Leveraging connected vehicle data for user-centred and equitable urban traffic control strategies

Roozbeh Mohammadi
Leveraging connected vehicle data for user-centred and equitable urban traffic control strategies

Roozbeh Mohammadi

A doctoral thesis completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Engineering, at a public examination held at the lecture hall Otakaari 1, A123 A1 on 2 August 2022 at 12 PM.

Aalto University
School of Engineering
Department of Built Environment
Transport and Traffic Operations Research Group
Supervising professors
Assistant professor Claudio Roncoli
Department of Built Environment
Aalto University School of Engineering
Finland

Thesis advisors
Assistant professor Miloš Mladenović
Department of Built Environment
Aalto University School of Engineering
Finland

Preliminary examiners
Professor Klaus Bogenberger, Technical University of Munich, Germany
Professor Aleksandar Stevanovic, University of Pittsburgh, Pennsylvania, US

Opponent
Professor Klaus Bogenberger, Technical University of Munich, Germany
For many decades, urban traffic management systems have been vehicle-dominated. That is not only because of a lack of attention to users, but conventional data collection tools are powerless to collect individual vehicle data as well as vehicle users data. Connected vehicles (CV), as an emerging technology, can collect and transmit real-time vehicle and its users data. This ability facilitates the development of user-centred traffic management strategies in urban transport networks. However, there are some challenges yet to be addressed to convert raw CV data to efficient input for traffic controls. Moreover, achieving a fully connected environment is not possible in near future due to various limitations. Accordingly, this dissertation aims at developing a traffic management strategy based on CV data that improves user-related performance measures at signalized intersections. Furthermore, this dissertation assesses the effect of CV data accuracy on traffic controllers and presents a method to compensate lack of CVs in the urban environment to deploy in traffic management strategies.

In this dissertation, we research two vital aspects of traffic signal control which are signal timing optimization and data. For signal timing optimization, First, using CV data, We develop a user-based signal timing optimization strategy where the objective of the controller is to maximize the user throughput of a signalized intersection. Second, We present a user-based Transit signal priority strategy where the objective of the controller is to reduce users average delay and bus scheduled delay by providing priority for buses that are behind the schedule and with a higher number of passengers on board. Moreover, secondary effects of the current transit priority systems and the proposed transit signal priority are compared, by considering the concept of total social cost. In the data section, first, the impact of CV data accuracy on the performance of signal controllers is investigated. Second, We develop a data-driven vehicle estimation method to make limited CV data usable for a signal controller.

The results of this dissertation show that implementing proper signal timing optimisation-based CVs data improves user and vehicle performance measures at signalized intersections. Moreover, a CV-based transit signal propriety that considers users of buses as well as other motorists can improve current Transit signal priority strategies while the delay of other motorists would not be increased. Moreover, the proposed transit signal priority strategy can reduce other social costs such as emission and fuel consumption. According to this dissertation's findings, data collection tools' accuracy can affect signal timing performance in some circumstances. Furthermore, the potential of a data-driven method to compensate lack of CVs has been presented in this dissertation.

Keywords Traffic control, connected vehicle, Traffic management, Data driven traffic state estimation, Traffic signal timing, Transit signal priority,
I am not sure that when you see this dissertation, it is 2022 or 2050 but when I am writing this preface, we are passing an awful worldwide pandemic that lasted for around two years, and there is a war on the east side of Europe, petrol price is growing fast and robots recently started to deliver grocery in Helsinki. In the second year of my PhD study, 176 innocent people were killed in *Ukraine International Airlines Flight 752* in the sky of Tehran. I cannot never forget this disaster and accordingly, I would like to dedicate this dissertation to the victims and their loved ones.

If you had asked me five years ago that in which country you will get your PhD?, Finland was not my answer at all. If you ask me now, I have a concrete answer: “Finland”. I do not like to call it an accident or a chance because I tried hard to find a good PhD position. I found Professor Claudio Roncoli’s profile interesting and asked him about any opportunity in his research group. He replied that I will open a position next week that you may apply for it. Professor Claudio Roncoli was not only my supervisor but also he was one of my first friends in Finland. I believe that his vision and his supervision taught me many crucial tools and skills but at the same time, he gave me enough freedom to find my path. I have learned from him that I can be relaxed but serious and motivated and I can enjoy my work but prepare a high-quality work for approaching deadlines. Thank you for all opportunities that you have provided for me.

I would like to give a big thanks to my advisor Professor Miloš N Mladenović. His advice, especially in the first two years of my PhD, was really helpful. I learned from him to have a critical view and not be naive about technology and advertisement. Moreover, He shared many useful resources with me and also helped me in improving my English writing skills.

My sincerest thanks go to the pre-examiner of this dissertation, Professor Aleksandar Stevanovic for inspiration and the approval he gave for my dissertation and I would like to give a enormous thanks to Professor Klaus Bogenberger for not only agreeing to act as a pre-examiner but also to act as an opponent in the public defence. I appreciate the time you gave to review this dissertation and also your presence in my
dissertation public examination.

**Dr Shaya Vosough** and **Samira dibaj**, our reunion here at Aalto after 10 years was one of the best memory I have from my PhD study and I appreciate your presence here and the good times we had in the office and during the lunch times. Thank you also **Dr Shaya Vosough** for helping me and being my co-author in my last paper. Moreover, thanks to all my other lovely office-mates who made this journey more pleasant for me: **Dr Sipetas Charalampos**, **Dr Cafer Avci**, **Dr Weiming Zhao**, **Ze Zhou**, **Serio Agriesti**, **Francesco Vitale**, **Oya Duman** and **Babak Fooladi**. Moreover, I am very thankful for **Dr Sanaz Bozorg** and **Farzam Tajdari** who helped me a lot in my first days at the office and Finland. I give a big thank you to **Milad Omidi** who helped me in my research and also he has always been a very good friend and also a great tennis partner.

My dear parents, you have been my first teachers and my first friends in life. **Shahrzad**, I learned from you to be passionate about my goals but I learned from my father **Mostafa** to be calm in my path. This can be a magic mix for success. You were with me during these four years even from thousands of kilometres away. I appreciate your support and you have initiated anything I have achieved.

Lastly, **Mahroo**, I am very lucky and happy to have you in my life. You gave me self-confidence and you kept me motivated in the up and down of my PhD. You endured my unhappy face when my paper was rejected or when I had bad experimental results while you had many things to be worried about them in your PhD study. I am also very proud of you and your great job in your PhD.

