



**HAL**  
open science

# Estimation and control of urban road traffic with communicating vehicles

Cyril Nguyen van Phu

► **To cite this version:**

Cyril Nguyen van Phu. Estimation and control of urban road traffic with communicating vehicles. Modeling and Simulation. Université Gustave Eiffel, 2022. English. NNT: 2022UEFL2024 . tel-03957203

**HAL Id: tel-03957203**

**<https://theses.hal.science/tel-03957203>**

Submitted on 26 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ GUSTAVE EIFFEL

Mémoire pour l'obtention de la  
THÈSE DE DOCTORAT

Ecole doctorale MSTIC 532, Spécialité : Informatique

# Estimation and control of urban road traffic with communicating vehicles

présenté par

Cyril Nguyen Van Phu  
Université Gustave Eiffel / Cosys / Grettia

Soutenu le 15 septembre 2022 devant le Jury composé de:

**Président du jury :**

Fabien Moutarde - Professor - Ecole Nationale Supérieure des Mines de Paris

**Rapporteurs :**

Ioannis Papamichail - Professor - Technical University of Crete

Maria Nadia Postorino - Professor - University of Bologna

**Examineurs :**

Gurcan Comert - Associate Professor - Benedict College, Columbia, SC.

Nadir Farhi - Researcher - Université Gustave Eiffel, thesis supervisor

Fabien Moutarde - Professor - Ecole Nationale Supérieure des Mines de Paris

Antoine Tordeux - Professor - University of Wuppertal

## GRATITUDE AND DEDICATIONS

I would like to thank all the people and all the conditions which are connected in any way, near or far, to this document. This is due to the people around that this document is made possible, and they are often users of transportation systems. I express my deep and respectful gratitude to everyone and everything connected to this thesis which is dedicated to them, for their most benefit. In particular, I dedicate this work to my mother Pascaline, my father Rémy, my brother Marc, my family, my friends and my colleagues. May this work benefit to the most people possible.

## GRATITUDE ET DÉDICACES

Je tiens à remercier toutes les personnes et toutes les conditions qui sont connectées à ce document. C'est grâce à toutes les personnes qui m'entourent que ce document est rendu possible, et elles sont souvent des utilisateurs des systèmes de transport. J'exprime ma profonde et respectueuse gratitude à tous et à tout ce qui est lié à cette thèse, qui leur est dédiée, pour leur plus grand bénéfice. En particulier, je dédie ce travail à ma mère Pascaline, mon père Rémy, mon frère Marc, ma famille, mes amis et mes collègues. Puisse ce travail bénéficier au plus grand nombre.

# Contents

<b>Dedications</b>	<b>2</b>
<b>Short Summary</b>	<b>7</b>
<b>Summary</b>	<b>11</b>
<b>I Introduction</b>	<b>23</b>
<b>1 Introduction and state of the art</b>	<b>25</b>
1.1 Introduction . . . . .	25
1.1.1 Context and stakes of road traffic science . . . . .	25
1.1.2 Intelligent Transportation Systems (ITS) . . . . .	27
1.1.3 Assumptions and tools . . . . .	29
1.1.4 Outline and contributions . . . . .	30
1.2 State of the art . . . . .	32
1.2.1 Estimation of road traffic . . . . .	32
1.2.1.1 Sensing road traffic with fixed sensors . . . . .	32
1.2.1.2 Sensing road traffic with probe vehicles . . . . .	33
1.2.1.3 Estimation of road traffic state with fixed sensors (historical approaches) . . . . .	34
1.2.1.4 Estimation of road traffic state with probe vehicles . . . . .	36
1.2.2 Road traffic signal control with connected and automated vehicles . . . . .	40
1.2.2.1 Actuated traffic signal control . . . . .	40
1.2.2.2 Planning based traffic signal control . . . . .	42
1.2.2.3 Platoon based traffic signal control . . . . .	44
1.2.2.4 Transit priority control . . . . .	46
1.2.2.5 Signal vehicle coupled control (SVCC) . . . . .	48

<b>II</b>	<b>Estimation of urban road traffic</b>	<b>49</b>
<b>2</b>	<b>Estimation of 2-lanes traffic</b>	<b>53</b>
2.1	Abstract . . . . .	54
2.2	Introduction . . . . .	54
2.2.1	State of the art . . . . .	54
2.2.2	Paper contribution and organization . . . . .	55
2.3	Problem statement . . . . .	56
2.3.1	Assumptions . . . . .	57
2.3.1.1	Road network topology . . . . .	57
2.3.1.2	Traffic demand . . . . .	57
2.3.1.3	Probe vehicles . . . . .	57
2.3.1.4	Turn ratios . . . . .	58
2.3.2	Notations . . . . .	59
2.4	Traffic state estimation . . . . .	59
2.4.1	Primary parameters estimation . . . . .	59
2.4.2	Queue length estimation . . . . .	62
2.4.2.1	Distribution probability law of the 2-lanes without having the information provided by the probes . . . . .	63
2.4.2.2	Distribution probability law of the 2-lanes queue lengths with the information provided by the probe vehicles . . . . .	63
2.4.2.3	Estimators . . . . .	65
2.4.3	Traffic light control and optimal assignment of vehicles onto the lanes . . . . .	66
2.5	Simulation results, examples and discussion . . . . .	67
2.5.1	Primary parameters estimation . . . . .	68
2.5.2	Probability distributions . . . . .	71
2.5.2.1	Example: probability distribution of the two queues lanes without and with the informa- tion provided by the probe vehicles . . . . .	71
2.5.2.2	Results for the scenarios of TABLE 2.2 . . . . .	73
2.5.3	Traffic light control and vehicles assignment onto the lanes . . . . .	76
2.5.4	Communication network performances . . . . .	81
2.6	Conclusion and perspectives . . . . .	81
2.7	Appendix : Calculus of the bias of $p$ estimator . . . . .	82
<b>3</b>	<b>Estimation of multi-lanes traffic</b>	<b>85</b>
3.1	Abstract . . . . .	86
3.2	Introduction and state of the art . . . . .	86
3.2.1	Introduction . . . . .	86
3.2.2	State of the art . . . . .	87

3.2.2.1	Road traffic estimation with fixed sensors (historical approaches) . . . . .	87
3.2.2.2	Road traffic estimation with probe vehicles . . . . .	88
3.2.3	Assumptions and main contributions . . . . .	90
3.3	Road traffic state estimation . . . . .	91
3.3.1	Primary parameters . . . . .	93
3.3.1.1	Estimation model for the penetration ratio $p$ . . . . .	93
3.3.1.2	Estimation model for the total arrival rate $\lambda$ . . . . .	93
3.3.1.3	Estimation model for the matrix $W$ . . . . .	94
3.3.1.4	Estimation of the arrival rates per lane $\lambda_i$ . . . . .	95
3.3.2	Traffic state estimation for roads composed of three lanes . . . . .	95
3.3.3	Application to roads composed of any number of lanes . . . . .	102
3.4	Numerical experiments . . . . .	103
3.4.1	Estimation of the primary parameters . . . . .	103
3.4.2	Road traffic state estimation . . . . .	105
3.5	Conclusion and perspectives . . . . .	109
<b>III</b>	<b>Control of urban road traffic</b> . . . . .	<b>111</b>
<b>4</b>	<b>Semi-decentralized control</b> . . . . .	<b>115</b>
4.1	Abstract . . . . .	116
4.2	Introduction . . . . .	116
4.3	A short review of TUC . . . . .	118
4.4	Semi-decentralization . . . . .	120
4.4.1	The dynamics . . . . .	121
4.5	Numerical example . . . . .	122
4.5.1	Model implementation and Simulation Tools . . . . .	123
4.5.2	Network configuration . . . . .	124
4.5.3	Preliminary results . . . . .	125
4.6	Preliminary conclusions . . . . .	129
<b>5</b>	<b>V2I traffic control algorithm</b> . . . . .	<b>131</b>
5.1	Abstract . . . . .	132
5.2	Introduction . . . . .	132
5.2.1	Introduction . . . . .	132
5.2.2	State of the art . . . . .	133
5.2.3	Paper organization . . . . .	133
5.3	Connected traffic light signal control . . . . .	133
5.3.1	Algorithm description . . . . .	133
5.3.1.1	Assumptions . . . . .	134
5.3.1.2	Dynamic Maps . . . . .	134
5.3.1.3	Election . . . . .	135

5.3.1.4	Action . . . . .	135
5.3.2	Properties of the algorithm . . . . .	137
5.3.2.1	Property 1 . . . . .	137
5.3.2.2	Property 2 . . . . .	137
5.3.2.3	Property 3 . . . . .	137
5.3.2.4	Property 4 . . . . .	137
5.3.3	Implementation . . . . .	137
5.4	Simulation results . . . . .	138
5.4.1	One Junction with connected TLS . . . . .	138
5.4.1.1	Scenario for one junction . . . . .	138
5.4.1.2	Simulation measurements . . . . .	139
5.4.2	Traffic control at the road network level . . . . .	143
5.4.2.1	Scenario for the road network . . . . .	143
5.4.2.2	Simulation results . . . . .	147
5.5	Conclusions . . . . .	148
<b>IV</b>	<b>Conclusion</b>	<b>149</b>
5.6	Conclusion . . . . .	151
5.7	Implications and consequences . . . . .	153
	<b>Bibliography</b>	<b>155</b>
	<b>List of figures</b>	<b>171</b>
	<b>List of tables</b>	<b>173</b>

## SHORT SUMMARY

This thesis addresses the topic of estimation and control of urban road traffic. The goal is to improve the conditions and the consequences of road traffic on the economical and ecological domains, on a short, mid and long term perspectives.

Concerning the estimation of road traffic, we present a method for the estimation of traffic state at road junctions controlled with traffic lights. We assume mixed traffic where a proportion of vehicles are equipped with communication resources. The estimation of road traffic state uses information given by communicating vehicles. The method proposed is built upon a previously published method which was applied to estimate the traffic in the case where roads are composed of one lane. In this thesis, we consider the case where roads are composed of two lanes and three lanes, and it is shown that these solutions can address the general case, where roads are composed of any number of lanes. The geometry of the road junction is assumed to be known, as well as its connections between incoming and outgoing lanes and roads. Using the location data provided by the communicating vehicles, first, we estimate some primary parameters including the penetration ratio of the probe vehicles, as well as the arrival rates of vehicles (equipped and non-equipped) per lane by introducing the assignment onto the lanes. Second, estimations of the queue length of the multiple-lanes road, without and with the additional information provided by the location of the communicating vehicles in the queue, are given. The proposed models are illustrated and discussed with numerical simulations.

Concerning the control of road traffic, first, a semi-decentralized approach for urban traffic control, based on the TUC (Traffic responsive Urban Control) strategy, is presented. We assume that the control is centralized as in the TUC strategy, but a contention time window inside the cycle time is introduced, where the traffic light is yellow for all the directions, and where antagonistic phases alternate a priority rule. The idea of introducing this time window is to reduce the red time inside the cycle, and by that, increase the capacity of the network junctions. In practice, the priority rule could be applied using vehicle to vehicle (v2v) communications. The vehicles having the priority pass almost normally through the junction, while the others reduce their speed and yield the way. A model for the dynamics and the control of such a system is given. The model is still formulated as a linear quadratic problem, for which the feedback control is derived analytically and off-line, and applied in real time. The model is implemented using the Simulation of Urban MObility (SUMO) tool in a small regular (american-like) network configuration. The results are presented and compared to the classical TUC strategy.



Second, we present a new algorithm for urban traffic light control with mixed traffic (communicating and non communicating vehicles) and mixed infrastructure (equipped and unequipped junctions) which explicits the implementation of the contention time window of the semi-decentralized control. We call equipped junction here a junction with a traffic light signal (TLS) controlled by a road side unit (RSU). On such a junction, the RSU manifests its connectedness to equipped vehicles by broadcasting its communication address and geographical coordinates. The RSU builds a map of communicating vehicles approaching and leaving the junction. The algorithm allows the RSU to select a traffic phase, based on the built map. The selected traffic phase is applied by the TLS; and both equipped and unequipped vehicles must respect it. The traffic management is in feedback on the traffic demand of communicating vehicles. The vehicular traffic as well as the communications are simulated. The two simulations are combined in a closed loop with visualization and monitoring interfaces. Several indicators on vehicular traffic (mean travel time, ended vehicles) and IEEE 802.11p communication performances (end-to-end delay, throughput) are derived and illustrated in three dimension maps. The traffic control is then extended to an urban road network where the number of equipped junctions are also varied. Other indicators are shown for road traffic performances in the road network case, where high gains are experienced in the simulation results.

## RÉSUMÉ COURT

Cette thèse aborde le thème de l'estimation et du contrôle du trafic routier urbain. L'objectif est d'améliorer les conditions et les conséquences du trafic routier sur les domaines économiques et écologiques, sur des perspectives à court, moyen et long terme.

Concernant l'estimation du trafic routier, nous présentons une méthode d'estimation du trafic aux carrefours contrôlés par des feux tricolores. L'estimation du trafic routier utilise les informations fournies par les véhicules communicants, dont le taux de pénétration est variable. La méthode proposée est basée sur une méthode précédemment publiée qui a été appliquée pour estimer le trafic dans le cas où les routes sont composées d'une voie. Dans cette thèse, nous considérons le cas où les routes sont composées de plusieurs voies. La géométrie de la jonction routière est supposée connue, ainsi que ses connexions entre les voies entrantes et les routes sortantes. A partir des données de localisation fournies par les véhicules communicants, dans un premier temps, des paramètres primaires sont estimés, dont le taux de pénétration des véhicules connectés, ainsi que les taux d'arrivée des véhicules par voie, en introduisant l'affectation sur les voies. Deuxièmement, des estimations des longueurs des files d'attente de la route à voies multiples, sans et avec les informations fournies par la localisation des véhicules communicants, sont dérivées. Le modèle proposé est discuté et illustré avec des simulations numériques.

Concernant le contrôle du trafic routier, nous présentons premièrement une approche semi-décentralisée pour le contrôle du trafic urbain, basée sur la stratégie TUC (Traffic Urban Control). Nous supposons que le contrôle est centralisé comme dans la stratégie TUC, mais avec l'introduction d'une fenêtre temporelle concurrente à l'intérieur du temps de cycle, où le feu est jaune pour toutes les directions, et où les phases antagonistes alternent une règle de priorité. Pendant la fenêtre temporelle concurrente, les véhicules prioritaires passent presque normalement le carrefour, tandis que les autres réduisent leur vitesse et cèdent le passage. Le modèle pour la dynamique et le contrôle d'un tel système est dérivé. Le modèle est toujours formulé comme un problème linéaire quadratique, pour lequel le contrôle par rétroaction est dérivé analytiquement et hors ligne, et appliqué en temps réel. Le modèle est implémenté à l'aide de l'outil Simulation of Urban MObility (SUMO) dans une petite configuration de réseau régulier. Les résultats sont présentés et comparés à la stratégie TUC classique. Deuxièmement, nous présentons un nouvel algorithme pour le contrôle des feux de circulation avec trafic mixte (véhicules communicants et non communicants) et infrastructure mixte (jonctions équipées et non équipées) qui explicite l'implémentation de la fenêtre temporelle concurrente du contrôle semi-décentralisé. On appelle carrefour équipé, un carrefour avec feu tricolore

piloté par une borne communicante. La borne construit une carte des véhicules communicants approchant et sortant de la jonction. Une phase de trafic est sélectionnée en fonction de la carte construite. La phase de trafic sélectionnée est appliquée par le feu. La gestion du trafic est en rétroaction sur la demande de trafic des véhicules communicants. La circulation des véhicules ainsi que les communications sont simulées. Plusieurs indicateurs sur le trafic véhiculaire (temps de trajet moyen, véhicules servis) et les performances de communication IEEE 802.11p (délai, débit) sont dérivés. Le contrôle du trafic est étendu à un réseau routier urbain où le nombre de carrefours équipés est variable. D'autres indicateurs sont présentés pour les performances du trafic routier dans le cas du réseau routier, où des gains élevés sont constatés dans les résultats de simulation.

## SUMMARY

Transportation systems enable the mobility of goods and people. The design of efficient transportation systems with a satisfactory quality of service is a major challenge for our modern societies. Indeed, transportation systems are complex systems whose design extends over the long term (several decades), and whose effects are lasting and significant. Among the effects of transportation systems, we distinguish effects at the individual or societal scales. Mobility, movement, are immutable characteristics of reality, and the stakes of transportation systems concern the economy, the ecology, the public health. Transportation systems include transportation demand, the need for mobility, and transportation supply, the service enabling mobility. There is an imbalance when demand exceeds supply, and then congestion occurs. Nowadays, the economic consequences of congestion are major. For example, in the United States, in 2020, the time spent during transportation was estimated to 4.3 billion hours. In the same year, 3.8 billion liters of fuel were used. Gaseous emissions associated with transportation represented 18 millions of tons. The limits of mobility demand are very high, which explains why congestion occurs. On the other hand, the transportation supply is limited. Building new roads requires significant amounts of time and money. In the city, space is limited and increasing the supply of transportation is difficult. Consequently, regulating transportation systems in order to limit imbalances is simpler than increasing transportation supply. The regulation of road traffic can be done for example with the guidance of vehicles, the adequate control of traffic lights, the application of adequate public policies, etc.

New information and communication technologies allow vehicles to communicate, to be localized, and to be automated. Wireless communication is based on communication protocols specific to vehicular communication networks, or not specific. Among vehicle localization systems, GPS satellite localization is common, but other techniques are possible. Finally, the automation of vehicle driving makes it possible to open up new perspectives in terms of traffic control mechanisms, since each vehicle can become an actuator of the regulatory mechanism. The concept of Intelligent Transportation Systems (ITS) proposes to use these new technologies in order to regulate transportation systems, and improve their efficiency in relation with societal issues. Vehicles communicating their GPS position wirelessly can improve road traffic. In order to regulate road traffic, traffic models are needed. They rely on the knowledge of traffic state, which is given by traffic estimation models. This thesis deals with the estimation and control of road traffic with communicating vehicles.

