

4.5. Algorithm structure and implementation

The methodology was implemented with an algorithm developed in SQL, which was applied to the 2013 Andante dataset from STCP described in Section 4.2. The results featured in Nunes, Galvão Dias, and Falcão e Cunha (2016), which this chapter is largely based on, relate to the application of the algorithm to a 2010 Andante dataset from STCP, carried out before newer data became available through the *Seamless Mobility* project (described in more detail in Section 6.4). Figure 4.3 illustrates the algorithm. The percentages shown in brackets represent the estimation rates obtained from the implementation applied to the 2013 dataset. The algorithm has linear complexity; its execution time is proportional to the number of transaction records selected. The algorithm goes through the dataset sorted firstly by travel card serial number and secondly by the transaction timestamp. The first decision of the algorithm is to verify if the transaction record is a duplicate, in which case will be used for spatial validation, but its destination will not be estimated.

The following decision (Figure 4.3) is to check if the boarding transaction record is the last or the only stage of a single daily journey for that travel card serial number. Two aspects are highlighted here. The first is the day interval definition. The dataset used revealed significant transaction levels around midnight, dropping steadily to minimums between 3:00 am and 5:00 am (Figure 4.4). Transaction records between midnight and 3:00 am appear to be largely related passenger journey chains of the previous day. Coincidentally, the shift from night time to daytime services happens around 5:00 am. Therefore, it was decided that in terms of daily journey chains a day starts at 5:00 am and ends at 4:59 am of the following morning.

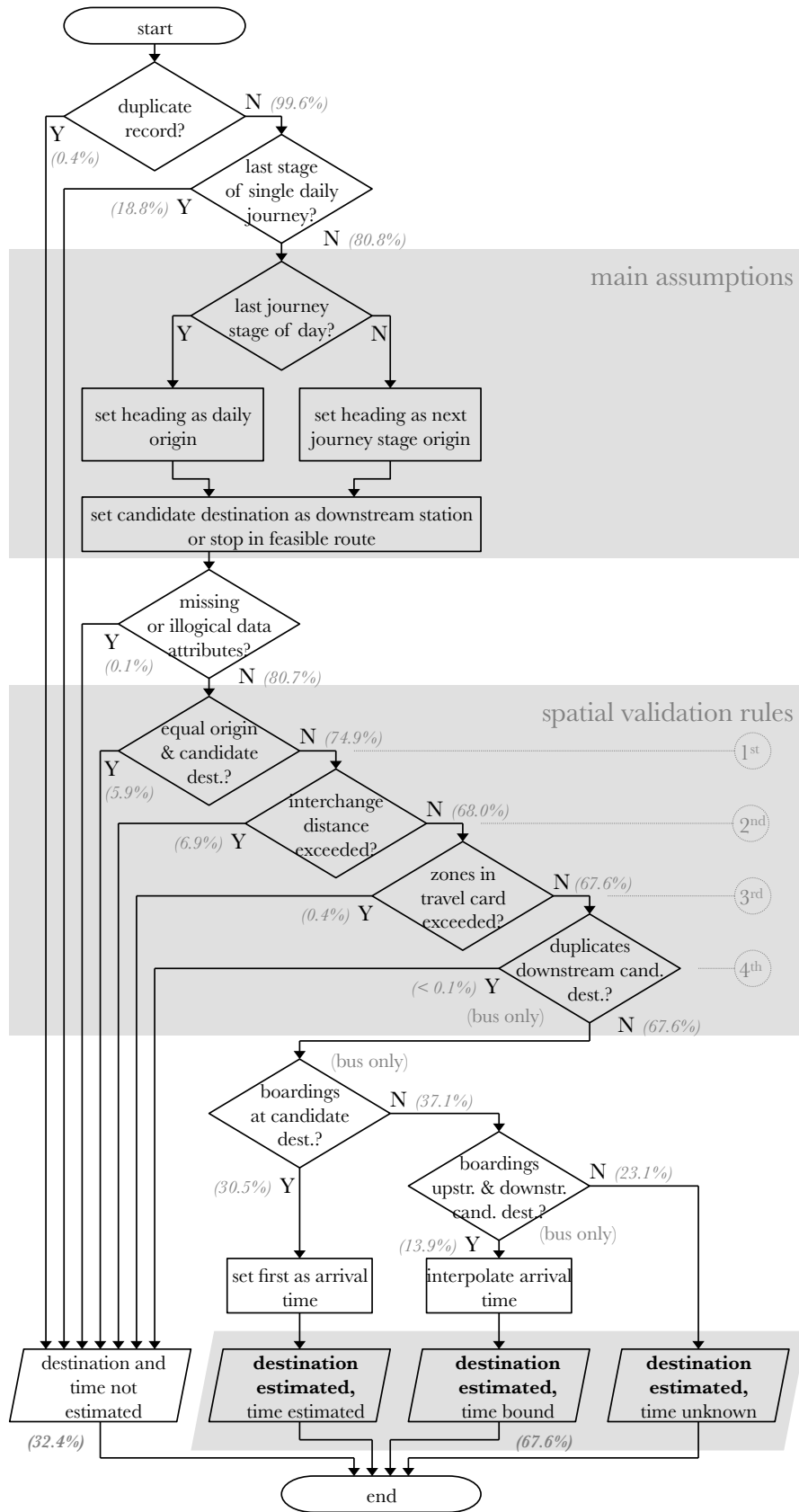


Figure 4.3 Methodology flowchart. Adapted from © 2016 IEEE with new data.

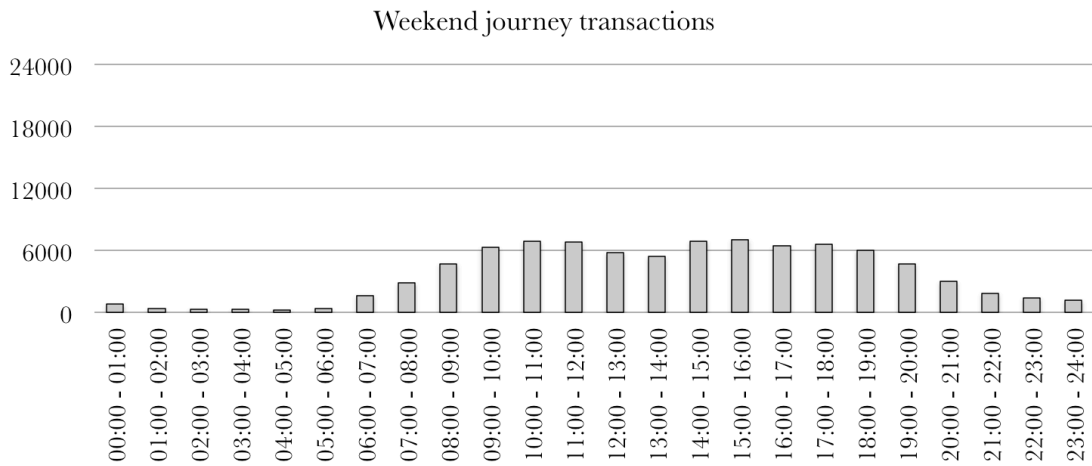
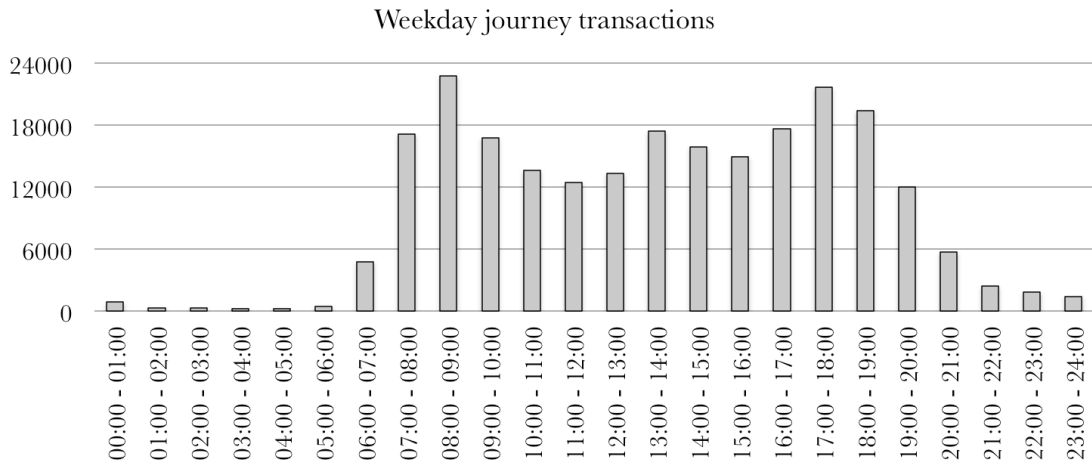


Figure 4.4 Daily average journey transactions throughout the day

The second aspect relates to the distinction between journey stages and complete journeys. Passengers often have to change between public transport routes to reach their destination and, in the case of the Andante system, tap their travel card on a reader every time they board a different vehicle. Each of those boarding transaction records relate to a stage of their complete journey. The difference matters in cases when there is a single daily journey for a passenger. If that journey is single staged, it is trivial that a destination cannot be estimated due to lack of information to determine a candidate destination. But if that journey has several stages, it is arguably most likely that the last journey stage was to reach a destination other than the daily origin (left in Figure 4.5), otherwise the passenger would be travelling in a circle (right in Figure 4.5). Both of these scenarios are possible in theory, but simply assuming the latter for every instance carries great risk of estimating the destination incorrectly. Therefore, a candidate for that destination is not determined rather than assuming it to be the daily origin as seen in previous literature. This is to support the highest accuracy of estimates. It is reminded that the stages of a complete journey are defined by the time-based Andante rules for pay-per-use passengers, which set out maximum journey durations according to the number of travel zones.

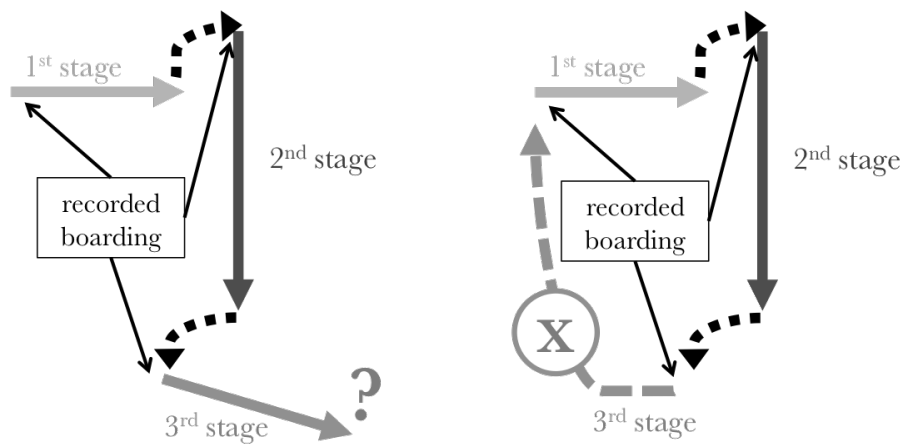


Figure 4.5 Single daily journeys with multiple stages (three stage example). Adapted from © 2016 IEEE.

The following decision (Figure 4.3) is to determine whether the boarding transaction record relates to the last stage of the day for that passenger or not. This determines

which of the key assumptions, continuity of daily travel or circularity of daily journey chains, should apply for setting the candidate destination. This is followed by verification if the origin and candidate destination station or stop codes are both present and are logical in the boarding transaction record, or else its destination cannot be estimated. This cannot be fulfilled earlier because the candidate destination is not determined prior to the application of key assumptions. If all required data attributes are present, the boarding transaction record goes through the four endogenous spatial validation rules described, which deal respectively with the inequality of origin and destination, the maximum interchange distance, the adequacy of the number of zones in the travel card, and the inexistence of duplicate records downstream from the candidate destination (only applicable to bus journeys). When a boarding transaction record survives all four spatial validation rules, the destination of its journey stage subsequently estimated with great confidence.

Although the estimation of arrival time is not paramount for the purposes of this methodology, the next step, which is only applicable to bus journeys, is to verify if there are boarding transaction records in the same service at the estimated destination stop. If this is true, the earliest timestamp of those transaction records is assigned as the arrival time. If not, an additional step verifies if there are boarding transaction records in the same bus service both between the origin and the estimated destination, and after the estimated destination. The presence of such transaction records allows the definition of arrival time upper and lower bounds, and the interpolation of an estimated arrival time using the number of stops in between as a weighing factor. The absence of enough boarding transaction records in the same bus service renders the estimation of the arrival time unfeasible without exogenous AVL data.

4.6. Discussion of results

The initial implementation of the algorithm was executed in a computer with a 1.6 GHz Intel Core i5 processor and 4 GB of Random Access Memory (RAM). The results were obtained for the entire 2013 Andante dataset from STCP, but with

live data, it can be executed incrementally on a daily basis in approximately 30 minutes for the magnitude of the dataset. Application of the methodology to the dataset yielded a percentage of estimated destinations of 67.6% of transaction records. This result is partially dictated by the nature of the STCP bus data, particularly in terms of the amount of single daily journeys that are an expression of travel behaviour. But it is heavily influenced by the application of strict validation rules to identify individual records for which the key assumptions, relating to the continuity of daily travel and to the circularity of daily journey chains, do not hold.

To put the above results into perspective, if the endogenous spatial validation rules described in Section 4.4 were not applied, and if the proposed distinction between journey stages and complete journeys were ignored, the methodology would yield a percentage of estimated destinations of 84.8%, generally in line with some of the previous works reviewed. However, the accuracy of the additional estimates (approximately 17% of the total) would be compromised. This highlights one contribution of the proposed methodology, avoiding estimation errors by identifying and dealing cautiously with single daily journeys with multiple stages. This approach favours highest accuracy over the percentage of estimated journey destinations.

Table 4.4 breaks the results down, and demonstrates the absence of major variations on a monthly basis. This paragraph looks into the aggregate results over the three-month period. Boarding transaction records that survive validation total 67.6%, and subdivide into those where the arrival time was estimated or bound (44.5% of total), and those for which there was not enough information to estimate an arrival time (23.1% of total). The remaining transaction records total 32.4%. The majority of these are either the only or last stage of single daily journeys (18.8% of total), many fail spatial validation (13.1% of total), and some have data attribute errors (0.1% of total) or are duplicates (0.4% of total). The methodology applies the spatial validation rules in the order specified above, otherwise some boarding transaction records could fail more than one validation simultaneously. An example would be a boarding transaction record for which the maximum interchange distance was exceeded and so were the number of zones in the travel card. In such case, the former spatial validation rule is considered the cause for the failure in estimating a destination.

Table 4.4 Results of the methodology applied to the Andante dataset

Result	% Transaction records			
	Sep 2013	Oct 2013	Nov 2013	Total
Boarding transaction records with estimated destination	67.9%	67.8%	67.1%	67.6%
Arrival time estimated	28.8%	31.7%	31.0%	30.5%
Arrival time bound	13.6%	14.0%	14.1%	13.9%
Arrival time unknown	25.5%	22.1%	22.0%	23.1%
Transaction records with destination not estimated	32.1%	32.2%	32.9%	32.4%
Duplicate transaction records	0.5%	0.4%	0.4%	0.4%
Single stage of single daily journeys	14.8%	15.1%	15.8%	15.2%
Last stage of single daily journeys	3.7%	3.5%	3.6%	3.6%
Missing or illogical data attributes	<0.1%	0.1%	0.1%	0.1%
Spatial validation, 1 st rule: inequality of origin and destination	5.8%	5.9%	5.9%	5.9%
Spatial validation, 2 nd rule: maximum interchange distance	6.9%	6.9%	6.9%	6.9%
Spatial validation, 3 rd rule: number of zones in travel card	0.5%	0.4%	0.4%	0.4%
Spatial validation, 4 th rule: duplicate records downstream destination	<0.1%	<0.1%	<0.1%	<0.1%
(Boarding transactions with estimated destination that have duplicates and fail the 4 th spatial validation rule)	(1.2%)	(1.2%)	(1.7%)	(1.3%)

Adapted from © 2016 IEEE with new data

The variations in key results between months were not very expressive. The maximum variation was 0.8% between September 2013 and November 2013 in the percentage of boarding transaction records that survived validation. In terms of the subdivisions of this results, the maximum variations were 3.3% between September 2013 and October 2013 in the percentage of boarding transaction records with an estimated or bound arrival time, and 3.5% between September 2013 and November 2013 in the percentage of boarding transaction records not having enough information to estimate an arrival time.

Sensitivity analysis of the cut-off Euclidean distance parameter was carried out for the boarding transaction records from October 2013, and is shown in Table 4.5. If that distance were made stricter by reducing it from 640 m to 400 m as in Zhao, Rahbee, and Wilson (2007), the percentage of boarding transaction records failing the second spatial validation rule would increase to 9.7% and the percentage of estimated destinations would drop to 65.0%. Conversely if the parameter were made more liberal by increasing it to 1000 m as in Munizaga and Palma (2012), those percentages would respectively drop to 5.1% and increase to 69.6%. The variation is

significant, however the parameter set was felt to be an adequate trade-off between the risks of rejecting true positives (rejecting a correct candidate destination) that is greater with a short cut-off distance, and accepting false positives (accepting an incorrect candidate destination) that is greater with a longer cut-off distance.

Table 4.5 Sensitivity analysis of the cut-off Euclidean distance parameter

Parameter value	% Boarding transaction records from October 2013	
	Failed interchange distance	Destination estimated
400 m	9.7%	65.0%
640 m	6.9%	67.8%
1000 m	5.1%	69.6%

Adapted from © 2016 IEEE with new data

Another contribution of this methodology is the introduction of the third and fourth spatial validation rules. Their main purpose is testing at a maximum disaggregation level, which is a single journey stage, if the two key assumptions underlying most O-D estimation methodologies in the literature apply to a given case study. Each public transport system has specific usage patterns; therefore the fit of key assumptions may vary and should be tested on a case-by-case basis. Previously, those assumptions had mostly been validated at an aggregate level using external data from O-D surveys and exit counts. However, there is a risk concerned with the errors tending to average out when the assumptions are tested at aggregate level and there is no bias. For example, if the assumptions were flawed, an overestimate of journeys from origin A to destination C may compensate an underestimate of journeys from origin B to destination C. Despite this risk, endogenous and exogenous validation rules should ideally have been combined. It was not possible for the Andante dataset due to the inexistence of a recent O-D survey or Automatic Passenger Count (APC) devices in STCP buses. The latter had reportedly been trialled, but the operator found the APC technology unreliable and decided against its use.

According to the results shown previously in Table 4.4, the newly introduced spatial validation rules supported the validity of the key assumptions for the vast majority of journeys in the 2013 Andante dataset from STCP. The percentage of boarding transaction records that fail the third and fourth spatial validation rules is reduced. Particularly the fourth spatial validation rule that deals with duplicate transactions

may at first seem unnecessary for the dataset used, since it highlighted less than 0.1% of boarding transaction records that survived previous validation rules. But this result must be read with caution. One should look at the percentage of boarding transaction records that failed the duplicate record validation rule in relation to those where the destination was estimated and had at least one duplicate: 666 out of 49626, approximately 1.3%. This low value provides evidence that no reasons exist to believe that the key assumptions are not valid for the vast majority of boarding transaction records. This claim applies as long as the first and second spatial validation rules dealing with the inequality of origin and destination, and with the maximum interchange distance are carried out beforehand.

It is noted that the methodology estimates the destination of passenger journeys at maximum disaggregation level, which allows for the construction of O-D matrices at any desired level of aggregation and geographic coverage. Lastly, a characteristic of the dataset worth mentioning is that complete journeys were composed by 1.27 stages on average.

4.7. Multi-modal upgrade

The destination estimation methodology was eventually incorporated into the *Seamless Mobility* project (described in more detail in Section 6.4) to provide an O-D data source. New requirements emerged in that context for the algorithm to be upgraded to deal with multi-modal data, and to be rewritten in Java scripting language. The upgrade was carried out in collaboration with other members of the *Seamless Mobility* project team. The multi-modal data consisted of 2013 Andante transaction records from STCP buses and from the light-rail operator Metro do Porto (www.metrodoporto.pt). A multi-modal dataset should expectedly reduce the number of transaction records that are the only or last stage of single daily journeys, and the number of boarding transaction records that fail the first spatial validation rule. The reason for this is that further journeys may exist for those passengers within the Andante system with other operators. From this point of view, applying the

methodology to a multi-modal dataset is expected to yield an increased percentage of boarding transaction records that survive validation.

The multi-modal Java implementation of the algorithm has been executed and results obtained for a subset of the 2013 Andante multi-modal dataset. The multi-modal implementation yielded a percentage of boarding transaction records for which a destination was estimated of approximately 65%, based on a cut-off Euclidean distance parameter of 640 m. This result seems unexpectedly lower than the equivalent on the STCP operator data alone. At least three factors may have influenced this result. First, the multi-modal dataset includes passengers that are not STCP users, and the nature of their travel behaviour with Metro do Porto may lead to a lower percentage of estimated destinations that offsets advantages of a multi-modal dataset. Second, while the multi-modal dataset includes the largest bus and light-rail operators in Porto, the Andante system includes nine additional operators for which data has not been made available. Boarding records with these operators contain parts of the travel patterns of STCP and Metro do Porto users, influencing the results. Third, a known data issue existed relating to duplicate station and stop identifiers across the two operators. It impacted on the accuracy of some destination estimates and possibly prevented an improvement to the results. It is noted that the issue only applied to the multi-modal implementation.

4.8. Chapter conclusions

This chapter described a methodology for estimating the destination of passenger journeys from AFC data. It builds on previous work found in the literature by replicating key assumptions, but applies specifically to the case of entry-only systems with a distance-based fare structure, which had not been addressed before. The ultimate purpose of the methodology for this thesis is to provide an enriched journey dataset to feed the second tool for the model described in Chapter 5. But, additionally, the methodology itself makes two specific contributions to the literature.

The first contribution consists of the new endogenous spatial validation rules. These additional rules deal with the number of zones or stages in a travel card, which is

specific to distance-based fares, and with the existence of duplicate transaction records. Their purpose is to test the validity of key assumptions regarding continuity of daily travel and circularity of daily journey chains, on a single case basis and at maximum disaggregation level. For the dataset used, the spatial validation rules were not prolific in the identification of false positives that remained unspotted from previous validation steps, but did support the validity of the key assumptions. The second contribution consists of improved reliability of estimation results. The methodology refines previous work by distinguishing between journey stages and complete journeys and subsequently not estimating the destination of the last stage of single daily passenger journeys with multiple stages. Such instances otherwise introduce a great deal of uncertainty to the estimation results.

Andante AFC system data was introduced and has been a new addition to the O-D matrix estimation literature (Nunes, Galvão Dias, and Falcão e Cunha 2016). The methodology proved effective to estimate the destination of journeys at disaggregate level and to detect instances where the candidate destination obtained from the application of key assumptions is likely incorrect. The approach towards these instances is conservative; their destinations are not estimated. The percentage of estimated destinations is partially dictated by the nature of the Andante AFC system data, but mostly by the strictness of validation rules seeking the highest accuracy of estimates. Future work may focus on exogenous validation of the methodology once up-to-date O-D survey results become available. Future improvements to the methodology may perhaps include an additional validation rule based on an interchange time interval as found in Gordon (2012). This has not been done before due to limited usefulness, given the low percentage of trips with estimated arrival time and multiple stages.

5. Temporary User-Centred Networks (TUNs)

This chapter describes the concept of Temporary User-Centred Networks (TUNs) for passenger transport systems, and its application to the enriched AFC system dataset obtained from the algorithm described in Chapter 4. TUNs are the second tool and core element of the user collaboration model (see Figure 3.2), linking users based on circumstantial collaboration potential. This chapter describes the mathematical formulation and the algorithm that performs the required calculations for the dynamic creation of TUNs. This chapter also reflects the initial and primary interest of this thesis in the urban public transport domain. Alternative data sources and passenger transport domains (e.g. private car use and cycling) may require adjustments to specific aspects relating to the implementation of the proposed algorithm.

This chapter is organised as follows. Section 5.1 introduces the concept of TUNs and the notion of spatiotemporal relevance between users of a passenger transport system. Section 5.2 provides the mathematical formulation of TUNs, drawing upon the topic of interpersonal affinity in urban passenger transport previously introduced in Chapter 3. Section 5.3 describes the structure of the algorithm, while Section 5.4 provides a description of the subset of data used to evaluate the potential of the user collaboration model. Section 5.5 describes the initial implementation of the algorithm, followed by results of its application to the subset of data in Section 5.6. Section 5.7 describes efficiency improvements applied to the initial implementation of the algorithm, carried out in the context of the *Seamless Mobility* project. Lastly, Section 5.8 summarises the conclusions drawn.

This chapter is partially based on Nunes, Galvão Dias, et al. (2016), “Temporary User-Centred Networks for Transport Systems”, Transportation Research Part C: Emerging Technologies.¹

¹ Reprinted from Transportation Research Part C: Emerging Technologies, 62, António A. Nunes, Teresa Galvão Dias, Chris Zegras, and João Falcão e Cunha, Temporary user-centred networks for transport systems, 55-69, Copyright (2016), with permission from Elsevier.

5.1. Concept

TUNs aim to identify circumstantial user-based collaboration opportunities to facilitate the diffusion of information spread across a passenger transport system in real-time. TUNs materialise spatiotemporally dependent ties between users, by means of affinity measures that are tangible and quantifiable. TUNs differ from existing ICT-based social networking services where the spatial element is absent or, in fewer cases, stationary (e.g. to connect guests at a local attraction or event). TUNs reveal latent social structures that are typically invisible to their users, in which interpersonal affinity is represented by spatiotemporal equivalence of travel patterns. The designation TUNs reveals two key characteristics of these networks. First, they are temporary, existing for the duration of a journey, and dynamically changing as users join and leave the passenger transport system. Second, they are centred on users, because users have unique networks based on their individual travel behaviours in real-time.

TUNs represent affinity as spatiotemporally dependent ties between users in a transport system. TUNs offers a solution to the problem of selecting, for any user, other users (peers) with the highest collaboration potential within a passenger transport system at a specific instant. The creation of a TUN therefore requires the identification and ranking of relevant peers of a user in real-time. In this context, relevance is an indicator of affinity strength between peers. Section 2.5 of this thesis previously argued that this problem shares some basic characteristics to that of searching for information on the Internet, but noted that the inherently subjective notion of relevance is shifted from the actual information to the ranking of users who generate that information.

Two key underlying premises guide the creation of TUNs. The first premise is that *a user will likely derive greater benefits from collaborating with peers that either have similar travel behaviours, or experience feasible alternatives to achieve the same travel goal.* The reasoning behind this premise is that a user should benefit the most from information relating to a planned route or from alternatives to that route (e.g. getting simultaneously a disruption warning for the usual commuting route and conditions of alternative

routes, to inform a travel choice). It implies that a user will be relevant to a peer travelling simultaneously if at least a part of their journeys is bounded by two shared locations, irrespective of their paths. The second premise is that *the larger the proportion of the journey of a user that is bounded by two locations shared with the journey of a peer, the higher is the relevance of that peer*. It implies that relevance between peers is asymmetric, because the users may likely be spanning different distances in their journeys. The relevance of the longer journey to the shorter journey is greater than the opposite.

This second premise alone renders some similarity metrics reviewed in Section 2.5 unusable for assessing similarity of journeys in the creation of TUNs. For example, a longer trip would be deemed dissimilar to a shorter one contained by it, if similarity were estimated based on Euclidean distance, Hausdorff distance, or Fréchet distance. The first and second premises combined render the clustering methods reviewed in Section 2.5 unusable in the creation of TUNs. Two journeys may neither share origins nor destinations, and follow different paths, and still remain significantly relevant to each other. Yet, a clustering algorithm would hardly identify relevance in such circumstances. Furthermore, the first premise reduces the usefulness of the POI-based similarity metric reviewed in Section 2.5 (Hwang, Kang, and Li 2005), because it would fail to cater for feasible alternatives to achieve a travel goal. Lastly, whilst the reviewed approaches to the carpool service problem (Section 2.5) are significantly closer to a solution for the creation of TUNs, they are computationally demanding for handling the real-time dynamics of these networks. All limitations of existing methods highlighted in this paragraph will become clearer in the following Section 5.2, which proposes and defines two measures of affinity that added together estimate spatiotemporal relevance. Calculating these measures in real-time is computationally less demanding, as they leverage the fact that urban passenger transport journeys are not unrestricted in space, having to follow predefined network grids.