Espoo, July 1, 2022,

Roozbeh Mohammadi
Contents

Preface 2

Contents 4

List of Publications 6

Author’s Contribution 7

List of Figures 8

List of Tables 9

Abbreviations 10

1. Introduction 11
   1.1 Research objectives ............................ 13
   1.2 Structure of the dissertation ....................... 13

2. Literature review and research gap 15
   2.1 User-centred traffic control in a CV environment .... 15
   2.2 Impact of CV data quality on traffic controller performance 17
   2.3 Traffic state estimation in a partially CV environment .. 19
   2.4 Research gaps .................................. 21

3. Research methodology 23
   3.1 How CVs data can be implemented in user-centred and equitable traffic management strategies? (RQ1) .......... 23
   3.2 How CVs can enhance the abilities of the current TSP strategies in improving social and environmental performance measures? (RQ2) ................................. 25
   3.3 What is the impact of CV data accuracy on the performance of different CV-based traffic controllers? (RQ3) ........ 25
   3.4 How limited CVs data can be used to estimate traffic state in a low penetration rate of CVs? (RQ4) ................. 26
4. Scientific contribution

4.1 Publication I ......................................................... 28
4.2 Publication II ....................................................... 30
4.3 Publication III ...................................................... 31
4.4 Publication IV ...................................................... 31

5. Discussion and future directions

5.1 Summary of the outcomes ........................................ 33
5.2 Transferability and practicality .................................. 34
5.3 Recommendations for future research .......................... 35

References ................................................................. 36

Publications .............................................................. 43
This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.


**Publication II** Mohammadi, Roozbeh; Shaghayegh Vosough; Roncoli, Claudio. User-based transit signal priority in a connected vehicle environment accounting for schedule delay and social cost evaluation. *Journal of Intelligent Transportation Systems*, Under review.


Author’s Contribution

Publication I: “Signalised intersection control in a connected vehicle environment: User throughput maximisation strategy”

Study conception and design: Roozbeh Mohammadi, Claudio Roncoli and Miloš N. Mladenović; programming and data collection: Roozbeh Mohammadi; analysis and interpretation of results: Roozbeh Mohammadi and Claudio Roncoli; manuscript preparation: Roozbeh Mohammadi, Claudio Roncoli and Miloš N. Mladenović;

Publication II: “User-based transit signal priority in a connected vehicle environment accounting for schedule delay and social cost evaluation”

Study conception and design: Roozbeh Mohammadi, Shaghayegh Vosough and Claudio Roncoli; programming and data collection: Roozbeh Mohammadi; analysis and interpretation of results: Roozbeh Mohammadi, Shaghayegh Vosough and Claudio Roncoli; manuscript preparation: Roozbeh Mohammadi, Shaghayegh Vosough and Claudio Roncoli;

Publication III: “Assessment of connected vehicle information quality for signalised traffic control”

Study conception and design: Héctor Del Pino Verona, Roozbeh Mohammadi and Claudio Roncoli; programming and data collection: Héctor Del Pino Verona and Roozbeh Mohammadi; analysis and interpretation of results: Héctor Del Pino Verona, Roozbeh Mohammadi and Claudio Roncoli; manuscript preparation: Héctor Del Pino Verona, Roozbeh Mohammadi and Claudio Roncoli;

Publication IV: “Towards data-driven vehicle estimation for signalised intersections in a partially connected environment”

Study conception and design: Roozbeh Mohammadi and Claudio Roncoli; programming and data collection: Roozbeh Mohammadi; analysis and interpretation of results: Roozbeh Mohammadi and Claudio Roncoli; manuscript preparation: Roozbeh Mohammadi and Claudio Roncoli;
List of Figures

2.1 Relationship of research questions, research gaps and publications of this dissertation ............... 22

3.1 Simulation framework setup ........................................ 24
3.2 Process framework for assessing the impact of GPS error on traffic signal control ............................. 26
3.3 Simulation framework setup ........................................ 27
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Summary of research on traffic state estimation using CV data</td>
<td>20</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary of contributions and findings of all publications</td>
<td>29</td>
</tr>
</tbody>
</table>
Abbreviations

**ANN** Artificial neural network

**CV** Connected vehicle

**RAIM** Receiver autonomous integrity monitoring

**GPS** Global Positioning System

**TSP** Transit signal priority

**UBSTO** user-based signal time optimisation
1. Introduction

Intersections are the main potential bottlenecks of traffic flow in urban areas, where usually traffic congestion is formed. Accordingly, intersections are controlled by traffic control signals to reduce overall vehicle delay, increase the level of safety and minimise the environmental impact of congestion [74]. Implementing the traffic signals is the most common way to control the congestion at intersections. Accordingly, various traffic signal systems have been developed to improve the functionality of intersections e.g, fix-timed controller, actuated controller and adaptive controller [34]. According to the U.S. Department of Transportation report, in 2004, there were more than 330000 traffic signals in U.S. [33].

Traditionally, the essential input of a traffic controller is provided by a discrete point detector or an area detection system, using technologies such as inductive loops or video image processing [58]. However, implementations of these data collection tools have several deficiencies. First, the installation and maintenance cost of these technologies is noticeably high [84]. Second, point detectors can not detect a vehicle unless the vehicle passes a predetermined position. Third, some types of tools, e.g., radar and traffic camera may lose their functionality in various exogenous factors, including adverse weather conditions, improper lighting, or signal interference [1].

Implementation of the aforementioned data collection tools has caused two main challenges for traffic controllers. Firstly, recently developed adaptive traffic controllers can not perform effectively in exceptional circumstances such as over-saturated data conditions when sufficient information might not be provided for the controller by a point data collector. Secondly, as only vehicle data can be transformed to the controller by these tools, the traffic control strategies are mainly vehicle-centred, focusing on maximising vehicular throughput or minimising vehicular delay, stops, or queue lengths. However, control strategies that apply the transit signal priority (TSP) strategies consider the different vehicle types and differentiate public transit users from private vehicle users. Nevertheless, there are several typical shortcomings in TSP strategies, include increasing
total network delay, inability to handle conflicting priority requests, and negative externalises for non-transit users [20].

Connected vehicle (CV), as an emerging technology, enables many opportunities for traffic control signal systems [75]. CVs can provide sufficient information for vehicle control systems as well as traffic signal controllers. The main benefit of CV data rather than conventional collection tools for traffic control is that CVs can transmit the information at any location [71] while data collection by the conventional tools is limited to the location of the detector. Accordingly, novel traffic control methods are being developed based on CV data [31]. In addition to vehicle data, CV technology may provide user-related information. For instance, number of users of a vehicle can be collected via seat belt fastness Detection. However, such data have not been used to its full extent for the development of user-centred control strategies [60].

Despite the benefits of using CV data in traffic control, there are some challenges yet to be addressed. The accuracy of CV data is highly dependent on the embedded sensors of CVs data such as the positioning system. Consequently, transmitted data from CVs to the controller may not be accurate enough. This error can adversely affect the controller performance. However, the impact can differ based on the structure of the controller. Another challenge with using CV data is the penetration rate of the CVs in traffic flow. In fact, by 2030, it is expected the penetration rate of 40% and 62% for CVs [75]. Accordingly, implementing control strategies based on a fully CV environment would be a challenge if data of conventional vehicles can not be imputed effectively, considering different CV-based traffic control strategies. Accordingly, there is a need to develop novel methods to compensate low penetration rate of CVs during this transition period to a fully CV environment. Various methods are developed to provide the required data for CV-based traffic controls, considering the low penetration of CVs. Recently, data-driven methods have been gaining popularity in traffic engineering applications thanks to their ability to identify patterns and correlations by learning them from available data [65]. Data-driven can identify complex relations between the parameters in some cases where model-driven methods are powerless. However, for training an effective data-driven model considerable amount of data is needed.