Traditionally, the estimation of road traffic used input data provided by fixed sensors, placed next to, or on the roadway. Communicating vehicles

can be used as mobile road traffic sensors to estimate travel times and queue lengths. With these mobile sensors, new traffic estimation methods are being developed. A distinction is made between deterministic and stochastic methods. This thesis proposes to extend a stochastic method from the literature, in order to estimate various parameters of urban traffic in the case of multi-lane roads. This method first published by Comert et al. in 2009 uses the probability distribution calculation of queue lengths. These probability distributions are conditioned by the information of the number of communicating vehicles present in the queue, as well as the information of the position of the last communicating vehicle in the queue. The length of the queue is estimated with the expectation of this probability distribution. It is shown that the queue length estimation only depends on the position of the last communicating vehicle. In subsequent articles, the authors add the estimation of primary parameters such as the penetration ratio of communicating vehicles, the arrival rate. The case where a residual queue is present at the start of a red light is also considered.

Note that GPS localization systems are not accurate enough to detect the lateral position of vehicles on a multi-lane road : it is not possible to determine on which lane a vehicle is located. Thus, it is not possible to duplicate the existing methods for one-lane roads, to the case of multi-lane roads. This thesis proposes to generalize the estimation of road traffic with communicating vehicles to the case of multi-lane roads. The part of the thesis on traffic estimation is composed of two articles, one of which studies the case of roads formed by two lanes and the other, the case of roads formed by three lanes. It is shown that the method is general and that the case where three-lane roads are considered also makes it possible to solve the case of roads with a greater number of lanes. Thus, the thesis proposes a general traffic estimation method for roads composed of any number of lanes.

It is assumed that the proportion of communicating vehicles is variable, and that the transportation demand is low or moderate, so that the vehicle arrival process can be modeled by a Poisson process. The topology of the intersection is also assumed to be known, in particular the connections between the incoming lanes and the roads leaving the intersection. The turn ratios of the vehicles are assumed to be known. One method for estimating turn ratios is to track the direction where communicating vehicles are heading as they exit the intersection. We propose a vehicle assignment model on the different lanes where the hypothesis considers that the queues per lane tend to balance each other, as much as possible. Two estimators of the penetration ratio of communicating vehicles are proposed, and their performances evaluated in simulation. The total vehicle arrival rate, as well as the arrival rate for each lane, are formulated analytically in dependence on the vehicle assignment model established previously. Assuming that the arrival process is modeled by a Poisson process, a formulation of the probability distribution of queue lengths is given. This formulation depends on

the arrival rate per lane, and on the red time, and does not use the information communicated by the communicating vehicles. An estimator of the number of communicating vehicles present on each lane is proposed. Three other formulations of the probability distribution of queue lengths are analyzed. These are conditional probabilities on the position of the last vehicle connected as well as on the total number of vehicles connected on the multi-lane road. One of the formulations uses the estimator of the number of communicating vehicles present on each lane.

The assumptions are discussed using several simulation tools. First, the simulation results are produced by a bi-directional simulator combining microscopic road traffic simulator and discrete event communication simulator. The simulation of bi-directional communicating vehicles is characterized by a simulation of road traffic which depends on the simulation of the communication, and conversely, by the simulation of the communication which depends on the simulation of road traffic. Microscopic road traffic and communication simulators interact with each other in real time. Second, in order to overcome the defaults of the assignment model of the road traffic microscopic simulator, simulations are carried out with a discrete-event communication simulator. The vehicles are represented by communication packets. Each vehicle is assigned to a queue associated with one of the lanes of the road. The traffic light extracts vehicles from the queues at the saturation rate during the green light phase. The simulation results show that the model can represent the cases of imbalance between the queue lengths on the different lanes. Various performance indicators show the improvement in the estimation of the queue lengths when using the location data communicated by the communicating vehicles. As expected, the performance of the estimators is increased with the increase in the penetration ratio of communicating vehicles.

The estimation of the road traffic enables the control of the road traffic. In this thesis, two contributions to the theory of road traffic control are presented. The first contribution extends one method of the literature, the Traffic Urban Control (TUC). TUC is a centralized traffic control that can be applied in real time on large road networks, including in the case of high demand and congestion. A simple model considers the dynamics of queues on each road in the network, which depends on the duration of the green light phase. An optimization method seeks to minimize a criterion depending on the number of vehicles present on each road, and on the duration of the green lights. The solution of the optimization problem comes down to solving the Riccati equation, which gives the optimal control as a function of the traffic state on each road.

A contribution to this TUC control is made in this thesis. A new phase of the traffic light managing the intersection is proposed. During this phase, the communicating vehicles cooperate with each other with wireless inter-vehicle communication, in order to cross the junction. This time window

introduces a concurrent access space, where vehicles compete with each other to cross the intersection while avoiding collisions. The formulation of the TUC control, which takes into account such a time window is given. The new control introduces a time window where access control is distributed, resulting in a semi-decentralized control. Performance evaluations of the new semi-decentralized control with microscopic road traffic simulation tool are carried out. The scenario uses an American-style, grid-like road network. The concurrency time window is implemented with the simulator with a parameter that represents the conflict between vehicles. The Riccati equation is solved off-line and traffic light control is applied in real time to the various intersections. In this scenario, the simulation results show that the new semi-decentralized control can resolve congestion faster than the centralized TUC control. The results show the sensitivity of the control on the parameter representing the conflict between the vehicles. The choice of this parameter could be the subject of further studies.

We then propose a traffic light control algorithm using vehicle to infrastructure communication and which could implement the contention time window of the semi-decentralized control. This algorithm assumes that some of the vehicles are equipped with wireless communication and localization capabilities. Some traffic lights are associated with a wireless road side unit, which allow communication between vehicles and the infrastructure. The connected traffic light builds a dynamic virtual map of the intersection using the trajectories of communicating vehicles. Then, the traffic light chooses a traffic phase according to the state of the map of the intersection. Once the signaling phase is applied, communicating and non communicating vehicles respect the control. The algorithm is implemented in a simulator of communicating vehicles using the WAVE communication architecture based on the IEEE 802.11p protocol. Bi-directional VEINS simulation implements road traffic dependence on communication, as well as communication dependence on vehicle mobility. A performance study is done on the same grid road network used for the semi-decentralized control simulation. In this scenario, the performance of the algorithm is evaluated for the communication protocol as well as for the transport of vehicles.

The thesis proposes a comprehensive approach to the estimation and control of road traffic. The proposed traffic estimation model and the semi-decentralized traffic control assume that a proportion of vehicles are equipped with localization systems and wireless communication. The road traffic estimation model generalizes a method proposed by Comert et al., while the road traffic control generalizes the method Traffic Urban Control (TUC). A traffic light control algorithm which uses communication between vehicles and infrastructure shows how to implement the contention window of the semi-decentralized control. The simulations allow the study of the performances of the proposed methods. The simulation studies use microscopic traffic simulator and communication simulator. Bi-directional simulation of

road traffic and communication is also used to evaluate the proposed methods. The results show the relevance of the assumptions and models, since gains are observed compared to conventional methods. Transportation demand is assumed to be moderate and the case where the demand is high could be the subject of further work.



## RÉSUMÉ

Les systèmes de transport permettent la mobilité des biens et des personnes. La conception de systèmes de transport efficaces et à la qualité du service rendu satisfaisante est un enjeu majeur de nos sociétés modernes. En effet, les systèmes de transport sont des systèmes complexes dont la conception s'étend sur le temps long (plusieurs dizaines d'années), et dont les effets sont durables, et significatifs. Parmi les effets des systèmes de transports, on distingue les effets à l'échelle de l'individu ou de la société. La mobilité, le mouvement, sont des caractéristiques immuables du réel, et les enjeux sont économiques, écologiques, sanitaires. Les systèmes de transport comportent une demande en transport, le besoin de mobilité, et une offre de transport, le service permettant la mobilité. Il y a un déséquilibre lorsque la demande excède l'offre, et alors la congestion se manifeste. De nos jours, les conséquences économiques de la congestion sont majeures. Par exemple, aux États-Unis, en 2020, le temps passé dans les transports était estimé à 4.3 milliards d'heures. La même année, 3.8 milliards de litres de carburant étaient utilisés. Les émissions gazeuses associées au transport représentaient 18 millions de tonnes. Les limites de la demande de mobilité sont très élevées ce qui explique que la congestion se manifeste. En revanche, l'offre de transport est limitée. Construire de nouvelles routes nécessite des quantités de temps et d'argent importantes. En ville, l'espace est limité et augmenter l'offre de transport est difficile. En conséquence, réguler les systèmes de transport afin de limiter les déséquilibres, est plus simple qu'augmenter l'offre de transport. La régulation du trafic routier peut se faire par exemple avec le guidage des véhicules, le contrôle adéquat des feux de circulation, l'application de politiques publiques adéquates, etc.

Les nouvelles technologies de l'information et de la communication permettent aux véhicules de communiquer, d'être localisés, et d'être automatisés. La communication sans fils repose sur des protocoles de communication spécifiques aux réseaux de communication véhiculaires, ou non spécifiques. Parmi les systèmes de localisation des véhicules, la localisation satellitaire GPS est commune, mais d'autres techniques sont possibles. Enfin, l'automatisation de la conduite des véhicules permet d'ouvrir de nouvelles perspectives quant aux mécanismes de contrôle du trafic, puisque chaque véhicule peut devenir actionneur du mécanisme de régulation. Le concept de Systèmes de Transports Intelligents (STI) propose d'utiliser ces nouvelles technologies afin de réguler les systèmes de transport et d'améliorer leur efficacité en rapport avec les enjeux sociétaux. Les véhicules communiquant leur position GPS sans fils, peuvent permettre l'amélioration du trafic routier. Afin de réguler le trafic routier, des modèles de trafic sont nécessaires. Ils s'appuient sur une connaissance de l'état du trafic, qui est

donnée par des modèles d'estimation du trafic. Cette thèse a pour sujet l'estimation et le contrôle du trafic routier avec des véhicules communicants.

Traditionnellement, l'estimation du trafic routier utilisait des données d'entrée fournies par des capteurs fixes, placés à côté, ou sur la chaussée. Les véhicules connectés peuvent être utilisés comme capteurs mobiles du trafic routier afin d'estimer des temps de parcours, des longueurs de files d'attente. Avec ces capteurs mobiles, de nouvelles méthodes d'estimation du trafic sont mises au point. On distingue les méthodes déterministes et stochastiques. Cette thèse propose d'étendre une méthode stochastique de la littérature, pour estimer divers paramètres du trafic urbain dans le cas des routes à plusieurs voies. Cette méthode publiée pour la première fois par Comert et al. en 2009 utilise le calcul de distribution de probabilités des longueurs de files d'attente. Ces distributions de probabilités sont conditionnées par l'information du nombre de véhicules connectés présents dans la file d'attente, ainsi que l'information de la position du dernier véhicule connecté dans la file d'attente. La longueur de la file d'attente est estimée avec l'espérance de cette distribution de probabilité. Il est démontré que l'estimation de la longueur de file d'attente ne dépend que de la position du dernier véhicule connecté. Dans des articles ultérieurs, les auteurs ajoutent l'estimation de paramètres primaires comme le taux d'équipement de véhicules connectés, le débit d'arrivée. Le cas où une file d'attente résiduelle est présente en début de feu rouge, est également considéré.

On remarque que les systèmes de localisation GPS ne sont pas suffisamment précis pour détecter la position latérale des véhicules sur une route multi-voies : on ne peut pas déterminer sur quelle voie un véhicule est positionné. Ainsi, il n'est pas possible de dupliquer les méthodes existantes pour les routes à une voie, au cas des routes multi-voies. Cette thèse propose de généraliser l'estimation du trafic routier avec des véhicules communicants au cas des routes multi-voies. La partie de la thèse sur l'estimation du trafic est composée de deux articles, dont l'un étudie le cas des routes formées de deux voies et l'autre, le cas des routes formées de trois voies. On démontre par la suite que la méthode est générale et que le cas où l'on considère des routes à trois voies permet aussi de résoudre le cas des routes à nombre de voies supérieur. Ainsi, la thèse propose une méthode générale d'estimation du trafic pour les routes composées d'un nombre de voies quelconque.

On suppose que la proportion de véhicules connectés est variable, et que la demande en transport est faible ou modérée afin que le processus d'arrivée des véhicules puisse être modélisé par un processus de Poisson. Sont également supposées connues la topologie de l'intersection, avec notamment les connections entre les voies entrantes et les routes sortantes de l'intersection. Les pourcentages directionnels des véhicules sont supposés connus. Une méthode pour estimer les pourcentages directionnels consiste à suivre la direction que prennent les véhicules connectés à la sortie de l'intersection. On propose un modèle d'affectation des véhicules

sur les différentes voies où l'hypothèse considère que les files d'attente par voie tendent à s'équilibrer autant que possible. Deux estimateurs du taux d'équipement des véhicules connectés sont proposés et leurs performances évaluées en simulation. Le débit total d'arrivée des véhicules, ainsi que le débit d'arrivée pour chaque voie est formulé analytiquement en dépendance du modèle d'affectation des véhicules établi précédemment. En supposant que le processus d'arrivée est modélisé par un processus de Poisson, une formulation de la distribution de probabilité des longueurs de files d'attente est donnée. Cette formulation dépend du débit d'arrivée par voie et du temps de feu rouge, et n'utilise pas l'information communiquée par les véhicules connectés. On propose un estimateur du nombre de véhicules connectés présents sur chacune des voies. Trois autres formulations de cette distribution de probabilité des longueurs de files d'attente sont analysées. Ce sont des probabilités conditionnelles sur la position du dernier véhicule connecté ainsi que sur le nombre total de véhicules connectés sur la route multi-voies. L'une des formulations utilise l'estimateur du nombre de véhicules connectés présents sur chaque voie.

Les hypothèses sont discutées au moyen de plusieurs outils de simulation. Premièrement, les résultats de simulation sont produits par une simulation bi-directionnelle associant simulateur microscopique de trafic routier et simulateur de communication à événements discrets. La simulation des véhicules connectés bi-directionnelle se caractérise par une simulation de trafic routier qui dépend de la simulation de la communication, et inversement, par la simulation de la communication qui dépend de la simulation du trafic routier. Les simulations de trafic routier microscopique et de la communication interagissent l'une avec l'autre en temps réel. Deuxièmement, afin de s'affranchir du modèle d'affectation du simulateur de trafic routier microscopique, des simulations sont réalisées avec un simulateur de communication à événements discrets. Les véhicules sont alors représentés par des paquets de communication. Chaque véhicule est affecté à une file d'attente associée à une des voies de la route. Le feu de circulation extrait au débit de saturation les véhicules des files d'attente pendant la phase de feu vert. Les résultats de simulation montrent que le modèle peut représenter les cas de déséquilibre entre les longueurs de files d'attente sur les différentes voies. Différents indicateurs de performances montrent l'amélioration de l'estimation des longueurs de files d'attente lorsque sont utilisées les données de localisation communiquées par les véhicules connectés. Comme attendu, les performances des estimateurs sont augmentées avec la croissance du taux d'équipement des véhicules connectés.

L'estimation du trafic routier permet la régulation du trafic routier. Dans cette thèse, on présente deux contributions à la théorie du contrôle routier. La première contribution consiste à étendre la méthode de la littérature Traffic Urban Control (TUC). TUC est un contrôle du trafic centralisé qui peut être appliqué en temps réel sur de vastes réseaux routiers, y compris

en cas de forte demande et de congestion. Un modèle simple considère la dynamique des files d'attente sur chaque route du réseau, qui dépend de la durée de la phase de feu vert. Une méthode d'optimisation cherche à minimiser un critère fonction du nombre de véhicules présents sur chaque route, et de la durée des feux verts. La solution du problème d'optimisation revient à résoudre l'équation de Riccati, qui donne le contrôle optimal en fonction d'une mesure de l'état du trafic sur chaque route.

Une contribution à ce contrôle TUC est apportée dans cette thèse. On propose une nouvelle phase du feu de signalisation gérant l'intersection. Durant cette phase, les véhicules connectés coopèrent entre eux avec de la communication inter véhiculaires sans fils, pour traverser le carrefour. Cette fenêtre temporelle introduit un espace d'accès concurrent, où les véhicules concourent entre eux à traverser le carrefour tout en évitant les collisions. La formulation du contrôle TUC qui prend en compte une telle fenêtre temporelle est donnée. Le nouveau contrôle introduit une fenêtre temporelle où le contrôle d'accès est distribué, ce qui donne un contrôle semi-décentralisé. Des évaluations de performances du nouveau contrôle semi-décentralisé avec des outils simulations de trafic microscopique sont réalisées. Le scénario utilise un réseau routier de type américain, en forme de grille. La fenêtre temporelle d'accès concurrent est implémentée avec le simulateur avec un paramètre qui représente le conflit entre les véhicules. L'équation de Riccati est résolue hors ligne et le contrôle des feux appliqué en temps réel sur les diverses intersections. Dans ce scénario, les résultats de simulation montrent que le nouveau contrôle semi-décentralisé peut résorber la congestion plus rapidement que le contrôle centralisé TUC. Les résultats montrent la sensibilité du contrôle au paramètre représentant le conflit entre les véhicules, et dont le choix pourrait faire l'objet d'études ultérieures.