5.2. Mathematical formulation

Affinity in the context of TUNs is the spatiotemporal degree of equivalence between the travel patterns of users in the passenger transport system. The expression ‘travel pattern’ generalises three types of instances, formally defined by the mathematical Intervals 5.1 to 5.3, which differ in relation to their time frame:

1. Travel history: past journeys (Interval 5.1);
2. Ongoing journeys: either occurring or imminently about to start (Interval 5.2); and
3. Travel intentions: tentative future journeys that have been predicted or somehow been expressed by travellers (Interval 5.3).

$$]-\infty, T - b[\quad (5.1)$$

$$[T - b, T + a] \quad (5.2)$$

$$]T + a, +\infty[\quad (5.3)$$

where,

T present time instant;

b period set before T for journeys with stages that started but are not known to be completed (constant, zero if the end time of journeys is known); and

a period set after T for journeys that are imminently about to start (constant).

The travel history may be captured, for example, from usage of AFC media, GPS traces, Electronic Toll Collection (ETC) records, or any other additional means reviewed in Section 2.4. The same applies to ongoing journeys, although in some cases it may be required predicting the destination of uncompleted journeys based on travel history. Travel intentions, which are tentative, may be obtained or inferred, for example, from travel ticket purchases, usage of journey planning applications, or from geo-referenced appointments recorded in electronic agendas. The travel history is useful both for predicting the destination of ongoing journeys if not known (e.g. rail tickets are often purchased for a specific O-D pair, in which case both origins and destinations are known in principle), and to predict future journeys. The availability of all instances of travel pattern data provides the necessary information for the

creation of TUNs based on the identification and ranking of relevant peers to a user in real-time. While in real-time this needs to account for the specificities of a passenger transport system, for the purpose of evaluating the potential of the model, all types of travel pattern instances are simulated from historical data by considering ‘present time’ to be a specific time instant in the past (see Section 5.4).

It is noted that some potential uses of TUNs may be exceptions to the above, not requiring their creation in real-time, in which case travel history data is likely enough. Examples include the identification of recurrent spatiotemporal clusters of car drivers to find potential demand for shared transport routes, and of cyclists for organising group rides and finding cycling buddies. In contrast, other potential uses of TUNs are also exceptions for relying almost solely on the prediction of future journeys. An example is the identification of imminent user clusters, prior to their occurrence, for the deployment of DRT vehicles. Apart from these exceptions, TUNs were developed with the primary interest of facilitating circumstantial collaboration between passenger transport users in real-time, which indeed requires all aforementioned instances of travel pattern data.

The relevance score between two users simultaneously in the system is proposed as the sum of two affinity measures, called journey similarity (SIM) and journey substitutability (SUB), which are defined to be mutually exclusive over a given portion of a journey path. This reflects the first key underlying premise set out above in Section 5.1.

Journey similarity (SIM)

SIM represents the portion of simultaneous journey paths of two users that share spatial characteristics (Figure 5.1, top). SIM measures the degree of similarity between journey paths. In public transport, similar journey paths are those of two users travelling along a common portion of the same route and in the same direction. In the context of private transport, whether by motorised vehicle or bicycle, similar journey paths are those of two users who share the same routing between two locations. As per the second key underlying premise set out above in Section 5.1, the proposed SIM measure is asymmetrical. It looks at the portion of a journey path that is contained by

another similar journey path, and therefore it is a function of the overall length of the first. For example, the path of user 1 may be entirely similar to the path of user 2 but the converse may not hold. This would be the case where the path of user 1 is longer and contains the path of user 2. In public transport, for example, passenger 1 would have the information about the entire length of the journey of passenger 2, whereas passenger 2 would only be partially relevant to passenger 1.

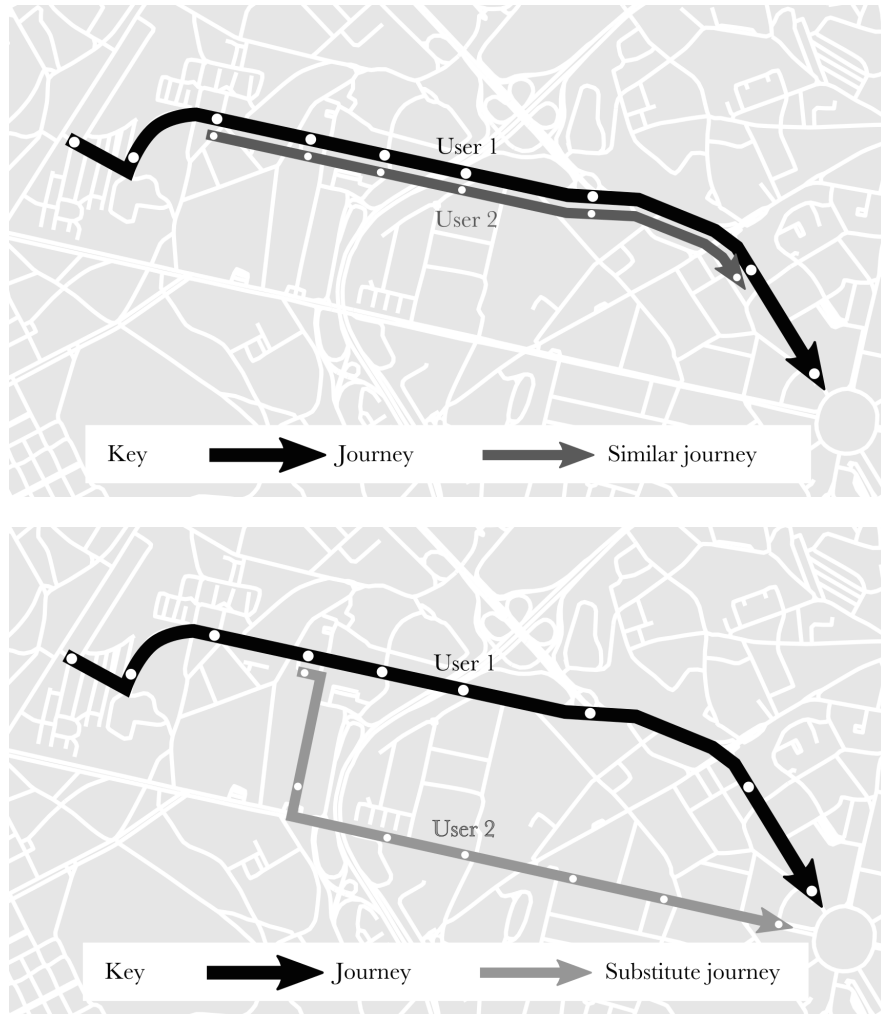


Figure 5.1 Journey similarity (SIM) and journey substitutability (SUB). Adapted from Nunes, Galvão Dias, et al. (2016).

The SIM between users in a passenger transport system is represented by a square matrix \mathcal{S} , the dimensions of which are given by the number of users simultaneously in the system. The entries of matrix \mathcal{S} at a given instant t are determined by Equation

(5.4). The SIM of one user to another equals the common length of their journey paths divided by the total length of the first. Hence matrix \mathcal{S} is asymmetric. The SIM of two journey paths is a ratio that varies between 0 (dissimilar) to 1 (perfectly similar). Lengths can be expressed in actual distance, Euclidean distance, or in journey segments (the span between two stations or stops), depending on the available data, purpose, and domain of application.

$$s_{ij} = \frac{c_{ij}}{l_i} \quad (5.4)$$

where,

- i journey path of the user;
- j journey path of a peer;
- s_{ij} SIM of j to i ;
- c_{ij} common length between i and j ; and
- l_i length of i .

Journey substitutability (SUB)

SUB represents the portion of simultaneous journey paths of two users where SIM is zero, but they serve at least two of the same potential Origin-Destination (O-D) pairs (Figure 5.1, bottom). SUB measures the degree of path substitutability. In public transport, substitute journey paths are those of two users travelling along a different route or direction but having at least two Shared Potential Origins or Destinations (SPODs). In the context of private transport, substitute journey paths are alternative routings between two SPODs. As per the second key underlying premise set out above in Section 5.1, the proposed SUB measure is also asymmetrical, for the same reasoning applied to the SIM measure. Simultaneous journey paths for two users may present a combination of SIM and SUB, however portions with a SIM score cannot have a SUB score and vice-versa. In other words, affinity measures are defined to be mutually exclusive over a given portion of a journey path, as illustrated by Figure 5.2: the first journey stage of user 2 is similar to a portion of the single staged journey of user 1; the second journey stage of user 2 is a substitute of a downstream portion of the single staged journey of user 1.

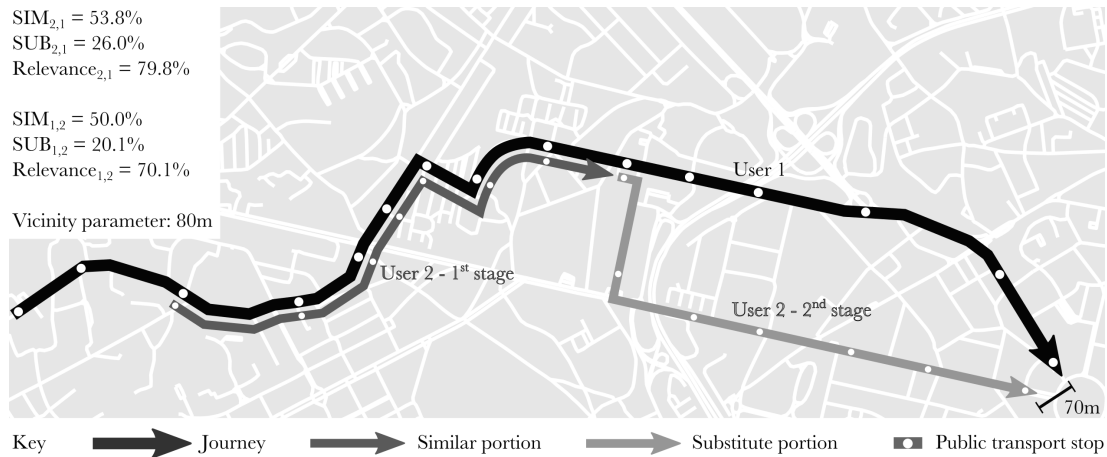


Figure 5.2 Journey with SIM and SUB scores. Adapted from Nunes, Galvão Dias, et al. (2016).

The SUB between users in a passenger transport system is represented by a square matrix \mathbf{B} , the dimensions of which are given by the number of users simultaneously in the system. Calculating the entries of the SUB matrix \mathbf{B} at a given instant t is more complex, for two reasons.

The first reason for the additional complexity in calculating SUB is the need to define precisely the concept of Shared Potential Origin or Destination (SPOD) mentioned earlier. This need is best illustrated with an example. Consider alternative routes for a public transport passenger in two different modes, bus and train. The bus departs from a stop adjacent to the train station. The departing locations, while not the same, are near enough to be considered a shared origin, so both routes are true substitutes. This instance can be generalised for private transport since alternative routes may not overlap yet still have nearby passing points (i.e. SPODs). The extent to which alternative routes actually share a potential origin or destination, depends upon the proximity of their respective origins or destinations. The calculation of SUB takes account of proximity based on the notion of vicinity, akin to a catchment area. *Vicinity is a fixed parameter representing a cut-off distance within which two points may be considered to represent a SPOD.* Nonetheless, the real substitutability of alternative routes decreases as the distances between points in the SPOD grow, reducing the degree to which the points share origin or destination potential and, thus, reducing the substitutability of the routes. So, even if two points are in the vicinity of each other,

the resulting SUB of route alternatives is inversely proportional to the distance between the points. For that reason, the proposed SUB measure incorporates a penalty factor, a function of distance among points within that vicinity.

The second reason for the additional complexity in calculating SUB is that substitute journey paths may have more than two SPODs. In such cases, various combinations have to be considered, because considering just the first and the last SPODs between two journey paths could underestimate SUB. This calculation is sensitive to the penalty factor, described above. Thus, SUB is influenced by a trade-off between lengths of the substitute portions and penalty factors (see Figure 5.3). The entries of matrix \mathbf{B} at a given instant t are determined by Equation (5.5).

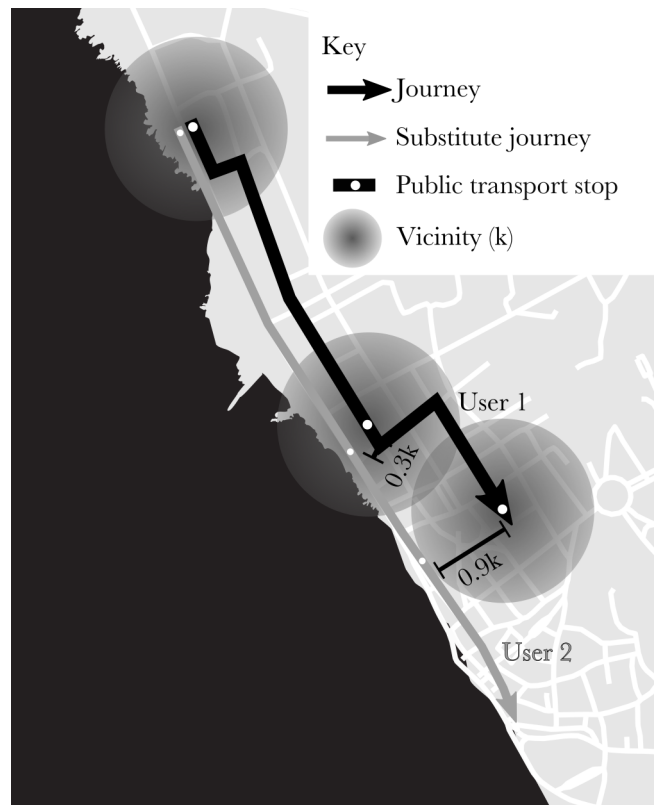


Figure 5.3 Journey paths with maximum SUB at shorter substitute portion. Adapted from Nunes, Galvão Dias, et al. (2016).

$$b_{ij}: \text{Max} \left[\frac{p_{ij}}{l_i} * \underbrace{\left[1 - \frac{(d_{o_{ij}} + d_{d_{ij}})}{2k} \right]}_{\text{penalty factor}} \right] \text{ subject to } d_{o_{ij}} \leq k, d_{d_{ij}} \leq k \quad (5.5)$$

where,

- i journey path of the user;
- j journey path of a peer;
- b_{ij} SUB of j to i ;
- p_{ij} substitute portion between i and j ;
- l_i length of i ;
- $d_{o_{ij}}$ distance between origins;
- $d_{d_{ij}}$ distance between destinations; and
- k vicinity parameter.

The SUB between users is given by a ratio multiplied by the penalty factor, as shown in Equation (5.5) The ratio is the substitute portion of the journey paths of the users divided by the total length of the first. Hence matrix \mathbf{B} is also asymmetric. The ratio depicts substitution potential, the portion of a journey that the can be substituted by the alternative. The penalty factor is the sum of distances between the location points at the SPODs of the substitute portion, divided by vicinity doubled and subtracted from 1. The penalty factor tends to 0 as the distances between location points at the SPODs get close to the set vicinity parameter, k , and equals 1 when those distances are null. This accounts for the reduction of the convenience of an alternative due to the necessity to detour. The multiplication in Equation (5.5) ensures that the penalty is applied proportionally to the substitution potential of journey paths. When the penalty factor equals 1, no penalty is applied. Figure 5.4 illustrates how SUB varies according to the distances between location points within SPODs for various levels of substitution potential.

Considering the first and the last SPODs between two journey paths may not return the highest value for SUB. The mathematical maximisation shown in Equation (5.5) avoids underestimating the true SUB by considering all candidate pairs of SPODs between two journeys that may yield the optimum SUB. The SUB of two journey paths varies between 0 (not substitute) to 1 (perfect substitute). As in SIM, lengths can be expressed in actual distance, Euclidean distance, or in journey segments (the span

between two stations or stops), depending on the available data, purpose, and domain of application.

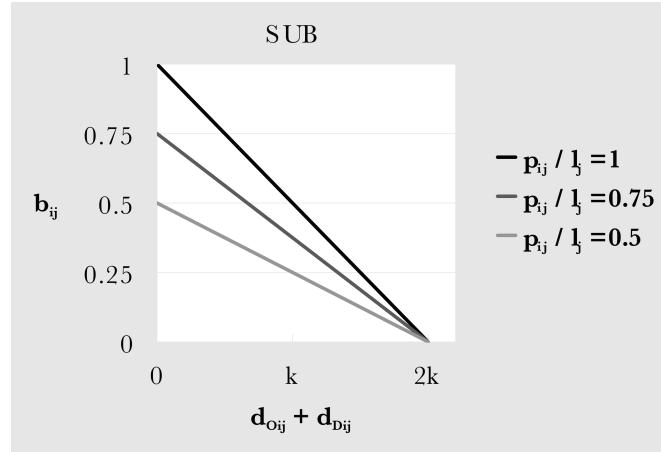


Figure 5.4 SUB according to distance between location points within SPODs for various levels of substitution potential. Adapted from Nunes, Galvão Dias, et al. (2016).

Relevance

Relevance in TUNs provides an indicator of the spatiotemporal equivalence between simultaneous journey paths of two users. Relevance estimates the degree to which simultaneous travel patterns of two users either share or have alternative spatial characteristics. The relevance score between users in a passenger transport system is represented by a square matrix \mathbf{R} , the dimensions of which are given by the number of users simultaneously in the system. The relevance score is obtained for a given instant t as the sum of the affinity measures SIM and SUB, calculated separately, as shown in Equation (5.6). Being the sum of asymmetric matrices \mathbf{S} and \mathbf{B} , \mathbf{R} is asymmetric too. Since SIM and SUB are mutually exclusive over a given portion of a journey path, the relevance score will range between 0 (not relevant) to 1 (perfectly relevant).

$$\mathbf{R} = \mathbf{S} + \mathbf{B} \quad (5.6)$$

where,

- \mathbf{R} relevance matrix;
- \mathbf{S} SIM matrix; and
- \mathbf{B} SUB matrix.

Relevance is fundamental for creating TUNs. In fact, for users of a passenger transport system, TUNs are created based on the highest relevance scores. Different cut-off rules and their parameters may be defined for this. A cut-off rule may be a minimum relevance score or a maximum number of users ranked by their relevance to the user, or a combination of both. The score and number of users are the parameters to be defined. The magnitude of a TUN will reduce with the strictness of the cut-off rules. For example, a TUN for a user may be defined to include all other users with relevance higher than 0.5, but limited to the top ranking 100 users. This process selects the spatiotemporally most relevant peers to be included in a TUN, and must be continuously updated to consider the inherent dynamics of the urban passenger transport system.

The relevance score reveals potentially useful, yet latent, social structures based on spatiotemporal equivalence of travel patterns. SIM instances may sometimes be visible over time to commuters, those who undertake regular or often repeated trips typically between home and work or school. Commuters may notice, for example, familiar faces in public transport. But many other SIM instances are not recurrent and most are likely unnoticeable. SUB is a different case, more likely invisible to commuters because the alternative routes do not share space, only SPODs. Hence, the combined relevance score represents fluid socialities (Büscher, Urry, and Witchger 2011) based on actual, yet largely invisible affinity, enabling collaboration opportunities, and facilitating user engagement.

Note, however, that the quantification of affinity through measuring relevance will not suffice; ties and communication channels between users must also be created to fulfil the potential of TUNs. Therefore, materialising engagement requires social ICT-based applications such as that in the case study *Journata* (the focus of Chapter 6), which either enable information sharing or reveal other collaboration opportunities between highly relevant users in real-time, whilst preserving their safety and privacy. Such collaboration between users may be voluntary or automatic, depending on whether they actively provide information or passively allow information extraction from sensors attached to them, like accelerometer data from the smartphones of travellers. Nonetheless it is stressed that the case study *Journata*

mentioned above is a single instance of application. The user collaboration model, both in terms of its conceptual foundations and implementation tools, is intended as a basis to facilitate the development of further services of its kind.

5.3. TUNs algorithm structure

This section presents the structure of the algorithm designed to calculate the spatiotemporal relevance of all users who are simultaneously present in a passenger transport system in real-time. The output of the algorithm is the relevance matrix \mathbf{R} . This output enables the definition and ranking of users to be included in each TUN. Relevance based ranking may be useful in some cases to predict the usefulness of information relayed to users, and order it accordingly. The proposed algorithm is divided into four main activities (Figure 5.5), which are detailed in the following subsections. The first of these activities consists of obtaining and filtering real-time data required for creating and updating TUNs. The second and third are performed in parallel and consist respectively of calculating the SIM and SUB matrices \mathbf{S} and \mathbf{B} . The fourth consists of compiling relevance to obtain matrix \mathbf{R} .

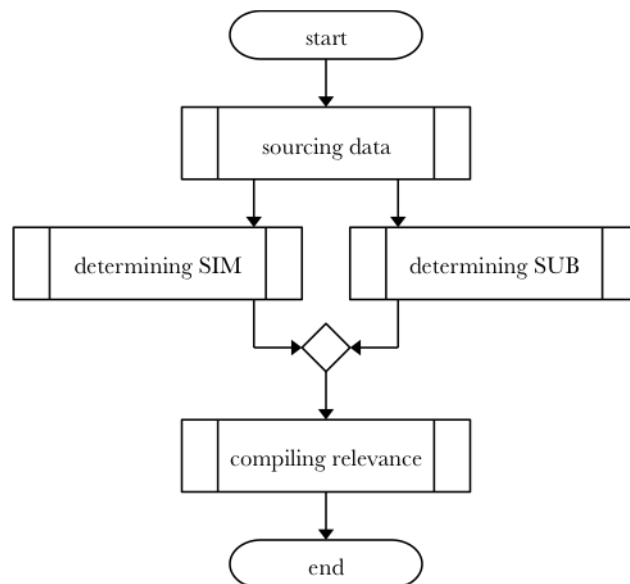


Figure 5.5 Flowchart of main activities

Sourcing data

The selection of data initiates with defining the time instant T , featured in Intervals (5.1) to (5.3). The time instant T may be in the past, be the present, or be in the future. As described in the previous section, this choice depends on the objective set for a given application of TUNs. But in any case, as T increments over time, travel intentions turn into ongoing journeys, and eventually become travel history (see Figure 5.6, these instances were defined in Section 5.2).

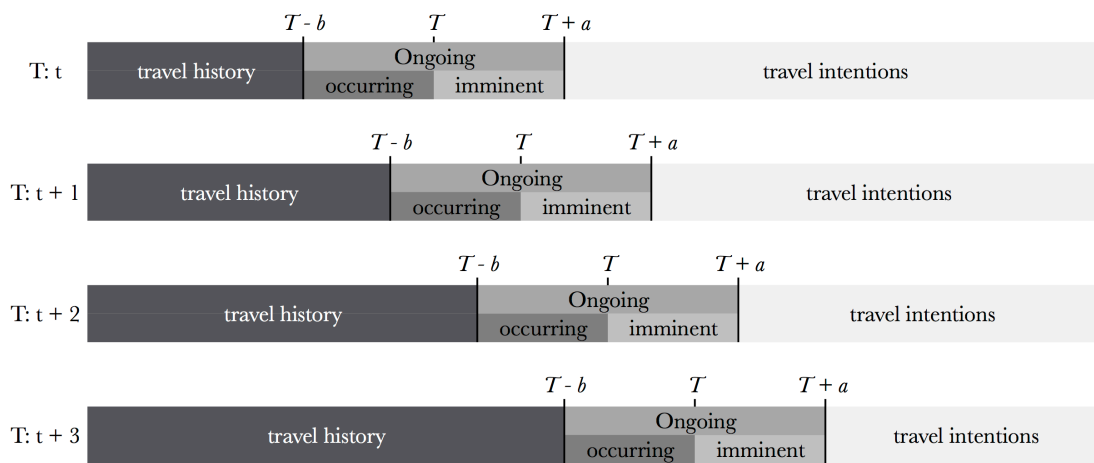


Figure 5.6 Time chart

Sourcing data is the first main activity of the algorithm and is detailed in Figure 5.7. It focuses solely on the extraction of ongoing journeys at the defined time instant T as per Interval (5.2). First, the algorithm goes through the source of journey data to identify ongoing journeys paths. Next it extracts all stages of ongoing journeys. This ensures that any stages that were completed prior to the ongoing journeys interval, but are part of them, are sourced from the travel history to be considered in the creation of TUNs. In public transport, a journey stage has already been said to consist of a portion of a journey occurring in a specific route, direction, and vehicle. In private transport modes such as private cars and cycling, a stage relates to a portion of a journey along a street or road. Hence a journey may comprise one or multiple stages. The extraction of journey stages will later allow the identification of similar portions between journey paths. Lastly for this activity, the algorithm

computes the length of journey paths to avoid duplication of tasks, because those will be necessary in both of the subsequent main activities running in parallel (see Equations (5.4) and (5.5), and Figure 5.5) described in the two next sub-sections.

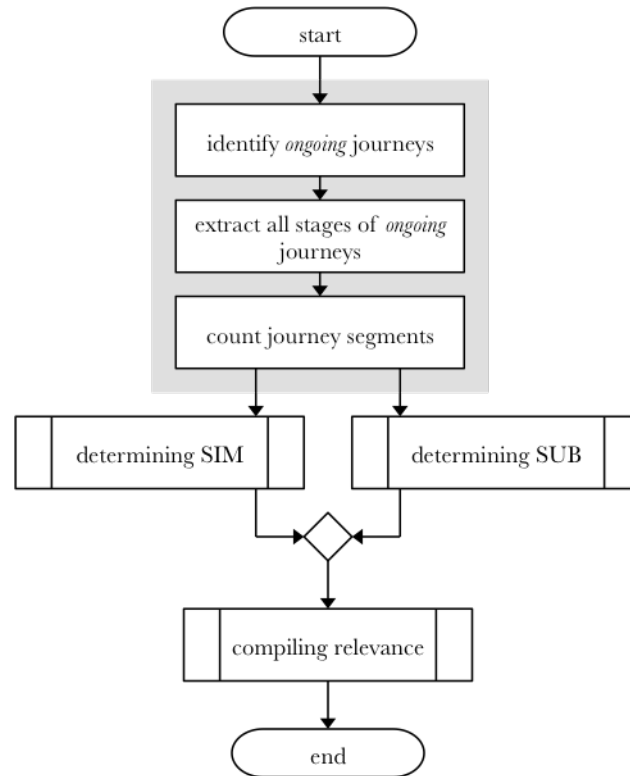


Figure 5.7 Flowchart of activity ‘sourcing data’

Calculating SIM

Calculating SIM is the second main activity of the algorithm, running in parallel with the third, and is detailed in Figure 5.8. The output of this activity is the square SIM matrix \mathcal{S} . This second main activity consists of three nested loops that iterate over the entire set of data sourced in the previous activity, which must be sorted primarily by journey and secondarily by journey stage. Once a journey stage is selected, the algorithm goes through all stages of the remaining ongoing journeys to identify similar stage candidates. It does so by querying first, if the similar stage candidate is on the same route or road and direction of travel, and second, if the path of the similar stage candidate overlaps with it. That second query is desirable to reduce computational effort, because a similar stage candidate may be on the same route or road and direction but in distinct parts of it, hence dissimilar. If the answer to both of these queries is affirmative, the algorithm identifies the beginning and end of the similar portions of the journeys stages under analysis. Then, it calculates the similarity of the stage candidate in relation to the entire journey length using the formula depicted by Equation (5.4). Lastly, the algorithm adds that value to the journey pair SIM, because the journey may consist of various similar stages. Once that calculation is complete, the algorithm iterates over the three loops until all similar stage candidates of all journey stages have been analysed.