Emerging of CVs will improve the efficiency of current urban traffic management systems. For this purpose, first, the potential of CVs in implementing user-centred strategies instead of vehicle-centred strategies is yet to be studied. Second, deficiencies of the current signal control system, i.e., TSP, can be solved by unique features of CVs. Third, the impact of CV data accuracy on the performance of signal timing has to be studied for different types of controllers. Fourth, in transition time to a fully CV environment, Complementary methods have to be applied to compensate for incomplete CVs data in traffic flow.
1.1 Research objectives

In this dissertation, we develop and assess traffic signal control strategies in a CV environment. In contrast to conventional vehicle-based traffic controllers, this dissertation proposes user-based signal timing optimisation in a CV environment. Accounting for users in traffic signal control can include both public transit users and passenger car users where we measure the traffic indicators and the social and environmental impacts. However, implementation of an effective traffic signal control is not possible unless sufficient data is available. Accordingly, we aim at four main primary objectives in this dissertation. The first objective is to develop a user-based signal timing optimisation based on the ability of CVs in the transmission of user data as well as vehicle data. The second objective is to solve current challenges of TSP strategies by accounting for bus schedule delay, other motorist delay, and social and environmental impacts in the signal timing optimisation and evaluation. The third objective of this study is to assess the impact of CV data accuracy on the performance of different CV-based control methods. As the last objective, this dissertation studies potentials of data-driven estimation methods to provide sufficient input for the traffic signal controller in a low penetration rate of CVs.

In way of the mentioned objectives, four research questions have been defined, which will be addressed in this dissertation:

- **Research question 1 (RQ1):** How CVs data can be exploited by developing user-centred and equitable traffic management strategies?

- **Research question 2 (RQ2):** How CVs can enhance the abilities of the current TSP strategies in improving social and environmental performance measures?

- **Research question 3 (RQ3):** What is the impact of CV data accuracy on the performance of different CV-based traffic controllers?

- **Research question 4 (RQ4):** How CVs data can be used to estimate traffic state in a low penetration rate of CVs?

1.2 Structure of the dissertation

This dissertation is built upon four peer-reviewed publications. In each publication, research objectives and the relevant research question have been addressed. RQ1 is studied and answered in Publication I and Pub-
lication II while RQ2 is particularly addressed in Publication II. RQ3 is studied in Publication III where the effect of CV data accuracy on signal timing performance is studied. Publication IV tries to answer RQ4 by developing a data-driven vehicle estimation model based on limited CV data.

The outline of this dissertation is as follows. Chapter 2 presents the state-of-the-art of CV-based traffic controllers, CV-based TSP, CV data quality and traffic state estimation in a CV environment. Chapter 3 presents the proposed methodology to address identified research gaps. In Chapter 4, the scientific contribution of this dissertation is presented, by elaborating on each publication. Finally, Chapter 5 offers a summary and a conclusion to the discussed research as well as suggestions for future research.
2. Literature review and research gap

2.1 User-centred traffic control in a CV environment

Based on the advanced communication capabilities and accurate detection technologies embedded in CVs, varying traffic signal controllers can be developed, such as improved signal arterial coordination, TSP, or signal-vehicle coupled control [49, 41, 86, 12, 32, 68].

However, even with the possibility of transmitting vehicles user data by CV, novel traffic control systems are mainly developed to improve vehicle-related performance measures. Recently, some research has tried to include user perspective in the design of traffic signal control. Accordingly, person delay has been introduced as a metric, mainly as a component to improve TSP strategies in a multi-modal signal timing optimisation. TSP is a signal timing strategy to decrease the travel time of public transit vehicles at a signalised intersection that leads to schedule adherence and increasing public transport reliability [23]. Typically, there are three types of TSP in literature which are a passive priority, active priority and adaptive priority [53]. In passive priority, bus presence is not considered but the priority is provided based on the bus arrival flow rate distribution or bus schedule. On the other hand, in the implementation of active control, the prescence of the bus is vital. In this case, priority is provided to public transit based on the estimation of bus arrival time to the intersection from the time that the bus has been detected. The third type of TSP is an adaptive signal priority. The main difference between the adaptive priority with passive and active priority is that first, detection in adaptive priority continues and is not limited to a predetermined spot. Second, adaptive priority tries to reduce all other motorist delays as well public transit vehicle [51]. In general, TSP can enhance public transportation in various ways, such as improving bus schedule reliability, mitigating transit vehicle emissions, and increasing public transit attractiveness for users [24]. However, ignoring other motorists to implement TSP strategies
may cause some typical shortcomings such as high total network delay, conflicting priority requests, and negative impact on non-transit users [20]. Moreover, even in adaptive TSP, the number of users on the non-transit vehicle has not been considered by traffic control.

Besides the effectiveness of TSP strategies in improving intersection performance, TSP strategies might be able to improve environmental and social performance measures. For example, [54] assessed the impact of the TSP strategies on the intersection of emission production and fuel consumption. However, finding the most effective TSP strategy is not simple, as TSP signal timing optimisation is a multi-criteria decision-making problem where many objectives should be evaluated, e.g., travel time, schedule delay, emission, and fuel consumption. Monetizing all objectives would be a practical solution to make different strategies comparable. The concept of total social cost is a practical way to calculate the total cost of transportation including travel time, emission and fuel consumption. Total social cost of transportation has been considered by several works [28, 22, 57, 43]. For example, [13] proposed a total social cost function to calculate the total social cost of congestion in freeways composed of time, fuel, pollution emissions, and on-road exposure components. In the field of traffic control, [66] assessed the effect of the transition between signal timing plans to minimise the social costs (delays, air pollution emissions and fuel consumption). To the best of our knowledge, no study has evaluated the total social cost of traffic controls, in particular for TSP strategies.

CVs abilities can address some of the mentioned challenges of TSP. In particular, a controller has been proposed by Hu et al. to solve the conflicting priority requests while priority-eligible modes such as public transport vehicles are equipped with V2I connection and passenger cars can be detected by point data collection tools. The proposed controller can reduce bus delays by up to 24.9% and pedestrian delays by up to 14% in high traffic flow. The method has been extended adding green time re-allocation and signal coordination for isolated and coordinated intersections. The developed controller minimises total passenger delay considering buses and passenger car occupancy. [37, 38, 39]. Wu et al. [36] could decrease bus delay up to 24.2% by developing an integrated optimisation approach that suggests a speed to transit vehicle by considering bus idling times at stops and signal timing. [85]. A series of works focused on the development of traffic signal control strategies for minimising person delay in a conventional traffic control framework. Firstly, a real-time adaptive TSP for an isolated intersection was proposed by Christofa et al. [18]. This controller can calculate signal settings to grant priority for buses while total person delay is minimised. Additionally, the controller minimises the adverse impact of other motorist delays by considering vehicles occupancy in the optimisation problem. in [16, 17] by improving the mathematical formulation for unsaturated conditions, adding further experimental sce-
narios, and considering other performance measures, such as the number of stops and emissions. As a further step, Christofa et al. [15] extended the same framework to arterial conditions, considering multiple public transport lines travelling in conflicting directions and platooned vehicle arrivals. In addition, they introduced a weight factor to consider adherence of public transport vehicles to the schedule. This framework has been implemented in some subsequent studies for flexible cycle lengths [90], for phase rotation [89], and for evaluating public transport preferential treatments strategy [25]. Note that, in all these works, the underlying model predicts the delay for each of the lane groups and not for each vehicle, thus the total person delay was calculated by multiplying average vehicle occupancy by total delay. Hence, the method is not developed based on CV capabilities and is assumed to work with data from conventional sources such as inductive loop detectors, for all vehicles, and automated vehicle location systems for public transport vehicles.