On propose dans un second temps un algorithme de contrôle des feux utilisant la communication véhicule-infrastructure et qui pourrait implémenter la fenêtre temporelle concurrente du contrôle semi-décentralisé. Cet algorithme suppose qu'une partie des véhicules est équipée de systèmes de communication sans fils et de localisation. Certains feux de signalisation sont associés à une borne de communication sans fils, et permettent une communication entre les véhicules et l'infrastructure. Le feu de signalisation connecté construit une carte virtuelle dynamique de l'intersection en utilisant les trajectoires des véhicules connectés. Ensuite, le feu de signalisation choisit une phase de trafic en fonction de l'état de la carte de l'intersection. Une fois la phase de signalisation appliquée, les véhicules connectés et non connectés respectent le contrôle. L'algorithme est implémenté dans un simulateur de véhicules communicants utilisant l'architecture de communication WAVE basée sur le protocole IEEE 802.11p. La simulation bi-directionnelle VEINS implémente la dépendance du trafic routier en fonction de la communication ainsi que la dépendance de la communication en fonction de la mobilité des véhicules. Une étude de performances est réalisée sur le même réseau routier

en forme de grille utilisé pour la simulation du contrôle semi-décentralisé. Dans ce scénario, les performances de l'algorithme sont déclinées pour le protocole de communication ainsi que pour le transport des véhicules.

La thèse propose une approche complète de l'estimation et la régulation du trafic routier. Le modèle d'estimation du trafic et le contrôle routier semi-décentralisé proposés supposent qu'une proportion des véhicules est équipée de systèmes de localisation et communique en réseau sans fils. Le modèle d'estimation du trafic routier généralise une méthode proposée par Comert et al. tandis que le contrôle du trafic routier généralise la méthode de la littérature Traffic Urban Control (TUC). Un algorithme de contrôle des feux utilisant la communication entre véhicules et infrastructure montre comment implémenter le contrôle semi-décentralisé. Les simulations permettent l'étude des performances des méthodes proposées et utilisent simulateur de trafic microscopique et simulateur de communication. La simulation bi-directionnelle du trafic routier et de la communication est également utilisée pour évaluer les méthodes proposées. Les résultats montrent la pertinence des hypothèses et des modèles puisque des gains sont constatés par rapport aux méthodes classiques. La demande en transport est supposée modérée et le cas où la demande est forte pourrait faire l'objet de travaux ultérieurs.

“In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of Université Gustave Eiffel’s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to [http://www.ieee.org/publications\\_standards/publications/rights/rights\\_link.html](http://www.ieee.org/publications_standards/publications/rights/rights_link.html) to learn how to obtain a License from RightsLink. If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.”



**Part I**

**Introduction**





# Chapter 1

## Introduction and state of the art

### 1.1 Introduction

#### 1.1.1 Context and stakes of road traffic science

In this document, we propose models, methods and algorithms which have the potential to improve the road traffic transportation conditions. Road traffic transportation is a major concern in the cities nowadays and its economical and ecological costs are critical in our societies. For example, in 2020, Schrank et al. [125] have estimated the travel delay in the USA to 4.3 billion hours. The same year, the fuel wasted was estimated to 1.7 billion gallons, the gas emission to 18 million tons and the congestion cost to 101 billion of dollars. In addition, the number of people living in the cities tends to increase and the perspective is a growing demand for transportation. Indeed, in 2014, 53% of the world population lived in the cities and a growth of 14% is expected by 2050 [101]. It is clearly shown that the cost of the transportation has severe impacts on road traffic conditions but also major consequences on the environmental, economical and social fields [130]. Among the effects of transportation systems, we cite the pollutants emitted by the road traffic, the noise, the energy consumption needed to move the vehicles, all having important consequences on the ecological health of the planet, for today and the future. The ecological effects of transportation systems are often balanced with their economical stakes. At first glance, the fast economical growth may seem contradictory with the ecological sustainability needs on a long term perspective. Indeed, the transportation systems enable the offer of mobility services to the society for a multitude of needs which range a wide bandwidth of motivations. By crossing different domains, by addressing various problems, by responding to various motivations and by implying various protagonists, the transportation systems are

complex and central to our modern societies, and the stakes are high.

We define the transportation demand as the need for mobility. The need for mobility raises many important questions and problems for the society and the individual. The economical and ecological problems exposed above may increase in the future due to the persistence of the demand and its evolutions. The transportation demand is an important factor of the transportation systems. The demand can be modeled with the static four step travel model [107]. During the generation phase, the four step model determines the number of trips. The trip generation phase distributes origins with destinations; coupled with the generation phase, it assigns a number of trips to any origin-destination pair. The result of this phase is an origin-destination matrix. The mode choice phase assigns the number of trips per origin-destination pair to a mode of transportation (road, public mass transit, pedestrians, etc...). Finally, the route assignment phase determines the choice of the route for each particular mode according to various principles of equilibrium. The four step level model is widely used for determining the transportation forecasts since the 1950's.

On another hand, the transportation supply is the mobility service which can be offered to the users. It depends on the capacity of the roads and the capacity of the vehicles. Congestions occur when the transportation demand exceeds the supply. Improving service by expanding the supply and building new roads is costly, time consuming and frequently not possible in the cities. Some decades are needed to build new road infrastructures, because the process involves the planification, tactical and operational levels. Furthermore, in the cities, the place available to build new road infrastructures is reduced and very costly. This is why optimizing the actual road network by road traffic management is an interesting alternative. For example, controlling road traffic with new traffic lights algorithms has the potential to improve the transportation system efficiency.

The properties of the equilibrium between the demand and the supply are managed within various processes targeting different horizons of time. A planification process estimates the long term needs of the users. Planification studies for the estimation of the long term demand are based on low frequency data, i.e. the studies are scarcely repeated, for example every 4-5 years for the city of Lyon [41]. On a medium horizon of time, the tactical level of transportation systems management validates the adequation between the supply and the demand. The tactical level updates the variables control of the transportation system every 1-2 year. Finally, the operational level is applied on very short horizons of time (or in real time) to ensure the daily functionality of the transportation system [17]. For example, concerning the road traffic, the operational level adjusts the timings of the traffic lights based on real time data input.

With the new communication technologies such as smartphones and communicating vehicles, the transportation demand can be measured with

greater accuracy than before and the users become more demanding on the quality of the service provided by the transportation systems [161]. The space time granularity and the amount of data given by communicating vehicles are some orders of magnitude higher than before, and the big data era is promised to flourish in the future. For example, in 2025, it is estimated that 463 exabytes (an exabyte is  $10^{18}$  bytes) of data will be exchanged per day. For reference, the total number of words pronounced by humanity until today is estimated to 5 exabytes [41]. The big data era represents an opportunity to improve the transportation systems.

In order to improve road traffic conditions and road traffic effects on the individual, social, environmental and economical fields, we need to manage the road traffic. The road traffic management relies on traffic models. The models can be dynamic and depend on an observation process; through the observation process, we get experimental data used into the modeling and understanding process. Furthermore, this observation process will give input data to the control process which uses the model in order to attain desirable road traffic conditions and road traffic effects. Among the three processes which are observing, modeling, controlling, we see that observing the road traffic can not be neglected. A good observation of the road traffic will allow a better understanding of it; hence, better control decisions towards the desired road traffic states could be taken. In the meantime, minimizing the errors in the road traffic observation will lead to less undesired road traffic conditions and effects. The quality and efficiency of the transportation models rely on the input data.

### 1.1.2 Intelligent Transportation Systems (ITS)

Input data is needed for the estimation and the control of road traffic. Sensors are designed to provide these data and throughout history different kinds of sensors were functionally convenient. In the state of the art section, we will review the different technologies of sensors used since the beginning of road monitoring until today. We distinguish between fixed sensors and mobile sensors. An example of fixed sensor is a magnetic loop placed on the road, or a video camera placed next to the road. An example of mobile sensors are modern vehicles used to travel : connected and automated vehicles (CAVs). A probe vehicle is a vehicle equipped with localization and communication capabilities. A connected and automated vehicle (CAV) is a probe vehicle which is partially or fully automated in the transportation service it renders to its users.

The National Highway Traffic Safety Administration defines 6 levels of automation for the driving task, starting from level 0 which is manual driving up to level 5 which is fully automated driving [5]. Road traffic will include CAVs; furthermore, the penetration ratio of CAVs in the road traffic is expected to increase in the future [113]. Gradually, vehicles will communicate

with other vehicles or devices in Vehicular Ad Hoc Networks (VANETs) [38]. VANETs is a category of Mobile Ad Hoc Networks (MANET) composed of communicating vehicles which use radio transceivers to provide ubiquitous connectivity between vehicles and the road infrastructure. The VANETS communication architectures are various and we distinguish vehicle to vehicle (V2V), vehicle to infrastructure (V2I), vehicle to pedestrians (V2P), vehicle to anything (V2X), and infrastructure to infrastructure (I2I) [38] categories.

V2V enables direct or multi-hop communication between moving vehicles without the need for a fixed infrastructure. V2I enables communication between moving vehicles and a fixed infrastructure such as road side units, cellular base stations, or access points. V2P enables communication between moving vehicles and vulnerable users such as pedestrians or cyclists. V2X combines V2V and V2I multi hop architectures and extends the communication network composed of both moving vehicles and infrastructure. I2I enables communication between different infrastructure units such as road side units (RSU) or backbone networks.

The wireless communication technologies can be divided into the traditional and vehicular communication technologies [31]. Concerning the traditional technologies, we distinguish WIFI [82], WIMAX [55], Bluetooth [19], Zigbee [129], Infrared [104] and Millimeter Waves [23] communication protocols. Each protocol has a unique application space determined by its properties like the radio range, the battery life, the communication rate. Although the traditional communication protocols can be used for VANETS, specific communication protocols are existing. Among the specific vehicular communication protocols, we refer to dedicated short range communication (DSRC) like WAVE communication stack [68], ETSI-5G [45] and CALM protocols [84]. Intelligent Transportation Systems (ITS) [36] use communication, data processing, control technologies in order to improve the efficiency of the transportation systems by reducing the congestion, the gas emissions and the various induced costs for the society or the individual. The large amount of data provided by the connected vehicles [8] has the potential to improve the accuracy of the estimation, the efficiency of the control, the safety, and the ecological footprint of the Intelligent Transportation Systems.

Various applications of ITS are considered, which can be divided into safety, efficiency and comfort applications. Concerning the safety applications, we distinguish between collision warning systems, safe distance warning road obstacle warning, cooperative driving, intersection collision warning, and lane changing assistance [15] [153] [60] [85] [137] [30]. The safety applications are demanding short communication delays and robust communication links because the stakes are high as they concern the health of the individuals. The efficiency applications aim at improving the road traffic conditions and emissions by using VANETS for road traffic management. The application range is addressing “smart traffic lights, variable speed lim-

its, parking management, traveler information systems” [43]. The efficiency application of VANETs is the main topic of this thesis where contributions to urban road traffic estimation and control are proposed. Finally, comfort and entertainment applications tend to increase the services to the drivers and passengers such as multimedia entertainment, in order to make them enjoy the travel [89].

Connected vehicles can send their trajectories data to another vehicle or to the infrastructure. It is an opportunity to use probe vehicles as mobile sensors in order to monitor road traffic state. Combined with models, this information sent by probe vehicles can be used to estimate road traffic state and conditions.

### 1.1.3 Assumptions and tools

This thesis proposes models for the estimation and control of road traffic. “Mathematical modeling of the dynamics of a complex system is done by determining or identifying a dynamic system describing its evolution.” [52]. Road traffic models can be divided as deterministic or stochastic depending on whether there is an uncertainty, discrete or continuous depending on the types of the variables involved, linear or non linear depending on the type of algebra, etc. Road traffic can also be represented as queueing networks such as in the backpressure control algorithm [65]. In this thesis, we consider for the estimation of the road traffic that some queues are formed at junctions controlled by traffic lights. The main purpose is to estimate the length of the queues. Concerning the road traffic control, models can be divided into the macroscopic, microscopic and mesoscopic models according to the scale of the phenomena observed. Macroscopic models consider large aggregation of vehicles and their global behaviour, although microscopic models consider individual vehicles. The hybrid mesoscopic models consider individual vehicles and the behaviour of their global aggregation [37]. The variables describing the traffic are different among the three kinds of models. Among the macroscopic variables there are the car density, the car flow, the average speed. Among the microscopic variables concerning the individual vehicles, there are the individual car speed, the space car headway, the relative speed, the car acceleration, the driver behaviour, etc.

We also notice that the road traffic models depend on whether the road traffic is on a highway or in an urban center. In the case of the urban road traffic, there are some particularities such that the frequent presence of traffic lights and junctions, although in the case of highways, ramp metering, tolls, or varying speed limits are distinctive properties. In this thesis, based on a road traffic model, we consider urban road traffic estimation and control.

The society needs experimental tools in order to study the VANETS and ITS, and some simulation frameworks enable numerical studies of such vehicular communication systems. As we do not have experimental data,

simulation studies to evaluate and discuss the performances of the contributions, have been used. Simulation studies were conducted with Veins framework [131] which combines the microscopic road traffic simulator SUMO [86] and the communication simulator OMNET++ [142]. This framework enables the simulation of VANETS by coupling a road traffic and a communication simulator in a bidirectional way. With this kind of coupling, the communication simulation depends on the mobility of the communicating nodes, and the vehicular mobility depends on the communication between the vehicles. iTETRIS is the other framework [124] (less frequently used than VEINS) which aims at simulating VANETS : it combines road traffic SUMO and the NS3 communication [121] simulators.

The main assumptions of this thesis are the presence of probe vehicles with a varying penetration ratio. The GPS localization system of the probe vehicles is assumed to be not accurate enough to enable the assignment of probe vehicles onto the lanes, i.e. we can not know on which lane the probe vehicles are located. This is a critical point especially for the estimation of road traffic since this prevents from using existing methods from the literature. This GPS inaccuracy is the reason why a contribution to existing methods can be proposed, by addressing the case where roads are composed of many lanes. Concerning the control of the road traffic, it is assumed that the queue lengths are inferior to the road length and that the vehicles cooperate with V2V communication in order to pass the junction during contention time windows. A possible implementation of the contention time window is detailed in the last chapter of this document.

#### 1.1.4 Outline and contributions

After the introduction which has presented the context and stakes of road traffic, we present here the organization of the manuscript. In section 1.2.1, we give a literature review of methods which use the data given by the sensors in order to estimate road traffic state. The methods rely on the type of input data. Historical methods were using fixed sensor data (section 1.2.1.1) although new methods use mobile sensors data (section 1.2.1.2). The historical methods are reviewed in section 1.2.1.3. Among the new methods using connected vehicles data (section 1.2.1.4), we distinguish between deterministic and stochastic methods.

We detail in part II methods which enable the estimation of the road traffic with probe vehicles in the context of ITS. As a contribution to the state of the art on the estimation of road traffic state, this thesis extends a stochastic method published by Comert and Cetin [28] in 2009. Comert and Cetin have addressed the estimation of road traffic state in the case of roads composed of one lane. The objective of part II is to consider roads which are composed of many lanes. The part is composed of two articles published in 2021 [111] and 2022 [112]. This part addresses the case where

roads are composed of two lanes in chapter 2, and the general case where the roads are composed of any number of lanes in chapter 3.

On another hand, based on the road traffic state measurements and estimations, road traffic control can be used in order to improve the experience of transportation, either from a global or individual point of view. One interesting way of controlling road traffic is by controlling the traffic light signals which manage road junctions. We distinguish several methods for controlling the traffic with traffic lights signals : fixed time, planning based, and signal vehicle coupled controls. By using probe vehicles as sensor input, these new methods improve traffic light control algorithms and the transportation system efficiency. In this thesis, two methods for controlling road traffic lights with the data given by probe vehicles, are proposed. These methods build upon existing centralized and distributed control methods and are the material for part III. The parameters of the traffic light (cycle time, phase specification, etc.) can be fixed, or controlled in response to measurements of the road traffic in an adaptive way. In section 1.2.2, a literature review on road traffic control by using traffic lights is given. Different fields are reviewed. We consider actuated traffic signal in section 1.2.2.1, where the traffic light timings are set in response to road traffic state. Compared to actuated traffic light control, planning based traffic signal control (section 1.2.2.2) introduces some prediction on the future road traffic state. Vehicles form platoons and the identification of such groups of vehicles can lead to better traffic light controls. Platoon based traffic signal control (section 1.2.2.3) uses the identification or the formation of platoons of vehicles in order to optimize the timings of the traffic light. Giving priority to buses is also an option to improve the average delay per passenger. Transit priority control (section 1.2.2.4) sets the timings of the traffic light by giving the priority to public transit. Finally, coordinated control (section 1.2.2.5) which couples the control of the traffic light and the control of the connected and automated vehicles is reviewed.

In part III, two published conference papers which address the control of traffic light problem are given. In chapter 4, a conference paper [53] extends a centralized approach (the Traffic Urban Control method [33, 35, 34]) by introducing some contention window where the vehicles communicate with each other to pass the junction. In chapter 5, a conference paper [139] details how such a contention window could be implemented with communicating vehicles. Finally, we conclude in part IV by summarizing the work and giving perspectives.

The list of the publications included in this thesis is given below.

- Chapter 2 : ©2021 IEEE. Reprinted, with permission, from Cyril Nguyen Van Phu and Nadir Farhi. Estimation of urban traffic state with probe vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 22(5):2797–2808, 2021



- Chapter 3 : ©2022 IEEE. Reprinted, with permission, from Cyril Nguyen Van Phu and Nadir Farhi. Estimation of road traffic state at a multilanes controlled junction. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–0, 2022
- Chapter 4 : Nadir Farhi, Cyril Nguyen Van Phu, Mouna Amir, Habib Haj-Salem, and Jean-Patrick Lebacque. A semi-decentralized control strategy for urban traffic. *Transportation Research Procedia*, 10:41–50, 2015. 18th Euro Working Group on Transportation, EWGT 2015, 14-16 July 2015, Delft, The Netherlands
- Chapter 5 : ©2017 IEEE. Reprinted, with permission, from Cyril Nguyen Van Phu, Nadir Farhi, Habib Haj-Salem, and Jean-Patrick Lebacque. A vehicle-to-infrastructure communication based algorithm for urban traffic control. In *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 651–656, 2017

Excepted for chapter 4 (semi-decentralized control) which is the result of an equilibrated collaborative effort, Cyril Nguyen Van Phu is the main contributor to the mathematical and modeling parts as well as the simulation parts, for all the chapters. Concerning chapter 4, Cyril Nguyen Van Phu has done the simulation program and experiments, and contributed to the algorithm implementing the time window, while Nadir Farhi has done the theoretical part, especially the dynamics of the linear quadratic regulator.

