

TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Echtzeitsysteme und Robotik

Analysis, Planning, Control and Surveillance of
Traffic Performance Defining Components for
Robust, Sustainable and Efficient Road
Transportation Systems

Jordan Ivanchev

Vollständiger Abdruck der von der Fakultät der Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Alexander Pretschner

Prüfer der Dissertation: 1. Prof. Dr.-Ing. habil. Alois Knoll

2. Prof. P. Christopher Zegras

Die Dissertation wurde am 27.07.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 23.05.2017 angenommen.

Abstract

This thesis presents efficient computational techniques for optimizing road transportation systems using the city of Singapore as a case study. In view of urban commuting systems, the two most important determinants of traffic, which can be subjected to optimization, are identified to be the road network and the routing choices of the population. Alterations in each of them and their impact on traffic conditions are studied to evaluate the efficiency and robustness of possible optimization approaches. On one side, an optimal road network capacity redistribution is computed reducing overall travel time. Respectively, the developed BISOS algorithm provides a fast and efficient computation of near system optimum paths and achieves a significantly greater decrease of total travel time compared to the road network alteration approach and overall fuel consumption reduction. The robustness importance concept for any planned infrastructure is introduced. Traffic demands are repetitively perturbed in order to simulate long term varying traffic conditions. An embodiment of this concept is presented as a method for optimal sensor placement performing equally well for various degrees of perturbation ensuring robustness against traffic demand changes. The individual findings in the thesis are combined into a structured optimization process, split into four steps: 1) an analysis step defining measures used to identify *super-sensitive* and highly dynamical topological locations in the road network used as strategic steering tools; 2) a planning step applying hypothetical changes at identified problematic locations and evaluating the respective outcomes using a modelling and simulation approach; 3) a routing control step, executed by state of the art system optimum computation algorithm; and finally, 4) a surveillance step solving the optimal sensor placement problem, which maximizes information gain with respect to the routing choices of the driver population. The suggested methodologies and systems in this thesis can be used in order to ensure robust, sustainable and environment friendly operational state of future road transportation systems.

Zusammenfassung

Die vorliegende Dissertation beschäftigt sich mit effizienten Berechnungsmethodiken zur Optimierung von Straßenverkehrssystemen. Hierzu dient der Stadtstaat Singapur als Fallstudie. Hinsichtlich des städtischen Verkehrs gibt es primär zwei Aspekte, die optimiert werden können: Das Straßennetz, sowie die Routenwahl der Fahrer. Diese Arbeit untersucht, inwiefern diese Aspekte beeinflusst werden können, und wie sich diese Änderungen auf den Verkehrsfluss auswirken, um somit mögliche Optimierungsansätze bezüglich ihrer Effizienz und Robustheit zu bewerten.

Im ersten Schritt wird zunächst zum Zweck der Reduzierung der globalen Fahrtzeit eine optimale Neuverteilung der Straßenkapazität berechnet. Im Vergleich dazu, erreicht der hier vorgestellte BISOS-Algorithmus durch eine performante und effiziente Berechnung von nahezu optimalen Routen signifikant bessere Ergebnisse. Außerdem konnte gezeigt werden, dass der globale Benzinverbrauch durch diesen Ansatz gesenkt werden kann.

Ein weiterer Kernaspekt dieser Arbeit ist die Einführung des Konzepts der Robustheit für neu-geplante Infrastrukturen. Zur Evaluierung wurden diesbezüglich Schwankungen im Verkehrsaufkommen simuliert, damit die Robustheit gegenüber sich langfristig verändernden Verkehrsverhältnissen bewertet werden kann. Diese Methodik kann z.B. für die optimale Platzierung von infrastrukturegebundenen Verkehrsüberwachungssensoren benutzt werden, um den Nutzen des Sensornetzes auch unter sich ändernden Verkehrsaufkommen zu maximieren.

Die einzelnen Teilergebnisse dieser Arbeit werden anschließend in einen strukturierten Optimierungsprozess integriert, der in vier Schritte unterteilt werden kann: 1) Im ersten Analyseschritt werden Metriken definiert, um *super-sensitive* und hochdynamische Teile des Straßennetzes zu identifizieren. Diese können dann als Stellschrauben zur Verkehrsbeeinflussung benutzt werden. 2) Im zweiten Schritt werden theoretische Änderungen an diesen Stellschrauben mit Hilfe von Modellierung und Simulation bewertet. 3) Im nächsten Schritt werden die geplanten Routen der einzelnen Fahrzeuge mit Hilfe eines modernen Systemoptimierungsalgorithmus neu bestimmt. 4) Im letzten Überwachungsschritt wird das Problem der Sensorplatzierung gelöst, mit dem Ziel das gesammelte Wissen über die Routen der Fahrer zu maximieren.

Zusammengefasst dienen die in dieser Arbeit vorgeschlagenen Methodiken dazu, zukünftige Straßenverkehrssysteme robuster, nachhaltiger und auch umweltfreundlicher zu gestalten.

Acknowledgements

This work is a result of my three-year-long research conducted at TUM Create Ltd, which was financially supported by the Singapore National Research Foundation under its Campus for Research Excellence And Technological Enterprise (CREATE) programme.

I am deeply thankful to Prof. Dr. -Ing. habil. Alois Knoll for his interest in my topic, for providing me the freedom to exploit different fields and yet keeping my research on track. Prof. Dr. Christopher Zegras has allocated a considerable amount of time and effort in order to be my second examiner and I am extremely thankful for our discussions and his valuable comments, which helped me to further improve my dissertation. Prof. Dr. Alexander Pretschner has served as a chairman at my defence and insured this very important event of my life proceeds smoothly for which I am very thankful.

I would also like to thank Dr. Heiko Aydt and Prof. Dr. Mike Lees for their guidance and insightful discussions during my research phase. Furthermore, I would like to thank Dr. Vaisagh Viswanathan for his immense efforts to make my research understandable and coherent. The feedback obtained from our countless discussions has helped me tremendously in shaping this dissertation. I would like to express my deepest gratitude to my colleague Daniel Zehe for enabling the utilization of all computational power used for this research, without which any of this would not be possible. I would also like to thank my friend Dr. -Ing. Vladimir Todorov for his constant support and conventional wisdom.

Additionally, I would like to extend my gratitude to all my colleagues at TUM Create, which were directly or indirectly part of my research.

Finally, I would like to thank my father Stanislav and my late mother *Vessela* for always supporting me and all my close friends for their help and understanding.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | Motivation | 1 |
| 1.2 | Problem Statement | 2 |
| 1.3 | Sensitivity of Traffic Conditions to Traffic Components | 3 |
| 1.3.1 | Sensitivity to Infrastructure Changes | 3 |
| 1.3.2 | Sensitivity to Traffic Demand Changes | 4 |
| 1.3.3 | Sensitivity to Routing Changes | 4 |
| 1.3.4 | Sensitivity Summary | 5 |
| 1.4 | Systematic Optimization Approach | 5 |
| 1.4.1 | Analysis of Traffic Data | 6 |
| 1.4.2 | Planning of Infrastructural Alterations | 8 |
| 1.4.3 | Routing Control | 9 |
| 1.4.4 | Traffic System Surveillance | 10 |
| 1.5 | A Holistic Complex System Optimization Approach | 11 |
| 1.6 | Thesis Structure | 12 |
| 2 | Model and Simulation Approach | 13 |
| 2.1 | Data and Methods | 13 |
| 2.1.1 | Overview | 13 |
| 2.1.2 | Data Sets | 14 |
| 2.1.2.1 | HITS | 15 |
| 2.1.2.2 | QI Data | 15 |
| 2.1.3 | Model and Simulation | 16 |
| 2.1.3.1 | Agent Generation | 17 |
| 2.1.3.2 | Routing | 20 |

CONTENTS

| | | |
|----------|---|-----------|
| 2.1.3.3 | Traverse Time Calculation | 21 |
| 2.1.3.4 | Assumptions | 22 |
| 2.1.3.5 | Extraction of Free Flow Speeds | 23 |
| 2.1.3.6 | Calibration | 23 |
| 2.1.3.7 | Validation | 24 |
| 3 | Identification of Mismatched Sensitive Road Network Locations and Effect of Optimal Infrastructure Design on Traffic Performance | 29 |
| 3.1 | Overview | 29 |
| 3.2 | Introduction | 31 |
| 3.2.1 | Background and Motivation | 31 |
| 3.2.2 | Demand-Infrastructure Mismatch | 32 |
| 3.2.3 | Temporal Variation of Demand and Dynamic Intersections | 33 |
| 3.2.4 | Inter-system Comparisons | 36 |
| 3.2.5 | Information Synthesis | 36 |
| 3.3 | Literature Review | 37 |
| 3.4 | Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes | 40 |
| 3.4.1 | Calculation of Deviation and Mismatch Measures | 40 |
| 3.4.2 | Calculating the Dynamic Factor of a Node | 46 |
| 3.4.3 | Deviation Measure Calculation for Singapore Case Study: | 49 |
| 3.4.4 | Dynamic Factor Calculation for Singapore Case Study | 53 |
| 3.5 | Application of Recommended Changes in Infrastructure | 57 |
| 3.5.1 | Methods | 57 |
| 3.5.1.1 | Identification of Locations to be Fixed | 57 |
| 3.5.1.2 | Calculation of Optimal Lane Distribution and Road Network Changes at Chosen Locations | 58 |
| 3.5.1.3 | Calculation of Effects of Mismatch Fix | 59 |
| 3.5.2 | Results | 60 |
| 3.5.2.1 | Improvement of Traffic in the Time Domain | 61 |
| 3.5.2.2 | Measures of Sensitivity | 62 |
| 3.5.2.3 | Spatial Distribution of Saved Time | 64 |
| 3.5.2.4 | Correlation Analysis of Measures | 64 |
| 3.6 | System Optimum Lane Distribution Problem | 67 |

| | | |
|----------|--|-----------|
| 3.6.1 | Results | 70 |
| 3.7 | Chapter Summary | 74 |
| 3.7.1 | Traffic Demand-Infrastructure Mismatch Analysis | 74 |
| 3.7.2 | Dynamic Locations | 75 |
| 3.7.3 | Fixing the Road Network | 76 |
| 3.7.4 | Optimal Lane Distribution Problem | 76 |
| 4 | Identification of Harmful Roads and Routing Control for Efficient System Optimum Traffic Assignment | 77 |
| 4.1 | Overview | 77 |
| 4.2 | Introduction and Existing Literature | 79 |
| 4.2.1 | Motivation | 79 |
| 4.2.2 | Price of Anarchy | 80 |
| 4.2.3 | User Equilibrium and System Optimum Traffic Assignment | 80 |
| 4.2.4 | Braess' Paradox | 81 |
| 4.2.5 | Towards a Centralized Routing Control System | 82 |
| 4.3 | Systematic Road Removal | 84 |
| 4.3.1 | Heuristic Design for Braess Paradox Detection | 88 |
| 4.3.2 | Invariance of Total Travel Time to Choice of Agents to be Re-routed | 90 |
| 4.3.3 | Spatial Distribution of Effects of Full Road Closures | 91 |
| 4.3.4 | Equilibria Analysis | 91 |
| 4.4 | Socially Optimal Routing | 94 |
| 4.4.1 | Existing System Optimum Computation Algorithms | 96 |
| 4.4.2 | Proposed Algorithm: Backwards Incremental System Optimum Search (BISOS) | 103 |
| 4.4.3 | Results | 107 |
| 4.4.3.1 | Quality of Solution Compared to SO Solution and Speed of Convergence | 107 |
| 4.4.3.2 | Performance for Various Step Sizes | 109 |
| 4.4.3.3 | Scaling of Algorithm with Population Size | 111 |
| 4.4.3.4 | Fuel Consumption Model and Evaluation | 112 |
| 4.4.3.5 | Participation Rate Effect on Average Travel Time | 116 |
| 4.5 | Chapter Summary | 117 |
| 4.5.1 | Sensitivity of Traffic Conditions to Road Removal | 117 |

CONTENTS

| | | |
|----------|---|------------|
| 4.5.2 | System Optimum Computation Using BISOS Algorithm | 118 |
| 5 | Robust Route Information Maximising Sensor Placement | 121 |
| 5.1 | Overview | 121 |
| 5.2 | Introduction | 122 |
| 5.2.1 | Motivation | 122 |
| 5.2.2 | Choice of Information Measure | 123 |
| 5.2.3 | Sensing Intersections vs. Road Segments | 123 |
| 5.2.4 | Robustness | 124 |
| 5.2.5 | Combination of Demand and Topology Information | 125 |
| 5.3 | Literature Review | 125 |
| 5.4 | Measuring Importance of Nodes | 129 |
| 5.5 | Achieving Robustness Against Changes in the OD Matrix | 135 |
| 5.5.1 | Methodology for Altering the OD matrix | 136 |
| 5.5.2 | Strategy for Robust Placement | 136 |
| 5.5.3 | Strategy for Finding the Optimal Number of Sensors | 138 |
| 5.6 | Optimal Sensor Placement for Singapore Case Study | 140 |
| 5.6.1 | Importance Analysis | 140 |
| 5.7 | Chapter Summary | 145 |
| 5.7.1 | Sensor Placement | 145 |
| 5.7.2 | Robustness | 145 |
| 6 | Conclusion | 147 |
| 6.1 | Main Messages | 147 |
| 6.2 | Theoretical Implications | 148 |
| 6.3 | Policy Implications | 149 |
| 6.4 | Limitation of the Study and Future Research | 150 |
| 6.5 | Final Remarks | 151 |
| A | Appendix | 153 |

List of Figures

| | | |
|------|--|----|
| 1.1 | Statistical temporal trends in extra fuel spent yearly due to congestion in billions of gallons and average fuel efficiency in miles per gallon. | 2 |
| 1.2 | Flow chart of interactions between ITS optimization modules (blue) and traffic determining components (red). | 6 |
| 2.1 | Distribution of starting times of trips according to HITS data. | 17 |
| 2.2 | Extraction of free flow velocities for different road categories. | 23 |
| 2.3 | Outcome of the calibration process. | 26 |
| 2.4 | Velocity profile according to QI data of the most congested links from simulation results. | 26 |
| 2.5 | Comparison of congestion maps from the simulation 2.5d and Google Maps typical traffic service. | 27 |
| 3.1 | Simplistic examples of a matched (top) and a mismatched (bottom) intersection. | 33 |
| 3.2 | Dynamic demand change example with temporal flows. | 35 |
| 3.3 | Visualisation of the difference between the deviation measure and the absolute mismatch measure. | 44 |
| 3.4 | Distributions of deviation and mismatch measures. | 50 |
| 3.5 | Temporal profile of deviation measure. | 52 |
| 3.6 | Spatial profile of mismatched intersections. | 54 |
| 3.7 | Histogram depicting the distribution of Dynamical Factor values in the Singapore road network. | 55 |
| 3.8 | Turning probabilities of the most dynamic node. | 55 |
| 3.9 | Spatial distribution of dynamic factor of intersections in Singapore. | 56 |
| 3.10 | Comparison of time lost at intersections before and after fixes. | 61 |

LIST OF FIGURES

| | | |
|------|--|-----|
| 3.11 | Distribution of saved time at intersections. | 62 |
| 3.12 | Distribution of saved time by percentiles. | 63 |
| 3.13 | Spatial distribution of saved time per lane. | 65 |
| 3.14 | Progress of the optimization algorithm as a function of the iteration count. | 70 |
| 3.15 | Spatial distribution of lane difference between original and optimal solution. | 72 |
| 3.16 | Distribution of the lane differences in various types of roads. | 73 |
| 3.17 | Distribution of the lane differences for roads with various lengths. | 73 |
| | | |
| 4.1 | Histogram of road closure effects | 87 |
| 4.2 | Map of changes occurring in traffic conditions due to closures. | 88 |
| 4.3 | Number of road segment removals to be evaluated in order to find the first n most harmful road segments. | 90 |
| 4.4 | Comparison of road closure effects. | 93 |
| 4.5 | Zoomed version of Fig. 4.4 in order to observe in more detail a sensitive area. | 94 |
| 4.6 | Convergence of BISOS. | 108 |
| 4.7 | Computation time - quality of solution trade-off. | 108 |
| 4.8 | BISOS performance for various step sizes. | 110 |
| 4.9 | Trade-off between number of computations and quality of SO solution for differ- ent values of the failed attempt limit value. | 111 |
| 4.10 | Effects on convergence of BISOS for increasing population size. | 113 |
| 4.11 | Effects on travel time for increasing population size. | 114 |
| 4.12 | Distribution of difference of saved fuel for each agent for system optimum routing traffic assignment. | 115 |
| 4.13 | Participation rate effect on average travel time for system optimum and user equilibrium traffic assignment. | 116 |
| | | |
| 5.1 | Example of suggested sensor positioning strategy. | 133 |
| 5.2 | Diagram example of calculating importance of nodes. | 134 |
| 5.3 | OD perturbation diagram. | 137 |
| 5.4 | Averaged importance value of nodes for a full day. | 141 |
| 5.5 | Effects of OD perturbation. | 141 |
| 5.6 | Performance of optimally robust sensor placement. | 142 |
| 5.7 | Utility function construction. | 144 |

List of Tables

| | | |
|-----|---|-----|
| 2.1 | Notation Table for Traverse Time Calculation | 22 |
| 2.2 | Comparison between simulation and real world data | 24 |
| 2.3 | Calibrated Parameters and their values | 25 |
| 3.1 | Notation used to derive deviation, mismatch and dynamic factor measures | 41 |
| 3.2 | Correlation coefficients between chosen sensitivity metrics, defined measures and traffic flow. | 66 |
| 4.1 | Variation coefficients for chosen links for different percentages of closure | 92 |
| 4.2 | Points of Equilibrium | 94 |
| 4.3 | BISOS comparison with convex combination method for different step sizes and threshold values. | 109 |
| 5.1 | Notation for importance of node measure derivation. | 131 |

Chapter 1

Introduction

1.1 Motivation

Statistics show road traffic has been rising at an increasing rate over the past few decades, leading to a progressive deterioration of traffic conditions in metropolitan areas worldwide. As reported in [1], the travel time and excess fuel wasted in congestion have significantly increased. The 2014 annual cost of congestion was estimated at \$160 billion in the United States. In fact, urban traffic caused 6.9 billion hours of delay and 3.1 billion gallons of extra fuel wasted. Advances in automotive fuel-saving technologies have not been enough to mitigate the rising costs of congestion [2] as shown in Fig. 1.1. This highlights the current state of inefficiency of transportation systems and the need for an effective optimization strategy.

On the other end of the efficiency spectrum are the innovative optimization algorithms and modelling strategies harvesting the rapid increase in available computational power. Therefore a potential strategy, which addresses the two main challenges, which transportation systems are facing: 1) efficient real-time traffic management in the short term; and 2) forward-looking development of the transportation infrastructure, in view of satisfying the time-volatility of traffic demand in the long term, can be tested and identified by a computational model of urban transportation systems.

Most broadly stated, the motivation behind this thesis and respectively its objective is to build on top of state of the art analysis and optimization methods from the field of computer science and apply them to road transportation systems. The resulting strategies can bring insight into the future of large cities and point out efficient economically-sustainable solutions, robust against volatility in long term alterations of traffic demand. Fortunately, a considerable

1. INTRODUCTION

amount of data was made available for this research coming from the city of Singapore. This allowed the testing of all suggested methodologies in a realistic environment enabled by the collected data, backed by modelling and simulation techniques.

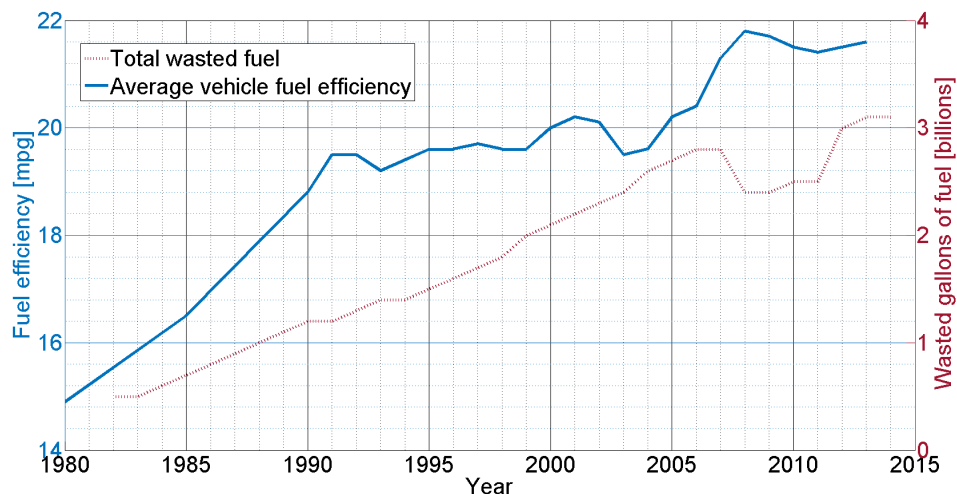


Figure 1.1: Statistical temporal trends in extra fuel spent yearly due to congestion in billions of gallons and average fuel efficiency in miles per gallon. Data used from [1] and [2].

1.2 Problem Statement

Congestion is perceived as the main problem of transportation systems. It occurs because of the heterogeneous nature of traffic demand in both time and space. Most commuters travel at the same time during the morning and evening, creating rush hours, which are defined by the working day temporal frame. Similarly, in space, there are central roads, which attract many drivers and become congested, while other roads remain underutilized. A related but qualitatively different reason for the onset of congestion is the slower pace at which the road infrastructure can be altered compared to the changes of commuting patterns. Therefore, road infrastructure is constantly lagging in its attempts to match the faster changing traffic demands and in turn, locations where demand and infrastructure are mismatched exist in the system. The latter leads to congestion.

There are two main tools, which can be used to tackle these problems. One is to redistribute traffic in a more homogeneous way in space, and the other is to change the road infrastructure for it to match more precisely the traffic demands and make those changes robust against

1.3 Sensitivity of Traffic Conditions to Traffic Components

possible future demand alterations. The problem to be resolved in this thesis is how to minimize congestion on a system level, in the sense of minimizing overall population travel time, in a robust and efficient manner, by using routing and/or infrastructure changes. Furthermore, the various approaches must be compared and the one most efficient and feasible for implementation should be identified.

1.3 Sensitivity of Traffic Conditions to Traffic Components

In order to model traffic, there are three main factors that should be taken into account, or respectively, three main questions that need to be answered. Traffic demand - from where, to where, and when do people want to go, traffic infrastructure (road network) - what medium is used to get from origin to destination, routing choices - how do people get from origin to destination provided the traffic network. This thesis examines how changes in these three components affect traffic conditions. The motivation for this study, except from pure scientific curiosity, is to identify the most beneficial component to be altered and controlled by the respective authorities in order to improve traffic conditions.

1.3.1 Sensitivity to Infrastructure Changes

The sensitivity of traffic conditions to alterations in the infrastructure is measured in two ways. First, in an identification of sensitive locations study the changes of total travel time, as a result of simple lane redistributions at single intersections, can be measured. For the morning rush hour period in Singapore, which is simulated, it is shown that the lane redistribution at the most sensitive location can save 500 hours from the total travel time of the population corresponding to 0.25% decrease. If all selected 500 sensitive intersections are fixed the total saved time on a daily basis reaches more than 4,500 hours (2.25%). Second, a systematic analysis for the identification of harmful roads further tests the sensitivity of traffic conditions to changes in the infrastructure. In this case, the small infrastructural change is the removal of a single road segment from the road network. The results demonstrate that in the extreme cases the removal of a single road can lead to up to 6,400 hours (3.2%) saved on a daily basis, or in the other extreme to 15,700 hours (7.85%) lost in the case of removing a vital for the network road. Changing the traffic infrastructure is a direct way for transportation officials to optimize traffic conditions, but it is limited by the relatively high costs of physical implementation.

1.3.2 Sensitivity to Traffic Demand Changes

The sensitivity of traffic conditions to changes in the commuting demands has been indirectly measured in a study that aims at finding optimal positions for sensor placement. The traffic demand is perturbed in order to create a robust placement solution. As a result of the perturbation approach that was undertaken the routing choices at intersections have been recalculated for different degrees of perturbation of the original demand. It was observed that the difference in the importance of the intersections, which is a combination of traffic volume and entropy of the routing choices varies very slightly with a coefficient of variation of 0.005 for a 5% perturbation. The changes of traffic conditions as a result of traffic demand alterations have not been examined in great detail in this work for two main reasons. First, they seem to be less significant than the changes induced by the other two traffic defining factors (road network and routing) based on this indirect measurement. Second, the demand does not depend completely on the free will of people, in the sense that decisions regarding housing and work locations are subject to external factors. In contrast to that, the decisions on how to get from origin to destination can be seen as a free will decision since there are no real constraints imposed on the drivers. It must also be noted, that the chosen mode of transportation can be influenced by officials by promoting public over private transit options, thus reducing the demand volume on the road network.

1.3.3 Sensitivity to Routing Changes

A significant contribution of this thesis is the finding that there is a high level of sensitivity of traffic conditions to changes in the routing of the population. The studies related to system optimal traffic assignment in this work demonstrate that in the case of Singapore a 70% decrease of population travel time can be achieved if drivers take the paths computed by the proposed algorithm. It is well-known that system optimal routing solutions in general increase the total travel distance by the population compared to the shortest paths initial traffic distribution. Therefore, it is desirable to find out whether the fuel consumption also increases as a result of the increase of overall travel distance. A fuel consumption calculation model is implemented and it is demonstrated that although the population covers more distance, the fuel consumption drops as a result of the decreased level of congestion in the city. To be more precise, in total the population uses 15% less fuel, when system optimal routing is applied.

It is important to note that algorithms for computing system optimum traffic assignment has existed for more than 30 years. Due to the fact that such traffic assignment has been

considered infeasible in practice, those algorithms are mostly used to determine a theoretical best utilization of the network rather than a real assignment that can be achieved. As a result of this, the constraints of those optimization problems are weaker and, in fact, lead to infeasible solutions that cannot be applied to real life systems. More precisely, there are no constraints that the flows along the roads should be integer values.

Although, such an additional constraint may seem easy to take into consideration, it significantly changes the performance of the already existing algorithms. Furthermore, the algorithm proposed in this work ensures that at all times of the computation every driver has a path from origin to destination that is known, while current state of the art either does not provide any information about the actual path of a driver or may split the driver between several possible paths. The approach of current state of the art algorithms is therefore considered impractical since if a centralized control system for routing control is designed it needs to supply a real route for each and every commuter. Furthermore, the algorithm suggested in this work converges about 15 times faster than previous ones and has smaller memory requirements.

1.3.4 Sensitivity Summary

In conclusion, the three components that can be subject to change can be addressed in 2 aspects: effect potential and ease of alteration. Road network is easy to change but costly and the results from a small change are not that significant. Traffic demand conditional on the origins and destinations of commuters can be considered almost impossible to change, or extremely challenging at least. The change of the way agents choose their routes is free of charge and does not restrict when from and where to people can go. It finds a more socially beneficial way for all of drivers to take. Furthermore, it brings the most significant decrease of overall travel time and fuel consumption by saving both time and money, while reducing the negative congestion induced effects on the environment. Therefore, routing is a promising approach to enabling more stable and sustainable growth for future transportation systems. For this reason, the work in this thesis has an emphasis on social optimal routing strategies as a way to achieve those goals.

1.4 Systematic Optimization Approach

The methods used to analyse the sensitivity of traffic to the three main components are combined with additionally developed methodologies forming a systematic optimization process for

1. INTRODUCTION

intelligent transportation systems consisting of four main steps. The first step of examining not just traffic but any functioning system is the analysis of already existing data extracted from it. After the analysis step, future changes in the system can be planned in order to create necessary conditions for more efficient operation. Followed or concurrently with the planning phase, an attempt may be made to control the system (if it is controllable) or steer it into a more beneficial and efficient state of operation. The fourth step determines the most important indicators of system performance and deploys efficient surveillance sub-systems to monitor the selected indicators efficiently. The data from the surveillance step is fed into a new analysis step and the system continues to improve over time. The process flow is depicted on Fig. 1.2.

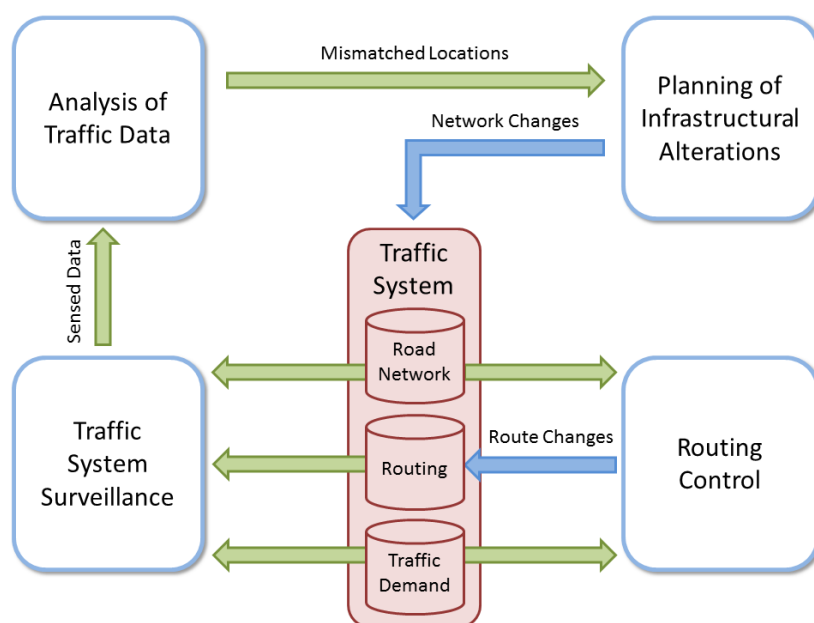


Figure 1.2: Flow chart of interactions between ITS optimization modules (blue) and traffic determining components (red). Green arrows represent flow of data, while blue arrows represent changes applied to traffic defining components generated by optimization modules.

1.4.1 Analysis of Traffic Data

The analysis step of this thesis consists of the identification of problematic regions in the topology of the road infrastructure and quantifying the inefficiencies associated with them. As the road network is a defining part of the whole transportation system, the approach of locating critical parts of it first seems reasonable as a starting point. The typically used measure for the importance, or criticality, of a road segment in literature is the number of vehicles that

are utilizing it. This work's contribution in this context is the examination of road network as an additional factor for the criticality of a road. Due to the fact that traffic infrastructure can change in a much slower pace than traffic demand, it may be expected that parts of the network no longer match the path patterns of commuters and therefore create unfavourable traffic conditions. The analysis step contributes to current state of the art analysis techniques with a measure definition for this mismatch and its computation for realistic traffic conditions in a large city scenario. The hypothesis of large amounts of discrepancy between traffic demand and infrastructure is confirmed, showing hundreds of locations where changes are advisable.

As a continuation of the analysis part based on matching the infrastructure to the demand, the optimal distribution of road widths within the city itself is computed. To the best of my knowledge, this is the first work, which defines and solves such an optimization problem on a large city scale. Although impractical, since the widths of all roads cannot be altered easily, the difference between the optimal solution and the existing infrastructure can be seen as a good measure of the overall mismatch of the city infrastructure with respect to the traffic demands. The aim of the optimization problem is to calculate the number of lanes on each road that minimize the overall travel time given a certain demand pattern. The only constraint in the problem is that the overall length of the road network in terms of lane meters must stay the same. In other words, road capacities can only be redistributed within the system. In a sense the solution to the optimization problem is what the network should look like if capacities had to be assigned on every road in the present moment from scratch in terms of number of lanes for each road.

As an additional step in the analysis phase and a prelude to the planning and control parts, the Braess paradox has been examined. The paradox [3] states that the addition of new roads to an existing infrastructure network may lead to an overall decrease of the performance of the system. Seen from a reverse perspective, the removal of roads from the existing network may lead to a performance increase. A systematic approach of measuring the impact of the removal of a single road has been developed in order to quantify the changes in travel time as a result of this type of network alteration. The Braess paradox has been intensely studied in previous works, however, the systematic approach presented in this thesis, which is on a full city level, is the first of its kind, and therefore a significant step in confirming the existence of the paradox in realistic conditions.

This analysis step can help planners by identifying road segments that, in practical terms, are harmful for the transportation system. Furthermore, the quantification of those problematic

1. INTRODUCTION

roads is also easily performed by examining the change of overall travel time of the commuting population. Naturally, an extensive search on the whole network, removing roads one by one is computationally extensive. It has been performed for the case study of Singapore and the results are presented in this work, however, a strategy for faster identification of those segments using a heuristic method presents interest for the academic community. An additional contribution of this work is a heuristic based on the difference in traffic flows between, naturally occurring traffic and the system optimum solution for traffic assignment. The systematic analysis of road closures results is utilized as a validation step for the defined heuristic.

1.4.2 Planning of Infrastructural Alterations

The identification of important, sensitive or critical places is not enough from the prospective of a traffic official, who might be able to fix only a few problematic areas by altering the road network. A prioritization procedure, needs to be suggested as well, thus directly providing quantified predicted outcomes of the suggested changes. Such a method is also presented in this work, where the identified problematic locations are altered according to the results from the analysis step. More precisely, the changes in the road network that are applied consist of optimizing the number of lanes at intersections in order for the turning options capacities to be in agreement with the traffic demands. Effects of such changes are easier to evaluate, since they do not alter the structure of the network in the sense of roads and intersections but only change their widths. This implies that the shortest path between two points do not change with the alterations. If it is assumed that the speed limits are not altered as well, it can be deduced that the fastest paths do not change either.

The metric that is widely used throughout the thesis for determining the impact of various changes in the traffic system is the saved travel time. There are two reasons for this. First, the majority of people consider time as one of their most valuable assets [4, 5] and second, the time spent in congestion is proportional to the extra fuel consumed, [6, 7] and therefore, transportation cost. By calculating the time saved from each change, the respective officials can easily pick the most beneficial locations to be addressed first. The novelty of this planning approach is that the impact of the changes are evaluated holistically in the context of the whole system, while typically in transportation literature, localized evaluations of various planning approaches are performed.

1.4.3 Routing Control

After the identification of harmful roads in section 1.4.1, either by a systematic search or with the help of a heuristic, it is clearly highlighted that the system in its original state is not functioning optimally. More precisely, it is demonstrated that commuters do not choose their routes in a way that minimizes overall system travel time. A road closure forces drivers to take other routes and in the case where a road closure is beneficial the only possible reason for that is that the routes before the road was removed were not socially optimal to begin with. The ratio between travel time computed with selfishly chosen routes and optimal routes is referred to as the Price of Anarchy [8].

Depending on the degree of non-linearity of the function that describes the relationship between the number of vehicles on a road and the level of congestion, the price of anarchy grows asymptotically as $\Phi(\frac{d}{\log d})$ [9]. In congested scenarios and non-linear traffic conditions this can easily mean doubling of the total travel time of the commuting population. It is, therefore, highly beneficial to have a centralized control system that is able to adequately distribute traffic and compute routes for drivers in order to minimize this negative effect.

Algorithms for system optimum routing computation exist for decades, however, they require great computational power and storage capacity, due to the high number of routings that need to be computed and the amount of viable paths that have to be stored. Furthermore, working on better algorithms for system optimum computation has been on the sidelines in the past since the actual realization of such a strategy in real life has been viewed as unreasonable. With the increased presence of technology in transportation networks in the sense of route guidance systems, computational efficiency and power, and autonomous vehicles, the actual implementation of such a system does not seem so far-fetched. It appears feasible that sooner rather than later, people will not be as involved in their routing choices as they are now. Therefore, the computation of system optimum routing solution has become an important problem with serious application potential. The requirements of fast and efficient computation, however, remain due to the highly dynamic nature of traffic. Efforts in improving the efficiency of such routing algorithms should be appreciated.

The contribution of the control part of this thesis is the design, implementation and analysis of such a system optimum computation algorithm, which reduces computational time and memory requirements that already existing algorithms have, while preserving the accuracy of the final routing solution that is reached. Realistic traffic scenarios can easily involve networks with hundreds of thousands of nodes and millions of drivers on them that need to be routed.

1. INTRODUCTION

Even small improvements of computational efficiency can therefore save large amounts of time and bridge the gap between theoretical approaches and real world applications. Furthermore, the positive effects of system optimum routing have been quantified for a realistic scenario of a full scale large city using the case study of Singapore. It is demonstrated how the saved total travel time percentage increases when the population of the city increases as well. It is shown that if the system optimum strategy is performed, the population of the city can increase twice, while keeping the same total population travel time.

1.4.4 Traffic System Surveillance

The final step of the process outline is traffic surveillance. As shown in the analysis, planning and control phases, it seems that routing has the highest potential to improve traffic conditions. Therefore, the object of the surveillance must be the routing choices commuters make. A method for sensor placement is designed that aims at minimizing the entropy (uncertainty) of the choices drivers make when they commute. This information theoretic approach to solving the sensor placement problem is able to generically identify the most important locations to be sensed, while avoiding the typical combinatorial problems that current state of the art methods exhibit. It must be noted that in order to maximize the information gain with regard to routing choices, static sensors are needed as opposed to participatory mobile sensing techniques.

The search for the set of optimal locations turns into an easily tackled problem, which basically consists of picking one by one the most important locations. The defined importance measure, based on entropy, possesses an intrinsic property that all locations are independent from each other. This property helps to avoid the problem of redundant sensor positions and thus allows immediate calculation of optimal sensor positions. Furthermore, the information that can be extracted from such positioned sensor network is fundamental in its nature since it captures traffic characteristics at the intersections rather than at the roads themselves. The completeness of the data collected at the examined intersections allows for the computation of all secondary traffic characteristics.

Traffic demand conditions change rapidly due to the fast pace at which new buildings and roads are constructed and the dynamic lifestyle of the commuter population involving frequent changes in working locations and housing. The deployed sensor placement cannot be moved once installed. Therefore, the optimal set of locations has to be robust against such type of traffic demand changes. A novel concept for transportation systems, which is similar to what

1.5 A Holistic Complex System Optimization Approach

is used for training data in machine learning, is presented in the thesis with direct application to robust optimal sensor placement problems.

Perturbations are introduced in the typical traffic demand consisting of commuters exchanging their housing locations, which aims at making the perturbations more realistic, since it simulates the process of people moving from one place to another. Following that, an optimal sensor placement solution is computed based not only on the current traffic demand but also on the perturbed one, so that the information gained is maximised. This procedure is analogous to training machine learning algorithms where noise is added to the data in order to make the trained entity more robust and increase its prediction or classification accuracy. The concept of robustness of traffic planning can be applied not just to sensor placement problems but in all city planning techniques, as it is a solid step towards solving the issue with the slower pace of traffic infrastructure changes compared to the traffic demands by allowing the infrastructure to change in a manner that is able to anticipate the future alterations of the traffic demand.

1.5 A Holistic Complex System Optimization Approach

As part of this introduction the author would like to address the way of tackling problems involving complex systems such as transportation. It might seem easier and practical to provide solutions and proofs of concept on a local scale, i.e. neighbourhoods, intersections, highways etc. It must be noted, however, that the changes at those locations although beneficial locally can be harmful for other parts of the system. Minimizing a non-convex function with several local minima can be used as an intuitive analogy on a higher level of abstraction of such cases. Considering only a portion of the parameter space and finding the corresponding minimum for this region, might seem optimal for the localized problem. Examining the whole domain, however, might show that the localized solution is orders of magnitude worse than the actual global minimum for the function. It is strongly believed that problems involving complex systems should be studied on a global level and that the whole system should be modelled and not be split in parts. The case study city Singapore that was chosen for this work is ideal for this type of approach since it is a large city allowing for emergence of complexity and most importantly an island city, which makes the whole system rather insulated and allows for its analysis. In all studies proposed in this work the suggested methods, techniques, analysis and optimization have been performed for the whole city thus allowing for a complete view of the effects and for higher degree of certainty for the efficiency of the described approaches.

1.6 Thesis Structure

The structure of this thesis is designed so that the reader can follow a complete story of how and why the research directions were chosen. In Chapter 2 the modelling and simulation approach is described in detail, including the SEMSim traffic assignment model from raw data, the routing strategies, travel time computations, calibration and validation of the model. Chapter 3 performs an analysis of the mismatch between traffic demand and traffic infrastructure and suggests heuristics for identifying super-sensitive locations and algorithms for computing optimal number of lanes for each road. Chapter 4 demonstrates the paradox of the existence of harmful roads in a city by performing a systematic search over the whole city of Singapore and identifying every harmful road and quantifying its degree of harmfulness. Following an interesting discovery about re-routing certain agents from congested roads, a novel system optimal routing algorithm is described that overcomes a significant portion of challenges current state of the art algorithms are facing. In Chapter 5, the problem of optimally positioning sensors in order to maximise the information about the routing choices of the commuters is solved with an information theoretic approach, which also guarantees robustness against long term traffic variations of the sensor placement solution. Chapter 6 offers a summary of the results, concluding remarks and future work directions.

Chapter 2

Model and Simulation Approach

This chapter aims at setting the groundwork for all the experiments and studies that will follow. It is based on the methodology parts of the author's contributions in [10, 11, 12, 13]. It will describe the three main components needed for traffic modelling and simulation, namely the road network, the traffic demand and the routing of agents (vehicles), and how they are estimated from the available data. The agent generation used in this work is part of the nano-scope traffic simulation platform SEMSim. Since hundreds of thousands of simulation runs need to be evaluated for the various experiments in the next chapters, a macrosimulation approach has been chosen for this thesis as it provides a fast computation of travel times from the already designated routes. Furthermore, most of the experiments concentrate on improving traffic conditions during rush hour conditions and therefore the model was calibrated with real data for such periods. Some experiments, however, make use of commuting patterns for a whole day period, which can also be modelled in the proposed method by splitting the day into smaller time segments and examining them separately.

2.1 Data and Methods

2.1.1 Overview

The road network used in this work is an unidirectional graph where nodes represent splitting or merging points at which drivers can take decisions. An intersection can be represented as a collection of nodes. Links represent road segments that connect two nodes, however, due to the nature of the acquired data there are some links with just one successor and one predecessor. In order for a vehicle to traverse between its origin and destination, a route needs to be calculated

2. MODEL AND SIMULATION APPROACH

based on the provided graph. The paths of all commuters are calculated using a preference-based routing approach. Every driver has a preference for path choice, based on speed, distance or comfort, with assigned calibrated probabilities.

The travel time of every commuter is the sum the traverse times of all links included in its route. Those delay times are calculated using a variation of the Bureau of Public Roads (BPR) function [14]. Realistic traffic is modelled by synthesizing a sufficiently large vehicle population based on Origin-Destination data that has been collected for the simulated city. Free flow velocities v_f , which are needed for the BPR function calculation, are extracted from collected GPS tracking data. The further needed parameters α^s and β^s for the BPR function are calibrated for different types of roads depending on their speed limits using both GPS tracking data and a travel time distribution of the population for the simulated time of day period.

The case study in this work examines the city of Singapore with population of 5.4 million people and around 1 million registered vehicles including taxis, delivery vans and public transportation vehicles [15]. It is an island city, which further simplifies the scenario since the examined system is relatively closed. Publicly available data has been used to acquire the unidirectional graph of Singapore, which comprises of 240,000 links and 160,000 nodes representing the road system of the city. The number of lanes, speed limit and length of every link is available allowing the extraction of information about its capacity.

For the purposes of this model two separate data sets have been used. The first one is the Household Interview Travel Survey (HITS) conducted in 2012 in the city of Singapore. It studies commuting habits of the population. Information about the origin-destination pairs, their temporal nature, and travel time distribution during rush hour periods is extracted from it. The second data set consists of GPS trajectories of a 20,000 vehicle fleet for the duration of one month, providing information about recorded velocities at various locations in the city during different times of the day.

2.1.2 Data Sets

The two used data sets are described in terms of the underlying challenges, taken assumptions and methods used to evaluate the degree of feasibility of the model, which is calibrated and validated using those data sets.

2.1.2.1 HITS

The first available data set is the Household Interview Travel Survey (HITS). It consists of a significantly large set of questions that aim at studying the travelling habits of Singapore’s population. The survey covers slightly more than 0.67% of the population, which amounts to 35715 participants. Each household representative has answered 108 questions about demographics, commuting preferences and motoring capabilities of his/her household. The questions of interest for this work are the ones, which deal with the commuters’ travel patterns. Every participant was asked to describe his/her trips for the whole day prior to the day the survey was taken. This information is described in the following format:

| Origin Postal Code | Destination Postal Code | Time of Start (hh:mm) | Duration (mins) | Means of Transportation |
|--------------------|-------------------------|-----------------------|-----------------|-------------------------|
|--------------------|-------------------------|-----------------------|-----------------|-------------------------|

The origin and destination locations are specified by a postal code. It should be noted that Singapore has a 6 digit postal code system, which allows for every building to have a unique postal code. Therefore the indicated origins and destinations can be pinned down with high precision. The column titled “means of transportation” can include various travel model such as private cars, taxi, public transportation, motorbike etc. Since the aim is to model the car population in Singapore and its dynamics, the entries that are of interest are the ones, which create traffic. In other words, the entries that put an extra vehicle on the road should only be examined. All surveyed people that use public transportation are excluded from the data set since public transportation runs regardless of the number of people that use it. Moreover, the entries of passengers in private cars are also excluded, in order not to count a vehicle multiple times. The trips that are left after the filtering process are used later for the agent generation step.

Furthermore, information about user estimated or recalled duration of the trips, which is also available, is used to create travel time distributions of the population, which is utilized in the calibration process. It is important to note that the survey was conducted on Singaporean residents from different age groups, ethnicities, professions etc. It is, therefore, safe to assume that the extracted data from the results is representative to an acceptable extent of the travel patterns in the city.