Zeng et al. [91] considered individual vehicle occupancy to minimise the total person delay, assuming a CV environment with cars and buses. Two separate models for queuing vehicles and arriving vehicles were proposed, to predict the arrival time of each vehicle at the intersection and calculate delay. Considering average occupancy for each passenger car and bus, the proposed framework can reduce person delay up to 11% and bus delay up to 39% compared to SYNCHRO optimisation in undersaturated traffic conditions. This model also showed decent performance with a low CV penetration rate. However, the model performance was measured in an undersaturated condition and using a fixed signal cycle time. In another research, a passenger-based adaptive controller was proposed [14]. The control mechanism extends the green time by considering the arrival time of vehicles, the number of passengers, as well as pollution and fuel consumption. In addition to signal timing optimisation, the concept of users has been considered to increase the capacity of intersections from passengers' perspective in [55] for the design of lane markings, exclusive bus lanes, and passive bus priority signal settings.

2.2 Impact of CV data quality on traffic controller performance

Collecting traffic data by CVs may have other challenges. The essential source of CVs data is Global Positioning Systems (GPS). Data obtained via GPS might be not perfect. GPS position accuracy depends on five factors: ionospheric errors, tropospheric errors, signal obstruction and multi-path errors, the geometric configuration of satellite errors, and other minor errors [11]. GPS devices can be categorized into three different types based on the level of accuracy: High Accuracy GPS, Standard GPS with a Receiver Autonomous Integrity Monitoring (RAIM), and mobile
GPS. High Accurate GPS relies on ground stations to obtain very high position accuracy. It has been demonstrated that high accurate GPS networks can reach centimetre-level accuracy. However, installing three or more fixed stations in the system is vital, which makes high accurate GPS networks a very costly source of data\[11\]. RAIM is a technology to assess the integrity of GPS signals that allows for evaluating the GPS signal data consistency and reducing positioning errors \[42\]. As a result, it is commonly integrated into safety-critical GPS applications, such as in aviation or marine navigation. Finally, mobile GPS is an inexpensive technology, which is frequently used in many devices, such as smartphones. The standard GPS accuracy is annually collected and published by the US government \[70\]. It is fairly known that mobile GPS is less accurate than traditional GPS receivers. Nonetheless, they can still represent a great opportunity since it is reasonable to assume that, even nowadays, there is at least a mobile GPS device in every vehicle, resulting in a potentially large availability of data. Some data accuracy was presented in \[72\] for several mobile devices.

Considering the required input type for different the traffic signal controllers, we categorize into two types: aggregated input controller and disaggregated input controllers. The conventional traffic signal controllers that are currently used are aggregated input controllers, where the controllers require aggregated data, such as queue lengths and vehicle flows. Various aggregated input controllers are currently being implemented on signalized intersections such as fully actuated controllers, semi-actuated controllers and Max-Pressure or back-pressure controllers \[64, 82\]. Max pressure requires as input the queue length on all approaches of the intersection. Max pressure controller has been widely studied as an effective adaptive signal control due to its simple implementation, lower communication requirements, and computational burdens \[82, 30, 45\]. On the other hand, a disaggregated input controller is assumed to operate with information on individual vehicles as input. CV capabilities are functional to deploy disaggregated input controllers since real-time and accurate data of vehicles are transmitted via V2I (vehicle-to-infrastructure) communication systems. Improved detection and communication capabilities offered by CVs provide various opportunities for controlling signalized intersections, such as improved signal arterial coordination, Transit signal priority or signal-vehicle coupled control \[50, 86, 31, 61\]. The aim of these controllers is usually to maximise vehicle delay or maximise vehicle throughput by using CV accurate and real-time data to estimate the arrival time of each vehicle to the intersection. However, it is reasonable to assume that the performance of these controllers is highly dependent on the accuracy of the data. GPS error in collecting vehicle position data could lead to inaccurate vehicle arrival time estimation or inaccurate queue length estimation. However, such a decrease in signal performance efficiency as a function
of CV data quality has not been widely studied. Existing relevant work includes, for example, the influence of the loss of information in communications [87].

2.3 Traffic state estimation in a partially CV environment

Several estimation methods are being developed to compensate for the limited number of CVs\(^1\) during the transition period before reaching a fully connected environment. On the one hand, being the required input for many traffic signal control methods, queue length estimation has been the focus of several previous works. As an example, [69] proposed a queue length estimation method by integrating CV data with shockwave theory, using data mining techniques; the robustness of the method was tested on the NGSIM data, showing promising accuracy in queue length estimation. The recent study [78] developed a cycle-based queue length estimation by fusing historical and real-time data from CVs, using a maximum likelihood estimation method. Other studies in this area include [21, 9, 47, 92, 81]. Another stream of works dealt with the estimation of total vehicle counts, which include both queuing and moving vehicles. The issue of incomplete CV data availability is typically addressed by applying data fusion techniques that integrate infrastructure-based sensor data and CV data. For instance, [44] proposed a method to fuse traffic camera data and CV data to estimate traffic state in urban streets; while [73] employed a data fusion method considering CVs and loop detectors, where to solve the problem of low CV penetration rate, a probability-based approach is applied to estimate the position of the queue tail. Similar data fusion approaches have been employed in other studies, including, e.g., [19, 48]. More recently, in [95] a traffic volume estimation method was proposed by assuming a time-dependent Poisson process with a constant arrival flow rate of vehicles, while the authors of [94] estimated queue length and traffic volume by applying probability theory on CV data. Furthermore, a Kalman filter-based method was developed in [6] for vehicle count estimation at signalised intersections, relying only on CV data; the method is applied for a system where the traffic flow conservation equation is used as a state equation, while the measurement equation is defined based on hydrodynamic relations of traffic flow. Moreover, a multi-lane vehicle estimation method has been proposed in [4].