## 1.2 State of the art

### 1.2.1 Estimation of road traffic

The estimation of road traffic state depends on the type of the input data. Various sensors exist which can be classified as fixed or mobile sensors. For each type of sensors, the methods differ. Historically, methods were using fixed point detectors although in the future, connected vehicles are promising new paradigms for estimating road traffic state. In this review, the main categories of sensors and the associated methods to estimate road traffic state are described. We distinguish between fixed and mobile sensors, and between the associated historical methods and contemporary methods.

#### 1.2.1.1 Sensing road traffic with fixed sensors

**Intrusive sensors (placed on the road)** Traditionally, road traffic was observed from the outside by fixed sensors which can be classified as intrusive or non intrusive. “The intrusive methods basically consist of a data recorder and a sensor placing on or in the road.”, say Leduc et al. [90]. Among the intrusive sensors, Leduc et al. [90] cite :

- Pneumatic road tubes which use the principle of air pressure change in order to detect vehicles passing over the tube.
- Piezoelectric sensors which convert mechanical energy (induced by the mechanical deformation when a vehicle is passing over the sensor) into electrical energy. This kind of sensor can measure weight and speed.
- Magnetic loops which generate a magnetic field. When a car is passing over the sensor, the magnetic field is modified. This kind of equipment has been widely deployed over the last decades, even if its costs can be expensive.

**Non intrusive sensors (placed out of the road)** Non intrusive sensors are placed out of the road. Among the non intrusive sensors, the authors of [90] cite :

- Manual counting using count boards and paper.
- Passive and active infra-red sensors. It is said in [90] that “the presence, speed and type of vehicles are detected based on the infrared energy radiating from the detection area. The main drawbacks are the performance during bad weather, and limited lane coverage.”
- Passive magnetic : these sensors are placed next to the road bed. A passive magnetic field detects the vehicles passing on the road.
- Microwave radar which use the Doppler effect principle. These sensors can detect the speed of the vehicles and count them.
- Ultrasonic and passive acoustic sensors which use the time-to-travel of the acoustic wave between the sensor and the vehicles. They are placed next to the road. These sensors can detect the speed of the vehicles and count them.
- Video image detection which use video camera and image processing techniques in order to count the vehicles, estimate queue lengths at junctions, detect vehicle types.

These intrusive and non intrusive sensors are monitoring road traffic from the outside since these sensors are usually not moving and do not participate directly to the transportation of goods or passengers.

### 1.2.1.2 Sensing road traffic with probe vehicles

Using probe vehicles as mobile sensors, road traffic can be observed from the inside. The localization and communication techniques used by such vehicles are variable.

Localization techniques using GPS localization system are the most deployed. It is also possible to use cellular mobile phones of the passengers and triangulation techniques, in order to detect the localization of vehicles. In this case the accuracy is very low, around 300m. The algorithms involved are sophisticated, but there is a huge number of devices available [90]. Access to cellular mobile phones is restricted to few entities and companies, which face privacy issues.

Different communication techniques can be used by the vehicles. Each of these communication techniques has specific communication range and connectivity properties. Beside cellular communication, short range communication such as the IEEE 802.11p protocol [78] has emerged as standard communication protocol. On top of IEEE 802.11p protocol, higher communication layers such as WAVE [2] or ETSI ITS-G5 [45] standards for inter vehicular communication (IVC) complete the communication landscape. Within this framework, a connected vehicle can communicate with another vehicle (V2V), the infrastructure (V2I), or any other device such as mobile phones (V2X).

### 1.2.1.3 Estimation of road traffic state with fixed sensors (historical approaches)

In this section, we provide an overview of the main historical approaches that have been used to estimate queue lengths. Although the potential of fixed point detectors is declining, magnetic loops are still very used today. Hence, the approaches which rely on fixed point detectors can be categorized as historical approaches. The historical approaches rely on the modeling of the car arrivals as well as the dynamics of the queuing process, and we distinguish in this section between the cases where the overflow queue is taken into account, or not.

Concerning the undersaturated traffic demand where the overflow queue is not considered, in 1936, Adams [4] was one of the first to propose that the Poisson process probability model describes the inter-arrival times of cars at a road junction. Experimental measurements of the arrivals of cars at a road junction have confirmed the adequacy of modeling cars arrivals as a Poisson process random series. In 1955, Beckmann et al. [13] gave an analytic formulation of the mean queue size of vehicles waiting at a stop sign or at a traffic light with fixed cycle. The authors assumed that the cars arrivals at the road junction is following a binomial process. The authors consider a junction with two incoming roads : a minor and a major road. In the case of a stop sign at the road junction, the time gaps where vehicles of the minor road can cross the junction are considered as green times, as opposed to the red times where the time gaps are not long enough for the vehicles to cross the junction. The authors assumed that there is no overflow queue : at the beginning of the green time, the queue has cleared such

that the queue does not grow indefinitely and an equilibrium is established. Addressing the various cases involved with these assumptions, mathematical expressions of the mean queue size at any time point were given by the authors. In 1968, McNeil [108] extended the approach from Beckmann et al. by giving an analytic formulation of the mean delay for vehicles waiting at a traffic light in the case of more general arrivals distributions (including the simple Poisson process). However, McNeil's solution needs the formulation of the average overflow queue which is not easy to estimate and which he gives only a close upper bound.

In this paragraph, we review the methods which take into account the cases where the overflow queue is also considered. In 1958, Webster [148] gave results for the estimation of the queue length at signalized junctions. Its estimations are done using random input flows for the saturated and unsaturated cases, with the additional information of the signal timings and delays induced when vehicles cross the considered arterial. The objective was to determine the delay at the road junction; in order to achieve this goal, the estimation of the queue length at the traffic light was performed. The mathematical expression of the delay was computed to fit numerical simulations using the knowledge of the inflows. In 1968, Miller [110] gave an expression of the average overflow queue, in the case of undersaturated traffic demands (i.e. when the arrival flow exceeds the intersection capacity). In 1980, Akcelic [7] extended the work of Miller to the cases of oversaturated traffic demands (i.e. when the arrival flow exceeds the intersection capacity). Among the input-output class of methods, Viti and Van Zuylen [145] present a probabilistic model for the estimation of queue lengths at signalized junctions, which can capture spillback and gridlock phenomena. Viti and Van Zuylen propose to establish a probabilistic model to estimate queue lengths and delays at junctions which relax some common assumptions made by former works. For this purpose they use Markov chains theory and "the specification of the input demand and service rates within a cycle in terms of probability distributions, and an initial queue state." They derive the back of the queue expression and validate their results with VISSIM microscopic traffic simulator. In the similar way, by using input output methods, in 2008, Vigos et al. [143] used at least three loop detectors combined with a Kalman filter in order to produce estimates on the number of vehicles driving on a signalized link. In 2009, Liu et al. [102] took advantage of shockwave traffic theory, combined with fixed detector and signal timings data input, in order to estimate queue lengths. They also consider the case of congested arterials, when the detector is not able to sense vehicle arrivals (since the detector is located in the queue in this case). Their method relies on the identification of break points which are times where traffic conditions change with discontinuity. Their research team has implemented their method in the field; a firm has also implemented and evaluated their method. The authors state that "It is very clear that the proposed models successfully describe queue

forming and discharging processes.”.

#### 1.2.1.4 Estimation of road traffic state with probe vehicles

With the introduction of probe vehicles, new methods have been published which use the information from the mobile sensors (i.e. probe vehicles). The data provided by these mobile sensors are quite different in their nature from the data provided by fixed sensors. With probe vehicles, we get data concerning individual sample vehicles trajectories, rather than global information on traffic state at a fixed location. This difference in the nature of the information furnished, raises the opportunity to develop new methods for queue length estimation. In the following paragraphs, we give an overview of the methods that have been developed upon the information given by these new mobile sensors, namely, probe vehicles. In [67], the authors provide a timely survey on traffic information collection and state estimation methods published in the last decade, which use the data provided by connected and automated vehicles (CAVs). They classify the different traffic observation methods which use CAVs data by distinguishing between deterministic and stochastic approaches. In this paragraph the same outline is applied.

**Estimation of queue lengths with probe vehicles : deterministic approaches** In this paragraph, the deterministic methods to estimate queue lengths using probe vehicles are reviewed. Among the methods based on shockwave theory, Ramezani and Geroliminis [118] expose a method to estimate queue lengths with probe vehicles as the single source of information. Position and instantaneous speed of probe vehicles are the input data of their method. Their method relaxes some common assumptions made in the literature, such as the knowledge of signal timings or arrival process distribution. The shockwave theory based on LWR model is the key model used in their queue length estimation method.

In [10], Ban et al. used intersection travel times in order to estimate queue lengths and delays at junctions. These intersection travel times are measured when probe vehicles cross virtual trip lines (VTL), located upstream and downstream relatively to the intersection. The benefits of using intersection travel times are : respect the privacy of the users, the flexibility in defining the virtual trip lines (as they are virtual locations), and the pliancy which enables other sensor inputs such as Bluetooth Mac address matching, and other travel times collection systems. The authors introduced “the concept of Queue Rear No-delay Arrival Time which is related to the non-smoothness of queuing delay patterns and queue length changes.” Queue Rear No-delay Arrival Time is a time unit measurement and enables computation of the minimum and maximum queue lengths in a cycle. Then, the authors assume that the queue length is linear between these two critical

points, which allows them to finally conclude for the queue length estimation during a cycle.

In 2015, Tiaprasert et al. [136] proposed to estimate the queue length at a road junction with probe vehicles. Their method does not need the knowledge of the traffic light timing neither of the arrival process. In addition, the method is useful both for undersaturated and saturated volumes of traffic demands. The speed of probe vehicles is used as data input and this enables to bound the queue length with a minimum and a maximum. The least mean square error algorithm is then used to give an estimate between these two bounds. In order to correct the estimation error done in the case of low penetration ratio of probe vehicles, the authors filter the queue length data. They discriminate low and high frequencies in the queue length estimation, which are respectively assumed to be the signal and the noise. The authors have used the Discrete Wavelet Transform (DWT) method for denoising the queue length data.

In 2011, Cheng et al. [22] proposed to identify critical points in the queuing process with a dedicated algorithm. The critical points are related to the queue formation and dissipation. In the time-space coordinates, critical points may relate to points when a vehicle decelerate because of the red signal, a vehicle stops and joins the queue, a vehicle accelerates because of a green light, a vehicles slows down because of a discharging queue. In their work, the authors filter negligible critical points and focus on four categories of critical points. Using shockwave theory and the critical points, the signal timings of the traffic light and different queue length estimates are formulated. The promising results are discussed based on the evaluation of the work done with three data sets.

In 2018, Yin et al. [159] proposed a similar method than [22] : they identify critical points in the time space coordinates. The knowledge of these critical points combined with shockwave theory gives an estimation of the queue length. The queue length estimation is finally processed with kalman filtering. They have tested their method on simulated and real world data. “The proposed method is tested in both real-world case and simulation scenarios, and the MAE/MAPE of the maximum queue length is 18.3 m/11.2% in real-world situation at a penetration rate of 7.4%.”

In 2017, Li et al. [93] proposed a method which estimates queue lengths without the common assumptions which are the knowledge of the signal timings, the arrival patterns and the penetration ratio of probe vehicles. Their method is also based on the shockwave theory. Compared to the method from Ramezani and Geroliminis [118], Li et al. [93] method proves to be more robust and more accurate.

**Estimation of road traffic state with probe vehicles : stochastic approaches** In addition to the deterministic approaches to estimate the

queue lengths with probe vehicles, we review in this section the stochastic methods. In 2013, Hao et al. [71] define a vehicle index as “the position of vehicles in the departure process of the cycle”. That paper has proposed a method for estimating these vehicle indices which are described as a basic and primordial information that can be provided by probe vehicles. For example, knowing the index of a vehicle gives its position in the queue. With the intent of solving some privacy issues, their method relies only on intersection travel times as input data. They derive the intersection travel times from the arrival time and departure time of probe vehicles into virtual areas (Virtual Trip Lines, VTL) respectively upstream and downstream the intersection. They model the arrival process as a time-dependent Poisson process; they use a log-normal distribution to model the departure headways for every vehicle index. In this framework, the authors use a Bayesian Network in order to estimate vehicle indices.

Vehicle indices are some basic information that can be used as input data for estimating queue length at junctions. Indeed in 2014, Hao et al. [70] have naturally pursued their work on vehicle indices with the estimation of queue lengths at intersections. They classify traffic states in three main cases, each case being subdivided into two or three sub-cases. Then, using as input data the intersection travel times and the vehicle indices, as determined in their former 2013 paper, they estimate queue lengths using a stochastic model based on Bayes theorem. The method is evaluated with field experiment data and produces results described as “best estimates when compared to the linear fitting method (Ban et al., 2011) and queue location method (Hao et al., in press) developed previously in the literature, especially in term of the success rate.”

In 2009, Comert and Cetin [28] proposed a method for the estimation of queue length using the location data provided by probe vehicles. They have assumed that probe vehicles indices are available as input data. The results are presented for an isolated intersection and undersaturated traffic flow input. Assuming that the probability distribution of the queue length is given, they compute a conditional probability distribution of the queue length, knowing the locations of the probe vehicles in the queue. They show that for the 1-lane case, the location of the last probe vehicle in the queue is the only one needed. The estimation of the queue length is performed by computing the expectation of this conditional probability. Comert and Cetin introduce the error random variable which is the difference between the queue length and its estimated. They derive an analytic formulation of the variance of the error, depending on the penetration ratio of probe vehicles. Furthermore, numerical analyses are performed, where the arrival processes models and arrival processes intensities are varied. The results show that obviously the error diminishes as  $p$  increases and that “the percentage error relative to the mean decreases significantly for  $p$  values larger than 10%.” This work “appears to be the first attempt to formulate the problem of

estimating the queue length from probe vehicle data.”

In 2011, Comert and Cetin [25] have continued their work by using the time when the last probe vehicle joins the queue. With this new data given by probe vehicles, they have proposed an estimation of the queue length assuming Poisson process arrivals. They have given analytical formulations of the variance of the error in estimating the queue length, depending on time joining and location of the last probe vehicle for the case with no overflow queue. In the case where the overflow queue is considered, the analytical formulation of the variance of the error requires the marginal distribution of the overflow queue as input. The authors give an approximation for this the marginal distribution of the overflow queue and discuss their mathematical formulations with simulation results. They also have analyzed “the relationship between the percentage of probe vehicles  $p$  and the accuracy of the estimates.”

In 2013, Comert [26] derived queue length, last probe location and queue joining times probability distributions, with or without overflow queue (residual queue at the end of the red time). Mean and variance for the queue length estimators are given. It is shown that these estimators depend on probe proportion, red duration and arrival process properties. In addition, a simple estimator for the probe vehicles proportion has been proposed.

In 2016, Comert [27] goes a step further by studying the cases with unknown probe proportions and unknown arrival rates. He gives analytical formulations for these primary parameters (proportions of probe vehicles and arrival rate), as well as various queue length estimators with or without overflow queue. Derivation of its estimators errors are also given, and numerical analysis performed with VISSIM microscopic simulator are presented.

In 2017, Zheng and Liu [164] estimated traffic volumes for low penetration ratio of equipped vehicles. The method proposed to use as input data “vehicle trajectories approaching to an intersection as well as traffic signal status”. The trajectories of equipped vehicles are used to detect if a probe vehicle has stopped at the traffic light and its stopping position. With these information, the arrival rate is estimated and bounds for this arrival rate are given. Zheng and Liu have used for their estimation a time dependent Poisson arrival process and the Expectation Maximization (EM) algorithm. They have tested their method with data sets from an experiment where around 2800 probe vehicles were deployed in the city of Ann Arbor, and from data provided by commercial navigation service in China.

Many methods require the knowledge of the penetration ratio of probe vehicles and of the queue length distribution. In 2021, Zhao et al. [163] propose a maximum likelihood estimation method for the estimation of these two primary parameters. They solve the estimation problem with the expectation-maximization algorithm. Their validation results show that these parameters can be estimated accurately.



In order to predict traffic flow, Lv et al. [105] used a data driven machine learning method : the stacked auto encoders model combined with the data provided by 15000 individual detectors deployed across California. The authors declare : “Thus, the effectiveness of the SAE method for traffic flow prediction is promising and manifested.”

### 1.2.2 Road traffic signal control with connected and automated vehicles

The estimation of road traffic state is an input to road traffic control and management. In [67], Guo et al. give a survey on road traffic control using connected and automated vehicles. This review is completed with recent papers, and with the same organization for the state of the art on urban road traffic control using traffic lights. The control of traffic light signal can be organized in five paragraphs :

- actuated signal control, where the traffic light control adapts in real time to traffic conditions but does not predict traffic state in a future horizon of time
- planning based signal control, which introduces some prediction on the road traffic state
- platoon based traffic signal control focuses on the identification of platoons of vehicles in order to improve the efficiency of the traffic light control
- transit priority control aims at setting the traffic light timings in a manner which favours public transit (buses)
- signal coupled control is a relatively new topic where trajectories of connected and automated vehicles are controlled in a coordinated manner with the traffic light signal

The methods generally tend to solve an optimization problem where a disutility function such as the average vehicle delay should be minimized. Some regular trade offs involve distributed versus centralized approach, isolated junction versus arterial control. The penetration ratio of connected and automated vehicles can be a variable, or fixed at 100%. The road traffic simulators are microscopic simulators among SUMO, VISSIM, AIMSUN. In the next paragraphs, the state of the art of traffic light control with connected and automated vehicles according to these highlights is reviewed.