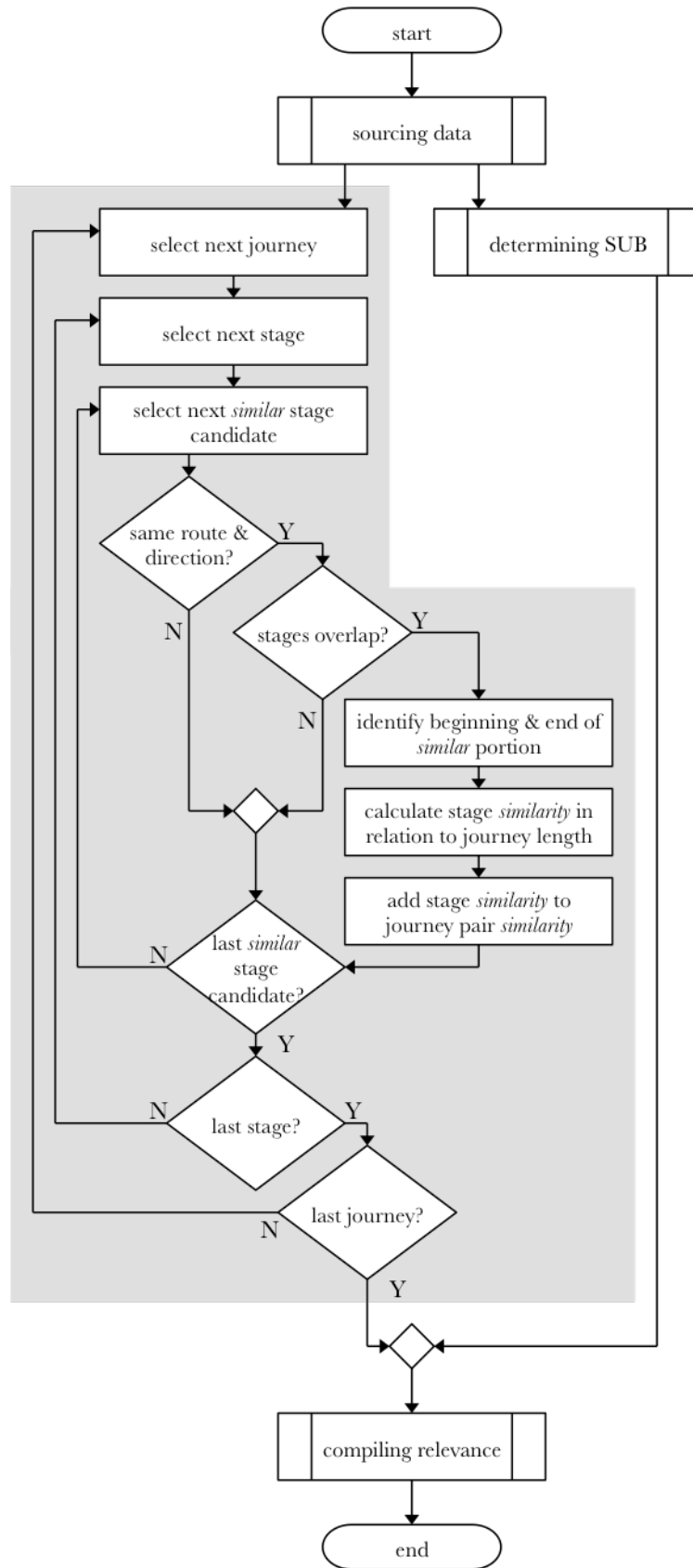


Figure 5.8 Flowchart of activity 'calculating SIM'

Calculating SUB

Calculating SUB is the third main activity of the algorithm, running with parallel with the second, and is detailed in Figure 5.9. The output of this activity is the square SUB matrix \mathbf{B} . This activity differs significantly from its SIM counterpart. Instead of looking for overlapping journey portions, the calculation of SUB is solely concerned with the length between pairs of SPODs. Hence the algorithm has to subdivide further each journey stage into journey segments. Segments are the span between two adjacent nodes that are linked in the passenger transport network grid. In public transport, a segment is a single span between two adjacent stations or stops of a route. In private transport modes such as private cars and cycling, a segment is the single span between two road junctions. This third main activity consists of two nested loops that iterate over all journeys and identify their substitute journey candidates, and two additional nested loops within the former to calculate the SUB of suitable candidates.

Once a journey stage and its suitable substitute journey candidate are selected, the algorithm verifies if they share at least two SPODs. These cannot be in the same route or road and direction of travel; otherwise they are accounted for in the parallel SIM calculation. It is reminded that SPODs may be within a set catchment area called vicinity (see Section 5.2), so the actual points may be slightly apart. If the answer is affirmative, the algorithm proceeds to identify all pairs of SPODs between the journey and suitable substitute journey candidate. Next, two nested loops consider all possible combinations of pairs of SPODs for the beginning and end of the substitute journey portion that may yield the optimal solution, and calculate SUB using the formula depicted by Equation (5.5). This avoids underestimation of SUB, because smaller portions of substitute journey paths with lower penalty factors due to closer proximity between SPODs may yield greater SUB. Some combinations of pairs of SPODs may be excluded to reduce computational effort; Equation (5.5) reveals that it is not worth considering a downstream origin SPOD or an upstream destination SPOD unless the proximity of points within SPODs becomes shorter. This exclusion is not mandatory and therefore not made explicit in the flowchart (Figure 5.9). Once the optimal SUB value is found for the journey and all of its suitable substitute journey candidates, the algorithm moves on to the next journey until all have been analysed.

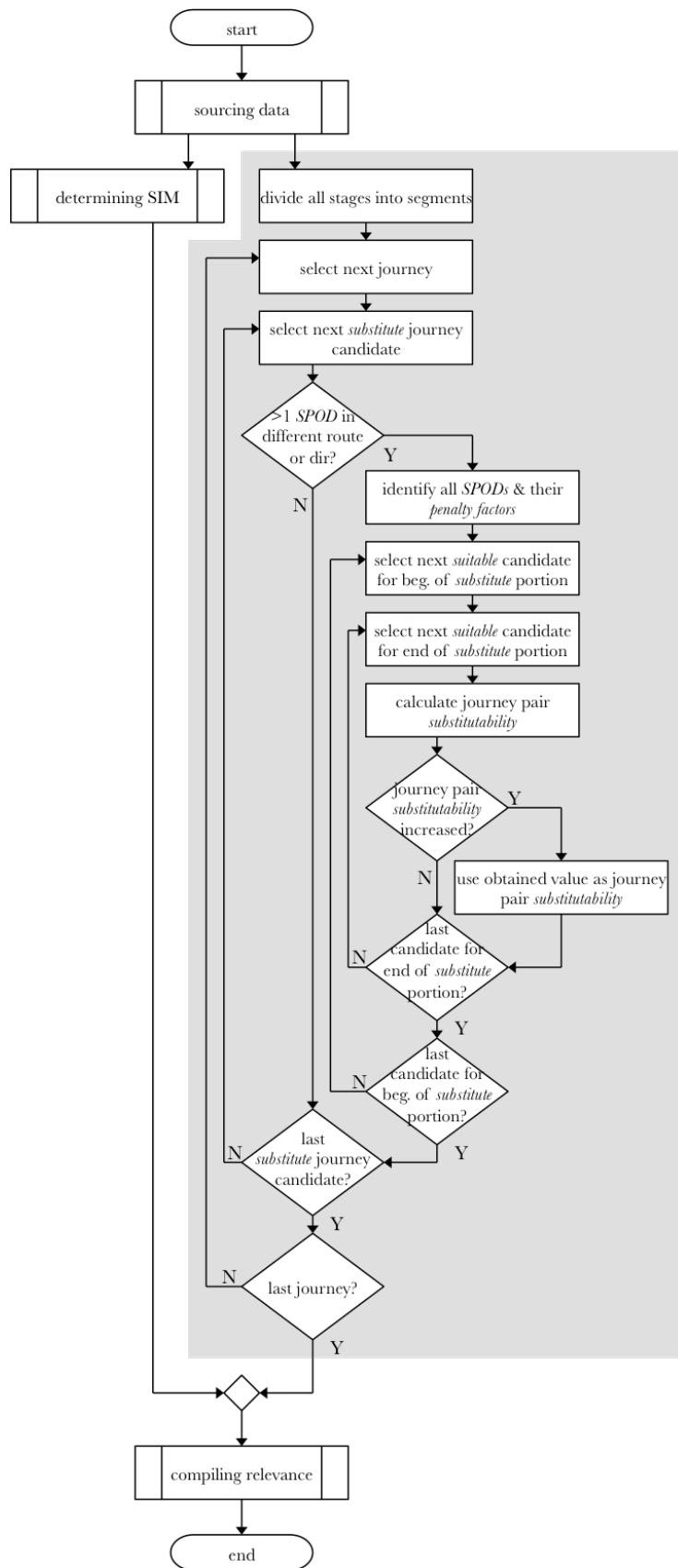


Figure 5.9 Flowchart of activity ‘calculating SUB’

Compiling relevance

Compiling relevance is the fourth main activity of the algorithm, executed once both second and third have been completed. It is detailed in Figure 5.10. The output of this activity is the output of the algorithm, the square relevance matrix R . This fourth main activity consists of a loop that iterates between all journey pairs, and in most cases relates to the straightforward sum of SIM and SUB of the journey pair.

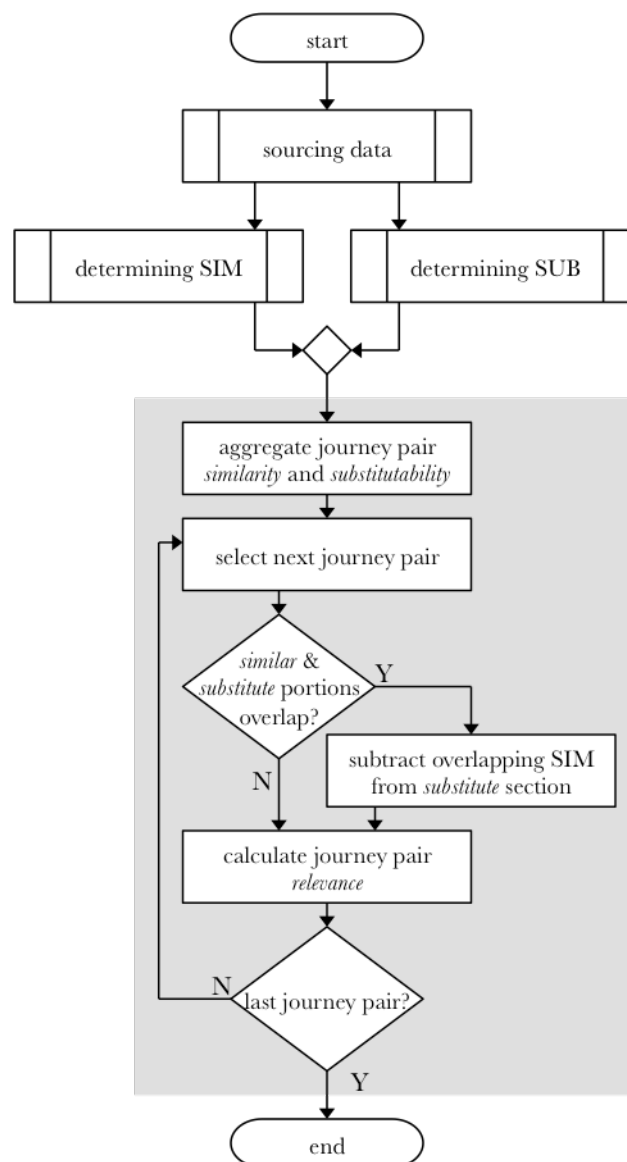


Figure 5.10 Flowchart of activity 'compiling relevance'

There are however some exceptions, relating to journey pairs with overlapping similar portions within a substitute portion as shown in the example of Figure 5.11. In those exceptional cases, the algorithm subtracts the overlapping SIM to avoid accounting twice for it and overestimating relevance for that journey pair. Once all journey pairs have been analysed, the relevance between them is determined, matrix R is complete, and the algorithm terminates.

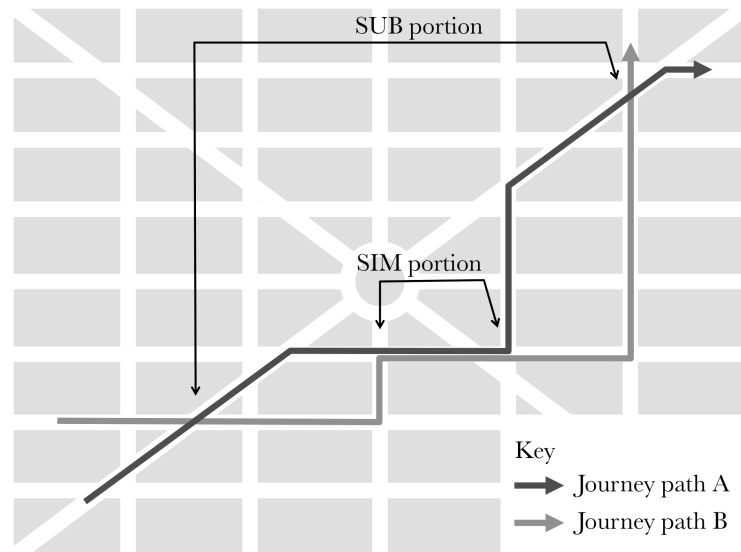


Figure 5.11 Similar portion within substitute portion

5.4. Data subset description

This thesis illustrates the process of creating TUNs and evaluates the potential of the user collaboration model using data from the urban public transport domain. This domain has clear relevance because, typically, many passengers simultaneously populate a public transport system at any moment in time. These passengers have individual journey purposes and plans, making the behavioural dynamics of a public transport system complex. Users often have numerous travel alternatives, ranging from different modes to different routes for the same mode, and even to various scheduled services on the same route. Since passenger may rely on more than one route to reach their destinations, a large number of available combinations for their travel plans exist. Within this complexity exists a rich potential source of collective

intelligence geographically distributed across the system. Harnessing that collective intelligence can potentially generate information of value for travellers and transport providers as discussed in Section 3.2 of this thesis.

The data used in this research study is a subset of the enriched 2013 Andante dataset from STCP obtained from the destination estimation methodology described in Chapter 4. Part of the dataset is used to illustrate the process of creating TUNs and evaluate the potential of the user collaboration model. The TUNs users in this case are passengers who may benefit from collaboratively generated information. A specific date and time were selected from the dataset to simulate what would be the present time instant in a real-time environment. That time was repeated in 10 minutes increments to capture the evolution of TUNs. For this experiment, Wednesday, 2nd October 2013 at 11:00 was selected as the initial time instant, with subsequent increments of 11:10, 11:20, and 11:30. The date represents a typical mid-week working day and a regular school day, with a total number of journey stage transaction records within a 4% difference of the weekday average for the three-month period of O-D data (excluding a one-day STCP staff strike within that period). The time chosen represents the inter-peak period, with the lowest level of journey stage transaction records for that particular day (Figure 5.12). This aims to provide the worst-case scenario in terms of user collaboration potential for the period of interest bounded by the morning and evening peak times.

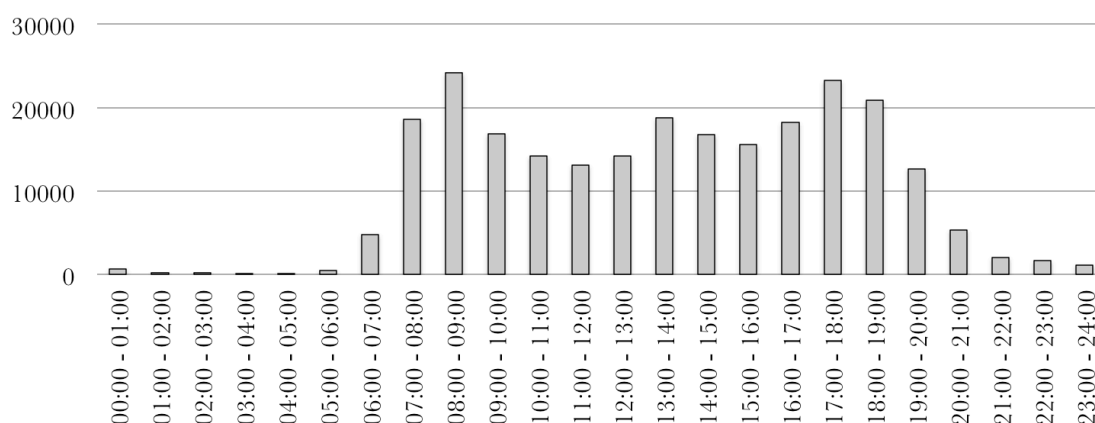


Figure 5.12 Journey stage transaction records throughout the day. Adapted from Nunes, Galvão Dias, et al. (2016).

Ongoing journeys are simulated for each time instant considering all journeys that have at least one stage that started in the previous 30 minutes and are not known to having been completed (the enriched 2013 Andante dataset from STCP contains journeys with estimated destination yet some without arrival time, see Section 4.6). These 30 minutes correspond to constant b in Intervals (5.1) and (5.2). In real-time, this may also require predicting the destination using the travel history of passengers, because ongoing journeys have not been completed. Passengers who are about to start their journeys should also be included, to potentially benefit from collaboratively generated information that allows them to adjust travel plans according to service status. These are imminent journeys, which effectively are intentions soon enough to be considered ongoing journeys, which in real-time must either be captured or predicted from the travel history of passengers. Imminent journeys are simulated as those starting within 15 minutes of each time instant, which correspond to constant a in Intervals (5.2) and (5.3). To summarise, the data selected for the simulation of real-time consists of the enriched 2013 Andante transaction records from STCP that represent ongoing journeys, either occurring or imminently about to start (Figure 5.13).

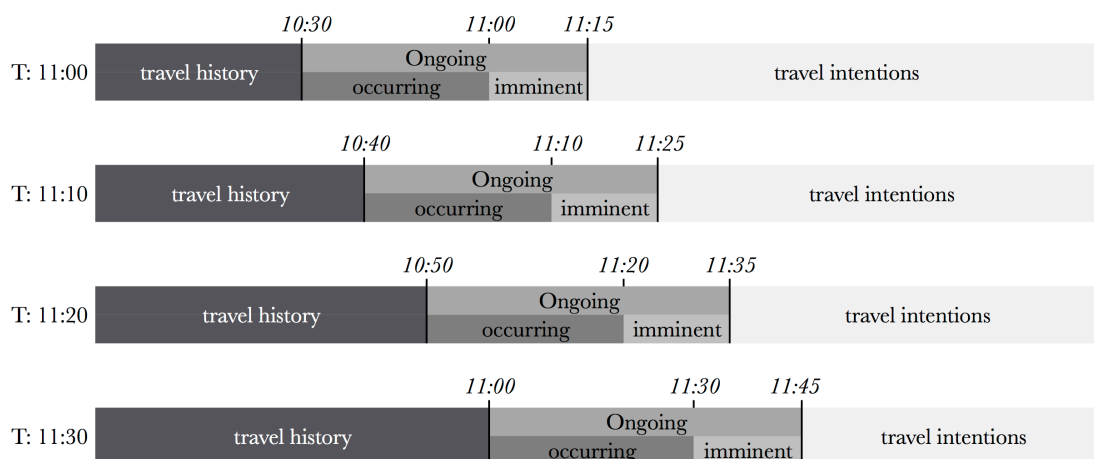


Figure 5.13 Time chart of ongoing journeys

Figure 5.14 represents a snapshot of ongoing journeys in the data subset at 11:00 over a map of the Andante system. The points represent bus stops in their geographic location, sized proportionally to the number of departures and arrivals. The lines represent passenger journeys between the respective origin and destination stop pairs. The thickness of lines depicts the number of passenger journeys between each pair. The Andante zones relevant to the present case study are labelled and shaded. The zone corresponding to the city centre of Porto (C1) concentrates a major share of passenger journeys. Not only does this area have higher activity levels and urban density, it is also where most STCP routes pass by. Table 5.1 provides the key statistics of the data subset at each of the time instants considered.

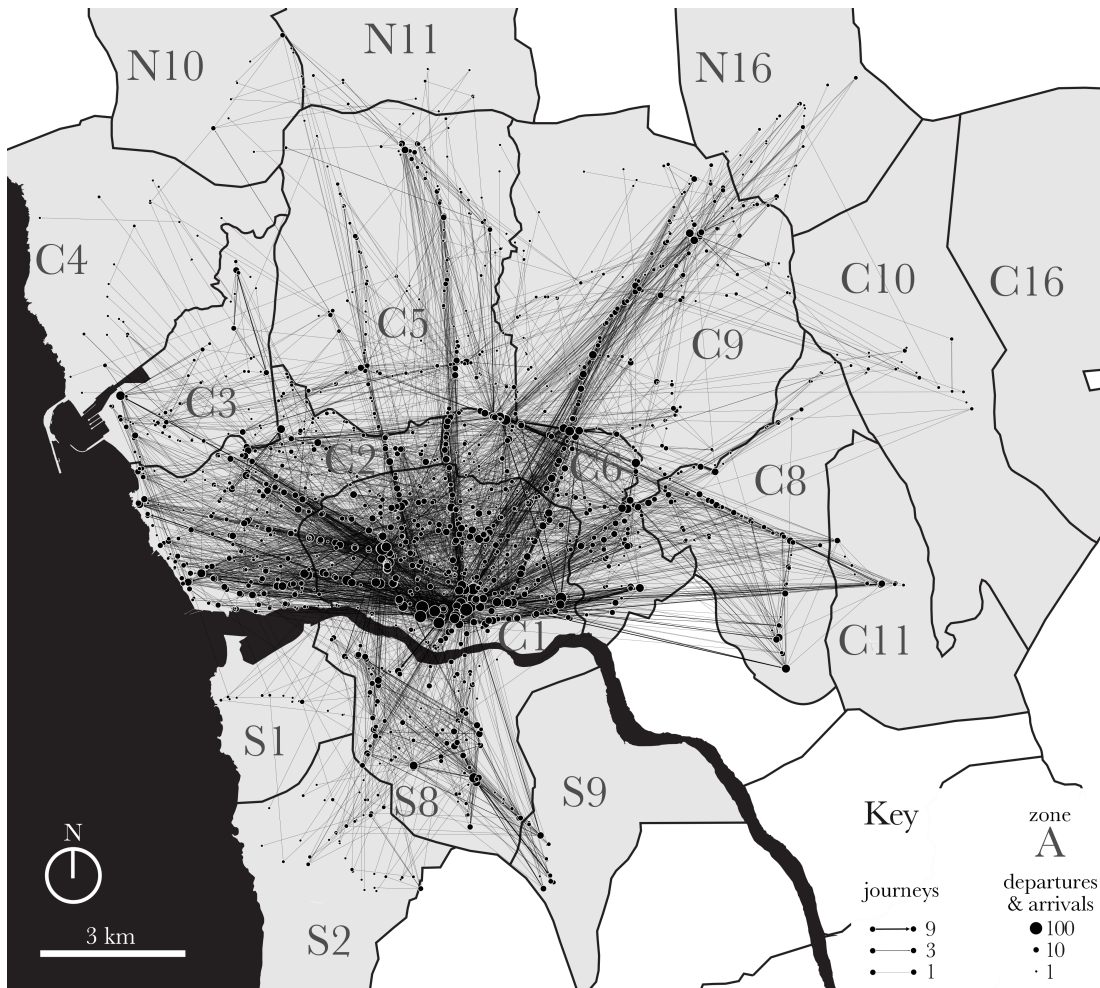


Figure 5.14 Ongoing journeys at 11:00. Adapted from Nunes, Galvão Dias, et al. (2016) with new data.

Table 5.1 Data subset summary

	11:00	11:10	11:20	11:30
Journey stage transaction records	8191	8136	8091	7992
Passengers	5736	5727	5733	5675
Bus routes with transaction records	58	59	59	58
Vehicles with transaction records	283	282	276	273
Stops with arrivals and/or departures	1636	1616	1626	1601

5.5. Implementation

An initial implementation of the TUNs algorithm was developed in SQL. The algorithm has exponential complexity; the relevance of each user is calculated against all other users simultaneously in the system, therefore each user adds one unit to the dimensions of output square matrix \mathbf{R} . The main purpose this initial implementation was to materialise the TUNs concept and evaluate the potential of the user collaboration model in light of random network theory as reviewed in Section 2.3. Yet, the execution time of the algorithm in real-time became a greater concern upon its incorporation into the *Seamless Mobility* project. This project (described in more detail in Section 6.4) adopted the TUNs algorithm to enable collaborative exchanges of information between its users in a public transport context. This requires TUNs to be created and persistently updated in real-time. The algorithm was then upgraded in collaboration with other members of the project team. This upgrade fundamentally consisted of being partly rewritten in Transact-SQL, and incorporating a method for indexing data that reduces the amount of calculations that are repeated to reduce execution time. The following section (Section 5.6) describes and discusses the results obtained from executing the initial implementation with the subset of data described in Section 5.4. Section 5.7 will describe the indexing method for improving efficiency and outline some preliminary conclusions from the upgrade.

5.6. Discussion of results

The initial implementation of the algorithm was executed in a computer with a 1.6 GHz Intel Core i5 processor and 4 GB of RAM. The results were obtained for all passengers in the subset of data described in Section 5.4. TUNs were created for the initial time instant at 11:00. The process was repeated for the subsequent time instants at 11:10, 11:20, and 11:30. The dimension of the resulting square relevance matrices varied slightly, ranging between 5675 and 5736 for the inter-peak time instants considered, as per Table 5.1. This number represents simultaneous passengers in the enriched 2013 Andante dataset from STCP, for whom the destination of journeys was successfully inferred using the methodology described in Chapter 4. Given that not all (86.2%) STCP transaction records in the three-month period considered were Andante, information on the complete set of journeys would expectedly yield relevance matrices of slightly greater dimensions. A multi-modal dataset (e.g. including trains) would likely significantly increase the number of simultaneous passengers.

Figure 5.15 summarises the evaluation of the potential of the user collaboration model based on the initial time instant at 11:00. As shown in Table 5.1, there were 58 bus routes in operation served by a total of 283 vehicles, with 1636 bus stops as the origin or destination of 8191 journey stages of 5736 passengers. The relevance threshold was set at 50%, meaning that the TUN of a passenger includes all peers with relevance of at least 50%. Lengths were expressed in journey segments (the span between two bus stops). The vicinity parameter was set to a Euclidean distance of 80 m, representing an approximate walk catchment area of 1 minute at 4.8 km/h. The vicinity parameter must be set taking into account the density of the passenger transport network and walkability of its surrounding areas. Too low, it may underrepresent the actual extent of a SPOD. Too high, it may overestimate the willingness of passengers to detour on foot, and will increase the number of stations or stops within SPODs adding computational effort to the creation of TUNs.

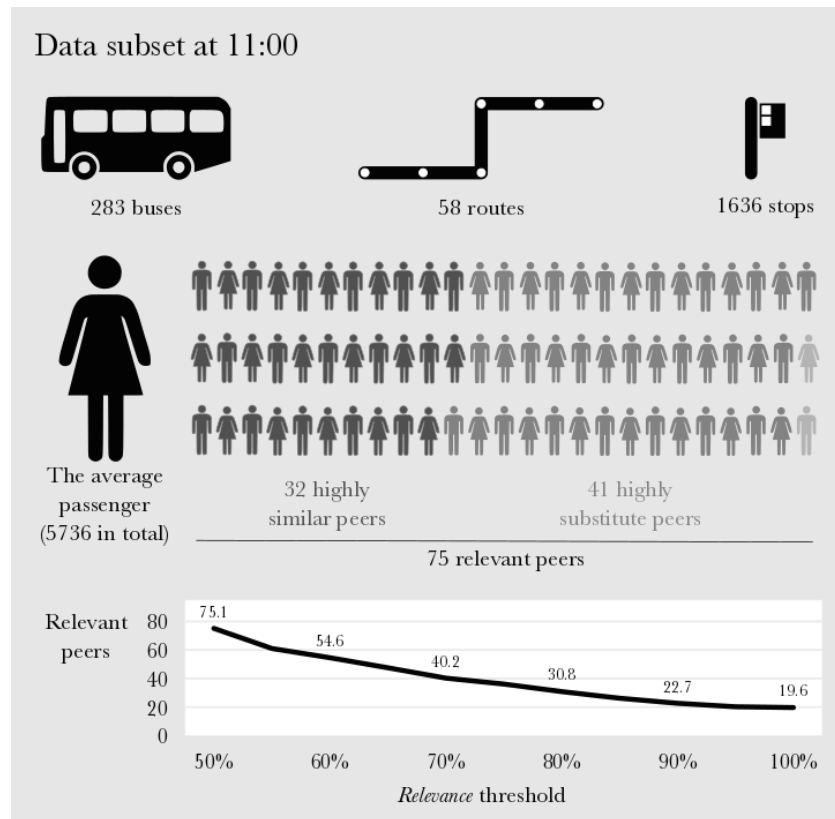


Figure 5.15 Aggregate results at 11:00. Adapted from Nunes, Galvão Dias, et al. (2016) with new data.

At 11:00, each passenger had on average 75 peers in their network. Of these, 32 had similar journey paths and 41 had substitute journey paths, with at least 50% of SIM and SUB respectively. If the entire set of journeys were considered, not just those captured by Andante, the number of routes and vehicles in operation would likely remain unchanged, whereas the number of bus stops at the origin or destination of journeys could be slightly higher. However, the average number of relevant peers in each TUN would certainly increase, given that more users would be candidates for the TUN of each passenger. Regarding the sensitivity of the relevance threshold, if raised to 60% the average number of peers per TUN in the data subset at 11:00 would drop from 75 to 55, to 40 for 70%, to 31 for 80%, to 23 for 90%, or to 20 for 100% (perfectly relevant peers). See bottom of Figure 5.15. Stricter thresholds reduce the magnitude of resulting TUNs.