2.1.2.2 QI Data

The second data set that used in this work is a GPS trajectory data from a commercial fleet tracking system. The size of the fleet is roughly 20,000 vehicles. It comprises mainly of goods

2. MODEL AND SIMULATION APPROACH

vans, trucks and small lorries, however there is also a small portion of data included from car leasing companies and personal trackers installed on private vehicles. The information about trip duration, origins and destinations, therefore, cannot be used to extract travel patterns reliably, since it is not representative for the commuting population but rather for the servicing sector. It can be used, however, to estimate average speeds on the roads with a good coverage of the whole network since, after all, the vehicles are sampling points of traffic conditions at the locations they visit. The available data is for the duration of two months in 2014. Each entry has the following format:

| Track id | Latitude | Longitude | Heading | Ground speed | Time stamp |
|----------|----------|-----------|---------|--------------|------------|
|----------|----------|-----------|---------|--------------|------------|

The time difference between two consecutive sent signals (sampling period) from the same vehicle can vary between 1 second and 30 minutes. Typically there are more data points when a vehicle is turning and less sent messages (lower sampling rate) when the vehicle is going in a straight line. Vehicles usually have lower speeds at turns and move faster when going in a straight line. Therefore, there possibly exists a slight bias towards lower driving speeds being recorded. In case of congestion, however, the vehicles typically go very slow in a straight line. As a result of this, samples from congested roads may be smaller in number than expected and therefore congestion may be underestimated when looking at such a data set. Since the trackers send information 24 hours a day, there are time periods, throughout which the vehicles are parked but still send out data. In order to exclude those samples all data points, where there is no change in the position and velocity in the last 15 samples, are removed. The size of the final working data set is around 120 million points. A map matching algorithm [16] is used on every trajectory in order to assign every sample point to an actual link in the routing graph of Singapore. All samples are then grouped according to links and time stamp in order to get a picture of the velocity profile of the city throughout the day.

2.1.3 Model and Simulation

This next part of the methods description deal with the way agents are generated, the traffic assignment of routes modelling, the design of the time delay function and the calibration and validation of the parameters so that the generated traffic is as realistic as the provided data allows. This section describes closely the traffic generation procedure in the nano-scopic simulation SEMSim.

2.1.3.1 Agent Generation

Most of the studies in this thesis aim at simulating traffic conditions during rush hour. Between the morning and evening typical commute times during weekdays, the morning traffic peak is preferred for analysis since the commuting intensity is more concentrated compared to evening rush hour and the region of time during which it appears is more distinct as in the evening people typically leave their working spaces at varying times and furthermore their destinations might also vary depending on external conditions. The simulated fragment of the day needs to be large enough in order to have enough data points from the HITS data to generate agents realistically, however, the traffic conditions during the time period have to be as homogeneous as possible in order for the assumptions of the traffic model to hold. As a result of these considerations a one hour period from 7 to 8 is chosen, which is centred around the time when most agents are starting their trips, which is around 7 : 30. Fig. 2.1 represents the distribution of trip starts throughout the day. The selected period from the morning rush hour presents both the peak of trips starts and a homogeneous trip generation rate, which is a predisposition for static traffic conditions [17].

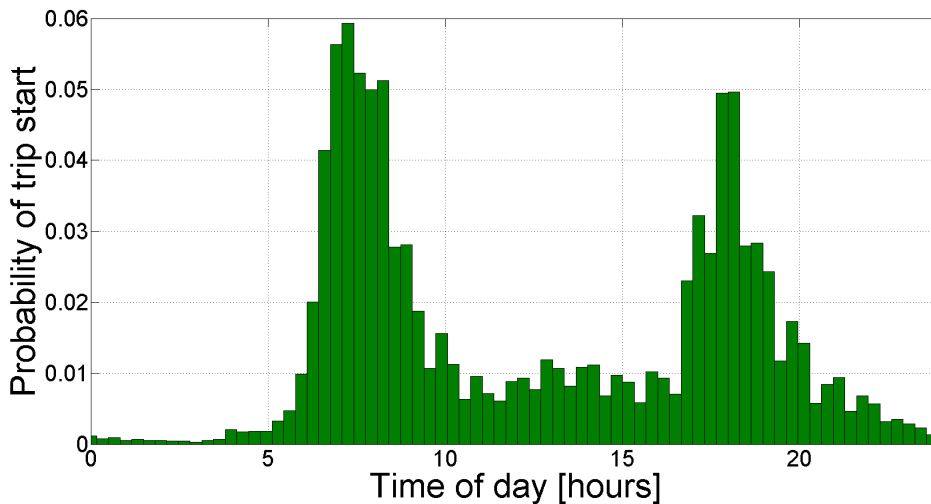


Figure 2.1: Distribution of starting times of trips according to HITS data.

The total number of agents that need to be generated in order to have a quantitatively good representation of the traffic situation needs to be estimated as well. Assuming that HITS data is representative of the portion of the population that uses taxis or personal vehicles to commute it can be stated that,

2. MODEL AND SIMULATION APPROACH

$$\text{Car Users} = \text{Population} \times \frac{\text{Car Users in HITS}}{\text{Surveyed People in HITS}} \quad (2.1)$$

This provides a way to estimate the number of people in Singapore that commute using a personal car or taxi (actively creating traffic) by knowing the percentage of people in HITS data that do so. The number of people who use cars and the total number of people in the HITS data are extracted from the itineraries by examining only trips with start time in the period 7 – 8 a.m. This gives us roughly 309,000 vehicles¹ that should be generated during the examined time period.

After the total number of agents to be generated is computed, a way to assign each of them an origin and destination should be designed. The description below describes the approach used in SEMSim for OD assignment at the time this work has been conducted.

As described in Section 2.1.2.1 a list of trips that actively create traffic on the roads has been extracted from the HITS data set. Using this list a distribution of postal codes being chosen as origins or destinations respectively is created, according to their intensities in the actual itineraries. A Bayesian estimation approach is used with a prior uniform distribution assumption. The process is as follows:

All existing postal codes in Singapore get an initial count of 1 and for every postal code that appears in the filtered trip data throughout the time period of interest the counter is incremented by 1. Using those counters distributions for origin and destination postal codes are constructed such that the probability of a postal code being chosen as an origin or destination is proportional to its counter value. A full description of the traffic generation procedure is formalized in Algorithm 1

This OD matrix construction approach is chosen in order to homogenize the origins and destinations of population extracted from the HITS data set in order to represent reasonably well the traffic demands of the city. As a result of this the origins and destinations of all agents that need to be generated are determined. Next, the routes of all agents from their origins to

¹This number does not take into consideration the expansion factors provided in the survey.

destinations should be computed also included in Algorithm 1.

Data:

| | |
|--------------------|---|
| G^i | Road network graph with weights according to preference i |
| PC | Set of all postal codes (PC) in the city |
| $Itineraries$ | Set of all trip itineraries extracted from HITS |
| $Agents$ | Set of all agents to be generated |
| $SampleRegion$ | Itineraries \rightarrow $\langle Originregion, Destinationregion \rangle$ |
| $SamplePostalCode$ | Region distribution \rightarrow Postal code |
| $SamplePreference$ | Preference probabilities \rightarrow Preference |
| $ComputeRoute$ | Origin PC \times Destination PC \times Graph \rightarrow Route |
| BPR | Link \times Flow \rightarrow Traverse time |

Result: Set of traverse times along every link $s - t_s$

Step 0: Set all postal origin and destination code counters to 1

```

foreach  $c \in PC$  do
    |  $c^o \leftarrow 1$  // set counter for PC  $c$  as origin to 1
    |  $c^d \leftarrow 1$  // set counter for PC  $c$  as destination to 1
end
    
```

Step 1: Increase counter according to the intensity in the trip itinerary from HITS

```

foreach  $i \in Itineraries$  do
    |  $i^{oo} \leftarrow i^{oo} + 1$  // Increment the counter for PC the origin of trip  $i$  as origin
    |  $i^{dd} \leftarrow i^{dd} + 1$  // Increment the counter for PC the destination of trip  $i$  as destination
end
    
```

Step 2: Construct Distribution.

```

foreach  $l \in PC$  do
    |  $p_l^o \leftarrow \frac{l^o}{\sum_{k \in PC^{io}} k^o}$  // Calculate probability of code  $l$  picked as origin in its region
    |  $D_{l^o}^o \leftarrow D_{l^o}^o \cup p_l^o$  // Add the probability of code  $l$  to the distribution of the region
    |  $p_l^d \leftarrow \frac{l^d}{\sum_{k \in PC^{id}} k^d}$  // Calculate probability of code  $l$  picked as destination in its
    | region
    |  $D_{l^d}^d \leftarrow D_{l^d}^d \cup p_l^d$  // Add the probability of code  $l$  to the distribution of the region
end
    
```

Step 3: Assign origins and destination to agents and compute routes

```

foreach  $a \in Agents$  do
    |  $\langle R^{ao}, R^{ad} \rangle \leftarrow SampleRegion(Itineraries)$  // Sample origin and destination region
    | pair from itineraries
    |  $a_o \leftarrow SamplePostalCode(D_{R^{ao}}^o)$  // Sample a PC from region
    |  $a_d \leftarrow SamplePostalCode(D_{R^{ad}}^d)$  // Sample a PC from region
    |  $P \leftarrow SamplePreference(p_t, p_d, p_c)$  // Assign a routing preference
    |  $a_r \leftarrow ComputeRoute(a_o, a_d, G^P)$  // Compute route of agent according to preference
end
    
```

Step 4: Calculate the flows along every link and traverse times.

```

foreach  $a \in Agents$  do
    |  $route \leftarrow a_r$  // Get route of agent  $a$ 
    | foreach  $s \in route$  do
    | |  $f_s \leftarrow f_s + 1$  // Increment flow of link  $s$ 
    | end
end
foreach  $s \in G$  do
    |  $t_s \leftarrow BPR(s, f_s)$  // Compute traverse time of link  $s$ 
end
    
```

Algorithm 1: Traffic Generation

2. MODEL AND SIMULATION APPROACH

2.1.3.2 Routing

Since the aim of the modelling step is to represent reality as much as possible the routing of the generated agents is preference-based. Some people prefer the shortest path, some the fastest and some prefer comfort rather than speed or time. Furthermore, drivers seem to lack a significant degree of adaptation to changes in traffic conditions.

The findings in [18] demonstrate that the routes, which drivers choose are predominantly static. The study examines data from a considerable amount of GPS trackers through a period of 18 months. The results show that drivers have strongly preferred routes that stay static, in the sense that they do not adapt them to traffic conditions changes in time. The routes themselves have been compared with the optimal (informed of traffic) routes and in more than half of the cases they do not coincide. Furthermore, in 34% of the cases the chosen routine routes are described as not even comparable to the optimal routes. This points out that people do not always minimize time but possibly other factors such as distance and comfort as modelled in our preference based routing. Furthermore, in [19] it was found that distance is the attribute most likely to be minimized, indicating good spatial perception, while the same cannot be said of journey times. This further highlights the need for a preference based routing approach.

In addition to that in [20] the authors found that one-third of the respondents, whose routing habits are studied deviate more than 10% from minimum travel time routes. Approximately 60% of surveyed drivers stated that they would use the same route during peak and off-peak conditions, which demonstrates the lack of adaptability to traffic information. In [21] it is shown that although 50% of driver population listen to live traffic reports 70% of them do not change their routes. Those results are in sound agreement with the findings in [22] where it was found that the path chosen on a trip was quite sensitive to the location of the origin and destination and that the chosen path most often differed considerably from the shortest time path across the network. Paths for trips made by the same driver were reported to be very consistent over time.

There are two main messages that can be taken from the collection of results from data collected describing routing behaviour. First, in a significantly large portion of the cases, drivers do not minimize their travelling time. Second, drivers will stick to their favoured route regardless of the traffic situation. For those reasons the preference based static routing was chosen over the standard user equilibrium traffic assignment. Furthermore the authors believe the assumption that commuters have full knowledge of the traffic conditions, which is fundamental to the user equilibrium state, would not produce realistic results. The traffic conditions simulated using the

suggested preference based routing have been calibrated and validated with recorded data and show to be in agreement with reality. User equilibrium computation takes significantly more time than preference based routing. Provided that hundreds of thousand of such assignments have to be performed, at this point of time it is not feasible to utilize user equilibrium traffic assignment.

We therefore there are 3 distinct ways to calculate the routes. The various routing types are realized by calculating the weights on the routing graph according to the respective preferences. After that, a shortest path algorithm that minimizes the sum of the weights for a path between origin and destination is used. The three types of defined weights are:

- $w_d = \text{road length}$ - minimizing distance
- $w_t = \frac{\text{road length}}{\text{road speed}}$ - minimizing time
- $w_c = \frac{\text{road length}}{\text{road speed} \times \text{number of lanes}}$ - maximizing comfort

After the generation of every agent one of the three preferences is chosen at random with probabilities p_d , p_t and p_c respectively. The probability values are calibrated since the preferences of routing choices vary depending on the city of choice. When the type of preference is chosen the corresponding route is calculated.

For each experiment run in this thesis a high performance cluster node was used. Each simulation ran on 32 threads on two Intel Xeon E5 (@ 2.60GHz) CPUs. The entire system has 192 GB of memory. A bi-directional Dijkstra implementation from the SEMSim traffic nano-scopic traffic simulation is used for route computation[23]. Since routing requests can be paralleled for each trip, the performance of the simulation benefited from the large number of threads. As 3 different metrics for weight calculation are used (distance, travel time and comfort), each thread has to load all three routing graphs in order to ensure maximum performance.

2.1.3.3 Traverse Time Calculation

Some notation must be defined in order to proceed to describing the calculation of traverse times. After calculating the routes of all agents, the number of vehicles that must pass through every road segment can be extracted. The time needed to traverse a link t_i for the link i is calculated using an extended version of the Bureau of Public Roads (BPR) function:

2. MODEL AND SIMULATION APPROACH

| Variable | Description |
|------------|---|
| t_i | time it takes to traverse road segment i [s] |
| l_i | length of road segment i [m] |
| v_f^s | free flow velocity on segment with speed limit s [m/s] |
| F_i | flow on segment i |
| w_i | number of lanes on road i |
| t | simulation time from which the flow is calculated [hours] |
| v_{min} | minimum flow velocity at link i at extreme congestion levels [m/s] |
| $I(i)$ | function that checks if there is an intersection at the end of link i |
| d^s | intersection added delay for roads with speed limit s [s] |
| α_s | parameter from the BPR function for roads with speed limit s |
| β_s | parameter from the BPR function for roads with speed limit s |
| $S(i)$ | number of successors of road segment i |
| $P(i)$ | number of predecessors of road segment i |

Table 2.1: Notation Table for Traverse Time Calculation

$$t_i = \min \left(\frac{l_i}{v_f^s} \left(1 + \alpha^s \left(\frac{F_i}{2000w_i t} \right)^{\beta^s} \right), \frac{l_i}{v_{min}} \right) + I(i)d^s \quad (2.2)$$

and

$$I(i) = \begin{cases} 1 & \text{if } S(i) + P(i) > 2 \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

2.1.3.4 Assumptions

There are four assumptions that are made regarding the traffic model:

1. Agents are rational in the sense that they would choose the best possible route with respect to their preferences.
2. There is no re-routing in the model.
3. In order to model the traverse time on every link using the capacity and the estimated flow, traffic is assumed to be homogeneous. This may lead to a reduction of congestion levels since homogeneous traffic flows are a best case scenario.
4. The minimum possible velocity at a link in extreme congestion is set to $5km/h$. The BPR function that is used to estimate traverse times is known not to represent realistically

extremely congested situations as the traverse time exponentially goes to infinity when the flow gets bigger. This is why a minimum possible speed is set for all links, which means that in all cases agents keep moving forward with an average velocity of at least 5km/h .

Please note that some of the studies in this thesis use the classical BPR function without the minimum velocity assumption and the added intersection timings for simplicity.

2.1.3.5 Extraction of Free Flow Speeds

All links are split according to their class into 3 categories with speed limits $s = [50, 70, 90]$ km/h. The free flow velocities for the three classes of roads are extracted from the QI data set, where the time variation of average velocities on all roads with the respective speed limits is calculated as shown on Fig.2.2. The maximum average velocity for each group of roads throughout the day is taken and set to be the free flow velocity v_f^s .

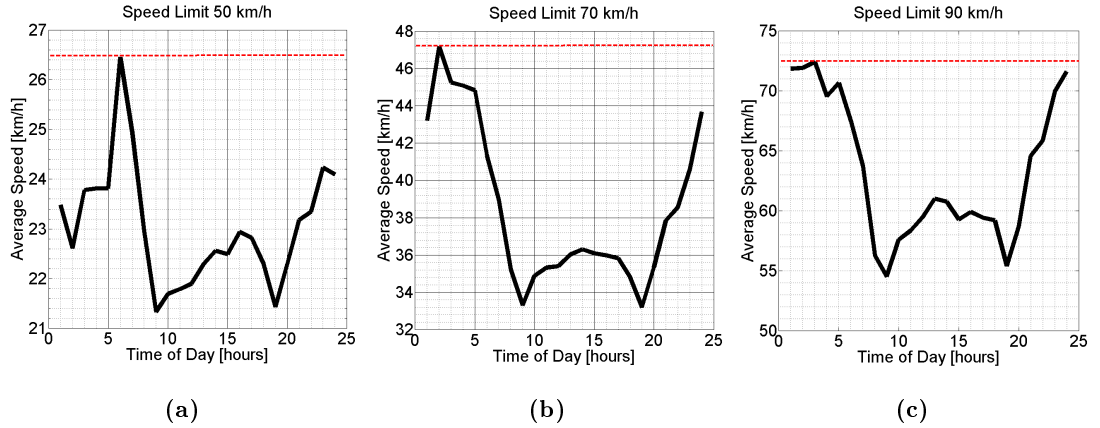


Figure 2.2: Extraction of free flow velocities for different road categories. Fig. 2.2a),b) and c) show average velocities throughout the day for roads with speed limit 50, 70 and 90 km/h respectively. The red dotted lines are used to mark the maximum velocity, which are considered to be also the free flow velocity for the respective type of roads.

2.1.3.6 Calibration

In order to calibrate the parameters of the simulation, real world data from the HITS and QI data sets is used. As a first step, the travel time distribution of commuters' trips who start their journeys within the time period of interest is constructed. The aim of the calibration process is to minimize the difference between this distribution and the one acquired from the

2. MODEL AND SIMULATION APPROACH

| Parameter | Value in Simulation | Value from Data |
|-----------------|---------------------|-----------------|
| μ_t | 2.99 | 2.90 |
| σ_t | 0.851 | 0.881 |
| v^{50} [km/h] | 22.4 | 22.8 |
| v^{70} [km/h] | 39.1 | 35.5 |
| v^{90} [km/h] | 64.3 | 59.3 |

Table 2.2: Comparison between simulation and real world data

simulation. Next, average velocities for all road classes are extracted from the QI data. Once again, the difference between those velocities and the ones calculated in the simulation should be minimized. This multi-objective optimization problem is solved using grid search.

The first parameters to be calibrated are the α and β parameters of the BPR function. Their values can vary widely depending on the road conditions and drivers' behaviour, which is why they have to be calibrated for a specific population and infrastructure profile. The values that have been acquired after the calibration step fall well into the range of accepted values in literature [24].

The next set of parameters that are calibrated are the preference probabilities. The values mentioned in [25] are used as a starting point. The final calibrated values show that each of the preferences is chosen with roughly an equal probability of one third.

The last set of parameters are the delays due to intersections for the three road classes. The calibrated values show that the most time on average is lost at major road intersections (usually due to traffic lights), while small roads and highways do not exhibit such large delays.

The calibrated parameters and their respective values are noted in Table 2.3. On Fig. 2.3 the comparison between real and simulation data is presented. The specific values are also shown in Table 2.2. It can be observed that there is a slight tendency for lower velocities in results from the QI data set. This, as already mentioned in the data set description, may be due to the sampling algorithm employed in order to collect the data and its tendency to exhibit a higher sampling rate when vehicles are turning and therefore have lower velocities.

2.1.3.7 Validation

In order to validate the results of the calibration step, the three most congested road segments according to the simulation are chosen. The velocity on those segments, which is calculated using the traverse function either reaches the critical preset minimum of 5 km/h or is very

| Parameter | Calibrated Value |
|---------------|------------------|
| α^{50} | 0.8 |
| β^{50} | 2 |
| α^{70} | 1 |
| β^{70} | 3 |
| α^{90} | 1.2 |
| β^{90} | 5 |
| p_d | 0.31 |
| p_t | 0.33 |
| p_c | 0.36 |
| d^{50} [s] | 1 |
| d^{70} [s] | 4 |
| d^{90} [s] | 1 |

Table 2.3: Calibrated Parameters and their values

close. All three examined roads have a speed limit of 90 km/h. All samples of vehicles that have traversed those roads between 7 and 8 am on weekdays are extracted from the QI data. Their velocity profile is shown on Fig. 2.4. It can be observed that in reality as well as according to the simulation results those road segments are experiencing heavy congestion and low average speeds.

It should be noted that in cases of severe congestion vehicles are mostly still, which results in a decrease in the number of samples for such periods of time. Therefore, it is possible that the velocity profiles of the examined road segments in reality might show even higher degree of congestion. The graph demonstrates that all three road segments, which are severely congested in the simulation seem to be congested in reality as well according to the GPS tracking data, which means that the traffic assignment strategy and traverse time function have provided appropriate approximations of reality for the desired level of detail.

Furthermore, in Fig.2.5 a congestion map produced by the simulation is presented and compared to typical traffic pictures from Google Maps for the desired period of time. Since this service presents traffic averaged over 10 minutes intervals, three different pictures are shown from the beginning, middle and the end of the examined period. The congestion map represents closely what is provided by the real data estimations of the Google Maps service, which further demonstrates that the applied simulation approach produces results which are not far from reality.

2. MODEL AND SIMULATION APPROACH

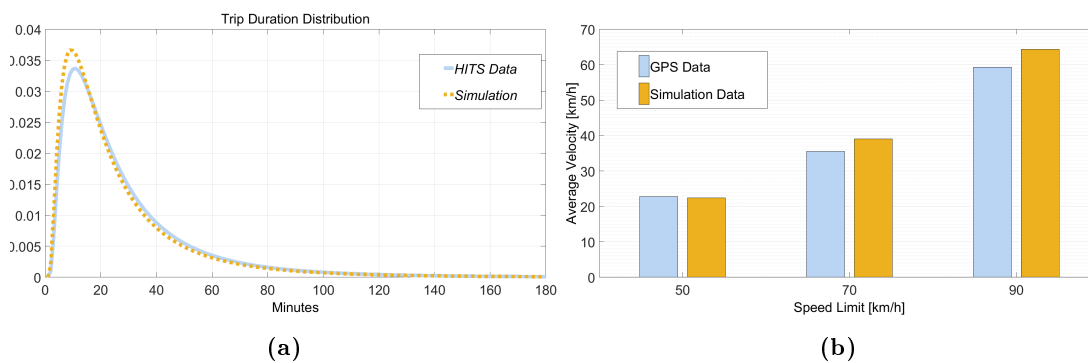


Figure 2.3: Outcome of the calibration process. On Fig.2.3a) a log normal distribution can be observed with the parameters extracted from the HITS data compared to a log normal distribution with the parameters extracted from the simulation. Fig. 2.3b) depicts a comparison of the average speeds of the three groups of roads extracted from QI data and from the simulation.

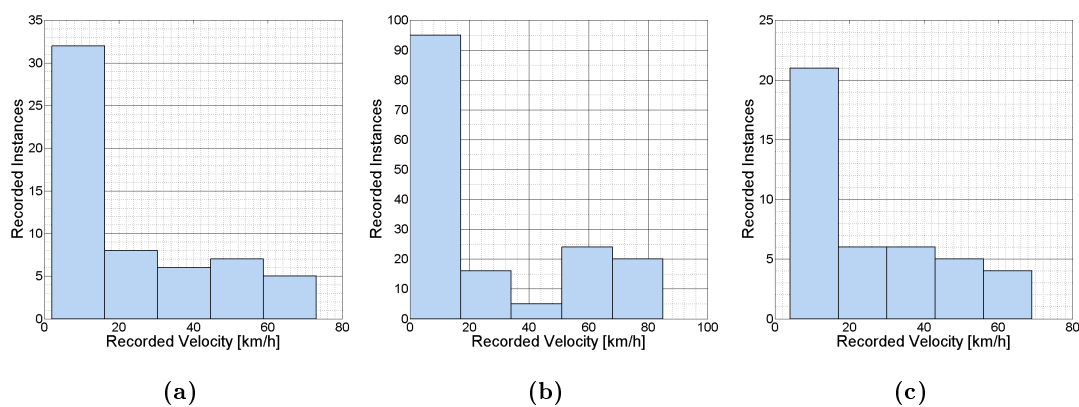
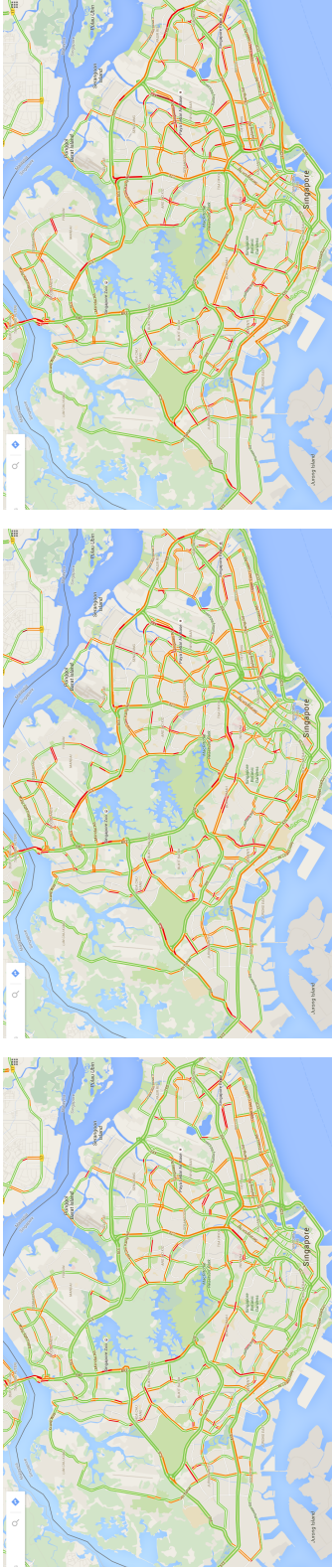


Figure 2.4: Velocity profile according to QI data of the most congested links from simulation results. The velocity samples are taken for the period between 7 and 8 a.m. for weekdays.



(c) 8:15 a.m.

(b) 7:40 a.m.

(a) 7 a.m.



(d) 7-8 a.m. from simulation

Figure 2.5: Comparison of congestion maps from the simulation 2.5d and Google Maps typical traffic service for 7 a.m. 2.5a, 7 : 40 a.m. 2.5b, and 8 : 15 a.m. 2.5c

2. MODEL AND SIMULATION APPROACH

Chapter 3

Identification of Mismatched Sensitive Road Network Locations and Effect of Optimal Infrastructure Design on Traffic Performance

3.1 Overview

This chapter studies the interactions between two of the three main traffic defining components, namely the traffic demand and road network. It presents analysis tools in the form of defined measures to detect and identify critical locations with high impact on traffic and sensitivity to changes. The first part of the chapter combines the author's contributions in [12, 11].

The mutually induced adaptations of traffic demand and infrastructure are present in the evolution dynamics of every large city. As a result of those processes, roads and intersections that, at the time of their construction have been steered by current traffic demands, can become mismatched when those demands change in time. This chapter will deal with the identification of locations that exhibit such suboptimal traffic conditions and the analysis of the effects of those mismatches on the overall commuting dynamics.

The generated traffic from the model described in Chapter 2 enables the calculation of turning probabilities on every node of the graph, which provide an overview of the differences between traffic demand characteristics and planned physical roads' capabilities.

A set of measures, which quantify the mismatch between demand and infrastructure is

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

defined. Different variations of the measures can be used to estimate deviations on single intersections and on a city level. Furthermore, normalised measures have been designed in order to enable the comparison between different cities. The objective of the defined metrics is the quantification of mismatches ensuring prioritization of possible changes that need to be applied to the road infrastructure so that it supplies adequate support for the population.

After a spatial and temporal picture of the demand-infrastructure mismatch is constructed, a redistribution of turning lanes on the most problematic intersections is performed in order to verify that the measures indeed locate high impact locations that are sensitive to changes. It must be noted that the redistribution consists of changing the number of lanes corresponding to different turning options at an examined node, which means that no extra lanes should be built. They should rather be shifted from one road segment successor to another. The analysis of the changes in population travel time as a result of the applied infrastructure changes indicates that a significant amount of congestion reduction can be achieved especially during rush hour periods. Furthermore, it has been observed that changes of the same magnitude at various locations can have significantly different implications on overall traffic conditions. High impact locations, labeled as *super-sensitive*, have been studied and a metric for their detection has been identified. Such locations are important for transportation systems, since they can be utilised as steering tools by traffic officials.

A more general optimization problem for determining the optimal lane distribution on the already existing road network has been formulated and solved. The main constraint of the problem states that the network topology should not be altered in the sense of construction of new roads or the removal of existing ones. Instead, the road capacities can be varied, in the sense of the number of lanes, under the constraint that the total number of lane meters in the network should be kept constant. The results show that the optimal lane distribution would reduce by 36% the total travel time for the population, when shortest time routes are chosen by the drivers. Although, impractical in the sense that such alterations to the whole road network cannot be realistically performed, the mathematical distance between the current road network and the one with optimal lane distribution, can be used as a global measure of the mismatch between road infrastructure and the traffic demand.

Fixing such type of mismatches can and will increase efficiency of the transportation system, however, the problem of the time-varying demand on a daily basis is harder to tackle. As the complexity of traffic conditions in large cities increases it becomes important and highly desirable to be able to analyse adequately the temporal nature of such systems. Due to the

high degree of heterogeneity of traffic demand and road network topology in space [26] and time (rush hour conditions), it becomes plausible that highly dynamical “hot spots” exist in a network, which may present a challenge for city planning and conventional intersection control methods.

An additional contribution of this work is the identification of intersections, which exhibit dynamically changing demand conditions in time and to quantify the level of volatility at those locations. The measure, which is defined using turning probabilities variation in time is called *dynamic factor* of an intersection. Highly dynamic intersections are identified and observed, thus demonstrating that such locations in fact exist and further contribute to the heterogeneous dynamic profile of the road network. The immediate application of this measure is that it points to intersections that require specialised control such as smart traffic lights or other intelligent transportation system techniques.

3.2 Introduction

3.2.1 Background and Motivation

Traffic demand and road network topology are two of the three main factors that frame to a large extent, the commuting picture in a city. They are strongly connected and a change in one of them will inevitably affect the other. First, consider the example where the road network is changed by building a new highway or extending an already existing road. The commuters will adapt to this change and make use of the newly built road while reducing the traffic on others. Naturally, entrepreneurs might also construct new business or industrial centres in proximity to the road, attracted by its high throughput thus shifting the traffic demand by adding an external attractor factor. In a similar fashion, although at a relatively slower pace, in case of an overly large demand in a certain area action might be taken to change the road network and satisfy those demands and consequently reduce congestion.

In order to formalise the description of the first factor, the traffic demand of the population, transportation engineers and officials usually use the Origin Destination (OD) Matrix. It is used to present information about the typical origins, destinations and trip starting times of the commuters in a structured way. This also allows for its mathematical integration in various transportation models. The OD matrix answers the questions from where, to where and when does the city population want to commute. There are two main distinct ways to estimate the real origin destination profile of the population. The first one is by surveying a representative

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

part of the population about their traffic habits and then scaling the results for the whole population similarly to what has been done with the HITS data set described in Chapter 2. The second methodology makes use of information about traffic flows acquired from sensors on various road segments. Given the traffic flows, there are techniques to estimate the OD pairs and their intensities as in [27].

The second traffic determining factor is the road infrastructure, which can be perceived as the medium enabling traffic demands to be met. The topology of the network is a result of the evolution of the city in time. Recently, it has become relatively easy to acquire a connected road network of almost every city. It is important to note, however, that some vital attributes of a network such as the number of lanes for every road segment are troublesome to obtain. As the reader might recall, the number of lanes on a road is used for the calculation of road capacity, which is part of the delay function calculation. The absence of easily accessible lane information, therefore might be the reason why until now there is no full scale analysis of overall network capacity performance in big cities described in literature.

3.2.2 Demand-Infrastructure Mismatch

As a result of the constant interaction between demand and infrastructure, the road network changes incrementally in time. Although such an incremental change is always intended in the direction of improving traffic conditions, it does not guarantee optimality. It is possible that some already constructed roads turn out not to meet the demands any more or it is even possible that those demands no longer exist. In other words, some roads may become obsolete as a result of better alternatives appearing with time and changing demands. This ever altering nature of network topology-traffic demand dynamics creates the need of a measure that can evaluate whether traffic demands correspond to the potential the network has and vice versa.

At this point the reader's attention should be brought to the importance of the perfect match between the road network capacity and the demand. Intuitively, when the capacity is lower than demand, the undesired situation of traffic congestion propagating through the whole network is reached. Counter-intuitively, however, in the case of much higher capacity than demand, congestion might increase as well, because of the higher willingness of commuters to travel due to improved traffic conditions [28]. This induces more journeys, which not only slow down overall traffic but also increase the harmful effects to the environment produced by fuel emissions.

Therefore, choosing just the right capacity of every road should be considered as the proper way to ensure the smooth operation of a transportation system. A highly desirable trait of such a system is that traffic is spread onto the network in sound agreement with the network capacity and its topology. As intersections are the places in a network where drivers make decisions, which define the traffic conditions, it might be beneficial to evaluate the deviation profile of the network. This is done by examining the deviations at the decision points as shown in Fig. 3.1 and studying their temporal and spatial distribution.

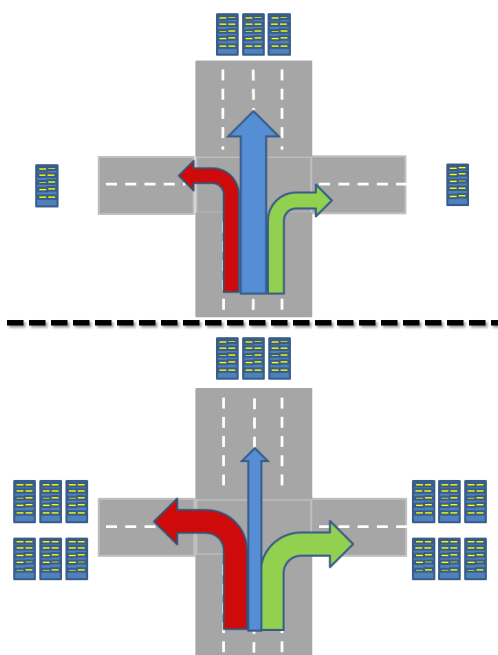


Figure 3.1: A simplistic example showing a properly performing intersection in agreement with the traffic demand on top and an intersection where the infrastructure is significantly deviating from the traffic demand as a result of construction of new business or housing areas. In order to fix the mismatch, the roads corresponding to left and right turns should be assigned more lanes and the road corresponding to the vehicles going straight should have a reduced number of lanes. For the sake of simplicity of the visualisation it is assumed that there is no other incoming traffic for the intersection.

3.2.3 Temporal Variation of Demand and Dynamic Intersections

It might sometimes be impossible to match the capacity of a road precisely to the traffic demands because of the temporal fluctuations of traffic conditions throughout the day and the static nature of the road network. Dynamically changing demands over parts of the traffic

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

infrastructure present challenges to city planners and force them to find various solutions for their optimal control by planning new infrastructure developments [29], control strategies [30], novel policies [31], etc.

Major city intersections are typically controlled by traffic lights. Highly congested systems might also benefit from the deployment of smart traffic lights [32], that adapt in real time to changes in the traffic system and strive to ensure smooth flow of traffic. Although, there are numerous discussed strategies for traffic light control, there is a fundamental element that is not present. How can the locations where those smart traffic lights should be installed be identified? The installation costs of such type of technology are not negligible and therefore it is not practical to simply deploy one at every intersection. Furthermore, some intersections exhibit static behaviour throughout the day and in such cases a static frequency controlled traffic light system is sufficient. There is, however, a need for a method that identifies the intersections that have the biggest demand for adaptive control due to their rapidly changing dynamics.

Although daily patterns do not seem to vary excessively, as observed in [33], morning and evening rush hour traffic patterns might exhibit significant qualitative differences. This is due to the fact that the OD pairs are reversed when people return home compared to when they go to work. The locations where traffic merges, therefore become locations where traffic is being split later on. In such cases a universal timing solution might just be inapplicable and even lead to further congestion as shown in Fig. 3.2. This problem does not occur only when the dynamics of the intersection have a bimodal nature due to morning and evening conditions. It might be the case that even higher degrees of variations and abrupt changes in the drivers demands at an intersection occur during the day as a result of the complexity of the system.

As it can be seen in Fig. 3.2, even if there is a successfully implemented dynamic control over the flow of vehicles, the varying volumes of cars taking different turns at intersections will require a changing capacity of the respective roads. The number of lanes and therefore the capacity of roads is, however, strictly fixed. Consequentially, the number of lanes must also be optimized according to the traffic demand. If the worst case scenario is always considered, roads will be planned with too many lanes and space will not be utilised optimally. Furthermore, the construction of such broad roads might not be possible at all times. If the average case is considered, also referred to as daily optimal, due to the extreme traffic variability at the intersection, at some point the flow will have a radically higher value than the capacity of the road, leading to an avalanche of congestions that will spread throughout the whole network.

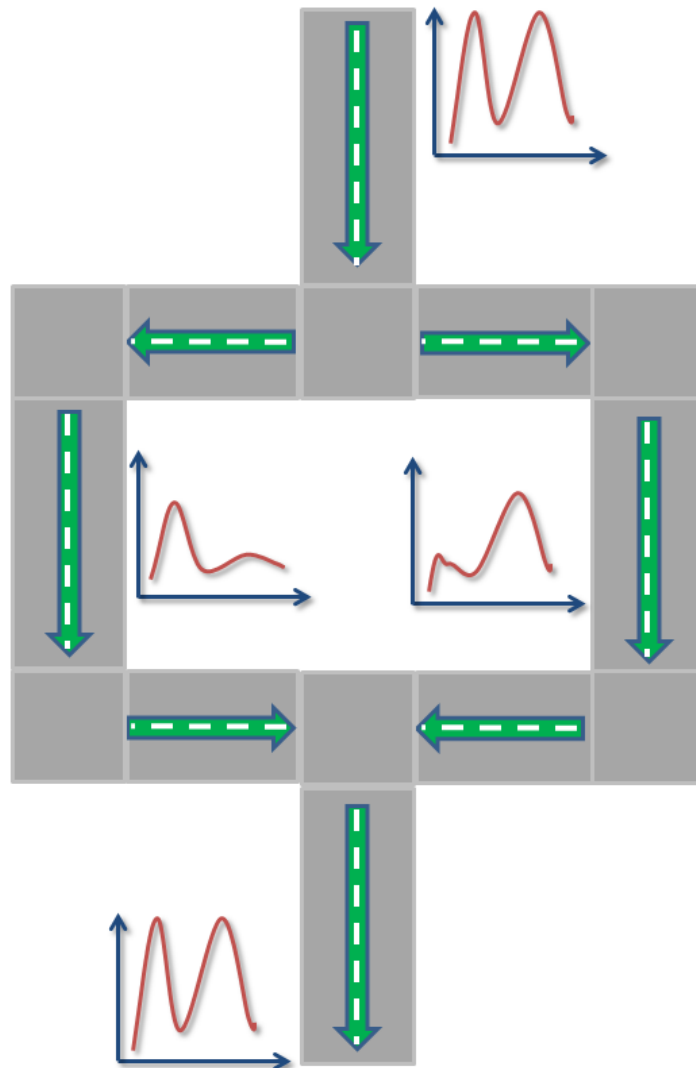


Figure 3.2: Diagram describing a simplistic example of change in demand depending on the time of day. Next to every road a flow vs time chart can be observed. In this example during the morning, most of the agents make a right turn, while in the evening most of the agents would make a left turn. This creates a dynamic intersection that experiences varying traffic demand throughout the day. As it can be seen after the roads merge again there is no variation. A static optimal control strategy would be to have equal amount of cars go to the left and to the right throughout the whole day. In this case congestions will occur for the cars making a right turn in the morning and a left turn in the evening.

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

In summary, highly dynamic intersection can be difficult to resolve with just traffic lights and proper road width planning and seem like a source of imminent problems. It is, therefore, desirable that such locations are to be avoided in the first place by careful road structure planning. The existence of such highly dynamic intersections implies an improper network utilisation. Other methods can be considered to fix such locations if they already exist by either using traffic lights at other intersections to redirect flows or construction of new roads in order to relax the variability.

3.2.4 Inter-system Comparisons

Dealing with complex social systems requires that the analysed entity is considered in its whole rather than examined only at specific locations. In the case of transportation systems, fixing or optimising one intersection with dynamic behaviour might produce another one at a distant intersection due to the high degree of interconnectedness that traffic network exhibit. Therefore evaluating the dynamic profile of all network intersections and comparing it to this of other networks can give precious insight into their dynamics and the severity of their problems. As already underlined, volatility and very dynamic and rapid changes in traffic demand at intersections makes their control challenging. It is, therefore, it is desirable to have less such intersection in a city in order to reduce congestion. In this line of thought, the level of such dynamics can be used as an optimisation heuristic that needs to be minimised so that the network utilisation efficiency is improved.

3.2.5 Information Synthesis

Existing traffic measures are governed mostly by demand and routing choice descriptors such as link flow, link average speed, etc. Measures connected to the topology of the infrastructure are centrality, heterogeneity, entropy. They are, however, defined in a purely topological sense [34]. It is strongly believed that both information contained in the OD matrix and network topology should be employed in order to come up with a more useful analysis measures of traffic networks.

The main contributions of this chapter are:

- Definition of a measure of intersection capacity deviation from demanded capacities extracted from turning option probabilities
- Definition of an overall network utilisation factor

- Case study with real world data for the city of Singapore identifying problematic intersections and providing spatial and temporal analysis of road network utilisation.
- Numerical study evaluating the effects of lane redistribution approach at nodes on population travel time.
- Correlation analysis for finding the best measure for identifying *super-sensitive* locations in the network.
- Optimization problem definition and solution for finding the optimal lane distribution for a whole network under constant road length constraints.
- Definition of the dynamic factor measure for a node in a traffic network used to identify dynamic intersections.
- Definition of the dynamic factor of a network in order to compare it to other networks.
- Analysis of heterogeneity of dynamic factor measures for a real world system.

3.3 Literature Review

The type of intersections, which attract most of the attention of researchers and city planners, are referred to as “critical” and are identified by their vehicle throughput. Naturally an intersection is critical if the flow of drivers through it is high as discussed in [35] and [36], where traffic management strategies for such locations are also discussed. A critical traffic volume has been defined in [37] in order to decide between deploying a traffic light or leaving an intersection un-signalised assuming all approaching traffic streams have the same prioritization. When there is a highly dynamical behaviour on a system level, which is typical for large cities, traffic conditions may further benefit from a control system based on self-organizing intersection control as described in [38].

Modelling such type of intersections has been studied for more than thirty years. In [39] a model based on queue dynamics as a function of demands and intersection characteristics is described and a real time control strategy is tested. Kirchhoff’s law for traffic flows at nodes are utilised to model intersection performance in [40]. Un-signalised intersections have also been modelled in [41] [42], where the traffic flows at non-signalised T-intersections in order are described using Petri nets. The study described in [43] performs a statistical analysis on critical intersections using data from various congested intersections in Shanghai during peak

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

hour. The study shows that the characteristics of intersections varies evidently from site to site. It must be pointed that most of the existing literature models single intersections, which as already discussed is not good practice when studying a complex system.

In order to be able to plan or simulate an intersections one may choose to evaluate the turning probabilities of the traffic participants. This has been done in [44], where a road density prediction method is described using the turning probabilities extracted from data. It should be noted that most methods found in literature are concerned with analysing and simulating intersections for rush hour conditions, while neglecting the temporal nature of traffic, the change that the flows might exhibit throughout the day and their effects. Most of the studies in this chapter also deal with rush hour conditions, however, the dynamic factor measure captures the temporal nature of traffic dynamics throughout a full day.

Modelling traffic volumes in a city, which indirectly can be used to determine flows at intersections, has been performed in [45] by using a Gaussian mixtures approach. A crucial network performance determining factor is how much the traffic demand on it varies in time. In [46] the variations of traffic on a daily and weekly scale are examined using cluster analysis techniques.

Road criticality can also be viewed as a combination of three factors: road flow and capacity, which looks at V/C ratio, path properties, examining estimation of path travel time, and network centrality, examining the percentage of OD pairs that use a certain road. Those components are encompassed in [47], where the authors demonstrate using traffic indicators that importance of road segments is mainly determined by the network structure and the flows. The majority of methods dealing with intersection analysis, however, focus mainly on the flows, and all approaches are solely defined by examining the traffic demands. An alternative or at least a complementary way of studying critical intersections in a network is to analyse it on a topological level. There are purely graphical measures defined in order to point out to “interesting” roads or intersections in a network.

In [48] critical links are identified using a network robustness index based on link flows, link capacity and network topology. In [49] the most vital links or nodes are defined as the first n links or nodes whose removal will lead to the biggest increase in average shortest path distance. While in [50] the importance of roads is simply defined to be proportional to the traffic load on them. In [51] three measures of centrality for a street are suggested: closeness, betweenness and straightness and their correlation to various economic activities in surrounding areas are examined.

Furthermore, the network itself can have some properties that are usually based on the system's structure rather than on local properties of its elements. In [52] the development of the Swiss road and railway network during the second half of the 20-th century is investigated. It is observed that the spatial structure of transportation networks is very specific, which makes it hard to analyse using general methods developed for complex networks. In [53] existing measures of heterogeneity, connectivity, accessibility, and interconnectivity are reviewed and three supplemental measures are suggested and described, including measures of entropy, connection patterns, and continuity. Entropy is used in order to determine the heterogeneity of the network regarding a chosen parameter.

Please note that, the topology of a network, itself, holds an enormous amount of information. Using it, insights into the structure of the roads can be gained. In [54] it is shown that transportation networks are organized hierarchically. In [55] the efficiency and accessibility in Paris and London based on the network connectedness is measured. Furthermore, topology information can be utilized in order to reconstruct agent's trajectories from GPS signals as in [56].

A family of graph measures based on entropy are summarized in [34]. It consists of measures from chemical structural analysis and social network analysis. The survey examines the overall connectedness of graphs such as the topological information content and the entropy of the weights of the edges. Furthermore, a measure of local features' such as entropy of nodes is defined as well, based on length of links connected to it. The centrality measure of links is also defined. Most measures deal with evaluating the information content in the graph itself. It is interesting to note that, those measures are highly uncorrelated, which means that they capture different aspects of graphs, so the proper measure for each specific application should be chosen with great care.