Despite the novelty and effectiveness of the aforementioned methods, some limitations may prevent their usage in practice. First, methods delivering only queue estimation may not provide sufficient inputs for some

\(^1\)Some of the cited papers utilise the term “probe vehicle” instead of “connected vehicle”; as in this dissertation we do not deal with automation, we consider these two terms as interchangeable. Therefore, for the sake of consistency, we are using the term CV throughout the entire paper.
Literature review and research gap

Table 2.1. Summary of research on traffic state estimation using CV data

<table>
<thead>
<tr>
<th>Research work</th>
<th>Estimated quantities</th>
<th>Spatial resolution</th>
<th>Time resolution</th>
<th>Utilised data</th>
<th>Estimation (main) method</th>
<th>Validation data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ramezani et al. [69]</td>
<td>queue profile</td>
<td>link</td>
<td>signal cycle</td>
<td>only CV data</td>
<td>shockwave analysis; data mining</td>
<td>real data</td>
</tr>
<tr>
<td>Zheng et al. [95]</td>
<td>traffic volumes</td>
<td>lane</td>
<td>10 min - 1 h</td>
<td>and signal status; traffic volume</td>
<td>maximum likelihood</td>
<td>real data</td>
</tr>
<tr>
<td>Zhao et al. [94]</td>
<td>queue length; traffic volume</td>
<td>link</td>
<td>1 h</td>
<td>only CV data</td>
<td>probability theory</td>
<td>real and simulated data</td>
</tr>
<tr>
<td>Ramezani et al. [69]</td>
<td>queue profile</td>
<td>link</td>
<td>signal cycle</td>
<td>only CV data</td>
<td>shockwave analysis; data mining</td>
<td>real data</td>
</tr>
<tr>
<td>Gao et al. [29]</td>
<td>queue length</td>
<td>lane</td>
<td>signal cycle</td>
<td>only CV data</td>
<td>shockwave sensing and neural network</td>
<td>simulated data</td>
</tr>
<tr>
<td>Aljamal et al [3, 4, 7, 6, 5]</td>
<td>traffic density</td>
<td>lane</td>
<td>variable</td>
<td>CV and detector data</td>
<td>combination of ANN and RF, K-NN, Kalman filter, and adaptive kalman filter, and simulated data</td>
<td>non-linear Particle filter</td>
</tr>
<tr>
<td>Nguyen Van Phu et al. [81]</td>
<td>penetration rates; vehicles arrival rate; turning ratios; queue lengths</td>
<td>lane</td>
<td>second</td>
<td>only CV data</td>
<td>probability theory</td>
<td>simulated data</td>
</tr>
</tbody>
</table>

signal control strategies, such as signal timing methods that require an estimate of the arrival time for each vehicle, including, e.g., [35] and [62]; as well as strategies that require total vehicle densities or flow such as SCATS and SCOOT [77]. Furthermore, to our best knowledge, all model-driven methods aimed at estimating total vehicle counts, i.e., both queuing and moving vehicles, require, in addition to CV data, at least an infrastructure-based point detector that provides vehicle arrival flow rate or employs strict assumptions on vehicles arrival rates.

In contrast, data-driven methods have gained recent popularity thanks to their ability to allow identifying complex patterns and correlations by learning them from available data [8, 65, 79, 26, 56]. Nevertheless, data-driven models typically need a large amount of data used for training, which may not be easy and inexpensive to collect [93]. To the best of our knowledge, the only data-driven method proposed for traffic density estimation of urban signalised links based on CV data is [7], where the authors develop a method based on artificial neural networks (ANN), random forest (RF), and k-nearest neighbours algorithm (K-NN) for traffic density estimation employing CV data. The proposed method was trained and tested on synthetic data, produced via microscopic simulation.

We present in Table 2.1 a collection of the most relevant research works on traffic estimation for urban signalised links using CV data. In summary, previous studies mainly focused on estimating total vehicle counts or queue lengths by fusing data from multiple sources, such as CV data and
point detectors. Fewer studies utilised only CV data, where the proposed methods are based on various mathematical models derived from, e.g., traffic flow theory and probability theory; to deal with data incompleteness, such methods require more or less strict assumptions, e.g., on the arrival flow rate, the arrival patterns of CVs, or their penetration rate. However, these methods are only capable to estimate queue length at lane level or the number of vehicles at the link level. Moreover, there are no existing methods that are providing estimates with a higher spatial granularity, e.g., intra-lane vehicle counts.

2.4 Research gaps

According to literature review and state of the are which has been presented, four main challenges relevant to this dissertation have been identified as following:

- **Research gap 1 (RG1):** User perspective has been considered rarely in traffic control strategies even in a CV environment. Person-based strategies are implemented only to grant priority to public transport vehicles, without considering the actual number of passengers on board other motorists. In particular, maximizing user throughput has not been the objective of any research. Besides, there are limited works that consider the over-saturated condition, in particular when a user-based controller is deployed. This research gap is addressed by **RQ1** and is reflected in **Publication I**.

- **Research gap 2 (RG2):** CV-based TSP strategies can mitigate the adverse impact on the other motorists than conventional TSPs. However, passenger occupancy has been always included as a weight factor in signal timing optimisation to grant priority to transit vehicles by applying different delay models for cars and buses. Moreover, the impact of deploying TSP strategies on the non-transit users on the also societal and environmental effects of TSP are yet to be studied. This research gap is addressed by **RQ2** and is reflected in **Publication II**.

- **Research gap 3 (RG3):** CVs can be equipped with different sensors in terms of accuracy. Accordingly, the quality of CV data is highly dependent on the accuracy and performance of the sensors. In the literature, no study has considered the impact of CV data accuracy on the different types of controller performance. This research gap is addressed by **RQ3** and is reflected in **Publication III**.

- **Research gap 4 (RG4):** In the near future, reaching a fully connected
environment is not expected. Accordingly, many methods have been
developed to estimate traffic state in the vicinity of signalized inter-
sections in a partially CV environment. However, these methods are
limited to estimate queue length at lane level or the number of vehicles
at the link level. Moreover, there is no method to estimate with a higher
spatial granularity, e.g., intra-lane vehicle counts. This research gap is
addressed by **RQ4** and is reflected in **Publication IV**.

Considering the presented research gap, the main contribution of this
dissertation is to address some of the current challenges of developing
CV-based control strategies. The identified research gaps originated from
two vital parts of a traffic control system which are signal timing and input
data. Accordingly, considering identified gaps and the research questions,
Figure 2.1 presents the relevance of research questions, research gaps,
publication and the traffic control system in a CV environment. In the
following, we describe the applied research methodology to answer each of
the research questions one by one.

![Figure 2.1. Relationship of research questions, research gaps and publications of this dissertation](image-url)
3. Research methodology

In this section, we present the proposed research methodology to address each of the identified research gaps.