These results provide insight into the potential of the user collaboration model. But an understanding of the spatiotemporally dynamic nature of TUNs requires analysis in greater depth. To illustrate this, the results from a single typical passenger journey are presented. The selected passenger journey consisted of two stages using two bus routes, and a total of 24 journey segments, 18 in the first stage and 6 in the second stage. Table 5.2 lists the relevant peers included in the TUN of this passenger in each time instant considered, their SIM, SUB, and relevance score. Each peer is identified via a two-letter label for anonymity and ease of reading.

Table 5.2 Relevant peers to selected passenger

Passenger	SIM	SUB	Relevance	10:00	10:10	10:20	10:30
AA	0	0.619	0.619	✓	✗	✗	✗
AB	0	0.667	0.667	✓	✗	✗	✗
AC	0.762	0	0.762	✓	✗	✗	✗
AD	0.571	0.095	0.666	✓	✓	✗	✗
AE	0	0.762	0.762	✓	✓	✗	✗
AF	0	0.619	0.619	✓	✓	✗	✗
AG	0	0.762	0.762	✓	✓	✗	✗
AH	0.714	0	0.714	✓	✓	✗	✗
AI	0	0.762	0.762	✓	✓	✗	✗
AJ	0	0.619	0.619	✓	✓	✓	✗
AK	0	0.762	0.762	✓	✓	✓	✗
AL	0	0.667	0.667	✓	✓	✓	✓
AM	0	0.619	0.619	✓	✓	✓	✓
AN	0	0.857	0.857	✓	✓	✓	✓
AO	0.667	0.048	0.715	✓	✓	✓	✓
AP	0.667	0	0.667	✓	✓	✓	✓
AQ	0.667	0	0.667	✓	✓	✓	✓
BA	0.619	0	0.619	✗	✓	✓	✓
BB	0.619	0	0.619	✗	✓	✓	✓
BC	0.619	0	0.619	✗	✓	✓	✓
DA	0.762	0	0.762	✗	✗	✗	✓
			Total	17	17	11	10

Adapted from Nunes, Galvão Dias, et al. (2016) with new data

The number of relevant peers in this TUN was considerable lower than the average, ranging from 10 to 17 across the time instants considered. Selection of this passenger journey was intended to simplify the graphical and numerical illustration of TUNs with a limited number of peers. Some peers were purely similar, others were purely substitutes, and a few were a combination of both. Figure 5.16 illustrates the evolution of this TUN. Each node is a relevant peer, labelled accordingly, and the edges are the connections in their own TUNs. The nodes are coded in different

shades: white to represent peers that joined the TUN in that time instant, grey for those that remain from a previous time instant, and black for those who left because their journeys were no longer ongoing. The sizes of nodes evolve dynamically with the number of connections to other nodes within the TUN, meaning that the largest nodes have the most connections in common with the selected passenger in a specific time instant. The gradient of edges is based on relevance between peers, with darker meaning higher. This representation highlights the temporal volatility of TUNs.

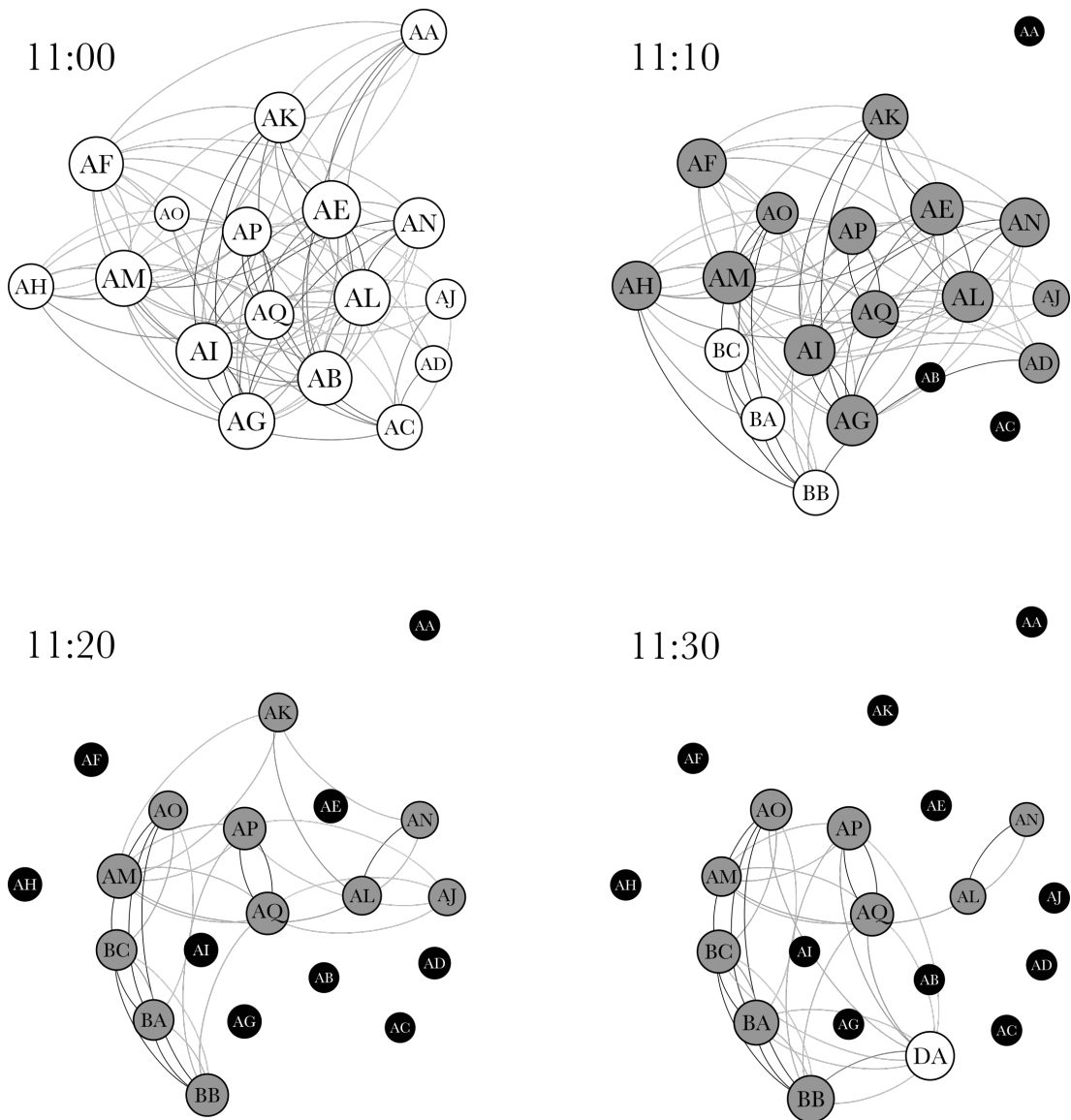


Figure 5.16 TUN evolution for selected passenger. Adapted from Nunes, Galvão Dias, et al. (2016) with new data.

Figure 5.17 provides a geographical representation of journeys of the selected passenger, shown in black, and of travel peers with highest relevance (greater than 0.7) at 11:00. Short dotted grey shaded paths represent predominantly similar peers; long dotted grey shaded paths represent predominantly substitute peers. This representation provides geographic evidence of the validity of the algorithm for creating TUNs. It clearly shows the spatiotemporal equivalence of journey paths from automatically identified peers.

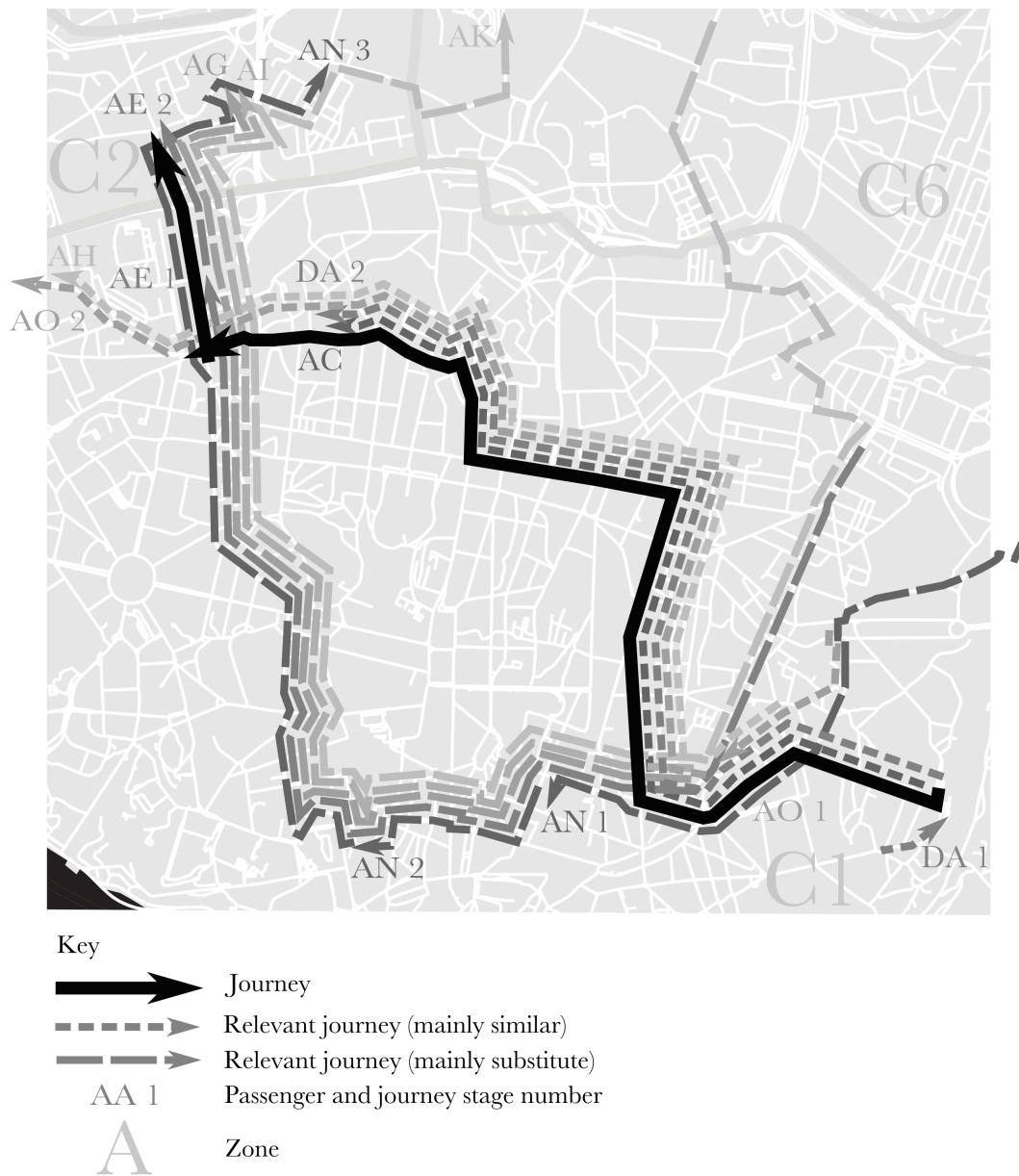


Figure 5.17 Journey paths of selected passengers and highly relevant peers at 11:00. Adapted from Nunes, Galvão Dias, et al. (2016) with new data.

The potential of TUNs and of the user collaboration model is promising. Although the research study only used the Andante data subset and focused on a period with a relatively small number of journey transaction records, the average number of passengers in each TUN is considerable. Even assuming that only 10% of the average of 75 peers per TUN would engage in some form of collaboration using their mobile devices, the property found by Erdős and Rényi (1960) that was reviewed in Section 2.3 would still apply: the average degree $\langle k \rangle$ would be greater than 1, meaning that large clusters should emerge. Figure 5.16 exhibited this clustering potential with the density of connections between peers of the selected passenger. Hence, information may travel quickly across TUNs. Furthermore, the results indicate a good balance between predominantly similar and substitute peers, meaning that passengers may effectively gain access to information about their selected journey paths inasmuch as feasible alternatives.

Deeper analysis of the structure of TUNs obtained, based on various levels of the cut-off *relevance* threshold throughout the time instants considered, demonstrates that their structure is resilient to node removal. Table 5.3 shows that the average degree $\langle k \rangle$ is consistently high, even at high levels of the cut-off threshold. The percolation threshold obtained is relatively low and the expected minimum number of nodes to keep connectivity is significantly lower than the number of nodes \mathcal{N} found across TUNs. These empirical results show that the TUNs obtained are robust and strongly connected. It is observed that, at a specific time instant, raising the cut-off *relevance* threshold causes the number of edges (E) to drop much faster than the number of nodes (\mathcal{N}). As Albert and Barabási (2002) point out, the topology of the network can determine the robustness of the network. Since the removal of a node implies removing all of its edges, node removal inflicts more damage than edge removal.

Table 5.3 Analysis of TUNs across time intervals at various relevance threshold levels

		Relevance threshold	0%	50%	60%	70%	80%	90%	100%
11:00	Number of nodes	\mathcal{N}	5736	5728	5686	5591	5397	5059	4688
	Number of edges	E	1801446	430606	313036	230848	176892	129941	112704
	Average degree	$\langle k \rangle$	314.060	75.176	55.054	41.289	32.776	25.685	24.041
	Maximum degree	$\max k$	1497	432	352	350	348	348	348
	Percolation threshold	$p_c = 1/\langle k \rangle$	0.003	0.013	0.018	0.024	0.031	0.039	0.042
	Min. nodes to keep G	$\mathcal{N} * p_c + 1$	19	77	104	136	166	198	196
11:10	Number of nodes	\mathcal{N}	5727	5712	5678	5577	5391	5070	4686
	Number of edges	E	1760825	428134	310900	228569	175185	128618	112561
	Average degree	$\langle k \rangle$	307.460	74.953	54.755	40.984	32.496	25.368	24.021
	Maximum degree	$\max k$	1468	433	365	365	329	329	329
	Percolation threshold	$p_c = 1/\langle k \rangle$	0.003	0.013	0.018	0.024	0.031	0.039	0.042
	Min. nodes to keep G	$\mathcal{N} * p_c + 1$	20	77	105	137	167	201	196
11:20	Number of nodes	\mathcal{N}	5733	5719	5691	5584	5393	5067	4698
	Number of edges	E	1699944	421192	309391	229027	175730	130142	114487
	Average degree	$\langle k \rangle$	296.519	73.648	54.365	41.015	32.585	25.684	24.369
	Maximum degree	$\max k$	1420	430	356	354	289	282	282
	Percolation threshold	$p_c = 1/\langle k \rangle$	0.003	0.014	0.018	0.024	0.031	0.039	0.041
	Min. nodes to keep G	$\mathcal{N} * p_c + 1$	20	79	106	137	167	198	194
11:30	Number of nodes	\mathcal{N}	5675	5661	5617	5507	5334	4994	4640
	Number of edges	E	1633520	406013	296687	219374	168270	123765	109724
	Average degree	$\langle k \rangle$	287.845	71.721	52.819	39.835	31.547	24.783	23.647
	Maximum degree	$\max k$	1375	420	364	328	328	328	295
	Percolation threshold	$p_c = 1/\langle k \rangle$	0.003	0.014	0.019	0.025	0.032	0.040	0.042
	Min. nodes to keep G	$\mathcal{N} * p_c + 1$	21	80	107	139	170	203	197

Adapted from Nunes, Galvão Dias, et al. (2016) with new data

Generalising these results across passenger transport domains, TUNs may effectively uncover an invisible layer of social engagement, which facilitates collaboration-based value co-creation. TUNs offer opportunities for circumstantial user collaboration, overcoming limitations of existing ICT-based social networking services to deal with spatiotemporally dynamic information. The sensitivity analysis of the relevance threshold provides evidence that those opportunities are greater with larger numbers of simultaneous users in a passenger transport system. This is due not only to the breadth of information distributed, but also to the chance of applying stricter relevance thresholds in larger systems and still obtaining adequate numbers of connections per TUN.

While the concept of TUNs is applicable across passenger transport domains, this research has focussed on the development of an ICT-based application for mobile devices called *Journata* that enables collaborative exchanges of information among passengers, which is specific to the urban passenger transport domain. It resembles

an ICT-based social networking service whereby users interact with peers in their TUNs by rating aspects relating to their journeys. *Journata* is the focus of Chapter 6.

5.7. Indexing method

Given the exponential complexity of the TUNs algorithm, the calculation of the square relevance matrix \mathbf{R} for high numbers of users and journey stages is a computationally demanding task. This prompted an attempt to minimise the algorithm execution time for a real-time implementation in the context of the *Seamless Mobility* project, carried out in collaboration with other members of the project team. Drawing from the notion of indexes in computer science, the Journey Pair Relevance Index (JPRI) was created to eliminate redundant calculations. In conceptual terms, those indexes are data structures that organise information to facilitate retrieval. Web search engines such as Google partly rely on text-based indexes to generate results in a fraction of a second (Brin and Page 1998). The JPRI is instead a journey-based index containing pre-processed relevance calculation results.

Given the complexity of urban passenger transport systems, users have a very large set of alternatives (with varying degrees of efficiency) to get from an origin to a destination (see also Section 5.4). However, only a fraction of observed journeys are distinct. A distinct journey in this context is characterised by the origin and destination of all of its constituent stages. For example, users often repeat their daily commute patterns, and many different users follow the same route indicated by a journey planner. Analysis carried out over a 25-day period of an enriched Andante dataset (1 to 25 April 2010, STCP only), obtained from the destination estimation methodology described in Chapter 4, has demonstrated that the number of distinct new journeys steadily decreases over time. Compared against the total of hourly journeys (Figure 5.18), towards the end of the period only 10% approximately are distinct (Figure 5.19). This provides ample scope for the effectiveness of an indexing method.

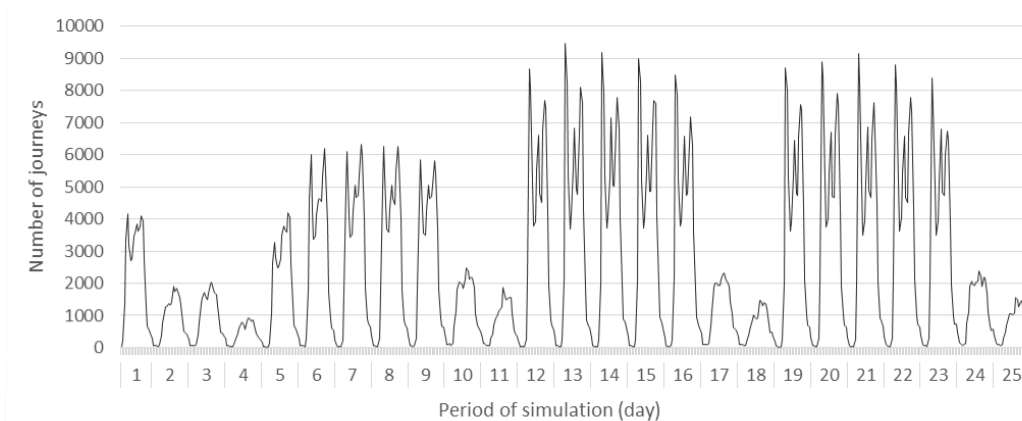


Figure 5.18 Journeys hourly across simulation period (chart courtesy of Tânia Fontes)

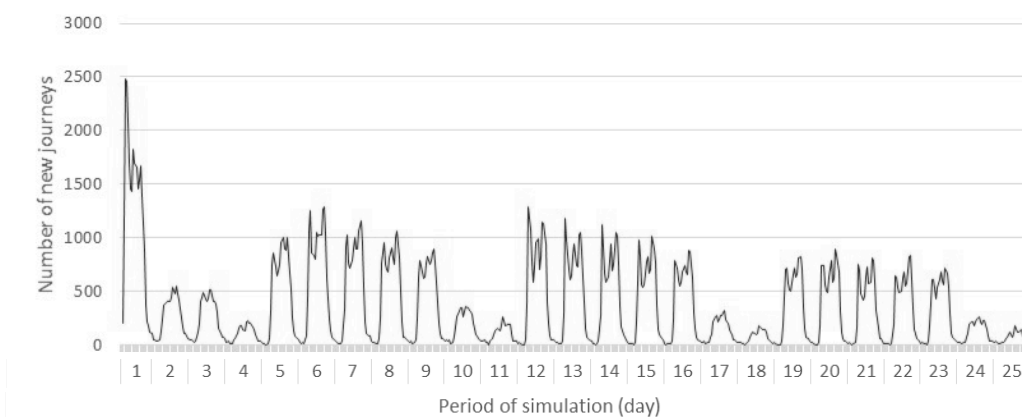


Figure 5.19 New distinct journeys hourly across simulation period (chart courtesy of Tânia Fontes)

The JPRI is incrementally updated with additional information every time a new distinct journey occurs in the passenger transport system. That additional information consists of the relevance between that journey and all other journeys occurring simultaneously that already part of the index. With the JPRI, the TUNs algorithm only needs to calculate the relevance between pairs of journeys that are not in the index yet. That is only done once for each pair of journeys because the results get subsequently stored in the JPRI. The upgrade of the TUNs algorithm in Transact-SQL carried out in the context of *Seamless Mobility* had not been finalised upon the time of writing. Yet, some provisional results obtained were indicative that the creation of TUNs in real-time is feasible, owing to the combination of a more efficient algorithm with the use of the JPRI.

5.8. Chapter conclusions

This chapter described the concept of TUNs as the second tool for implementing the user collaboration model. This included the mathematical formulation, the structure of an algorithm to perform the calculations, an implementation of the algorithm, and the results from applying it to the subset of 2013 Andante data. TUNs aim to overcome limitations of existing ICT-based services that have inhibited user collaboration-based value co-creation to occur more widely in urban passenger transport domains. These networks introduce the ability to identify circumstantial user-based collaboration opportunities to facilitate the diffusion of information, dynamically and in real-time.

The creation of TUNs for the subset of 2013 Andante data demonstrated the potential of the user collaboration model. Even with a subset of passenger journeys and under conservative assumptions, large TUNs of highly relevant peers were created for those passengers. Therefore, the potential exists for information to travel quickly across networks. The results additionally revealed a good balance between predominantly similar and substitute peers, although this is highly dependent on the density of the passenger transport system. Denser systems are likely to increase the portion of substitute peers. Sensitivity analysis of the relevance threshold indicated that circumstantial user collaboration opportunities across transport domains will grow with the number of simultaneous users in a system. Preliminary results obtained from the upgraded algorithm incorporating the JPRI are indicative that it can feasibly be executed in a real-time environment. Along with the conceptual foundations of the user collaboration model and first tool described in the two previous chapters, TUNs provided the last element for fulfilling the second objective of this thesis.

6. Case study: *Journata*

This chapter describes the evolution of the case study *Journata* from its inception as a non-functional prototype, through to its implementation as a commercial prototype service (Figure 6.1). *Journata* is an instance of the user collaboration model, enabling collaborative exchanges of information in real-time among public transport users and providers. The service combines three main components: a mobile application, a web service, and a database. *Journata* was named after the interplay of three key words that are intimately connected with its essence: journey, journal, and data. The way it is pronounced is intended to resemble translations of the word journey into romance languages. The *Journata* logo combines the letter J with a minimalistic representation of a public transport vehicle (Figure 6.2).

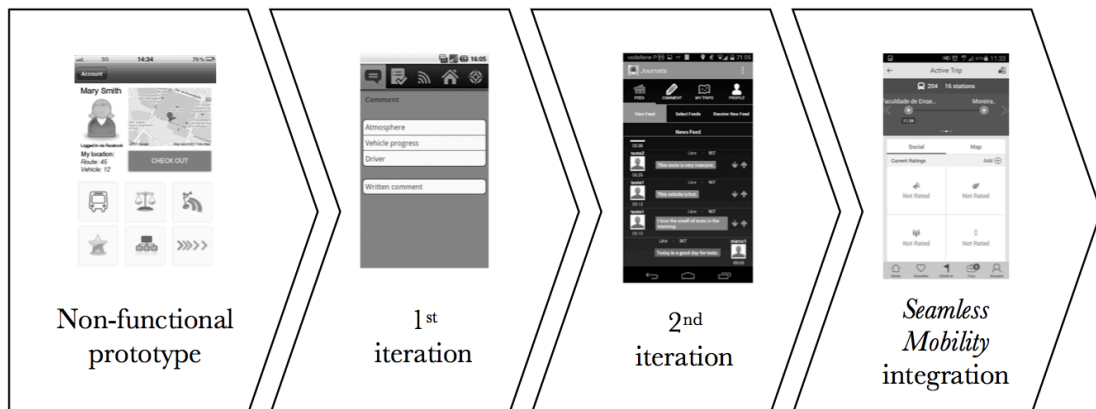


Figure 6.1 *Journata* prototype iteration evolution



Figure 6.2 *Journata* logo in alternative shade combinations (originally in grey tones with a single accent colour, orange or blue)

Section 6.1 describes the concept that originated the development of *Journata*, and a preliminary non-functional prototype of its mobile application interface. Section 6.2 describes the first iteration of a functional prototype of *Journata*, developed in collaboration with a Masters student in Informatics and Computing Engineering at the Faculty of Engineering, University of Porto (FEUP), and a usability test of the mobile application interface carried out in a realistic travel environment. Section 6.3 describes the second iteration of a functional prototype of *Journata* developed in collaboration with another Masters student in Informatics and Computing Engineering at FEUP, and the results of usability evaluation of the mobile application interface. Section 6.4 details the integration of *Journata* into the *Seamless Mobility* project, and presents considerations from this integration. Lastly, Section 6.5 summarises the conclusions drawn.

6.1. Non-functional prototype: the concept

The initial concept for a mobile-based service was grounded on a nested mobile application within an existing social networking application such as Twitter or Facebook (Nunes et al. 2011). It aimed to let public transport users and operators exchange information of interest to each other in a structured way and in real-time, opening up a new communication channel between them. The nested mobile application would leverage existing user accounts to encourage public transport users to try it without downloading a separate application, and while using that main social networking application. A database would store information either provided manually by users or automatically collected from sensors in their mobile devices. Then, a web service would use that information to create personalised travel profiles, and generate information of interest for passengers and transport providers. The initial concept put forward the principles of a social networking model that would rely on the location, travel patterns, and travel intentions of its users, which eventually evolved into TUNs (Chapter 5). Furthermore, it identified the need for validation and incentive mechanisms to safeguard the reliability and availability of information, which became elements of the user collaboration model (Section 3.5).

The non-functional prototype of the mobile application interface was designed for illustrative purposes. It was based on prototyping techniques, initially hand-drawn sketches, and subsequently using Balsamiq Mockups (balsamiq.com) for creating wireframe mock-ups, and MS PowerPoint for animating the interaction. Figure 6.3 shows some interface screen mock-ups from the non-functional prototype of *Journata*. Around the same time, a selection of interface screen mock-ups were enhanced to a visually more elaborate standard edited in high-resolution with Adobe Photoshop (Figure 6.4) to be featured in a scientific publication that described the concept (Nunes et al. 2011).



Figure 6.3 Non-functional prototype interface screen mock-ups



Figure 6.4 Interface screen mock-ups created in Photoshop (source: Nunes et al. 2011)

The prototype mobile application interface home screen (top left, Figure 6.3; left Figure 6.4) displayed the user name, current location, and, whenever applicable, route and vehicle. The identification of route and vehicle was envisaged to either be based on pairing GPS locations of users and vehicles or on a manual check-in process. The various buttons would lead to other screens where the user could comment on aspects of a journey (top centre, Figure 6.3), rate information provided by others as a form of validation (top right, Figure 6.3; centre Figure 6.4), read information from others in spatiotemporally equivalent journeys (bottom left, Figure 6.3), check their usage score and available rewards as participation incentives (bottom centre, Figure 6.3; right Figure 6.4), observe the location of other users (anonymised) in the same user network (bottom right, Figure 6.3), and plan journeys. The types of information to be exchanged would include aspects relating to the travel environment such as temperature, noise, available seating, and crowding; to the travel time, such as punctuality and incidents; and to the skilfulness and courtesy of public transport drivers. Additional details on the complete set of interactions in this non-functional prototype of *Journata* can be found in Nunes et al. (2011). Those adopted into the functional versions are most relevant to this thesis and are described in greater detail in the following Section 6.2.