Road networks are subject to an incremental process of evolution. As societies change and cities grow the traffic demands and the road network itself changes with a high degree of self-organization and spontaneous organization of hierarchies occurs. This phenomenon observed in [57] leads to imperfections of the once planned infrastructure as the traffic demands have changed. Temporal variations in the relative importance of parts of the network have been further observed as well. In [58] the evolution over 200 years of a North Milan road network is followed. Two main processes are described in order to explain the collected observations. The first one is the densification of the road network around the main roads and the second is the emergence of new roads as a results of urbanisation. The interconnectedness between those two

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

processes is eminent and has been already mentioned in the introduction as the adaptation of the road network to the changes in the demand and vice versa. It can be, therefore, concluded that every location of the network which has been optimally planned for the time it has been constructed will probably become suboptimal as a result of changes in both the travel patterns and the network itself.

In conclusion, it is to be expected that there are mismatched locations naturally occurring in cities with dynamic traffic conditions due to either qualitatively different traffic throughout the day or the faster pace of evolution of traffic demand compared to the infrastructure. In order to identify such locations currently there are separate measures for criticality in terms of demand and topology, while a measure, which combines the two has not been developed.

3.4 Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes

In this section the measure of deviation between node capacity and traffic demand is introduced. A mathematical formulation of the deviation of a node and an overall deviation of a transportation network is defined and justified. Furthermore, an absolute measure of mismatch of an intersection that is derived from the number of lanes that need to be redistributed is defined. Along with the deviation and mismatch measures, the dynamic factor of a node is introduced.

Table 3.1 introduces the notation that will be used:

3.4.1 Calculation of Deviation and Mismatch Measures

The deviation of a node and the degree of mismatch between the network and the traffic demand are calculated in the following steps:

1. Calculate turning probabilities:

Let N_{ij} be the number of cars that pass through the i -th node and after that through the j -th node and let P_l be the path of agent l . Let the function f_{ij}^l be defined as:

$$f_{ij}(P_l) = \begin{cases} 1 & \text{if nodes } ij \text{ are in } P_l \\ 0 & \text{otherwise} \end{cases} \quad (3.1)$$

Then:

3.4 Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes

| Variable | Description |
|-----------------------|---|
| N_{ij} | number of vehicles that moves from node i to node j for the whole day |
| P_l | the path of the l -th agent |
| f_{ij}^l | function that is 1 if the sequence of nodes ij is in the path of agent l and 0 otherwise |
| A | a set containing all the agents |
| p_{ij}^t | probability that an agent that is at node i will continue on to node j during time period t |
| q_{ij}^t | turning probability that would be a perfect match for the infrastructure |
| S_i | set of nodes that are successors to node i |
| N_{ij}^t | number of cars that pass sequentially through node i and j during time period t |
| T | number of regions the day is split into |
| L | length of a time period |
| R | number of road segments in the network |
| w_{ik} | number of lanes on the road between nodes i and k |
| r_{ik} | ideal number of lanes between nodes i and k based on turning probabilities |
| m_i | absolute mismatch between number of lanes at an intersection i |
| Δ_{ij}^t | deviation measure of road from node i to node j during time period t |
| Δ_i^t | deviation measure of node i during time period t |
| Δ_{ij} | daily deviation measure of road from node i to node j |
| Δ_i | daily deviation measure of node i |
| $\hat{\Delta}_{ij}^t$ | corrected deviation measure of road from node i to node j during time period t |
| $\hat{\Delta}_i^t$ | corrected deviation measure of node i during time period t |
| $\hat{\Delta}_{ij}$ | corrected daily deviation measure of road from node i to node j |
| $\hat{\Delta}_i$ | corrected daily deviation measure of node i |
| $\hat{\Delta}_t$ | corrected overall deviation measure of whole network for time period t |
| $\hat{\Delta}$ | corrected overall deviation measure of whole network |
| C_{ij} | capacity of road from node i to road j per hour |
| G_{ij}^t | congestion factor for road from node i to node j for time period t |
| c_v | coefficient of variation |
| V_{ij} | variation of traffic on the road segment between nodes i and j |
| V_i | variation of traffic at node i |
| D_{ij} | dynamic factor on road segment between nodes i and j |
| D_i | dynamic factor at node i |
| C_i | capacity of a node |
| w_i | average number of lanes associated with a node |
| \hat{D}_{ij} | normalised dynamic factor on road segment between nodes i and j |
| \hat{D}_i | normalised dynamic factor at node i |
| M | dynamic factor of a city |

Table 3.1: Notation used to derive deviation, mismatch and dynamic factor measures

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

$$N_{ij} = \sum_{l=1}^{|A|} f_{ij}^l(P_l) \quad (3.2)$$

where $|A|$ is the number of agents.

Let p_{ij}^t be the probability that an agent at node i continues to node j during time period t . Let S_i be the set of nodes that are successors of node i . Then the turning probability is defined as the ratio between the number of drivers that pass through node i and then proceed to node j and the total number of vehicles that pass through node i :

$$p_{ij}^t = \frac{N_{ij}^t}{\sum_{k \in S_i} N_{ik}^t} \quad (3.3)$$

2. Calculate intersection demand deviation value:

The objective of this measure is to quantify the degree to which the road infrastructure at the intersection matches the traffic demand. The turning probabilities at every intersection for every period of the day have already been defined. The metric of the degree of discrepancy between the demand and actual roads is computed by comparing the ideal ratios between the roads' capacities computed from the turning probabilities and the physical number of lanes of the respective roads.

Let us first calculate what would be the best demand distribution such that the existing road width ratios optimally fulfil the traffic needs. The ideal turning probability q_{ij} from node i onto a certain successor road ij can be calculated by dividing the width (number of lanes) of the successor road by the total width of all possible successors as shown in Equation 3.5. It must be noted that, the optimal turning probabilities can only assume values of fractions of integers, since the number of lanes w_{ij} is an integer.

The next step is to take the difference between the real turning probability and the ideal turning probability with respect to the already existing road infrastructure as shown in Equation 3.6. Please note that, it is possible that the deviation value is non-zero even if the distribution of lanes on the successor roads is optimal due to the integer fractions that constitute the ideal turning probability, see Fig.3.3. Due to this fact, an absolute mismatch measure m_i is also defined in Equation 3.4, which states how many lanes must be redistributed within an intersection in order to achieve optimal performance. This measure computed by finding the difference between the actual number of lanes of all

3.4 Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes

successors w_{ij} and the ideal number r_{ik} , where the total number of lanes coming out of the intersection is kept constant.

$$m_i = \frac{\sum_{k \in S_i} \|w_{ik} - r_{ik}\|}{2} \quad (3.4)$$

Finally, in order to calculate the intersection demand deviation value Δ_i^t the average of all possible successors' deviations is taken as shown in Equation 3.7.

$$q_{ij}^t = \frac{w_{ij}}{\sum_{k \in S_i} w_{ik}} \quad (3.5)$$

$$\Delta_{ij}^t = \|p_{ij}^t - q_{ij}^t\| \quad (3.6)$$

$$\Delta_i^t = \frac{\sum_{k \in S_i} \Delta_{ik}^t}{\|S_i\|} \quad (3.7)$$

3. Weigh the deviation by the temporal flow profile of the node:

The time dependent deviation measures Δ_{ij}^t and Δ_i^t can be utilised to evaluate the degree of change of the turning probabilities throughout the day and consequentially optimal lane ratios. In order to make the analysis more complete, a deviation value that represents the whole day rather than just one time period will be defined as well.

An averaging technique that turns the time period observations into a representative metric for the whole day is, therefore, needed. Naturally, the values during some periods are more important than others since a mismatch would be more harmful during rush hours. Therefore, this performance is measured by the flow of vehicles through the respective road segment ij relative to the overall flow for the whole day. As shown in Equation 3.8 the deviation of the turn for period t is simply weighed by the ratio between the vehicles that have passed through it during this time period and the whole flow throughout the day. In a similar fashion as in Equation 3.7 the total daily deviation of the node is calculated by averaging the deviations of all possible turns as shown in Equation 3.9.

$$\Delta_{ij} = \left\langle \Delta_{ij}^t \frac{N_{ij}^t}{\sum_{k=1}^T N_{ij}^k} \right\rangle_t \quad (3.8)$$

$$\Delta_i = \frac{\sum_{k \in S_i} \Delta_{ik}}{\|S_i\|} \quad (3.9)$$

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

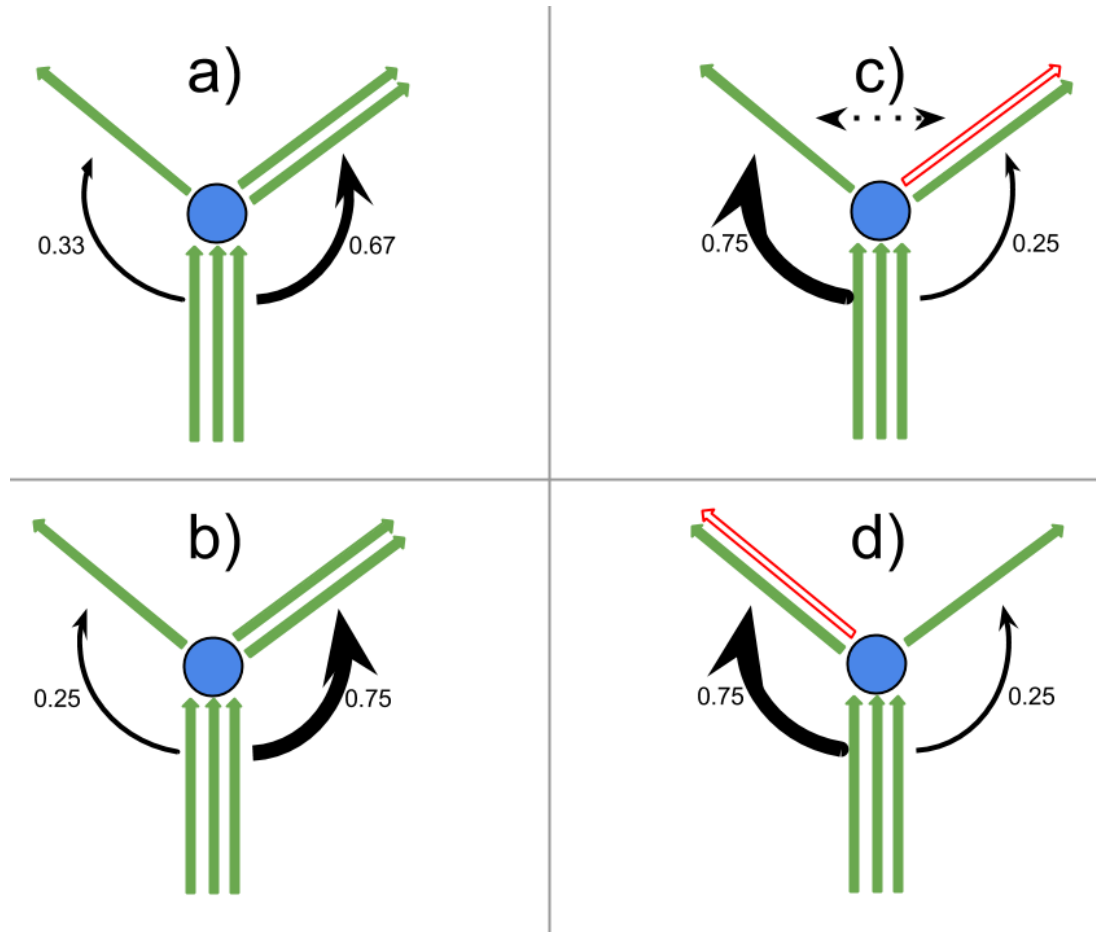


Figure 3.3: Visualisation of the difference between the deviation measure and the absolute mismatch measure. In a) a perfect agreement between the turning probabilities and the number of lanes is observed, since both ratios are 2 : 1. In b), however the ratio of the turning probabilities is 3 : 1, while the lane ratio is still 2 : 1. In this case the deviation measure Δ_i will be non-zero. The absolute mismatch measure m_i , however will still be 0 since given those turning probabilities and the total number of lanes to be distributed, which is 3, the optimal ratio of lanes is still 2 : 1. In c) the absolute mismatch measure is non-zero. The turning probabilities are in ratio 1 : 3, however, the lanes are in ratio 2 : 1. In this case one lane should be moved from the successor road on the right to the successor road on the left as shown in d). The absolute mismatch measure is therefore $m_i = 1$.

3.4 Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes

4. Extension for computation of overall deviation of the network infrastructure from traffic demand

In order to extend the presented methodology to compute a measure that represents the mismatch between the whole network infrastructure and the traffic demand, the already defined measures should be normalised in a way that enables comparison among distinct cities. The term that, in practice, needs normalisation is the number of vehicles that pass through every road segment ij . This is because it will obviously be of different magnitude in every examined city. In order to be able to compare one city to another a more global measure must be used. It must take into consideration the performance of a road rather than the absolute value of its throughput. The usage of the congestion factor measure that constitutes of the flow through the node over its capacity is proposed as shown in equation 3.11. The capacity of a segment ij , C_{ij} is defined as the number of lanes multiplied by a standard number of vehicles that maximize the throughput of the road segment per lane per hour and is usually set to 2000 [59].

$$C_{ij} = 2000Lw_{ij} \quad (3.10)$$

$$G_{ij}^t = \frac{N_{ij}^t}{C_{ij}} = \frac{N_{ij}^t}{2000Lw_{ij}} \quad (3.11)$$

Furthermore, in order to get a better picture of the overall road network deviation from the traffic demand, the individual intersection deviation values should be weighed by the already computed congestion factor. In this way, if an intersection has a high deviation value and is congested, it will receive a higher weight than an intersection that has the same degree of deviation but the traffic conditions on it are still on a satisfactory level. Also consider the example where an intersection has a very high flow of vehicles, but due to its high capacity, the congestion factor is still low and another intersection has a much lower flow, but is experiencing congestion. Assuming equivalent deviation measures, the second intersection will get a higher normalized deviation value. From one side, the deviation at the first intersection affects more people, however, from another side the effect is not as strong as it is for the congested intersection with smaller flow. The choice of measure to use should, therefore, be made with great care depending on the problem at hand. The congestion factor is included as a correction in all the already computed node deviations expressions as shown in Equations 3.15 to 3.17

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

$$t_{max} = \max_t G_{ij}^t \quad (3.12)$$

$$G_{ij} = G_{ij}^{t_{max}} \quad (3.13)$$

$$\hat{\Delta}_{ij}^t = \Delta_{ij}^t G_{ij}^t \quad (3.14)$$

$$\hat{\Delta}_{ij} = \Delta_{ij} G_{ij} \quad (3.15)$$

$$\hat{\Delta}_i^t = \frac{\sum_{k \in S_i} \hat{\Delta}_{ik}^t}{\|S_i\|} \quad (3.16)$$

$$\hat{\Delta}_i = \frac{\sum_{k \in S_i} \hat{\Delta}_{ik}}{\|S_i\|} \quad (3.17)$$

After the measures are normalised, the overall deviations of the network can be computed. The generalized mean with power factor of $\alpha = 2$ is used over all $\hat{\Delta}_i$ as in equation 3.19 and 3.18. The generalised or power mean is used in order to put an emphasis on the extreme values in the distributions that exhibit higher deviations. The deviation value for the whole city is also calculated for every time period separately in order to study the dynamics of the deviation value throughout the day.

$$\hat{\Delta}^t = \left(\frac{1}{N} \sum_{i=1}^N (\hat{\Delta}_i^t)^\alpha \right)^{\frac{1}{\alpha}} \quad (3.18)$$

$$\hat{\Delta} = \left(\frac{1}{N} \sum_{i=1}^N (\hat{\Delta}_i)^\alpha \right)^{\frac{1}{\alpha}} \quad (3.19)$$

3.4.2 Calculating the Dynamic Factor of a Node

A node is defined as dynamic if many vehicles pass through it and if the choices that drivers make at this node vary abruptly in time. In order to measure the variation of commuter choices the rate of change of turning probabilities of agents in time should be examined. After that the rate of change is weighed with the flow through it. In this way it can be measured how big and fast are the variations at the nodes are combined with how central their role in the traffic is. The following steps are taken to calculate the dynamic factor of a node and of a whole transportation system. The dynamic factor calculation follows closely the logic used for the deviation measure calculation. As a prerequisite, the turning probabilities should be calculated following the definitions described in step 1. The rest of the procedure is described in the following points.

3.4 Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes

1. Calculate the variation of turning probabilities at nodes:

The variation of the turning probabilities is calculated using a slowness measure similar to [60]. The derivative of the probability over time is examined. High values of the absolute value of this derivative implies high degree of variation of the turning probabilities and vice versa. This can be seen in Equation 3.20. Note that the term $\langle \rangle_t$ is used to depict averaging over t .

It is interesting to note that the derivative was chosen to determine the degree of variation rather the variance, which would be the more natural choice. This is due to the temporal properties of the used turning probabilities. The change of turning probabilities in time is observed and the sequence in which these alterations occur has a determining effect over traffic conditions. Therefore, a measure that takes into account this factor such as computing the derivative is preferred over the variance, which is intrinsically order invariant. Furthermore, the object of interest is the dynamics of the changes of behaviour at the nodes rather than their deviation, and naturally the changes in a time series are observed by examining their derivatives.

Since nodes are being examined, all the successors of a node should be taken into consideration in order to evaluate the variation of the node itself. Therefore, the variation of a node is defined as the average of the variations of the turning probabilities associated with it as shown in equation 3.21. An average is used instead of a simple summation in order to avoid cases of highly connected nodes, which exhibit static behaviour getting a high dynamic factor.

$$V_{ij} = \langle \|\dot{p}_{ij}^t\| \rangle_t \tag{3.20}$$

$$V_i = \frac{\sum_{k \in S_i} V_{ik}}{\|S_i\|} \tag{3.21}$$

2. Weight the variation of every node with the number of drivers that pass through it:

In order to differentiate between nodes with a variation value, which have different levels of traffic throughput, the variation of every node is weighed by the number of vehicles

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

utilising it. The dynamic factor of a node i is defined as in equation 3.23 and intuitively can be perceived to represent the speed at which the activity at this node is changing. More precisely, it represents how rapid and diverse are the changes in the typical choices that agents make at this node. In other words, the nodes that experience rapid changing dynamics and are critical in the sense of traffic demand will receive a high dynamic factor value.

The change in traffic demand throughout one time period is weighed by the number of drivers that pass through the node during this time. In case of a big change that does not affect many agents the dynamic factor is still small. Moreover, the logarithm of the flow is taken since, the main interest is the change of the turning probabilities. The term which takes into consideration the volume of vehicles, is added in order to distinguish between busy intersections and ones that have very small throughput since the latter may not have such a big effect on global traffic conditions. In case the dynamic factor is calculated with the absolute value of the flow, busy intersections that do not have that much variation in turning probabilities will get very high dynamic factors, which is undesirable.

$$D_{ij} = \left\langle \left\| \dot{p}_{ij}^t \right\| \log \sum_{j \in S_i} N_{ij}^t \right\rangle_t \quad (3.22)$$

$$D_i = \frac{\sum_{k \in S_i} D_{ik}}{\|S_i\|} \quad (3.23)$$

3. Extension for Calculating the Dynamic Factor of a Network

Similarly to the case with the deviation measure a normalized measure is included by using the congestion factor as shown in equation 3.26.

$$w_i = \frac{\sum_{j \in S_i} w_{ij}}{\|S_i\|} \quad (3.24)$$

$$C_i = 2000Lw_i \quad (3.25)$$

$$G_{ij}^t = \frac{N_{ij}^t}{C_i} = \frac{N_{ij}^t \|S_i\|}{2000L \sum_{j \in S_i} w_{ij}} \quad (3.26)$$

Dynamic factor measures are redefined with the normalized congestion factor term in equation 3.28. Since the traffic flow is already normalised the logarithm should not be taken anymore.

3.4 Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes

$$\hat{D}_{ij} = \left\langle \left\| \dot{p}_{ij}^t \right\| \sum_{j \in S_i} G_{ij}^t \right\rangle_t \quad (3.27)$$

$$\hat{D}_i = \frac{\sum_{k \in S_i} \hat{D}_{ik}}{\|S_i\|} \quad (3.28)$$

In order to compare the dynamic factor profiles of the traffic conditions of two or more cities one can compare the dynamic factor distributions of all nodes. A dynamic factor of a city M is defined as the generalised mean with power factor $\alpha = 2$ of the distribution of \hat{D}_i as in equation 3.29. The generalised or power mean, as in the case with the deviation measure, is used in order to put an emphasis on the extreme values in the distribution that make the network more dynamic.

$$M = \left(\frac{1}{N} \sum_{i=1}^N \hat{D}_i^\alpha \right)^{\frac{1}{\alpha}} \quad (3.29)$$

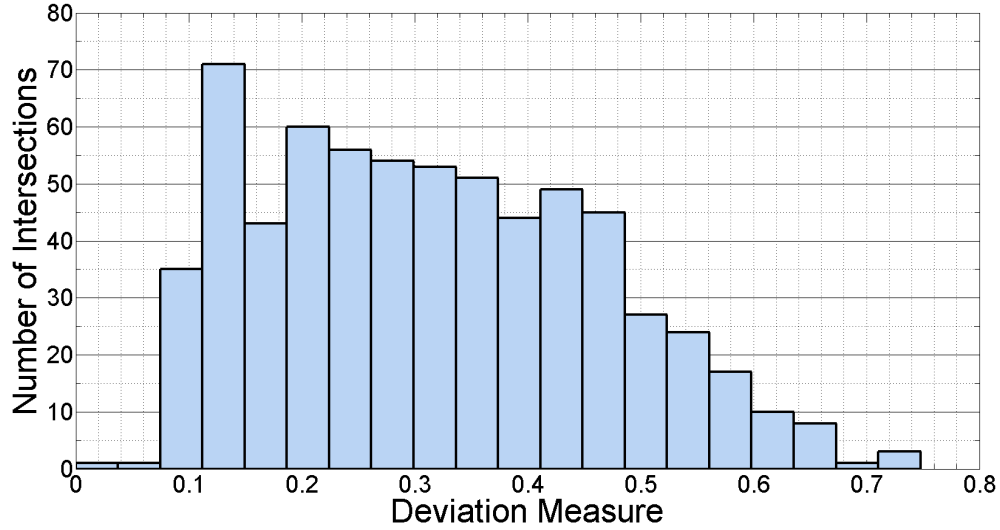
3.4.3 Deviation Measure Calculation for Singapore Case Study:

In order to apply the designed measures in a realistic scenario and supply the needed information for their computation, the traffic generation procedure described in Chapter 2 is used. Furthermore, in order to ensure easier assimilation of the results only nodes with a throughput higher than 10000 vehicles per day are examined. First, the deviation factors (Δ_i) of every node that satisfies the flow constraints is calculated and their distribution is shown in Fig. 3.4a

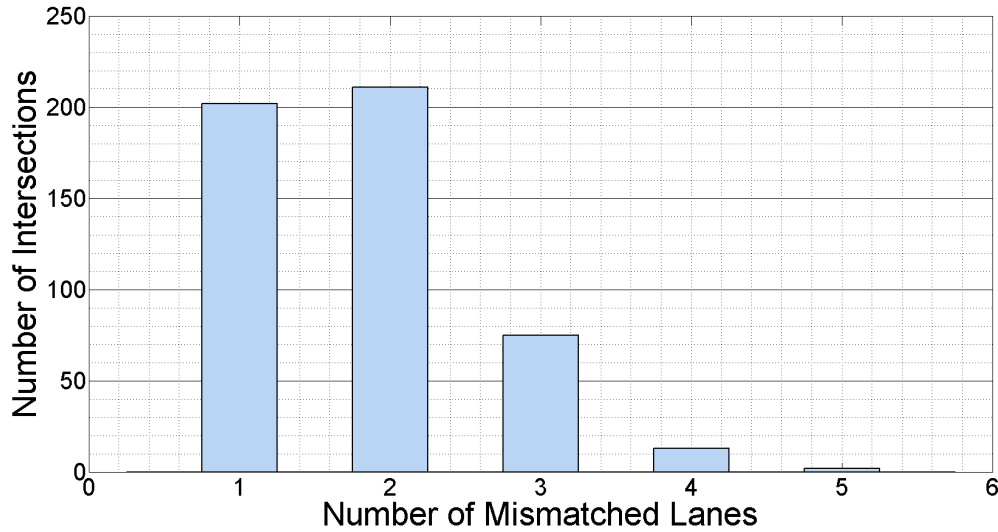
On Fig. 3.4a it can be observed that the distribution resembles a log normal distribution with a peak at around 0.1. The maximum deviation of a node in the examined case study is 0.74. It must be pointed out that the maximum possible value of Δ_i is 1. An example case of how the value can reach one is when all drivers passing through a certain node systematically (throughout the whole day) turn onto a one lane road while no drivers are turning onto a road with more than one lane coming from the same intersection. It can also be observed that the intersections that have a deviation value of 0 (perfect match) or lower than 0.1 are only two while the ones with a value around the peak, which is at 0.1 sum up to almost 70.

Although most of the intersections do not perfectly match the demand, the lane ratios between the successors might still be optimal. In other words, the distribution of lanes leading to the successor roads can still be optimal even if the deviation value is not exactly zero as shown in Fig.3.3. It is necessary to examine the absolute mismatch measure m_i as well in order

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE



(a)



(b)

Figure 3.4: Distributions of the measure Δ_i and the mismatch of number of lanes between road capacities and traffic demand. Fig. 3.4a presents the distribution of the deviation measure for every intersection in the city of Singapore that has a throughput higher than 10,000 vehicles per day. Fig. 3.4b is the distribution of the number of number of mismatched lanes, based on the deviation of the road infrastructure and the traffic demand, on intersections with throughput higher than 10,000 in cases where this number is bigger than 0. A value of 1 means that one lane should be moved from one successor road of the node to another in order for the road structure to be in optimal agreement with the traffic demand.

3.4 Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes

to estimate how many intersections do not have an optimal lane distribution. The measure represents the number of lanes that should be transferred from one road to another in order to optimally satisfy the demand. A distribution of this measure for the same set of intersections can be seen in Fig.3.4b, however only the intersection with $m_i > 0$ are included in the distribution.

Out of 711 examined high throughput intersections, 211 have an optimal lane distribution. It can be observed that there are around 200 intersections that can benefit from one lane being moved from one road to the other and there is one intersection that can benefit from redistributing five lanes. The existence of places with such extreme deviations can be explained as being caused by the heterogeneous nature of deviation between infrastructure and demand, which can manifest in highly sensitive locations existing on the road network.

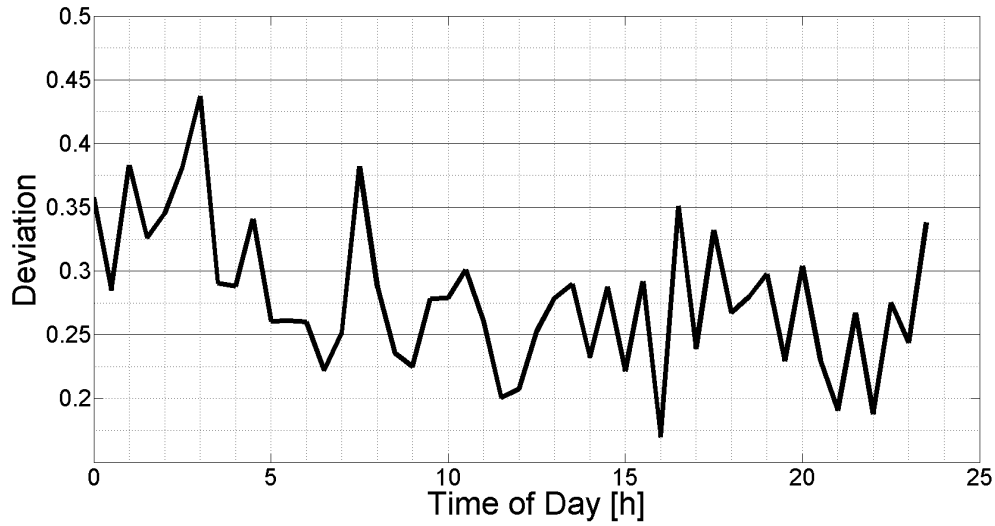
Next, let us examine the temporal nature of the measures. On Fig. 3.5a the evolution in time of the measure Δ^t is depicted, which summarizes the deviation of the whole network in the form of a single value for every time period. It can be observed that the deviation has a stable nature with a coefficient of variation $c_v = \frac{\sigma}{\mu}$ of only 0.0019. Therefore, the overall deviation of the network is time invariant according to the suggested metric.

On Fig.3.5b the evolution in time of the measure Δ_i^t for the intersection that shows the highest overall deviation is depicted. It can be observed that there is no significant variation as the coefficient of variation c_v is 0.0470, however, the curve is relatively low throughout the morning (7 : 30 – 9 : 30) and evening (17 : 30 – 19 : 30) rush hours, which means that at this particular intersection, the mismatch between demand and road capacities is getting smaller with increasing traffic volumes.

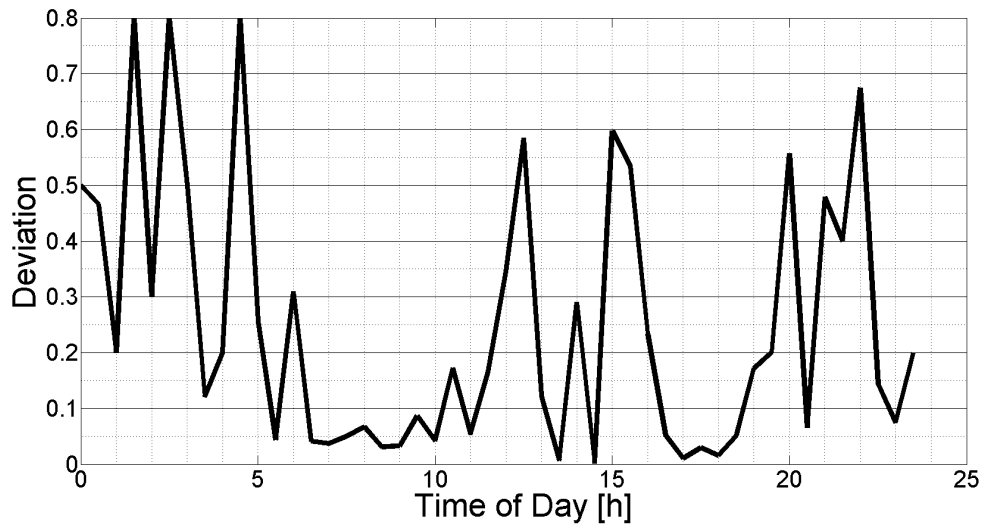
This is the expected result since the intersections have most likely been optimised to perform best during rush hour, since the deviations at those periods of time are the ones that can lead to traffic jams. A high deviation value at midnight may not be of such concern since even if the road capacities do not match the demands, the vehicle flows are not big enough for a tangible effect on the traffic conditions to be sensed. A significant mismatch during morning or evening traffic peak, however, will inevitably lead to a setting of congestion and overall reduced network performance.

Fig. 3.6 presents the spatial distribution of mismatched intersections. The biggest cluster of mismatches is observed at the central business district (south central part), as it is the most dynamically changing location in the city. The high pace of emergence of new buildings and businesses, which attract new employees or move existing ones to new places inevitably results in a mismatch of the traffic demand with the relatively slowly changing road structure.

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE



(a)



(b)

Figure 3.5: Evolution of the deviation measures in time. Fig. 3.5a shows the measure Δ^t for different values of t . This is the deviation of the road infrastructure of the whole network from the traffic demands. Fig.3.5b depicts the evolution in time of the measure Δ_i^t of the intersection with highest degree of deviation.

3.4 Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes

It can also be observed that the other most mismatched intersections (red and dark orange) are positioned along major roads that connect the down-town area with the east, west and northern parts of the city. Those areas have been growing extensively in the past decade as a result of the government's government attempt to relax traffic demand in the down-town area by building self sufficient districts in various parts of the island. As a result of this fast OD alteration, the road network is lagging in its development and thus such levels of mismatch between demand and infrastructure capabilities are not surprising to observe.

3.4.4 Dynamic Factor Calculation for Singapore Case Study

Similar to the deviation and mismatch case, the dynamic factors are presented only of the nodes that have a significant throughput, which is set to be more than 10,000 vehicles per day. After calculating the dynamic factor of every node that satisfies the constraints, a distribution of the dynamic factors throughout the network is acquired as shown in Fig. 3.7

The observed distribution resembles a log-normal distribution peaking at dynamic factor value of 0.75. A plateau can be noticed between 1 and 1.5, just after the peak. This might be due to standard degrees of variation at the nodes with varying flows of vehicles. The tail part contains very few intersections with high dynamic factors. It seems like the distribution has a fat tail and the nodes contained in the far right part exhibit abruptly and dynamically changing turning probabilities and high traffic of vehicles. These are the intersections which present the biggest challenges for traffic control and road infrastructure planning. Fig. 3.8 presents the evolution in time of the turning probabilities of the most dynamic node in the network. It can be observed that all three options that the drivers can choose from are varying abruptly in time and that every option is the most preferred one during at least one period of the day.

The highly heterogeneous distribution of the dynamic factor can also be observed in its spatial distribution on Fig.3.9. The majority of indicated locations are either intersections between major roads or are connecting residential and business areas. The latter case presents high dynamic factors due to qualitatively different demands in the morning and evening. The first type is created as a result of changes due to a steady dynamic nature of events at those intersections. A cluster of points with high dynamic factor at the central business district (south central part) can also be observed, which can be due to the high concentration of business offices and constant movement of commuters within this area.

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE



Figure 3.6: Spatial distribution of mismatched intersections in the city of Singapore according to the m_i measure. Big and red dots represent high number of mismatching lanes (maximum is 6), while green and small dots represent intersections with lower number of mismatched lanes (minimum is 1).

3.4 Defining the Measure of Deviation Between Network Capacity and Traffic Demand and Dynamic Factor of Nodes

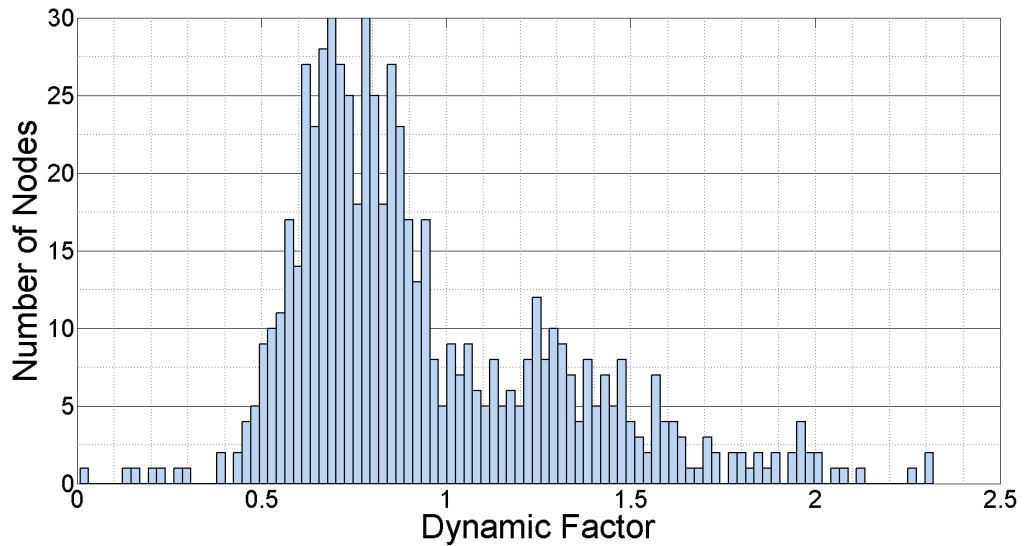


Figure 3.7: Histogram depicting the distribution of Dynamical Factor values in the Singapore road network. Only intersections with daily throughput higher than 10,000 vehicles are taken into consideration.

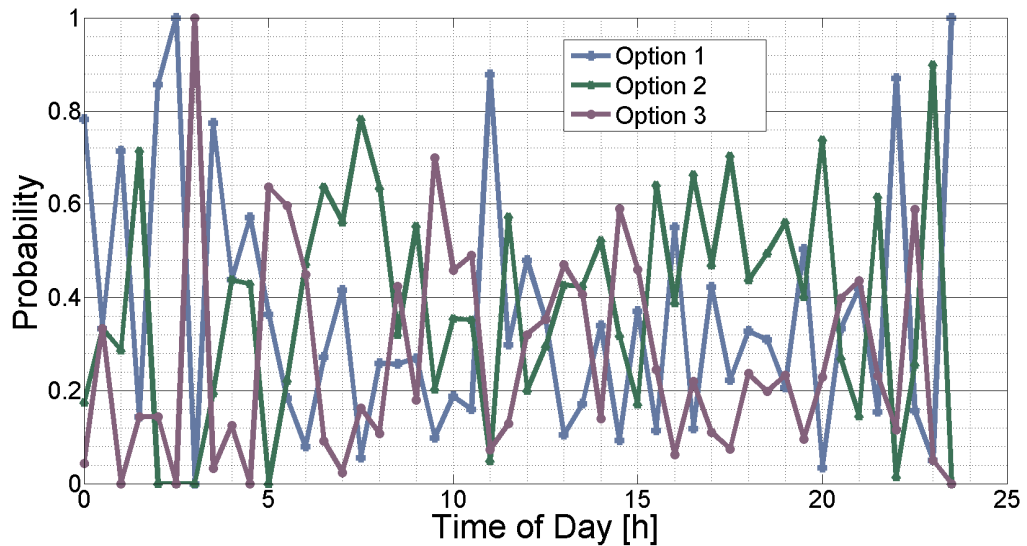


Figure 3.8: Turning probabilities of the most dynamic node according to the model in the city of Singapore. The distinct time series represent the probability that a driver would choose the corresponding outgoing road from the examined node throughout the day.

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

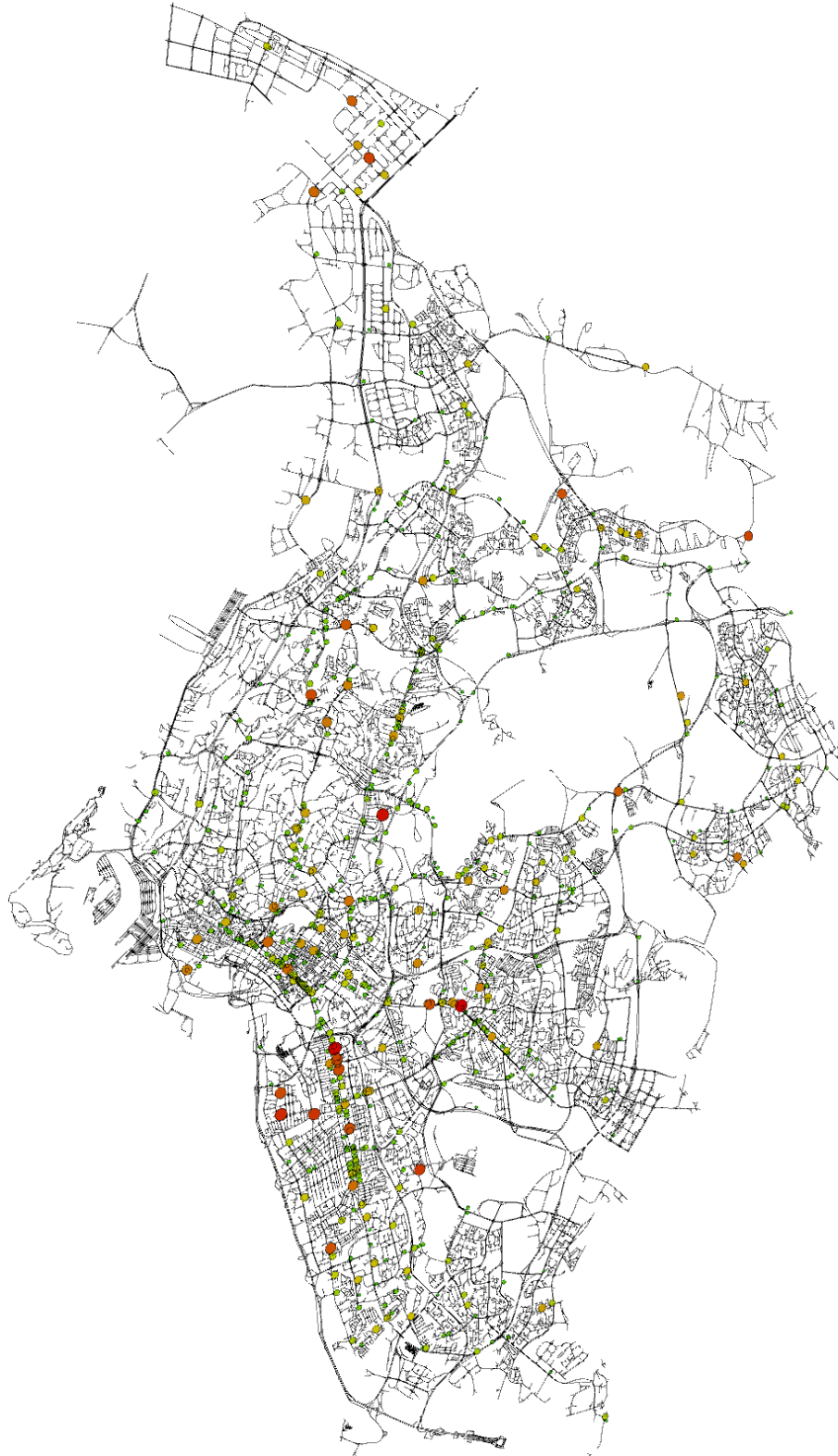


Figure 3.9: Spatial distribution of dynamic factor of intersections in Singapore according to simulated agent routes. Big red dots mean high dynamic factor and small green dots mean low dynamic factor. Only intersections with throughput higher than 10,000 cars per day are considered.

3.5 Application of Recommended Changes in Infrastructure

ture

After the identification of mismatched intersections assured by the defined mismatch and deviation measures, the analysis is extended by further investigating and validating those findings. A simple and generic way of performing this task would be to apply the recommended changes to the road network and observe the changes in traffic conditions stemming from them. The measure of improvement is chosen to be the overall change in travel time of the commuting population. Furthermore, it is important to address the quantitative accuracy of the defined measure or, in other words, to answer the question, whether “fixing” the most mismatched intersection brings the biggest improvement in overall travel time. In order to do this, the correlation between the defined measures and the changes of system performance are examined.

3.5.1 Methods

The procedure of applying changes to the road network comprises of three main steps:

- Choose mismatched locations identified by the defined measures. This step provides a starting set of intersections for the study filtering out the locations that do not need to be examined.
- Calculate the optimal distribution of lanes at the chosen locations. Then the appropriate changes are applied to the road network in order to evaluate the effect on the population commuting time.
- Calculate the travel times of the population before and after the changes are applied to the road network. This step involves the macro-simulation approach described in Chapter 2 in order to acquire the desired traverse times on every road segment.

3.5.1.1 Identification of Locations to be Fixed

There are two main choices about the measure to be used to initially select the locations to be examined: the deviation measure δ_i or the mismatch measure m_i . It must be noted that the goal of this step is to identify all the intersections that perform in a suboptimal way due to a discrepancy between infrastructure and demand. It would therefore be pointless to select intersections that exhibit problematic behaviour, which, however, cannot be fixed by lane redistribution, which is the approach undertaken in this work. Therefore, it is reasonable to use

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

the mismatch measure since it corresponds to exactly the intersections that can benefit from lane redistribution and also orders them by the number of lanes that should be exchanged. As a reminder, the mismatch measure filters out all locations that have a non-zero deviation measure, but still possess optimal lane ratios between the turn options. At this point a fundamental difference between the deviation and mismatch measures should be pointed out.

The deviation measure captures the absolute discrepancy between infrastructure and demand in the sense that it performs a weighted average of the absolute difference between optimal and real infrastructure. This property is valuable for analysing the performance of the system as a whole, however, it might be misleading if used for localized cases. The reason for this is the fact that the mismatch can occur in both directions, namely during one part of the day, the optimal lane ratios might be in one extreme while in another portion of the day another extreme distribution might be beneficial. In such cases, it might happen that the initial lane distribution is optimal in the long term and changes will lead to worsening of traffic conditions. This observation does not mean that the deviation measure has no use. The existence of such places is undesired in a road network, since they become bottlenecks during both rush hour periods. Those locations have quantitatively similar discrepancies in the sense of created congestion but qualitatively different reasons for those congestions. While the mismatch measure cannot differentiate between them, the deviation measure can be used as a heuristic together with the dynamic factor for locating precisely such cases.

As in the previous section, however, only locations with a daily flow of more than 10,000 vehicles were examined since only then, theoretically, congestion can be observed. To summarize the chosen intersections are:

$$\forall i : m_i > 0 \wedge \sum_t \sum_{j \in S_i} N_{ij}^t > 10,000 \quad \forall t$$

3.5.1.2 Calculation of Optimal Lane Distribution and Road Network Changes at Chosen Locations

In order to apply the changes to the network to eliminate the mismatch, one has to know the optimal number of lanes of every successor of node i , which is defined as r_{ik} as noted in the previous section. The calculation of r_{ik} involves taking a weighted average for the number of vehicles choosing every option in time and then calculating the weighted probabilities for the whole day and the respective optimal lane distribution. In order to obtain the optimal probabilities, the real flows are weighted by the throughput at the node i for time period t .

3.5 Application of Recommended Changes in Infrastructure

The reason behind this approach is the non-linear relationship between flow and time. In other words, as the number of vehicles on the road increases the travel time increases in a non-linear fashion and thus periods with high throughput are potentially significantly more important than periods with small volumes of vehicles passing through.

$$\hat{N}_{ij}^t = N_{ij}^t \sum_{k \in S_i} N_{ik}^t \quad (3.30)$$

Then the total weighted flow is calculated by summing up the total weighted flows and the optimal probabilities are calculated according to the respective ratios.

$$\hat{N}_{ij} = \sum_t \hat{N}_{ij}^t \quad (3.31)$$

$$p_{ij}^{\hat{}} = \frac{\hat{N}_{ij}}{\sum_{k \in S_i} \hat{N}_{ik}} \quad (3.32)$$

Following this the desired number of lanes r_{ik} is simply computed by choosing the appropriate number of lanes corresponding to the optimal probabilities $p_{ij}^{\hat{}}$:

$$r_{ik} = \min(1, \lfloor p_{ij}^{\hat{}} \sum_{l \in S_i} w_{il} \rfloor) \quad (3.33)$$

Next, the number of lanes in the network description structure is changed in order to prepare for the final travel time calculation step.