3.1 How CVs data can be implemented in user-centred and equitable traffic management strategies? (RQ1)

As previously stated in the RG1, current traffic control strategies are mainly vehicle-dominated even in a CV environment while users data can be transmitted as well as vehicle data. To address this limitation, a user-based signal timing (UBSTO) has been developed that has been presented in Publication I. The main assumption of this research is that the number of users on-board of each CV and vehicle data including speed, and distance to stop-bar is available on demand. Based on these assumptions, an analytical model is developed to estimate the arrival time of vehicles to the stop bar and then the output of this model is used to maximise the user throughput of the intersection within a cycle. Due to the complexity of the problem and making the solution applicable for real-time applications, we use a genetic algorithm to solve the optimisation problem. Using a heuristic solution method may lead to sub-optimal results since the entire feasible solution space is not explored and the objective function of the problem is throughput (not delay), there is the possibility of assigning unusable high green time to a phase just to satisfy the constraints of the problem. This may happen more often in low traffic flow conditions, where the last detected vehicle may pass the intersection a considerable amount of time before the end of the assigned green time. Accordingly, a cycle adaption module is applied to trim the excess green time at the end of each phase.

In this research, we test the controller performance at a synthetic isolated signalized intersection in VISSIM environment which is a well-known microscopic simulation software [67]. However, signal timing optimisation for each cycle can be solved in any programming software such as MATLAB.
Implementation of user-supportive strategies such as UBSTO in practice depends on the development of data collection, communication, and controller tools. Firstly, UBSTO needs the number of users on board each vehicle in addition to vehicle data, however current development of CVs focuses mainly on vehicle-related data. For this purpose, using an existing in-vehicle sensor such as seat belt and weighting sensors can be used to collect user data. Secondly, UBSTO needs real-time data as well as other real-time signal timing strategies. In our proposed algorithm, we assumed that optimisation is completed when the closest detected vehicle arrives at the stop-bar. Accordingly, the information needs to be received by the controller in advance and the detection range depends on communication speed and controller processor power. The limitation of conventional data collection tools might be the main reason for limited attention to users in signal timing strategies, but novel technologies such as CV can provide sufficient data to implement user-based signal timing. In this research, we assumed that number of users on each vehicle is available as well as other vehicle-related such as speed and position. However, vehicles data are the critical data for our algorithm. Thus, UBSTO can work when the number of user of a vehicle is missing but vehicle data is available. To achieve this, we can assume the number of users in each vehicle based on, e.g., statistical data [46]. Alternatively, in the case of missing data for the number of users, our strategy can be switched to just vehicle-based mode which maximises vehicle throughput. The proper performance of the
vehicle-based has been also illustrated by the above results. In either case, the experimental setup established in this research has aimed at following the best practice of simulation-based evaluation. Demand patterns used for the case intersection, including intersection layout, are in line with similar research, e.g., [59, 88, 52] and general signal timing guidelines [80].

3.2 How CVs can enhance the abilities of the current TSP strategies in improving social and environmental performance measures? (RQ2)

**RG2** states that in CV-based TSPs other motorist based on number of users on-boards have not been considered in signal timing optimisation which can cause excessive delay for non-public transit users. Moreover, current TSP strategies cannot effectively perform in case of conflict priority request. Accordingly, we propose a user-based TSP in CV environment which has been presented in the Publication II. User-based TSP is a signal timing optimisation that tries to minimise user delay as well as bus schedule delay. The core of the analytical model that has been developed in this research is similar to analytical model of Publication I. In this research, first, the problem of conflicting request can be solved by granting priority to buses with higher number of passengers and higher bus schedule delay and second, other motorists can be prioritized based on number of users on board. Furthermore, we propose a social cost function to evaluate environment and social performance of TSP strategies.

3.3 What is the impact of CV data accuracy on the performance of different CV-based traffic controllers? (RQ3)

CV data availability leads to develop various adaptive traffic controllers. However, the accuracy of this data may affect the controller performance. **RG2** shows the impact of CV data quality on the controller performance yet to be studied. To address this research gap, we measure the effect of CV data accuracy on two types of traffic controllers in Publication III. For this purpose, we assume that CVs can be quipped by three different types GPS receivers which are High accurate GPS, standard GPS with RAIM and mobile GPS. Then, two traffic controllers are considered to test the effect of each data quality level on the traffic controller performance which is Vehicle-based signal timing and max pressure. The vehicle-based signal timing is the same traffic controller which has been developed in Publication I and Publication II while max pressure is a traffic controller.
Research methodology

which optimises signal timing to minimise queue length in the approaches of the intersection\cite{82}. We choose these two controllers to have different structures. Vehicle-based signal timing works based on the estimation of the arrival of each vehicle to the intersection while max pressure only considered queue length in link-level. Each combination of traffic controller and GPS accuracy levels has been tested under under-saturated, saturated and over-saturated traffic conditions at an isolated signalized intersection. Figure 3.2 presents the summary of the method that has been implemented to answer this research question.

![Process framework for assessing the impact of GPS error on traffic signal control](image)

**Figure 3.2.** Process framework for assessing the impact of GPS error on traffic signal control

### 3.4 How limited CVs data can be used to estimate traffic state in a low penetration rate of CVs? (RQ4)

Considering this research question and RG4, we develop a data-driven vehicle estimation to model to estimate the total number of vehicles and number of vehicles between a pair of CVs in a CV environment. This data-driven model is developed in particular to estimate the number of vehicles in the vicinity of a signalized intersection. Figure 3.3 presents the framework of the developed data-driven model. Usually, data-driven models require a considerable amount of data which is difficult and expensive to collect in most cases, The merit of our method is that the model can be trained by a limited amount of real-world data. First, a high-resolution dataset can be collected from signalized intersections. Then, a microsimulation software is calibrated based on the collected data. Finally, the needed data for training can be generated by the microsimulation software
To test the performance of this method, we use real-world vehicle data related to a signalised link, collected in Rotterdam, the Netherlands, provided in [83]. The road segment covered by the video, which starts from the east side of the tower, is 180 m in length, of which about 125 m are located upstream of the traffic light. A traffic signal regulates the traffic flows of two two-lane roads that merge into a two-lane road. Vehicle trajectories are extracted from the videos by considering frames at a frequency of 15 Hz via image processing techniques. The processed dataset includes the longitudinal position, lane, speed, and acceleration of all vehicles. More details about the dataset have been presented in the Publication IV. As the trajectory of a single vehicle is constructed by processing the images at each timestep, there is a possibility of finding many partial trajectories of the same vehicle; in this regard, the dataset was cleaned and extra trajectories for each vehicle were removed. Moreover, since the method used to extract vehicle data from images produced position errors, filtering techniques were applied.
4. **Scientific contribution**

This chapter presents a summary of the contribution and the main findings of each research question that are presented in four peer-review publications. Table 4.1 presents a summary of contributions and findings of each publication of this thesis. The details of each publication have been elaborated on in the following.

4.1 **Publication I**

This paper addresses the RQ1, considering the RG1. In this publication, we present a user-based signal time optimisation (UBSTO) strategy. UBSTO aims at finding the optimum signal timing to maximise user flow throughput within a signal cycle, as opposed to vehicular flow throughput. The main features of UBSTO are:

- Addressing RG1, UBSTO accounts for known (measured or estimated) number of users on-board, as opposed to average occupancy rates

- UBSTO can estimate of each individual users arrival time at the intersection based on CV data, unlike previous works that are based on total (average) approach delay.