6.2. Functional prototype, first iteration

This section is partially based on Nunes, Gonçalves, and Galvão (2013), “A Prototype for Public Transport Service Co-creation Using Social Media: Results from Usability Testing”.

The first iteration of *Journata* as a functional prototype was developed and evaluated in collaboration with a Masters student in Informatics and Computing Engineering, who carried out the implementation for his own thesis with the author of this thesis as co-advisor (Gonçalves 2012). It materialised the majority of aspects previously outlined as a concept (see Section 6.1), with two main exceptions. First, the mobile application was implemented in a stand-alone format. Second, only a simplified version of the social networking model was implemented, linking all users travelling simultaneously in the same public transport route.

The development of the first iteration of *Journata* followed the four basic activities of a user-centred design approach proposed by Preece, Rogers, and Sharp (2002, 168): “identifying needs and establishing requirements, developing alternative design that meet those requirements, building interactive versions so that they can be communicated and assessed, and evaluating them”. These activities are meant to be carried out iteratively as a cycle. Regarding the first basic activity in the context of *Journata*, potential final users were involved early in the design process at a brainstorming session that was held for eliciting requirements. The session lasted approximately two hours, and brought together a group of six public transport users, and a representative of a public transport operator, of varied age groups and backgrounds. The heterogeneity of the brainstorming group yielded a rich outcome in terms of requirements. The outcomes of the session were the refinement of the types of information that could usefully be exchanged among passengers, and a formal requirement specification consisting of twenty-three use cases, detailed in Gonçalves (2012). These were prioritised based on their importance for evaluating the concept, with the highest priority ones being chosen for implementation in the first iteration of *Journata*. The next two paragraphs provide a summary of key functionalities in the formal set of use cases.

After logging into their accounts, users check-in when they start a journey, either manually or automatically. The latter scenario can be based on integration with a mobile-based public transport ticketing system. The check-in triggers the creation of a network of users with spatiotemporally equivalent journeys. At the end of a journey, check-out is also manual or automatic. In the latter scenario, a comparison between the GPS coordinates of the vehicle and user can reveal when the check-out should be triggered. While checked-in, users may share information with their peers (other users) in their network with the mobile application. That information may be shared either in a categorised or in a written format. The first type relates to information that fits within a predefined set of categories and can be scored subjectively in a rating scale, such as the crowding of a vehicle or the skilfulness of a driver. The second type relates to information that is not included in a category or cannot be scored in a rating scale, such as the report of a traffic incident. Besides supplying information, users are randomly prompted to rate information provided by peers in their network, or in their network and vehicle for some categories, in a scale of perceived correctness, to safeguard that it is valid and worthy of being distributed mode widely.

Users accumulate points for providing information that is rated as correct, which may be used towards claiming rewards such travel card discounts, as defined by the transport provider. Users can monitor their score and available rewards in the mobile application. Users can read a news feed consisting of information provided and validated by others in their network. Users can plan journeys and decide to be checked-in automatically when that journey is about to start. Furthermore users can configure their profile and privacy settings, and observe other users and their location overlaid on a map (the level of detail subject to the privacy setting of each user). Lastly, users receive notifications with relevant new information while the mobile application is running in the background.

The abovementioned user score was aimed at turning the *Journata* into a serious game with real rewards, remunerating passengers actively engaged in the co-creation of value in the public transport service. Yet, as long as the reward system parameters

are well defined, potential benefits for the transport provider in terms of increased patronage may offset expenses related to rewarding passengers.



Figure 6.5 First iteration functional prototype interface screens (Nunes, Gonçalves, and Galvão 2013)

Going back to the basic activities of interaction design, regarding the second (developing alternative design) and third (building interactive versions) the process had already been initiated with various iterations of the non-functional prototype of *Journata* (see Section 6.1). Those were revisited and evolved into the first version of the interactive functional prototype of the mobile application. Figure 6.5 shows some of the main mobile application screens. The screen on the top left allows users to

provide categorised comments. The screen on the top centre allows users to rate information provided by others for validation before being distributed more widely. The screen on the top right presents the user network news feed. The screen on the bottom left allows users to monitor their score, claim rewards, and edit their personal and privacy settings. The screen on the bottom centre is the journey planner. The screen on the bottom right is the notification centre that alerts users of new information, of requests for rating information provided by others, and reminds them of scheduled journeys.

Finally, in terms of the fourth basic activity of interaction design (evaluating the interactive versions) regular reviews were held with the thesis advisors, and the prototype was eventually tested with potential users in a realistic setting, as described later in this section.

Implementation

The service consisted of three main components that were implemented using the following technologies respectively: Android for the mobile application, REST for the web service, and MySQL for the database. A secondary component consisted of a MS Windows application for querying the database using the Android Cloud to Device Messaging Framework (C2DM). Figure 6.6 illustrates the architecture of the prototype service. It has been mentioned that use cases were prioritised, some of which were left out of this first iteration of *Journata*. Those were felt not to compromise the purposes of validating the underlying concept, and testing the usability of the prototype mobile application interface. The first, already mentioned, was the creation of user networks centred on routes instead of users based on the spatiotemporal equivalence of their journeys (i.e. TUNs were not implemented in the first iteration). The second was the mapping of users. Besides the use case specification, in this first iteration journey planning was restricted to journeys with a maximum of two stages, the mobile application did not capture information from device sensors (e.g. GPS tracking to trigger automatic check-out), and the database was only fed with bus service data so the prototype service was mono-modal.

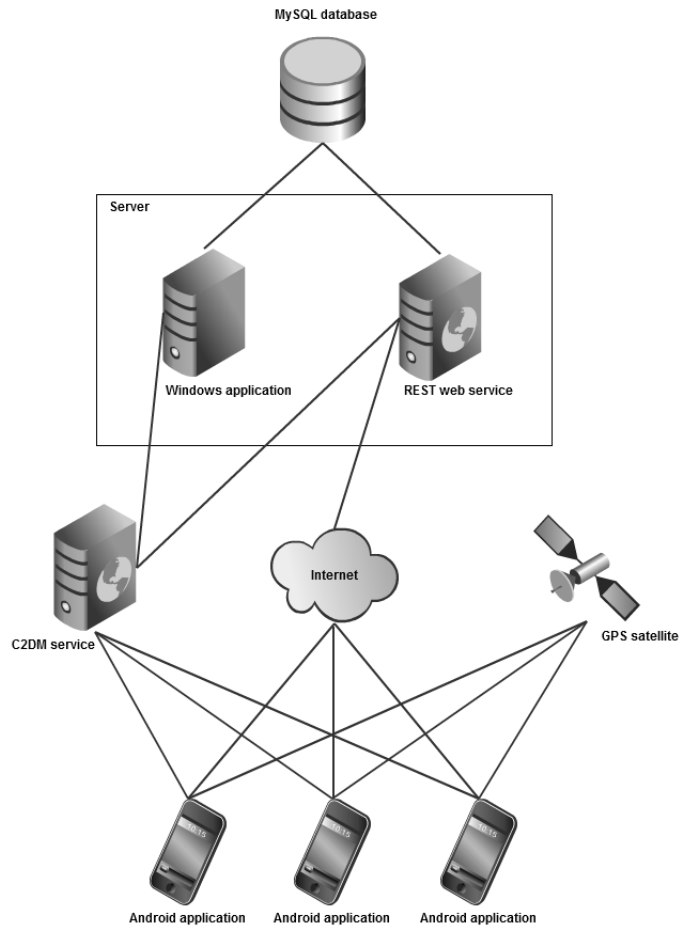


Figure 6.6 Architecture of the *Journata* prototype (source: Gonçalves 2012)

Evaluation

For the purposes of evaluating the prototype service *Journata* with a usability test, the database was fed with public transport data from TfL. London was chosen for various reasons, namely that rich sets of TfL data are publicly available for developers, that TfL has one of the most comprehensive urban public transport systems in the world, and that it consequently offered potentially greater exposure to *Journata* in the future. The first iteration of the prototype service was tested with potential users in a realistic context riding TfL buses.

The main objective of the experiment was to evaluate the usability of the prototype mobile application interface. A secondary objective has been to validate the underlying concept of *Journata*, particularly if the notion of dynamic networks fits in

with the existing mind-set of social networking service users. Participants for the usability test were recruited in two steps, firstly among acquaintances of the author of this thesis living in London, and secondly via a non-probabilistic technique called snowball sampling. The latter consisted of the first recruited participants recruiting others among their own acquaintances. The goal has been to recruit more than five participants to exceed recommendations from Jakob Nielsen (1989) and Robert A. Virzi (1992). According to these authors, five users are generally enough to identify 80% of usability problems, but the broad profile of public transport users may justify the adoption of a more conservative approach. Ten participants were eventually recruited, but only nine were able to attend the usability test. Whilst this number of participants may be adequate for the usability test, for the secondary objective of validating the underlying concept the results obtained must be read with caution. Table 6.1 shows the characteristics of the participant sample.

Table 6.1 Participant characteristics

Variable	Number of participants
Gender	
Female	3
Male	6
Age	
$20 \leq \text{age} < 30$	2
$30 \leq \text{age} < 40$	7
Weekly average public transport journeys	
$0 \leq \text{journeys} < 10$	5
$10 \leq \text{journeys} < 20$	1
$20 \leq \text{journeys} < 30$	3
Smartphone (*) user	
Yes	8
No	1
Smartphone type (**)	
Android	4
iOS (Apple)	5
N/A	1
Notes:	
(*) Mobile phone with mobile operating system; and	
(**) Total exceeds number of participants since one had two types of smartphones.	

Adapted from Nunes, Gonçalves, and Galvão (2013)

An effort was made to simulate as closely as possible the foreseen context of use of the prototype service. The first part of the experiment was conducted while waiting for a bus to arrive, at a London bus stop, and a second part was conducted whilst riding the bus (Figure 6.7). Each participant was given an Android smartphone with the *Journata* mobile application prototype, whereas the facilitator held a similar Android smartphone for exchanging real-time data needed for the usability test. The experiment was supported by pre-test and post-task questionnaire-based surveys, and a post-test questionnaire-based interview. The pre-test questionnaire simply captured the participant characteristics summarised in Table 6.1. The post-task questionnaire was meant to capture the perceived ease-of-use associated with a set of tasks scripted for the usability test, in a 5-point Likert scale ranging from ‘difficult’ to ‘easy’ (Table 6.2). The facilitator also captured the time it took each user to complete each task. The post-test questionnaire-based interview was geared towards the secondary objective of validating the concept (Table 6.3). Participants were prompted to answer each question in a 5-point Likert scale and then verbally expand on their responses. It tried to capture the perceptions of usefulness, ease of use, ease of comprehension, and safety, in the own words of participants.

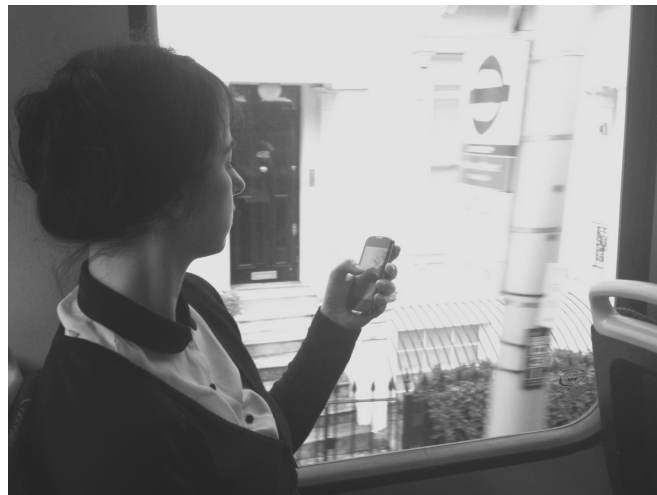


Figure 6.7 Experiment environment (Nunes, Gonçalves, and Galvão 2013)

Table 6.2 Post-task questionnaire

Task	Task difficulty (in a scale of 1: difficult to 5: easy)
1	Login with the following credentials [user is given unique credentials]
2	Through the Android Context Menu, check into the current route and direction using the GPS check-in functionality
3	Share a written comment about this trip
4	Share a driver-related categorised comment
5	Check the news feed
6	Rate a categorised comment
7	Check the current number of points in your account
8	Check the current route of travel
9	Change your nickname
10	Change your profile visibility to public
11	Checkout through the Android Context Menu
12	Plan a journey between Oxford Circus Station and Victoria Bus Station to start in 2 minutes. Select the first option displayed.
13	Accept being added to the journey you have just planned
14	Through the Android Context Menu, manually check-in at Oxford Circus Station. Select the first option displayed.
15	Logout through the Android Context Menu

Adapted from Nunes, Gonçalves, and Galvão (2013)

Table 6.3 Post-test questionnaire-based interview

Question	Description Evaluation
1	What is your overall opinion on the usefulness of such an application? Useless 1 2 3 4 5 Useful
2	Was it easy or difficult to interact with? Difficult 1 2 3 4 5 Easy
3	As a concept, how easy or difficult was it to grasp? Difficult 1 2 3 4 5 Easy
4	How safe would you feel using such an application? Unsafe 1 2 3 4 5 Safe
5	How likely would you be using such an application? Unlikely 1 2 3 4 5 Likely

Adapted from Nunes, Gonçalves, and Galvão (2013)

Results

Figure 6.8 summarises the post-task questionnaire results. It shows, for each, task, how many participants found it to be ‘easy’, ‘relatively easy’, ‘medium’, ‘mildly difficult’, or ‘difficult’ to complete, and the average time of completion. All participants found all of the tasks to be between ‘medium’ and ‘easy’ to complete. All participants found tasks numbered 2 and 15 relating to the check-in and check-out to be ‘easy’. At the opposite end, one or two participants found tasks numbered 3, 5, and 6 relating to sharing, reading, and rating comments, and tasks numbered 8 and 13 relating to checking the current route and to joining a user network based on their journey, to be of ‘medium’ difficulty. All remaining tasks were perceived by all participants to be ‘easy’ or ‘relatively easy’ to complete. Two participants were not able to complete task 6 due to an intermittent fault with the prototype mobile application.

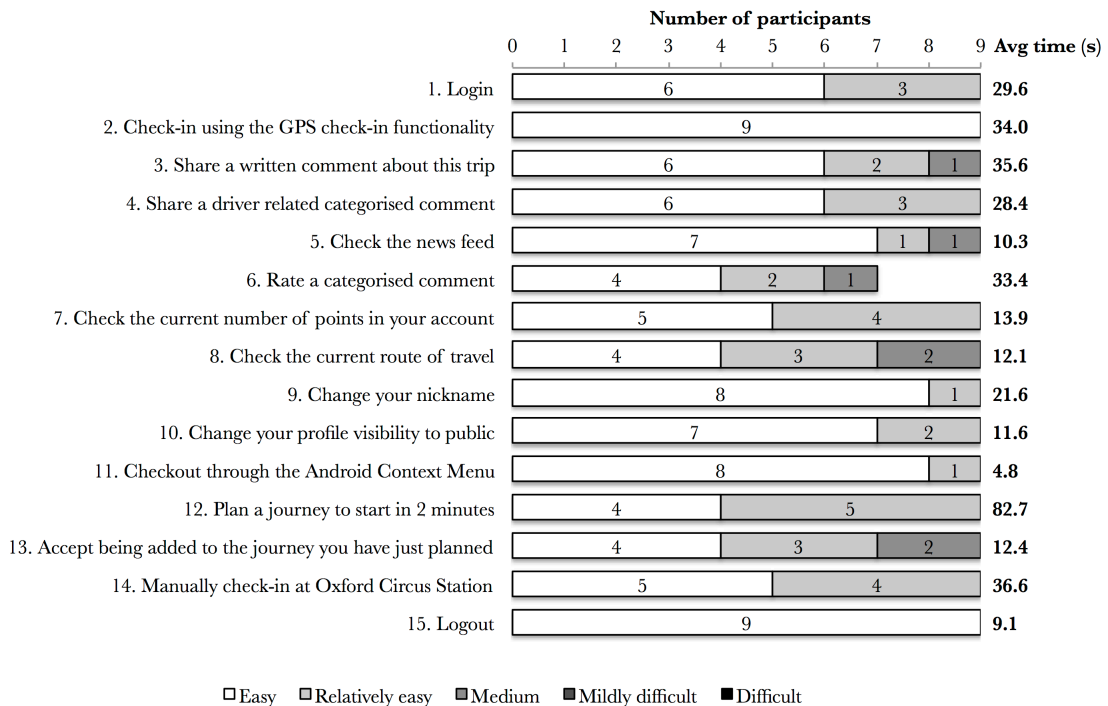


Figure 6.8 Post-task questionnaire results. Adapted from Nunes, Gonçalves, and Galvão (2013).

The overall ease of use was indicative of its intuitiveness and absence of major usability issues. The participants also had the opportunity to provide the facilitator with some insight as to why they found some tasks less easy to perform. The least encouraging results were the comparatively higher difficulty of some fundamental interaction tasks, such as sharing and rating categorised comments. This was identified as a key aspect for improvement in the following iteration of *Journata*. Conversely, the ease of checking-in using GPS was a success, and so was the simplicity for users to configure their privacy settings. The latter is important to raise user confidence on the service. The time it took participants to complete the tasks was recorded to provide some indication if the average was reasonable, and to identify significant variability among participants for the same task. Some variability was present, but mainly related with the familiarity of users with the Android system making it difficult to extract any elaborate conclusions. Overall, the results did not raise concerns. All but one task, relating to journey planning, were completed on average between 5 and 37 seconds by participants who were having their very first contact with the mobile application. Journey planning took on average just over 80 seconds, which is not surprising, given the requirement for users to input various parameters of the journey.

Figure 6.9 summarises measurable results from the post-task questionnaire. It only depicts the frequencies of qualitative assessment responses. It harvested interesting insights in terms of validating the concept. All participants found the mobile application prototype to be ‘useful’ or ‘somewhat useful’, which was indicative of its potential to foster value co-creation in urban passenger transport. All participants found it ‘easy’ or ‘relatively easy’ to interact with, suggesting a high usability level. The ease of comprehension of the concept was the main concern, yet seven participants found it ‘easy’ or ‘relatively easy’ to grasp, and two others perceived ‘medium’ difficulty. This result encouraged further work on the development of *Journata* and highlighted a need to focus on making it clearer for potential users. All participants considered the prototype mobile application to be ‘safe’ or ‘somewhat safe’ to use, suggesting that the privacy options were likely adequate, understandable, and easy to configure. Eight participants indicated that they were ‘likely’ or ‘somewhat likely’ to use *Journata*, whereas one was ‘neutral’ about it. This suggested

that further iterations of the prototype should also focus on making it appealing for potential users.

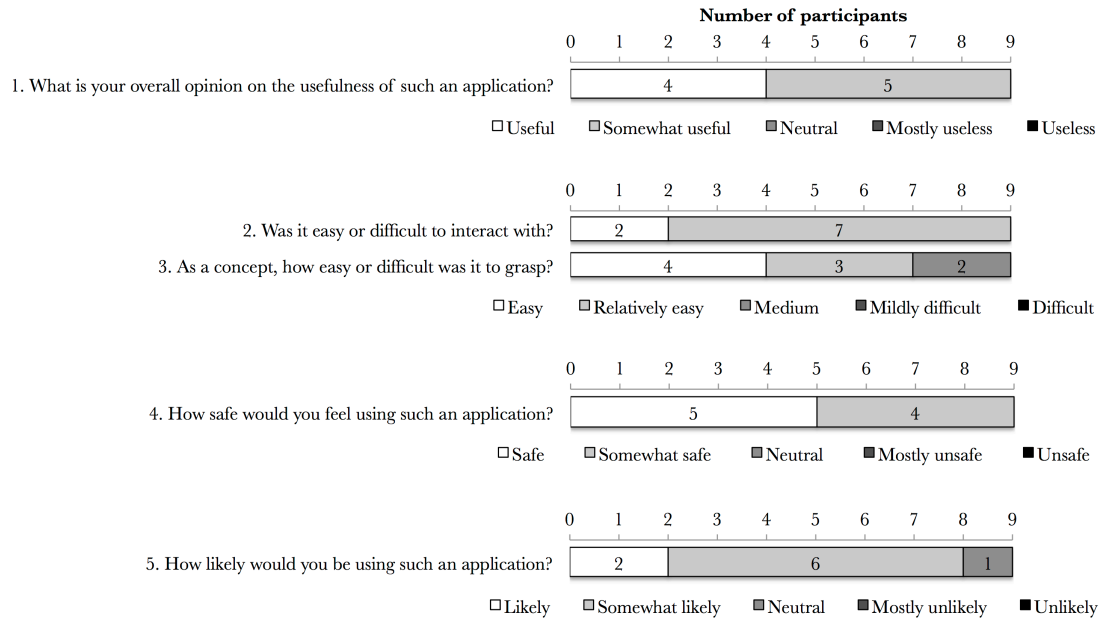


Figure 6.9 Post-test questionnaire results. Adapted from Nunes, Gonçalves, and Galvão (2013).

It is emphasised that the above results must be read with caution. Unlike the usability test, the number of participants was reduced for generalising conclusions. The participants revealed some ideas in their interviews for enhancing the appeal of *Journata*. Those ideas included reinforcing the gaming nature of the mobile application, providing a first time use tutorial emphasising potential usage benefits, and integrating the application with existing social networking services to avoid the creation of user accounts, leverage their popularity, and facilitate the provision of contextual information. Additionally, some participants questioned the feasibility of the validation mechanism based on rating comments prior to being more widely distributed. This issue not only delays the distribution of potentially relevant time-specific information (e.g. report of an incident), but also it may be difficult to enforce outside peak hours due to lower occupancy of public transport vehicles.

6.3. Functional prototype, second iteration

The second iteration of *Journata* as a functional prototype was developed in collaboration with another Masters student in Informatics and Computing Engineering, who overhauled the prototype mobile application interface and carried out its evaluation for his own thesis with the author of this thesis as co-advisor (Amador 2014). This iteration was based on the existing architecture and focussed on bringing the mobile application interface up-to-date, whilst addressing known interaction limitations. The second iteration of the mobile application remained implemented in a stand-alone format. In terms of the web service and database components, the scope of modifications was minor.

The development of the second iteration of *Journata* consisted fundamentally of another loop around the four activity cycle of user-centred design (Preece, Rogers, and Sharp 2002) described in Section 6.2. The first, dealing with the identification of requirements, consisted of the involvement of potential final users at a focus group. The session was not intended to be as exploratory as the brainstorming held in the development of the first iteration. It was organised around a predefined set of use cases for participants to focus on, with the objective of refining the requirement specification. Preliminary interface design options were prepared in mock-ups beforehand for some use cases (Figure 6.10), to be presented during the focus group and stimulate active discussion on their advantages and disadvantages. The session lasted approximately two hours and brought together a group of six participants and a moderator, of varied age groups, levels of proficiency with mobile use, and backgrounds. It included public transport users and transport specialists. The session yielded a variety of ideas, some of which required a major redesign from the previous iteration of the mobile application prototype.



Figure 6.10 Mock-up samples of preliminary interface designs (source: Amador 2014)

Two key issues emerged from the focus group that were not accounted for in the first iteration of *Journata*. The first was the requirement from users to start receiving information about their journeys in advance to assist travel choices. In the first iteration, receiving information required a prior check-in into a specific route, which may be unknown prior to entering a vehicle. The second was the desire to receive information about more than one journey at the same time, allowing users to monitor other routes in which they may need to travel in the near future. Again, the first iteration of *Journata* would only let users receive information about the route in which they were checked-in. Moreover, the concept of TUNs had evolved significantly since the first iteration with clear impacts on the interface. For example, since users receive information not only about their planned route (similar journeys), but also about feasible alternatives (substitute journeys), it became essential to identify clearly in the news feed where that information originates.

Along with the aforementioned issues, a suggestion that emerged from the focus group and was adopted, consisted of introducing news feed subscriptions. Essentially, this entailed a dissociation of the notions of check-in and subscription. A user may have several active subscriptions for receiving information, each of those associated with a journey between two locations, and optionally with specific dates and times. However, for the purposes of providing and rating information, the user needs to be actually checked-in into a journey, ensuring that user actually knows about it. The

implementation of multiple feed subscriptions requires creation of a TUN for each simultaneous subscription of each user.

Apart from the focus group, another change that was required from the previous iteration of *Journata* related to the validation mechanism. A solution found was not to require validation prior to distributing information more widely, yet allow users to rate it. In practical terms this offers a clear benefit regarding the timeliness of information. But also in conceptual terms, it emphasises trust on users, who will be assumed to provide reliable information unless they prove themselves untrustworthy. Two rating options were considered, one that only allows users to ‘report’ inaccurate information, akin to Facebook or Instagram for example, and another that allows users to ‘upvote’ or ‘downvote’ information, akin to Quora (quora.com) or Stack Overflow (stackoverflow.com) for example. The latter was adopted. This type of validation mechanism facilitates the ordering of information in news feeds based on the reputation of users who provide it, and even block the distribution of information for any extreme cases of poor reputation.

Regarding the second main activity of user-centred design (developing alternative designs), the process had already been initiated with the production of preliminary interface design options that were presented at the focus group session (Figure 6.10). That work was continued afterwards, to provide options for all mobile application screens and incorporating conclusions drawn from the first activity. The work involved in the development of alternative designs was extensive and detailed, and considered and documented all aspects of user interaction with the mobile application (Amador 2014).

Regarding the third main activity (building interactive versions) the process led towards implementation of the second version of the interactive functional prototype of the mobile application in Android. Figure 6.11 shows some of the main mobile application screens. Note that the number of tabs grouping actions was reduced from five to four, largely as a result of aggregation between the news feed and rating of comments. The screen on the top left relates to the news feed and allows users to view, select, and search for feeds associated with specific journeys. The screens on the

top centre and right relate to providing comments and respectively allow users to submit written and categorised comments. The screen on the bottom left relates to the journey planner and allows users to set favourite trips and schedule new ones. The screens on the bottom centre and right relate to the user profile and respectively allow users to check their usage statistics and claim rewards. This second version of the interface introduced new features over and above those previously described, including journey favourites, journey scheduling made more explicit, and multiple feed selections. Whilst multiple feed subscriptions can be selected, information regarding all of them is visually presented in a single screen for ease of reading. Each item is referenced to a specific route to facilitate identification.



Figure 6.11 Second iteration functional prototype interface screens (source: Amador 2014)

The fourth basic activity of interaction design (evaluating the interactive versions) consisted of two stages subsequently described.

Evaluation

The first evaluation stage consisted of a usability test with potential users of the mock-up interface design. The second evaluation stage consisted of evaluation by experts of the implemented functional prototype interface. The usability test in the first stage was structurally equivalent to that of the previous prototype iteration, with pre-test and post-task questionnaires, and two questions being asked to participants after completing the test. Six participants, who were public transport users with smartphones and were aged between 18 and 25 years old, attended the usability test. The post-task questionnaire captured the perceived ease-of-use of a set of tasks scripted for the usability test, in a 5-point Likert scale ranging from ‘difficult’ to ‘easy’ (Table 6.4). At the end, the participants were asked if they would use such a mobile application as part of their daily routines, and which aspects they found most appealing about it.

Table 6.4 Post-task questionnaire

Task	Task difficulty (in a scale of 1: difficult to 5: easy)
1	Add a news feed subscription
2	Filter the visible news feeds
3	View a news feed information
4	Remove a news feed subscription
5	Rate a comment from another user
6	Submit a written and a categorised comment
7	Check the favourite journeys list
8	Check scheduled journeys and add a journey to that list
9	Check the user profile and the rewards list

Adapted from Amador (2014)

Regarding the second stage, a group of five experts of interaction design and development of transport-related mobile and web applications were provided with the same set of tasks to complete (Table 6.4) as a starting point. A briefing session was held beforehand to explain the concept behind *Journata*. After completing the set of tasks, the experts were asked to explore the functional prototype of the mobile

application freely. The database was fed with hypothetical data to simulate the experience of the application in a real environment.

Results

Figure 6.12 presents average results of the first evaluation stage usability test. It shows that all tasks were perceived to be easy to grasp and to complete, with the average score ranging from 4.2 to 5 in a scale of 1 ('difficult') to 5 ('easy'). Filtering feeds (task 2) and checking the user profile (task 5) were perceived to be easiest with all participants giving it a score of 5 ('easy'). Adding a news feed subscription (task 1) and viewing news feed information (task 3) were perceived to be the hardest. This was largely due to initial difficulties from participants to understand the underlying concept.