3.5.1.3 Calculation of Effects of Mismatch Fix

In order to evaluate the effects on congestion of the alterations suggested in the previous section, a comparison between the travel times before and after the changes in the infrastructure should be performed. Therefore, a way of transforming the flows on the road segments into travel times is desired. The usual approach in transportation science is to use one of the flow-time relationships and, in this work, the Bureau of Public Roads (BPR) function was chosen. It is a non-linear function that depends on the flow and the capacity of the road (number of lanes mostly) in order to predict what will be the transit time along a road segment. The calibration of the parameters of the BPR function with real world data and the validation of the results can be found in more detail in Chapter 2. The BPR function is formalized as:

$$t_{ij} = \frac{l_{ij}}{v_f^s} \left(1 + \alpha^s \left(\frac{F_{ij}}{2000w_{ij}t} \right)^{\beta^s} \right) \quad (3.34)$$

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

where t_{ij} is the time to traverse link ij , v_f^s is the free flow velocity at a link with speed limit s , α^s and β^s are coefficients to be calibrated for different types of roads, F_{ij} is the flow on link ij , w_{ij} is the number of lanes on ij and t is the time for which the flow is observed. The main measure that will be used in the following results chapter is the difference between the travel time before and after the changes in the road network multiplied by the number of people utilizing the location in question.

$$I_{ij} = F_{ij} \left(\frac{l_{ij}}{v_f^s} \left(1 + \alpha^s \left(\frac{F_{ij}}{2000w_{ij}t} \right)^{\beta^s} \right) - \frac{l_{ij}}{v_f^s} \left(1 + \alpha^s \left(\frac{F_{ij}}{2000r_{ij}t} \right)^{\beta^s} \right) \right) \quad (3.35)$$

$$I_{ij} = F_{ij} \alpha^s \left(\frac{F_{ij}}{2000(w_{ij} - r_{ij})t} \right)^{\beta^s} \quad (3.36)$$

$$(3.37)$$

In order to calculate the overall improvement of fixing the mismatched intersection all the improvements of successors of node i must be summed up:

$$I_i = \sum_{j \in S_i} I_{ij} \quad (3.38)$$

More robust metrics of improvement will be defined in the next section.

3.5.2 Results

This section presents the results of the simulated traffic conditions as a result of the recommended changes to the traffic network. Besides the examination of the pure improvement of traffic time, two other questions need to be addressed. First, how sensitive are the intersections to changes in their lane distributions and are there locations that are *super-sensitive* to such alterations. The identification of such places allows for effective steering of traffic conditions with minimal effort, which is a highly desirable goal for a transportation network. Second, is the sensitivity of intersections correlated with the defined measures. The deviation measures that were defined perform well in identifying locations that exhibit high degrees of discrepancy between demand and infrastructure, thus pointing to areas that are problematic but cannot necessarily be fixed locally. Similarly, the dynamic factor identifies locations that are likely to exhibit problems throughout the day due to highly fluctuating traffic conditions, which as well cannot necessarily be fixed. Finally, the mismatch measure identifies only locations that will benefit from local concrete changes of the road network. It will be interesting to see, with which measure do the sensitivity of locations correlate to the most.

3.5.2.1 Improvement of Traffic in the Time Domain

Although it is intuitive that the recommended changes in the road network will benefit the system, it must be examined where those improvements are most notable during the day. Fig. 3.10 shows the total saved time by applying all recommended changes in the system at once during the different periods of the day. It can be noticed that very little effect is observed during the non-rush hours as the before and after curves are virtually the same. During peak traffic conditions, however, it can be observed that there is a significant difference between the total travel times of the commuting population.

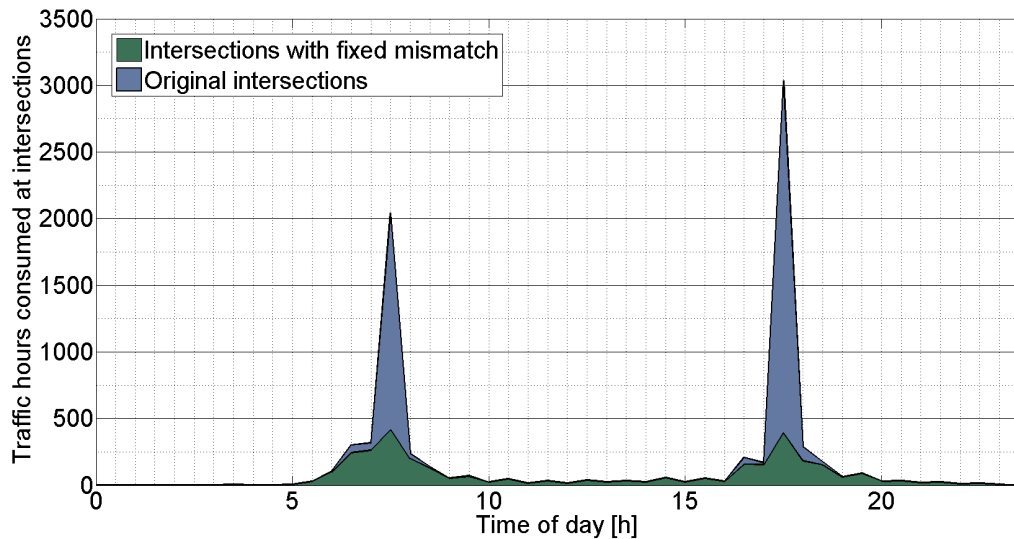


Figure 3.10: Comparison of time lost at chosen intersections with throughput higher than 10,000 vehicles between original intersection lane distribution and after recommended mismatch fix was applied to the road network.

This rather extreme difference between the performance of the network alteration during the different periods of the day is due to the highly heterogeneous traffic distribution over time (the existence of morning and evening rush-hours) and the non-linearity of the flow-time diagram, which introduces big changes in travel times for smaller in magnitude changes in the flows of vehicles on a given road segment. It can be observed that the peaks during morning and evening commute time are much less pronounced after the road network alteration, which sums up to more than 4,600 hours of travel time saved on a daily basis from redistributing the lanes at about 500 intersections.

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

3.5.2.2 Measures of Sensitivity

Measuring saved time, although useful in general terms of transportation system optimization, does not give much insight about the local influence of changes. Furthermore, some intersections may have the ability (flexibility) to redistribute more lanes than others. As redistribution effort is correlated with the number of intersections it will be informative to see, what is the time improvement per redistributed lane:

$$U_i = \frac{I_i}{m_i} \quad (3.39)$$

To further normalize the measure of sensitivity can also be defined as the time saved per lane per meter. It might happen that in some cases the road segment that receives an extra lane is much longer than the one that loses a lane. In order to exclude such influences, the time difference for every successor is divided by its length.

$$\hat{U}_i = \frac{\sum_{j \in S_i} F_{ij} \alpha^s \left(\frac{F_{ij}}{2000 (w_{ij} - r_{ij}) t} \right)^{\beta^s}}{m_i l_{ij}} \quad (3.40)$$

Fig. 3.11 shows histograms of the three measures of sensitivity.

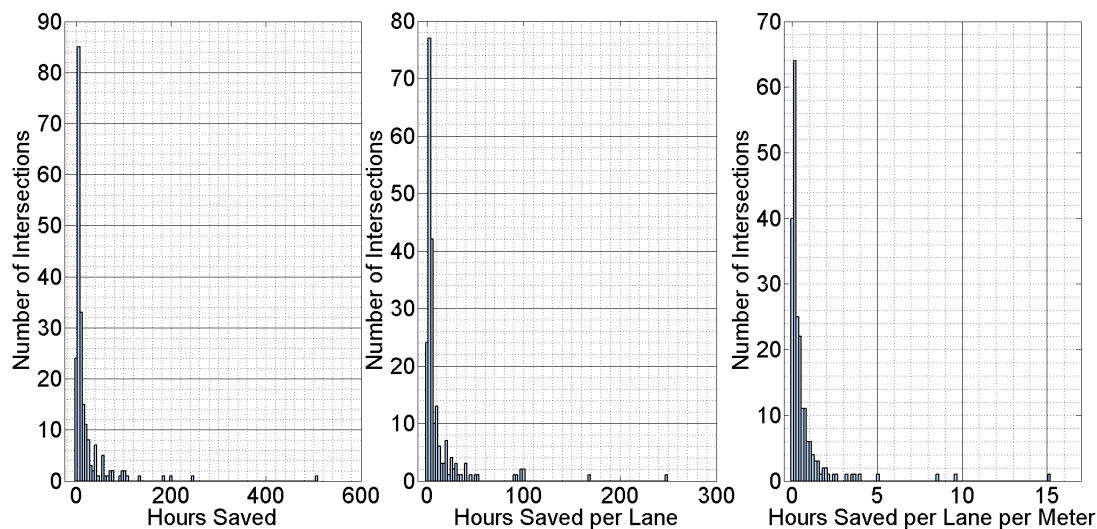


Figure 3.11: Distribution of saved time measures of the chosen intersections with throughput higher than 10,000 vehicles. From left to right: total time saved, time saved per exchanged lane, time saved per exchanged lane per meter.

3.5 Application of Recommended Changes in Infrastructure

From the figure it can be observed that the majority of intersections do not bring much improvement as the histograms resemble exponential distributions. There is a small number of intersections that form a fat tail of the distribution, which are responsible for most of the saved time. Even the fully normalized measure of time saved per lane per meter exhibits highly heterogeneous form, which points to the existence of *super-sensitive* locations, whose alteration may lead to significant change in overall traffic conditions.

In order to visualize the heterogeneous distribution of saved time in a better way Fig. 3.12 pictures the distribution of saved time according to percentiles of the set of locations ordered by their impact.

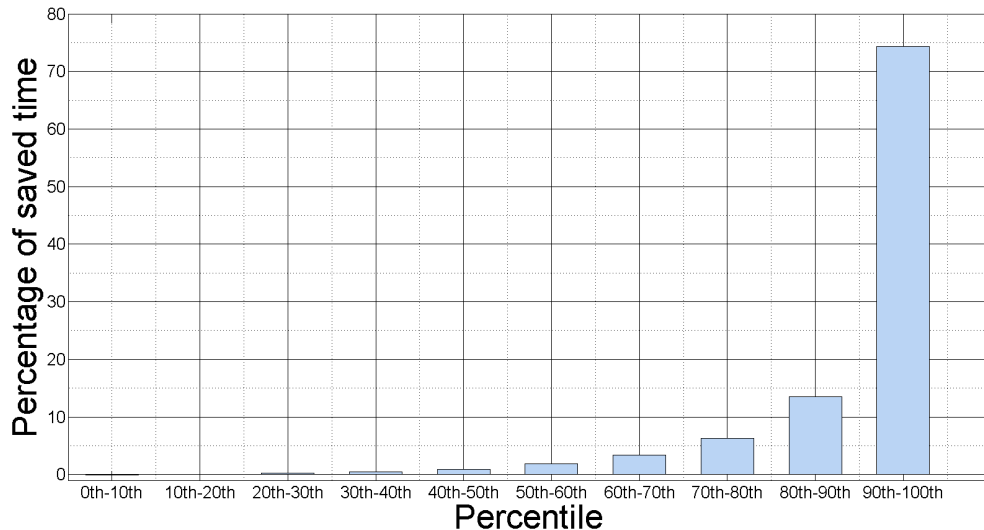


Figure 3.12: Distribution of the saved time by percentiles grouped according to total saved time of the chosen intersections with throughput higher than 10,000 vehicles.

The examined locations are ordered according to the time saved after lane redistribution. It can be seen that locations in the top percentile (90th to 100th percentiles) are responsible for nearly 75 percent of the overall saved time. These are the locations that will benefit the most from a redistribution, corresponding to roughly 50 intersections, which save 3 times more time than all 450 other intersections combined. Fig. 3.11 and 3.12 clearly show that some locations are far more critical than others. Intuitively, it makes sense that since traffic is heterogeneously distributed on the road network, some locations have far greater flows than others; this results in changes at those places having more effect than changes on others. Please be reminded,

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

however, that in this case all examined locations have high throughputs (more than 10,000 vehicles per day). Furthermore, the correlation coefficient between the traffic flow and time saved is 0.22 as seen from Table 3.2, which refutes the intuitive hypothesis that busier locations can save more time by lane redistribution.

3.5.2.3 Spatial Distribution of Saved Time

Fig. 3.13 illustrates the positioning of sensitive locations geographically in the city of Singapore. It must be pointed out that due to the highly heterogeneous nature of the time saved distribution the third root of the saved time per lane measure had to be taken in order for more intersections to be visible on the map.

Singapore's financial district is located in the south-central part of the city, which as can be seen is a zone with a high concentration of mismatched intersections. However, the most mismatched intersections are not in the region itself, but rather on the roads connecting it to other parts of the city (like the eastern and northern regions) which are mostly residential zones. Furthermore, it can be noticed that occasionally the locations are along a single major road. This can be explained by the phenomenon suggested in the introduction of construction plans reacting to infrastructure changes (like building of new roads) thus producing extra demand for initially empty locations that were not taken into consideration in the initial road system planning.

3.5.2.4 Correlation Analysis of Measures

As a final step of this study, let us examine the correlations between the various defined metrics and measures in the chapter represented in Table 3.2.

Attention should be given to the realization that sensitivity measures show weak correlation with the traffic flow, which indeed suggests that the examined locations do not necessarily save more time because they exhibit higher flows of vehicles. The metric that presents the most interest is the saved time per lane per meter since it captures in the most robust way the sensitivity of a location. It turns out that it is weakly correlated with both the deviation and the dynamic factor measures. It is, however, important to note that the normalized deviation measure defined in order to be able to compare traffic conditions between different sized cities shows a much better correlation to the saved time per lane per meter metric. It is therefore, advisable that precisely this measure is used in order to spot the *super-sensitive* locations in a city. One reason for this outcome may be the formulation of the normalized deviation measure.



Figure 3.13: Spatial distribution of saved time per lane. The third root was taken from the actual measure in order for more intersections to be observable due to the highly unbalanced distribution of saved time among intersections.

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

Table 3.2: Correlation coefficients between chosen sensitivity metrics, defined measures and traffic flow.

| | Saved Time | Saved Time/Lane | Saved Time/Lane/m | Traffic Flow | Dynamic Factor | Deviation Measure | Normalized Deviation Measure |
|------------------------------|------------|-----------------|-------------------|--------------|----------------|-------------------|------------------------------|
| Saved Time | 1 | 0.8856 | 0.5046 | 0.2235 | 0.0625 | 0.2424 | 0.4814 |
| Saved Time/Lane | 0.8856 | 1 | 0.6776 | 0.2312 | 0.0504 | 0.2235 | 0.5427 |
| Saved Time/Lane/m | 0.5046 | 0.6776 | 1 | 0.2672 | 0.0848 | 0.2154 | 0.6321 |
| Traffic Flow | 0.2235 | 0.2312 | 0.2672 | 1 | 0.4559 | 0.4144 | 0.3420 |
| Dynamic Factor | 0.0625 | 0.0504 | 0.0848 | 0.4559 | 1 | 0.1473 | 0.0743 |
| Deviation Measure | 0.2424 | 0.2235 | 0.2154 | 0.4144 | 0.1473 | 1 | 0.5728 |
| Normalized Deviation Measure | 0.4814 | 0.5427 | 0.6321 | 0.3420 | 0.0743 | 0.5728 | 1 |

3.6 System Optimum Lane Distribution Problem

As a reminder, the idea behind it is to not include absolute values of the flows on the roads but rather work with congestion factors, which turn the absolute flows into a ratio between the flow and the capacity of the road. This is the only measure that actually takes into consideration the number of lanes on the road segments and therefore represents the potential congestion infused time delays. The disadvantage of this approach is that this measure might show a high value for a smaller mismatched intersection that is not going to have a big effect on the overall traffic performance. It turns out, however, that such roads are hard to come upon, since no such cases were observed in the study performed.

3.6 System Optimum Lane Distribution Problem

In line with the findings from the previous sections that demonstrate the existing mismatch between traffic demand and the infrastructure a more generalised approach can be undertaken in order to evaluate the degree of discrepancy. This section presents the problem of finding the optimal number of lanes on every existing road for minimizing the overall population travel time subject to a set of constraints. In order to compute the travel times the BPR function will be used once more, turning the calculated flows on every road segment to travel times of commuters.

As pointed out in the introduction of this chapter, it is vital that traffic demands are met by the infrastructure and that roads are neither under nor over utilized. As shown by the fundamental law of traffic [28], building more road infrastructure can have negative effects on the overall system performance. The constraints of the optimization problem will be that road lanes can be only redistributed. In other words, the total length of the roads multiplied by the number of lanes of every segment should stay constant. The optimization problem can be formalized as follows:

$$\min_{\mathbf{w}} T(w) = \sum_i t_i(w_i)F_i = \sum_i \frac{l_i}{v_f^s} \left(1 + \alpha^s \left(\frac{F_i}{2000w_it} \right)^{\beta^s} \right) F_i \quad (3.41)$$

subject to

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

$$\sum_i l_i w_i \leq \sum_i l_i w_i^0 \quad (3.42)$$

$$w_i \geq 1 \quad \forall i \quad (3.43)$$

$$w_i \leq 10 \quad \forall i \quad (3.44)$$

$$(3.45)$$

where $t_i(w)$ is the BPR function for traverse time dependent on the number of lanes w_i . The constraint 3.43 defines the upper bound for the total length of the road network and constraint 3.44 ensures that all existing roads will have at least one lane allocated to them. Finally, constraint 3.45 ensures that there are not segments with more than 10 lanes. Currently in the Singapore network, the number of lanes on the widest road is 9. Typically, the objective function in this problem is minimized with respect to the flows F_i in order to obtain the system optimum path distribution. This is a separate problem that will be discussed in more detail in Chapter 4.

It must be noted that the objective function $T(w)$ is non-linear, however, convex due to the fact that it is a sum of convex functions $t_i(w_i)$. Furthermore, the constraints are linear inequalities, which ensures that the feasibility space of the problem is also convex. Therefore, any minimum that is found by the optimization algorithm is guaranteed to be unique.

In order to solve the optimization problem the Interior Point Method [61] is used. The barrier function is:

$$B(\mathbf{w}, \lambda) = T(w) - \sum_i \lambda_i c_i \quad (3.46)$$

where c_i are the constraints in the form $c_i \geq 0$ and λ_i are the Lagrange multipliers such that $c_i(w)\lambda_i = \mu \forall i$, where μ is a small positive scalar that converges to 0 when a solution is reached. The gradient of $B(\mathbf{w}, \lambda)$ therefore becomes:

$$\nabla T(\mathbf{w}) - A^T \lambda = 0 \quad (3.47)$$

where A is the Jacobian of the constraints $c(w)$. Next, Newton's method is applied to get the update directions (d_w and d_λ) for λ and \mathbf{w} :

$$\begin{pmatrix} H_B & -A^T \\ \Lambda A & C \end{pmatrix} \begin{pmatrix} d_w \\ d_\lambda \end{pmatrix} = \begin{pmatrix} -\nabla T(\mathbf{w}) + A^T \lambda \\ \mu^1 - C \lambda \end{pmatrix}$$

3.6 System Optimum Lane Distribution Problem

where H_B is the Hessian of the barrier function, Λ is the diagonal matrix of λ , C is the diagonal matrix with the values of $c(\mathbf{w})$ and μ^1 is a column vector with all values μ .

The update rule is:

$$(\mathbf{w}, \lambda) \rightarrow (\mathbf{w} + \alpha d_w, \lambda + \alpha d_\lambda) \quad (3.48)$$

Since the constraints are linear the Hessian of B is a diagonal matrix containing the second partial derivatives of the barrier function:

$$\begin{bmatrix} \frac{\partial^2 T(\mathbf{w})}{\partial^2 w_1} & 0 & 0 & \dots & 0 \\ 0 & \frac{\partial^2 T(\mathbf{w})}{\partial^2 w_2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{\partial^2 T(\mathbf{w})}{\partial^2 w_N} \end{bmatrix}$$

and in the case of the BPR function:

$$\frac{\partial^2 T(\mathbf{w})}{\partial^2 w_i} = \frac{l_i}{v_f^s} \alpha^s \left(\frac{F_i}{2000t} \right)^{\beta^s} F_i \beta^s (\beta^s + 1) \left(\frac{1}{w_i} \right)^{(\beta^s + 2)} \quad (3.49)$$

The Jacobian of the constraints A is:

$$\begin{bmatrix} -l_1 & -l_2 & -l_2 & \dots & -l_N \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

and the gradient of the objective function $\nabla T(\mathbf{w})$ in the i -th direction is:

$$\nabla T(\mathbf{w}_i) = -\frac{l_i}{v_f^s} \alpha^s \left(\frac{F_i}{2000t} \right)^{\beta^s} F_i \beta^s \left(\frac{1}{w_i} \right)^{(\beta^s + 1)} \quad (3.50)$$

Given the sparse nature of all the matrices, it is relatively easy to compute the update steps numerically despite the large number of segments $N = 240,000$.

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

3.6.1 Results

The optimization problem was numerically solved in Matlab using the *fmincon* function with provided analytical expressions for the Hessian of the barrier function and the gradient of the objective function in order to speed up the process. After 68 iterations with initial step size parameter $\alpha = 0.1$ the optimization stopped after satisfying the function tolerance condition set to 10^{-3} . Fig. 3.14 shows the progress of the optimization algorithm as a function of the iteration count. It can be observed that the optimal lane distribution leads to a 36% population travel time decrease.

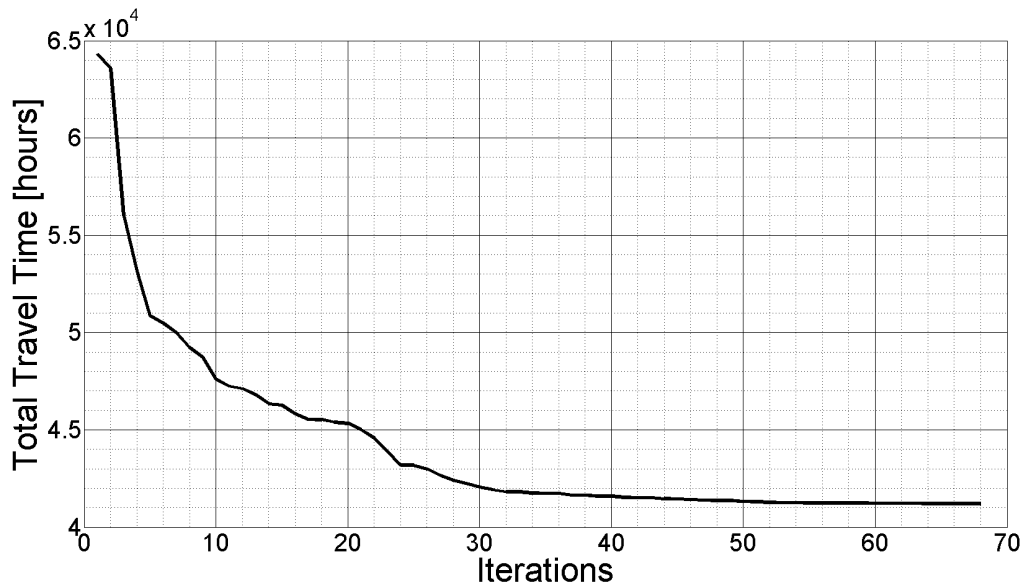


Figure 3.14: Progress of the optimization algorithm as a function of the iteration count.

Fig. 3.15 depicts the difference between the optimal solution and the current lane distribution in Singapore. It can be observed that the major road arteries are allocated more lanes in the optimal solution, while the more minor roads give away lanes. It should be noted that the roads leading to the central business district are the ones that receive the biggest amount of lanes, which is also intuitive since they exhibit the highest traffic flows. In total around 31,000 road segments do not already have their optimal number of lanes, which means that about 87% of the road network segments are optimal with respect to their lane count. There is an abnormality occurring at the south west part. It can be seen that there is a specific route that is heavily utilized for some reason, while the streets next to it remain empty. This is most

3.6 System Optimum Lane Distribution Problem

probably not a realistic scenario and might be due to the fact that there is a strictly shorter path there (lack of any viable alternatives). Furthermore, there also might be a heavily represented origin or destination at the region that influences the agent generation process. Abnormalities in the same region can also be observed in all other spatial profiles in the chapter. At this point the reader is asked ignore the results for this area since they might not be representative of the reality.

These observations can be further validated by exploring the distribution of difference in lane count between different types of roads with respect to their class (express-ways, major, minor) (Fig. 3.16) and their length (Fig. 3.17). The segments that experience the biggest traffic demand and congestion, which are the express-ways, receive the biggest amount of lanes, which amounts to about 50% relative increase in the number of lanes. Major roads have about 40% relative increase and the minor roads are the ones that give away lanes with a relative loss of 6%. It must be noted that the number of lanes given away does not equal the lanes gained since the constraint of the optimization problem is to keep the total length of all lanes combined and not their count, thus keeping the infrastructure capacity constant.

To further observe this, Fig. 3.17 presents the distributions of lane exchange within various road groups based on segment length. It can be clearly observed that long roads give away lanes and medium sized segments mostly redistribute within the group. The really short segments, which are usually associated with turns at intersections, linking the two intersecting main roads, receive large amounts of lanes for the optimal solution and almost double in number with a 89% relative increase. Those roads are evidently bottlenecks in the current road network and despite their small length, the network can benefit significantly from their expansion.

In conclusion the suggested generalised measure of mismatch W is defined as the total length of the road segments that need to be redistributed. Once more, this approach is preferred against counting the number of redistributed lanes since it captures the effort required to achieve optimality rather than calculating the distance between the current and the optimal solutions:

$$W = \sum_i \frac{l_i |w_i - w_i^{opt}|}{2} \quad (3.51)$$

W is measured in meters, which might vary in magnitude from system to system. In order to compare the mismatch of one city to another a relative measure can be defined, namely:

$$\hat{W} = \frac{W}{\sum_i l_i w_i} \quad (3.52)$$

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

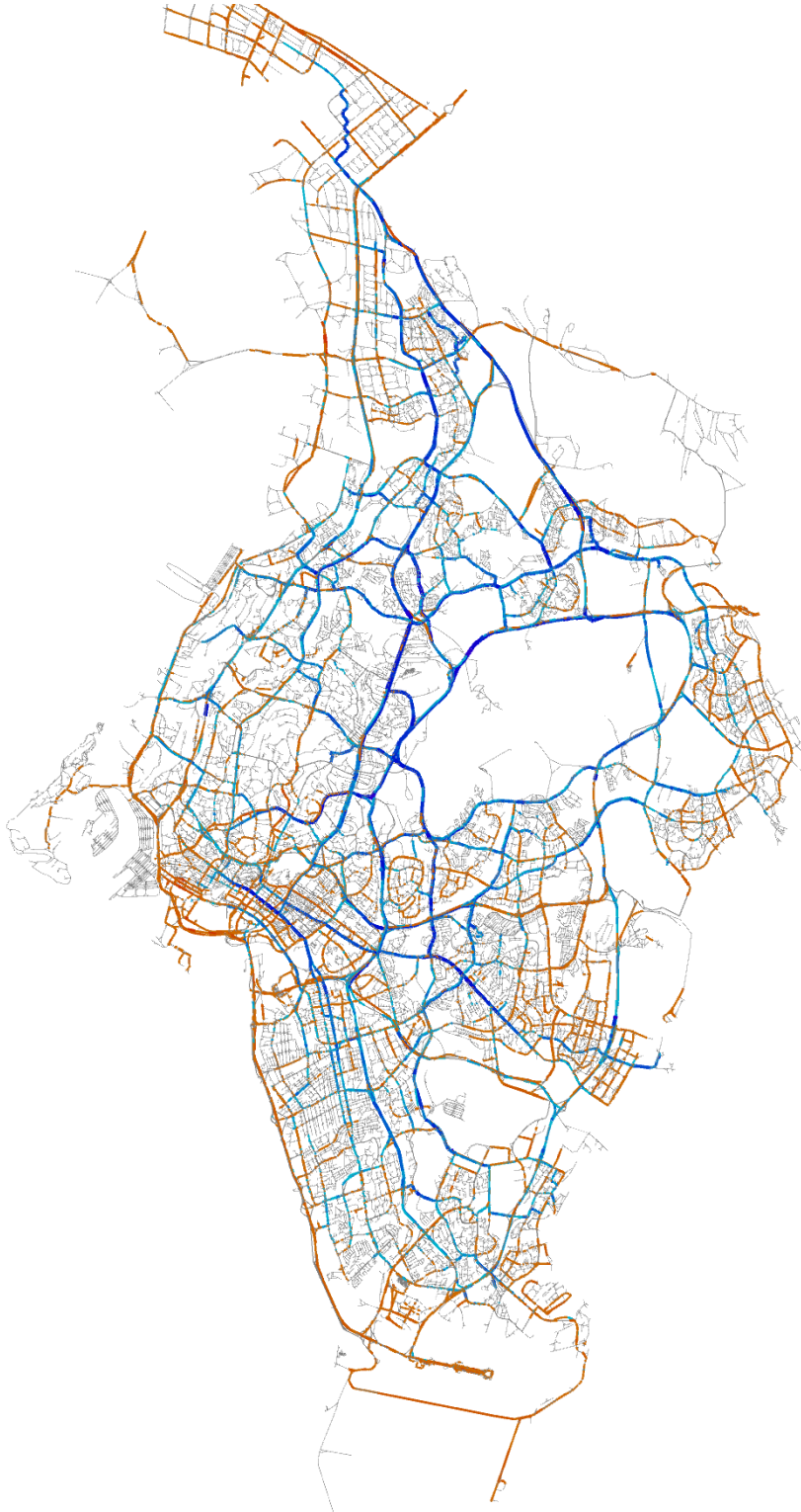


Figure 3.15: Spatial distribution of the difference in lanes between the current infrastructure and the optimal lane allocation found by the optimization problem. The blue gamma depicts road segments that have received lanes, while the red gamma depicts the roads that have given away lanes. The darker the colour the bigger the absolute value of the lane exchange.

3.6 System Optimum Lane Distribution Problem

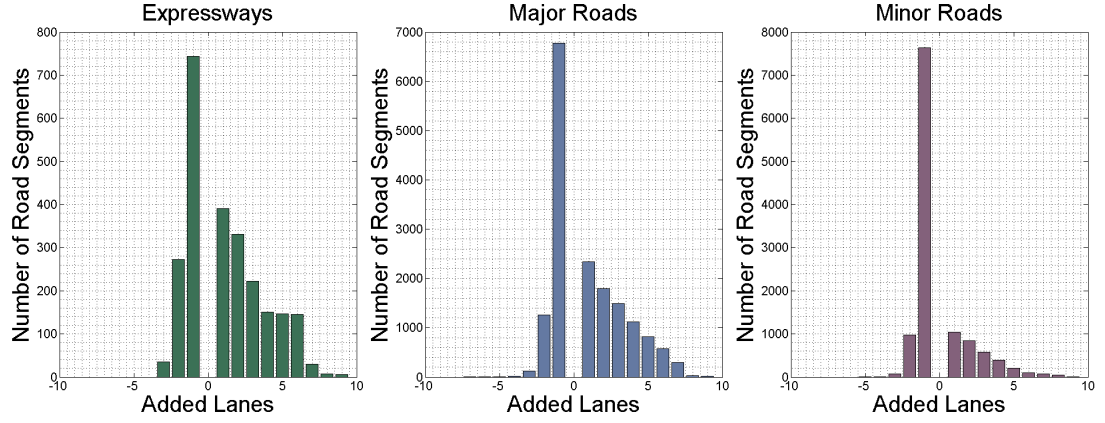


Figure 3.16: Distribution of the lane differences in various types of roads. From left to right: Expressways, Major roads, Minor Roads. All segments with no lane changes recommended have been excluded for better visibility of the distributions.

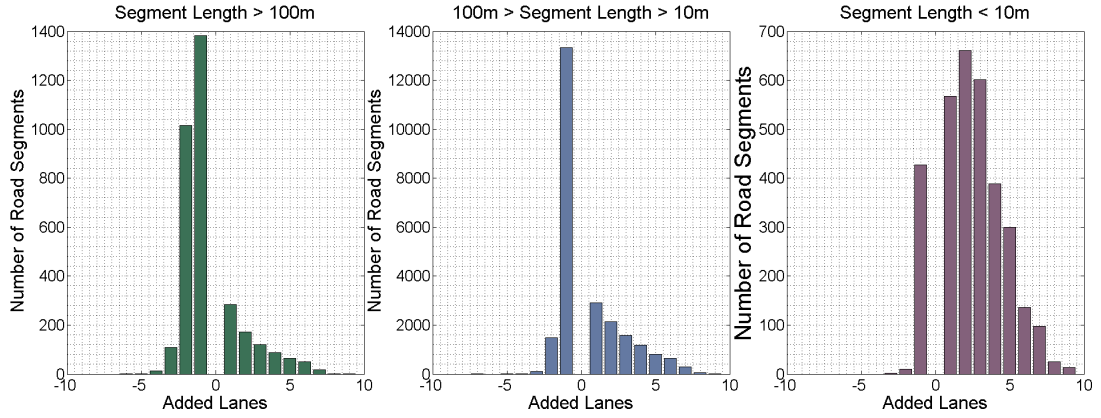


Figure 3.17: Distribution of the lane differences for roads with various lengths. From left to right: Long segments $l > 100m$, Medium segments $100m > l > 10m$, short segments $l < 10m$. All segments with no lane changes recommended have been excluded for better visibility of the distributions.

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

This metric represents the percentage of road length that needs to be redistributed. It is interesting to note that the flows, which are considered constant for the optimization problem are actually the shortest time flows computed with free flow velocities. Another approach of computing the flows would be to compute the user equilibrium solution, which represents the state of the system when all drivers have perfect information of current traffic conditions and no user would prefer to change his path. There is also a third possible way the routing can be performed, which is referred to as the system optimum solution, when the paths of all commuters are chosen so that the total travel time on the network is minimized. In summary, there are three possible routings that can be performed: routing according to free flow traverse times, user equilibrium where no agent would change his route, and system optimum which minimizes total travel time of the system. From system's perspective one would obviously choose the last option, however, it might result in unfair routes for some drivers. Ideally, the network should be designed in such a way that the fastest paths (routing option 1) coincide with the system optimum and with the user equilibrium. This is the reason why, the shortest paths were taken in order to define the flows in the system in this optimization problem.

3.7 Chapter Summary

3.7.1 Traffic Demand-Infrastructure Mismatch Analysis

Chapter 3 points out to the fact that traffic demand and road infrastructure are mismatched as a result of the faster pace of change of the demand and confirms the hypothesis, by locating severely mismatched locations in the studied traffic network. A set of measures is defined, based on turning probabilities acquired from the traffic assignment methodology application, which are used to identify the most mismatched intersections. Furthermore, measure of the total mismatch of the network are defined together with normalized version of the initially suggested metric using the concept of congestion factors, in order to enable the comparison of degree of mismatch between different cities.

It has been observed that the distribution of the deviation between infrastructure and traffic demand has a fat tail, meaning that there are locations, which exhibit strong disagreement between demand and infrastructure. The absolute mismatch measure that has been defined points out to an intersection, where 5 lanes have to be redistributed in order for it to match the demand in an optimal way. The measure of overall deviation in time, demonstrates that

the total city deviation has small variation coefficient and is therefore concluded to be time-invariant. Examining the location with the highest deviation value, however, shows that the deviation value drops during rush hours, which is expected since the intersection has possibly been optimized for such traffic conditions.

The spatial distribution of the intersections with high degrees of mismatch from the traffic demand is also examined. It is observed that the connections to and from regions with high degree of dynamics, such as business districts and fast growing sub-cities are experiencing the highest levels of mismatch between infrastructure capacity and traffic demand. This further strengthens the hypothesis that those deviations occur in cases where the traffic demands of the population change faster than the road topology can adapt.

3.7.2 Dynamic Locations

As a second step of evaluating the close fit that must exist in a transportation system between demand and infrastructure, a problem, which cannot be resolved in a simple infrastructure construction manner has been pointed out. The possible existence of intersections, which exhibit strongly dynamic behaviour throughout the day is hypothesised and demonstrated for the examined city using the dynamic factor measure, which is also defined in the chapter. The qualitatively varying congestion problems at such locations cannot be resolved, with a change in the infrastructure since they occur on a daily basis, level on which, the road infrastructure is strictly static.

Furthermore, the dynamic factor measure is also defined for the whole city in order to enable the comparison between different systems. It has been shown that there are highly dynamic intersections residing in the fat tail of the dynamic factor distribution, which is similar to the case with mismatched intersections described in the first part of the chapter. It has also been shown that the highest spatial concentration of highly dynamic intersections is in the business district area, which is most likely due to the high volatility of people movements in the area.

One additional practical application of the dynamic factor measure would be as a quantifier of the need for smart traffic regulation on intersections. It points out to the locations, which exhibit very abrupt and strong changes of qualitative demand, and furthermore quantifies how dynamic their behaviour is, thus enabling prioritization of such locations for traffic officials. Furthermore, the reduction of the overall dynamic measure of a city can be used as a heuristic for the transportation structural optimality.

3. IDENTIFICATION OF MISMATCHED SENSITIVE ROAD NETWORK LOCATIONS AND EFFECT OF OPTIMAL INFRASTRUCTURE DESIGN ON TRAFFIC PERFORMANCE

3.7.3 Fixing the Road Network

The second part of Chapter 3 implements the suggested changes to the mismatched locations identified by the designed measures and evaluates their effects. The results demonstrate that “fixing” those locations can benefit the traffic conditions in the city especially during rush hour periods. Furthermore, it has been shown that the main contribution to the increased quality of traffic conditions comes from a very small number of locations. In other words, there are places in the road network, which can greatly reduce the total travel time of the whole population if small alterations are applied to them.

Those locations are referred to as *super-sensitive* and a correlation analysis study is performed in order to find, which of the defined metrics can be used with the highest degree of success in order to locate such *super-sensitive* places, which present great interest of transportation officials as they require a more intelligent management approach. It has been shown that counter intuitively, the the amount of flow through an intersection is not correlated to its criticality. The metric with the highest correlation to the sensitivity metric is the normalized deviation measure exploiting the congestion factor of the examined location.

3.7.4 Optimal Lane Distribution Problem

Finally, a more general problem has been addressed and solved, namely the optimal distribution of lanes on the road network under the condition that the total length of the road network in lane meters is kept constant. The optimization problem is solved using the internal point method and the results demonstrate that the optimal distribution of lanes can save 36% of the overall population travel time and that 13% of the road segments currently do not have an optimal number of lanes. It is further, observed that the road types, which “receive” lanes are the highways and short segments road segments, which are likely to create bottlenecks. A generic measure is therefore defined, which looks at the mathematical distance between the current lane distribution and the optimal one in order to evaluate the overall degree of capacity distribution inefficiency of the network. A sensitivity analysis of the optimal lane distribution solution can be performed as a continuation of this line of research in order to evaluate how sensitive is the optimal solution to changes in the traffic demand. The next chapter will deal in more detail with the traffic assignment problem and specifically with the system optimum computation algorithms.

Chapter 4

Identification of Harmful Roads and Routing Control for Efficient System Optimum Traffic Assignment

4.1 Overview

The previous chapter has discussed the existence of mismatched, dynamic and *super-sensitive* locations in a complex transportation system and road network alterations in the sense of lane redistribution. This chapter will build on road alterations that might improve traffic conditions. This is done by examining the impact of the more severe measure of removing road segments and evaluating the performance of the changed road network. This allows for an extended sensitivity analysis of the system against road network alterations and also offers insights into traffic assignment problems. As every commuter chooses the most optimal route from his/her own perspective, traffic distribution on the road network becomes heterogeneous. This results in a small number of roads, which are largely overpopulated, while others remain underutilized [62]. In order to achieve a more homogeneous road utilization and thus reduce congestion levels, drivers should take more socially aware routes. As this is not observed in reality, a simple hypothetical situation is presented in this chapter, where roads are removed from the network as a way to force drivers into choosing more socially beneficial paths. It is demonstrated that the removal of certain road segments can redistribute traffic in a socially

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

beneficial way, leading to a decrease of total travel time of the population. This phenomenon has been known for decades as the Braess Paradox, first mentioned in [3]. This chapter, presents, the first systematic analysis on a large city level, of the existence of the paradox and quantifies the change in total travel time of the single closure of every road segment in the network. The removal of the most harmful segment, out of the 240,000 segments comprising the road network of the examined city, leads to a 4% decrease of total travel time. The first part of the chapter is based on the author's contributions in [10].

As the Braess paradox has been studied for a long time by transportation researches, a natural question is, whether it can be detected by using some heuristic function instead of testing the removal of every single road. This chapter demonstrates that road segments, which should be removed from the network exhibit lower level of utilization for system optimum traffic assignment when compared against the utilization using the standard selfish routing approach. Therefore, using such a heuristic one can filter out most of the links in the network, which are not likely to be harmful. As the objective of this work is to minimize overall travel time in the system, removing a road in order to reduce overall travel time can be considered a tool in achieving this. It is, however, in no way a complete solution. Furthermore, it is rather impractical to close an existing road for the whole commuting population. This chapter also presents the Backwards Incremental System Optimum Search (BISOS) also described in the author's contribution [63], a very practical approach to achieving system optimum traffic assignment by incrementally closing roads only for a chosen set of agents.

The described approach redistributes the traffic homogeneously among the city and converges much faster than any other method for system optimum computation in literature. Although it does not guarantee convergence to the exact optimum solution it gets significantly close to it for all practical purposes. Furthermore, as previous methods have been developed for theoretical purposes, their final solutions need not be practically feasible or do not provide explicit paths for the population. In contrast, the BISOS algorithm works only with feasible solutions and preserves the information about the exact paths of all commuters, throughout the whole process of computing the system optimum. Furthermore, in the context of this thesis, the results from the chapter point clearly towards routing as the best tool to be used in order to improve system performance. A decrease of 70% in total population travel time is achieved by employing the BISOS routing method, which is higher in effect than the infrastructure alterations studied in the previous chapter. As system optimum routing solutions increase the total distance traveled by the population, the fuel consumption is modelled as well in order

to check if the time saved has a price in terms of higher commuting cost. The results clearly demonstrate that on top of saving a considerable amount of time, fuel consumption also drops as a result of optimal routing, due to the decrease of congestion levels in the city, thus making centralized system optimum routing systems a strong candidate for ensuring the sustainability of future transportation systems.

4.2 Introduction and Existing Literature

4.2.1 Motivation

Technological advancement and growing demand for optimal functioning mechanisms imposes high standards of efficiency on both existing and future systems. The most studied such systems are usually large and depend to a great extent on the human factor. This increases their complexity and thus decreases the ability to model them effectively and improve their performance. The presence of people in a system may introduce a disorganized manner of operation, which can lead to induced flaws that can, in theory, be fixed by a centralised control system. It is also crucial that inefficiencies at such a level are resolved with minimal amount of actions thus minimising the probability of spawning further problems. The first part of this chapter can be considered on an abstract level to locate butterfly effect events [64], in which small changes to initial conditions can lead to performance changes that are much bigger in magnitude [65] and use it as an efficient steering tool, thus in a way exploiting the complexity of the system.

Transportation systems as a type of social complex system have a sparse and heterogeneous structure, which makes them harder to steer into an optimal operating state, compared to homogeneous and dense systems [66]. Numerous control techniques have been described in literature, which achieve increase in traffic performance. These include self-organizing traffic lights [67, 68, 69], or information dissemination techniques as [70, 71, 72, 73, 74, 75, 76], where traffic participants receive real time information about congestion in the network and adapt their routes accordingly.

The rapid advancement of Intelligent Transportation Systems [77] is enabled by broader distribution of personal smart devices, which provide higher data availability and thus a more complete view of the network, which leads to faster coordination [78]. As, nowadays, drivers use as support the advice provided by their navigation tool [79] they can be used to implement more efficient and robust traffic control strategies. The potential for intensive interaction between

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

a commuting system and an ITS, which exists in present days can be used to steer traffic participants in a beneficial way.

4.2.2 Price of Anarchy

Often societies are governed by non-coordinated actions performed by individuals, which aim at optimizing their own state. The evolution of systems exhibiting such dynamics is studied in [80]. Even though, every participant in such a system follows a self-defined optimal strategy, the collective behaviour of the group is often suboptimal. In the context of transportation systems traffic assignment, the ratio between the population travel time computed for the user equilibrium state, where every commuter chooses an optimal path, and the system optimum, which minimizes total travel time is called the “price of anarchy” (POA) and is indicative of the inefficiency of the system due to decentralization. The user equilibrium traffic assignment is referred to as “selfish routing” and studied in [81], however, such behaviour also exists in other complex networks such as the Internet [82].

Furthermore, measuring and reducing the price of anarchy has been the subject of numerous studies such as [81] and [83]. In [84] a useful general theory is developed for bounding the price of anarchy. A middle ground between centrally enforced solutions and completely unregulated anarchy is sought after in order to achieve stability in [85].

4.2.3 User Equilibrium and System Optimum Traffic Assignment

At this point it is important to recollect the difference between system optimum and user equilibrium states of the system. As described first in [86], user equilibrium (UE) occurs when all agents have perfect information about the traffic situation in the system and distribute on it so that no agent would be willing to change his/her path. This state of the system can be perceived as an analogy of the Nash equilibrium also known as the Wardrop’s equilibrium, which for a long time has been used in order to calculate the expected distribution of traffic in transportation networks. In [87] various modelling techniques for user equilibrium are reviewed including individual choice theory, interacting choice theory, effects on travel information affecting individual choice and interacting behaviour.

On the other hand, system optimum (SO) aims at minimizing the total travel time of the system defined in [17]. Although subtle, there is a difference between the two formulations. In the case of UE all alternative paths with flows on them have the same length in order to ensure that no one has incentive to switch their route. This constraint does not exist in the SO

formulation, which is concerned solely with everyone arriving at the destination and minimizing the total travel time of the population. The ideal case would be that the UE and SO solutions coincide. In other words, the commuters are “steered” into taking the most optimal routes by different types of incentives.

The difference between the performance of UE and SO has been thoroughly studied and evaluated in [88], where the total travel time at user equilibrium is bounded from above as twice the traffic routed in an optimal way. In [89] an upper bound is given to the inefficiency of stochastic user equilibrium (SUE). The price of anarchy was coined to characterize this inefficiency in [90].