#### 1.2.2.1 Actuated traffic signal control

In actuated traffic signal control, the parameters of the traffic light are set in response to the road traffic state changes. This kind of control performs

better than fixed time traffic light control since it adapts in real time to the road traffic conditions. Historically, the input data was given by loop detectors located at a fixed point on the incoming road of the junction. With probe vehicles, new methods solve the defaults of the fixed point location of the measurements. Indeed, with probe vehicles, the measurements are the trajectories of individual vehicles, which enable new controls. Compared to planning based traffic signal control, actuated traffic signal control does not predict road traffic demand in a future time horizon.

In 2016, Day and Bullock [32] have investigated the use of probe vehicles as a substitute for loop detectors. In their study, they have varied the penetration ratio of probe vehicles. By doing this, they first have studied road traffic state measurement given some probe vehicles penetration ratio. Then, they have applied this estimation of the arrivals to control traffic lights on a corridor. The authors have shown that “acceptable performance was obtained even at relatively low values of  $p$ .”

In 2013, Goodall et al. [61] have proposed an algorithm which adjusts the timings of a traffic light using trajectories input from probe vehicles. The algorithm minimizes an objective function which is computed with microscopic simulation. The authors have tested their algorithm on a road test network composed of four intersections. The authors have declared that “The algorithm showed much greater improvements during unexpected demands”.

In 2007, Gradinescu et al. [64] have proposed a system which uses dedicated short range communication between vehicles and infrastructure in order to control the traffic light of a junction. The authors have developed a simulator which couples microscopic road traffic and communication simulators in order to evaluate the performance of their system. The performances have proven to have “clear benefits compared to adaptive systems based on sensors or cameras.”

In 2014, Kari et al. [83] have proposed an adaptive traffic light control strategy based on the data sent by probe vehicles. Their algorithm is a multi agent algorithm which aims at minimizing some criteria computed with the queue lengths data at the traffic light. They have evaluated their work with SUMO [86] and they have shown that compared to control strategies based on point detectors, “the proposed strategy exhibits significant savings of 4-61% in reducing travel time and 2-32% in reducing system-wide fuel economy for the varied demand scenario. The results clearly demonstrate the robustness of system to significant variation in traffic demand.”

In 2015, Wu et al. [152] have formulated the intersection management problem with connected vehicles as a variant of the classic mutual exclusion problem. They have proposed to solve the new variant of the mutual exclusion problem with two different algorithms : one centralized and the other distributed. Finally, the authors have declared that “performance evaluation by simulations shows that our algorithms can handle various traffic

cases very well, with very little message cost.”

In 2016, Younes and Boukerche [11] have proposed two algorithms which use information given by connected vehicles in order to control traffic light at an isolated junction and to control traffic lights on a corridor. They have used ns-2 communication simulator and SUMO microscopic road traffic simulator in order to evaluate the performances of their algorithms. They have shown that the throughput is increased by 30% on a junction and by 70% on an arterial street.

To conclude, actuated traffic signal control can take advantage of the trajectories data provided by connected vehicles compared to the historical point detectors. The timings parameters of the traffic light at an isolated junction can be adjusted in real time to adapt to the road traffic demand. Some papers have also addressed the arterial case where a serie of traffic lights is synchronized. However, actuated traffic light signal control does not address the road traffic state prediction for a future time.

### 1.2.2.2 Planning based traffic signal control

In planning based traffic signal control, road traffic at an individual vehicle level is predicted on a future time horizon in order to set the timings parameters of the traffic lights. We distinguish centralized and distributed systems. Among the historical centralized systems, we cite SCOOT [77, 18], SCATS [103], RHODES [74], MOTION [20], and TUC [33]. For historical distributed responsive urban traffic controls, we cite UTOPIA [39], PRO-DYN [46], OPAC [58].

In 2013, Lee et al. [91] have proposed an algorithm to control traffic light signal at an isolated junction using cumulative travel time input data. The cumulative travel time is estimated at various equipment penetration ratios with Kalman filtering. The phases of the traffic light are set according to this cumulative travel time estimation. The authors have evaluated their algorithm under various traffic demand levels. The authors declare “that at least 30% market penetration rates are needed to realize the benefits of the CTR algorithm.”

In 2015, Feng et al. [54] have proposed an algorithm which enables real time adaptive phase allocation in a connected vehicles environment. The problem is formulated as a two layers optimization problem. The data input is given by a prediction of the arrivals of vehicles, both connected and not connected. They have evaluated their method on a real world intersection using VISSIM simulator. “Results showed that the proposed algorithm reduced total delay significantly under high penetration rates and was comparable to actuated control under low penetration rates.”

In 2016, Chen and Sun [21] have proposed “an adaptive control method for urban isolated signalized intersection”. Based on the estimation of the vehicle arrivals with an upstream detector, they have solved an optimization

problem with dynamic programming. The results show that their algorithm outperforms a fixed time control. They plan to incorporate the use of data given by connected vehicles in the future.

In 2019, Li and Ban [94] have formulated the traffic light signal optimization setting as a mixed integer non linear program. They have assumed that all the vehicles are connected vehicles (100% penetration ratio) which send by broadcast their trajectory data. Then, they have approximated the problem as a dynamic programming formulation and have solved the problem using a branch and bound method. Compared with another solver, the authors have noticed that their model is more favorable when dealing with large scale problems.

Planning based traffic signal control can be extended from the case of an isolated junction to the case of an arterial road. In 2017, Beak et al. [12] provide two algorithms respectively for the control of an isolated junction in a connected vehicles environment and also for the control of an arterial. At the junction level, dynamic programming is used to solve the green time allocation problem although a mixed integer linear program adjusts the offsets at the corridor level. Their model was evaluated with VISSIM simulator in comparison to a coordinated actuated signal control. The results indicate that their model reduces the delay and the number of stops on a corridor

Concerning the distributed control strategies, in 2017, Islam and Hajbabaie [81] have proposed a solution to the distributed coordinated problem for the network signal control in a connected vehicles environment. They have assumed that all the vehicles are connected and that the junctions can share information. The optimization problem consists in maximizing the throughput and penalizing the queue lengths. By using a decentralized approach, the complexity of the optimization problem is significantly reduced. The authors have used VISSIM road traffic simulator to compare their method with a coordinated fixed time and a coordinated actuated signal setting. The results show that the throughput is increased and the travel times are decreased in all the cases which were studied.

In 2020, Islam et al. [80] have extended their previous work by proposing two algorithms. In addition to connected vehicles, the authors have assumed the presence of a fixed point detector. By using the fixed point detector and a car following model, the first algorithm estimates the trajectories of un-equipped vehicles. The second algorithm provides estimates of the vehicular occupancy on network links. Each junction controller solves an optimization problem coordinated with adjacent junctions in a distributed strategy. Extensive simulation results confirm the relevance of this distributed approach.

In 2020, Rafter et al. [116] have also proposed a distributed method which uses a combination of data given by a fixed point detector and connected vehicles in order to control traffic lights. The method proposed by Rafter et al. can adapt to environment with a low penetration ratio of connected vehicles. They have also tested their control strategy with a framework

which is using the city of Birmingham as testbed. Their algorithm was proven to be robust in cases of heavy demand and non ideal communication environments.

Among the distributed control strategies, we cite the backpressure routing algorithm [150] [66] which was initially designed for the communication networks. The pressure is defined as the current flow rate of the traffic movement, weighted by the difference between the number of vehicles on the two traffic movements from the considered and the following junctions. The algorithm computes the pressure for each phase and selects the phase with the highest pressure. The system needs only the information from the adjacent junction. It has been proven that the maximum throughput is achieved at a network level with this distributed algorithm.

In conclusion, planning based traffic signal control introduces the prediction of road traffic state in a future horizon, which may improve the performance of the control. However, as the optimization problem tends to get more complex, the computation costs can be significant and penalize the method. Distributed control strategies help to reduce the cost of such planning based control such as in the backpressure algorithm which has been proven to reach the maximum network throughput.

### 1.2.2.3 Platoon based traffic signal control

Similarly to planning based control, platoon based traffic signal control needs a prediction of road traffic conditions. Platoons are groups of vehicles which can pass the junction at the same time, increasing the traffic throughput. Traditionally, platoons were detected with fixed location sensors such as in [77, 18] but with connected vehicles, the platoons can be recognized or formed with more accuracy.

In 2017, Lioris et al. [100] have considered the case where platoon formation is supported by inter vehicular communication and adaptive cruise control (ACC), that would reduce the headway which separates following vehicles. In their paper, it is considered that all vehicles would be connected vehicles. Within this framework, the authors have shown that platoon formation enabled by connected vehicles can increase the junction capacity by multiplying the saturation flows rates by a factor 2 or 3. The authors have analyzed this proposition with simulation on a road network composed of 16 junctions and they have shown that without changing the control, the signalized road network can handle a traffic demand multiplied by a factor 2 or 3.

In 2013, Pandit et al. [114] have addressed the isolated junction control problem with connected vehicles. With the help of the trajectories data sent by wireless communication, vehicles are grouped into equal sized platoons. Then, the authors apply to the platoons an algorithm referred as oldest job first (OJF) in order to minimize the average delay for vehicles

crossing the junction. The authors have evaluated their algorithm with simulation coupling SUMO road traffic microscopic simulator and OMNET++ communication simulator. They have tested their algorithm with various arrival rates and penetration ratio of connected vehicles. It was shown that their algorithm reduces the delay for light and medium traffic demands, and performs the same as an actuated traffic signal control under heavy demand.

Although Pandit et al. [114] have addressed the isolated junction case, in 2012, He et al.[72] have proposed a formulation to minimize the delay of platoons of vehicles crossing an arterial. They have assumed a connected vehicles environment. Connected vehicles are representing a proportion of the road traffic vehicles and the model proposed by the authors, combined with the exchanged data, enable to recognize platoons of vehicles. The objective is to minimize the total delay of vehicles crossing the arterial, with respect to constraints such as the strategy to coordinate the platoons passing along the different junctions, or platoon splitting when needed. The optimization problem is formulated as a mixed integer linear program with a rolling horizon. The authors have evaluated their work with VISSIM simulator and have shown that their algorithm significantly reduces the delays, both for undersaturated and oversaturated demands.

In 2018, Liang et al. [95] have shown that grouping vehicles into platoons can also reduce the computational cost of the various control strategies. In their work, they assume a mixed traffic composed of autonomous, connected and unequipped vehicles. After grouping the vehicles into platoons, a traffic signal optimization is performed in order to reduce the vehicles delay. A control strategy of the trajectory of the lead autonomous vehicles of the platoons has also been proposed in order to reduce the number of stop maneuvers. The simulation results have shown that the performances of the algorithm increase with the penetration ratio of equipped vehicles, until the fleet is composed of 40% of connected and automated vehicles.

In 2011, Xie et al. [154] have proposed a scheduling algorithm in order to control road network in real time. The data given by sensors is used to group vehicles into platoons and anticipated queues. Based on the information given by the junctions of the road network, the self scheduling algorithm performs a choice between the decision of extending or terminating the green phase. The algorithm tends to favour “green waves” along arterials in the road network. The authors have implemented their work within SUMO road traffic microscopic simulator and they have compared it with a pretimed plan and two adaptive strategies with the best results.

In 2020, Yao et al. [158] have proposed a dynamic platoon dispersion model in order to predict the arrival times of connected vehicles. The authors have formulated the signal setting problem as an optimization problem where the average vehicle delay should be minimized. They have solved the optimization problem with a genetic algorithm which enables real time adaptive control. Successful simulation experiments were performed with

VISSIM simulator for various penetration ratios of connected vehicles.

In 2020, Liang et al. [96] have developed a method to optimize the phase durations and sequences of traffic lights at isolated junctions. They use the data given by some vehicles which are communicating their trajectories wirelessly. With this information, platoons of vehicles (connected and not connected) are identified in the traffic stream. The optimization problem consists in minimizing the average delay of all vehicles. In order to solve this optimization problem, the authors use various heuristics such as tree search, genetic algorithms in order to reduce the computational time required.

#### 1.2.2.4 Transit priority control

Transit priority control is a special case of planning based control where the priority is given to transit vehicles, in the context of multi modal traffic. As giving the priority to transit vehicles can affect the traffic of individual vehicles, the question of avoiding the decrease in the utility cost of individual vehicles was addressed. For example, in 2000, Balke et al. [9] have addressed the transit priority control problem, assuming that the buses can send their GPS location to a controller every second. The priority is given to the bus by extending the green time, add an extra green phase or return to the green phase earlier. The authors have performed simulation studies with hardware-in-the-loop to evaluate their concept. It was proven that their method decreases significantly the bus travel time with only minor increases in the individual vehicles travel times.

In 2009, Ekeila et al. [42] have addressed the problem assuming that there are “a check-in detector located 50 m from the approach stop line and a check-out detector placed at the stop line.” These sensors combined with an arrival time prediction model and a strategy to avoid stops of transit vehicles, have enabled the study of a control which can adapt itself in real time to road traffic and transit conditions. The authors have evaluated their work with VISSIM simulator and have shown that their algorithm outperforms conventional transit signal priority controls for reducing transit trip travel time.

In 2007, Liao and Davis [97] consider vehicles communicating with dedicated short-range communication (DSRC) 802.11p protocol to develop an adaptive signal priority strategy. The buses communicate with the signal controller (V2I communication) in order to request the priority. After having estimated the arrivals of the buses, a strategy to control the traffic light signal has been proposed in order to extend the green time or truncate the red time. The authors have performed a simulation study on an avenue of the city of Chicago with AIMSUN microscopic road traffic simulator which has confirmed the relevance of the algorithm.

It is noticeable that the benefit gained in an upstream junction can be wasted in subsequent junctions in the case where there is no coordination

between the traffic signals. Hence, in 2015, Hu et al. [76] have considered the transit priority control over a corridor in order to coordinate successive signalized intersections. They have formulated the problem as a binary mixed integer linear program solved by branch and bound method. The objective was to minimize per person delay. The transit priority control is granted only if the bus is behind schedule and if the control causes no extra delay. Their research has been successfully evaluated with analytical and microscopic simulation studies on four scenarios.

In 2016, Wu et al. [151] do not take into account corridors, but they have included in their approach the bus holding time at stations and bus speed in addition to signal timings control. They aim at minimizing the total delay, including bus delay and delay of other vehicles nearby the junction. They have assumed that only the buses are equipped with communication capabilities. Experimental results confirm the minimal consequences on general traffic and that the proposed method is significant compared with no priority and signal priority strategies.

In 2019, Yang et al. [157] have used transit priority control to minimize total delay of all passengers in a connected environment. The algorithm is designed to coordinate with the bus stops and the bus schedule. The algorithm activates the priority control only if it is optimal from a system point of view. The authors have performed a sensitivity analysis and they have shown that their algorithm is robust relatively to bus occupancy, location of bus stops and dwell time information.

In 2015, Zeng et al. [162] include the vehicles speed and location information sent by wireless communication to compute vehicle arrivals. By adding the on board passenger information, their model computes the delay for every vehicle passing the junction. Hence, the basis for the optimization problem was proven to be more accurate.

In order to improve the effectiveness of such transit priority control strategies, one direction of research concerns the conflicting priority requests scheduling. In 2012, Zlatkovic et al. [165] have shown that the first come first served policy is not the best solution and can even perform worse than a control with no priority. The authors have proposed an algorithm which helps resolve the problem of conflicting priority requests. They have tested their proposal with VISSIM micro simulator which gave promising results without negative impact on vehicular traffic. Another problem concerns the concurrent objectives between priority transit control and actuated signal controls. In 2014, He et al. [73] have addressed this problem in the context of V2I communication. They have formulated the problem as a mixed integer linear program with an objective function which favours signal priority coordination. Their method also enables the management of conflicting priority requests. Simulation experiments have shown that the proposed control model can reduce the delay, especially in the case of congested traffic conditions. Another way to improve the performances of transit priority control



is by changing the speed and dwell times of buses, which was proposed in 2014 by Serebinski and Khadraoui [127]. This idea introduces us to signal vehicle coupled control which considers the vehicles and traffic signals in an interdependent way.

### 1.2.2.5 Signal vehicle coupled control (SVCC)

In the actuated and planning based traffic signal control, strategies set the traffic light parameters in real time in response to road traffic conditions. However, the traffic light control and the control of connected and automated vehicles by adjusting their speeds and trajectories, could also be addressed in a coordinated way. This possibility is enhanced by the automation and communication technologies embedded in the vehicles. For example, in 2014, Guler et al. [79] have proposed an algorithm assuming that a percentage of vehicles composing the traffic are connected and autonomous vehicles. The algorithm optimizes an objective function by enumerating the sequences of cars crossing the junction. They have shown that with autonomous vehicles, the traffic light signal could change quicker than when relying on the driver reaction. This would improve the performances and decrease a disutility function such as the delay. In 2017, Sun et al. [132] maximize the capacity of the road junction by grouping automated vehicles into platoons and assigning vehicles to their optimal lanes. Despite their method is limited to theoretical analysis because of the absence of a sufficient amount of automated vehicles, their method does not need 100% of connected and automated vehicles. In 2017, Xu et al. [155] have proposed a cooperative method between traffic light and vehicles. The coordination of signal timings with vehicle engine power and brake management enables the improvement of the transportation efficiency on one side, as well as fuel consumption on the other side. In 2016, Yang et al. [156] extend the work by Guler et al. [79] by integrating trajectory design for the vehicles and by reducing the delay and stops. The authors have also replaced the initial enumeration method by a branch and bound method in order to search the optimal departure sequence and the optimal speed of vehicles. In addition, the authors have used a Kalman filter to reduce the impact of measurement errors. In 2018, Yu et al. [160] have formulated the problem of dependently setting the traffic light signal and the speeds of vehicles as a mixed integer linear problem. The traffic light timings, the trajectories, the lane changing maneuvers and the grouping of vehicles into platoons were considered. Their method demonstrates advantages compared to actuated and planning based traffic light controls with improvements in the junction capacity, vehicles delay and fuel consumption. We can see with this short review that although signal vehicle coupled control is promising, it is a new topic with open perspectives.