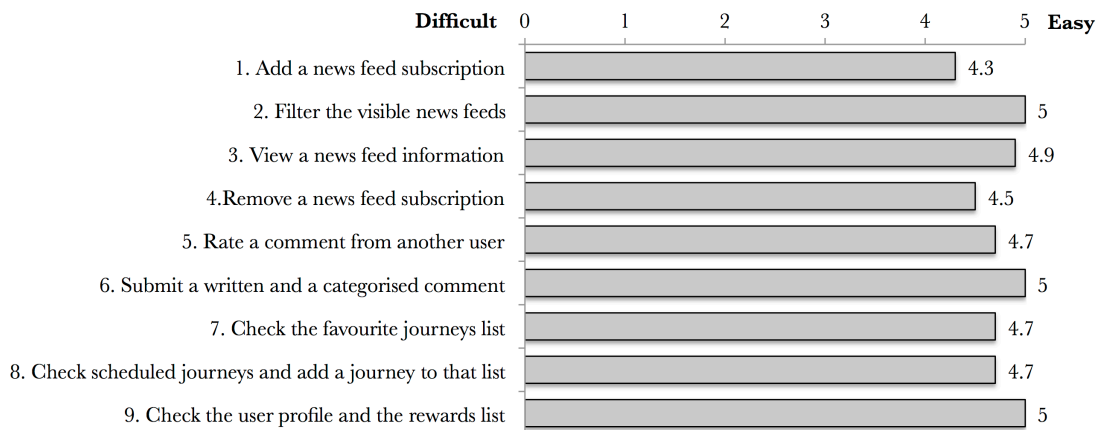


Figure 6.12 Post-task questionnaire results. Adapted from Amador (2014).

Several contributions emerged from evaluation with experts, who generally found *Journata* easy to use and intuitive. Whilst the interaction was found to be fluid, most concerns raised related to underlying concepts. A key concern was the ability to identify the vehicle from which a user provides information. A proposed solution was the integration of check-in with a mobile-based ticketing system, which had already been identified in the first iteration, and later materialised into part of *Seamless Mobility* (see Section 6.4). The experts were also concerned with the complexity in the notion of news feeds, and suggested that it could be made simpler by limiting to a

single journey stage instead of entire journeys. Adoption of this suggestion would however hinder the potential of TUNs in terms of providing users with information on travel alternatives. The experts praised the simplicity of the comment rating feature, but warned that an adopted solution should clearly convey agreement or disagreement rather than like or dislike. Its goal is ultimately about safeguarding reliability of information. Lastly, the experts endorsed some form of aggregation of comments to ease readability. Besides the above concerns, the experts identified finer interaction issues relating to navigation, feedback, swipe gestures, reduction of screen taps, consistency between actions, auto-completion features, and to the spacing, sizing, colouring, and prominence of items. These finer aspects are detailed in Amador (2014).

6.4. *Journata in Seamless Mobility*

The evolution of *Journata* eventually triggered interest outside academia, and the concept was included in a successful tender for a government-funded project. The project relates to the implementation of a system for public transport users featuring three components: journey planning, mobile payments, and social interaction. The project is being developed by an institutional consortium consisting of two companies with a vast experience in the delivery of transport-related technologies, STCP and Metro do Porto as the two main public transport operators in Porto, and FEUP. The development of the social component has been the main responsibility of the FEUP team, which included the author of this thesis. The social component is based on *Journata*.

The integrated *Seamless Mobility* service is planned to have the following features. A journey planner component that combines public transport schedules and real-time tracking data, allowing users to identify the nearest stations and stops, the next departures from a station or stop, and the best routes for a journey using their mobile devices. A mobile payments component to allow users to purchase and validate travel tickets and seasonal subscriptions using their mobile devices. It will save users from having to go to a ticket machine and from carrying a travel card, and will facilitate

the identification of the zone fare from indication of the intended journey origin and destination. Furthermore, this mobile payments component will enable the introduction of advanced commercial policies to stimulate customer attraction and retention. In the longer-term it will also promote greater efficiency of the transport system by reducing operational and maintenance costs associated with the ticketing infrastructure. Usage of the mobile payments component will generate unique passenger travel profiles in real-time, which will eventually act as a source of information for creating TUNs for the social component (see Chapter 5). All components will be accessible to users via a single mobile application available for iOS and Android devices called *One Ride*.

The social component of *One Ride* will essentially be a toned-down version of *Journata*. The design had to be simplified for integration with the combined set of features of the various components, whilst maintaining the mobile application easy to use. The social component will allow users to share information with their peers in a categorised format at the beginning, but likely also in a written format in the future. The social component will eventually learn individual travel patterns over time, to allow passengers to start receiving information that may support their travel decisions prior to the beginning of journeys. Key aspects that were initially left out in relation to the functional prototypes of *Journata* include the lack of validation and incentive mechanisms, and the absence of multiple feeds for a single user. Additionally, the set of comment categories will be smaller. However, validation and incentive mechanisms will likely be implemented in the future.

Figure 6.13 shows the current *One Ride* screens relating to the social component. The screen on the left allows users to observe aggregated ratings provided by others, as icons, regarding their active journeys. It also provides access to the screen in the centre, which allows users to rate their active journeys on crowding, noise level, punctuality, and cleanliness. The same screen is shown on the right with a selection of categorised comment ratings to be provided (originally colour coded green, amber, and red, reflecting the intensity of the rating selection).

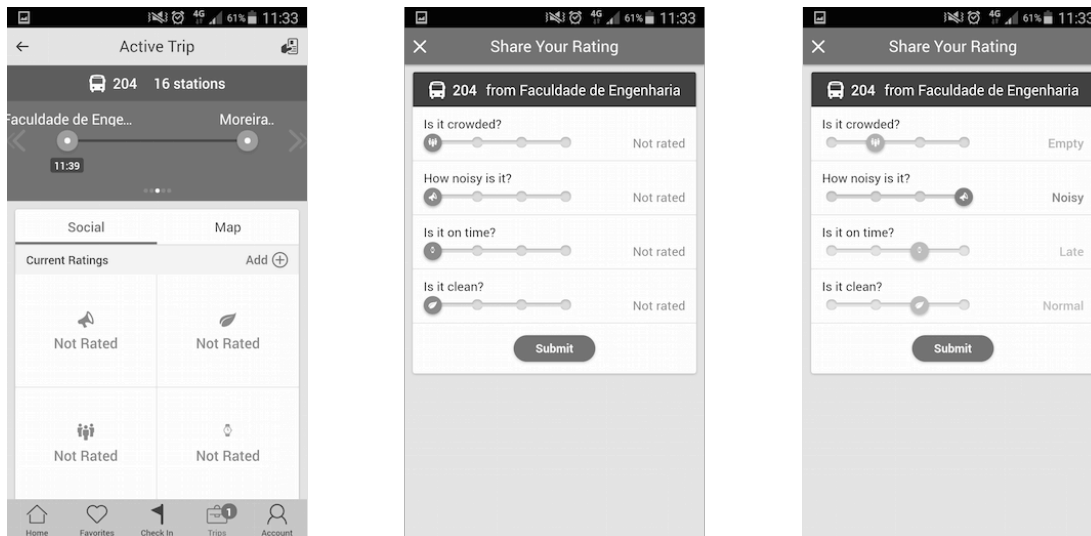


Figure 6.13 Social component interface screens of *One Ride* (source: *Seamless Mobility*)

Seamless Mobility is expected to change how passengers and transport providers interact and co-create value in an urban passenger transport service setting. Ultimately, the project is expected to contribute towards the improvement of experience associated with more sustainable travel alternatives, in line with the objectives of the user collaboration model. Large scale testing of *Seamless Mobility* has started in early 2016 as part of a field trial that will involve a large number of participants who are urban public transport users in Porto. This will offer an exciting opportunity for the empirical validation of *Journata* and of the user collaboration model.

6.5. Chapter conclusions

This chapter described the evolution of the case study *Journata* from its inception as non-functional prototype through to its implementation as a commercial prototype service. The design and evaluation of *Journata* fulfilled the last objective of this thesis (RO.3, see Section 1.2), consisting of the experimental development part of the methodology adopted. It provided a valuable feedback loop for the iterative refinement of conceptual foundations of the user collaboration model. The questionnaire-based interviews carried out alongside usability tests with potential users provided some indication of the usefulness of such a service. Those initial

observations were largely validated afterwards, firstly from the opinion of experts on the second iteration of the functional prototype, and later from the commercial interest that it generated amongst urban transport stakeholders.

The integration of *Journata* within *Seamless Mobility* is in line with the original plans and suggestions from participants in the evaluation stage, in terms of nesting the application within a broader service. However, instead of being integrated with an existing ICT-based social networking application as foreseen at that time, *Journata* became a component of a service that aims to ease and encourage urban public transport use. This was an unpredicted but fortunate outcome, since the purpose of the *Seamless Mobility* service is closely related to the main goal of this thesis, which is to promote the sustainability of urban passenger transport (Section 1.2). Furthermore, the integration of *Journata* within *Seamless Mobility* further proved the concept behind it, through its recognition amongst industry experts. Whilst *Journata* has been an output on its own, its main contribution for this thesis was empirical validation of the user collaboration model.

7. Conclusions

This thesis put forward and evaluated a model of collaboration between users of urban passenger transport. Firstly, in Chapter 2, it framed the topic theoretically with a multidisciplinary review of existing literature. Secondly, in Chapter 3, it established the conceptual foundations of the model. Thirdly, in Chapters 4 and 5, it set out the algorithmic tools for implementing the model and empirically tested them. Lastly, Chapter 6 described the evolution of the case study *Journata*, as an instance of the user collaboration model.

7.1. Findings and contributions

Three main research objectives were set out in the introduction of this thesis (Section 1.2). The first research objective was concerned with an assessment of the potential of user collaboration to improve resource utilisation and enhance travel experience. This objective was accomplished with a thorough review of the literature that informed where the potential lies. This thesis is believed to have intersected various topics in a pioneering way, and to have generated new theory as such. It initially drew from service science theory that has had limited expression in passenger transport domains, to extract key factors influencing travel experience and gather precedents of value co-creation across service domains. Then it reviewed ICT-based collaboration mechanisms against passenger information requirements identified in the literature. While all topics that were reviewed had previously been applied to transport, finding the overlap between them to inform the user collaboration model was innovative. The main conclusion drawn from that intersection to accomplish the first objective is that the potential of user collaboration for passenger transport is great and that it largely remains to be explored. It may potentially enhance travel experience making use of existing resources, which include both transport provider infrastructure and personal mobile communication devices of passengers. As a mechanism to generate collective intelligence, user collaboration was demonstrated

to address passenger information requirements, and consequently to positively influence travel experience associated with more sustainable travel alternatives.

The second research objective was concerned with the definition of conceptual foundations and tools for the user collaboration model. This objective was accomplished across the qualitative and quantitative stages of the applied research part (see Figure 1.2, reprinted here as Figure 7.1 for reading convenience). Initially, the qualitative stage established the conceptual design of a collaboration model, capable of acting upon travel experience factors by addressing the passenger information requirements identified in the literature (see Table 2.1 and Table 2.2 respectively). This research established that collaboratively generated information not only enhances travel experience per se, but may also provide additional layers of information to help travellers make and adjust travel plans to meet their unique preferences and needs. The design and evaluation of *Journata* later confirmed these points. The model hinges upon the creation of dynamic social networks, and requires the active participation of users of a passenger transport system as providers, consumers, and validators of spatiotemporally structured information.

Afterwards, the second objective was entirely accomplished with the quantitative stage of the applied research part (see Figure 1.2 reprinted as Figure 7.1). This stage was concerned with the development of two tools for the user collaboration model. The first of these consisted of the implementation of a methodology for estimating the destination of passenger journeys from AFC data to produce an O-D dataset. This was an enabling step for implementing the model, which reflected the primary interest in the urban public transport domain and the characteristics of the AFC system where the experimental data for this research came from. Therefore, this first tool may not apply to some public or private urban passenger transport domains, or may require adaptations to be defined on a case-by-case basis. The methodology built upon previous work reviewed in the literature but introduced specific considerations that apply specifically to entry-only AFC systems with distance-based fare structures. The methodology itself proved to make two contributions to the literature, relating to the introduction of endogenous spatial validation rules to test

the validity of key assumptions, and to the improved reliability of destination estimation results.

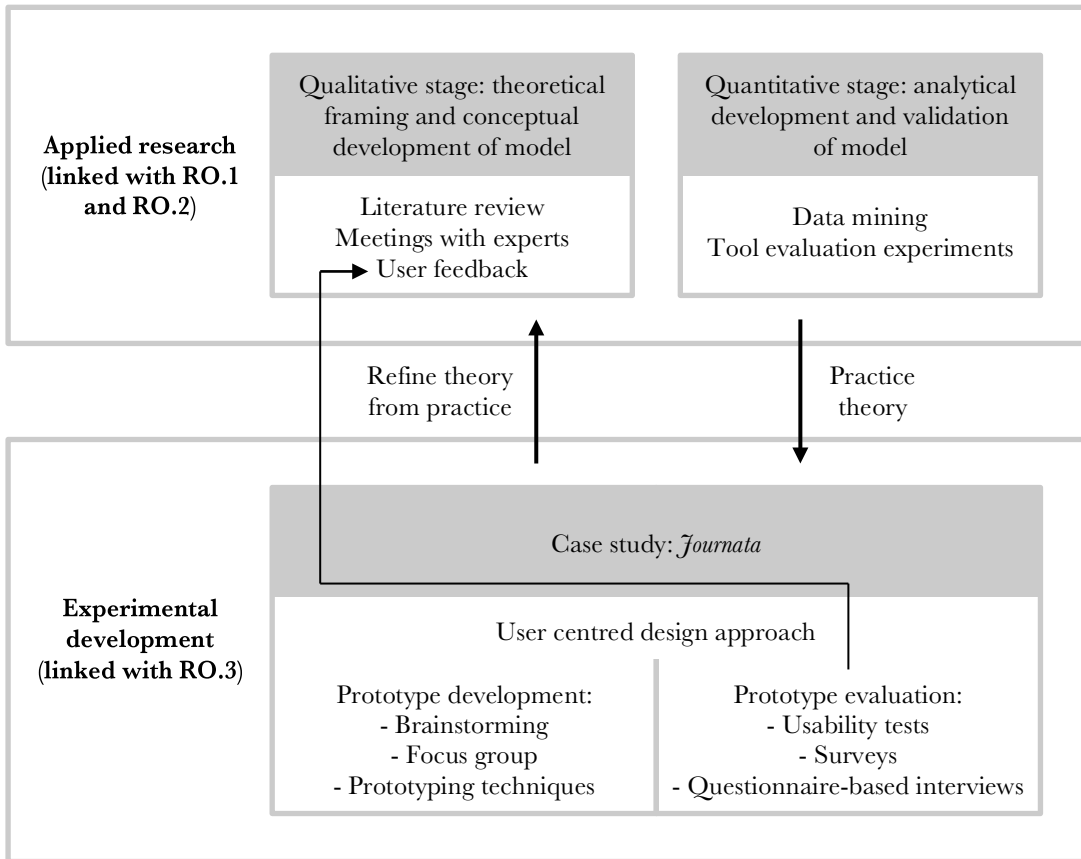


Figure 7.1 Methodology (reprinted)

The second tool consisted of implementing the concept of TUNs for passenger transport domains, which link users based on circumstantial collaboration potential. This comprised the mathematical formulation of TUNs, the structuring of an algorithm to perform required calculations, and its implementation. TUNs represent affinity as spatiotemporally dependent ties between users in a transport system. TUNs aim to facilitate the diffusion of information spread across a passenger transport system in real-time, to overcome limitations of ICT-based services that have inhibited user collaboration-based value co-creation practices across urban transport domains. TUNs deal with the transient nature of transport-related information and enable its timely diffusion. Testing the creation of TUNs with Andante data demonstrated their feasibility in light of random graph theory and

reinforced the conclusion that the user collaboration model has ample potential. Even with a subset of passenger journeys and under conservative assumptions, testing led to the conclusion that large TUNs of highly relevant peers can be created for most passengers. Owing to its original nature, TUNs are themselves a key contribution to the literature. In relation to this research, TUNs were the final element to accomplish the second research objective.

The third research objective was concerned with designing and evaluating the prototype service *Journata* based on the user collaboration model. This objective was accomplished iteratively with collaborative input from various researchers, as *Journata* evolved from a non-functional prototype towards its integration into a commercial prototype service. The last objective was entirely associated with experimental development part of the research (see Figure 1.2, reprinted as Figure 7.1). Owing mostly to its integration into the *Seamless Mobility* service, *Journata* has been an output of this thesis on its own. Still, its main contribution to this thesis has been the empirical validation of the user collaboration model.

The three research objectives were combined to address the research problem set out in the introduction of this thesis. The research problem is concerned with defining an approach to facilitate collaboration between users of an urban passenger transport system, leveraging the ubiquity of personal mobile devices and widespread availability of sensor-based data to ease their journeys. It is considered that all research objectives were successfully accomplished as described in the previous paragraphs, hence that this thesis has fully addressed the research problem, and that the overarching goal of contributing towards the sustainability of urban passenger transport is achievable.

Lastly in terms of conclusions, the main contributions of this thesis are reminded. The first main contribution is to the state-of-the-art, and consists of the intersection of theory topics to bring a service science perspective that has sparsely been associated with transport. The second main contribution is also to existing theory, and consists of the user collaboration model. At a more detailed level, each of the tools for the model also contributed to the state-of-the-art as mentioned in the

previous paragraphs. The third main contribution is the most tangible in its nature, and consists of the design and evaluation of *Journata*. It is the corollary of this thesis in the shape of value brought back to society, which justifies all the efforts dedicated to this research.

7.2. Future research directions

Future research directions identified can be divided into four main topics. The first are improvements and further evaluation of the passenger journey destination estimation methodology described in Chapter 4. The second are improvements and further evaluation of the concept of TUNs described in Chapter 5. The third is further development and evaluation of *Journata*. The fourth is deeper empirical evaluation of impacts of value co-creation practices in urban passenger transport, from large-scale usage of new services such as *Journata*.

Regarding the first topic, future work may focus on exogenous validation of the methodology once up-to-date O-D survey results become available from STCP. Improvements to the methodology itself may include an additional validation rule based on an interchange time interval as found in Gordon (2012), taking account of topography and barriers in the urban fabric to refine the maximum interchange distance spatial rule, and improvements to the multi-modal upgrade. Additional work relating to this methodology may include testing it with other AFC systems, in specific those with alternative distance-based fare structures (e.g. stages instead of zones), to evaluate transferability and benchmark the percentage of estimated destinations.

Regarding the second topic, future work may focus on the continuing development of algorithms for creating TUNs in real-time to reduce execution time further and reduce lags in the persistent update of networks. This is already ongoing in the context of the *Seamless Mobility* project. Further work may also focus on extending the concept of TUNs to include affinity measures other than SIM and SUB. An opportunity already identified is to add another measure of substitutability of locations instead of journey paths, which could help travellers find alternative

destinations to reach the same journey goal. For example, an alternative for a disrupted route for a shopping trip may be a route to another store that sells the same products or similar ones. Lastly, as with the previous topic, future work may cover additional case studies, in order to evaluate the transferability of TUNs to other public transport systems and even to other transport domains.

Regarding the third topic, the continuous improvement and evaluation of *Journata* may include further tests with users in a real environment. Pending success of its ongoing field trial, future work may stem from a more comprehensive release of *Journata* to the market, into the analysis of the distribution of information flows to provide a better understanding of the market potential and usefulness of TUNs. This type of analysis will allow the evaluation in greater depth of the benefits actually realised with the user collaboration model, both for urban transport users and transport providers. It is also hoped that future work will include the rollout of *Journata* to other urban passenger transport systems elsewhere in the world.

Regarding the fourth and last topic, the materialisation of value co-creation practices in the urban passenger transport domain will soon enable the deeper exploration of their impacts. To what extent does value co-creation raise the collective awareness of transport systems? What is the de facto impact of that relationship on personal travel choices and on society more broadly? Answering these research questions will likely be possible after the rollout and commercial maturing of services such as *Journata*, which will become data-rich case studies should they be adopted by users at a large-scale.

References

- Abou-Zeid, Maya, and Moshe Ben-Akiva. 2012. "Travel Mode Switching: Comparison of Findings from Two Public Transportation Experiments." *Transport Policy* 24: 48–59. doi:10.1016/j.tranpol.2012.07.013.
- Abou-Zeid, Maya, Regina Witter, Michel Bierlaire, Vincent Kaufmann, and Moshe Ben-Akiva. 2012. "Happiness and Travel Mode Switching: Findings from a Swiss Public Transportation Experiment." *Transport Policy* 19: 93–104. doi:10.1016/j.tranpol.2011.09.009.
- Adler, Jeffrey L., and Victor J. Blue. 1998. "Toward the Design of Intelligent Traveler Information Systems." *Transportation Research Part C* 6: 157–72. doi:10.1016/S0968-090X(98)00012-6.
- Aguiléra, Anne, Caroline Guillot, and Alain Rallet. 2012. "Mobile ICTs and Physical Mobility: Review and Research Agenda." *Transportation Research Part A: Policy and Practice* 46: 664–72. doi:10.1016/j.tra.2012.01.005.
- Alben, Lauralee. 1996. "Quality of Experience." *The ACM Interactions Design Awards Criteria* 1113: 6. doi:10.1145/235008.235010.
- Albert, Réka, and Albert-László Barabási. 2002. "Statistical Mechanics of Complex Networks." *Reviews of Modern Physics* 74 (1). American Physical Society: 47–97. doi:10.1103/RevModPhys.74.47.
- Alexander, Matthew, and Elina Jaakkola. 2011. "Exploring Value Co-Creation within Networks: Actor-to-Actor Service Provision within a Public Transport Service System." In *Industrial Marketing and Purchasing Conference*. Glasgow.
- Alford, John. 2009a. *Engaging Public Sector Clients: From Service-Delivery to Co-Production*. Houndmills: Palgrave Macmillan.
- . 2009b. *Public Value from Co-Production by Clients*.
- Alves, Helena. 2012. "Co-Creation and Innovation in Public Services." *The Service Industries Journal* iFirst: 1–12. doi:10.1080/02642069.2013.740468.
- Alves, Luiz Philipe Serrano, Ana Paula Chaves, and Igor Steinmacher. 2011. "Um Aplicativo Baseado Em Inteligência Coletiva Para Compartilhamento de Rotas Em Redes Sociais." In *Anais Do VIII Simpósio Brasileiro de Sistemas Colaborativos (SBSC)*.
- Amador, Marco André Moreira. 2014. "Journata - Designing the Interface of a Mobile Application for Exchanging Public Transport Information among Travellers." Faculdade de Engenharia da Universidade do Porto.
- Anable, Jillian, and Birgitta Gatersleben. 2005. "All Work and No Play? The Role of Instrumental and Affective Factors in Work and Leisure Journeys by Different Travel Modes." *Transportation Research Part A: Policy and Practice* 39: 163–81. doi:10.1016/j.tra.2004.09.008.

- Antrop, Marc. 2004. "Landscape Change and the Urbanization Process in Europe." *Landscape and Urban Planning* 67: 9–26. doi:10.1016/S0169-2046(03)00026-4.
- Atev, Stefan, Osama Masoud, and Nikos Papanikolopoulos. 2006. "Learning Traffic Patterns at Intersections by Spectral Clustering of Motion Trajectories." In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4851–56. doi:10.1109/IROS.2006.282362.
- Austin, John. 2010. "Use of Social Networking to Promote Public Transport and Sustainable Travel." *Local Transport Today, ITS Review*.
- Babbar, Sunil, and Xenophon Koufteros. 2008. "The Human Element in Airline Service Quality: Contact Personnel and the Customer." *International Journal of Operations & Production Management* 28 (9): 804–30. doi:10.1108/01443570810895267.
- Banks, John, and Sal Humphreys. 2008. "The Labour of User Co-Creators: Emergent Social Network Markets?" *Convergence: The International Journal of Research into New Media Technologies* 14 (4): 401–18. doi:10.1177/1354856508094660.
- Bar-Gera, Hillel. 2007. "Evaluation of a Cellular Phone-Based System for Measurements of Traffic Speeds and Travel Times: A Case Study from Israel." *Transportation Research Part C: Emerging Technologies* 15: 380–91. doi:10.1016/j.trc.2007.06.003.
- Barabási, Albert-lászló. 2003. *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. New York: Plume.
- Barry, James J., Robert Freimer, and Howard Slavin. 2009. "Use of Entry-Only Automatic Fare Collection Data to Estimate Linked Transit Trips in New York City." *Transportation Research Record: Journal of the Transportation Research Board* 2112: 53–61. doi:10.3141/2112-07.
- Barry, James J., Robert Newhouser, Adam Rahbee, and Shermeen Sayeda. 2002. "Origin and Destination Estimation in New York City with Automated Fare System Data." *Transportation Research Record: Journal of the Transportation Research Board* 1817: 183–87.
- Batty, Michael, Elena Besussi, and Nancy Chin. 2003. "Traffic, Urban Growth and Suburban Sprawl." *UCL Working Papers Series* 70: 1–12.
- Beirão, Gabriela, and J. A. Sarsfield Cabral. 2007. "Understanding Attitudes towards Public Transport and Private Car: A Qualitative Study." *Transport Policy* 14: 478–89. doi:10.1016/j.tranpol.2007.04.009.
- Berry, Leonard L., and Lewis P. Carbone. 2007. "Build Loyalty Through Experience Management." *Quality Progress* 40 (September): 26–32.
- Berry, Leonard L., Lewis P. Carbone, and Stephan H. Haeckel. 2002. "Managing the Total Customer Experience." *MIT Sloan Management Review* 43 (3): 85–89.
- Bitner, Mary Jo, Amy L. Ostrom, and Felicia N. Morgan. 2008. "Service Blueprinting: A Practical Technique for Service Innovation." *California Management Review* 50 (3): 66–95. doi:10.2307/41166446.

- Black, William Richard. 2010. *Sustainable Transportation: Problems and Solutions*. New York: Guilford Press.
- Bolla, Raffaele, and Franco Davoli. 2000. "Road Traffic Estimation from Location Tracking Data in the Mobile Cellular Network." In *IEEE Wireless Communications and Networking Conference, WCNC*, 3:1107–12. doi:10.1109/WCNC.2000.904783.
- Boschmann, E. Eric, and Mei-Po Kwan. 2008. "Toward Socially Sustainable Urban Transportation: Progress and Potentials." *International Journal of Sustainable Transportation* 2 (3): 138–57. doi:10.1080/15568310701517265.
- Boyd, danah m., and Nicole B. Ellison. 2008. "Social Network Sites: Definition, History, and Scholarship." *Journal of Computer-Mediated Communication* 13: 210–30. doi:10.1111/j.1083-6101.2007.00393.x.
- Bregman, Susan. 2012. *TCRP Synthesis 99: Uses of Social Media in Public Transportation*. Washington, DC.
- Brin, Sergey, and Lawrence Page. 1998. "The Anatomy of a Large-Scale Hypertextual Web Search Engine." *Computer Networks and ISDN Systems* 30: 107–17. doi:10.1016/S0169-7552(98)00110-X.
- Bujari, Armir, Bogdan Licar, and Claudio E. Palazzi. 2012. "Movement Pattern Recognition through Smartphone's Accelerometer." In *IEEE Consumer Communications and Networking Conference (CCNC)*, 502–6. doi:10.1109/CCNC.2012.6181029.
- Büscher, Monika, John Urry, and Katian Witchger. 2011. *Mobile Methods*. Abingdon: Routledge.
- Caceres, N., J.P. Wideberg, and F.G. Benitez. 2007. "Deriving Origin–destination Data from a Mobile Phone Network." *IET Intelligent Transport Systems* 1 (1): 15–26. doi:10.1049/iet-its:20060020.
- Calabrese, Francesco, Massimo Colonna, Piero Lovisolo, Dario Parata, and Carlo Ratti. 2011. "Real-Time Urban Monitoring Using Cell Phones: A Case Study in Rome." *IEEE Transactions on Intelligent Transportation Systems* 12 (1): 141–51. doi:10.1109/TITS.2010.2074196.
- Calabrese, Francesco, Giusy Di Lorenzo, Liang Liu, and Carlo Ratti. 2011. "Estimating Origin-Destination Flows Using Mobile Phone Location Data." *IEEE Pervasive Computing* 10 (October-December): 36–44. doi:10.1109/MPRV.2011.41.
- Calabrese, Francesco, Mi Diao, Giusy Di, Joseph Ferreira, and Carlo Ratti. 2013. "Understanding Individual Mobility Patterns from Urban Sensing Data: A Mobile Phone Trace Example." *Transportation Research Part C: Emerging Technologies* 26: 301–13. doi:10.1016/j.trc.2012.09.009.
- Cantwell, Mairead, Brian Caulfield, and Margaret O'Mahony. 2009. "Examining the Factors That Impact Public Transport Commuting Satisfaction." *Journal of Public Transportation* 12 (2): 1–21.
- Carbone, Lewis P., and Stephen H. Haeckel. 1994. "Engineering Customer Experiences." *Marketing Management* 3 (3): 8–19.