- UBSTO is responsive to traffic demand through a flexible cycle length

To test the performance of the proposed controller, a comparison to a conventional fully actuated controller is implemented in microscopic simulation software. Four vehicle types based number of users on board are considered in the simulation. Three performance measures which are throughput, average delay and the average number of stops are measured. The result shows that UBSTO improves all user-related performance measures compared to the fully actuated controller, especially in high traffic demands. Additionally, the delay of other non-transit motorists has
### Table 4.1. Summary of contributions and findings of all publications

<table>
<thead>
<tr>
<th>Publication</th>
<th>Main contributions</th>
<th>Main findings</th>
</tr>
</thead>
</table>
| Publication I | - A user-based traffic control strategy  
- Accounting for number of users of CVs  
- Considering each CV arrival time  
- Responsive to traffic demand through a flexible cycle length | - Granting priority for Vehicles with higher number of users  
- Effective in low penetration of CVs  
- Outperforming the vehicle-based signal timing and the conventional controllers, considering user-related and vehicle-related performance measures |
| Publication II | - A user-based TSP strategy  
- Accounting for number of users of all motorists and bus schedule delay in signal timing optimisation  
- Social cost evaluation | - Granting priority to bus with higher schedule delay and higher number of users  
- Reduction of total social cost compared to a baseline state-of-the-art TSP strategy |
| Publication III | - Investigating the effect of CV data quality on the CV-based adaptive signal controller performance | - Disaggregated input controllers are less sensitive to data error compared to aggregated input controllers  
- Low quality CV data leads to larger performance degradation in aggregated adaptive controllers  
- Disaggregated input controllers outperforms aggregated input controllers using mobile phone GPS data |
| Publication IV | - A data-driven estimation methods based on to estimate the number of vehicles at signalised intersections by using a limited amount of CVs  
- A novel method to address the need for large amount of data for data-driven based estimation models | - A limited high resolution data can effective in training the estimation model  
- Descent performance in estimation of total number of vehicles and number of vehicles between a pair of CV |
been ordered based on the number of users on board. In another word, for example, vehicles with 4 passengers experience less delay than vehicles with passengers.

To the best of our knowledge, maximising user throughput using an adaptive controller in a CV environment has not been investigated in previous research, as most of the research is focused on vehicle-centred strategies. Additionally, in contrast with previous works, we consider different passenger vehicle classes based on the number of users on board, as opposed to assuming the same occupancy for all passenger vehicles. The finding of this publication shows that first, CV can provide sufficient data for developing a user-based signal timing. Second, UBSTO prioritizes vehicles with a higher number of users on board opposing conventional vehicle-based traffic controllers. Third, the computational time of UBSTO is low enough to implement as real-time signal timing. Fourth, UBSTO can also improve user-related performance measures even with a low penetration rate of CVs.

4.2 Publication II

Following Publication I, Publication II contributes to RQ2 and RG2. In this paper, UBSTO is enhanced by a novel user-based TSP designed for a CV environment. The proposed signal timing design in this paper is similar to the developed signal timing in Publication I while accounting for bus priority. Here, the objective signal timing optimisation is to minimise total user delay and minimise the bus schedule delay. According to RG2, we address the problem of conflicting priority requests by considering each bus schedule delay and the number of riders on each bus. The proposed method in Publication II facilitates the mobility of high occupancy vehicles at the signalized intersection by accounting for the exact number of users in the signal timing optimisation. Moreover, in this publication, we propose a social cost function to evaluate the environmental and social performance of TSP strategies.

We tested the performance of proposed TSP strategies on two adjacent intersections, located in Helsinki, Finland where the experiment was deployed, using a microscopic simulation software. Result shows that both proposed TSP strategies can improve the performance of the controllers compared to a baseline controller. In general, both controllers show effective performance in providing priority for buses while performance of passenger cars are improved as well. Moreover, we show that proposed strategies can improve bus schedule adherence. Overall, the proposed strategies are able to solve conflicting TSP request by considering bus schedule delay and number of users of all vehicles in the signal timing optimisation. Moreover, proposed TSP strategies can reduce total social cost
in the test site. In average, total social cost is decreased more than 10% by proposed TSP strategies compared to baseline. Similar improvement can be seen in all component of the total social cost such as fuel consumption and emission.

4.3 Publication III

This paper tries to address RQ3 and RG3 by testing the effect of CV data quality on two types of controllers which are max pressure (in a modified version adapted to operate in a CV environment) and VST, as representative of aggregated input controller and disaggregated input controller, respectively. For this purpose, An isolated four-approach intersection is considered three traffic conditions and by applying three error distributions associated with different GPS technologies. The simulation results show that disaggregated input controllers are less sensitive to data error compared to aggregated input controllers. In another word, Our finding shows that errors in CV data may lead to higher performance degradation in aggregated adaptive controllers.

A type of GPS device that is not very accurate is already embedded in most of the smart cellphones which, at least one, is found in every vehicle. Accordingly, mobile GPS data can be used as input for the controller. In this paper, we showed that disaggregated input controllers can outperform aggregated input controllers if mobile GPS data is used as the source of input. Although high accurate GPS may increase CV data quality but equipping all the vehicles with this type of GPSs is considerably expensive and unfeasible.

4.4 Publication IV

This publication focuses on the input of a traffic controller in a CV environment. According to RQ4 and RG4, we propose a data-driven vehicle estimation method based on a limited number of CVs. Accordingly, two data-driven estimation methods are developed by applying machine learning principles, to estimate the number of vehicles approaching a signalised intersection, which can be used, e.g., for operating adaptive signal timings. First, a method that is denoted as “aggregated” in this paper, is designed to estimate the total number of vehicles in a signalised link. Second, a method that is denoted as “disaggregated” in this paper, is developed for a more granular estimation, to estimate the number of non-connected vehicles upstream and downstream of each CV. By using the two developed estimation models, the total vehicle count can be estimated, as well as the number of vehicles between each pair of CVs. Consequently, a higher resolution of
vehicle estimation can be expected concerning existing approaches. However, providing a sufficient amount of data to train the mentioned data-driven models is a challenge. In this regard, we propose a novel method which can reduce the need for a large amount of data. In our proposed method, relatively small real data is collected offline, e.g., from a fixed camera or an unmanned aerial vehicle (i.e., drone). Then, microscopic traffic simulation software such as Vissim or Aimsun can be employed to expand the training data without compromising the estimation accuracy.

To the best of our knowledge, this paper is one of the first data-driven efforts in estimating vehicle counting, relying only on limited CV data in the vicinity of signalized intersections. The findings of this paper can be summarized as follows:

• Data-driven methods, e.g., on machine learning models, can provide useful estimations in estimating traffic variables during the transition period to a fully connected environment.

• The proposed method allows vehicle counting needless of infrastructure-based sensors, such as loop detectors, even with a limited number of CVs.

• We showed that using a limited amount of real data can be successfully employed for training the estimation model by using a synthetic dataset generated by a calibrated simulation tool.

• The results show that the estimation error of all models in the disaggregated method is widely affected by the distance between CVs in a pair. Additionally, more accurate estimation is achieved when the speed difference within a CV pair is lower.