Furthermore, efforts have also been made for the design of networks where the UE and SO coincide. In [91] efficient methods for selfish network design are examined in the case of linear latencies and specific network topologies in polynomial times. In [92] it is shown that bases of matroids are maximal structures in which Braes paradox does not occur.

In literature, the user equilibrium is more thoroughly studied since it is easier to compute and more realistic [93, 94], however, the need for a system-wide view of performance has long been recognised [95, 96, 97]. With the advancement in technology, namely GPS devices, increased computational capabilities, and the rise of autonomous vehicles, it is natural that researchers should be looking in the direction of system optimum computation since it does not seem so unrealistic anymore.

4.2.4 Braess’ Paradox

The first part of this chapter deals with evaluating the sensitivity of traffic conditions to changes in the road network. Similar to [76] and [71], which use information dissemination, an alteration of the road network, closing a single road segment in this case, affects the route choices of the population. It is shown that if the right road segment is closed, the commuters are generically steered towards choosing more socially optimal routes.

Although unconventional, removing a road from the traffic infrastructure may lead to improved commuting conditions. The Braess paradox first mentioned in 1968 [98, 99], states that adding extra capacity to a network where drivers act selfishly, can, in some cases, decrease performance. A generalisation of this paradox [100] states that removing edges for large networks can produce an arbitrarily large improvement. It was further shown that the paradox can exist in all varieties of LOS networks as well [101]. Even the development of the human brain has a

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

mechanism called synaptic pruning during which synapses (connections between neurons) are being removed in order to achieve more optimal learning [102].

There are numerous studies in real life cities that confirm the existence of the Braes paradox as in Stuttgart [103] and New York [104], where streets were closed for renovation or on purpose and better traffic conditions were observed. There are 70 more case studies from 11 countries that examine such conditions summarized in [105, 106].

Due to its peculiar nature and the fact that it points out apparent problems of the topology of the network, the Braess paradox has been studied intensively since its discovery. The paradox is found in semiconductor networks in [107]. The work in [108] defines and studies the paradox in networks with pricing and shows that under monopoly prices the paradox does not occur. Attempts for a detection methodology for the paradox can be found in [106, 109, 110]. In [111] it is shown, however, that the construction of Braess paradox free networks is NP hard; it is also stated that the paradox cannot be detected efficiently.

4.2.5 Towards a Centralized Routing Control System

After presenting empirical proof for the existence of significant inefficiencies in the way drivers choose their paths, the second part of this chapter will provide a tool for minimizing such inefficiencies in the sense of a centralized routing system. One way of achieving such a task can be realized by using guidance systems, which are widely available in a large portion of the vehicles, and through suggested routes that “steer” drivers into socially optimal routes. In [112], the authors suggest reactive guidance that aims at using information of current conditions to calculate recommendation. Another type of guiding system has been suggested in [113], where an anticipatory system is utilized, which predicts future demands and gives recommendations with respect to that.

A more complex approach has been adopted more recently in [114], where a fair system for user equilibrium is designed. It is based on distributed multi-agent traffic model made of vehicles with routing guidance systems. Vehicles, which are assumed as selfish, negotiate over auction with OD agents and OD agents negotiate over auctions with intersection agents. The end result is a better performance than traditional user equilibrium conditions with small exceptions of certain OD pairs.

Another possible way to actively control routing choices is by using financial incentives. Toll taxes have been studied for a long time; research efforts have intensified in the recent decades due to the emergence of more efficient charge collecting and detection tracking systems. In

[115] a differentiated congestion pricing strategy is used, which calculates charges according to travel characteristics and attributes. An incentive program is further designed to mitigate the privacy concerns since the method minimizes the revenue from the tolls. The morning commute problem described by [116] is extended in [117] for more modes of transportation. Time dependent prices exist to achieve system optimum using wish curves that identify the departures curves for various modes of transportation.

In [118] the formulation of the cell transmission model (CTM) is extended to allow more general non-linear flow time relationship. It is shown that under the proposed pricing scheme the user equilibrium coincides with the system optimum under dynamic conditions. Neither CTM [119] nor FDA (Finite Difference Approximation) [120] based models, however, nor their solution include any explicit expressions or variables for the link travel times experienced by users, which makes traffic analysis more challenging. In [121] a price design mechanism is studied that is based on quadratic structure, which aims at minimizing the price of anarchy.

A more practical approach is taken in [122] where network optimal tolls are computed with constraints on the available edges. Due to the usual lack of complete data for the system, the strategy in [123] is suggested to optimize road pricing with unknown demand and cost functions based on a trial and error approach adjusting the link toll charges.

Finally, a third approach might be the method of [10] that will be suggested in the first part of the chapter, where information about very congested areas can be supplied to portions of the commuters with the goal of avoiding them. In this way, the targeted system state can be achieved. This approach, assuming its correct implementation, will guide the system into an optimal state. The partial provision of information to the society instead of full knowledge (assumed in the case of user equilibrium) is the root of the different traffic performance between minimal commuting time and Nash equilibrium with full information. This approach can be considered as “fooling” the commuting population, or at least, part of it, however, an analysis of such implications are beyond the scope of this work. It must be noted, that an autonomous vehicle fleet scenario will not have to deal with such hurdles as the route choice is seamlessly taken away from the human passenger.

Finally, it must be noted that using a pricing system that charges people traveling on certain roads, performs precisely the same function as the method described in [10] at an abstract level. The only difference is that, one is based on limiting traffic based on financially imposing restrictions and the other is based on a well designed information dissemination technique.

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

Regardless, of the means, the second part of this chapter aims at solving the more general problem of the optimal traffic distribution.

The contributions in this chapter can be summarized as:

- First systematic analysis of the impact of road removal of every road segment in a realistic large city scenario.
- Identification and quantification of all harmful road segments in a city.
- Design of a heuristic for Braess paradox detection based on system optimum solution.
- Numerical experiment demonstrating the invariance of total population travel time to a constant size set choice of drivers to be re-routed from a certain road.
- Design and implementation of system optimum search algorithm, BISOS, with explicit paths and integer valued flow solutions, converging more than one order of magnitude faster than current methods.
- Systematic analysis of the BISOS algorithm parameters.
- Evaluation of system optimum routing strategies for increasing population size.
- Fuel consumption model, evaluating the fairness and cost of system optimum routing solutions.

4.3 Systematic Road Removal

The first set of studies that will be presented deal with evaluating the effects of a single road segment closure on the system's performance. A systematic approach is taken for the identification of road segments whose closure would result in improved traffic performance. It consists of examining all 240,000 links one by one and removing them from the routing graph. For every link removal a separate traffic assignment run is performed and the routes and travel times of the population are recalculated according to the new road network. The results are compared to the initially simulated scenario in accordance with the model described in Chapter 2, while the origins and destinations of all drivers are kept the same for all simulation runs.

The procedural sequences of actions for the systematic road closures are formalized in a step-by-step manner below in Algorithm 2. It describes the process of closing every link in the

Data:

| | |
|----------------------|--|
| G | Road network graph consisting of nodes and links |
| A | Set of all agents in the population |
| L | Set of all links in G |
| $ComputeRoutes$ | Set of agents \times Graph \rightarrow Set of routes |
| $ComputeTravelTimes$ | Graph \times Routes \rightarrow Set of travel times |
| $RemoveLink$ | Link \times Graph \rightarrow Graph |

Result: Set of average population travel times for respective link closures at 100% - t^{100}

```

// Compute routes  $R^A$  and travel times  $T^A$ 
 $R^A \leftarrow ComputeRoutes(A, G)$ 
 $T^A \leftarrow ComputeTravelTimes(G, R^A)$ 
 $t^0 \leftarrow mean(T^A)$ 
 $t^{100} \leftarrow t^{100} \cup t^0$ 

foreach  $l \in L$  do
     $G^l \leftarrow RemoveLink(l, G)$  // Remove link  $l$  from the road network
     $A^l \leftarrow \forall a \in A : l \subseteq R^a$  // Identify agents that pass through link  $l$ 
    // Re-calculate routes of affected agents and population travel times
     $R^{A^l} \leftarrow ComputeRoutes(A^l, G^l)$ 
     $R_l \leftarrow R^{A^l} \cup R^{A \setminus A^l}$ 
     $T^A \leftarrow ComputeTravelTimes(G, R_l)$ 
     $t^l \leftarrow mean(T^A)$ 
     $t^{100} \leftarrow t^{100} \cup t^l$  // Store the computed population average travel time
end

```

Algorithm 2: Quantifying population travel time change for partial and full closure of links

road network individually and re-simulating the traffic conditions, in terms of computing the new traverse times for every link based on the new flows.

The results of the experiment confirm the existence of the Braess paradox in the examined network. Furthermore the harmful effect of every link in terms of time saved as a result of its closure can be quantified. In 21 cases the closure of a link in the network leads to a decrease of one minute or more in the average travel time, which corresponds to 3.73% overall system performance increase. The most harmful link brings a 74.25 second decrease in overall trip duration translating to 6400 saved hours for the driver population on a daily basis, solely from the morning rush hour period. Although the existence of harmful roads has been predicted by Braess, it is counter-intuitive that some of the most inefficient road segments are major roads.

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

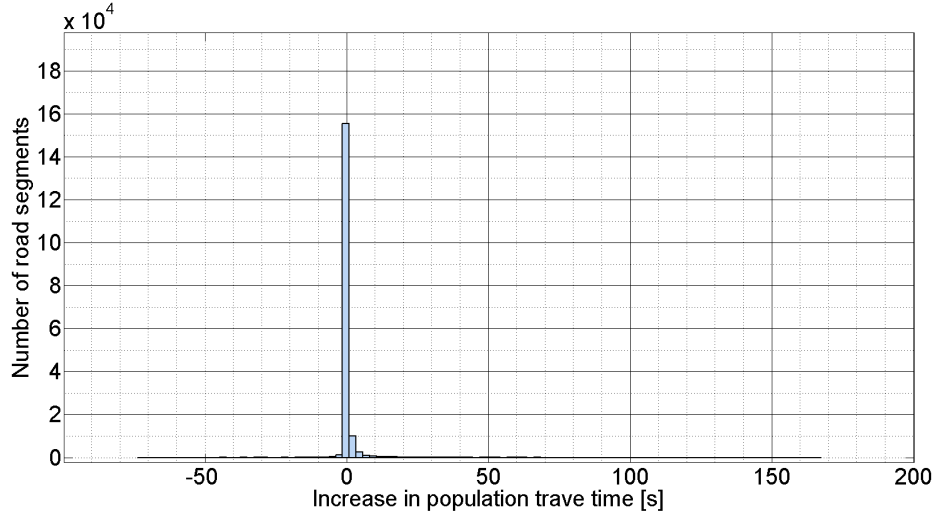
Although part of the backbone of the network in a topological sense, the removal of certain major road segments, would decrease overall travel time.

Fig. 4.1a shows that the full removal of the majority of road segments in the network would have almost no effect on the average travel time of the population. The removal of a single road segment does not present a drastic change to the network, therefore it is expected that the effect of such a removal is minuscule. The reader is, however, probably more interested in the counterexamples of this, which can be observed in more detail by removing the unaffected links from the results set. This leaves out only the road segments that have a more significant effect on traffic (illustrated on Fig. 4.1b). This distribution can be helpful not just for identifying harmful links (segments on the far left side of the distribution) but also to locate roads, which are crucial for the road network (segments on the far right side of the distribution). Such segments are the weak spots of the infrastructure in the sense that their incapacitation would result in significant decrease of traffic performance.

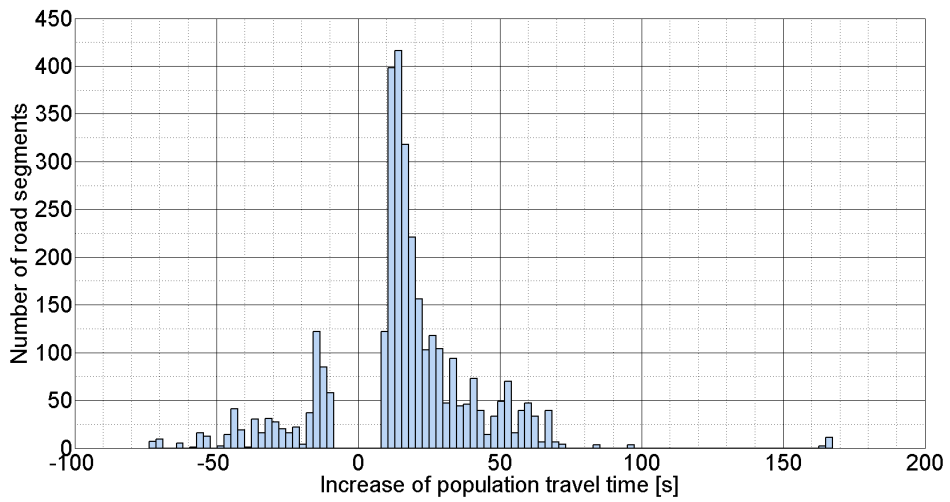
It seems that in most cases traffic that has to be reassigned has to go through a smaller road than the initial one and congests the system even further. This does not happen in only 639 out of the 240,000 examined cases, corresponding to a 99.73% probability of worse traffic conditions arising from a road closure. Intuitively, after considering the complexity of a large city road network and the hundreds of thousands of vehicles on the roads, decreasing connectivity of the road infrastructure will rather increase congestion than relieve it. As noted in the results, the likelihood of observing commuters being forced into more socially beneficial paths given a road closure is indeed small, but, as demonstrated, not impossible. It seems that in such cases the additional bottlenecks, if any, introduced by the road network alteration are less harmful than the traffic stress that is relieved from the initial state of the system.

One can consider two main reasons for the occurrence of the observed phenomenon. The first one is selfish routing whose effects can be diminished if a certain road is closed leaving the drivers no alternatives other than choosing a more socially optimal path. The second reason is the nature of transportation networks evolution, which adapt in an incremental manner in time to the changes in traffic demand. As a natural consequence of that as new roads are added, rarely old ones are removed. And although, evolutionary processes undeniably offer working solutions, their optimality is not guaranteed. It is, therefore, not improbable that some of the old roads become obsolete or even harmful.

A detailed view at the implications of closing a harmful road, which depicts the flow differences before and after the removal of the road can be seen in Fig. 4.2. The closure of the



(a)



(b)

Figure 4.1: Results of Study 1 and 2 summarized in histograms. Fig. 4.1a shows the distribution of links according to their effect on the average population travel time. Fig. 4.1b shows the same distribution, however all links that have an effect smaller than 10 seconds in magnitude are excluded from the distribution.

examined segment can be observed to reduce the amount of traffic on the roads in its proximity since many existing paths become non-viable. The collective length of roads receiving traffic is far greater than that of the roads that experience reduced flow volumes. Since no vehicles are lost and their number is conserved, the vehicles that are taken from the closed road and the

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

ones in its vicinity are spread over a larger portion of the city thus increasing the homogeneity of traffic flows. A reduction in average travel time is observed and it can be concluded that the streets, which receive traffic produce less additional time (have a smaller marginal cost) than what is gained by the streets that give away traffic. Due to the non-linear nature of the traverse time vs. flow relationship taking an agent from a severely congested street and putting it on a less congested one, reduces overall trip duration. Therefore, as a result of the road closure traffic from a group of highly congested roads is distributed along less populated parts of the network thus relieving traffic conditions.



Figure 4.2: A map representing the changes that occur in traffic due to the closing of the road indicated in blue. Green and red bars represent reduced and increased traffic respectively. The height of the bars illustrates the magnitude of the change.

4.3.1 Heuristic Design for Braess Paradox Detection

An attempt has been made in this work, to design a heuristic, which would speed up the process of identifying Braess paradoxes in a network. Instead of removing all road segments one by one from the network and recalculating the routes of the affected agents, it might be more beneficial to compute the system optimum solution and use it as guidance towards which segments to

remove first in search of the paradox. The logic behind this approach is the following: since the system optimum flows minimize travel time, it is implied that if there is a harmful road on the network, no user will be going through it in the system optimum solution. Therefore, by identifying roads that experienced heavy traffic in the initial traffic assignment, but have much smaller system optimum flows, a Braess paradox road can be located. In other words, when the system optimum solution shows that a road should not be used and there is heavy utilization of it in the initial traffic assignment, there is a high probability that this road is harmful and its removal would lead to improvement of traffic conditions.

Numerous heuristics based on the difference between system optimum solution and initial assignment were designed in search of an optimal approach. The simplest one, however, proved to be the most successful. The heuristic h_i for segment i is simply the difference between the initial flows and the system optimum flow on the respective road segment: $h_i = f_i^0 - f_i^{so}$. Ideally, the segment with highest value of h should also lead to the biggest improvement in total travel time when removed from the network. Assuming b_i is the time saved by the population when link i is removed, then for the best possible heuristic there must be perfect correlation between \mathbf{b} and \mathbf{h} , or $\rho_{\mathbf{b}\mathbf{h}} = 1$. In the case of the defined heuristic the correlation coefficient is equal to 0.6, which, provided the complexity of detecting Braess roads, is a solid result.

In order to practically evaluate how helpful such a heuristic would be, let us visualize how the already performed systematic experiment would have gone, if the heuristic was used instead of trying every single road in the city. Fig. 4.3 depicts the number of roads that should be explored using the heuristic values in order to identify the first 100 most harmful roads in the system. The graph can be understood by interpreting the value on the y axis as the number of roads that need to be examined using the heuristic in order to find the first n most harmful roads, where n is the respective value on the x axis. In summary it can be observed that the most harmful road is also the one, which has the highest value for the heuristic h . The 10 most harmful roads can be identified by examining 550 cases and the first 100 most harmful road can be identified within 1650 road segment examinations. Depending on the intrinsic goals of an experiment, one might aim at just identifying the most harmful road, removing it and then identifying the next most harmful segment in an incremental way. Another option would be to identify a group of potentially harmful roads and then apply an optimization algorithm to find the optimal combination of links to be removed. In any case, the suggested heuristic brings a significant decrease in the number of road removal evaluations required, by reducing this number from 240,000 to 1650, which is almost 150 times less.

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

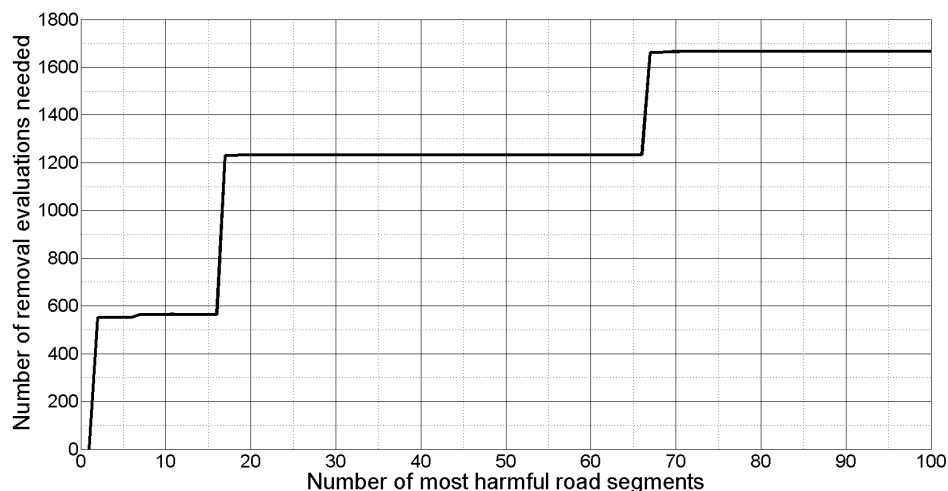


Figure 4.3: Number of road segment removals to be evaluated in order to find the first n most harmful road segments.

4.3.2 Invariance of Total Travel Time to Choice of Agents to be Rerouted

As the general objective of this work is to maximize utilization of the road network, it may be contradictory and inefficient to completely remove parts of it. Greater social benefits can be achieved with milder measures than complete removal. As the initial traffic generation chooses optimal trips assuming free flow conditions, the path that drivers are forced to reroute to will be sub-optimal for them. Therefore, it might be more appropriate to not fully dispose of already existing infrastructure, which can help avoid cases of extreme detours.

An additional study performed in [10], the most harmful roads are chosen and the degree of closure is varied, in the sense that a certain percentage of agents are allowed to still pass through while the rest were rejected access. The results of this study demonstrate that each link has an optimal percentage of the volume of vehicles that need to pass through it, which minimizes the total travel time of the population. The percentages of closure for every link can restrict the access to the point which coincides with the system optimum flow. The experiment also produced some other significant results. Since the set of agents that should be denied access to the link can be chosen at random, several runs for the same percentage are performed. The coefficient of variation for the results of those runs is unexpectedly small.

Four of the most harmful roads were selected to be examined. The group of agents to be

re-routed is sampled from all drivers that initially need to use the examined road segment. In order to measure the sensitivity of the results against the sampling of agents to be rerouted 10 runs are performed for each percentage of closure, where every time a different set of agents is redirected. It is intuitive to expect variations in the results since every driver has a distinct origin and destination and would affect the system differently after re-routing. The computed coefficient of variation σ / μ on those experiments is 4×10^{-4} . The measured coefficients of variation $\frac{\sigma}{\mu}$ is recorded in Table 4.1.

The small coefficient of variation suggests that the choice of agents, which need to find other routes is not a decisive factor. This finding can be used to simplify significantly an optimal routing solution computation. This is because the problem of finding the optimal set of drivers to reroute, which is considered NP hard, can be avoided. This finding may be explained with the fact that in a real world scenario there is a great variety of origin destination pairs. The apparent homogeneity of agents at this level of abstraction enables a simpler and faster solution to the system optimum problem which will be presented later in the chapter.

4.3.3 Spatial Distribution of Effects of Full Road Closures

Fig. 4.4 illustrates a spatial perspective of the performed experiment. The segments, which present a significant change to the overall travel time if closed (positive and negative) are coloured according to the magnitude of their effect. As already noted from the distribution in Fig. 4.1, most of the segments that lead to a significant change in population travel time, if closed, would have a negative impact on the system. Although, in smaller quantities, the harmful road segments seem to cover some of the backbone roads of the city. Topologically speaking, those roads are considered rather central and thus important for the system, however, according to the simulated results their removal can reduce congestion levels. Furthermore, there are regions of the city where an alternation on a single road of beneficial and harmful segments can be observed. One of those sensitive regions is shown in more detail on Fig. 4.5.

4.3.4 Equilibria Analysis

The gathered data from the experiment allows for the analysis of the way optimality is perceived at different hierarchical levels of the driver population. Consider a single road is closed and three sets of drivers. The first set is all drivers in the population, the second one is all drivers that pass through the segment and the third set is composed of all affected drivers, who, however, act selfishly so that user equilibrium is achieved. In order to demonstrate how different degrees

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

| Road ID | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% |
|---------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 1 | 4.3×10^{-4} | 4.5×10^{-4} | 3.1×10^{-4} | 3.3×10^{-4} | 0.0017 | 6.2×10^{-4} | 3.4×10^{-4} | 3.4×10^{-4} | 4.2×10^{-4} |
| 2 | 1.9×10^{-6} | 3×10^{-6} | 5.3×10^{-6} | 5.8×10^{-6} | 4.6×10^{-6} | 3×10^{-6} | 3×10^{-6} | 2.6×10^{-6} | 3×10^{-6} |
| 3 | 1.5×10^{-4} | 2.5×10^{-4} | 3.4×10^{-4} | 2.5×10^{-4} | 2.4×10^{-4} | 4.3×10^{-4} | 3.6×10^{-4} | 2.9×10^{-4} | 1.9×10^{-4} |
| 4 | 5.8×10^{-4} | 3.9×10^{-4} | 3×10^{-4} | 6.8×10^{-4} | 2.7×10^{-4} | 3.4×10^{-4} | 3.7×10^{-4} | 0.0016 | 1.4×10^{-4} |

Table 4.1: Variation coefficients for chosen links for different percentages of closure



Figure 4.4: Comparison of road closure effects. Road segments coloured in the red gamma represent increase in average population travel time. The thicker and redder road segments are represented the higher the increase of average population travel time. Road segments coloured in the blue gamma analogically represent a decrease of average population travel time. The closure of roads that are not coloured would result in a change of average travel time of less than 10 seconds.

of socially aware behaviour can affect traffic conditions, let us compute the optimum percentage of closure for the road segment in question for each of the three sets. The first set would choose the percentage, which minimizes the total travel time of the population. The second group will choose the percentage, which minimizes the total travel time of the group of people, which passes through the road segment. The set of selfish drivers would choose the Nash equilibrium point percentage where the travel time on or off the closed road would be the same. The calculated values of the optimal percentage of closing in the 3 cases can be found in Table 4.2. Four road segments have been examined and in 3 out of the 4 cases the optimal percentages of closure for the three groups are all different.

The above experiment demonstrates clearly the concept of Price of Anarchy and also includes a middle level of centralization in the sense of a group optimum (the second examined set). It must be noted, the difference in chosen optimal percentages of closing are not due to lack of information. All three groups are assumed to have perfect information about the road conditions. The reason for this mismatch is simply the different priorities that each of the

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT



Figure 4.5: Zoomed version of Fig. 4.4 in order to observe in more detail a sensitive area.

| Road | Social Optimum [%] | Nash Equilibrium [%] | Affected Group Optimum [%] |
|--------|--------------------|----------------------|----------------------------|
| Road 1 | 100 | 70 | 90 |
| Road 2 | 90 | 90 | 90 |
| Road 3 | 40 | 27 | 50 |
| Road 4 | 50 | 44 | 60 |

Table 4.2: Points of Equilibrium

examined sets has. It can be, therefore, concluded that it is vital that the system is always considered as a whole because the collection of local optimal solutions may not produce the expected result due to uncoordinated behaviour of the users.

4.4 Socially Optimal Routing

The results from the previous part of this chapter strongly suggest the existence of inefficiencies in the way people choose their routes. This is due to the lack of a centralized control system, which first can coordinate the population and second, is aware of the routing choice every driver is assigned. The main task of such a system would be to control the third traffic determining

factor, namely, the routing choices. Demands change in the long term faster than the network can cope, as evident from the mismatch study in Chapter 3. Furthermore, the demands change on a daily basis as observed from the existence of highly dynamic intersections (also shown in Chapter 3). The road network is static in the short term and therefore cannot adapt to such variations and react accordingly. The study done in the first part of the chapter, however, can be extended in order to enable the static infrastructure to change dynamically. Removing roads from the network, or making them very hard to access for only a certain group of drivers, in a sense allows the road infrastructure to choose the way it looks like in front of each individual user, thus making it appear to be dynamically changing perceived from the commuter's side.

This concept will be applied in order to solve efficiently the problem of system optimum routing. The problem consists of finding the paths for all drivers in a city, which minimize the total travel time of the population. By allowing the network to alter its appearance for the different drivers, a much faster approach for computing optimal routes can be achieved. The finding from the previous part of the chapter regarding the very small coefficient of variation of total travel time with respect to the subgroup of vehicles that have to avoid a certain road segment, allows the circumvention of the NP hard problem of choosing, which vehicles to reroute. In this way, using the BISOS algorithm that will be suggested later in this section, the computation of system optimum becomes more feasible than the standard Frank Wolfe algorithm [124] variations typically used until now, bringing it it closer to being utilized in real life, real time applications.

On a side note, interest in the traffic assignment field of research has been mainly concentrated on calculating user equilibrium solutions as this has been considered to be the only feasible situation on the road. The system optimum, historically, has been computed just as a theoretical minimum for the travel time of the population but never considered possible in real life conditions. Hence researchers have spent more time developing specialised algorithms for user equilibrium and rather neglected the computation of system optimum. Due to the rising thrust in guidance systems and the development of autonomous vehicles, research into fast and efficient system optimum algorithms has become a significantly relevant task.

Furthermore, all existing methods have rather weak constraints for the system optimum problem since it has been only used for theoretical purposes. There are two main additional constraints that the algorithm in this thesis has taken into account. If a centralized control system is developed it needs to supply every driver that wants to get from a predefined origin to his/her destination an explicit path, therefore the path for every agent on the road needs to be

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

known at all times of the computation of the optimal solution. The majority of existing work, provides, as final solution, the optimal flows on every road segment, which although proven to be feasible, are impractical in terms of assigning specific routes to drivers. In fact, it is NP hard to compute feasible routes from a set of feasible flows.

Second, there is no integer constraint on the number of vehicles per road segment. The flows based algorithms, therefore, are allowed to split drivers between as many roads as they want thus making the solution impractical. Furthermore, the path based algorithms that assign optimal path flows for every origin destination pair also do not have the constraint that the number of people on each path has to be an integer. The addition of those two constraints, namely, stored explicit path for every driver and integer number of vehicles passing through every segment, means that the already existing methods that rely on the convexity of the feasibility space do not present feasible solutions anymore. Next, the standard used algorithm for system optimum computation is presented and existing solutions to the system optimum problem are introduced.

4.4.1 Existing System Optimum Computation Algorithms

The work in [17] will be relied on in describing the system optimum problem and the user equilibrium (UE) algorithm described there has been re-tailored here for the system optimum problem. The classic formulation of the system optimum (SO) problem is:

$$\min_{\mathbf{F}} T(F_i) = \sum_i t_i(F_i) F_i \quad (4.1)$$

subject to

$$\sum_k p_k^{od} = q_{od} \quad \forall o, d \quad (4.2)$$

$$p_k^{od} \geq 0 \quad \forall k, o, d \quad (4.3)$$

where $t_i(w)$ is the BPR function for traverse time dependent on the flow F_i . The constraint 4.2 makes sure that the flow is conserved. p_k^{od} is the flow on path k between origin o and destination d and q_{od} is the number of vehicles that belong to this OD pair. Constraint 4.3 ensures that the flows on all possible paths are non-negative. The definitional constraints linking the path flows with the link flows are:

$$F(i) = \sum_o \sum_d \sum_k p_k^{od} \delta_{i,k}^{od} \quad \forall i \quad (4.4)$$

where $\delta_{i,k}^{od}$ is 1 if path k between o and d passes through i and 0 otherwise. Please note the lack of constraints that F_i is an integer. The necessary condition for a minimum are the first-order conditions for a stationary point of the Lagrangian:

$$L(\mathbf{p}, \lambda) = T(F_p) + \sum_{od} \lambda_{od} \left(q_{od} - \sum_k p_k^{od} \right) \quad (4.5)$$

The minimum of the Lagrangian with respect to p is constrained by:

$$p_k^{od} \geq 0 \quad \forall k, o, d \quad (4.6)$$

The variable λ_{od} is the vector of Lagrange multipliers associated with the flow conservation constraints for OD pair od . The first order conditions for a stationary point at the optimal solution are then:

$$p_k^{od} \frac{\partial L(\mathbf{p}, \lambda)}{\partial p_k^{od}} = 0 \quad \forall k, o, d \quad (4.7)$$

$$\frac{\partial L(\mathbf{p}, \lambda)}{\partial p_k^{od}} \geq 0 \quad \forall k, o, d \quad (4.8)$$

$$\frac{\partial L(\mathbf{p}, \lambda)}{\partial \lambda^{od}} = 0 \quad \forall o, d \quad (4.9)$$

$$p_k^{od} \geq 0 \quad \forall k, o, d \quad (4.10)$$

$$(4.11)$$

Constraints 4.10,4.11 simply restate the conservation flow and the non-negativity conditions. Conditions 4.8 and 4.9 consist mainly of the partial derivative of the Lagrangian, which can be expressed as:

$$\frac{\partial L(\mathbf{p}, \lambda)}{\partial p_k^{od}} = \frac{\partial}{\partial p_l^{mn}} T(F_p) + \frac{\partial}{\partial p_k^{od}} = \frac{\partial}{\partial p_l^{mn}} \sum_{od} \lambda_{od} \left(q_{od} - \sum_k p_k^{od} \right) \quad \forall m, n, l \quad (4.12)$$

The expression on the right hand is simplified to: $-\lambda_{mn} \quad \forall l, m, n$. By using the chain rule and Equation 4.4 on the left becomes:

$$\sum_i \delta_{i,l}^{mn} \left(t(F_i) + F_i \frac{dt(F_i)}{dF_i} \right) \quad \forall l, m, n \quad (4.13)$$

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

If $t(\hat{F}_i) = t(F_i) + F_i \frac{dt(F_i)}{dF_i} \quad \forall i$, the $t(\hat{F}(i))$ basically is the marginal time that an additional traveller on link i produces for the system as described in [125]. Therefore the partial derivative of the objective function can be rewritten as:

$$\frac{\partial}{\partial p_l^{mn}} T(F_p) = \sum_i \delta_{i,l}^{mn} t(\hat{F}_i) = c_l^{\hat{mn}} \quad \forall l, m, n \quad (4.14)$$

Where $c_l^{\hat{mn}}$ is the marginal total travel time on path l connecting OD pair mn . The first order conditions of the SO problem can now be rewritten as:

$$p_l^{mn} (c_l^{\hat{mn}} - \lambda_{mn}) = 0 \quad \forall l, m, n \quad (4.15)$$

$$c_l^{\hat{mn}} - \lambda_{mn} \geq 0 \quad \forall l, m, n \quad (4.16)$$

$$\sum_l p_l^{mn} = q_{mn} \quad \forall m, n \quad (4.17)$$

$$p_l^{mn} \geq 0 \quad \forall l, m, n \quad (4.18)$$

$$(4.19)$$

The resulting constraints simply state that the marginal total travel times on all used paths connecting an OD pair have to be equal. Please note that in the case of UE computation the total travel times on all used paths connecting an OD pair have to be equal. This theoretical reminder has been shown here in order to see the importance of the marginal costs on the links. Those costs are a manifestation of the non-linearity of the delay function. The algorithm that has been presented in the next section relies heavily on those marginal costs and calculates them in terms of a congestion factor in order to guide the optimization process.

SO problems are usually solved using the convex combination algorithm proposed by [124] or an improved version called Partan, which was suggested in [126] and discussed in [127, 128]. A brief explanation of the convex combination algorithm applied on the SO problem is provided next for completeness.

The approach of the convex combination algorithm can be intuitively explained as a linearization of the objective function at the current point and optimal computation of the step size that needs to be performed in the gradient minimizing direction. The linear program to be solved with respect to all feasible \mathbf{y} at every iteration of the algorithm is:

$$\min_{\mathbf{y}} T^n(\mathbf{y}) = \nabla T(\mathbf{F}^n) \cdot \mathbf{y}_{\top} = \sum_i \left(\frac{\partial t(F(i)^n)}{\partial F_i} F_i + t(F_i) \right) y_i \quad (4.20)$$

In this notation the upper index n is used to denote the iteration number. For example, $\nabla T(\mathbf{F}^n)$ is the gradient of the objective function at the current solution \mathbf{F} on iteration n . It can be noticed that the marginal cost appears once more in the equation as the coefficients of the minimization variable \mathbf{y} and therefore the problem can be rewritten as:

$$\min_{\mathbf{y}} T^n(\mathbf{y}) = \sum_i t(\hat{F}_i^n) y_i \quad (4.21)$$

subject to

$$\sum_k p_k^{od} = q_{od} \quad \forall o, d \quad (4.22)$$

$$p_k^{od} \geq 0 \quad \forall k, o, d \quad (4.23)$$

$$(4.24)$$

The solution of this linear problem is simply computing the shortest paths of all commuters by using as weights the marginal cost of every link, calculated at the previous iteration. The descent direction is therefore, $\mathbf{d}^n = \mathbf{y}^n - \mathbf{F}^n$. The flow solution of the next iteration is:

$$\mathbf{F}^{n+1} \rightarrow \mathbf{F}^n + \alpha(\mathbf{y}^n - \mathbf{F}^n) \quad (4.25)$$

The step size of the descent α is computed so that it minimizes the objective function in the given direction. The derivative of the objective function with respect to α is:

$$\frac{\partial}{\partial \alpha} T(\mathbf{F}^n + \alpha(\mathbf{y}^n - \mathbf{F}^n)) = \sum_i (y_i - F_i) t(F_i + \alpha(y_i - F_i)) + (F_i + \alpha(y_i - F_i)) \frac{\partial t(F_i + \alpha(y_i - F_i))}{\partial \alpha} \quad (4.26)$$

This function can be evaluated if an analytical expression is provided for the delay function $t(F_i)$ and a root finding method such as the bisection method can be applied to find the optimal step size. Let us consider the first iteration of the algorithm where clearly \mathbf{F}^0 and \mathbf{y}^0 are both integer valued solutions, since they present the flows for the computed shortest paths. It must be noted, however, that the step size α is not chosen such that the resulting new flows \mathbf{F}^1 are integer-valued. In fact the condition that needs to be satisfied by α to ensure integer valued solution at the end of the iteration is that the admissible values for the step size should split the feasible region $[0, 1]$ as the greatest common divider of the difference between the current flows \mathbf{F}^0 and \mathbf{y}^0 , which is \mathbf{d}^0 . The condition can be summarized as follows:

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

$$\alpha = \frac{i}{GCD(\mathbf{d})} \quad \forall i \in \mathbb{N} \leq GCD(\mathbf{d}) \quad (4.27)$$

Where $GCD(\mathbf{d})$ represents the greatest common divider of the numbers in vector \mathbf{d} . It must be noted, that as long as there is a prime number in \mathbf{d} or at least one of the numbers in \mathbf{d} is equal to 1, the greatest common divider will always be 1. This means that the only admissible values for α are 0 and 1. On a more intuitive level, the existing algorithms say that points \mathbf{F}^0 and \mathbf{y}^0 are both feasible solutions and due to the convexity of the feasibility space, all points on the line connecting them are also feasible solutions. The additional constraint for integer values, however, states that only points on that line, which have all integer valued solutions are feasible. Although, it might not seem to be such a binding constraint, it, for all intents and purposes, makes the already existing algorithms incapable of finding good solutions. In the realistic scenario that has been studied in this work, there is not a single case where α has an admissible value, other than 0 or 1. Therefore, the additional integer constraint in Equation 4.27, turns the convex combination method into a simple all-or-nothing assignment procedure, which is known to not find an optimal solution. That being said, it would be unfair to compare the solutions of the presented algorithm in this work with such a method. As the convex combination method computes a theoretical system optimum, although, with unfeasible assumptions and thus solutions, it will still be used as a benchmark for the presented algorithm in its original form. The convex combination method for SO computation is summarized in Algorithm 3.

Step 0: Initialize flows.

Compute all routes on the basis of a preferred algorithm i.e shortest paths. The resulting vector is: \mathbf{F}^1 . Set $n = 1$

Step 1: Calculate new traverse times.

Set all $t^n(F_i) = t^n(\hat{F}_i^n)$

Step 2: Find the descent direction.

Compute shortest paths and new flows \mathbf{y}^n where the graph weights are equal to the marginal costs. $w(i) = t(\hat{F}_i^n)$

Step 3: Find appropriate step size.

α^n , which minimizes the objective function $T(\mathbf{F}^n + \alpha(\mathbf{y}^n - \mathbf{F}^n))$

Step 4: Calculate the flows for the next iteration.

Set $\mathbf{F}^{n+1} \rightarrow \mathbf{F}^n + \alpha^n(\mathbf{y}^n - \mathbf{F}^n)$

Step 5: Calculate new traverse times and test for convergence.

Set all $t^n(F_i) = t^n(\hat{F}_i^n)$

If the chosen convergence criterion is satisfied stop, else set $n \rightarrow n + 1$ and go to step 2

Algorithm 3: Convex combination method for calculating system optimum

The convex combination method will be implemented and used as a benchmark for the algorithm suggested in this work. According to [17] it converges to a satisfactory solution within 5

iterations. The Partan extension of the convex combination algorithm implementation shown in [129] uses a simulated network there were roughly 50,000 drivers for which using this method more than 320,000 paths must be computed. This translates to more than 6 paths per driver. Very large real networks storing so many paths can lead to memory problems [129]. It must be noted that in the suggested implementation this problem is non-existent since only one path is stored per commuter at all times.

Some existing algorithms work with path flows instead of link flows and thus employ a method called column generation, which reduces the size of the problem by concentrating on basic variables. This was first studied for user equilibria in [130]. Detailed description of UE and SO algorithms can be found in [17, 131]. Furthermore, computation of Wardrop's equilibrium or UE is discussed rigorously in [132] and algorithmic implementation are suggested in [133] that shifts flows from paths until all path costs equalize and in [134, 135, 136]. What makes the suggested algorithm different from existing ones is that it works with link flows and also that it stores the paths of all drivers at all times. In fact, new paths are computed only for certain vehicles and the old ones do not need to be stored, which makes it less demanding for memory.

Furthermore, there is no need for explicitly implementing column generation methods, since they naturally appear in the algorithm as it only deals with the most congested links in the system and thus the rest are not subject to attention. The algorithm can be perceived as a backwards version of the incremental assignment shown in [17]. Instead of assigning vehicles one by one and changing the weights at every step according to the current traffic situation, all commuters are assigned routes according to shortest path or other graph weight choice and then the vehicles are re-routed one by one from the most congested links in the network. This strategy can be seen as a greedy backwards incremental approach.

A different modelling approach has been demonstrated in [137] where a cell-based extension of the Merchant-Nemhauser model [138] is presented and it is shown that it becomes a linear program. It is also pointed out, however that marginal cost algorithms can be more efficient. In [139] a Nash equilibrium and a system optimum approaches are compared in an evacuation scenario. The system optimum approach is used using weights specified by the marginal travel time described in [125]. These are also used in the new algorithm described in this work.

Theoretical work for formulation of SO problems has been done in [140] where a single destination system optimum dynamic traffic assignment is formulated using the cell transmission model as a linear programming problem. It also shows that a sufficient condition for SO is that every unit of flow follows the time-dependent least marginal cost path to the destination.

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

No attempt was made to address computational requirements. Furthermore, in [141] Nash equilibrium is examined as selfish routing using non-convex functions as latency functions and it is demonstrated that the equilibrium is not optimal in a congested network. They use the user equilibrium defined in [86] as a starting point. The introduction of arc capacities creates multiple equilibria and it is shown that the price of anarchy may jump to infinity.

It is possible that in an SO solution some of the commuters might have to choose paths that are much longer than the shortest ones. This problem is addressed in [129], which deals with solving the SO problem by also minimising an unfairness measure that they define as the ratio between a recommended path to the shortest possible path. Additional constraints to the optimization problem are added so that no path exceeds a certain unfairness ratio. It is shown that the resulting solution is still better than the UE. In the presented algorithm, the unfairness is also computed, however, not in the sense of additional traveled distance but rather in the sense of extra fuel spent.

On a more theoretical side of the fairness problem, in [142] it has been proven that the optimization problem of finding the minimum of the maximum latency of flows in networks is NP-hard even with linear latency functions and one origin and destination pair. In this case, however all flow-carrying paths have the same length, which means that the optimal distribution is also fair, which is not true for non-linear latency functions. The Nash equilibrium, which can be computed efficiently is shown to be within a constant factor of the system optimum. It is, however, shown that the price of anarchy is unbounded even with a linear latency functions when there are multiple origins and one destination. It is also shown that the unfairness is bounded from above by a polynomial of the degree of the latency function.

An improved numerical approach for faster computation of the interior point method (IPM), which can be used to solve the SO problem is described in [143]. A distributed multi-vehicle routing algorithm is proposed that minimizes the network breakdown probability. Through distributed matrix decomposition techniques the algorithm is scalable and is able to perform well in large city scenarios. One advantage of this approach is that it allows for an easy extension to control only a portion of the traffic in an optimal way, which is a more realistic scenario. This algorithm is similar to the one used in Chapter 3 for optimal lane computation, however it does not exploit the fact that there is a road network and routing on it is faster than using the IPM.

Finally, an approach that might have larger implications in the future with the increase of computational power is described in [144]. A surrogate based optimization approach with

one-stage and two-stage models is used in order to speed up the objective function evaluation because a traffic simulation was used to evaluate the travel times rather than a latency function.

4.4.2 Proposed Algorithm: Backwards Incremental System Optimum Search (BISOS)

The algorithm presented in this section aims at resolving some of the inefficiencies underlined in the previous section of already existing algorithms and to provide, feasible and practical solution to the SO traffic assignment problem. First of all, it does not have to store in memory any of the possible paths for an OD pair. For every user there is exactly one path from origin to destination that needs to be stored at all times. If a large city needs to be analysed in the sense of an SO solution, the path storage can become problematic, when dealing with millions of vehicles that each has several possible paths. Second, no unnecessary paths need to be computed. Instead of recomputing paths for all users at every iteration step, the proposed algorithm re-routes only a portion of the vehicles that are on the most congested road in the current iteration. This amounts to a severe reduction in the number of paths to be calculated, which is in fact the most time consuming part of any SO algorithm. In this way the column generating methods, which are usually used, can be viewed as a natural consequence of the algorithm since only roads that are congested beyond a certain threshold are inspected. Finally, it is considerably simpler to implement this algorithm and furthermore many of the computations can be performed in parallel in order to gain an additional performance boost.

Intuitively the BISOS algorithm can be described in the following way. Similar to the incremental assignment [17] it deals with vehicles one by one, or in chunks, but not all at once. The incremental assignment method starts off with all weights of the graph equal to the free flow travel times on the links. Each group of vehicles that is assigned on the network computes its paths according to the shortest path algorithm. After each chunk is assigned routes, the weights are recalculated so that each weight represents the current travel time on the respective link. This algorithm was aimed at achieving user equilibrium, however, it is shown that such state is not reached by it.

The BISOS algorithm can be viewed to do a similar process but in a backwards manner. First, all routes are computed based on shortest path algorithm with weights given by the free flow traverse time of the links. After that the most congested road is identified, by looking at the marginal cost of the road segments. A predefined number of vehicles are removed from the road by changing the weight of that road to its current marginal cost. The new routes

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

are computed for those vehicles using the new routing graph with the updated weight. In a way increasing the weight to the marginal cost of the link can be perceived as closing the road segment for the rerouted vehicles since the marginal cost is significantly larger than the free flow cost. Once the link is determined to be unable to give away more vehicles (re-routing vehicles from it leads to an increase in overall population travel time), the next most congested link is explored. When all the links that have a flow higher than their predefined capacity are explored, the iteration is finished. Please note, that the most congested link is chosen by examining the congestion factor of the link, which is proportional to the marginal cost.

For this work the BPR function is going to be used in order to evaluate the congestion at morning rush hour in a static environment since in [17] it has been pointed out that during rush hours, traffic exhibits steady - state behaviour. The general version of the BISOS algorithm is formalized in Algorithm 4.