- Carreira, Rui, Lia Patrício, Renato Natal Jorge, and Chris Magee. 2014. "Understanding the Travel Experience and Its Impact on Attitudes, Emotions and Loyalty towards the Transportation Provider-A Quantitative Study with Mid-Distance Bus Trips." *Transport Policy* 31: 35–46. doi:10.1016/j.tranpol.2013.11.006.
- Carreira, Rui, Lia Patrício, Renato Natal Jorge, Chris Magee, and Qi Van Eikema Hommes. 2013. "Towards a Holistic Approach to the Travel Experience: A Qualitative Study of Bus Transportation." *Transport Policy* 25: 233–43. doi:10.1016/j.tranpol.2012.11.009.
- Carvalho, Sara, Luís Sarmiento, and Rosaldo J. F. Rossetti. 2010. "Real-Time Sensing of Traffic Information in Twitter Messages." In *4th Workshop on Artificial Transportation Systems and Simulation (ATSS @ IEEE ITSC 2010)*.
- Cassia, Fabio, and Francesca Magno. 2009. "Public Services Co-Production: Exploring the Role of Citizen Orientation." *International Journal of Quality and Service Sciences* 1 (3): 334–43. doi:10.1108/17566660911004249.
- Caulfield, Brian, and Margaret O'Mahony. 2007. "An Examination of the Public Transport Information Requirements of Users." *IEEE Transactions on Intelligent Transportation Systems* 8 (1): 21–30. doi:10.1109/TITS.2006.888620.
- Chathoth, Prakash, Levent Altinay, Robert James Harrington, Fevzi Okumus, and Eric S.W. Chan. 2013. "Co-Production versus Co-Creation: A Process Based Continuum in the Hotel Service Context." *International Journal of Hospitality Management* 32: 11–20. doi:10.1016/j.ijhm.2012.03.009.
- Chaves, Ana Paula, Igor Steinmacher, and Vaninha Vieira. 2011. "Social Networks and Collective Intelligence Applied to Public Transportation Systems: A Survey." In *VIII Simpósio Brasileiro de Sistemas Colaborativos (SBSC)*.
- Chen, Ching-Fu, and Fu-Shian Chen. 2010. "Experience Quality, Perceived Value, Satisfaction and Behavioral Intentions for Heritage Tourists." *Tourism Management* 31: 29–35. doi:10.1016/j.tourman.2009.02.008.
- Chen, Cynthia, Ling Bian, and Jingtao Ma. 2014. "From Traces to Trajectories: How Well Can We Guess Activity Locations from Mobile Phone Traces?" *Transportation Research Part C: Emerging Technologies* 46: 326–37. doi:10.1016/j.trc.2014.07.001.
- Cho, Eunjoon, Seth A. Myers, and Jure Leskovec. 2011. "Friendship and Mobility: User Movement In Location-Based Social Networks." In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1082–90.
- Chorus, Caspar G., Theo A. Arentze, Harry J. P. Timmermans, Eric J. E. Molin, and Bert Van Wee. 2007. "Travelers' Need for Information in Traffic and Transit: Results from a Web Survey." *Journal of Intelligent Transportation Systems* 11 (2): 57–67. doi:10.1080/15472450701293841.
- Chorus, Caspar, Eric Molin, and Bert Van Wee. 2006. "Use and Effects of Advanced Traveller Information Services (ATIS): A Review of the Literature." *Transport Reviews* 26 (2): 127–49. doi:10.1080/01441640500333677.

- Chorus, Caspar, Joan Walker, and Moshe Ben-Akiva. 2010. "The Value of Travel Information: A Search-Theoretic Approach." *Journal of Intelligent Transportation Systems* 14 (3): 154–65. doi:10.1080/15472450.2010.484746.
- Chu, Ka Kee Alfred, and Robert Chapleau. 2008. "Enriching Archived Smart Card Transaction Data for Transit Demand Modeling." *Transportation Research Record: Journal of the Transportation Research Board* 2063: 63–72. doi:10.3141/2063-08.
- Cirillo, Cinzia, Laura Eboli, and Gabriella Mazzulla. 2011. "On the Asymmetric User Perception of Transit Service Quality." *International Journal of Sustainable Transportation* 5: 216–32. doi:10.1080/15568318.2010.494231.
- Costa, Pedro Maurício, Jeremy Pitt, João Falcão e Cunha, and Teresa Galvão. 2012. "Cloud2Bubble: Enhancing Quality of Experience in Mobile Cloud Computing Settings. A Framework for System Design and Development in Smart Environments." In *MCS 2012 The Third International Workshop on Mobile Cloud Computing & Services*.
- Cotfas, Liviu Adrian, Mihai Cătălin Croicu, and Dumitru Cotfas. 2009. "A Collaborative GIS Solution for Public Transport." *Informatica Economică* 13 (2): 50–58.
- de Montjoye, Yves-Alexandre, César A. Hidalgo, Michel Verleysen, and Vincent D. Blondel. 2013. "Unique in the Crowd: The Privacy Bounds of Human Mobility." *Scientific Reports* 3 (1376): 1–5. doi:10.1038/srep01376.
- de Oña, Juan, Rocio de Oña, and Francisco J. Calvo. 2012. "A Classification Tree Approach to Identify Key Factors of Transit Service Quality." *Expert Systems with Applications* 39: 11164–71. doi:10.1016/j.eswa.2012.03.037.
- de Oña, Juan, Rocío de Oña, Laura Eboli, and Gabriella Mazzulla. 2013. "Perceived Service Quality in Bus Transit Service: A Structural Equation Approach." *Transport Policy* 29: 219–26. doi:10.1016/j.tranpol.2013.07.001.
- dell'Olio, Luigi, Angel Ibeas, and Patricia Cecin. 2011. "The Quality of Service Desired by Public Transport Users." *Transport Policy* 18: 217–27. doi:10.1016/j.tranpol.2010.08.005.
- dell'Olio, Luigi, Angel Ibeas, and Patricia Cecin. 2010. "Modelling User Perception of Bus Transit Quality." *Transport Policy* 17: 388–97. doi:10.1016/j.tranpol.2010.04.006.
- Dodge, Somayeh, Patrick Laube, and Robert Weibel. 2012. "Movement Similarity Assessment Using Symbolic Representation of Trajectories." *International Journal of Geographical Information Science* 26 (9): 1563–88. doi:10.1080/13658816.2011.630003.
- Eboli, Laura, and Gabriella Mazzulla. 2011. "A Methodology for Evaluating Transit Service Quality Based on Subjective and Objective Measures from the Passenger's Point of View." *Transport Policy* 18: 172–81. doi:10.1016/j.tranpol.2010.07.007.
- . 2012a. "Performance Indicators for an Objective Measure of Public Transport Service Quality." *European Transport* (51): 1–21.

- . 2012b. “Structural Equation Modelling for Analysing Passengers’ Perceptions about Railway Services.” *Procedia - Social and Behavioral Sciences* 54: 96–106. doi:10.1016/j.sbspro.2012.09.729.
- Eiter, Thomas, and Heikki Mannila. 1994. *Computing Discrete Frechet Distance*.
- Erdmenger, Christoph, and Vivien Führ. 2005. *Hidden Subsidies for Urban Car Transportation: Public Funds for Private Transport*. Freiburg.
- Erdős, Paul, and Alfréd Rényi. 1960. “On the Evolution of Random Graphs.” *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5: 17–61.
- Eriksson, Lars, Margareta Friman, and Tommy Gärling. 2013. “Perceived Attributes of Bus and Car Mediating Satisfaction with the Work Commute.” *Transportation Research Part A: Policy and Practice* 47: 87–96. doi:10.1016/j.tra.2012.10.028.
- Ettema, Dick, Margareta Friman, Tommy Gärling, Lars E. Olsson, and Satoshi Fujii. 2012. “How in-Vehicle Activities Affect Work Commuters’ Satisfaction with Public Transport.” *Journal of Transport Geography* 24: 215–22. doi:10.1016/j.jtrangeo.2012.02.007.
- Ettema, Dick, Tommy Gärling, Lars Eriksson, Margareta Friman, Lars E. Olsson, and Satoshi Fujii. 2011. “Satisfaction with Travel and Subjective Well-Being: Development and Test of a Measurement Tool.” *Transportation Research Part F: Traffic Psychology and Behaviour* 14: 167–75. doi:10.1016/j.trf.2010.11.002.
- European Commission. 2011. *White Paper: Roadmap to a Single European Transport Area - Toward a Competitive and Resource Efficient Transport System*. Brussels.
- Evans, Andrew W. 1991. “Are Urban Bus Services Natural Monopolies?” *Transportation* 18 (2): 131–50. doi:10.1007/BF00150469.
- Evans, Simon, Sarah Hills, and Judy Orme. 2012. “Doing More for Less? Developing Sustainable Systems of Social Care in the Context of Climate Change and Public Spending Cuts.” *British Journal of Social Work* 42: 744–64. doi:10.1093/bjsw/bcr108.
- Falcão e Cunha, João, and Teresa Galvão. 2014. “State of the Art and Future Perspectives for Smart Support Services for Public Transport.” In *Service Orientation in Holonic and Multi-Agent Manufacturing and Robotics*, edited by Theodor Borangiu, Damien Trentesaux, and Andre Thomas, 544:225–34. Studies in Computational Intelligence. Springer. doi:10.1007/978-3-319-04735-5.
- Faria, Ricardo, Pedro Moura, Joaquim Delgado, and Anibal T. De Almeida. 2012. “A Sustainability Assessment of Electric Vehicles as a Personal Mobility System.” *Energy Conversion and Management* 61: 19–30. doi:10.1016/j.enconman.2012.02.023.
- Farzin, Janine M. 2008. “Constructing an Automated Bus Origin-Destination Matrix Using Farecard and Global Positioning System Data in São Paulo, Brazil.” *Transportation Research Record: Journal of the Transportation Research Board* 2072: 30–37. doi:10.3141/2072-04.
- Ferreira, Marta Campos, and Teresa Galvão Dias. 2015. “How to Encourage the Use of Public Transport? A Multiservice Approach Based on Mobile Technologies.” *Lecture Notes in Business Information Processing* 201: 314–25.

- Ferris, Brian, Kari Watkins, and Alan Borning. 2010a. "Location-Aware Tools for Improving Public Transit Usability." *Pervasive Computing* (January-March): 2–9.
- . 2010b. "OneBusAway: A Transit Traveller Information System." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1807–16.
- . 2010c. "OneBusAway: Results from Providing Real-Time Arrival Information for Public Transit." In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 1–10.
- Fick, Gavin R., and J.R. Brent Ritchie. 1991. "Measuring Service Quality in the Travel and Tourism Industry." *Journal of Travel Research* 30 (2): 2–9. doi:10.1177/004728759103000201.
- Filippi, Francesco, Gaetano Fusco, and Umberto Nanni. 2013. "User Empowerment and Advanced Public Transport Solutions." *Procedia - Social and Behavioral Sciences* 87: 3–17. doi:10.1016/j.sbspro.2013.10.590.
- Fink, Daniel, Theodoros Damoulas, Nicholas E. Bruns, Frank A. La Sorte, Wesley M. Hochachka, Carla P. Gomes, and Steve Kelling. 2014. "Crowdsourcing Meets Ecology: Hemispherewide Spatiotemporal Species Distribution Models." *AI Magazine* 35 (2): 19–30.
- Fink, Daniel, Theodoros Damoulas, and Jaimin Dave. 2013. "Adaptive Spatio-Temporal Exploratory Models: Hemisphere-Wide Species Distributions from Massively Crowdsourced eBird Data." *27th AAAI Conference on Artificial Intelligence*.
- Fire, Michael, Dima Kagan, Rami Puzis, Lior Rokach, and Yuval Elovici. 2012. "Data Mining Opportunities in Geosocial Networks for Improving Road Safety." In *IEEE 27th Convention of Electrical and Electronics Engineers in Israel, IEEEI 2012*. doi:10.1109/IEEEI.2012.6377049.
- Fonzone, Achille. 2015. "What Do You Do With Your App? A Study Of Bus Rider Decision-Making with Real-Time Passenger Information." In *Transportation Research Board 94th Annual Meeting*.
- Foth, Marcus, and Ronald Schroeter. 2010. "Enhancing the Experience of Public Transport Users with Urban Screens and Mobile Applications." *Proceedings of the 14th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '10*, 33–40. doi:10.1145/1930488.1930496.
- Friedrich, Markus, Katrin Immisch, Prokop Jehlicka, Thomas Otterstätter, and Johannes Schlaich. 2010. "Generating Origin-Destination Matrices from Mobile Phone Trajectories." *Transportation Research Record: Journal of the Transportation Research Board* 2196: 93–101. doi:10.3141/2196-10.
- Friman, Margareta. 2004. "Implementing Quality Improvements in Public Transport." *Journal of Public Transportation* 7 (4): 49–65.
- Fronzizi, Risieri. 1971. *What Is Value? An Introduction to Axiology*. Illinois: Open Court.
- Frow, Pennie, and Adrian Payne. 2007. "Towards the 'perfect' Customer Experience." *Journal of Brand Management* 15: 89–101. doi:10.1057/palgrave.bm.2550120.

- Gal-Tzur, Ayelet, Susan M. Grant-Muller, Tsvi Kuflik, Einat Minkov, Silvio Nocera, and Itay Shoor. 2014. "The Potential of Social Media in Delivering Transport Policy Goals." *Transport Policy* 32: 115–23. doi:10.1016/j.tranpol.2014.01.007.
- Gal-Tzur, Ayelet, Susan M. Grant-Muller, Einat Minkov, and Silvio Nocera. 2014. "The Impact of Social Media Usage on Transport Policy: Issues, Challenges and Recommendations." *Procedia - Social and Behavioral Sciences* 111: 937–46. doi:10.1016/j.sbspro.2014.01.128.
- Gao, Huiji, Geoffrey Barbier, and Rebecca Goolsby. 2011. "Harnessing the Crowdsourcing Power of Social Media for Disaster Relief." *IEEE Intelligent Systems* 26 (May/June): 10–14. doi:10.1109/MIS.2011.52.
- Gao, Yunjun, Baihua Zheng, Gencai Chen, and Qing Li. 2009. "Algorithms for Constrained K-Nearest Neighbor Queries over Moving Object Trajectories." *GeoInformatica* 14: 241–76. doi:10.1007/s10707-009-0084-5.
- Gatersleben, Birgitta, and David Uzzell. 2007. "Affective Appraisals of the Daily Commute: Comparing Perceptions of Drivers, Cyclists, Walkers, and Users of Public Transport." *Environment and Behavior* 39 (3): 416–31. doi:10.1177/0013916506294032.
- Gault, Paul, David Corsar, Peter Edwards, John D. Nelson, and Caitlin Cottrill. 2014. "You'll Never Ride Alone: The Role of Social Media in Supporting the Bus Passenger Experience." In *2014 Ethnographic Praxis in Industry Conference Proceedings*, 199–212.
- Gebauer, Heiko, Mikael Johnson, and Bo Enquist. 2010. "Value Co-Creation as a Determinant of Success in Public Transport Services: A Study of the Swiss Federal Railway Operator (SBB)." *Managing Service Quality* 20 (6): 511–30. doi:10.1108/09604521011092866.
- Geetika, and Shefali Nandan. 2010. "Determinants of Customer Satisfaction on Service Quality: A Study of Railway Platforms in India." *Journal of Public Transportation* 13 (1): 97–113.
- Gehl, Jan. 2010. *Cities for People*. Washington DC: Island Press.
- Geiger, David, Michael Rosemann, and Erwin Fiel. 2011. "Crowdsourcing Information Systems – A Systems Theory Perspective." In *22nd Australasian Conference on Information Systems*.
- Gentile, Chiara, Nicola Spiller, and Giuliano Noci. 2007. "How to Sustain the Customer Experience: An Overview of Experience Components That Co-Crete Value With the Customer." *European Management Journal* 25 (5): 395–410. doi:10.1016/j.emj.2007.08.005.
- Glaeser, Edward L., and Matthew E. Kahn. 2004. "Sprawl and Urban Growth." In *Handbook of Regional and Urban Economics*, edited by J. Vernon Henderson and Jacques-François Thisse, 4:2481–2527. Amsterdam: Elsevier. doi:10.1016/S0169-7218(04)07056-X.
- Gonçalves, Tiago. 2012. "A Smartphone Application Prototype for Exchanging Valuable Real Time Public Transport Information among Travellers." Faculdade de Engenharia da Universidade do Porto.

- González, Marta C., César A. Hidalgo, and Albert-László Barabási. 2008. "Understanding Individual Human Mobility Patterns." *Nature* 453: 779–82. doi:10.1038/nature06958.
- Gordon, Jason B. 2012. "Intermodal Passenger Flows on London's Public Transport Network: Automated Inference of Full Passenger Journeys Using Fare-Transaction and Vehicle-Location Data." Massachusetts Institute of Technology.
- Granovskii, Mikhail, Ibrahim Dincer, and Marc A. Rosen. 2006. "Economic and Environmental Comparison of Conventional, Hybrid, Electric and Hydrogen Fuel Cell Vehicles." *Journal of Power Sources* 159: 1186–93. doi:10.1016/j.jpowsour.2005.11.086.
- Greengard, Samuel. 2011. "Following the Crowd." *Communications of the ACM* 54 (2): 20–22. doi:10.1145/1897816.1897824.
- Grönroos, Christian. 2008. "Service Logic Revisited: Who Creates Value? And Who Co-Creates?" *European Business Review* 20 (4): 298–314. doi:10.1108/09555340810886585.
- . 2011. "Value Co-Creation in Service Logic: A Critical Analysis." *Marketing Theory* 11 (3): 279–301. doi:10.1177/1470593111408177.
- Grönroos, Christian, and Päivi Voima. 2013. "Critical Service Logic: Making Sense of Value Creation and Co-Creation." *Journal of the Academy of Marketing Science* 41: 133–50. doi:10.1007/s11747-012-0308-3.
- Gudmundsson, Joachim, Patrick Laube, and Thomas Wolle. 2008. "Movement Patterns in Spatio-Temporal Data." *Encyclopedia of GIS*, 726–32.
- Gupta, Sunil, and Donald R. Lehmann. 2005. *Managing Customers as Investments*. Upper Saddle River, NJ: Wharton School.
- Güting, Ralf Hartmut, Thomas Behr, and Jianqiu Xu. 2010. "Efficient K-Nearest Neighbor Search on Moving Object Trajectories." *The VLDB Journal* 19: 687–714. doi:10.1007/s00778-010-0185-7.
- Habib, Khandker M. Nurul, Lina Kattan, and Md. Tazul Islam. 2011. "Model of Personal Attitudes towards Transit Service Quality." *Journal of Advanced Transportation* 45: 271–85. doi:10.1002/atr.
- Haeckel, Stephan H., Lewis P. Carbone, and Leonard L. Berry. 2003. "How to Lead the Customer Experience." *Marketing Management* (January/February): 18–23.
- Han, Ji, and Yoshitsugu Hayashi. 2008. "Assessment of Private Car Stock and Its Environmental Impacts in China from 2000 to 2020." *Transportation Research Part D: Transport and Environment* 13: 471–78. doi:10.1016/j.trd.2008.09.007.
- Harris, Kim, and Steve Baron. 2004. "Consumer-to-Consumer Conversations in Service Settings." *Journal of Service Research* 6 (3): 287–303. doi:10.1177/1094670503260132.
- Hemminki, Samuli, Petteri Nurmi, and Sasu Tarkoma. 2013. "Accelerometer-Based Transportation Mode Detection on Smartphones." In *SenSys '13 The 11th ACM Conference on Embedded Networked Sensor Systems*. doi:10.1145/2517351.2517367.

- Herring, Ryan, Aude Hofleitner, Saurabh Amin, Tania Abou Nasr, Amin Abdel Khalek, Pieter Abbeel, and Alexandre Bayen. 2010. "Using Mobile Phones to Forecast Arterial Traffic through Statistical Learning." In *Transportation Research Board Annual Meeting*.
- Heskett, James L., W. Earl Sasser Jr., and Leonard A. Schlesinger. 1997. *The Service Profit Chain: How Leading Companies Link Profit and Growth to Loyalty, Satisfaction, and Value*. New York: Free Press.
- Holbrook, Morris, ed. 1999. *Consumer Value: A Framework for Analysis and Research*. London: Routledge.
- Holub, Amalia R. 2014. "Street-Level Air Quality: Exploring the Feasibility of a Cyclist-Led, Crowdsourced Map in Singapore and Mexico." Massachusetts Institute of Technology.
- Howe, Jeff. 2006. "The Rise of Crowdsourcing." *Wired* 14 (6).
- Hu, Shou-ren, and Chang-Ming Wang. 2008. "Vehicle Detector Deployment Strategies for the Estimation of Network Origin–Destination Demands Using Partial Link Traffic Counts." *IEEE Transactions on Intelligent Transportation Systems* 9 (2): 288–300. doi:10.1109/TITS.2008.922883.
- Huang, Shih-Chia, Ming-Kai Jiau, and Chih-Hsiang Lin. 2014a. "A Genetic Algorithm-Based Approach to Solve Carpool Service Problem in Cloud Computing." *IEEE Transactions on Intelligent Transportation Systems* 16 (1): 352–64.
- . 2014b. "Optimization of the Carpool Service Problem via a Fuzzy Controlled Genetic Algorithm." *IEEE Transactions on Fuzzy Systems*. doi:10.1109/TFUZZ.2014.2374194.
- Hwang, Jung-Rae, Hye-Young Kang, and Ki-Joune Li. 2005. "Spatio-Temporal Similarity Analysis Between Trajectories on Road Networks." In *Perspectives in Conceptual Modeling*, edited by Jacky Akoka, Stephen W. Liddle, Il-Yeol Song, Michela Bertolotto, Isabelle Comyn-Wattiau, Willem-Jan van den Heuvel, Manuel Kolp, Juan Trujillo, Christian Kop, and Heinrich C. Mayr, 3770:280–89. Lecture Notes in Computer Science. Springer. doi:10.1007/11568346_30.
- Iryo, Takamasa, Kazuma Yamabe, and Yasuo Asakura. 2012. "Dynamics of Information Generation and Transmissions through a Social Network in Non-Recurrent Transport Behaviour." *Transportation Research Part C: Emerging Technologies* 20: 236–51. doi:10.1016/j.trc.2011.05.011.
- Isaacman, Sibren, Richard Becker, Ramón Cáceres, Stephen Kobourov, James Rowland, and Alexander Varshavsky. 2010. "A Tale of Two Cities." In *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications - HotMobile '10*. doi:10.1145/1734583.1734589.
- Jaakkola, Elina, and Matthew Alexander. 2014. "The Role of Customer Engagement Behavior in Value Co-Creation: A Service System Perspective." *Journal of Service Research* 17 (3): 247–61. doi:10.1177/1094670514529187.

- Jansen, Gijsbertus R.M. 1993. “Commuting: Home Sprawl, Job Sprawl, Traffic Jams.” In *A Billion Trips a Day: Tradition and Transition in European Travel Patterns*, edited by Ilan Salomon, Piet Bovy, and Jean-Pierre Orfeuil. Transportation Research, Economics and Policy. Dordrecht: Kluwer Academic Publishers.
- Järv, Olle, Rein Ahas, and Frank Witlox. 2014. “Understanding Monthly Variability in Human Activity Spaces: A Twelve-Month Study Using Mobile Phone Call Detail Records.” *Transportation Research Part C: Emerging Technologies* 38: 122–35. doi:10.1016/j.trc.2013.11.003.
- Jiang, Shan, Gaston A. Fiore, Yingxiang Yang, Joseph Ferreira, Emilio Frazzoli, and Marta C. González. 2013. “A Review of Urban Computing for Mobile Phone Traces: Current Methods, Challenges and Opportunities.” In *UrbComp’13*.
- Jiau, Ming Kai, Shih-Chia Huang, and Chih-Hsian Lin. 2013. “Optimizing the Carpool Service Problem with Genetic Algorithm in Service-Based Computing.” In *IEEE International Conference on Services Computing*, 478–85. doi:10.1109/SCC.2013.56.
- Johnson, Mikael, Heiko Gebauer, and Bo Enquist. 2010. “Value Creation in Public Transit Services.” In *The 11th International Research Seminar in Service Management*.
- Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47 (2): 263–91.
- . 1984. “Choices, Values, and Frames.” *American Psychologist* 39 (4): 341–50.
- Kalinowski, Tess. 2014. “TTC Puts Distance-Based Fares in No-Go Zone.” *Toronto Star*.
- Kalnis, Panos, Nikos Mamoulis, and Spiridon Bakiras. 2005. “On Discovering Moving Clusters in Spatio-Temporal Data.” In *Proceedings of the 9th International Symposium on Spatial and Temporal Databases (SSTD ’05)*, 364–81. doi:10.1007/11535331_21.
- Kang, Chaogui, Xiujun Ma, Daoqin Tong, and Yu Liu. 2012. “Intra-Urban Human Mobility Patterns: An Urban Morphology Perspective.” *Physica A: Statistical Mechanics and Its Applications* 391: 1702–17. doi:10.1016/j.physa.2011.11.005.
- Kanhere, Salil S. 2011. “Participatory Sensing: Crowdsourcing Data from Mobile Smartphones in Urban Spaces.” *12th IEEE International Conference on Mobile Data Management*, 3–6. doi:10.1109/MDM.2011.16.
- Kattan, Lina, and Baher Abdulhai. 2012. “Sensitivity Analysis of an Evolutionary-Based Time-Dependent Origin/Destination Estimation Framework.” *IEEE Transactions on Intelligent Transportation Systems* 13 (3): 1442–53. doi:10.1109/TITS.2012.2189437.
- Kleinberg, Jon M. 1999. “Authoritative Sources in a Hyperlinked Environment.” *Journal of the ACM* 46 (5): 604–32. doi:10.1145/324133.324140.
- Kotler, Philip. 2000. *Marketing Management, Millenium Edition*. Boston, MA: Pearson.

- Kusano, Kristofer D., and Hampton C. Gabler. 2012. "Safety Benefits of Forward Collision Warning, Brake Assist, and Autonomous Braking Systems in Rear-End Collisions." *IEEE Transactions on Intelligent Transportation Systems* 13 (4): 1546–55. doi:10.1109/TITS.2012.2191542.
- Lathia, Neal, Chris Smith, Jon Froehlich, and Licia Capra. 2013. "Individuals among Commuters: Building Personalised Transport Information Services from Fare Collection Systems." *Pervasive and Mobile Computing* 9: 643–64. doi:10.1016/j.pmcj.2012.10.007.
- Le Bel, Jordan L. 2005. "Beyond the Friendly Skies: An Integrative Framework for Managing the Air Travel Experience." *Managing Service Quality* 15 (5): 437–51. doi:10.1108/09604520510617293.
- Lévy, Pierre. 2010. "From Social Computing to Reflexive Collective Intelligence: The IEMML Research Program." *Information Sciences* 180: 71–94. doi:10.1016/j.ins.2009.08.001.
- Li, Daming, Yongjie Lin, Xinliang Zhao, Hongjun Song, and Nan Zou. 2011. "Estimating a Transit Passenger Trip Origin-Destination Matrix Using Automatic Fare Collection System." In *Database Systems for Advanced Applications*, edited by Jianliang Xu, Ge Yu, Shuigeng Zhou, and Rainer Unland, 502–13. Springer.
- Li, Daqing, Qiong Zhang, Enrico Zio, Shlomo Havlin, and Rui Kang. 2015. "Network Reliability Analysis Based on Percolation Theory." *Reliability Engineering & System Safety* 142: 556–62. doi:10.1016/j.res.2015.05.021.
- Li, Yuen-wah. 2003. "Evaluating the Urban Commute Experience: A Time Perception Approach." *Journal of Public Transportation* 6 (4): 41–67.
- Libai, Barak, Ruth Bolton, Marnix S. Bugel, Ko de Ruyter, Oliver Gotz, Hans Risselada, and Andrew T. Stephen. 2010. "Customer-to-Customer Interactions: Broadening the Scope of Word of Mouth Research." *Journal of Service Research* 13 (3): 267–82. doi:10.1177/1094670510375600.
- Line, Tilly, Juliet Jain, and Glenn Lyons. 2011. "The Role of ICTs in Everyday Mobile Lives." *Journal of Transport Geography* 19: 1490–99. doi:10.1016/j.jtrangeo.2010.07.002.
- Litman, Todd. 2011. *Valuing Transit Service Quality Improvements: Considering Comfort and Convenience In Transport Project Evaluation*.
- . 2015. *Evaluating Public Transit Benefits and Costs: Best Practices Guidebook*.
- Long, Jed A., and Trisalyn A. Nelson. 2013. "A Review of Quantitative Methods for Movement Data." *International Journal of Geographical Information Science* 27 (2): 292–318. doi:10.1080/13658816.2012.682578.
- Lovell, David J. 2001. "Accuracy of Speed Measurements From Cellular Phone Vehicle Location Systems." *Intelligent Transportation Systems* 6: 303–25. doi:10.1080/10248070108903698.
- Low, Nicholas. 2003. "Is Urban Transport Sustainable?" In *Making Urban Transport Sustainable*, edited by Nicholas Low and Brendan Gleeson, 1–22. Houndmills: Palgrave Macmillan. doi:10.1057/9780230523838preview.