## Part II

# Estimation of urban road traffic



This part is built upon two journal papers [111] [112] in the IEEE Transactions on Intelligent Transportation Systems journal. Cyril Nguyen Van Phu is the main author of these two papers which address road traffic state estimation at a controlled junction using probe vehicles. This part builds upon a method published by Comert and Cetin in 2009 [28] to consider the case where the roads are composed of multiple lanes.



## Chapter 2

# Estimation of urban traffic state with probe vehicles

# Estimation of urban traffic state with probe vehicles

## 2.1 Abstract

This paper presents a method to estimate urban traffic state with communicating vehicles. Vehicles moving on the links of the urban road network form queues at the traffic lights. It is assumed that a proportion of vehicles are equipped with localization and communication capabilities, namely probe vehicles. First, a method for the estimation of the penetration ratio of probe vehicles, as well as the vehicles arrival rate on a link, is proposed. Moreover, it is shown that turn ratios at each junction can be estimated. Second, assuming that the turn ratios at each junction are given, an estimation of the queue lengths on a 2-lanes link, by extending a 1-lane existing method, is given. The extension introduces vehicles assignment onto the lanes. Third, based on this approach, control laws for the traffic light and for the assignment of the arriving vehicles onto the lane queues, are proposed. Finally, numerical simulations are conducted with Veins framework that bi-directionally couples microscopic road traffic and communication simulators. The propositions are illustrated and discussed with the simulation results.

## 2.2 Introduction

### 2.2.1 State of the art

Different techniques are traditionally used to measure road traffic parameters; for example we can cite inductive loops or video cameras. There is nowadays an infrastructure-less technique to estimate traffic flow parameters such as queue lengths : GPS localization system coupled with communicating vehicles, namely probe vehicles. This kind of equipment penetration ratio is increasing and does not need heavy set up.

Probe vehicles were historically studied for measuring travel times [99]. They also helped to estimate penetration ratio and arrival rate of vehicles (equipped and non equipped) on a link. For example, the author of [27] derived these estimations from the estimation of queue lengths at junctions, queue lengths being estimated using the information provided by the probe vehicles. Thus, we can see that in order to characterize urban road traffic state and its primary parameters such as arrival rates or penetration ratio, estimating queue lengths at junctions is an important step. Furthermore, Varaiya [140] has modeled a road network as “a controlled store-and-forward (SF) queuing network” and proposed an algorithm to control this network of queues. Indeed, minimizing delays and waiting times can be done by minimizing queue lengths at junctions controlled with traffic light signals.

Hence, queue length estimation is a major measurement input data, used to control traffic light signals, and so transportation road networks.

Concerning queue lengths, in 1963, Miller [109] found an approximation of the average queue length at junctions. More recently, the authors of [117] used shockwave theory to refine queue length estimation. Some works also proposed to use probability distribution of the queues [75]. Other works used Markov chains to model the dynamics of queue lengths [144]. The authors of [28] and [128] have addressed the queue length estimation with probe vehicles by proposing a probabilistic analytical model. In [28], the authors have estimated queue length in under-saturated traffic conditions, with the “a priori knowledge of the marginal distribution of the queue length” and using “the location information of the last probe vehicle in the queue”. The authors of [128] have proposed a method to estimate the queue length, the incoming arrival rate, and the output flow, on a  $m$ -lanes link ( $m \geq 2$ ). The estimations are given for low or saturated demand with no requirement of information concerning the timings of the traffic light signal. In [128], all the lanes are assumed to be balanced (i.e. cars share the lanes of the link without any preference). Therefore, all the lanes would have the same length. In [164], the authors estimate arrival rate for low penetration ratio of equipped vehicles. The method proposed in [164] uses as input data “vehicle trajectories approaching to an intersection as well as traffic signal status”. The trajectories of equipped vehicles are used to detect if a probe vehicle has stopped at the traffic light and its stopping position. With these information, the arrival rate is estimated and bounds for this arrival rate are given. In [135], the authors proposed another method. They have lower-bounded the queue length by “the location of the last stopped connected vehicle” and upper-bounded it, when the bound exists, by the location of the “closest moving connected vehicle”. Once bounded, the queue length is estimated using the least-mean-square-error method and the noise is filtered using discrete wavelet transform. In 2015, the authors of [92] have addressed the two lanes case by combining discriminant models “based on time occupancy rates and impulse memories” from detectors. The proportions of total traffic volume in each lane are estimated with Kalman filter. In 2018, the authors of [123] have also addressed the two lanes case. They have measured “individual probe vehicles’ shockwave speed”. Then the lane each probe is moving on, is determined by discriminating the two lanes with data clustering methods. They have shown that a bivariate mixture model clustering gives the best results. Shockwave theory and LWR (Lighthill, Whitham and Richards) model [98, 120] refine the queue length estimation.

### 2.2.2 Paper contribution and organization

We present here an extension of an existing method that uses probe vehicles for the estimation of urban traffic state, including penetration ratio of



communicating vehicles, vehicular arrival rates, as well as the queue lengths of an urban link. The extension considers the general case where different destinations can be associated to the lanes, which produces different arrival rates to each lane of the urban link. We propose here to estimate the joint probability distribution of all the queue lengths of the urban link, instead of estimating only one queue length for the link, as done in [128, 164]. This distinction of the lane queues improves the estimation of the number of cars on the queues. Moreover, it gives the possibility to control the flows of each queue separately, and then ameliorates the traffic control on the junction. In addition control laws for balancing the queue lengths in a multi-lane link are proposed. The method considers here a link of two lanes. The ideas presented in this particular case could be adapted in order to address the general case ( $m$ -lanes, with  $m > 2$ ). Furthermore, the method proposed here could be used and extended in a decentralized manner to the network case because of the low computational effort needed for the one link case. The estimators and the control laws proposed here would permit to perform multi-level urban traffic control, as initiated in [115] (local control) and in [53] (semi-decentralized control).

In section 2.2 an introduction with the related works is given. In section 2.3 the problem statement and the notations are described. In section 2.4 estimators of traffic state parameters are proposed : penetration ratio of probe vehicles, vehicles arrival rate (subsection 2.4.1), and queue lengths in the case of two incoming lanes (subsection 2.4.2). The method estimates queue lengths at junctions with two lanes incoming roads, under the hypothesis of under-saturated traffic (moderate/low demand without overflow queue). It is also assumed that the GPS localization system is not able to determine which lane a vehicle is moving on because of a typically five meters accuracy [138]. The analytic model proposed in [28] is extended to the two-lanes case by introducing a vehicle assignment model onto the lanes. In subsection 2.4.3, a control of the traffic light and an optimal assignment of the vehicles onto the lanes is given, in order to balance the two lanes queue lengths. In section 2.5, numerical simulations are conducted with Veins framework [131] which bi-directionally couples microscopic road traffic and communication simulators. Finally, section 2.6 concludes the article.

## 2.3 Problem statement

In this section we describe the main assumptions of our work and the notations used.

### 2.3.1 Assumptions

#### 2.3.1.1 Road network topology

We consider a road network composed of junctions controlled by traffic light signals, and links between junctions. It is assumed that all the incoming and outgoing links to/from a signalized junction have maximum two lanes. It is assumed that the geometry of the road network is known. A typical junction is represented on Fig. 2.1. It is also assumed that the timings of

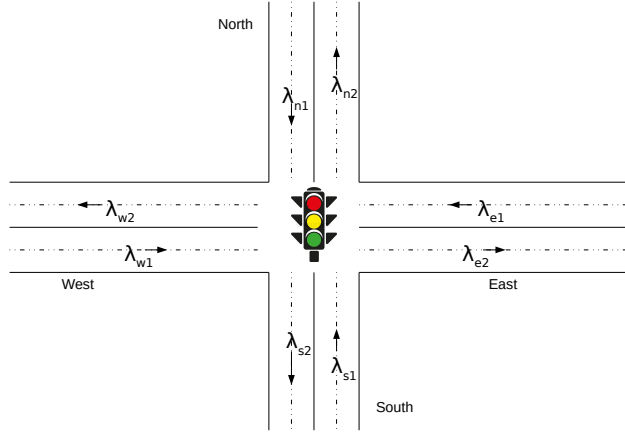


Figure 2.1: A signalized junction of the road network.

the traffic light signal are known, and specially the red times.

#### 2.3.1.2 Traffic demand

It is assumed that the travel demand is exogenous, which means that the demand is located only at the entry links of the network. It is assumed that the vehicles arrive onto each link  $l$  under a Poisson process of rate  $\lambda_l$ . We consider in this paper the low/moderate demand case where the Poisson arrival assumption is valid. In [27] the author discusses the Poisson arrivals assumption and recalls that this assumption is commonly used to describe arrivals at isolated intersections, specially in the case of low/moderate demand with no overflow queue. The vehicles form queues at junctions. Since it is assumed Poisson arrivals, we consider that the queues are empty at the beginning of each red time (no overflow queue).

#### 2.3.1.3 Probe vehicles

It is assumed that a ratio  $p$  (with  $0 \leq p \leq 1$ ) of vehicles are equipped with localization and communication systems and we name them probe vehicles.

The probe vehicles send their positions and speeds to a road side unit (RSU) coupled with the traffic light signal of the junction. It is assumed that the transmit power of the communication system embedded in every vehicle is strong enough, and that the sensitivity of the RSU is accurate enough, such that the RSU can detect every vehicle in every incoming or outgoing link of its associated junction. We consider the case where the localization system embedded in the vehicles is not accurate enough to discriminate the lane the vehicle is moving on.

### 2.3.1.4 Turn ratios

Fig. 2.2 represents the queues we consider, on a link of the road network. Probe vehicles are represented by full rectangles and unequipped vehicles

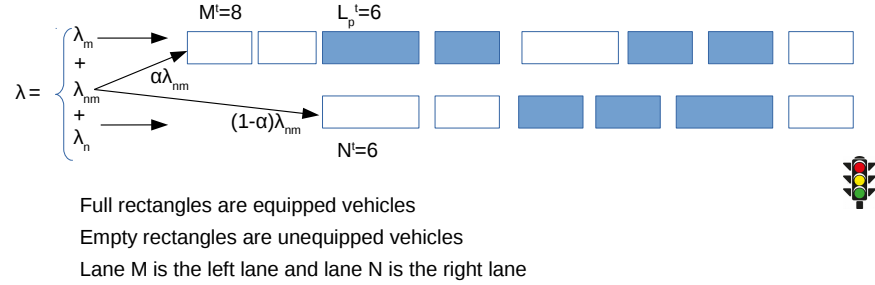


Figure 2.2: Queues in 2-lanes incoming link. Vehicles that can choose both lanes are assigned onto lane  $M$  with probability  $\alpha$  and onto lane  $N$  with probability  $(1 - \alpha)$ .

are represented by empty rectangles. Some vehicles are necessarily assigned to the queue on lane  $N$  (they turn right) and other vehicles are necessarily assigned to the queue on lane  $M$  (they turn left). Vehicles going straight can choose both lanes. It is assumed that a ratio  $\alpha$  (with  $0 \leq \alpha \leq 1$ ) of such vehicles going straight will choose the queue on lane  $M$ . So, on a 2-lanes link, it is assumed that the main flow  $\lambda$  is composed of three flows :

1. the flow with arrival rate  $\lambda_n$  which is necessarily assigned to lane  $N$  (vehicles turning right).
2. the flow with arrival rate  $\lambda_m$  which is necessarily assigned to lane  $M$  (vehicles turning left).
3. the flow with arrival rate  $\lambda_{nm}$  which can be assigned to both lanes  $N$  or  $M$  (vehicles going straight).

We consider that these three flows are independent and identically distributed (iid) stochastic arrivals, each one being a Poisson process. It is supposed that a fraction  $\alpha$  (with  $0 \leq \alpha \leq 1$ ) of the flow  $\lambda_{nm}$  is assigned to

lane  $M$  and the complement  $(1 - \alpha)$  of this same flow  $\lambda_{nm}$  is assigned to lane  $N$ . As denoted in TABLE 2.1,  $A^t$  is the random variable representing the assignment onto the two lanes. It is assumed that  $A^t$  is following a Bernoulli law such that  $P(A^t = 1) = \alpha$  and  $P(A^t = 0) = 1 - \alpha$ . Thus,  $\mathbb{E}(A^t) = \alpha$ . We define :

$$\mu_N(t) := r_N(t)(\lambda_n + (1 - \alpha)\lambda_{nm}) \quad (2.1)$$

$$\mu_M(t) := r_M(t)(\lambda_m + \alpha\lambda_{nm}) \quad (2.2)$$

It will be shown later in Proposition 1 that  $\mu_N(t)$  and  $\mu_M(t)$  represent the average arrival rate multiplied by the red duration on respectively lane  $N$  and lane  $M$ .

Also, it is assumed that the turn ratios are given. Indeed, it is easy to measure the turn ratios as following : the RSU detects all the probe vehicles in its radio range area. This is because the probe vehicles embed WAVE (Wireless Access in Vehicular Environments) [2] on OBU (on board unit). In [2], the basic safety messages (BSM) broadcast periodically the location and speed of probe vehicles. This is a default feature which is also implemented as a basic function in VEINS simulator [131] which was used in this paper. So, if we know at time  $t$  the location of each probe vehicle on a given link and its unique identifier, it is enough to look at a time  $t + t_x$  ( $t_x$  being a time shift), where those vehicles are located. With this method, it is possible to estimate the turn ratios  $l_n$ ,  $l_m$  and  $l_{nm}$  which are the proportions of the main flow  $\lambda$  on the incoming link that respectively turn right, left or go straight. We note :  $\lambda_n = l_n\lambda$ ,  $\lambda_m = l_m\lambda$ ,  $\lambda_{nm} = l_{nm}\lambda$ , with  $l_n + l_m + l_{nm} = 1$ . It is assumed that  $l_n$ ,  $l_m$ , and  $l_{nm}$  are given.

### 2.3.2 Notations

We will use the notations of TABLE 2.1.

## 2.4 Traffic state estimation

### 2.4.1 Primary parameters estimation

In this section, we give a method for the estimation of the primary traffic parameters  $p$  and  $\lambda$ . It is assumed that every probe vehicle in the RSU radio range area is assigned to an incoming or outgoing link to/from the junction. Thus, the total number of incoming probe vehicles  $x_p(t)$  in a given link to the junction is known.

We consider vehicles  $i$  moving at speed  $v_i(t)$  and at a distance  $\rho_i(t)$  (depending on time  $t$ ) from the traffic light. Let us consider the following definition.

Name	Definition
$L_V$	the average vehicle length
$G_V$	the minimum distance gap between vehicles
$R$	the total red time in one cycle
$r_N(t)$	the time since the beginning of the red phase for lane N (it is 0 if we are not in red phase at time $t$ ), $0 \leq r_N(t) \leq R$ .
$r_M(t)$	the time since the beginning of the red phase for lane M (it is 0 if we are not in red phase at time $t$ ), $0 \leq r_M(t) \leq R$ .
$\lambda_n, \lambda_m, \lambda_{nm}$	the average arrival rate in vehicles/second for vehicles that respectively turn right, left or go straight.
$\lambda = \lambda_n + \lambda_m + \lambda_{nm}$	the total arrival rate for the incoming link in vehicles/second.
$l_n, l_m, l_{nm}$	the proportions (turning ratios) of the main flow $\lambda$ on the incoming link that respectively turn right, left or go straight, with $\lambda_n = l_n \lambda$ , $\lambda_m = l_m \lambda$ , $\lambda_{nm} = l_{nm} \lambda$
$x(t)$	the total number of vehicles on all the lanes of the considered link at time $t$ .
$x_p(t)$	the number of probe vehicles on all the lanes of the considered link at time $t$ .
$p, 0 \leq p \leq 1$	the penetration ratio of probe vehicles.
$N^t$	the total number of vehicles in the queue at time $t$ and lane N. In this paper, $N^t$ is assumed to be a random variable.
$M^t$	the total number of vehicles in the queue at time $t$ and lane M. In this paper, $M^t$ is assumed to be a random variable.
$A^t$	the assignment of a vehicle entering the edge at time $t$ . $A^t = 1$ if the vehicle is assigned on lane M and $A^t = 0$ if it is assigned on lane N. $A^t$ is assumed to be a random variable.
$L_p^t$	the location (in number of vehicles) of the last probe in the queue, namely the last connected vehicle, at time $t$ . $L_p^t$ is assumed to be a random variable, taking value $l_p$ .
$N_p^t$	the total number of probe vehicles in the queue at time $t$ and lane N.
$M_p^t$	the total number of probe vehicles in the queue at time $t$ and lane M.
$c_p$	the total number of probe vehicles in all the lanes and all the queues at time $t$ .

Table 2.1: Notations

**Definition 1.** For a given threshold car-speed  $v^*$  and a given threshold car-distance  $\rho^*$  to the junction, the vehicles queue  $Q = Q(t, v^*, \rho^*)$  is defined by  $Q = \{i, v_i < v^* \text{ and } \rho_i < \rho^*\}$ .

$\rho^*$  is useful because if the queue would exceed the bound  $\rho^*$ , we could know that the assumption of low/moderate demand is not adequate. Furthermore,  $\rho^*$  is less than the edge length, so the queue keeps bounded. We then denote by  $Q_p$  the subset of  $Q$  that includes only probe vehicles,  $Q_p \subset Q$ . The total number of probes  $c_p$  in the queue is given by the cardinal (number of elements) of the set  $Q_p$ . It is assumed mixed vehicles (equipped and non equipped) with an average vehicle length  $L_V$  and minimum distance gap  $G_V$  between vehicles.  $\rho_0$  denotes the offset distance from the RSU to the stop line of the traffic light signal.  $l_p$  the last probe location in the unit of "number of vehicles" is computed as follows. We have :

$$\max_{i \in Q_p}(\rho_i) = \rho_0 + l_p L_V + (l_p - 1)G_V \quad (2.3)$$

Then,

$$l_p = [(\max_{i \in Q_p}(\rho_i) - \rho_0 + G_V)/(L_V + G_V)] \quad (2.4)$$

where  $[\cdot]$  denotes the round operator to the nearest integer.