• In the case of using the disaggregated method to estimate the total number of vehicles, estimation is not accurate as simply using the aggregated method. This is attributed to error accumulation, which is indeed more pronounced in the case of the high penetration rate of CVs.

• The aggregated method shows better performance as the penetration rate of CVs increases. This suggests that in case both the total number of vehicles and more granular estimations are needed, it is wiser to use a combination of the proposed methods.
5. Discussion and future directions

5.1 Summary of the outcomes

This dissertation aiming to address four research questions regarding urban traffic management, considering emergence of CVs. In a bigger picture, two important elements of a traffic controller which are signal timing and input data are considered in this research. Two first research question are formulated in particular to study user-based signal timing strategies in CV environment. The third questions, is in the borderline of signal timing and input data where impact of CV data quality on the controller performance is assessed. The last question belongs to input data to developed a method in order to estimate traffic state in vicinity of signalized intersection in a partially connected vehicle environment.

The first and the second research questions have been addressed by developing a user-based signal timing strategy. In Publication I and Publication II, we showed that how CVs can help to devolved a user-centering signal timing strategies opposing to the existing vehicle-dominated strategies. The developed signal timing method is able to improve user-related performance measures. Additionally, by applying this method, vehicles with higher number of users on-board experience less travel time in a signalized intersection. According to the Federal Highway Administration, in 2018 average vehicles occupancy rate was 1.7 in US [2]. A similar number is reported for EU in 2016 [27]. In fact, low vehicle occupancy is directly related to concerns about energy consumption and emissions, especially with the advancement of automated vehicle technology [76]. Accordingly, implementing user-based traffic management strategies can be a supporting mechanism for user-based mobility management strategies where vehicles with higher number of passengers are prioritized such as High-occupancy vehicle lane. Thus, by implementing user-based signal control strategies not only the right-of-way for high-occupancy and ride-sharing vehicles can be provided, but also it can be an effective strategy in supporting

---

33
behavioral shift to shared mobility services. For example, it has been studied that prioritizing of high-occupancy vehicles in traffic management strategies can affect users' willingness to share their rides [63]. Moreover, in publication II, we propose a social cost formula to measure the impact of TSP strategies on the environment and the society.

To address the third research question, effect of CV data quality on the two type of traffic controller has been studied, by assuming three CV data accuracy level. In this study, it has been showed that inaccurate CV data can adversely effect the performance of the traffic controller. However, this effect can be mitigated in a type of traffic controller where the input of the controller is not aggregated in the link level and each individual CV data is an input for the controller. Additionally, in this research we showed that even mobile GPS data can be sufficient to run a traffic control if the input of the controller is desegregated CVs data.

The fourth research question is mainly formulate to address the challenge of providing sufficient data for traffic controllers in a partially CV environment. Accordingly, a data-driven method has been proposed to estimate total number of vehicles in an approach of a signalized intersection based on limited CVs data. Moreover, a method has been developed to estimated number of non connected vehicle between a pair of CV. This research is one of the first data-driven studies in estimating vehicle counting, relying only on limited CV data. In this paper, we showed that data-driven methods, e.g, machine learning are able to be implemented as a significant estimator of traffic state in a partially CV environment. This finding of this research can be a start point for implanting traffic state estimation system that are needless of infrastructure-based sensors, such as loop detector, even with a low amount of CVs. Moreover, this research demonstrates that even by using a limited real-world data, training of data-driven methods is possible, by using advantages of other tools such as traffic simulation software.

5.2 Transferability and practicality

The proposed user-based signal timing methods in this dissertation can be implemented in practice on any signalized intersections if required data can be transmitted from CVs to the controller. In this case, passenger information in the CVs, can be collected using the current deployed sensors such as seat belt or weight sensor. Implementing of user-centred and equitable signal control strategies such as controllers that have been presented in the Publication I and Publication II can be an effective incentives to encourage shared mobility and public transportation and also they can be implemented as complimentary policy of existing policies such as high occupancy vehicle priority lane.
According to the World Health Organization\(^1\), 24% of total annual death in the world are linked to the environmental issues and transport plays a large role in this. Hence, there is a essential need for traffic signal control strategies that can mitigate air pollution levels and fuel consumption. The performance of the proposed TSP strategies show that implementation of these type of strategies in the real-world not only can enhance the performance for the current TSP strategies but also they can mitigate adverse impacts of the vehicle traffic on the society and environment.

However, it could possibly take 25–30 years to reach 95% CV penetration rate, in case of mandatory installation of data transferring tools on new light vehicles in the U.S. is \(^40\). In this case, applying estimation methods, e.g. data driven method proposed in the Publication 4, can be implemented to provided the required data for the controller. The proposed method is able to be implemented on any signalized intersection where limited high resolution traffic data is available, e.g. from CVs or collected by drones similar to pNEUMA traffic dataset \(^10\).

### 5.3 Recommendations for future research

Emerging of CVs is a greet opportunity to implement novel traffic management strategies by improving both the signal timing and the input data. According to findings of this dissertation, here the recommendations for future research are presented.

Regarding the signal timing, first, user-based signal timing strategy should be implemented on road networks with additional transport modes, such as trucks, emergency vehicles, bicycles and pedestrians. For instance, special weight can be assigned to travel time for cyclists and pedestrians - if these modes are to be prioritised further or emergency vehicle can be prioritized by assigning a high weight. Second, stop-bar passage time prediction model can be extended to a partially CV environment where data of non-connected vehicles can be estimated for example by the method presented in this dissertation. Third, the the user-based signal timing strategy could be tested on a larger networks with multiple intersections. In this case, other performance measures, e.g., queue lengths and direct emissions can take into account. Fifth, the signal timing propose din the dissertation can be improved by taking advantage of self-driving vehicles capabilities. Regarding the input of traffic control, considering the usage of filtering methods to improve estimation accuracy, by utilising time-series data in addition to instantaneous measurements can be a topic for future research.

---

1\(^1\)https://www.who.int/data/gho/data/themes/public-health-and-environment
References


References


References


[67] PTV. VISSIM. 11.00-03 user manual. 2018.


For many decades, urban traffic management systems have been vehicle-dominated. That is not only because of a lack of attention to users, but conventional data collection tools are powerless to collect individual vehicle data as well as vehicle users data. Connected vehicles (CV), as an emerging technology, can collect and transmit real-time vehicle and its users data. This ability facilitates the development of user-centred traffic management strategies in urban transport networks. However, there are some challenges yet to be addressed to convert raw CV data to efficient input for traffic controls. Moreover, achieving a fully connected environment is not possible in near future due to various limitations. Accordingly, this dissertation aims at developing a traffic management strategy based on CV data that improves user-related performance measures at signalized intersections. Furthermore, this dissertation assesses the effect of CV data accuracy on traffic controllers and presents a method to compensate lack of CVs in the urban environment to deploy in traffic management strategies. In this dissertation, we research two vital aspects of traffic signal control which are signal timing optimisation and data.