- Lucy, William H., and David L. Phillips. 1997. "The Post-Suburban Era Comes to Richmond: City Decline, Suburban Transition, and Exurban Growth." *Landscape and Urban Planning* 36: 259–75. doi:10.1016/S0169-2046(96)00358-1.
- Lüke, Karl-Heinz, Holger Mügge, Matthias Eisemann, and Anke Telschow. 2009. "Integrated Solutions and Services in Public Transport on Mobile Devices." In *9th International Conference on Innovative Internet Community Systems (I2CS)*, 109–20.
- Lusch, Robert F., Stephen L. Vargo, and Matthew O'Brien. 2007. "Competing through Service: Insights from Service-Dominant Logic." *Journal of Retailing* 83 (1): 5–18. doi:10.1016/j.jretai.2006.10.002.
- Ma, Xiao-lei, Yin-hai Wang, Feng Chen, and Jian-feng Liu. 2012. "Transit Smart Card Data Mining for Passenger Origin Information Extraction." *Journal of Zhejiang University SCIENCE C (Computers & Electronics)* 13 (10): 750–60. doi:10.1631/jzus.C12a0049.
- Maklan, Stan, and Philipp Klaus. 2011. "Customer Experience: Are We Measuring the Right Things?" *International Journal of Marketing Research* 53 (6): 771–92. doi:10.2501/IJMR-53-6-771-792.
- Manovich, Lev. 2009. "The Practice of Everyday (Media) Life: From Mass Consumption to Mass Cultural Production?" *Critical Inquiry* 35 (2): 319–31.
- Meyer, Christopher, and André Schwager. 2007. "Understanding Customer Experience." *Harvard Business Review*.
- Misra, Aditi, Aaron Gooze, Kari Watkins, Mariam Asad, and Christopher A. Le Dantec. 2014. "Crowdsourcing and Its Application to Transportation Data Collection and Management." *Transportation Research Record: Journal of the Transportation Research Board* 2414: 1–8. doi:10.3141/2414-01.
- Mohan, Prashanth, Venkata N. Padmanabhan, and Ramachandran Ramjee. 2008. "Nericell: Rich Monitoring of Road and Traffic Conditions Using Mobile Smartphones." In *SenSys '08 Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*, 323–36.
- Mohring, Herbert. 1972. "Optimization and Scale Economies in Urban Bus Transportation." *The American Economic Review* 62 (4): 591–604.
- Moore, Robert, Melissa L. Moore, and Michael Capella. 2005. "The Impact of Customer-to-Customer Interactions in a High Personal Contact Service Setting." *Journal of Services Marketing* 19 (7): 482–91. doi:10.1108/08876040510625981.
- Motta, Renata Almeida, Paulo Cesar Marques da Silva, and Augusto César de Mendonça Brasil. 2012. "Desafios Da Mobilidade Sustentável No Brasil." *Revista Dos Transportes Públicos Ano 34 (2º quadrimestre)*: 25–48.
- Munizaga, Marcela A., and Carolina Palma. 2012. "Estimation of a Disaggregate Multimodal Public Transport Origin–Destination Matrix from Passive Smartcard Data from Santiago, Chile." *Transportation Research Part C: Emerging Technologies* 24: 9–18. doi:10.1016/j.trc.2012.01.007.

- Munizaga, Marcela, Flavio Devillaine, Claudio Navarrete, and Diego Silva. 2014. "Validating Travel Behavior Estimated from Smartcard Data." *Transportation Research Part C: Emerging Technologies* 44: 70–79. doi:10.1016/j.trc.2014.03.008.
- Musser, John, and Tim O'Reilly. 2007. *Web 2.0 Principles and Best Practices*. Sebastopol, CA.
- Nanni, Mirco, and Dino Pedreschi. 2006. "Time-Focused Clustering of Trajectories of Moving Objects." *Journal of Intelligent Information Systems* 27: 267–89. doi:10.1007/s10844-006-9953-7.
- Nathanail, Eftihia. 2008. "Measuring the Quality of Service for Passengers on the Hellenic Railways." *Transportation Research Part A: Policy and Practice* 42: 48–66. doi:10.1016/j.tra.2007.06.006.
- Ndoye, Mandoye, Alan M. Barker, James V. Krogmeier, and Darcy M. Bullock. 2011. "A Recursive Multiscale Correlation-Averaging Algorithm for an Automated Distributed Road-Condition-Monitoring System." *IEEE Transactions on Intelligent Transportation Systems* 12 (3): 795–808.
- Neuhofer, Barbara, Dimitrios Buhalis, and Adele Ladkin. 2012. "Conceptualising Technology Enhanced Destination Experiences." *Journal of Destination Marketing and Management* 1: 36–46. doi:10.1016/j.jdmm.2012.08.001.
- Newman, Mark E.J. 2010. *Networks: An Introduction*. Oxford: Oxford University Press.
- Nicholls, Richard. 2010. "New Directions for Customer-to-Customer Interaction Research." *Journal of Services Marketing* 24 (1): 87–97. doi:10.1108/08876041011017916.
- Nielsen, Jakob. 1989. "Usability Engineering at a Discount." In *Designing and Using Human-Computer Interfaces and Knowledge Based Systems*, edited by G. Salvendy and M.J. Smith, 394–401. Amsterdam: Elsevier.
- . 1993. *Usability Engineering*. Interactive Technologies. Cambridge, MA: Academic Press.
- Nunes, António A., Teresa Galvão Dias, and João Falcão e Cunha. 2016. "Passenger Journey Destination Estimation from Automated Fare Collection System Data Using Spatial Validation." *IEEE Transactions on Intelligent Transportation Systems* 17 (1): 133–42. doi:10.1109/TITS.2015.2464335.
- Nunes, António A., Teresa Galvão Dias, Chris Zegras, and João Falcão e Cunha. 2016. "Temporary User-Centred Networks for Transport Systems." *Transportation Research Part C: Emerging Technologies* 62: 55–69. doi:10.1016/j.trc.2015.11.006.
- Nunes, António A., Teresa Galvão, and João Falcão e Cunha. 2014. "Urban Public Transport Service Co-Creation: Leveraging Passenger's Knowledge to Enhance Travel Experience." *Procedia - Social and Behavioral Sciences* 111: 577–85. doi:10.1016/j.sbspro.2014.01.091.
- Nunes, António A., Teresa Galvão, João Falcão e Cunha, and Jeremy V. Pitt. 2011. "Using Social Networks for Exchanging Valuable Real Time Public Transport Information among Travellers." In *IEEE Conference on Commerce and Enterprise Computing*, 365–70. doi:10.1109/CEC.2011.60.

- Nunes, António A., Tiago Gonçalves, and Teresa Galvão. 2013. "A Prototype for Public Transport Service Co-Creation Using Social Media: Results from Usability Testing." In *International Conference on Exploring Service Science 1.3*.
- O'Hear, Steve. 2015. "Public Transport App Trafi Closes \$6.5M Series A Led By Octopus Investments." *TechCrunch*.
- O'Reilly, Tim. 2005. "What Is Web 2.0." *O'Reilly Network*.
- O'Reilly, Tim, and John Battelle. 2009. *Web Squared: Web 2.0 Five Years On. Web 2.0 Summit*.
- Olson, Parmy. 2014. "Why Google's Waze Is Trading User Data With Local Governments." *Forbes*.
- Olsson, Lars E., Margareta Friman, Jörg Pareigis, and Bo Edvardsson. 2012. "Measuring Service Experience: Applying the Satisfaction with Travel Scale in Public Transport." *Journal of Retailing and Consumer Services* 19: 413–18. doi:10.1016/j.jretconser.2012.04.002.
- Ory, David T., and Patricia L. Mokhtarian. 2005. "When Is Getting There Half the Fun? Modeling the Liking for Travel." *Transportation Research Part A: Policy and Practice* 39: 97–123. doi:10.1016/j.tra.2004.09.006.
- Ostrom, Elinor. 2000. "Social Capital: A Fad or a Fundamental Concept?" In *Social Capital: A Multifaceted Perspective*, edited by Partha Dasgupta and Ismail Serageldin, 172–214. Washington, DC: The World Bank.
- Otto, Julie E, and J R Brent Ritchie. 1996. "The Service Experience in Tourism." *Tourism Management* 17 (3): 165–74. doi:10.1016/0261-5177(96)00003-9.
- Palmer, Adrian. 2010. "Customer Experience Management: A Critical Review of an Emerging Idea." *Journal of Services Marketing* 24 (3): 196–208. doi:10.1108/08876041011040604.
- Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry. 1988. "SERQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality." *Journal of Retailing* 64 (1): 12–40. doi:10.1016/S0148-2963(99)00084-3.
- Paulley, Neil, Richard Balcombe, Roger Mackett, Helena Titheridge, John Preston, Mark Wardman, Jeremy Shires, and Peter White. 2006. "The Demand for Public Transport: The Effects of Fares, Quality of Service, Income and Car Ownership." *Transport Policy* 13: 295–306. doi:10.1016/j.tranpol.2005.12.004.
- Pelletier, Marie-Pier, Martin Trépanier, and Catherine Morency. 2011. "Smart Card Data Use in Public Transit: A Literature Review." *Transportation Research Part C: Emerging Technologies* 19: 557–68. doi:10.1016/j.trc.2010.12.003.
- Phithakkitnukoon, Santi, Zbigniew Smoreda, and Patrick Olivier. 2012. "Socio-Geography of Human Mobility: A Study Using Longitudinal Mobile Phone Data." *PloS One* 7 (6). doi:10.1371/journal.pone.0039253.
- Pine II, B. Joseph, and James H. Gilmore. 1998. "Welcome to the Experience Economy." *Harvard Business Review* (July-August): 97–105.
- . 1999. *The Experience Economy: Work Is Theater & Every Business a Stage*. Boston: Harvard Business School Press.

- Pitt, Jeremy, Aikaterini Bourazeri, Andrzej Nowak, Magda Roszczynska-Kurasinska, Agnieszka Rychwalska, Inmaculada Rodriguez Santiago, Maite Lopez Sanchez, Monica Florea, and Mihai Sanduleac. 2013. "Transforming Big Data into Collective Awareness." *Computer* 46 (June): 40–45. doi:10.1109/MC.2013.153.
- Plato, and Allan David Bloom. 1991. *The Republic of Plato*. New York: Basic Books.
- Pollitt, Christopher. 2010. "Cuts and Reforms – Public Services as We Move into a New Era." *Society and Economy* 32 (1): 17–31. doi:10.1556/SocEc.32.2010.1.3.
- Prahalad, C.K. 2004. "The Cocreation of Value." *Journal of Marketing* 68 (1): 23.
- Prahalad, C.K., and Venkat Ramaswamy. 2004. "Co-Creation Experiences: The next Practice in Value Creation." *Journal of Interactive Marketing* 18 (3): 5–14. doi:10.1002/dir.20015.
- Prahalad, C.K., and Venkatram Ramaswamy. 2000. "Co-Opting Customer Competence." *Harvard Business Review* (January-February): 79–87.
- Preece, Jenny, Yvonne Rogers, and Helen Sharp. 2002. *Interaction Design: Beyond Human-Computer Interaction*. New York, NY: John Wiley & Sons.
- Pucher, John, and Anders Markstedt. 1983. "Consequences of Public Ownership and Subsidies for Mass Transit: Evidence from Case Studies and Regression Analysis." *Transportation* 11: 323–45. doi:10.1007/BF00150722.
- Rao, Bharat, and Louis Minakakis. 2003. "Evolution of Mobile Location-Based Services." *Communications of the ACM* 46 (12): 61–65. doi:10.1145/953460.953490.
- Reddy, Alla, Alex Lu, Santosh Kumar, Victor Bashmakov, and Svetlana Rudenko. 2009. "Entry-Only Automated Fare-Collection System Data Used to Infer Ridership, Rider Destinations, Unlinked Trips, and Passenger Miles." *Transportation Research Record: Journal of the Transportation Research Board* 2110: 128–36. doi:10.3141/2110-16.
- Reddy, Sasank, Min Mun, Jeff Burke, Deborah Estrin, Mark Hansen, and Mani Srivastava. 2010. "Using Mobile Phones to Determine Transportation Modes." *ACM Transactions on Sensor Networks* 6 (2): 13:1–13:27. doi:10.1145/1689239.1689243.
- Richardson, Harry Ward, and Chang-Hee Christine Bae. 2004. *Urban Sprawl in Western Europe and the United States*. Urban Planning and Environment. Aldershot: Ashgate.
- Rihova, Ivana, Dimitrios Buhalis, Miguel Moital, and Mary Beth Gouthro. 2013. "Social Layers of Customer-to-Customer Value Co-Creation." *Journal of Service Management* 24 (5): 553–66. doi:10.1108/JOSM-04-2013-0092.
- Rosado, Wade. 2014. "Data Detour: Analytics Will Move Transportation Forward." *Wired*.
- Roser, Thorsten, Alain Samson, Patrick Humphreys, and Eidi Cruz-Valdivieso. 2009. *Co-Creation: New Pathways to Value*. London.
- Rowley, Jennifer. 1999. "Measuring Total Customer Experience in Museums." *International Journal of Contemporary Hospitality Management* 11 (6): 303–8.

- Rowley, Jennifer, Beata Kupiec-Teahan, and Edward Leeming. 2007. "Customer Community and Co-Creation: A Case Study." *Marketing Intelligence & Planning* 25 (2): 136–46. doi:10.1108/02634500710737924.
- Russell, James A. 1980. "A Circumplex Model of Affect." *Journal of Personality and Social Psychology* 39 (6): 1161–78. doi:10.1037/h0077714.
- Saha, Debashis, and Amitava Mukherjee. 2003. "Pervasive Computing: A Paradigm for the 21st Century." *Computer* 36 (March): 25–31. doi:10.1109/MC.2003.1185214.
- SceneSusTech. 1998. *Car Systems in the City. Report 1*. Dublin.
- Schmitt, Bernd H. 2003. *Customer Experience Management: A Revolutionary Approach to Connecting with Your Customers*. Hoboken, NJ: John Wiley & Sons.
- Seaborn, Catherine, John Attanucci, and Nigel H. M. Wilson. 2009. "Analyzing Multimodal Public Transport Journeys in London with Smart Card Fare Payment Data." *Transportation Research Record: Journal of the Transportation Research Board* 2121: 55–62. doi:10.3141/2121-06.
- Shao, Fei, Songmei Cai, and Junzhong Gu. 2010. "A Modified Hausdorff Distance Based Algorithm for 2-Dimensional Spatial Trajectory Matching." *The 5th International Conference on Computer Science & Education*, 166–72. doi:10.1109/ICCSE.2010.5593666.
- Shaw, Colin, Qaalfa Dibechi, and Steven Valden. 2010. *Customer Experience: Future Trends & Insights*. Houndmills: Palgrave Macmillan.
- Shaw, Colin, and John Ivens. 2002. *Building Great Customer Experiences*. Houndmills: Palgrave Macmillan.
- Sheller, Mimi. 2004. "Automotive Emotions: Feeling the Car." *Theory, Culture & Society* 21 (4/5): 221–42. doi:10.1177/0263276404046068.
- Shibatani, Masayoshi, and Sandra A. Thompson. 1999. *Grammatical Constructions: Their Form and Meaning*. Oxford: Oxford University Press.
- Shirabe, Takeshi. 2006. "Correlation Analysis of Discrete Motions." In *Geographic Information Science*, edited by Martin Raubal, Harvey J. Miller, Andrew U. Frank, and Michael F. Goodchild, 4197:370–82. Lecture Notes in Computer Science. Springer.
- Shoshany, Maxim, Asaf Even-Paz, and Shlomo Bekhor. 2007. "Evolution of Clusters in Dynamic Point Patterns: With a Case Study of Ants' Simulation." *International Journal of Geographical Information Science* 21 (7): 777–97. doi:10.1080/13658810601169881.
- Simini, Filippo, Marta C González, Amos Maritan, and Albert-László Barabási. 2012. "A Universal Model for Mobility and Migration Patterns." *Nature* 484: 96–100. doi:10.1038/nature10856.
- Sinha, Gaurav, and David M. Mark. 2005. "Measuring Similarity between Geospatial Lifelines in Studies of Environmental Health." *Journal of Geographical Systems* 7: 115–36. doi:10.1007/s10109-005-0153-8.

- Slovan, Lynn, Sally Cairns, Carey Newson, Jillian Anable, Alison Pridmore, and Phil Goodwin. 2010. *The Effects of Smarter Choice Programmes in the Sustainable Travel Towns: Summary Report*. London.
- Smith, Aaron, Lee Rainie, Kyley McGeeney, Scott Keeter, and Maeve Duggan. 2015. *The Smartphone Difference*. Washington, DC.
- Smith, Shaun, and Joe Wheeler. 2002. *Managing the Customer Experience: Turning Customers Into Advocates*. Harlow: Prentice Hall.
- Sohn, Keemin, and Daehyun Kim. 2008. "Dynamic Origin–Destination Flow Estimation Using Cellular Communication System." *IEEE Transactions on Vehicular Technology* 57 (5): 2703–13. doi:10.1109/TVT.2007.912336.
- Song, Chaoming, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. 2010. "Limits of Predictability in Human Mobility." *Science* 327: 1018–21. doi:10.1126/science.1177170.
- Sreenivas, Ashok, and Girish Sant. 2008. "Unravelling Myths about Subsidies in Urban Transport." *Economic and Political Weekly*, 37–40.
- Stauffer, Dietrich, and Amnon Aharony. 2003. *Introduction To Percolation Theory*. London: Taylor & Francis.
- STCP. 2014. *Relatório E Contas STCP 2013*. Porto.
- Steg, Linda. 2005. "Car Use: Lust and Must. Instrumental, Symbolic and Affective Motives for Car Use." *Transportation Research Part A: Policy and Practice* 39: 147–62. doi:10.1016/j.tra.2004.07.001.
- Steinfeld, Aaron, Shree Lakshmi Rao, Allison Tran, John Zimmerman, and Anthony Tomasic. 2013. "Co-Producing Value through Public Transit Information Services." In *Advances in the Human Side of Service Engineering*, edited by James C. Spohrer and Louis E. Freund, 231–40. Boca Raton: CRC Press.
- Steinfeld, Aaron, John Zimmerman, Anthony Tomasic, Daisy Yoo, and Rafae Dar Aziz. 2011. "Mobile Transit Information from Universal Design and Crowdsourcing." *Transportation Research Record: Journal of the Transportation Research Board* 2217: 95–102. doi:10.3141/2217-12.
- Stradling, Stephen, Michael Carreno, Tom Rye, and Allyson Noble. 2007. "Passenger Perceptions and the Ideal Urban Bus Journey Experience." *Transport Policy* 14: 283–92. doi:10.1016/j.tranpol.2007.02.003.
- Stradling, Stephen G., Jillian Anable, and Michael Carreno. 2007. "Performance, Importance and User Disgruntlement: A Six-Step Method for Measuring Satisfaction with Travel Modes." *Transportation Research Part A: Policy and Practice* 41: 98–106. doi:10.1016/j.tra.2006.05.013.
- Stuart, Kenneth R., Marc Mednick, and Johanna Bockman. 2000. "Structural Equation Model of Customer Satisfaction for the New York City Subway System." *Transportation Research Record: Journal of the Transportation Research Board* 1735: 133–37. doi:10.3141/1735-16.

- Sullivan, Brian L., Christopher L. Wood, Marshall J. Iliff, Rick E. Bonney, Daniel Fink, and Steve Kelling. 2009. "eBird: A Citizen-Based Bird Observation Network in the Biological Sciences." *Biological Conservation* 142: 2282–92. doi:10.1016/j.biocon.2009.05.006.
- Susi, Tarja, Mikael Johannesson, and Per Backlund. 2007. *Serious Games – An Overview*. Skövde.
- Toledo, Tomer, and Tanya Kolechkina. 2013. "Estimation of Dynamic Origin–Destination Matrices Using Linear Assignment Matrix Approximations." *IEEE Transactions on Intelligent Transportation Systems* 14 (2): 618–26. doi:10.1109/TITS.2012.2226211.
- Toole, Jameson L., Carlos Herrera-Yaqué, Christian M. Schneider, and Marta C. González. 2015. "Coupling Human Mobility and Social Ties."
- Transport for London. 2010. *Transport Assessment Best Practice: Guidance Document*. London.
- Trépanier, Martin, Nicolas Tranchant, and Robert Chapleau. 2007. "Individual Trip Destination Estimation in a Transit Smart Card Automated Fare Collection System." *Journal of Intelligent Transportation Systems* 11 (1): 1–14. doi:10.1080/15472450601122256.
- Tyrinopoulos, Yannis, and Georgia Aifadopoulou. 2008. "A Complete Methodology for the Quality Control of Passenger Services in the Public Transport Business." *European Transport* 38: 1–16.
- United Nations. 2014. *World Urbanization Prospects: The 2014 Revision, Highlights*. New York. doi:10.4054/DemRes.2005.12.9.
- Utah Transit Authority. 2014. "UTA Begins Fare Policy Analysis Project."
- van der Hurk, Evelien, Leo Kroon, Gábor Maróti, and Peter Vervest. 2015. "Deduction of Passengers' Route Choices From Smart Card Data." *IEEE Transactions on Intelligent Transportation Systems* 16 (1): 430–40. doi:10.1109/TITS.2014.2333583.
- Vargo, Stephen L., and Robert F. Lusch. 2004. "Evolving to a New Dominant Logic for Marketing." *Journal of Marketing* 68 (January): 1–17.
- . 2006. "Service-Dominant Logic: What It Is, What It Is Not, What It Might Be." In *The Service-Dominant Logic of Marketing: Dialog, Debate, and Directions*, 43–56. Armonk, NY: M.E. Sharpe.
- . 2008. "Service-Dominant Logic: Continuing the Evolution." *Journal of the Academy of Marketing Science* 36 (1): 1–10. doi:10.1007/s11747-007-0069-6.
- Vargo, Stephen L., Paul P. Maglio, and Melissa Archpru Akaka. 2008. "On Value and Value Co-Creation: A Service Systems and Service Logic Perspective." *European Management Journal* 26: 145–52. doi:10.1016/j.emj.2008.04.003.
- Verhoef, Peter C., Katherine N. Lemon, A. Parasuraman, Anne Roggeveen, Michael Tsiros, and Leonard A. Schlesinger. 2009. "Customer Experience Creation: Determinants, Dynamics and Management Strategies." *Journal of Retailing* 85 (1): 31–41. doi:10.1016/j.jretai.2008.11.001.

- Virzi, Robert A. 1992. "Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough?" *Human Factors* 34 (4): 457–68. doi:10.1177/001872089203400407.
- Wang, Dashun, Dino Pedreschi, Chaoming Song, Fosca Giannotti, and Albert-lászló Barabási. 2011. "Human Mobility, Social Ties, and Link Prediction." In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. doi:10.1145/2020408.2020581.
- Wang, Wei, John P. Attanucci, and Nigel H.M. Wilson. 2011. "Bus Passenger Origin-Destination Estimation and Related Analyses Using Automated Data Collection Systems." *Journal of Public Transportation* 14 (4): 131–50.
- White, J., J. Quick, and P. Philippou. 2004. "The Use of Mobile Phone Location Data for Traffic Information." In *12th IEE International Conference on Road Transport Information & Control - RTIC 2004*, 321–25. doi:10.1049/cp:20040048.
- Windmiller, Sarah, Todd Hennessy, and Kari Edison Watkins. 2014. "Accessibility of Communication Technology and the Rider Experience: Case Study of Saint Louis, Missouri, Metro." *Transportation Research Record: Journal of the Transportation Research Board* 2415: 118–26. doi:10.3141/2415-13.
- Wise, Sean, Robert A. Paton, and Thomas Gegenhuber. 2012. "Value Co-Creation through Collective Intelligence in the Public Sector: A Review of US and European Initiatives." *VINE: The Journal of Information and Knowledge Management Systems* 42 (2): 251–76. doi:10.1108/03055721211227273.
- Wolfson, Ouri, and Bo Xu. 2010. "Spatio-Temporal Databases in Urban Transportation." *Bulletin of the Technical Committee on Data Engineering* 33 (2): 18–25.
- Woodall, Tony. 2003. "Conceptualising 'Value for the Customer': An Attributional, Structural and Dispositional Analysis." *Academy of Marketing Science Review* 12.
- Work, Daniel B., and Alexandre M. Bayen. 2008. "Impacts of the Mobile Internet on Transportation Cyberphysical Systems: Traffic Monitoring Using Smartphones." *National Workshop for Research on High-Confidence Transportation Cyber-Physical Systems: Automotive, Aviation, & Rail*.
- Xuan, Yiguang, Raja Sengupta, and Yaser Fallah. 2010. "Crowd Sourcing Indoor Maps with Mobile Sensors." In *Mobile and Ubiquitous Systems: Computing, Networking, and Services 7th International ICST Conference, MobiQuitous 2010*, 125–36.
- Yanagisawa, Yutaka, Jun-ichi Akahani, and Tetsuji Satoh. 2003. "Shape-Based Similarity Query for Trajectory of Mobile Objects." In *Proceedings of the 4th International Conference on Mobile Data Management (MDM '03)*, 63–77. doi:10.1007/3-540-36389-0_5.
- Yim, Y. B. Youngbin, and Randall Cayford. 2001. *Investigation of Vehicles as Probes Using Global Positioning System and Cellular Phone Tracking: Field Operational Test*. UCB-ITS-PWP-2001-9. *California PATH Working Paper*.
- Yuan, Yihong, Martin Raubal, and Yu Liu. 2012. "Correlating Mobile Phone Usage and Travel Behavior – A Case Study of Harbin, China." *Computers, Environment and Urban Systems* 36: 118–30. doi:10.1016/j.compenvurbsys.2011.07.003.

- Zhang, Yi, Xiao Qin, Shen Dong, and Bin Ran. 2010. "Daily O-D Matrix Estimation Using Cellular Probe Data." In *Transportation Research Board Annual Meeting*.
- Zhang, Zhang, Kaiqi Huang, and Tieniu Tan. 2006. "Comparison of Similarity Measures for Trajectory Clustering in Outdoor Surveillance Scenes." *The 18th International Conference on Pattern Recognition (ICPR'06)*. doi:10.1109/ICPR.2006.392.
- Zhao, Jinhua, Adam Rahbee, and Nigel H.M. Wilson. 2007. "Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems." *Computer-Aided Civil and Infrastructure Engineering* 22: 376–87. doi:10.1111/j.1467-8667.2007.00494.x.
- Zheng, Yu, Quannan Li, Yukun Chen, Xing Xie, and Wei-ying Ma. 2008. "Understanding Mobility Based on GPS Data." In *UbiComp '08 Proceedings of the 10th International Conference on Ubiquitous Computing*, 312–21.
- Zhou, Xuesong, and Hani S. Mahmassani. 2006. "Dynamic Origin–Destination Demand Estimation Using Automatic Vehicle Identification Data." *IEEE Transactions on Intelligent Transportation Systems* 7 (1): 105–14. doi:10.1109/TITS.2006.869629.
- Zomerdijk, Leonieke G., and Christopher A. Voss. 2010. "Service Design for Experience-Centric Services." *Journal of Service Research* 13 (1): 67–82. doi:10.1177/1094670509351960.
- . 2011. "NSD Processes and Practices in Experimental Services." *Journal of Product Innovation Management* 28: 63–80.