There are some concerns with such type of greedy optimization. The major one is that converging to the optimal solution is not guaranteed. The random selection of vehicles to be re-routed at every sub-iteration excludes the possibility of optimal solution. At this point the reader must be reminded of the finding discussed in the end of the first part of this chapter in Section 4.3.2. It has been shown that the variation $\frac{\sigma}{\mu}$ of the resulting traffic conditions with respect to choosing different set of agents to move from a specific road segment is less than 10^{-4} . Therefore, it can be argued that the optimal set of agents to be moved need not be computed explicitly. In the results section the final result of the suggested algorithm will be compared to the actual SO solution achieved by the convex combination algorithm in order to verify this assumption.

Furthermore, as the set of agents to be re-routed is chosen at random, it is possible although with small likelihood that, the randomly sampled agents, which have to be removed from a certain road have no viable alternatives. In this case, the total travel time will be increased and as a result of that, the road segment will not be explored again within the current main iteration. In order to avoid this, one might set a counter of how many times, the road segment can “fail” in the sense of not producing an improvement to traffic conditions when a set of agents is removed from it. The performance of the algorithm for different values of the counter will be examined in the results section.

Data:

| | |
|----------------------|---|
| G | Road network graph consisting of nodes and links |
| A | Set of all agents in the population |
| $ComputeRoute$ | Agent \times Graph \rightarrow Route |
| $ComputeTravelTimes$ | Graph \times Routes \rightarrow Set of travel times |
| $RemoveLink$ | Link \times Graph \rightarrow Graph |
| $RandomSample$ | Set of agents \times Number of Agents \rightarrow Set of agents |
| $CalculateFlows$ | Set of routes \times Graph \rightarrow Set of flows |

Result: Set of opimal path flows \mathbf{F} , set of routes of the population R^A

Step 0: Initialize flows.

Compute all routes on the basis of a preferred algorithm i.e shortest paths. The resulting initialized vectors are: the flows \mathbf{F} , the routes R^A , the graph of every agent a is set to the original graph $G^a \leftarrow G$.

$T \leftarrow CalculateTravelTimes(G, R^A)$ // Compute the total travel time

$E \leftarrow \emptyset$ // Initialize the set of links that should not be explored anymore

Step 1: Identify most congested link.

The congestion is defined as $c(i) = \alpha \left(\frac{F_i}{2000w(i)} \right)^\beta$

$U \leftarrow \forall i : \frac{F_i}{2000w(i)} > 1$ // identify all links with flows over the capacity

$U \leftarrow U \setminus E$ // Remove the already explored links from the list.

$m \leftarrow \max c(U)$ // Identify the next link to explore.

Step 2: Change weight and re-route a given percentage of passing agents determined by the step-size σ .

$A^l \leftarrow RandomSample(A^m, \sigma)$ // Randomly sample the indicated by σ number of vehicles from the set of all vehicles passing through link m A^m

$SetWeight(m, MarginalCost(m))$ // Set the weight of the most congested link to its marginal cost

foreach $v \in A^i$ **do**

$R^v \leftarrow ComputeRoute(G)$ // compute new route

$\tilde{R}^A \leftarrow R^v \cup R^{A \setminus v}$ // replace old route with new route

end

Step 3: Recalculate flows and travel times and check if there is an improvement.

$\mathbf{F} = CalculateFlows(R^A, G)$

$\tilde{T} = CalculateTravelTimes(G, R^A)$

if $\tilde{T} < T$ **then**

$T \leftarrow \tilde{T}$ // Update the minimum travel time.

$R^A \leftarrow \tilde{R}^A$ // Update the routes of the population.

else

$E \leftarrow E \cup m$ // Add the link to the already explored links list

if $U = \emptyset$ **then**

 Go to Step 4

end

end

Go to Step 1

Step 4: Test for convergence.

$E \leftarrow \emptyset$

$ResetWeights(G)$

// Reset all weights on the graph to the free flow times

If the chosen convergence criterion is satisfied stop, else set go to Step 1.

Algorithm 4: Backwards incremental system optimum search algorithm

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

The choice of which link to examine is one of the more challenging aspects of the suggested algorithm. It is very likely that the most congested link will not be able to give away all its commuters and that at some point, moving a vehicle from it to an alternative path will start to increase the overall population time rather than decrease it. However, the link might still have the highest congestion value. In order to solve this problem, once such a situation occurs, the link in question is excluded from the list of explorable links and its weight is not reset from the marginal cost value assigned to it.

This, however, gives rise to another problem. After traffic redistribution occurs at the lower congestion levels, the initially excluded very congested links might become viable redistribution options once again. For this reason, at every iteration of the algorithm, the set of unexplorable links is reset. In other words, starts as the empty set and all weights are set back to their free flow initial values. Then, for example, at the second iteration of the algorithm, the most congested link that was removed from the explorable set of links can be attempted again after all redistributions on the lower levels of congestion have been done.

In order to bridge the gap between the optimal solution and the solution provided by the BISOS algorithm, the threshold for congested links to be examined can be reduced to an arbitrary value. In the general implementation, the threshold is set to one, which is the capacity of the road and the most optimal state of operation in the sense that it maximizes the throughput of vehicles. This threshold can be reduced in order to further redistribute traffic on streets with lower congestion levels. The variation of the solution quality and computing time as a result of such change will be examined in the results section.

Another way to view the algorithm is through a simulated annealing [145] analogy. If the temperature of the system is the number of vehicle redistributions that have been performed, in order to find a good solution, one would need to cool down the system at a slow rate. This is one of the reasons why it is not beneficial to redistribute many vehicles from many roads all at once. Therefore, the step size (the number of agents to be rerouted at once) of the algorithm should be chosen with great care.

A standard stopping criterion is chosen for the suggested algorithm. If the relative difference between the solution of two consecutive iterations is less than a predetermined ratio, the algorithm halts. Please note, that the algorithm can only halt at the end of an iteration. One possible drawback of such a strategy is that once a very good solution is reached at the end of one iteration, the algorithm is forced to go through all over-capacitated links once again

and try to compute new paths for a large number of agents in vain, since there is not much redistribution to be done.

The proposed algorithm is also prone to providing more fair paths to the vehicles, since the weights of the routing graph are not being changed, therefore, drivers that should be rerouted find the second best shortest path that exists between the origin and destination. As a conclusion the suggested algorithm provides a reduction of memory requirements, reduction of computation time and an increase of fairness of the calculated paths.

4.4.3 Results

This section will examine the performance of the algorithm compared to other already existing methods. The performance will be evaluated with respect to the speed of the algorithm and the accuracy of the final solution. Furthermore, the performance for different step sizes will be examined. Finally, an experiment investigating the improvement that the algorithm brings to the overall traffic situation will be examined for different population sizes. The city of Singapore will be used as a case study for all experiments.

4.4.3.1 Quality of Solution Compared to SO Solution and Speed of Convergence

The SO solution for the traffic assignment was computed using the convex combination method and compared to the solution of the proposed algorithm. The basis of the comparison will be the number of paths to be computed since this is the most time-consuming part of SO computation algorithms and is independent of the machine used (unlike the runtime). The speed of convergence comparison of the BISOS algorithm and the convex combination algorithm with and without the integer constraint is shown in Fig. 4.6.

It can be observed that the BISOS algorithm converges much faster than the convex combination method. In fact, the BISOS algorithm needs only 87,000 route computations, which is about a quarter of the computations needed for just one iteration of the convex combination algorithm. Altogether, the BISOS algorithm converges 15 times faster. This is a significant result. Since the optimal solution is not guaranteed by BISOS, Fig. 4.7 examines the difference between the quality of the solutions presented by the BISOS algorithm with different congestion thresholds and the convex combination method with and without the integer constraint.

It can be observed that the BISOS algorithm is not only faster than the convex combination method but also provides a much better solution than the integer constrained version of the convex combination method. Furthermore, the final solution is only within 1% difference from the

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

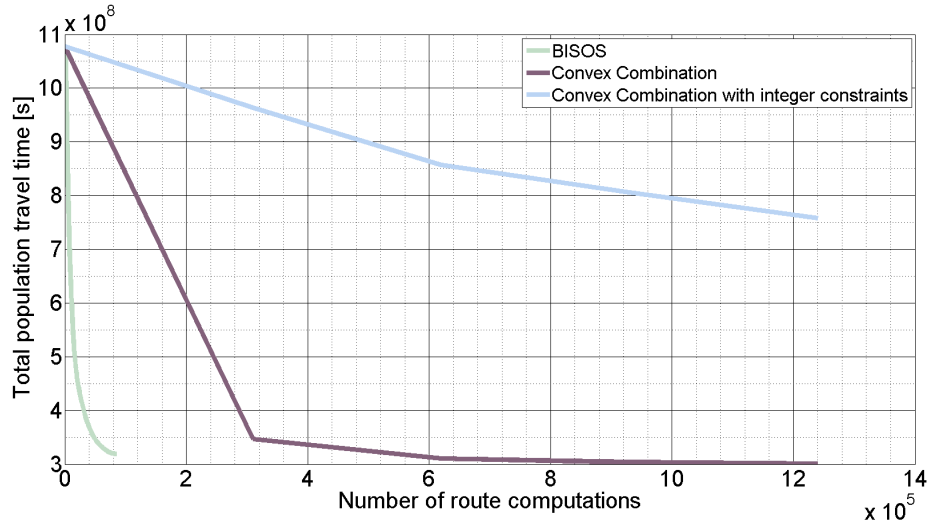


Figure 4.6: The convergence speed with respect to needed route computations for convergence for the BISOS algorithm and the convex combination algorithm with and without integer constraints.

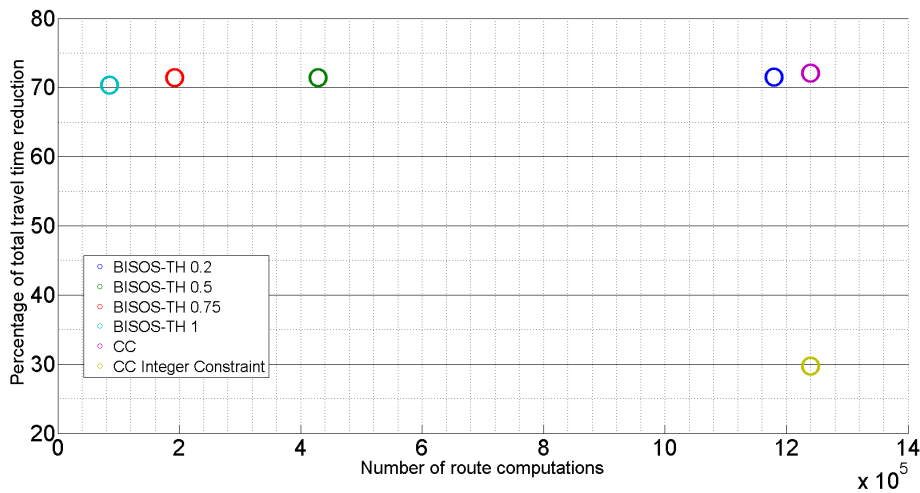


Figure 4.7: A demonstration of the trade-off between computing time and quality of system optimum solution and comparison between the final solution of BISOS algorithm and the convex combination method with and without integer constraints.

theoretical, although infeasible, system optimum solution. The trade-off between computation time and optimality can be observed as well. Decreasing the threshold congestion value, makes the algorithm examine more road segments, thus increasing the number of route computations, however distributing more traffic and further reducing congestion levels. The default value of

the congestion threshold th is kept at 1 for the rest of the experiments since the throughput of a road is maximized for this value. The results from Fig. 4.7 have been summarized in Table 4.3.

| Algorithm | Number of route computations | Final travel time improvement [%] | Number of sub-iterations |
|--------------------------------|------------------------------|-----------------------------------|--------------------------|
| BISOS $\sigma = 1, th = 1$ | 46,557 | 69.32 | 46,557 |
| BISOS $\sigma = 2, th = 1$ | 60,058 | 70.05 | 30,029 |
| BISOS $\sigma = 4, th = 1$ | 68,296 | 70.23 | 17,074 |
| BISOS $\sigma = 8, th = 1$ | 78,064 | 70.39 | 9,758 |
| BISOS $\sigma = 16, th = 1$ | 85,296 | 70.4 | 5,331 |
| BISOS $\sigma = 32, th = 1$ | 94,176 | 70.4 | 2,943 |
| BISOS $\sigma = 16, th = 0.2$ | 1,179,040 | 71.48 | 73,690 |
| BISOS $\sigma = 16, th = 0.5$ | 428,672 | 71.46 | 26,792 |
| BISOS $\sigma = 16, th = 0.75$ | 192,336 | 71.39 | 12,021 |
| CC | 1,240,000 | 72.07 | NA |
| $CC_{integer}$ | 1,240,000 | 29.75 | NA |

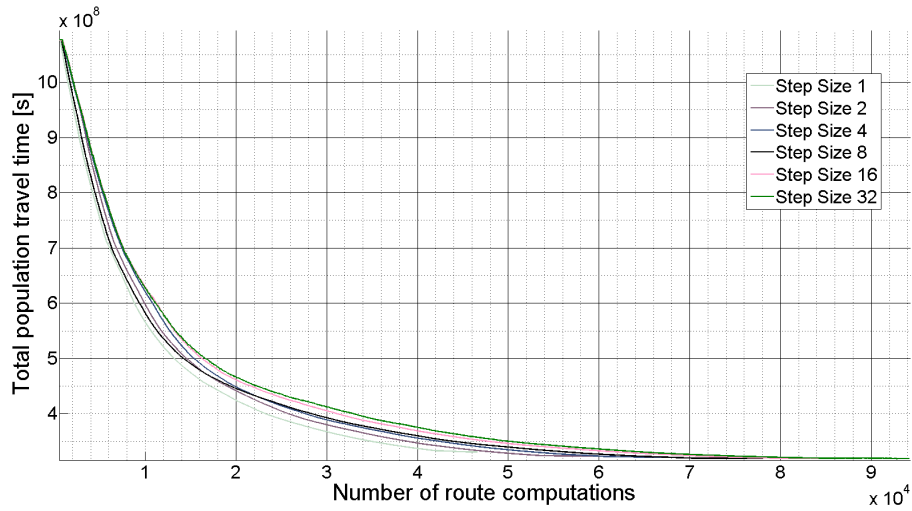
Table 4.3: BISOS comparison with convex combination method for different step sizes and threshold values.

4.4.3.2 Performance for Various Step Sizes

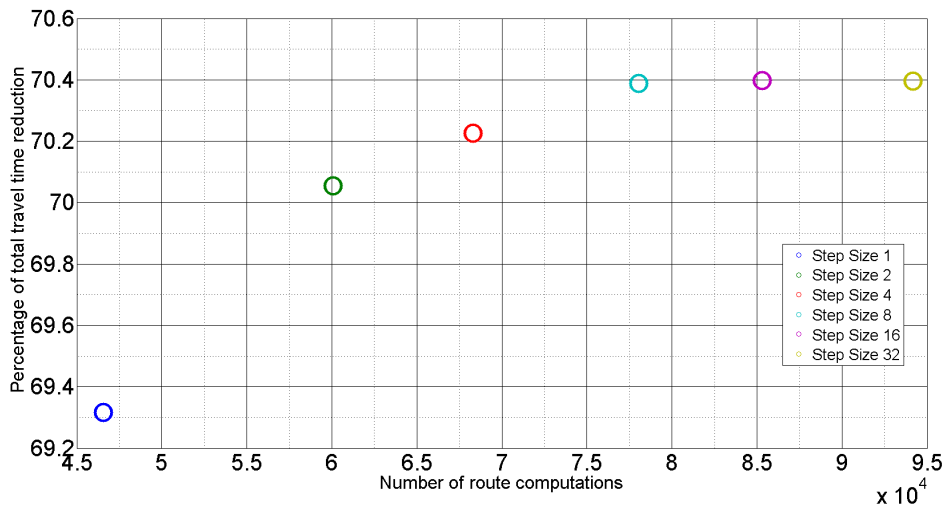
Varying the step size σ of the algorithm has two main aspects to be examined. The first one is the speed of convergence. Technically, the bigger the step size, the higher the opportunity to parallelize the algorithm. For example if the step size is set to 20 vehicles to be re-routed at one step, in the presence of 20 cores available the time for one iteration will be virtually the same as re-routing 1 vehicles at a time. The increase of the step size, however, might become practically unnecessary when a realistic maximum number of available cores is reached. The second aspect to be examined is the quality of the solution that is reached with the various step sizes. Fig. 4.8 depicts the step size influence on the quality of the final solution and the number of route computations for convergence.

A rather unexpected result, which can be observed on Fig. 4.8b and Table 4.3 is that the SO solution gets better with increasing step size. Intuitively, the opposite can be expected to be true since using a smaller step size would allow to find the point where a road cannot give away any more vehicles in a more precise way. The results can be explained with the

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT



(a)



(b)

Figure 4.8: BISOs algorithm evolution for difference step sizes Fig. 4.8a and convergence speed as a function of required route computations for convergence with respect to BISOs's step size 4.8b.

stochastic nature of the BISOs algorithm. Naturally, there exist agents in the set of drivers, which pass through a road segment, which should not be moved. When a big group is sampled, even if such agents exist others, counteract their negative influence of the total travel time and the re-routing step is still successful and the segment continues to be examined. When only one agent is sampled at a time, if its re-routing is not beneficial for the system, the algorithm

labels the examined segment unexplorable and moves on to the next most congested road, even though, there might be more traffic to be reassigned. In other words, a high step size ensures the existence of the finding in Section 4.3.2, while a small step size increases the probability of premature removal of the road from the examinable roads list. This concept is formalized in Appendix A.

In order to test the correctness of this statement, a new parameter is included in the algorithm, which allows a certain number of unsuccessful re-routings for each routing segment. This number, which is referred to as the “failed attempts limit” or FAL is varied in order to observe its effect on the quality of the SO solution. The results are shown on Fig. 4.9

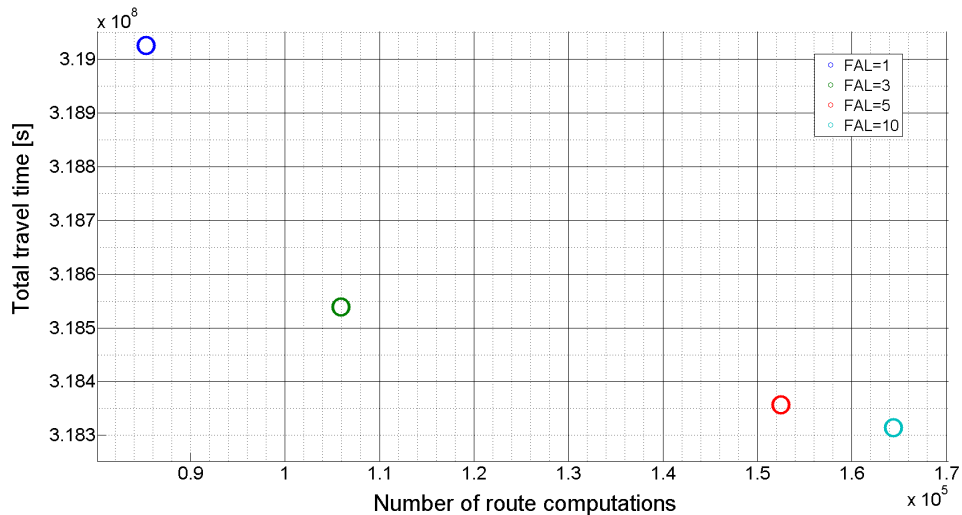


Figure 4.9: Trade-off between number of computations and quality of SO solution for different values of the failed attempt limit value.

It can be observed that, as predicted, with increasing the FAL the quality of the BISOS solution increases. Furthermore, the very exceedingly small difference between number of route computations for FAL= 5 and FAL= 10 indicates that saturation is reached and no further increase in the FAL value is required.

4.4.3.3 Scaling of Algorithm with Population Size

As the performance of the BISOS algorithm has been validated and proven exceedingly better than the complex combination method for the examined case study, it is important to evaluate how the performance of the algorithm scales with increasing population size. The first component to be evaluated is the quality of the solution, which will be once again compared

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

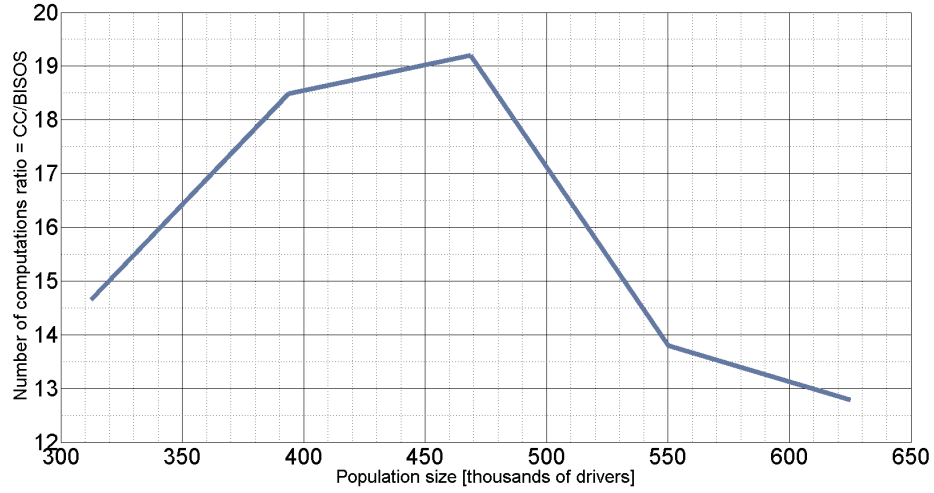
to the non-integer constrained convex combination method, despite the fact that it provides infeasible solutions. The population size has been increased in incremental steps to twice the current population and the system optimum was found using both algorithms. The quality of the solution is measured as in the previous section using the ratio between the percentage of improvement provided by the BISOS and the CC algorithms. It is important to note that the origins and destination of all drivers are identical for both algorithms. Next, the computational performance of both approaches was compared by looking at the ratio of route computations performed for each of the population sizes. The results can be found in Fig. 4.10.

It can be observed that as the population size is increased the solution computed by BISOS get closer to the theoretical minimum travel time computed by the CC algorithm. It can also be noticed that there is no definitive trend for the scaling of the speed of convergence ratio. The BISOS solution is found between 12 and 19 times faster than the CC solution. Additionally, since the aim of this part of the thesis is to evaluate the possible effects system optimal routing might have in the future Fig. 4.11 represents the system optimum travel time evolution as the population size is increased.

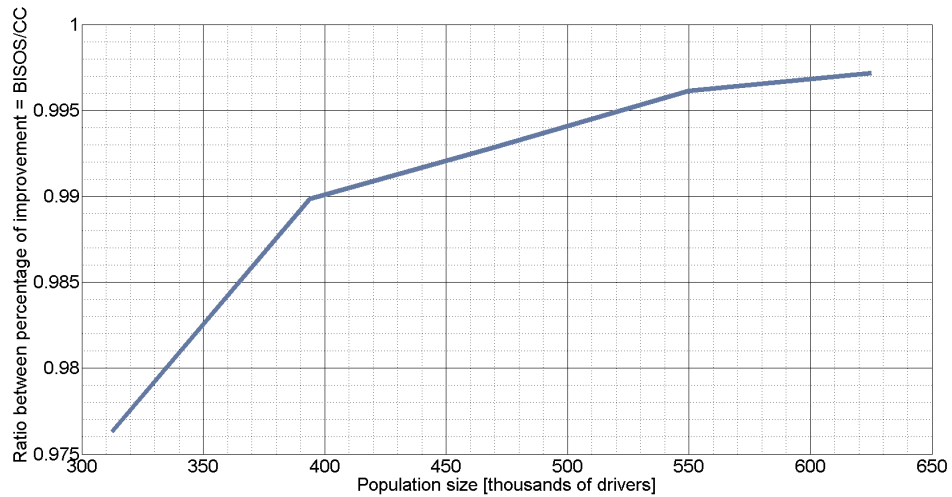
An expected exponential growth of the total population time can be observed as a result of both increased number of drivers and the non-linearity of the delay function. Fig. 4.11 also depicts the comparison between the total population travel time for 310,000 with selfish fastest path routing and the SO population travel times. It can be stated, that according to the results, if SO optimum routing is utilized the driver population of the examined city can be doubled in size and the total travel time of the population would still be the same as the selfish routing current one. At this point it should be mentioned that the total population travel time for SO in the case of 620,000 drivers, in fact, has twice as many trips. This means that if the average trip time of a driver is being compared, SO would allow doubling of the population size, while halving the average trip duration. The presented results strongly imply the importance of utilization of system optimum routing centralized system in the future of transportation systems.

4.4.3.4 Fuel Consumption Model and Evaluation

In order to evaluate the added cost of choosing optimal routes, which increase the total traversed distance by the population, the fuel efficiency of the BISOS solution is evaluated and compared against the initial traffic distribution. In this section it is shown that on top of saving travel time for the whole population, the traffic redistribution algorithm also saves fuel. This is due



(a)



(b)

Figure 4.10: Effects of increasing population size on the speed of convergence ratio between BISOS and CC for increasing population size Fig. 4.10a and on the ratio between relative improvement of total travel time for the BISOS and CC algorithms Fig. 4.10b

to the fact that vehicles are much more inefficient in congested environments. Therefore, when congestion is reduced, fuel consumption is decreased as well. It must be noted, however, that the fuel decrease happens despite the increase of overall path length of the users. There are two main changes that the new distribution brings, which affect the fuel consumption. First, the congestion levels are decreased, which reduces fuel consumption. Second, the overall path

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

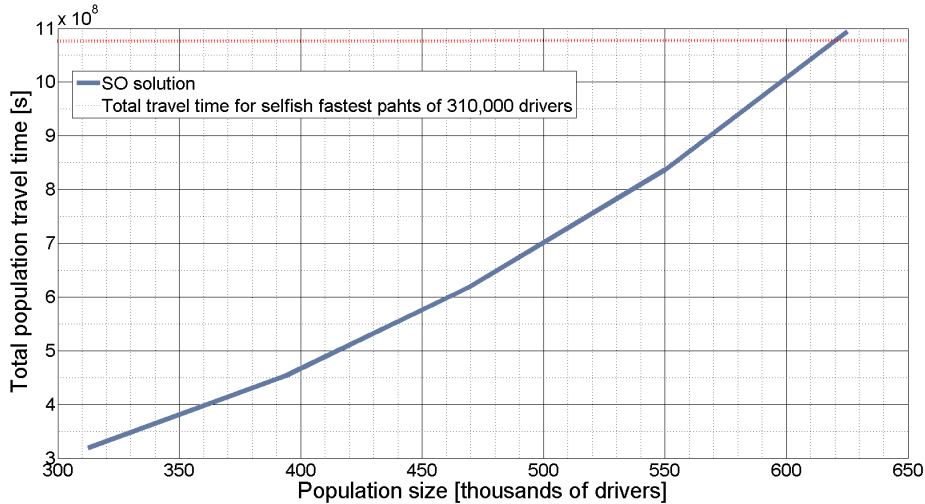


Figure 4.11: Comparison between total population travel times for different population sizes. The red dashed line, represents the total population travel time for the 310,000 initial population under fastest path selfish routing traffic assignment.

length of the commuters is increased, which increases fuel consumption. Apparently, the effect from the first factor is stronger and therefore, a reduction in the total amount of consumed fuel is observed.

The fuel consumption calculation is derived from an average speed fuel consumption model developed first in [146] and further discussed in [147, 148, 149, 150]. This model, also known as the elemental model, creates a relationship of fuel consumed by a measure of distance that is inversely proportional to the average velocity. It has been chosen for this study since it only requires information about the average velocity on a road segment and is thus very applicable to macrosimulation approaches like the one undertaken in this work. The model can be formalized as:

$$C = k_1 + \frac{k_2}{V} \quad (4.28)$$

Where C is the fuel consumed per 100 km in litres and V is the average velocity in km/h. The parameters k_1 and k_2 are fitted using linear regression using real world measurements of vehicle fuel consumption. The BPR delay function allows the calculation of the average speed on a road, given the flows. In order to get the total fuel consumption of the commuter population from the flows on every road segment, the following formula can be applied:

$$C_T = \sum_i F_i \left(k_1 + \frac{k_2}{t(F_i)} \frac{l_i}{100} \right) \quad (4.29)$$

Where C_T is the total number of fuel consumed in litres, l_i is the length of the segment i in km and $t(F_i)$ is the traverse time of link i with flow F_i in hours. Since the model only promises accurate representation of fuel consumption for average velocities higher than 10 km/h, the minimum speed of all segments is set to 10. This might underestimate fuel consumption in congested cases, however, the results still demonstrate very clearly that fuel is indeed saved. For the standard use case scenario of Singapore, on top of saving 70% time, the system optimum solution also saves roughly 15% fuel. The personal implications on the agent population, however, should also be calculated in order to check if there are users, which are forced into spending more fuel than needed. Fig. 4.12 depicts the distribution of fuel saved in litres for the agent population.

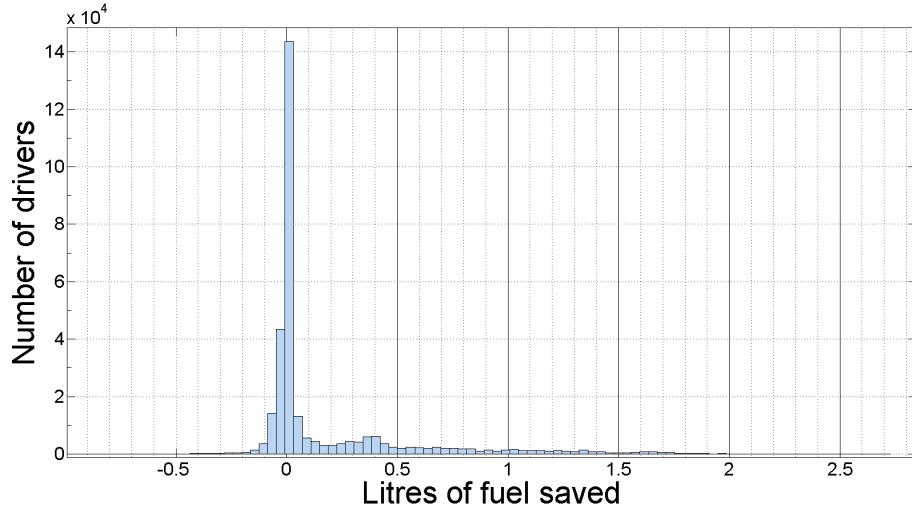


Figure 4.12: Distribution of difference of saved fuel for each agent for system optimum routing traffic assignment.

It can be observed that some of the drivers indeed use more fuel for the system optimum solution. This extra fuel, however, as can be observed from the distribution is at most half a litre more. This seems to be a reasonably small excess cost, compared to the social benefits brought forward by the system optimum solution.

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

4.4.3.5 Participation Rate Effect on Average Travel Time

The BISOS algorithm can be used as an engine empowering a city-scale routing platform, which can service the driving population of a city. If a system consisting of only autonomous vehicles is considered, then system optimum routing can be seamlessly integrated as the route choice can also be made automatically. If there are human-drivers in the system, however, they might not all utilize the platform and follow the suggested system optimum routes. Therefore, the participation rate effect on the system's performance should be studied.

In order to do that, the following scenario is set up. The percentage of drivers using the SO platform is varied from 0 to 100 in 10% intervals. Once the percentage is set, the participants (people following the advice of the SO platform) and non-participants (people not following the advice of the SO platform) are chosen at random from the agent population. The non-participants simply choose the fastest routes under free-flow traffic conditions while the participants query the SO platform and follow the system optimum route. Fig. 4.13 shows the simulated average travel time of the system for different percentages of participation of the driver population. Furthermore, a similar platform, which, however, computes user equilibrium routes was implemented as well in order to observe the difference of system performance between the two approaches. The user equilibrium platform basically represents a service very similar to *Google Maps*, where a person would query the platform with an origin and destination and will be provided with the fastest route given the current traffic conditions.

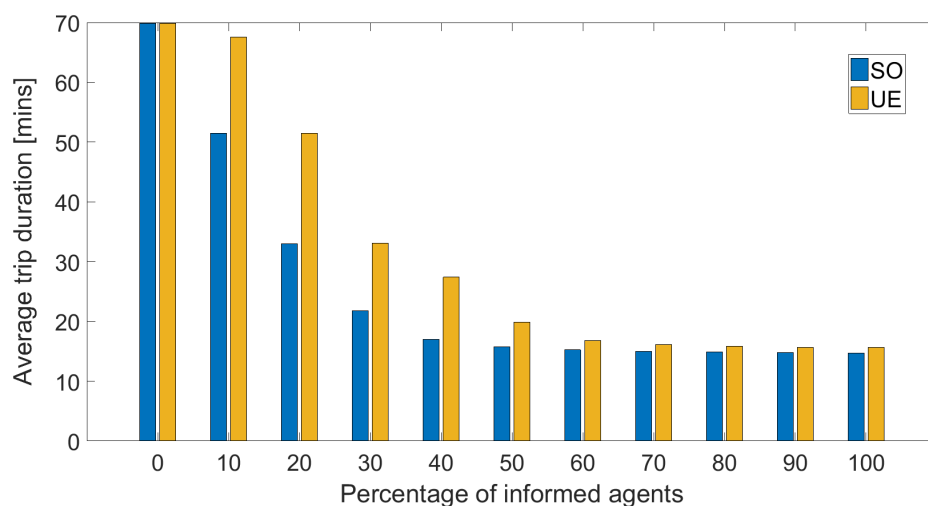


Figure 4.13: Participation rate effect on average travel time for system optimum and user equilibrium traffic assignment.

Two main observations can be made. First, the difference of performance between 100% and 40% participation is very small in the SO case, which means that only about 40% of the population need to be convinced to follow the advice of the SO routing platform. Second, less people are required to follow the SO system compared to the UE system in order to achieve the same average travel time. For example 30% participation rate in the SO platform produces almost the same average travel time as 50% participation rate for the UE platform. This would mean that providing the right routes in a routing platform can make a significant difference on traffic conditions and that SO routing platforms need, in general, less participants in order to function at a satisfactory level.

4.5 Chapter Summary

4.5.1 Sensitivity of Traffic Conditions to Road Removal

In Chapter 4 the sensitivity of the traffic system against single road segment removals is studied. It has been shown that the closure of a single road segment might reduce overall travel time for the whole commuting population by as much as 4% which corresponds to thousands of saved hours on a daily basis. This rather extreme outcome of a small change in initial conditions is a manifestation of the complexity of the system and is also known in transportation literature as the Braess paradox.

The contribution of this part of the chapter lies in the scale of the performed experiment. A realistic scenario is being simulated by taking a complete road network of a large city and populating it with agents according to collected real world data that is also used for the calibration and validation of the presented model. The completeness of the scenario ensures that the effects of the paradox are not only local but global. Exploiting the computational power available in present days will be a key tool for the planning, control and support of future smart cities. Simulation based methods such as adaptive and selective road closures can be used to ensure efficient utilization of resources and fast instantaneous adaptability to traffic demand changes.

The method of simulating outcomes of network changes can also be used for future infrastructure planning to avoid building roads that produce congestion and for personalized road pricing or exchange of road usage permits that can be used to balance traffic flows and achieve a social equilibrium state. It is expected that future smart cities will rely heavily on simulation approaches enabled by the increase of computational power availability. A futuristic ITS would

4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT

make use of a system optimum routing approach, which will enable it to dynamically change the road network and thus continuously steer the system dynamics into optimal states.

4.5.2 System Optimum Computation Using BISOS Algorithm

This is the motivation for the second part of the chapter, which deals with the efficient computation of such system optimum routings. The main contribution of this part of the thesis and possibly of the whole work is the BISOS algorithm, which aims at computing a system optimum routing solution for all agents in a transportation system. The suggested algorithm converges 15 times faster than current algorithms and furthermore, provides explicit paths for every single driver, which is something that to my best knowledge none of existing methods do.

The immense reduction of route computations needed for the algorithm to converge and the practicality of its functionality present a great step in the direction of a centralized routing control system. The key practical aspect of the algorithm except its speed, is that it can be halted at virtually any point of time, if time constraints require this, and would still be able to produce explicit paths for all agents in the system, which is a highly desired trait for real time operating systems. Such an algorithm may turn system optimum routing strategies from theoretical measures for estimating the utilization of a road system into practically used strategies for severe reduction of congestion levels.

This algorithm can also be used as a first step of a hybrid system optimum computation. Since it reaches a good solution order of magnitude faster than the existing one, in case the theoretical minimum should be computed, the first few iterations of the standard algorithms can be sped up by using the BISOS algorithm final solution as a starting point for the conventional methods.

Furthermore, on a different level, the reduction of total travel time of the population as a result of system optimum routing (roughly 70%) clearly demonstrates that the routing choice component of traffic determination is the most promising one to be controlled and presents the key to ensuring smooth traffic operation. A comparison has been made between the total travel time of the population with selfishly chosen fastest routes and a population double its size, which employs a system optimum routing. The results show that even though the population size has been doubled the average travel time of a driver will still be half of what it is in the case of selfish routing.

This chapter has demonstrated the importance of routing for the improvement of the performance of a transportation system. As the potential effect of changes in any of the main

traffic determining components is being measured in this thesis, the component with the highest potential to benefit transportation systems has been identified as the routing choices of the population. For this reason, the following chapter aims at maximizing the routing choice information collected from a sensor network by placing the sensors at the most optimal locations.

**4. IDENTIFICATION OF HARMFUL ROADS AND ROUTING CONTROL
FOR EFFICIENT SYSTEM OPTIMUM TRAFFIC ASSIGNMENT**

Chapter 5

Robust Route Information Maximising Sensor Placement

5.1 Overview

After the traffic situation has been analysed and actions are taken in order to plan for a more sustainable future of the transportation system, the most important traffic determining components should be monitored in order to observe accurately and efficiently the traffic conditions. As the routing choices of drivers have been shown to be an effective steering mechanism to ensure efficient state of operation of the transportation system, a surveillance system would aim at maximizing the information about them. This chapter introduces a new entropy based metric to help identify the most important, uncertain, elements of a transportation system, as a function of network topology and level of traffic. The presented work here is taken from the author's contribution in [13]. Consequentially the presented metric is used to solve the sensor placement problem, maximising the information gain in terms of drivers' routing choices by sensing the most uncertain areas in the system. It is demonstrated that utilising the proposed strategy makes the performance robust against short and long term variations of traffic patterns. A new concept of robustness of transportation systems is introduced, which is based on perturbing the current OD matrix of the demand and requiring that an optimal solution of any planning type performs well for various degrees of perturbation thus making it robust against such traffic demand alterations. Finally, a method for finding the optimal number of sensors to be installed in a city is proposed. It models and maximises the utility stemming from the trade-off between cost, performance, robustness and reliability of the sensor placement problem

solution.

5.2 Introduction

5.2.1 Motivation

Identifying the most important modules or elements of a complex system is a problem that is of great interest to engineers and researchers. Its most significant entities or subsystems are, depending on the system, either cautiously monitored, robustly controlled, or rigorously studied in order to gain deeper understanding of the system's dynamics. In the case of transportation systems, "important" parts of the network are usually sensed in order to get information about the overall traffic state. Engineers go even further by trying to change and control traffic parameters at such locations by planning new infrastructure developments [29], control strategies [30], novel policies [31], etc.

The aim of sensing traffic, until now has been mostly to determine the flows in a city. The problem of optimal placement of counting sensors in order to estimate an Origin - Destination (OD) matrix has been around for more than four decades [27]. Knowing the OD matrix, the flows can be extracted [151, 152] and knowing the flows, the traverse times on the respective roads can also be evaluated [153], thus providing aggregated information about the traffic situation.

Given the increased pace of introduction of new technologies to the market and growing availability of computing power, traffic sensing and city planning are getting more interdependent and strongly connected. There are methods that use sensed data in real time in order to apply changes to the traffic system [32]. Therefore, sensors may not be placed with the sole reason to observe traffic. Smart cities use their sensors' data streams in order to optimise their performance in real time. With the increased number of sensor types such as plate scanning, velocity measuring, emissions measuring, etc. and their reduced error rate, it is now a matter of great importance to shift the sensor placement problem toward a more active goal. The information stream coming out from the sensors should be utilised by control algorithms or long term planning strategies in order to "actively" sense the traffic by controlling it at the same time.

Fundamentally, sensors are put in such positions so that they maximize the information gain. The locations that need to be sensed in order to maximize the information gain are the ones of higher importance. In other words, the chosen locations to be sensed are usually the

ones that present the highest uncertainty with respect to a predefined information measure. In case the uncertainty of both nodes is of equal magnitude, the one that is more used is of greater importance. Therefore, the uncertainty of drivers' choices at every node should be weighted by the number of drivers that utilise it. Depending on the definition of information the placement problem can take different forms.

5.2.2 Choice of Information Measure

In most cases, information is considered to be a characteristic of the link, like throughput or flow velocity. If, however, the aim is to find intrinsically important locations, the places where the choices, that lead to those characteristics, are made should be examined. Average flows, velocities, densities on road segments are perceived as the factors that describe traffic conditions. The main factor that determines all those, however, is the routing choices that the drivers make. Routing choices can account for a significant difference of traffic performance as shown in [154], where a 70% improvement of average travel time for the whole system was demonstrated by just optimally selecting the paths of the drivers and also in Chapter 4. A novel and more efficient approach for sensing traffic would then be to try and maximise the information about the routing choices of commuters rather than the flows, speeds and densities on the roads.

By knowing the routing choices of drivers at key intersections, a detailed map of the specific flows on the links can be inferred, which is what usually sensor placement strategies aim at. In addition to this, however, information about where the flows come from is also available. Therefore, by gaining information about the routing choices, both the link and path flows can be simultaneously approximated. Furthermore, by examining differences between sensed flows and predicted ones from the routing choices, the OD matrix can also be estimated. Therefore, routing choices stand as the source of information that determines the usually addressed problems of estimating OD matrix, path flows and link flows.

5.2.3 Sensing Intersections vs. Road Segments

Intersections are the places where drivers make choices, and what is sensed at the connecting road segments are just the consequences of those choices. Therefore, instead of examining links in a traffic network, a more topologically central approach would be to examine nodes (intersections) instead. It is important to note that the uncertainties and the dynamics of each node are expected to be weakly correlated with respect to their spatial connections. Contrary to

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

the speed, flow, etc. measures of road segments, which are usually correlated in case of physical proximity, the uncertainty at nodes tends not to have a distance dependent correlation. This key difference diminishes the need for a combinatorial mutual information optimisation approach of high complexity in order to find the optimal sensor placement, because there is intrinsically no redundant information in the sensing network.

Sensing a node or a group of nodes representing an intersection in a road network boils down to tracking the decisions of the drivers that pass through it. This task can be performed by placing plate scanning sensors on the roads that lead to the intersection and out of it. This sensing architecture allows us to both evaluate the turning choices of the commuters as well as still collect information about flows on the separate links. More than that, since in order to properly “sense” a node, only the flows along its connecting edges are needed, sensors can be placed at any position on the edge. Thus, if positioned at a midpoint along the edge, the sensors will also be able to collect information about the cruising speed along it, therefore maximising the information input.

5.2.4 Robustness

Another pressing matter that has been ignored in the past is the robustness of a sensor placement solution. Usually, robustness is understood as the error rate or redundancy of a particular sensor placement. In this work, however, robustness has been examined from a different angle. Due to the fact that sensors are quite expensive and their installation consumes both time and resources, the need to move the sensors around (if at all possible) after they are once installed should be minimized. In this sense, robustness of a sensor placement can be defined as the property of the set of locations to stay important when the traffic demand conditions in the system are changed. Such changes may include short term changes in the OD matrix such as daily variations of traffic (evening rush hour vs. morning rush hour or weekday against weekend) and also long term changes in the network demand such as people moving around the city and changing living districts and jobs, building of new living complexes or business centres, etc. Such changes may severely alter the situation for a given sensor placement and thus make the investment for their installation obsolete. The robustness of a planned sensor network against variations in the traffic demand is of great importance, especially for constantly evolving large cities.

5.2.5 Combination of Demand and Topology Information

In a more fundamental aspect, existing measures concentrate mostly on the two out of three main components that determine traffic conditions separately: traffic demand (where and when do people want to go) and transportation network topology (the medium that allows the commuters to move). The existing measures governed mostly by the demand are link flows, path flows, link average speed, etc. Measures connected to the topology are centrality, heterogeneity, entropy, etc. There have been previous efforts to define entropy (uncertainty) of a node or a link but only in a purely topological sense [34]. There is a strong need of employing the information contained in the OD matrix, namely the traffic demand, as well in order to come up with a more useful definition and measure of the uncertainty of a node and the importance of it being sensed. The combination of demand and topology, in fact, further allows the measure to take into consideration the third main traffic determining component - the routing choices of the population. In this study the entropy of a node is defined and consequently its importance by using both information about the traffic demand and topological information about the network, which surely gives a better overview than basing the definition on just one of them. The information from both sources must be entangled since they are actively affecting each other. In this way a single measure that represents all the available information can be defined.

The main contributions of this work are:

- Definition of entropy of a network and importance of nodes.
- Characterization of the concept of robustness of planned infrastructure against long term changes in the OD matrix.
- Study on the robustness of the measure against changes in the OD matrix.
- Design of methods for finding the most robust optimal sensor placement against short and long term variations
- Design of a method for finding the optimal number of sensors to be placed in a given network.

5.3 Literature Review

Determining the importance of locations in traffic networks is crucial. There are two main branches of research that are interested in locating central spots in a network. The first one is

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

traffic sensing. In most sensor placement problems, the set of locations to be sensed is chosen such that, the resulting synthesis of data is the most informative, which boils down to sensing the important locations that describe the traffic demand. The other area is complex networks research. Importance is then defined and studied in a purely topological sense by examining the transportation network without considering any traffic demands. This review will cover both areas with an emphasis on the sensor placement studies.

There have been many attempts to find optimal sensor placement in order estimate an important traffic characteristic. One of the most comprehensive surveys [155] discusses and summarizes existing sensor location problems. It defines the traffic sensing problems into categories depending on sensor types (AVI sensors, counters etc.), prior information and flows of interest (link flows, route flows, OD flows). The optimisation problems are divided into two categories: Flow observability problems and flow estimation problems. Moreover, it describes different rules for optimisation and analyses methods such as flow intercepting, demand intercepting, independence of traffic counts (mutual information). This work is valuable because it summarizes and categorizes the various approaches by providing a unifying picture of existing strategies.