Given  $c_p, l_p$ , Comert [27] has derived many estimators for  $p$ , one of them being  $c_p/l_p$ , which is biased for  $p < 1$ . We follow here the same idea and propose a variation of the estimator of  $p$ . For the one lane case, we propose:

$$\hat{p} = (c_p - 1)/(l_p - 1), \text{ for } l_p > 1 \quad (2.5)$$

We have :  $N^t = l_p + 1/p - 1$ , where  $1/p - 1$  represents the average backlog of the queue behind the last probe. Then  $\hat{p} = c_p/N^t = c_p/(l_p + 1/p - 1)$ . Moreover, by following the same arguments of the proof in [27], it is easy to check that this estimator is unbiased for every  $p$ , i.e.  $\mathbb{E}(\hat{p}) = p, \forall p, 0 \leq p \leq 1$ .

For two lanes, we introduce :

$\kappa := \min(\mu_n, \mu_m)/\max(\mu_n, \mu_m)$ .  $\kappa$  depends only on turn ratios but not on the arrival rate for the link, because  $\kappa$  is a ratio.

$$\kappa = \frac{\min(r_N(t)(l_n + (1 - \alpha)l_{nm}), r_M(t)(l_m + \alpha l_{nm}))}{\max(r_N(t)(l_n + (1 - \alpha)l_{nm}), r_M(t)(l_m + \alpha l_{nm}))} \quad (2.6)$$

We consider queue lengths on lanes  $N$  and  $M$  respectively equal to  $n$  and  $m$ . We propose :

$$\hat{p} = \frac{c_p}{n + m} \quad (2.7)$$

By the way, in our case, the length  $n$  of queue  $N$  can be estimated with the number of arrivals on lane  $N$  during  $r_N(t)$  which is  $\mu_n$ . As  $\mu_n + \mu_m = \max(\mu_n, \mu_m) + \min(\mu_n, \mu_m)$  and by estimating  $\max(\mu_n, \mu_m) = l_p + 1/p - 1$ ,

where  $1/p - 1$  represents the backlog of the queue behind the last probe, we can write :

$$\mu_n + \mu_m = \max(\mu_n, \mu_m) \left( 1 + \frac{\min(\mu_n, \mu_m)}{\max(\mu_n, \mu_m)} \right) \quad (2.8)$$

$$\mu_n + \mu_m = (l_p + 1/p - 1)(1 + \kappa) \quad (2.9)$$

We introduce  $c_\kappa = c_p/(1 + \kappa)$  and replace (2.9) in (2.7). Finally, we get the following equation :

$$\hat{p} = \frac{c_\kappa}{l_p + (1 - p)/p} \quad (2.10)$$

Hence, solving in  $\hat{p}$ , by putting  $p = \hat{p}$  :

$$\hat{p} = (c_\kappa - 1)/(l_p - 1), \text{ for } l_p > 1 \text{ and } c_p > 1 \quad (2.11)$$

which extends (4.7) for the case of two lanes. Similarly, we can check that this estimator for two lanes is unbiased, by following again the same arguments as in [27]. In Appendix 2.7, another method to compute  $\hat{p}$  based on the calculus of an unbiased estimator gives the same result than formula (2.11).

The vehicles arrival rate  $\lambda$  is computed with formula (2.12) by simply accumulating probe vehicles on the entire radio range area of the RSU during red time, and using  $\hat{x} = x_p/p$ .  $\lambda$  should be computed when all the lanes of the considered link have a red light at the traffic light. This is formulated as following, where  $t_0$  is the starting time for the red light on both lanes and  $t_1$  is the time after which the set of the two lanes are not at red light.

$$\hat{\lambda} = \frac{x_p(t_1) - x_p(t_0)}{p(t_1 - t_0)} \quad (2.12)$$

### 2.4.2 Queue length estimation

Once  $p$  and  $\lambda$  are estimated with probe vehicles, the traffic state estimation (queue lengths) can be refined. In this section, the queue lengths associated to all the lanes on a link of the road network are estimated. A model that uses vehicular assignment onto the lanes, for links composed of two incoming lanes, is proposed. In a first step, an analytical probability distribution formulation of the queue lengths, without using the information from the probe vehicles, will be presented. Then, the information provided by the probe vehicles is used : while generalizing the work for 1-lane road done in [28] to the 2-lanes case, the analytical formulation is refined. We recall here that we can not directly detect the lanes on which the probe vehicles are moving because of insufficient accuracy of GPS localization system [138], which makes the problem not obvious.

### 2.4.2.1 Distribution probability law of the 2-lanes without having the information provided by the probes

First, we propose an estimation of the probability distribution  $P(N^t = n, M^t = M)$  without having any information from the probe vehicles.

#### Proposition 1.

$$P(N^t = n, M^t = m) = \frac{\mu_N(t)^n e^{-\mu_N(t)}}{n!} \frac{\mu_M(t)^m e^{-\mu_M(t)}}{m!} \quad (2.13)$$

*Proof.* The Poisson process of rate  $\lambda_{nm}$  common to the two lanes is subdivided. The common arrival of rate  $\lambda_{nm}$  is splitted with probability  $\alpha$  to lane  $M$  and probability  $(1 - \alpha)$  to lane  $N$ . The two produced flows are independent random flows each one following Poisson process of parameters respectively  $\alpha\lambda_{nm}$  for the flow assigned to lane  $M$  and  $(1 - \alpha)\lambda_{nm}$  for the flow assigned to lane  $N$ . Furthermore, the splitted Poisson processes are independent; see subdividing Poisson process in reference [57].

By combination, arrivals on lane  $N$  is the sum of two independent Poisson processes. Using the stationary property of Poisson processes, we can show that the number of arrivals in  $[0, r_N(t)]$  on lane  $N$  is a Poisson process of parameter  $\mu_N$ . Similarly, the number of arrivals on lane  $M$  in  $[0, r_M(t)]$  is a Poisson process of parameter  $\mu_M$ . As these two arrival flows on lanes  $N$  and  $M$  are independent, then the bivariate distribution probability law of the two queue lengths is the product of two Poisson Law of parameters  $\mu_N$  and  $\mu_M$ .  $\square$

### 2.4.2.2 Distribution probability law of the 2-lanes queue lengths with the information provided by the probe vehicles

We present here the conditional probability law of the two queue lengths, taking into consideration the information provided by the probe vehicles, specially the location of the last probe  $l_p$  and the total number of probes in the two lanes queues  $c_p$ .  $N_p^t$  and  $M_p^t$  are the number of probe vehicles respectively on the lane  $N$  and on the lane  $M$ , at time  $t$ .

#### Proposition 2.

- If  $l_p \leq \max(n, m)$  and  $c_p \leq n + m$ , then

$$P(N^t = n, M^t = m | L_p^t = l_p, N_p^t + M_p^t = c_p) = \frac{\binom{l_p - 1 + \min(l_p, n, m)}{c_p - 1} (1 - p)^{n+m} P(N^t = n, M^t = m)}{\sum_{\substack{j, k \geq 0 \\ \text{subject to} \\ \max(j, k) \geq l_p \\ j + k \geq c_p}} \binom{l_p - 1 + \min(l_p, j, k)}{c_p - 1} (1 - p)^{j+k} P(N^t = j, M^t = k)}.$$



- Otherwise,  
 $P(N^t = n, M^t = m | L_p^t = l_p, N_p^t + M_p^t = c_p) = 0.$

*Proof.* By Bayes' rule we have

$$P(N^t = n, M^t = m | L_p^t = l_p, N_p^t + M_p^t = c_p) = \frac{P(N^t = n, M^t = m, L_p^t = l_p, N_p^t + M_p^t = c_p)}{P(L_p^t = l_p, N_p^t + M_p^t = c_p)} \quad (2.14)$$

Then the numerator in (2.14) is written

$$\begin{aligned} P(N^t = n, M^t = m, L_p^t = l_p, N_p^t + M_p^t = c_p) &= \\ P(L_p^t = l_p | N_p^t + M_p^t = c_p, N^t = n, M^t = m) & \\ P(N_p^t + M_p^t = c_p | N^t = n, M^t = m) & \\ P(N^t = n, M^t = m) & \end{aligned} \quad (2.15)$$

We have

- $P(L_p^t = l_p | N_p^t + M_p^t = c_p, N^t = n, M^t = m) =$   
 $\binom{l_p - 1 + \min(l_p, n, m)}{c_p - 1} / \binom{n + m}{c_p}.$
- $P(N_p^t + M_p^t = c_p | N^t = n, M^t = m) =$   
 $\binom{n + m}{c_p} p^{c_p} (1 - p)^{n + m - c_p}.$

For the calculus of  $P(L_p^t = l_p | N_p^t + M_p^t = c_p, N^t = n, M^t = m)$ , we followed the same ideas as those of section 3 in [28]. Indeed, we will use the example of Fig. 2.2 where  $l_p = 6$ ,  $M^t = 8$ ,  $N^t = 6$  and  $c_p = 7$ . The probability is then computed by selecting the total number of events where  $L_p = l_p = 6$  divided by the sample space. The sample space, which is composed of all the last probe possible locations is given by  $\binom{n + m}{c_p} = \binom{14}{7}$ . For  $L_p = l_p = 6$  we must have all the probes in the preceding locations. The event space has a number of events corresponding to choosing  $(c_p - 1) = 6$  probes among  $l_p - 1 + \min(l_p, n, m) = 6 - 1 + 6 = 11$  positions available. Here, the event space has a total number of elements given by  $\binom{l_p - 1 + \min(l_p, n, m)}{c_p - 1} = \binom{11}{6}$ . This is why :

$$P(L_p^t = l_p | N_p^t + M_p^t = c_p, N^t = n, M^t = m) = \binom{l_p - 1 + \min(l_p, n, m)}{c_p - 1} / \binom{n + m}{c_p} \quad (2.16)$$

For the calculus of  $P(N_p^t + M_p^t = c_p | N^t = n, M^t = m)$  we have  $c_p$  probe vehicles among  $n + m$  total vehicles. The probability for a vehicle to be

a probe vehicle is  $p$  and the probability to be unequipped is  $(1 - p)$ . The configurations considered in this case are  $c_p$  vehicles equipped and  $(n + m - c_p)$  vehicles unequipped. The number of combinations of such configurations is  $\binom{n+m}{c_p}$ . This is why :

$$P(N_p^t + M_p^t = c_p | N^t = n, M^t = m) = \binom{n+m}{c_p} p^{c_p} (1-p)^{n+m-c_p} \quad (2.17)$$

So the numerator in (2.14) is given by :

$$P(N^t = n, M^t = m, L_p^t = l_p, N_p^t + M_p^t = c_p) = \binom{l_p - 1 + \min(l_p, n, m)}{c_p - 1} p^{c_p} (1-p)^{n+m-c_p} P(N^t = n, M^t = m)$$

The denominator in (2.14) is the marginal distribution probability of  $P(N^t = j, M^t = k, L_p = l_p, N_p^t + M_p^t = c_p)$  on  $(j, k)$ . Therefore, the ideas to compute this probability are the same as the ideas used to compute the numerator of (2.14). We notice that the last probe position (in the unit number of vehicles) is necessarily less than or equal to the maximum of the queue lengths, since the last probe is necessarily in one of the two lanes queues. Similarly, the total number of probes  $c_p$  is less than or equal to the total number of vehicles in the queues, since the probes are in the queues. Therefore, we can write :

$$P(L_p^t = l_p, N_p^t + M_p^t = c_p) = \sum_{\substack{j, k \geq 0 \\ \max(j, k) \geq l_p \\ j+k \geq c_p}} P(N^t = j, M^t = k, L_p = l_p, N_p^t + M_p^t = c_p).$$

□

### 2.4.2.3 Estimators

The distribution probability law of the couple  $(N^t, M^t)$  is known; see Proposition 2. As,  $\mathbb{E}(N^t, M^t) = (\mathbb{E}(N^t), \mathbb{E}(M^t))$ , one way to estimate the two queue lengths is to derive each queue length separately from the couple, by computing the expectation of  $N^t$  and  $M^t$  separately. We propose the following estimator for queue length on lane  $N$  :

$$\mathbb{E}(N^t | L_p^t = l_p, N_p^t + M_p^t = c_p) = \sum_{n \geq 0} n \sum_{k \geq 0} P(N^t = n, M^t = k | L_p^t = l_p, N_p^t + M_p^t = c_p) \quad (2.18)$$

Similarly for the queue length on lane  $M$ , we have :

$$\begin{aligned} \mathbb{E}(M^t | L_p^t = l_p, N_p^t + M_p^t = c_p) &= \\ \sum_{m \geq 0} m \sum_{j \geq 0} P(N^t = j, M^t = m | L_p^t = l_p, N_p^t + M_p^t = c_p) & \quad (2.19) \end{aligned}$$

### 2.4.3 Traffic light control and optimal assignment of vehicles onto the lanes

We are interested here in the equilibration of the two queue lengths with respect to the two parameters  $\alpha$  and  $\bar{r} := r_N/r_M$ . We exclude here the case  $r_N = r_M = 0$  where both lanes have green light, and where no queue is formed by assumption; and therefore, the assignment onto the queues is meaningless. We use notations  $\mathbb{E}_{(\alpha, r_N)}(N^t) := \mathbb{E}(N^t)$  and  $\mathbb{E}_{(\alpha, r_M)}(M^t) := \mathbb{E}(M^t)$  in order to emphasize the dependence of these two expectations on the parameters  $\alpha, r_N$  and  $r_M$ . Let us now define  $f(\alpha, \bar{r})$  as follows.

$$\begin{aligned} f(\alpha, \bar{r}) &:= |\mathbb{E}_{(\alpha, r_N)}(N^t) - \mathbb{E}_{(\alpha, r_M)}(M^t)| / \mathbb{E}_{(\alpha, r_N)}(N^t) \\ &= |\mu_N - \mu_M| / \mu_N \\ &= \frac{\lambda |r_N(l_n + (1-\alpha)l_{nm}) - r_M(l_m + \alpha l_{nm})|}{\lambda r_N(l_n + (1-\alpha)l_{nm})} \\ &= \frac{|\bar{r}(l_n + (1-\alpha)l_{nm}) - (l_m + \alpha l_{nm})|}{\bar{r}(l_n + (1-\alpha)l_{nm})}. \end{aligned}$$

$f(\alpha, \bar{r})$  is the relative difference between the two queue lengths. It is assumed that the two queue lengths tend to equilibrate. The minimization of  $f(\alpha, \bar{r})$  with respect to the two parameters  $\alpha$  and  $\bar{r}$  represents the equilibration of the two queue lengths. Let us use the notations.

$$r^*(\alpha) := \arg \min_{\bar{r}} f(\alpha, \bar{r}). \quad (2.20)$$

$$\alpha^*(\bar{r}) := \arg \min_{\alpha} f(\alpha, \bar{r}). \quad (2.21)$$

Proposition 3 and Proposition 4 below determines  $r^*(\alpha)$  and  $\alpha^*(\bar{r})$  respectively.

**Proposition 3.**  $\forall \alpha \in [0, 1], r^*(\alpha) = \frac{l_m + \alpha l_{nm}}{l_n + (1-\alpha)l_{nm}}$ , and  $f(\alpha, r^*(\alpha)) = 0$ .

*Proof.*  $\forall \alpha \in [0, 1], \bar{r} = \bar{r}_0 = \frac{l_m + \alpha l_{nm}}{l_n + (1-\alpha)l_{nm}}$  implies

$$\begin{aligned} \mathbb{E}_{(\alpha, r_N)}(N^t) &= r_N(\lambda_n + (1-\alpha)\lambda_{nm}) \\ &= (r_M \bar{r})(\lambda_n + (1-\alpha)\lambda_{nm}) \\ &= r_M(\bar{r}_0(\lambda_n + (1-\alpha)\lambda_{nm})) \\ &= r_M(\lambda_m + \alpha\lambda_{nm}) = \mathbb{E}_{(\alpha, r_M)}(M^t). \end{aligned}$$

Therefore,  $f(\alpha, \bar{r}_0) = 0$ . Thus,  $r^*(\alpha) = \bar{r}_0 = \frac{l_m + \alpha l_{nm}}{l_n + (1-\alpha)l_{nm}}$ .  $\square$

**Proposition 4.**  $\forall \bar{r} \geq 0$ ,

$$\alpha^*(\bar{r}) = \max \left( 0, \min \left( 1, \frac{\bar{r}l_n + \bar{r}l_{nm} - l_m}{l_{nm}(\bar{r} + 1)} \right) \right).$$

Moreover, if  $\bar{r} \in I := [\frac{l_m}{l_n+l_{nm}}, \frac{l_m+l_{nm}}{l_n}]$ , then

$$\alpha^*(\bar{r}) = \frac{\bar{r}l_n + \bar{r}l_{nm} - l_m}{l_{nm}(\bar{r} + 1)}, \text{ and } f(\alpha^*(\bar{r}), \bar{r}) = 0.$$

*Proof.* For any  $\bar{r} \geq 0$ ,  $\alpha^*(\bar{r})$  is simply the argument of the minimization of  $f(\alpha, \bar{r})$  with respect to  $\alpha$ , projected into the interval  $[0, 1]$ . In the case where  $\bar{r} \in I := [\frac{l_m}{l_n+l_{nm}}, \frac{l_m+l_{nm}}{l_n}]$ , we can easily check that the constraint  $\alpha^*(\bar{r}) \in [0, 1]$  is not activated, and then we do not need to project into the interval  $[0, 1]$ . Moreover, in this case,  $\alpha^*(\bar{r})$  cancels  $f(\alpha, \bar{r})$ .  $\square$

We notice here that the calculus of the optimal assignment proportion  $\alpha^*(\bar{r})$  of the vehicles going straight onto the lanes is done in deterministic at the macroscopic level (proportion of vehicular flow). The realization of the optimal assignment proportion  $\alpha^*(\bar{r})$  is done randomly at the microscopic level : every vehicle going straight is randomly assigned to the the lanes M and N with probabilities  $\alpha$  and  $(1 - \alpha)$  respectively, assuming that the vehicles going straight that can choose both lanes will choose the shortest queue. By equilibrating the two queues with  $r^*(\alpha)$  or  $\alpha^*(\bar{r})$ , spill-back onto the links of the network can be avoided, and by that the risk of congestion is reduced. For example, the optimal  $r^*(\alpha)$  given in Proposition 3 can be taken into account as an additional constraint in the optimization problem of the traffic light split on every intersection, in such a way that the optimal traffic light setting will automatically balance the queue lengths on the incoming links of the intersections, which should help to avoid spill-back at the network level.