One of the most standard traffic characteristic to be observed is the OD matrix. Estimating it from sensor data has become a central problem discussed in numerous studies [156, 157, 158]. In [159] the sensor location problem for OD matrix estimation is defined and a solution is suggested. The study deals with counting sensors, while other studies also include the possibility of using (AVI) Automated Vehicle Identification readers, which are more informative since they also collect information about the identity of the car, which allows for easier tracking and therefore path estimation [160]. In [161] both types of sensors are used in a method that places counting sensors and AVI readers to maximize the expected information gain for an OD demand estimation problem. It also takes into consideration uncertainty in historical demand information. A technique for calculating the optimal number and locations of plate scanning sensors for a given OD matrix is also presented in [162]. Those approaches are centred around the goal of estimating the OD matrix. In most of the cases they are applied on artificial networks as a proof of concept, however their high complexity might turn into a disadvantage if one tries to apply such a strategy for a real life large city. Therefore there is a need for sensor placement method that is less computationally intensive so that it is practically applicable.

Once the locations of sensors are fixed one might use a linear approximation technique in order to estimate the OD pairs using traffic counts offline such as the one described in [163].

In case plate scanning sensors are used a method for path reconstruction from such type of data can be used as in [160]. In [164] methods for extracting information from sensors data in order to estimate travel times are discussed, while also looking at sensor failure probabilities. Furthermore, due to the heterogeneous nature of collected data, information demands and the limited storage capacity of road side sensors, a maximum content dissemination strategy must be employed as well as done in [165]. Another issue that must be taken into account once the road segments that need to be sensed are determined is the feasibility of positioning a sensor there. More precisely, it must be verified if the sensor will be able to handle the volume in the sense of contact time and contact rate or more sensors must be placed at the same road segment in order to increase the accuracy of the extracted data. A rigorous analysis of these problems can be found in [166].

There are more universal approaches for choosing the most important locations to be sensed, which are based on maximizing information gain. There are information theoretic techniques such as [167], where a non-myopic strategy is used to find the most informative locations for sensors, [168] where a Kalman filtering structure is employed in order to solve a traverse time prediction problem via optimally placing sensors, and [169] where a method for target localization and tracking is presented, which computes the posterior target location distribution minimizing its entropy. Furthermore, in [170] the spatial and temporal correlation between the flows are used in order to feed an ant colony optimisation algorithm that finds optimal sensor locations.

Information theoretic approaches, however, may vary among each other. In [171] traffic phenomena are modelled as Gaussian processes. They discuss maximizing entropy for sensor locations and also mutual information between the locations and demonstrate that the mutual information approach performs better for certain type of scenarios. Moreover the method is extended to find robust placement against failures of sensors and uncertainties in the model and uses real world data sets. This is a generic method that can be applied to different types of sensors. It locates the most representative links in the network that reduce the uncertainty about the unobserved links. There is, however, no method that is able to determine the most important links in the sense of locations where drivers makes the choices that are later observed at the representative links.

With the advancement of technology some type of sensors now can be mobile instead of static, while granting better coverage. In [172] a mobile traffic surveillance method is presented. A routing problem is defined such that it computes the optimal paths for the mobile sensors and

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

show that in most cases it performs better than a static network. Mobile sensors provide several advantages such as bigger area of coverage, adaptability to changes in traffic patterns and are the better approach when there is no prior knowledge about the system. In [173] a strategy for a sensor placement for monitoring mass objects is described. By allowing the sensors to be mobile the sensing network can self-organize in order to achieve better coverage. An expectation-maximization algorithm is used in order to update the distributions of objects, which are then used to implement an adaptive sensor placement strategy for the desired tracking task. On the other hand, static sensor placements are easier to implement, cheaper to maintain and due to their static nature are able to use technologies that are more complex and precise. One more advantage of static sensor networks is that they are able to provide a better “instantaneous” picture of the traffic situation in the sense that mobile sensors collect samples that vary both spatially and temporally, while static sensors collect much larger number of samples for exactly the same time period. Especially in rush hour conditions with fast changing traffic patterns such temporal stability is valuable for a more precise estimation of traffic characteristics.

In [47] the authors demonstrate using traffic indicators that importance of road segments is mainly determined by the network structure and the flows. Even though, this statement is clearly known there is still no indicator of importance of road segments that fully utilizes the flow information and the topological properties of the network. As it can be seen most methods to determine important locations for sensor placement are based mostly on the flows; while in a separate part of literature people look at purely topological properties of transportation graphs.

Important locations can be determined based solely on the topology of a network. Some efforts deal with identifying critical links using a network robustness index based on link flows, link capacity and network topology as in [48]. In [49] the most vital links or nodes are defined as the first n links or nodes whose removal will lead to the biggest increase in average shortest path distance. While in [50] the importance of roads is simply defined to be proportional to the traffic load on them, in [51] three measures of centrality for a street are suggested: closeness, betweenness and straightness and their correlation to various economic activities in the respective areas are examined.

Moreover, the network itself can have some properties that are usually based on the structure of the system and not on local properties of its elements. In [52] the development of the Swiss road and railway network during the second half of the 20th century is investigated. It is observed that the spatial structure of transportation networks is very specific, which makes it hard to analyse using methods developed for complex networks. In [53] existing

measures of heterogeneity, connectivity, accessibility, and interconnectivity are reviewed and three supplemental measures are proposed, including measures of entropy, connection patterns, and continuity. Entropy is also used in order to determine the heterogeneity of the network regarding a chosen parameter.

The topology of a network holds an enormous amount of information. It may provide insights into the structure of the roads (transportation networks are organized hierarchically as shown in [54]). In [55] they measure the efficiency and accessibility in Paris and London based on the network connectedness. Moreover, this information can be utilized in order to reconstruct driver's trajectories from GPS signals as in [56]. There is also a family of graph measures based on entropy that are rigorously summarized in the survey [34]. It includes some measures from chemical structural analysis and social network analysis. The survey examines the overall connectedness of graphs such as the topological information content and the entropy of the weights of the edges. A measure of local features' such as entropy of nodes is defined as well, based on length of links connected to it. The centrality measure of links is also defined. Most of the measures deal with evaluating the information content in the graph itself. Those measures are highly uncorrelated, which means that they capture different aspects of graphs, so the proper measure should be chosen for each specific task.

Once a measure of importance is defined and the most informative locations are chosen, there is one more aspect that needs to be examined. The robustness of those choices depends on the evolution of both the topology of the network and on the evolution of the OD matrix as well. Those two factors are naturally also highly interdependent. In [57] the evolution of the topology of networks is observed. A high degree of self-organization and spontaneous organization of hierarchies is observed in the city of Indiana. Also variations in the relative importance of parts of the network are observed. In [58] the evolution over 200 years of a North Milan road network is observed. Two main processes can explain the developments that occur. Densification of the road network around the main roads and emergence of new roads as a results of urbanisation. An evaluation of the robustness against such type of long term network evolution for any type of sensor placement is lacking at the moment.

5.4 Measuring Importance of Nodes

In this section the measure of importance of nodes is introduced. A node is defined as important if many drivers pass through it and there is high uncertainty about the choices they make. In

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

order to get the uncertainty an entropy measure at the node is needed. Following that stream of logic, the measure is weighed by the throughput of drivers. In this way it can be measured how much this node adds to the overall uncertainty of the road network given an OD matrix. The notation that will be used throughout this chapter is introduced in Table 5.1.

Shannon's entropy is calculated using the transition probabilities between the states of the system. Let us assume that the state of an is its current link. The set of possible transitions from this state represents the set of actions of the turning on any of the links that are successors of the current link. The entropy of the node connecting those links is calculated using this information.

The following are the steps taken in order to calculate the importance of a node:

1. Calculate turning probabilities:

Let N_{ij} be the number of cars that pass through the i -th node and after that through the j -th node, where node j is a successor of node i in the directed graph describing the road network, and let P_l be the path of the l -th . Then let the function $f_{ij}(l)$:

$$f_{ij}(P_l) = \begin{cases} 1 & \text{if nodes } ij \text{ are in } P_l \\ 0 & \text{otherwise} \end{cases} \quad (5.1)$$

Then:

$$N_{ij} = \sum_{l=1}^{|A|} f_{ij}^l(P_l) \quad (5.2)$$

, where $|A|$ is the number of s.

Let p_{ij} be the probability that an at node i continues to node j

Let S_i be the set of nodes that are successors of node i . Then the turning probability can be defined as the ratio between the number of cars that pass through node i and then proceed to node j and the total number of cars that pass through node i :

$$p_{ij} = \frac{N_{ij}}{\sum_{k \in S_i} N_{ik}} \quad (5.3)$$

2. Calculate the entropy at every node:

The entropy of a node i , H_i , is calculated using Shannon's entropy definition. A state is represented as the current link an is on and the transition probabilities are the turning probabilities from this node to its successors. Then the entropy becomes:

5.4 Measuring Importance of Nodes

| Variable | Description |
|----------------|---|
| N_{ij} | number of cars that moves from node i to node j |
| P_l | the path of the l -th |
| $f_{ij}(l)$ | function that is one if the sequence of nodes ij is the that path of l |
| A | a set containing all the s |
| p_{ij} | probability that an that is at node i will continue on to node j |
| S_i | set of nodes that are successors to node i |
| H_i | entropy of node i |
| I_i | importance of node i |
| T | number of regions the day is split into |
| N_{ij}^t | number of cars that pass sequentially through node i and j during time period t |
| H_i^t | entropy of node i during time period t |
| I_i^t | importance of node i during time period t |
| \bar{I}_i | overall daily importance of node i |
| \bar{I}_i^d | the overall importance of node i for a degree of perturbation d |
| R | total reduced entropy |
| L | a set of sensor locations |
| L^d | locally optimal sensor placement for a degree of perturbation d |
| L_o | globally optimal robust sensor placement |
| $Var_d[I_i^d]$ | the mathematical variance of I_i across all possible values of the d coefficient |
| $E_d[I_i^d]$ | the mathematical expectation of I_i across all possible values of the d coefficient |
| g^d | a function that takes as argument a set of sensor locations L and return the total reduced entropy for a given degree of perturbation d |
| V_{L_o} | variation level of the importance values of sensor placement L_o |
| $M_{L_o}^d$ | percentage of mismatched sensors between locally optimal placement L^d and globally optimal robust placement L_o |
| M_{L_o} | the overall percentage of mismatched sensors for all degrees of perturbation |
| Q_{L_o} | performance measure of robust optimal solution L_o |
| K_{L_o} | cost of installing solution L_o |
| U_{L_o} | utility function value of solution L_o |

Table 5.1: Notation for importance of node measure derivation.

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

$$H_i = - \sum_{j \in S_i} p_{ij} \log p_{ij} \quad (5.4)$$

3. Weight the entropy of every node with the number of s that pass through it:

In order to differentiate between nodes that have a high entropy value that have high and low traffic throughput, the entropy of every node is weighed by the number of s utilising it. The importance of node i is defined as:

$$I_i = H_i \sum_{j \in S_i} N_{ij} \quad (5.5)$$

Although the main goal of this work is not to design the physical architecture needed to sense an intersection, an example architecture is provided on Fig. 5.1. Since sensing a certain intersection might require several sensors placed in close proximity it is important to comment on the redundancy of the positions of sensors.

When deploying sensors in order to maximise information about flows or velocities using common methodologies, the measured values are usually highly correlated in space due to the fact that what is being sampled is in fact continuous sequence of road segments that belong to the same road. In the case of sensing, based on the proposed importance of a node measure, the main value that determines the priority of sensing a certain intersection is the collection of turning probabilities at it. It is trivial to observe that those are not correlated in space.

Let us consider an intersection between two main roads. The importance value for this intersection is likely to be high and therefore a group of sensors should be placed there. It is not likely that there is such an intersection in close proximity to this one, since this would be considered as an inefficient design and even if there is another major intersection in close proximity, this would be a topological peculiarity of the road network rather than an intrinsic property of spatial correlation as in the case of sensor placement strategies that maximise information about the flows or average speeds.

Therefore, sensor placement based on importance values of intersections is intrinsically not prone to redundancy of sensor positions. If a situation occurs where two neighbouring intersections both possess a high importance value it is vital to understand that this is not a redundancy issue and in fact both intersection should be sensed in this case since the information acquired

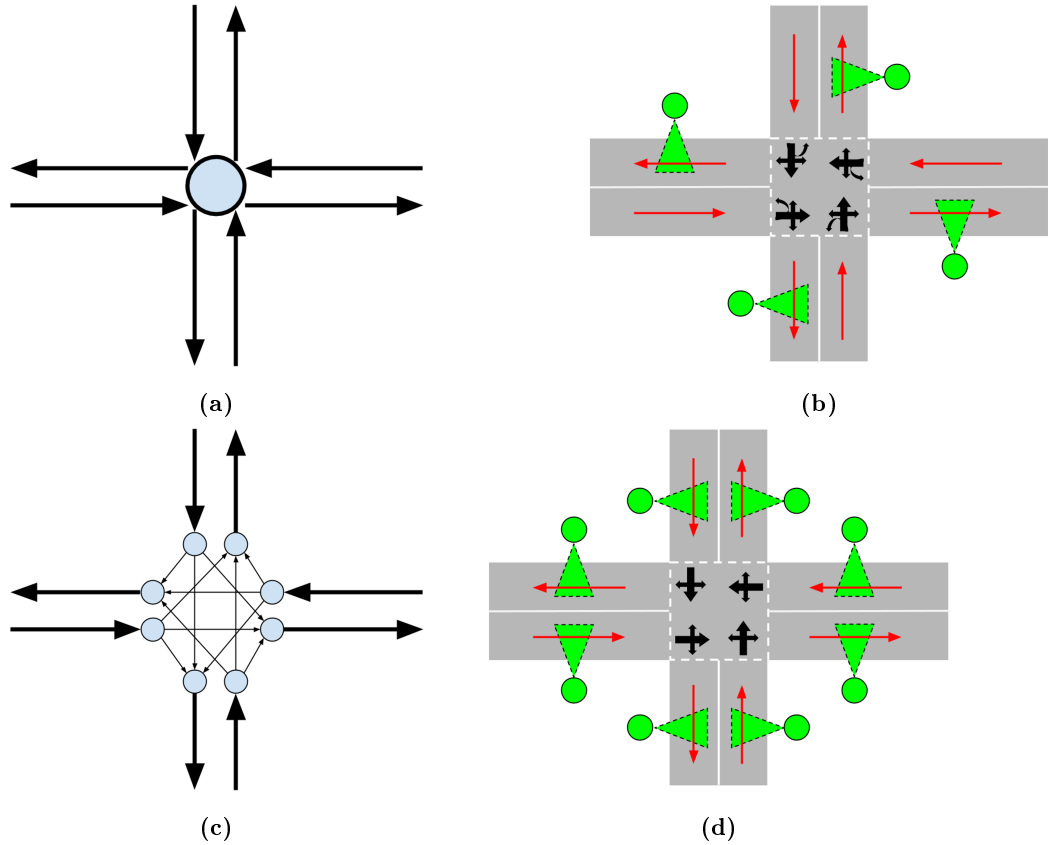


Figure 5.1: Illustration of a suggested positioning strategy for sensors for different types of intersections. On Fig. 5.1a the graph representation of an intersection with all allowed turns is presented. As it can be observed this scenario can be represented with a single node connecting all the edges leading to the intersection and leading out of it. On Fig. 5.1b the green circles represent plate scanning sensors that need to be installed in order to be able to obtain the necessary information to calculate all the turning probabilities. It should be noted that this is the simplest possible case and that the sensors positioned at the intersection do not even need to collect the id of the passing vehicles in order to evaluate the turning probabilities. Fig. 5.1c depicts a slightly more complicated case where U-turns are not allowed. In this case 8 separate nodes are needed in order to represent the intersection using an unidirectional graph. As observed in Fig. 5.1d all edges from and to the intersection need to be observed by the plate scanning sensors in order to extract the needed information to calculate the turning probabilities and the entropy of all nodes.

from each of them is different. Naturally, the edges connecting the two intersections should not be sensed twice.

An example of calculating turning probabilities, entropy and importance of nodes is given in Fig. 5.2. Furthermore, the exact positions of the sensors in the simple example network are shown depending on the number of sensors to be put.

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

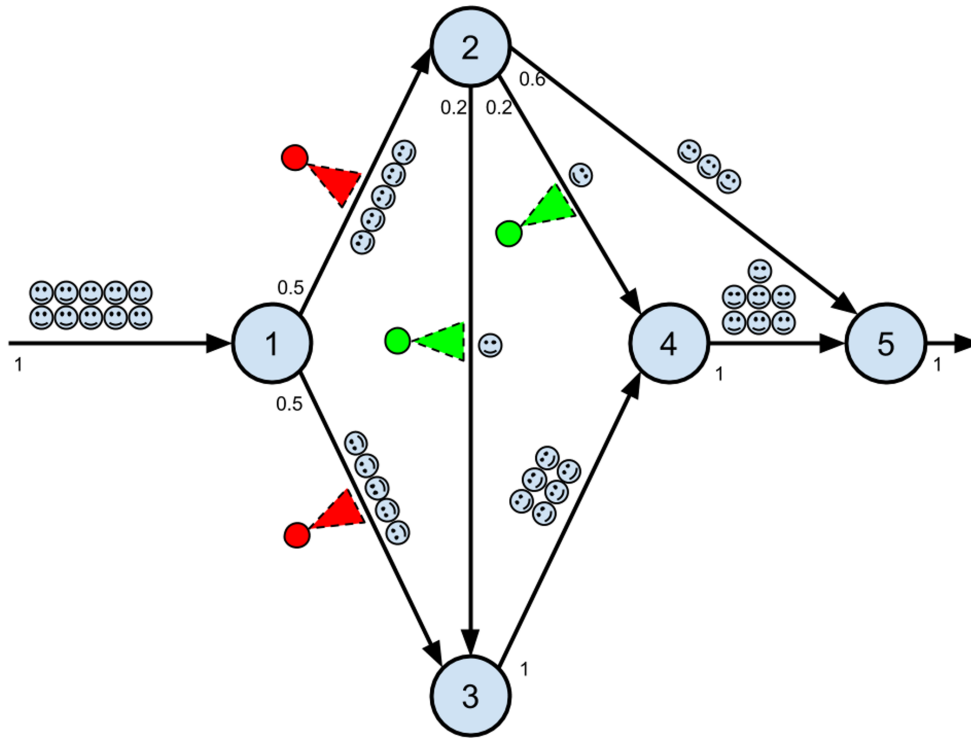


Figure 5.2: Diagram providing an example of calculating importance of nodes. The first thing to do is to calculate the entropy of every node. Nodes 3,4 and 5 have only one successor, which means that there is only one turning probability with value 1, which means that the entropy and therefore the importance of those nodes is 0. As it can be seen on the graphs, no choices are being made at those nodes, therefore they have a small importance value. The entropy of node 1 is calculated using Eq. 2 is $H_1 = 0.3$ and the importance from Eq.3 is $I_1 = 3$. Similarly for the other nodes $H_2 = 0.41$ and $I_2 = 2.05$. Therefore, the nodes that are of interest and in this case have non-zero values are nodes 1 and 2. In case sensors for one intersection are available the most important node (1) is sensed. The precise positions of the sensors are on the links L_{12} and L_{13} , which ensure complete knowledge of the choices made at the intersection. In case one more intersection can be sensed, obviously it should be intersection 2 and the green sensors represent the additional links that should be sensed. Since there already is a sensor placed at L_{12} , only sensors on L_{23} and L_{24} have to be added in order to gain complete information about the routing choices at node 2. Since the entropy of all other nodes is 0 after the placement of the red and green sensors, full information about the network is guaranteed.

5.5 Achieving Robustness Against Changes in the OD Matrix

Changing traffic demands over the course of a day results in the importance value of a node changing as well. Some nodes may experience high importance values during morning rush hour while having lower values during the evening. In case sensors are placed at nodes, whose importance value varies significantly throughout the day, they cannot be moved if some other nodes become more important. This is the reason why the nodes that overall, have the biggest importance values across the day, must be located.

Since it is also important to study the daily variation of importance let us examine the notation describing splitting the day into time-of-day (TOD) intervals:

N_{ij}^t - the number of s that go from node i to node j in period t

H_i^t - the entropy of node i during period t

I_i^t - the importance of node i during period t

Next step is to come up with an importance value representative for the whole day. Some regions of the day are of less interest than others simply because the amount of information that can be extracted is smaller. Typically, the factor that plays the largest role in this case is the amount of traffic. Therefore, the total importance of a node for the whole day is computed, using a weighted average of importance values of the node for different regions of the day. The weight function is governed by the number of s that pass through the node during the respective time region. Then, the overall daily importance of a node can be defined as:

$$\bar{I}_i = \sum_{t=1}^T I_i^t \frac{\sum_{k \in S_i} N_{ik}^t}{\sum_{k \in S_i} N_{ik}} \quad (5.6)$$

The second term in the sum is simply the number of cars that pass through the node throughout time region t over the total number of cars that pass throughout the whole day and T is the number of regions the day is split into. This definition of overall importance puts an emphasis on the nodes that are interesting during the important parts of the day. This weighting is included in order to avoid high importance values throughout periods of time where the node is not being utilised.

5.5 Achieving Robustness Against Changes in the OD Matrix

In reality, apart from the changes in the OD matrix over the course of the day, there is another process that alters the traffic demand in a less intense and more gradual way. This process

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

is a result of long term changes to both the population and the city structure. In order to demonstrate that the method is robust against such type of variations a generic way to “alter” or “perturb” the traffic demand is implemented.

5.5.1 Methodology for Altering the OD matrix

Let every have a list of itineraries, which is composed of separate trips. Every trip has an origin, destination and start time. In most cases an agent takes two trips per day: from home to work in the morning and from work to home in the evening. Let us take two agents. We assume that the first origin and the last destination in the itinerary of those agents is their place of residence. Then exchange those locations, as if the first now lives in the home of the second and vice versa. This is done for a predetermined percentage of all the s . This percentage is referred to as the degree of disturbance. By executing this strategy the number of people starting from or arriving at all the regions is not changed. This means that the intensity of people starting from any region is not changed and the intensity of people arriving at those regions is also not altered. The factor that is perturbed is precisely the OD matrix, since only the intensities of the connections between origins and destinations are varied.

This procedure is visualised in Fig. 5.3

5.5.2 Strategy for Robust Placement

In order to find locations that are optimal for performance and robust against variations in the OD matrix a measure that represents the importance of a set of nodes for different degrees of perturbation has to be found. This is the overall importance of the chosen locations. Every node i has an importance measure \bar{I}_i . Assume that a given number of sensors from all possible locations can be taken and then calculate the total reduced entropy in the network which is:

$$R = \sum_{i=1}^{|L|} \bar{I}_i \quad (5.7)$$

For every different degree of perturbation every node has a calculated importance value \bar{I}_i^d where d is the degree of perturbation.

Let the resulting reduced entropy from a set of locations L for different degrees of perturbation d be calculated by the function $g^d(L) = R$, let the optimal placement for a given degree of perturbation d be L^d and $g^d(L^d) = R^d$

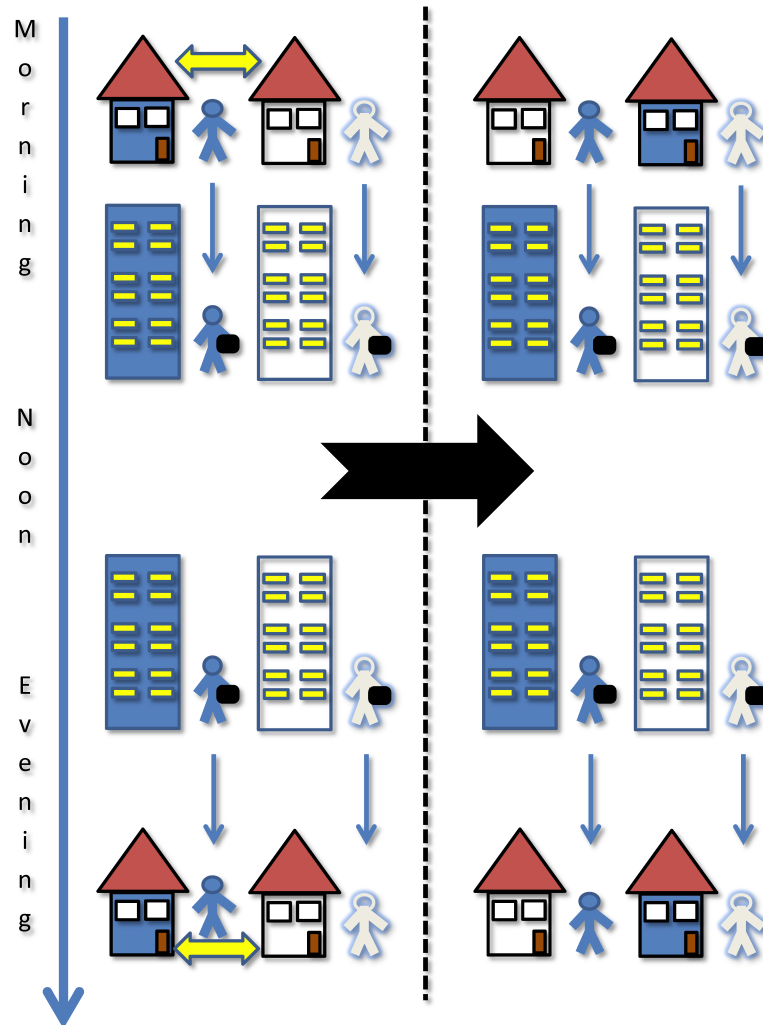


Figure 5.3: Diagram illustrating the exchange of origins or living locations of two s. Both s still have the same work locations however they switch their homes. In this way the OD pairs intensity is changed.

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

An optimal placement L_o is sought, which maximizes the reduced entropy relative to the local maximum across the various perturbations:

$$\max_{L_o} \sum_d \frac{g^d(L_o)}{g^d(L^d)} \quad (5.8)$$

5.5.3 Strategy for Finding the Optimal Number of Sensors

There are four aspects that should be taken into account when designing a utility function to be maximised in order to find the optimal number of sensors.

1. The variation of the importance value across the perturbations:

Every node has a different importance value across the perturbations \bar{I}_i^d . The degree of variation must be evaluated so that globally important nodes are located rather than nodes that have just one high importance value among various degrees of perturbation. In order to do that, the variance for every node i across different degrees of perturbations d : $Var_d[\bar{I}_i^d]$ is calculated. Then it is normalised by the average value across the perturbations so that this measure is comparable to others:

$$\frac{Var_d[\bar{I}_i^d]}{E_d[\bar{I}_i^d]} \quad (5.9)$$

In order to evaluate the total variation level of the importance for a sensor placement the average of the scaled variances is computed for all chosen locations:

$$V_{L_o} = E_i \left[\frac{Var_d[\bar{I}_i^d]}{E_d[\bar{I}_i^d]} \right] \quad (5.10)$$

The number of nodes to be included in the set of optimal locations L_o is varied; this can also be referred to as its cardinality: $|L_o|$. The goal is to minimise the variation of importance of the same node across the degrees of perturbation in order to ensure robustness of the placement.

2. The percentage of mismatched sensors:

Define $M_{L_o}^d$ as the percentage of sensors that are mismatched between the optimal sensor placement for a certain degree of perturbation L^d for a given number of sensors, and the robust optimal solution L_o . This is basically the cardinality of the difference between the two sets divided by the cardinality of the set:

$$M_{L_o}^d = \frac{|L_o \setminus L^d|}{|L_o|} \quad (5.11)$$

Then the overall percentage of mismatched sensors is just the average of this measure across all degrees of perturbation:

$$M_{L_o} = E_d[M_{L_o}^d] \quad (5.12)$$

This is a measure of distance between the optimal solution for d and the robust optimal solution for all degrees of perturbation. It can also be understood as a value signifying the percentage of sensors that need to be moved in order to reach the local optimal solution.

This measure should be minimised to ensure robustness of the placement. In other words, the sensor locations should be as universal as possible.

3. Performance measure of the robust optimal solution compared to the local optimal solutions:

This measure is used to describe how close is the robust optimal solution to perfectly match the locally optimal solutions.

$$Q_{L_o} = \sum_d \frac{g^d(L_o)}{g^d(L^d)} \quad (5.13)$$

This measure should be maximised since maximum performance is sought.

4. Cost of sensors:

A function that punishes high number of sensors is further included. For simplicity a linear function is used, which grows with the increase in number of sensors:

$$K_{L_o} = \alpha |L_o| \quad (5.14)$$

The utility function that needs to be maximised subject to the number of sensors or the cardinality of the set L_o then becomes:

$$\begin{aligned} \max_{|L_o|} U_{L_o} &= w_1 Q_{L_o} - w_2 V_{L_o} - w_3 K_{L_o} - w_4 M_{L_o} , \\ \text{where } \sum_{i=1}^4 w_i &= 1 \end{aligned} \quad (5.15)$$

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

All the separate functions are scaled to assume values between 0 and 1, however depending on the designers choice some measures can be given more weight by varying w_{1-4} . On Fig.5.5.3 all the separate functions and the utility function that determines the optimal sensor number can be seen.

5.6 Optimal Sensor Placement for Singapore Case Study

First, traffic is assigned for a whole day using the methodology described in Chapter 2.

5.6.1 Importance Analysis

At first, the entropies of every node of the network within each of the time periods that the day is split into is calculated. Following this the robustness against daily variations technique is executed. This results in the the overall daily importances of the nodes used to find the optimal sensors placement as described in section 5.4. Fig. 5.4 shows the Singapore road network and the importance of nodes. It can be observed that the sensors cover the city well with accents on the central business district (south central part), the highway intersections and intersection of highways with other large roads. Moreover, there are plenty of sensors in the residential areas(east and north central), which, however, have lower importance values due to the smaller number of cars that go through those intersections.

Next, the change in traffic demand as explained section 5.5.1 is simulated. Fig. 5.5 visualises the results of applying the change. Since it is not practical to visualise all OD pairs, in the visualisation only the intensities of OD pairs that have as origin the university area around the Nanyang Technical University (NTU) in the western part of the city are shown. We can observe the change in the destinations intensities as people increase their trips to the east part of the city while reducing the trips that stay within the western part.

The following step is finding the optimal sensor placement for Singapore that is robust against such type of variations in the OD matrix as described in section 5.5.2. In order to evaluate the performance of the robust placement the following set of actions is executed:

1. For each degree of perturbation run a set of 10 simulations in order to get an averaged value for all the required parameters. The number of simulations is determined so that the degree of variation is below a certain threshold as described in [174].
2. Using the simulation outputs, calculate the turning probabilities, entropies, and importance of all nodes



Figure 5.4: Averaged importance value of nodes for a full day. The sensor placement resulting from those values is optimally robust against daily traffic variations.

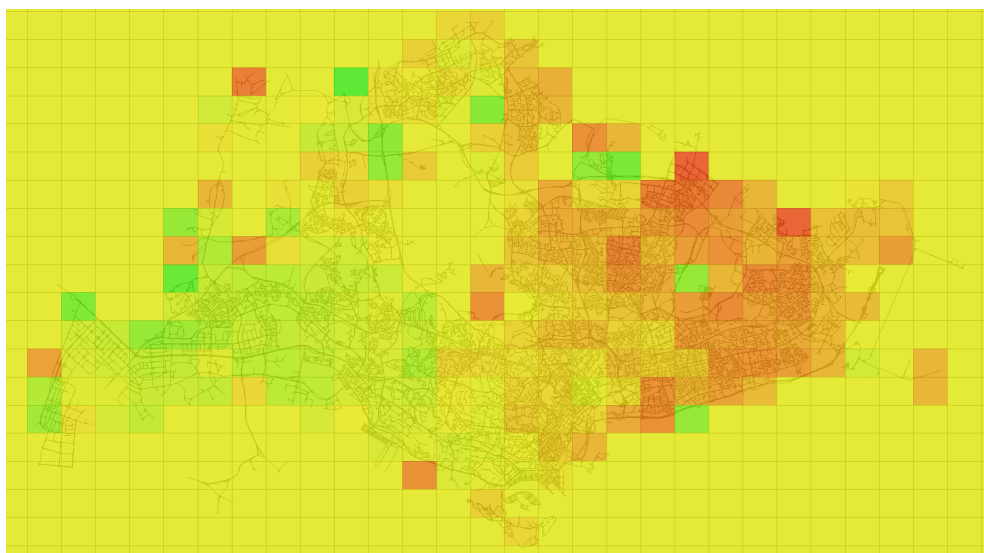


Figure 5.5: Visualising the effects of perturbing the OD matrix. A heat map showing the differences between the OD matrices (original and with 30 % perturbation). We have taken only the pairs that have as destination the area around the Nanyang Technical University (NTU) and plotted the intensities of the different origins. Yellow colouring means intensity has not changed, red means higher intensity in the second OD matrix and green means lower intensity in the second OD matrix.

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

3. Find the optimal placement of sensors for every degree of perturbation.
4. Using the optimisation strategy described above, calculate a robust sensor placement.
5. Compare the performance of the robust sensor placement to the performance of the locally optimal (in the sense of perturbation degree) sensor placements. The performance in this case is the ratio between the total reduced entropy R of the robust placement to the total reduced entropy of the locally optimal placements.

In Fig. 5.6 a comparison of the performance of the optimally robust method versus the locally optimal solutions for sensor placement can be observed.

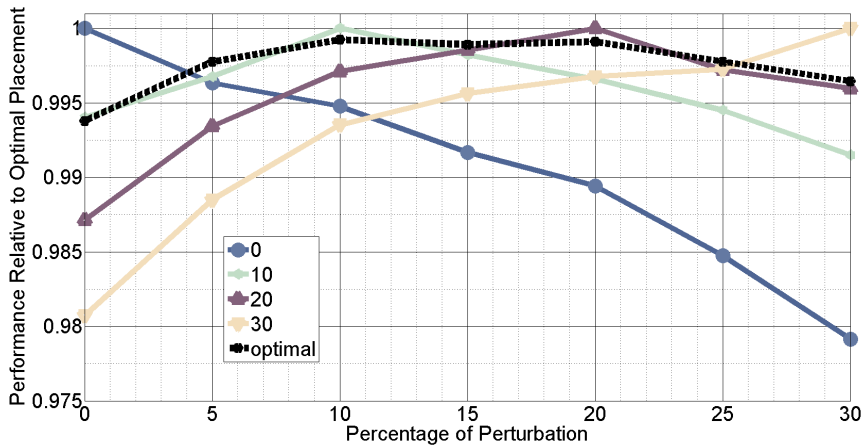


Figure 5.6: Comparison of the performance of the optimally robust method across the various degrees of perturbation for the OD matrix to the performance of the locally optimal sensor placements. It can be observed that the robust placement is performing better in the sense that it is more invariant to changes in the OD matrix. The performance value is stable throughout the changes of the traffic demands.

Following this, the performance of those sets of locations for other degrees of perturbations is computed. For example the blue line on Fig. 5.6 represents the optimal sensor placement for the original OD matrix. Naturally, since the sensor placement was made based on the traffic patterns in this scenario, the performance is 100%. We can then see that the performance of this sensor placement if the traffic is governed by the OD matrix perturbed by 5% decreases. The more the traffic demand is perturbed, the more the performance of the optimal sensor placement calculated from the original OD matrix decreases. The goal of the method is to achieve robustness in the sense that the performance stays consistently high as the OD matrix

5.6 Optimal Sensor Placement for Singapore Case Study

is varied. The performance of the robust sensor placement solution is also depicted on the graph as the black dotted line. It can be observed that the robust solution does not vary that much when the OD matrix is perturbed and is performing better than the rest. It can be concluded that the performance of optimal sensor placements calculated from a specific OD matrix vary more than the performance of the sensor placement strategy that uses not just one OD matrix but rather a set of perturbed variations of it. Moreover, the defined measure of overall importance proves to have very little degree of variation, which makes it a suitable candidate for a robust measure of importance of nodes.

Finally, the optimal number of sensors to be placed in Singapore are computed as described in section 5.5.3. For the sake of simplicity let all the discussed factors be equally important. In Fig. 5.7 the functions related to the process of finding the sensor count are plotted. On the last sub-graph the utility function whose maximum corresponds to the optimal number of sensors to be installed can be observed. For the case of Singapore this number is 582. Surely, if some factors are more important than others, they can be weighted differently and this will affect the optimal number of sensors.

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

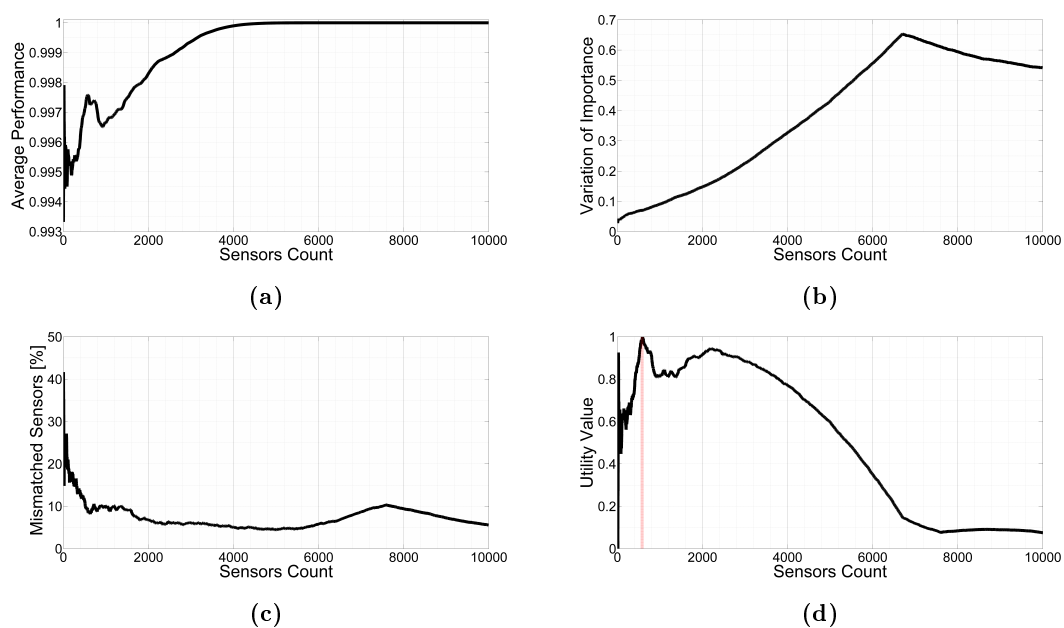


Figure 5.7: Functions that construct the utility function and the utility function itself that need to be maximized by the number of sensors: 5.7a performance of the sensor placement as a function of the sensor number, 5.7b Overall variation of the importance values of the sensor network as a function of sensor count, 5.7c Overall Percentage of mismatched sensors for the sensor network as a function of the sensor number, 5.7d Utility function that maximizes 5.7a and minimizes 5.7b, 5.7c and the number of sensors

5.7 Chapter Summary

5.7.1 Sensor Placement

Chapter 5 points out the need for an importance measure that is able to combine and refine information about the traffic demand contained in the OD matrix and the information about the topology of the network. In this way one can point to the locations in the network that hold the biggest amount of uncertainty related to drivers' routing choices; this is a crucial underlying factor that determines traffic conditions in a network. It has been discussed that nodes should be examined instead of links since the intersections are the places where decisions are made and the roads are the locations where the results of those choices are observed.

A measure of importance is defined, which satisfies the aforementioned conditions using information theory. More precisely, the measure is a combination of the flow through a node and the entropy of the node itself. The novel definition of entropy of a node that is provided in this chapter is dictated by the routing choices drivers made instead of by purely topological factors.

It has been observed that the importance of nodes can vary throughout the day due to changes in traffic patterns, moreover a method is designed to find the most robust sensor placement against such type of changes, which are referred to as short term traffic demand variations. Long term changes are also being addressed by this work. A method is designed to simulate long term city dynamics and their effect on the traffic demand in the city by realistically perturbing the OD matrix.

A technique, which allows designers to weigh various factors connected to their preferences regarding the sensor network and its functionality is designed, in order to determine the optimal number of sensors that need to be placed. The utility function consists of the variation factor of the sensor readings, the average percentage of mismatched sensors under varying traffic demand, the performance and the sensor installation and sustaining cost.

5.7.2 Robustness

The concept of making sensor placement, or any other planned infrastructure robust against such perturbations in the OD matrix is another significant contribution of this work. As pointed out in the introduction this technique is analogous to the addition of noise in the training data of a neural network or another entity in order to increase its performance (typically prediction or pattern recognition accuracy). Planned infrastructure, which is optimal with respect to all

5. ROBUST ROUTE INFORMATION MAXIMISING SENSOR PLACEMENT

various degrees of perturbation of the OD matrix will therefore, be much more robust against changes of this type. Naturally, a method is described in order to find a robust sensor placement against long term OD variations using the OD perturbations, providing certainty that sensors will not have to be moved once they are installed.

The heterogeneity of the nodes' importance as a network characteristic can be of great importance as well, as it is directly proportional to the utilization factor of the transportation network. In case of homogeneous importance values, there is lack of central points at which congestion is created. Homogeneity of the importance measure also means that drivers are evenly spread across the network and utilize fully its infrastructure. Heterogeneity, on the other hand, means that drivers' paths are very similar with the exception of several hub points through which everyone passes. This might bring imbalance of traffic on the network as some roads become congested while others stay empty. Following this argument it might be interesting to use the measure of heterogeneity of the importance measure of a network in order to either evaluate the traffic performance or optimise the routing of commuters leading to overall reduction of congestion.

The work described in this chapter represents the final step of surveillance in the four step suggested transportation system optimization strategy. The reader is encouraged to advance to the next chapter, which provides a general summary of the collection of efforts encompassed in the thesis.

Chapter 6

Conclusion

The final conclusion chapter of this thesis connects all findings in the previous chapters and discusses their purpose and collective impact for road transportation systems optimization. The main subject of discussion in this dissertation is traffic optimization enabled by computational power and efficient algorithms. As the vehicle population around the globe keeps increasing, congestion steadily but surely becomes a major factor leading to significant losses of time and money. Furthermore, inefficiencies in road transportation systems present harmful environmental consequences. The questions that were outlined in the beginning of the dissertation aim at identifying the sources of commuting system inefficiencies. This work also tries to find effective, sustainable ways for eliminating these sources by determining the most beneficial traffic defining components to be influenced and applying state of the art algorithms to control them. Aside from the contributions presented in the thesis there are 5 main messages, which emerge from the collection of results acquired in the various experiments, that can be taken out from this work as suggested directions of future efforts and insights into road transportation system dynamics.

6.1 Main Messages

Several messages can be taken out from the synergies of results acquired in this work. First, Chapter 3 demonstrates that traffic infrastructure is not in agreement with traffic demand, which is a main source of increased congestion levels and worsened traffic conditions. Second, the results from Chapter 3 indicate that there are some highly dynamical and *super-sensitive* locations in a city, which can be identified and optimized in order to achieve beneficial results on a system level. Furthermore, the first part of Chapter 4 demonstrates that small changes in

6. CONCLUSION

specific parts of the road network can have bigger effect on congestion levels of the whole system, further identifying potential steering tools for efficient traffic control. Third, the collective findings in all chapters indicate strongly that controlling the way agents choose their routes is a cost effective method (as no infrastructure needs to be built) for traffic optimization and has the greatest potential in terms of magnitude of improvement of performance. Fourth, as routing choices are the most influential component of a traffic system, surveillance methods, such as sensor placement should be routing-centric, in the sense that the information gain that must be maximised should be defined in terms of the information gathered about the routing choices of the drivers. Chapter 5 points out the need for an importance measure that is able to combine and refine information about the traffic demand contained in the OD matrix and the information about the topology of the network in order to maximize the information gain of a sensor network. Fifth, the concept of adding noise to the OD matrix in the sense of perturbations to traffic demands, presented in Chapter 5, demonstrates promising results and should be used in order to ensure robust planning and control strategies in order to mitigate the effects of fast varying traffic demands.

6.2 Theoretical Implications

The finding in Chapter 4, which allows the BISOS algorithm to bypass the NP hard problem of choosing set of drivers to be rerouted opens the door for more algorithms that seek performance at a very small cost in accuracy of finding the optimum solution. The significant speed up of the algorithm brings theoretical work one step closer to being applied in practice by showing that the relative difference between theoretical optimum and a “shortcut” solution such as BISOS algorithm decreases to nearly negligible values for the large city scenario, which was considered.

The holistic nature of the simulated road transportation system, provides the first demonstration of the existence of Braess paradox in a realistic city scenario. The road closing study in Chapter 4 has further theoretical implications as it presents a technique to be used for identification of both harmful and critical roads. This aspect of the thesis can be perceived as fundamental study of the surface of traffic projection on road networks since the experiment, viewed from a higher level of abstraction, computes the derivative of the traffic distribution with respect to the network.

The projection of traffic onto the road network is also central for the importance measure defined in Chapter 5, which combines the flow through a node and the entropy of the node itself.

The novel definition of entropy of a node captures the combination of demand and topology being dictated by the routing choices drivers made instead of purely topological factors. The main theoretical implication of this is that it captures the projection of traffic onto a topology, thus allowing further classification of locations and criticality analysis.

Apart from deepening theoretical understanding within the field, this work also serves as a conceptual bridge between the fields of transportation and machine learning. The concept of making sensor placement, or any other planned infrastructure robust against perturbations in the OD matrix is another significant contribution of this work. As pointed out in the introduction this technique is analogous to the addition of noise in the training data of a neural network or another entity in order to increase its performance (typically prediction or pattern recognition accuracy). Planned infrastructure, which is optimal with respect to all various degrees of perturbation of the OD matrix will therefore, be much more robust against changes of this type. This concept aims at bridging the gap between machine learning and transportation research by demonstrating that theoretical concepts from one field can be successfully applied in the other.

6.3 Policy Implications

The findings in this work demonstrate that routing control can lead to greater reduction of overall travel time for the commuting population. Therefore, efforts must be concentrated in developing efficient and fast ways to optimally guide the driver population and distribute traffic demand such as the BISOS algorithm, which has been presented. One strong trait of the BISOS algorithm is that the logic supporting it, is founded in actions that can be physically taken, rather than in inapplicable mathematical framework. The approach of providing different information to different drivers, thus enabling traffic load balancing, is viable and provided the high potential impact such a strategy can have, various policies enabling this techniques must be implemented. A key approach of such a task would be to make drivers perceive the marginal costs of their actions, thus self organizing in a socially optimum state.

Furthermore, if excessive traffic congestion is viewed as the marginal costs to society of congestion exceeding the marginal costs of efforts to reduce congestion (such as adding to road or other transport infrastructure), the system optimum routing approach presents a strategy that can both ease traffic conditions and is cost free in the sense that no construction of infrastructure is necessary. In a way the functioning of the BISOS algorithm can be viewed

6. CONCLUSION

as empowering the previously static road infrastructure to present itself in a different way to each of the traffic participants, thus making it dynamical and able to adapt to all possible traffic demand changes. The importance of this realization lies in the fact that routing control is demand invariant as it can adapt instantaneously to any changes presented by the drivers population, while road network changes have to be applied every time a significant change in traffic demand occurs. This further points out to the necessity of new policies that enable and encourage approaches that can affect the way people choose their routes.

6.4 Limitation of the Study and Future Research

One of the main limitations of the study is the amount of data that has been used in order to generate the driver population. As a result of this, the work is concentrated on the methodology of traffic optimization rather than stating with certainty precise changes that have to be implemented in the city's infrastructure. In order to validate the findings from the case study in Singapore, more complete data about the OD pairs and the way people choose their routes is needed. Provided the required data exists, the same experiments should be performed in more cities with varying topological and traffic characteristics, which will further strengthen the arguments made in this work.

The routing model described in Chapter 2 assumes that people behave rationally, which, as it is suggested by real life measurements, is not consistently true. Therefore, a possible and desired future project would be to implement a more elaborate user behaviour model describing both following instructions for optimal routing and route choice itself. In order to validate such a model, however, real data is needed. Furthermore, as it is unrealistic to think that all people might behave rationally and follow a suggested route, a study that determines what percentage of population is needed to follow directions can be performed.

A macroscopic simulation approach is used in this study as the number of separate simulations that need to be run in order to acquire the results is several million. With the increase of computational power available, however, using a micro and even nano-simulation can become feasible sooner rather than later and therefore all the studies can be performed with higher degree of detail in the future.

On a theoretical side, the empirical findings of this thesis should be formalized in a theory of *super-sensitive* locations from dynamic systems point of view, which can enable their fast detection using machine learning techniques.

Future work on the concept of robust planning would require the development of a more comprehensive tool for modelling long term changes in city dynamics, such as building new living or business areas, building new road segments etc. Implementation of such types of changes will bring qualitatively different type of OD matrix variations, since this processes will create completely new origins and destinations. Moreover the population growth should be modelled as well. Furthermore, this thesis deals explicitly with optimizing traffic conditions for vehicles. A natural future step would be to also include pedestrians and incorporate their utility functions as well, especially when dealing with intersection control.

6.5 Final Remarks

Knowing and understanding the general guidelines outlined in this final chapter, can play an important role in fighting the existing problems of increasing population, congestion cost, environmental damages and traffic demand variability. Ensuring efficient, highly intelligent and informed solutions are in place is key to providing the future commuting populations with a robust, sustainable and environment friendlier intelligent transportation system backed up by state of the art optimization algorithms, vast amounts of computing power and socially aware attitude.

6. CONCLUSION

Appendix A

Appendix

Assume that the improvement in total travel time of the removal of a single driver from a road is a Gaussian random variable \mathbf{X} , with mean $\mu_{\mathbf{X}}$ and deviation $\sigma_{\mathbf{X}}$. When a group of N agents is removed from a road, the random variable of the total time saved \mathbf{Y} can be expressed as the sum of all the individual variables:

$$\mathbf{Y} = \sum_{i=1}^N \mathbf{X}_i \quad (\text{A.1})$$

The mean and deviation of the variable \mathbf{Y} are then:

$$\mu_{\mathbf{Y}} = N\mu_{\mathbf{X}} \quad (\text{A.2})$$

$$\sigma_{\mathbf{Y}} = \sqrt{N}\sigma_{\mathbf{X}} \quad (\text{A.3})$$

$$(\text{A.4})$$

We are interested in comparing the probability that each of the two variables \mathbf{X} and \mathbf{Y} is smaller than 0 in which case the examined road segment will not be explored anymore. The probabilities can be calculated in the following way:

$$P(\mathbf{X} < 0) = \int_{-\infty}^0 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \quad (\text{A.5})$$

$$P(\mathbf{Y} < 0) = \int_{-\infty}^0 \frac{1}{\sqrt{N}\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-N\mu)^2}{2N\sigma^2}\right) dy \quad (\text{A.6})$$

$$(\text{A.7})$$

A. APPENDIX

Therefore the cumulative distribution function of the two variables should be examined. In the case of Gaussian distributions the CDF of a function is described using the error function erf and can be written as:

$$F_{\mathbf{X}}(x) = \frac{1}{2} \left[1 + erf \left(\frac{x - \mu_x}{\sqrt{2}\sigma_x} \right) \right] \quad (\text{A.8})$$

$$F_{\mathbf{Y}}(y) = \frac{1}{2} \left[1 + erf \left(\frac{y - N\mu_x}{\sqrt{2}\sqrt{N}\sigma_x} \right) \right] \quad (\text{A.9})$$

$$(\text{A.10})$$

We are interested in comparing the values of $F_{\mathbf{X}}(0)$ and $F_{\mathbf{Y}}(0)$, which are:

$$F_{\mathbf{X}}(0) = \frac{1}{2} \left[1 + erf \left(-\frac{\mu_x}{\sqrt{2}\sigma_x} \right) \right] \quad (\text{A.11})$$

$$F_{\mathbf{Y}}(0) = \frac{1}{2} \left[1 + erf \left(-\frac{N\mu_x}{\sqrt{2}\sqrt{N}\sigma_x} \right) \right] \quad (\text{A.12})$$

$$(\text{A.13})$$

Since the erf is monotonically increasing, we just need to find out, which of the two arguments inside the error function is bigger. Since $\mu_{\mathbf{X}} > 0$, since the examined road is congested, it is apparent that:

$$-\frac{\mu_x}{\sqrt{2}\sigma_x} > -\frac{N\mu_x}{\sqrt{2}\sqrt{N}\sigma_x} \quad (\text{A.14})$$

Therefore, the probability of \mathbf{X} being smaller than zero is higher than the probability of \mathbf{Y} being smaller than 0 and thus the probability to have a non-beneficial sampling sub-iteration is higher when the step size is smaller.

References

- [1] DAVID SCHRANK, BILL EISELE, TIM LOMAX, AND JIM BAK. **Urban Mobility Scorecard**. Technical report, Technical Report August, Texas A&M Transportation Institute and INRIX, Inc, 2015. 1, 2
- [2] NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION ET AL. **Summary of Fuel Economy Performance**. *US Department of Transportation, NHTSA*, 2011. 1, 2
- [3] PRIV-DOZ DR D BRAESS. **Über ein Paradoxon aus der Verkehrsplanung**. *Unternehmensforschung*, **12**(1):258–268, 1968. 7, 78
- [4] FRANK J CESARIO. **Value of time in recreation benefit studies**. *Land economics*, **52**(1):32–41, 1976. 8
- [5] TERENCE C LAM AND KENNETH A SMALL. **The value of time and reliability: measurement from a value pricing experiment**. *Transportation Research Part E: Logistics and Transportation Review*, **37**(2):231–251, 2001. 8
- [6] RÉMY PRUD'HOMME AND JUAN PABLO BOCAREJO. **The London congestion charge: a tentative economic appraisal**. *Transport Policy*, **12**(3):279–287, 2005. 8
- [7] OLA JABALI, T WOENSEL, AND AG DE KOK. **Analysis of travel times and CO2 emissions in time-dependent vehicle routing**. *Production and Operations Management*, **21**(6):1060–1074, 2012. 8
- [8] T. ROUGHGARDEN. *Selfish Routing and the Price of Anarchy*. MIT Press, 2005. 9
- [9] SEBASTIAN ALAND, DOMINIC DUMRAUF, MARTIN GAIRING, BURKHARD MONIEN, AND FLORIAN SCHOPPMANN. **Exact price of anarchy for polynomial congestion games**. In *STACS 2006*, pages 218–229. Springer, 2006. 9

REFERENCES

- [10] JORDAN IVANCHEV, SORINA LITESCU, DANIEL ZEHE, MICHAEL LEES, HEIKO AYDT, AND ALOIS KNOLL. **Determining the Most Harmful Roads in Search for System Optimal Routing**. Technical Report TUM-I1632, Technical University Munich, 2016. 13, 78, 83, 90
- [11] JORDAN IVANCHEV, HEIKO AYDT, AND ALOIS KNOLL. **Spatial and Temporal Analysis of Mismatch between Planned Road Infrastructure and Traffic Demand in Large Cities**. In *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on*, pages 1463–1470. IEEE, 2015. 13, 29
- [12] JORDAN IVANCHEV, HEIKO AYDT, AND ALOIS KNOLL. **On Identifying Dynamic Intersections in Large Cities**. In *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on*, pages 2907–2914. IEEE, 2015. 13, 29
- [13] JORDAN IVANCHEV, HEIKO AYDT, AND ALOIS KNOLL. **Information Maximizing Optimal Sensor Placement Robust Against Variations of Traffic Demand Based on Importance of Nodes**. *IEEE Transactions on Intelligent Transportation Systems*, **17**(3):714–725, March 2016. 13, 121
- [14] STELLA C DAFERMOS AND FREDERICK T SPARROW. **The traffic assignment problem for a general network**. *Journal of Research of the National Bureau of Standards, Series B*, **73**(2):91–118, 1969. 14
- [15] LAND TRANSPORT AUTHORITY. **Annual Vehicle Statistics 2014**, 2014. 14
- [16] ABHINAV SUNDERRAJAN, HEIKO AYDT, WENTONG CAI, AND AARON KNOLL. **Map stream: initializing what-if analyses for real-time symbiotic traffic simulations**. In *Simulation Conference (WSC), 2014 Winter*, pages 2896–2906. IEEE, 2014. 16
- [17] YOSEF SHEFFI. **Urban transportation network**. *Pretince Hall*, 1985. 17, 80, 96, 100, 101, 103, 104
- [18] ANTONIO LIMA, RADE STANOJEVIC, DINA PAPAGIANNAKI, PABLO RODRIGUEZ, AND MARTA C GONZÁLEZ. **Understanding individual routing behaviour**. *Journal of The Royal Society Interface*, **13**(116):20160021, 2016. 20
- [19] A VACCA AND I MELONI. **Route choice analysis using a GPS-based data collection system**. In *Rio de Janeiro, Selected Proceedings of the 13th World Conference on Transportation Research*, 2013. 20

-
- [20] EMILY PARKANY, JIANHE DU, LISA AULTMAN-HALL, AND RYAN GALLAGHER. **Modeling stated and revealed route choice: consideration of consistency, diversion, and attitudinal variables.** *Transportation Research Record: Journal of the Transportation Research Board*, (1985):29–39, 2006. 20
- [21] MOHAMED A ABDEL-ATY, KENNETH M VAUGHN, RYUICHI KITAMURA, PAUL P JOVANNIS, AND FRED L MANNERING. **Models of commuters’ information use and route choice: Initial results based on a southern California commuter route choice survey.** *California Partners for Advanced Transit and Highways (PATH)*, 1993. 20
- [22] OLIVER JAN, ALAN HOROWITZ, AND ZHONG-REN PENG. **Using global positioning system data to understand variations in path choice.** *Transportation Research Record: Journal of the Transportation Research Board*, (1725):37–44, 2000. 20
- [23] DANIEL ZEHE, ALOIS KNOLL, WENTONG CAI, AND HEIKO AYDT. **SEMSim Cloud Service: Large-scale urban systems simulation in the cloud.** *Simulation Modelling Practice and Theory*, pages –, 2015. 21
- [24] ALAN J HOROWITZ. **Delay/volume relations for travel forecasting based upon the 1985 Highway Capacity Manual.** *FHWA Report*, 1991. 24
- [25] SHLOMO BEKHOR, MOSHE E BEN-AKIVA, AND M SCOTT RAMMING. **Evaluation of choice set generation algorithms for route choice models.** *Annals of Operations Research*, **144**(1):235–247, 2006. 24
- [26] STEFAN LÄMMER, BJÖRN GEHLEN, AND DIRK HELBING. **Scaling laws in the spatial structure of urban road networks.** *Physica A: Statistical Mechanics and its Applications*, **363**(1):89 – 95, 2006. Information and Material Flows in Complex Networks Information and Material Flows in Complex Networks. 31
- [27] PIERRE ROBILLARD. **Estimating the O-D matrix from observed link volumes.** *Transportation Research*, **9**(2):123 – 128, 1975. 32, 122
- [28] GILLES DURANTON AND MATTHEW A TURNER. **The fundamental law of road congestion: Evidence from US cities.** *The American Economic Review*, pages 2616–2652, 2011. 32, 67

REFERENCES

- [29] E.A. BEUKES, M.J.W.A. VANDERSCHUREN, AND M.H.P. ZUIDGEEST. **Context sensitive multimodal road planning: a case study in Cape Town, South Africa.** *Journal of Transport Geography*, **19**(3):452 – 460, 2011. Special Issue : Geographic Information Systems for Transportation. 34, 122
- [30] FEI-YUE WANG. **Parallel Control and Management for Intelligent Transportation Systems: Concepts, Architectures, and Applications.** *Intelligent Transportation Systems, IEEE Transactions on*, **11**(3):630–638, Sept 2010. 34, 122
- [31] CHIEN-LIANG FOK, M. HANNA, S. GEE, Tsz-CHIU AU, P. STONE, C. JULIEN, AND S. VISHWANATH. **A Platform for Evaluating Autonomous Intersection Management Policies.** In *Cyber-Physical Systems (ICCPs), 2012 IEEE/ACM Third International Conference on*, pages 87–96, April 2012. 34, 122
- [32] KHALIL M YOUSEF, MAMAL N AL-KARAKI, AND ALI M SHATNAWI. **Intelligent Traffic Light Flow Control System Using Wireless Sensors Networks.** *J. Inf. Sci. Eng.*, **26**(3):753–768, 2010. 34, 122
- [33] Y. LI, D. JIN, P. HUI, Z. WANG, AND S. CHEN. **Limits of Predictability for Large-Scale Urban Vehicular Mobility.** *Intelligent Transportation Systems, IEEE Transactions on*, **15**(6):2671–2682, Dec 2014. 34
- [34] MATTHIAS DEHMER AND ABBE MOWSHOWITZ. **A history of graph entropy measures.** *Information Sciences*, **181**(1):57 – 78, 2011. 36, 39, 125, 129
- [35] ROBERT L. GORDON. **A Technique for Control of Traffic at Critical Intersections.** *Transportation Science*, **3**(4):279–288, 1969. 37
- [36] M. AYMERICH AND A. NOVO. **Madrid critical intersections antiblocking strategies.** In *Road Traffic Monitoring, 1992 (IEE Conf. Pub. 355)*, pages 70–, 1992. 37
- [37] MA DONGFANG, SONG XIANMIN, TAO PENGFEI, AND WANG DIANHAI. **Critical Traffic Volume Warrant of Signal Installing at Equal Weight Intersections.** In *Intelligent Computation Technology and Automation (ICICTA), 2011 International Conference on*, **1**, pages 388–391, March 2011. 37
- [38] CARLOS GERSHENSON AND CENTRUM LEO APOSTEL. **Self-organizing traffic lights.** In *Complex Systems*. Citeseer, 2005. 37

-
- [39] G. STEPHANOPOULOS AND P.G. MICHALOPOULOS. **Real time control of critical intersections based on queue dynamics.** In *Decision and Control including the 17th Symposium on Adaptive Processes, 1978 IEEE Conference on*, pages 1018–1024, Jan 1978. 37
- [40] STEFAN LÄMMER AND DIRK HELBING. **Self-control of traffic lights and vehicle flows in urban road networks.** *Journal of Statistical Mechanics: Theory and Experiment*, **2008(04):P04019**, 2008. 37
- [41] SHAO-XIN YUAN, YI-SHENG AN, XIANG-MO ZHAO, AND HAI-WEI FAN. **Modeling and simulation of dynamic traffic flows at non-signalized T-intersections.** In *Power Electronics and Intelligent Transportation System (PEITS), 2009 2nd International Conference on*, **2**, pages 1–5, Dec 2009. 37
- [42] HEATHER J RUSKIN AND RUILI WANG. **Modeling traffic flow at an urban unsignalized intersection.** In *Computational Science-ICCS 2002*, pages 381–390. Springer, 2002. 37
- [43] LIANGHUI QIAN AND XIAOHONG CHEN. **Peak Volume Characteristics of Urban Road Intersections: Case in Shanghai.** In *Computer and Network Technology (ICCNT), 2010 Second International Conference on*, pages 397–401, April 2010. 37
- [44] J. BUDIARTO, S. SULISTYO, I.W. MUSTIKA, AND A. INFANTONO. **Road density prediction: Updated methods of turning probabilities and highway capacities manual for achieving the best route.** In *Electrical Engineering and Computer Science (ICEECS), 2014 International Conference on*, pages 168–173, Nov 2014. 38
- [45] YU HU AND J HELLENDORRN. **Daily traffic volume modeling based on travel behaviors.** In *Networking, Sensing and Control (ICNSC), 2013 10th IEEE International Conference on*, pages 639–644. IEEE, 2013. 38
- [46] W.A.M. WEIJERMARS, INFRASTRUCTURE NETHERLANDS RESEARCH SCHOOL FOR TRANSPORT, AND LOGISTICS. *Analysis of urban traffic patterns using clustering.* TRAIL thesis series. Netherlands TRAIL Research School, 2007. 38
- [47] ERIK JENELIUS. **Network structure and travel patterns: explaining the geographical disparities of road network vulnerability.** *Journal of Transport Geography*, **17(3):234 – 244**, 2009. 38, 128

REFERENCES

- [48] DARREN M. SCOTT, DAVID C. NOVAK, LISA AULTMAN-HALL, AND FENG GUO. **Network Robustness Index: A new method for identifying critical links and evaluating the performance of transportation networks.** *Journal of Transport Geography*, **14**(3):215 – 227, 2006. 38, 128
- [49] H.W CORLEY AND DAVID Y SHA. **Most vital links and nodes in weighted networks.** *Operations Research Letters*, **1**(4):157 – 160, 1982. 38, 128
- [50] MICHAEL SCHRECKENBERG AND REINHARD SELTEN. *Human behaviour and traffic networks.* Springer, 2004. 38, 128
- [51] SERGIO PORTA, VITO LATORA, FAHUI WANG, SALVADOR RUEDA, EMANUELE STRANO, SALVATORE SCELLATO, ALESSIO CARDILLO, EUGENIO BELLI, FRANCISCO CARDENAS, BERTA CORMENZANA, ET AL. **Street centrality and the location of economic activities in Barcelona.** *Urban Studies*, **49**(7):1471–1488, 2012. 38, 128
- [52] ALEXANDER ERATH, MICHAEL LÖCHL, AND KAY W AXHAUSEN. **Graph-theoretical analysis of the Swiss road and railway networks over time.** *Networks and Spatial Economics*, **9**(3):379–400, 2009. 39, 128
- [53] FENG XIE AND DAVID LEVINSON. **Measuring the Structure of Road Networks.** *Geographical Analysis*, **39**(3):336–356, 2007. 39, 128
- [54] BHANU M YERRA AND DAVID M LEVINSON. **The emergence of hierarchy in transportation networks.** *The Annals of Regional Science*, **39**(3):541–553, 2005. 39, 129
- [55] L DA F COSTA, BAN TRAVENCOLO, MP VIANA, AND E STRANO. **On the efficiency of transportation systems in large cities.** *EPL (Europhysics Letters)*, **91**(1):18003, 2010. 39, 129
- [56] O. PINK AND B. HUMMEL. **A statistical approach to map matching using road network geometry, topology and vehicular motion constraints.** In *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*, pages 862–867, Oct 2008. 39, 129
- [57] FENG XIE AND DAVID LEVINSON. **Topological evolution of surface transportation networks.** *Computers, Environment and Urban Systems*, **33**(3):211 – 223, 2009. 39, 129

-
- [58] EMANUELE STRANO, VINCENZO NICOSIA, VITO LATORA, SERGIO PORTA, AND MARC BARTHÉLEMY. **Elementary processes governing the evolution of road networks.** *Scientific reports*, **2**, 2012. 39, 129
- [59] RP ROESS AND WR MCSHANE. **Changing concepts of level of service in the 1985 highway capacity manual: some examples.** *ITE Journal*, **57**(5), 1987. 45
- [60] LAURENZ WISKOTT AND TERRENCE J SEJNOWSKI. **Slow feature analysis: Unsupervised learning of invariances.** *Neural computation*, **14**(4):715–770, 2002. 47
- [61] S.P. BOYD AND L. VANDENBERGHE. *Convex Optimization.* Berichte über verteilte messsysteme. Cambridge University Press, 2004. 68
- [62] STEFAN LÄMMER, BJÖRN GEHLEN, AND DIRK HELBING. **Scaling laws in the spatial structure of urban road networks.** *Physica A: Statistical Mechanics and its Applications*, **363**(1):89–95, 2006. 77
- [63] J. IVANCHEV, D. ZEHE, V. VISWANATHAN, S. NAIR, AND A. KNOLL. **BISOS: Backwards Incremental System Optimum Search algorithm for fast socially optimal traffic assignment.** In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 2137–2142, Nov 2016. 78
- [64] E.N. LORENZ. *The Essence Of Chaos.* Jessie and John Danz lectures. Taylor & Francis, 1995. 79
- [65] STEVEN D GRIBBLE. **Robustness in complex systems.** In *Hot Topics in Operating Systems, 2001. Proceedings of the Eighth Workshop on*, pages 21–26. IEEE, 2001. 79
- [66] YANG-YU LIU, JEAN-JACQUES SLOTINE, AND ALBERT-LÁSZLÓ BARABÁSI. **Controllability of complex networks.** *Nature*, **473**(7346):167–173, 2011. 79
- [67] STEFAN LÄMMER AND DIRK HELBING. **Self-control of traffic lights and vehicle flows in urban road networks.** *Journal of Statistical Mechanics: Theory and Experiment*, **2008**(04):P04019, 2008. 79
- [68] C. GERSHENSON. *Design and Control of Self-organizing Systems.* Copit-Arxives, 2007. 79

REFERENCES

- [69] SEUNG-BAE COOLS, CARLOS GERSHENSON, AND BART D HOOGHE. **Self-organizing traffic lights: A realistic simulation.** In *Advances in applied self-organizing systems*, pages 41–50. Springer, 2008. 79
- [70] GIOVANNI PETRI, H JELDTOFT JENSEN, AND JOHN W POLAK. **Global and local information in traffic congestion.** *EPL (Europhysics Letters)*, **88**(2):20010, 2009. 79
- [71] JAVED ASLAM, SEJOON LIM, AND DANIELA RUS. **Congestion-aware traffic routing system using sensor data.** In *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*, pages 1006–1013. IEEE, 2012. 79, 81
- [72] SEJOON LIM, HARI BALAKRISHNAN, DAVID GIFFORD, SAMUEL MADDEN, AND DANIELA RUS. **Stochastic motion planning and applications to traffic.** *The International Journal of Robotics Research*, page 0278364910386259, 2010. 79
- [73] SEJOON LIM AND DANIELA RUS. **Stochastic motion planning with path constraints and application to optimal agent, resource, and route planning.** In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 4814–4821. IEEE, 2012. 79
- [74] SEJOON LIM AND DANIELA RUS. **Congestion-aware multi-agent path planning: distributed algorithm and applications.** *The Computer Journal*, **57**:825–839, 2013. 79
- [75] SALVATORE SCCELLATO, LUIGI FORTUNA, MATTIA FRASCA, JESÚS GÓMEZ-GARDEÑES, AND VITO LATORA. **Traffic optimization in transport networks based on local routing.** *The European Physical Journal B*, **73**(2):303–308, 2010. 79
- [76] SORINA LITESCU, VAISAGH VISWANATHAN, MICHAEL LEES, ALOIS KNOLL, AND HEIKO AYDT. **Information impact on transportation systems.** *Journal of Computational Science*, **9**:88–93, 2015. 79, 81
- [77] W. BARFIELD AND T.A. DINGUS. *Human Factors in Intelligent Transportation Systems.* Taylor & Francis, 2014. 79
- [78] DANIEL ENEMARK, MATHEW D MCCUBBINS, AND NICHOLAS WELLER. **Knowledge and networks: An experimental test of how network knowledge affects coordination.** *Social Networks*, **36**:122–133, 2014. 79

-
- [79] VERENA FRANKEN AND TRANSPORTATION RESEARCH BOARD. **Use of navigation systems and consequences for travel behaviour**. In *ECTRIFEHRLFERSI Young Researcher Seminar, Brno, Czech Republic*, 2007. 79
- [80] HYEJIN YOUN, MICHAEL T GASTNER, AND HAWOONG JEONG. **Price of anarchy in transportation networks: efficiency and optimality control**. *Physical review letters*, **101**(12):128701, 2008. 80
- [81] TIM ROUGHGARDEN. **How unfair is optimal routing?** In *Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 203–204. Society for Industrial and Applied Mathematics, 2002. 80
- [82] ERIC J. FRIEDMAN. **A generic analysis of selfish routing**. In *in Proceedings of IEEE CDC*, 2004. 80
- [83] TIM ROUGHGARDEN. **The price of anarchy is independent of the network topology**. *Journal of Computer and System Sciences*, **67**(2):341–364, 2003. 80
- [84] TIM ROUGHGARDEN. **The price of anarchy in games of incomplete information**. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 862–879. ACM, 2012. 80
- [85] ELLIOT ANSHELEVICH, ANIRBAN DASGUPTA, JON KLEINBERG, EVA TARDOS, TOM WEXLER, AND TIM ROUGHGARDEN. **The price of stability for network design with fair cost allocation**. *SIAM Journal on Computing*, **38**(4):1602–1623, 2008. 80
- [86] MARTIN BECKMANN, CB MCGUIRE, AND CHRISTOPHER B WINSTEN. **Studies in the Economics of Transportation**. Technical report, 1956. 80, 102
- [87] MARISKA VAN ESSEN, TOM THOMAS, ERIC VAN BERKUM, AND CASPAR CHORUS. **From user equilibrium to system optimum: a literature review on the role of travel information, bounded rationality and non-selfish behaviour at the network and individual levels**. *Transport Reviews*, pages 1–22, 2016. 80
- [88] TIM ROUGHGARDEN AND ÉVA TARDOS. **How bad is selfish routing?** *Journal of the ACM (JACM)*, **49**(2):236–259, 2002. 81

REFERENCES

- [89] XIAOLEI GUO, HAI YANG, AND TIAN-LIANG LIU. **Bounding the inefficiency of logit-based stochastic user equilibrium.** *European Journal of Operational Research*, **201(2)**:463–469, 2010. 81
- [90] CHRISTOS PAPADIMITRIOU. **Algorithms, games, and the internet.** In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 749–753. ACM, 2001. 81
- [91] DIMITRIS FOTAKIS, ALEXIS C KAPORIS, AND PAUL G SPIRAKIS. **Efficient methods for selfish network design.** *Theoretical Computer Science*, **448**:9–20, 2012. 81
- [92] SATORU FUJISHIGE, MICHEL GOEMANS, TOBIAS HARKS, BRITTA PEIS, AND RICO ZENKLUSEN. **Congestion games viewed from M-convexity.** *Operations Research Letters*, **43(3)**:329–333, 2015. 81
- [93] HS MAHMASSANI, TY HU, S PEETA, AND A ZILIASKOPOULOS. **Development and testing of dynamic traffic assignment and simulation procedures for ATIS.** Technical report, ATMS applications. Technical Report DTFH61-90-R-0074-FG, Center for Transportation Research, University of Texas at Austin, 1994. 81
- [94] MOSHE BEN-AKIVA, MICHEL BIERLAIRE, JON BOTTOM, HARIS KOUTSOPOULOS, AND RABI MISHALANI. **Development of a route guidance generation system for real-time application.** In *Proceedings of the IFAC Transportation Systems 97 Conference, Chania*. Citeseer, 1997. 81
- [95] C CHARBONNIER, JL FARGES, AND JJ HENRY. **MODELS AND STRATEGIES FOR DYNAMIC ROUTE GUIDANCE. PART C, OPTIMAL CONTROL APPROACH.** In *DRIVE Conference (1991: Brussels, Belgium). Advanced telematics in road transport*, 1991. 81
- [96] G BECCARIA AND ALFREDO BOLELLI. **Modelling and assessment of dynamic route guidance: the MARGOT project.** In *Vehicle Navigation and Information Systems Conference (3rd). Vehicle navigation & information systems: conference record of papers*, 1992. 81
- [97] ISAM KAYSI, MOSHE BEN-AKIVA, AND ANDRE DE PALMA. **Design aspects of advanced traveler information systems.** In *Urban Traffic Networks*, pages 59–81. Springer, 1995. 81

-
- [98] P. G. TAYLOR N. G. BEAN, F. P. KELLY. **Braess's Paradox in a Loss Network.** *Journal of Applied Probability*, **34**(1):155–159, 1997. 81
- [99] DIETRICH BRAESS, ANNA NAGURNEY, AND TINA WAKOLBINGER. **On a paradox of traffic planning.** *Transportation science*, **39**(4):446–450, 2005. 81
- [100] TIM ROUGHGARDEN. **On the severity of Braess's paradox: designing networks for selfish users is hard.** *Journal of Computer and System Sciences*, **72**(5):922–953, 2006. 81
- [101] NG BEAN, FP KELLY, AND PG TAYLOR. **Braess's paradox in a loss network.** *Journal of Applied Probability*, pages 155–159, 1997. 81
- [102] GAL CHECHIK, ISAAC MEILIJSON, AND EYTAN RUPPIN. **Synaptic pruning in development: a computational account.** *Neural computation*, **10**(7):1759–1777, 1998. 82
- [103] GINA KOLATA. **The New York Times.** [online], 2015. Available at <http://www.nytimes.com/1990/12/25/health/what-if-they-closed-42d-street-and-nobody-noticed.html>. 82
- [104] W. KNÖDEL. *Graphentheoretische Methoden und ihre Anwendungen.* Ökonometrie und Unternehmensforschung Econometrics and Operations Research. Springer Berlin Heidelberg, 2013. 82
- [105] SALLY CAIRNS, STEPHEN ATKINS, AND PHIL GOODWIN. **Disappearing traffic? The story so far.** *Proceedings of the ICE-Municipal Engineer*, **151**(1):13–22, 2002. 82
- [106] JN HAGSTROM AND RA ABRAMS. **Characterizing Braess paradox for traffic networks.** *Links*, **4**(1.32):1–17, 2001. 82
- [107] MG PALA, S BALTAZAR, PENG LIU, HERMANN SELLIER, BENOIT HACKENS, F MARTINS, VINCENT BAYOT, X WALLART, L DESPLANQUE, AND SERGE HUANT. **Transport inefficiency in branched-out mesoscopic networks: An analog of the Braess paradox.** *Physical review letters*, **108**(7):076802, 2012. 82
- [108] XIN HUANG, ASUMAN E OZDAGLAR, AND DARON ACEMOGLU. **Efficiency and Braess' Paradox under pricing in general networks.** *Selected Areas in Communications, IEEE Journal on*, **24**(5):977–991, 2006. 82

REFERENCES

- [109] KOOHYUN PARK. **Detecting braess paradox based on stable dynamics in general congested transportation networks.** *Networks and Spatial Economics*, **11**(2):207–232, 2011. 82
- [110] SAEED ASADI BAGLOEE, AVISHAI CEDER, MADJID TAVANA, AND CLAIRE BOZIC. **A heuristic methodology to tackle the Braess Paradox detecting problem tailored for real road networks.** *Transportmetrica A: Transport Science*, **10**(5):437–456, 2014. 82
- [111] TIM ROUGHGARDEN. **On the severity of Braess’s paradox: designing networks for selfish users is hard.** *Journal of Computer and System Sciences*, **72**(5):922–953, 2006. 82
- [112] MOSHE BEN-AKIVA, ANDRE DE PALMA, AND ISAM KAYSI. **The impact of predictive information on guidance efficiency: An analytical approach.** In *Advanced methods in transportation analysis*, pages 413–432. Springer, 1996. 82
- [113] I.A. KAYSI. *Framework and Models for the Provision of Real-time Driver Information.* Massachusetts Institute of Technology, 1992. 82
- [114] MARIN LUJAK, STEFANO GIORDANI, AND SASCHA OSSOWSKI. **Route guidance: Bridging system and user optimization in traffic assignment.** *Neurocomputing*, **151**:449–460, 2015. 82
- [115] MAHMOOD ZANGUI, YAFENG YIN, SIRIPHONG LAWPHONGPANICH, AND SHIGANG CHEN. **Differentiated congestion pricing of urban transportation networks with vehicle-tracking technologies.** *Transportation Research Part C: Emerging Technologies*, **36**:434–445, 2013. 83
- [116] WILLIAM S VICKREY. **Congestion theory and transport investment.** *The American Economic Review*, **59**(2):251–260, 1969. 83
- [117] ERIC J GONZALES AND CARLOS F DAGANZO. **Morning commute with competing modes and distributed demand: user equilibrium, system optimum, and pricing.** *Transportation Research Part B: Methodological*, **46**(10):1519–1534, 2012. 83
- [118] MALACHY CAREY AND DAVID WATLING. **Dynamic traffic assignment approximating the kinematic wave model: System optimum, marginal costs, externalities and tolls.** *Transportation Research Part B: Methodological*, **46**(5):634–648, 2012. 83

-
- [119] CARLOS F DAGANZO. **The cell transmission model, part II: network traffic.** *Transportation Research Part B: Methodological*, **29**(2):79–93, 1995. 83
- [120] CARLOS F DAGANZO. **A finite difference approximation of the kinematic wave model of traffic flow.** *Transportation Research Part B: Methodological*, **29**(4):261–276, 1995. 83
- [121] YING-JU CHEN AND JIAWEI ZHANG. **Design of price mechanisms for network resource allocation via price of anarchy.** *Mathematical programming*, **131**(1-2):333–364, 2012. 83
- [122] TOBIAS HARKS, INGO KLEINERT, MAX KLIMM, AND ROLF H MÖHRING. **Computing network tolls with support constraints.** *Networks*, **65**(3):262–285, 2015. 83
- [123] HAI YANG, WEI XU, BING-SHENG HE, AND QIANG MENG. **Road pricing for congestion control with unknown demand and cost functions.** *Transportation Research Part C: Emerging Technologies*, **18**(2):157–175, 2010. 83
- [124] MARGUERITE FRANK AND PHILIP WOLFE. **An algorithm for quadratic programming.** *Naval research logistics quarterly*, **3**(1-2):95–110, 1956. 95, 98
- [125] SRINIVAS PEETA AND HANI S MAHMASSANI. **System optimal and user equilibrium time-dependent traffic assignment in congested networks.** *Annals of Operations Research*, **60**(1):81–113, 1995. 98, 101
- [126] LARRY J LEBLANC, RICHARD V HELGASON, AND DAVID E BOYCE. **Improved efficiency of the Frank-Wolfe algorithm for convex network programs.** *Transportation Science*, **19**(4):445–462, 1985. 98
- [127] MICHAEL FLORIAN, JACQUES GUÁLAT, AND HEINZ SPIESS. **An efficient implementation of the Partan variant of the linear approximation method for the network equilibrium problem.** *Networks*, **17**(3):319–339, 1987. 98
- [128] Y AREZKI AND D VAN VLIET. **A full analytical implementation of the PARTAN/Frank–Wolfe algorithm for equilibrium assignment.** *Transportation Science*, **24**(1):58–62, 1990. 98

REFERENCES

- [129] OLAF JAHN, ROLF H MÖHRING, ANDREAS S SCHULZ, AND NICOLÁS E STIER-MOSES. **System-optimal routing of traffic flows with user constraints in networks with congestion.** *Operations research*, **53**(4):600–616, 2005. 101, 102
- [130] T LEVENTHAL, G NEMHAUSER, AND L TROTTER JR. **A column generation algorithm for optimal traffic assignment.** *Transportation Science*, **7**(2):168–176, 1973. 101
- [131] M. PATRIKSSON. *The Traffic Assignment Problem: Models and Methods*. Dover Publications, 2015. 101
- [132] JOSÉ R CORREA AND NICOLÁS E STIER-MOSES. **Wardrop equilibria.** *Wiley encyclopedia of operations research and management science*, 2011. 101
- [133] ROBERT B DIAL. **A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration.** *Transportation Research Part B: Methodological*, **40**(10):917–936, 2006. 101
- [134] M FLORIAN, I CONSTANTIN, AND D FLORIAN. **A New Look at the Projected Gradient Method for Equilibrium Assignment. CD-ROM.** *Transportation Research Board of the National Academics, Washington, DC*, 2009. 101
- [135] GUIDO GENTILE. **Linear User Cost Equilibrium: a new algorithm for traffic assignment.** *submitted to Transportation Research B*, 2009. 101
- [136] HILLEL BAR-GERA. **Traffic assignment by paired alternative segments.** *Transportation Research Part B: Methodological*, **44**(8):1022–1046, 2010. 101
- [137] YU MARCO NIE. **A cell-based Merchant–Nemhauser model for the system optimum dynamic traffic assignment problem.** *Transportation Research Part B: Methodological*, **45**(2):329–342, 2011. 101
- [138] DEEPAK K MERCHANT AND GEORGE L NEMHAUSER. **A model and an algorithm for the dynamic traffic assignment problems.** *Transportation science*, **12**(3):183–199, 1978. 101
- [139] GREGOR LÄMMEL AND GUNNAR FLÖTTERÖD. **Towards system optimum: Finding optimal routing strategies in time-dependent networks for large-scale evacuation problems.** In *KI 2009: Advances in Artificial Intelligence*, pages 532–539. Springer, 2009. 101

-
- [140] ATHANASIOS K ZILIASKOPOULOS. **A linear programming model for the single destination system optimum dynamic traffic assignment problem.** *Transportation science*, **34**(1):37–49, 2000. 101
- [141] JOSÉ R CORREA, ANDREAS S SCHULZ, AND NICOLÁS E STIER-MOSES. **Selfish routing in capacitated networks.** *Mathematics of Operations Research*, **29**(4):961–976, 2004. 102
- [142] JOSÉ R. CORREA, ANDREAS S. SCHULZ, AND NICOLÁS E. STIER MOSES. *Computational Complexity, Fairness, and the Price of Anarchy of the Maximum Latency Problem*, pages 59–73. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. 102
- [143] HONGLIANG GUO, ZHIGUANG CAO, JIE ZHANG, DUSIT NIYATO, AND ULRICH FASTENRATH. **Routing multiple cars in large scale networks: Minimizing road network breakdown probability.** In *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*, pages 2180–2187. IEEE, 2014. 102
- [144] XIQUN MICHAEL CHEN, LEI ZHANG, XIANG HE, CHENFENG XIONG, AND ZHIHENG LI. **Surrogate-Based Optimization of Expensive-to-Evaluate Objective for Optimal Highway Toll Charges in Transportation Network.** *Computer-Aided Civil and Infrastructure Engineering*, **29**(5):359–381, 2014. 102
- [145] BALRAM SUMAN AND PRABHAT KUMAR. **A survey of simulated annealing as a tool for single and multiobjective optimization.** *Journal of the operational research society*, pages 1143–1160, 2006. 106
- [146] LEONARD EVANS AND ROBERT HERMAN. **Automobile fuel economy on fixed urban driving schedules.** *Transportation Science*, **12**(2):137–152, 1978. 114
- [147] DC BIGGS AND R AKCELIK. **Estimation of car fuel consumption in urban traffic.** In *Proc. 13th ARRB Conf*, **13**, pages 123–132, 1986. 114
- [148] RAHMI AKCELIK AND MARK BESLEY. **Operating cost, fuel consumption, and emission models in aaSIDRA and aaMOTION.** In *25th conference of australian institutes of transport research (CAITR 2003)*, pages 1–15, 2003. 114

REFERENCES

- [149] WALEED F FARIS, HESHAM A RAKHA, RAED ISMAIL KAFIFY, MOUMEN IDRES, AND SALAH ELMOSELHY. **Vehicle fuel consumption and emission modelling: an in-depth literature review.** *International Journal of Vehicle Systems Modelling and Testing*, **6**(3-4):318–395, 2011. 114
- [150] HESHAM A RAKHA, KYOUNGHO AHN, KEVIN MORAN, BART SAERENS, AND ERIC VAN DEN BULCK. **Virginia tech comprehensive power-based fuel consumption model: model development and testing.** *Transportation Research Part D: Transport and Environment*, **16**(7):492–503, 2011. 114
- [151] ENNIO CASCETTA. **Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator.** *Transportation Research Part B: Methodological*, **18**(4):289–299, 1984. 122
- [152] HAI YANG, TSUNA SASAKI, YASUNORI IIDA, AND YASUO ASAKURA. **Estimation of origin-destination matrices from link traffic counts on congested networks.** *Transportation Research Part B: Methodological*, **26**(6):417–434, 1992. 122
- [153] B.S. KERNER. *Introduction to Modern Traffic Flow Theory and Control: The Long Road to Three-Phase Traffic Theory.* Springer Berlin Heidelberg, 2009. 122
- [154] SEJOON LIM AND DANIELA RUS. **Stochastic distributed multi-agent planning and applications to traffic.** In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 2873–2879. IEEE, 2012. 123
- [155] M. GENTILI AND P.B. MIRCHANDANI. **Locating sensors on traffic networks: Models, challenges and research opportunities.** *Transportation Research Part C: Emerging Technologies*, **24**(0):227 – 255, 2012. 126
- [156] IWAO OKUTANI AND YORGOS J STEPHANEDES. **Dynamic prediction of traffic volume through Kalman filtering theory.** *Transportation Research Part B: Methodological*, **18**(1):1–11, 1984. 126
- [157] M CREMER AND H KELLER. **A new class of dynamic methods for the identification of origin-destination flows.** *Transportation Research Part B: Methodological*, **21**(2):117–132, 1987. 126

-
- [158] CHUNG-CHENG LU, XUESONG ZHOU, AND KUILIN ZHANG. **Dynamic origin-destination demand flow estimation under congested traffic conditions.** *Transportation Research Part C: Emerging Technologies*, **34**:16–37, 2013. 126
- [159] LUCIO BIANCO, GIUSEPPE CONFESSORE, AND PIERFRANCESCO REVERBERI. **A Network Based Model for Traffic Sensor Location with Implications on O/D Matrix Estimates.** *Transportation Science*, **35**(1):50–60, 2001. 126
- [160] ENRIQUE CASTILLO, JOSÉ MARÍA MENÉNDEZ, AND PILAR JIMÉNEZ. **Trip matrix and path flow reconstruction and estimation based on plate scanning and link observations.** *Transportation Research Part B: Methodological*, **42**(5):455 – 481, 2008. 126, 127
- [161] XUESONG ZHOU AND GEORGE F. LIST. **An Information-Theoretic Sensor Location Model for Traffic Origin-Destination Demand Estimation Applications.** *Transportation Science*, **44**(2):254–273, 2010. 126
- [162] R. MÍNGUEZ, S. SÁNCHEZ-CAMBRONERO, E. CASTILLO, AND P. JIMÉNEZ. **Optimal traffic plate scanning location for OD trip matrix and route estimation in road networks.** *Transportation Research Part B: Methodological*, **44**(2):282 – 298, 2010. 126
- [163] T. TOLEDO AND T. KOLECHKINA. **Estimation of Dynamic Origin Destination Matrices Using Linear Assignment Matrix Approximations.** *Intelligent Transportation Systems, IEEE Transactions on*, **14**(2):618–626, June 2013. 126
- [164] XIAOPENG LI AND YANFENG OUYANG. **Reliable sensor deployment for network traffic surveillance.** *Transportation Research Part B: Methodological*, **45**(1):218 – 231, 2011. 127
- [165] YONG LI, XIANGMING ZHU, DEPENG JIN, AND DAPENG WU. **Multiple Content Dissemination in Roadside-Unit-Aided Vehicular Opportunistic Networks.** *Vehicular Technology, IEEE Transactions on*, **63**(8):3947–3956, Oct 2014. 127
- [166] Y. LI, D. JIN, P. HUI, AND S. CHEN. **Contact-Aware Data Replication in Roadside Unit Aided Vehicular Delay Tolerant Networks.** *Mobile Computing, IEEE Transactions on*, **PP**(99):1–1, 2015. 127

REFERENCES

- [167] ANDREAS KRAUSE AND CARLOS GUESTRIN. **Near-optimal Nonmyopic Value of Information in Graphical Models.** *CoRR*, abs/1207.1394, 2012. 127
- [168] TAO XING, XUESONG ZHOU, AND JEFFREY TAYLOR. **Designing heterogeneous sensor networks for estimating and predicting path travel time dynamics: An information-theoretic modeling approach.** *Transportation Research Part B: Methodological*, **57**(0):66 – 90, 2013. 127
- [169] HANBIAO WANG, KUNG YAO, AND DEBORAH ESTRIN. **Information-theoretic approaches for sensor selection and placement in sensor networks for target localization and tracking.** *Communications and Networks, Journal of*, **7**(4):438–449, 2005. 127
- [170] YANG LIU, NING ZHU, SHOUFENG MA, AND NING JIA. **Traffic sensor location approach for flow inference.** *IET Intelligent Transport Systems*, **9**(2):184–192, 2014. 127
- [171] ANDREAS KRAUSE, AJIT SINGH, AND CARLOS GUESTRIN. **Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies.** *J. Mach. Learn. Res.*, **9**:235–284, June 2008. 127
- [172] NING ZHU, YANG LIU, SHOUFENG MA, AND ZHENGBING HE. **Mobile Traffic Sensor Routing in Dynamic Transportation Systems.** *Intelligent Transportation Systems, IEEE Transactions on*, **15**(5):2273–2285, Oct 2014. 127
- [173] ZHEN GUO, MENGCHU ZHOU, AND GUOFEI JIANG. **Adaptive sensor placement and boundary estimation for monitoring mass objects.** *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, **38**(1):222–232, 2008. 128
- [174] LIANYU CHU, HENRY X LIU, JUN-SEOK OH, AND WILL RECKER. **A calibration procedure for microscopic traffic simulation.** In *Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE*, **2**, pages 1574–1579. IEEE, 2003. 140