## 2.5 Simulation results, examples and discussion

We present in this section the results of numerical simulations conducted with Veins framework [131] which combines the microscopic road traffic simulator SUMO [86] and the communication simulator OMNET++ [142]. The road network is one simple junction with links composed of two incoming lanes described in Fig. 2.1. The junction is controlled by a traffic light with a cycle duration of 90 s. The traffic demand is coming from West towards East, North and South. We vary the arrival rates and turn ratios depending on the scenarios, as we mentioned in TABLE 2.2. The messages used to detect the location of the vehicles are the Basic Safety Messages (BSM) that are sent in broadcast by every probe vehicle. Given the road network topology and the data provided by the probe vehicles, it is known on which incoming/outgoing link each probe vehicle is located.

### 2.5.1 Primary parameters estimation

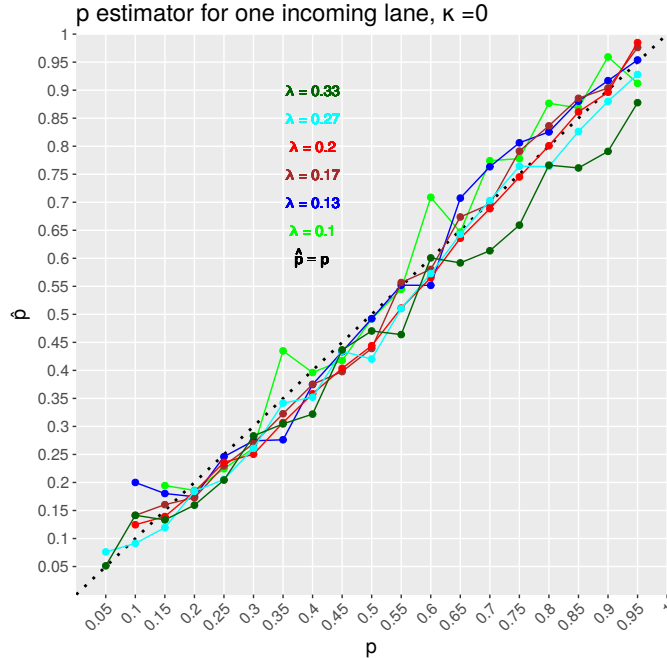


Figure 2.3: Estimated penetration ratio  $p$  for a one lane incoming link, depending on penetration ratio for various demand scenarios. The arrival demand levels ( $\lambda$ ) are given in (vehicles/s). Simulated time = 40 min.

In this part, we illustrate estimation of primary parameters proposed in section 2.4.1. Simulations are conducted for various arrival rates, each one being represented by one color as detailed on Fig. 2.3 2.4 2.5. Based on the penetration ratio given as an input of the simulation, SUMO simulator produces random samples of communicating vehicles, which form the input data for the various estimators. Fig. 2.3 represents the estimated penetration ratio  $\hat{p}$ , given by formula (4.7), associated in this figure to the real penetration ratio  $p$ , in the case of an incoming link of one lane. Fig. 2.4 represents the estimated penetration ratio  $\hat{p}$ , given by formula (2.11), associated in this figure to the real penetration ratio  $p$ , in the case of an incoming link of two lanes. Ideally,  $\hat{p} = p$ , forming a line of slope 1 drawn in discontinuous black on the figure.

Fig. 2.5 represents arrival rate estimated as given by formula (2.12) of section 2.4.1 in the case of an incoming link of two lanes for  $\kappa = 0.5$ . We can see here that the estimation is better when  $p$  gets higher. As there are more data and as the arrival rate is higher, the estimation of arrival rate is more accurate.

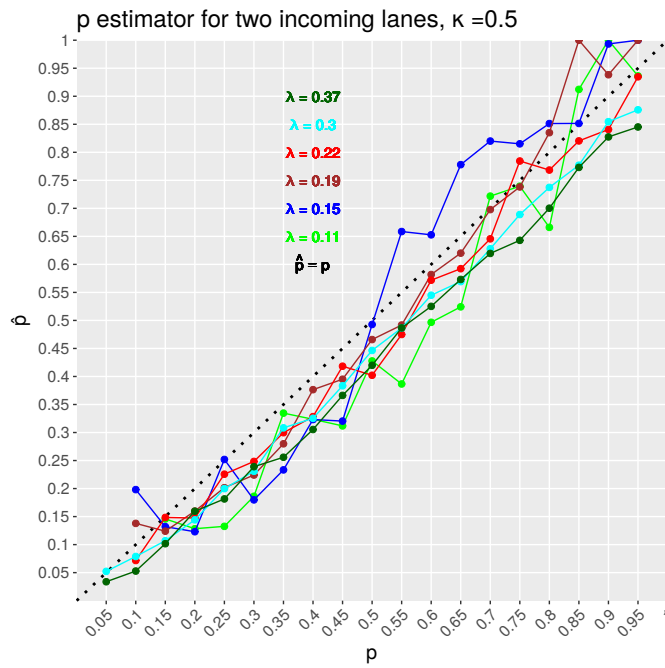


Figure 2.4: Estimated penetration ratio  $p$  for a two lanes incoming link, depending on penetration ratio for various demand scenarios. The arrival demand levels ( $\lambda$ ) are given in (vehicles/s). Simulated time = 40 min.  $\kappa = 0.5$ .

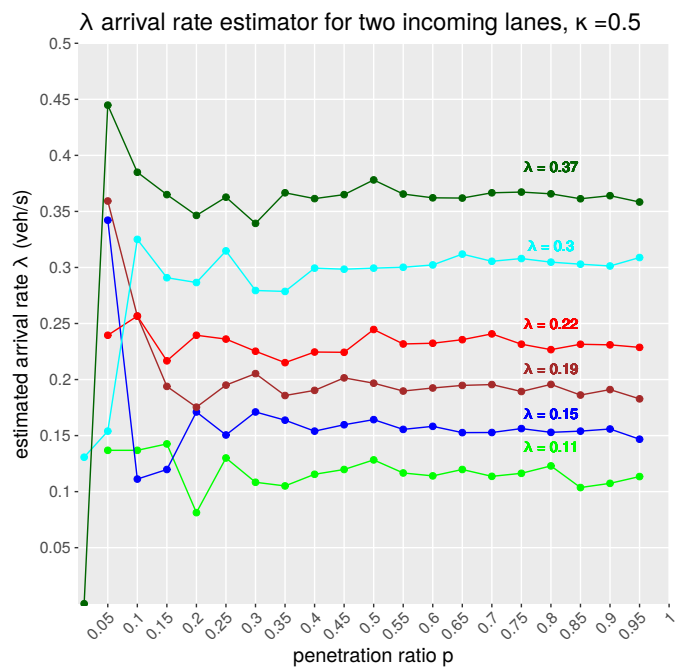


Figure 2.5:  $(\lambda_n + \lambda_m + \lambda_{nm})$  in vehicles/second for a two lanes incoming link, depending on penetration ratio for various demand scenarios. Simulated time = 40 min.  $\kappa = 0.5$ .

**2.5.2 Probability distributions**

In this part, we illustrate the probability distribution queue lengths as proposed in section 2.4.2. It is assumed that the demand is coming from the West of the junction as described in TABLE 2.2 with  $r_N = r_M$  i.e.  $\bar{r} = 1$ .

These different scenarios include different possibilities concerning the demand such as : symmetric (Scenario S3) or asymmetric arrivals (scenarios S1, S2, S4, S5). The arrival demand  $\lambda_n, \lambda_m$  and  $\lambda_{nm}$  is varied and we derive the optimal  $\alpha^*(1)$  by Proposition 4. The values  $\alpha^*(1) = [0.1, 0.25, 0.5, 0.75, 0.9]$  derived with Proposition 4, represent the assumption that the queues tend to equilibrate as much as possible. The cases where  $\bar{r} \neq 1$  are not considered.

Scenario	S1	S2	S3	S4	S5
$\alpha^*$	0.1	0.25	0.5	0.75	0.9
$\lambda_{nm}$	125	100	50	100	125
$\lambda_m$	200	125	200	75	100
$\lambda_n$	100	75	200	125	200
<b>Arrival rates</b>	<b>Amount of vehicles for 1200 s</b>				

Table 2.2: Demand for different scenarios (simulated time=1200 s) and for  $r_N = r_M$

**2.5.2.1 Example: probability distribution of the two queues lanes without and with the information provided by the probe vehicles**

We draw on top of Fig. 2.6 the probability distribution law  $P(N^t, M^t)$  (Proposition 1), and on bottom of Fig. 2.6 the conditional probability distribution law  $P(N^t, M^t | L_p, N_p + M_p)$  (Proposition 2) for scenario of Fig. 2.6. On top of Fig. 2.6, we can see that the total number of vehicles in the queue is estimated to  $(N^t = 6, M^t = 4)$ , for a total of 10 vehicles in the queue. There is an asymmetry in the distribution probability law because of the asymmetric demand and because the common flow is not strong enough to equilibrate the two queue lanes. We can see on bottom of Fig. 2.6 that the conditional distribution clearly discriminates the two queue lengths and keeps track of the asymmetry. In this case, the parameters are  $p = 0.55$ ,  $c_p = 8$ ,  $l_p = 9$ . We compute  $\kappa = 0.75$ . Following the same ideas as above, we have  $\hat{p} = (8/1.75 - 1)/(9 - 1) = 0.45$ . It is probable that there are not many cars behind  $l_p$ , maybe 1 vehicle. Therefore, as  $l_p = 9$ , the biggest lane should contain around 10 vehicles. Given the asymmetry of the distribution law  $P(N^t, M^t)$ , the conditional probability  $P(N^t, M^t | L_p, N_p + M_p)$  will favour the lane with the highest arrival rate (lane  $N$ ). Then the queue on lane  $M$  should contain very few vehicles and will be around the same lane length



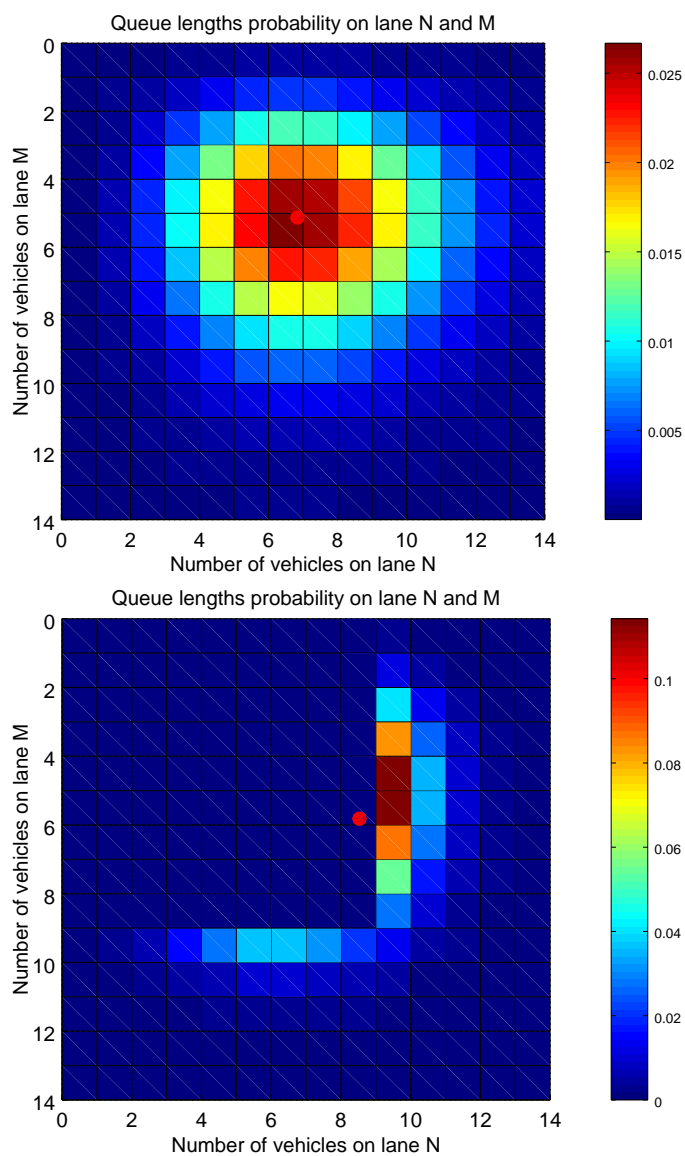


Figure 2.6: On the top : probability law of Proposition 1. On the bottom : probability law of Proposition 2. Scenario with  $\lambda_n = 1/6$ ,  $\lambda_m = 1/12$ ,  $\lambda_{nm} = 1/24$  (veh/s)(thus  $\alpha^*(1) = 1$ ),  $r(t) = 41$  s,  $p = 0.55$ ,  $c_p = 8$ ,  $l_p = 9$ , at time  $t = 760$ s. Red dot is the expectation of the probability distribution. Simulation time=1200s.

estimation as in the top figure. In this example, the conditional distribution probability calculus emphasizes the asymmetry of the two lanes.

### 2.5.2.2 Results for the scenarios of TABLE 2.2

For each scenario we measure the maximum and average queue lengths as estimated by SUMO microscopic road traffic simulator. Different seeds for the random number generator of the simulators have been used, taking the average value for TABLE 2.3 and 2.4. We notice that SUMO queue length is measured in such a way that any vehicle with a speed greater than 0.1 m/s is not considered in the queue.  $p \in [0, 1]$  is varied for each scenario and we compute :

- MAE(P2):=the mean absolute error between the estimated queue lengths as given by the estimator based on Proposition 2 and SUMO queue lengths on a subset of the data.
- MAE(P1):=the mean absolute error between the estimated queue lengths as given by  $\hat{N} = \mu_n$ ,  $\hat{M} = \mu_m$  (Proposition 1) and SUMO queue lengths on a subset of the data.
- MAE( $l_p$ ):=the mean absolute error between the estimated queue lengths as given by  $\max(\hat{N}, \hat{M}) = l_p$  and  $\min(\hat{N}, \hat{M}) = \kappa l_p$ , for the queue lengths N and M, and SUMO queue lengths on a subset of the data.

The results for the scenarios of TABLE 2.2 in TABLE 2.3 and TABLE 2.4 are given here. We comment and emphasize some tendencies on the results. We notice on these tables that the error is decreasing as  $p$  tends to 1 : Proposition 2 is getting more accurate as more data is given in input. Even if the estimator based on  $l_p$  (MAE( $l_p$ )) is getting more accurate as  $p$  increases, the performances of the estimator based on Proposition 2 are generally better. Furthermore, if we compare MAE(P1) and MAE(P2) we notice that the performances of Proposition 2 based estimator are in general more accurate than the ones of the other estimators, especially when  $p$  increases. Finally, MAE( $l_p$ ) gives in general less accurate results than the estimators based on Propositions 1 and 2. <sup>1</sup>.

The main source of difference between our estimations and SUMO queue lengths is the assignment model. Indeed, drivers do not always choose the shortest queue for their assignment in SUMO. In fact, in SUMO, there is a kind of thresholds on the difference between the queue lengths, beyond

---

<sup>1</sup>A significant source of error for this model is the fact that the estimation is done in real numbers, while the measured number of vehicles is done in integer ones. In the general case, the error due to discretization is about 0.5 vehicles. This is very big in the cases where the queue length is small, in particular at the beginning of the red time. If for example, the queue length is 0.5 vehicles in average, then we have 100% error, due only to the discretization.

Results \ Scenarios	S1	S2	S3	S4	S5
Average SUMO Queue Length	3.54	3.21	4.33	3.21	4.77
Max SUMO Queue Length	10.07	10.07	14.12	9.71	15.95
$p = 0.05$ MAE(P2)	1.59	1.11	1.38	1.19	1.84
MAE(P1)	1.40	1.14	1.45	1.20	1.97
MAE( $l_p$ )	3.91	3.71	4.38	3.71	5.61
$p = 0.10$ MAE(P2)	1.51	1.02	1.29	1.12	1.75
MAE(P1)	1.40	1.14	1.45	1.20	1.97
MAE( $l_p$ )	2.01	2.28	1.96	2.13	2.60
$p = 0.15$ MAE(P2)	1.52	1.04	1.28	1.06	1.58
MAE(P1)	1.40	1.14	1.45	1.20	1.97
MAE( $l_p$ )	1.77	1.52	1.64	1.53	1.95
$p = 0.20$ MAE(P2)	1.47	0.96	1.20	0.98	1.43
MAE(P1)	1.40	1.14	1.45	1.20	1.97
MAE( $l_p$ )	1.85	1.28	1.40	1.16	1.44
$p = 0.50$ MAE(P2)	1.47	0.80	0.98	0.79	1.10
MAE(P1)	1.40	1.14	1.45	1.20	1.97
MAE( $l_p$ )	2.33	1.14	1.24	0.99	0.89
$p = 0.70$ MAE(P2)	1.33	0.73	0.93	0.70	1.13
MAE(P1)	1.40	1.14	1.45	1.20	1.97
MAE( $l_p$ )	2.48	1.25	1.30	0.97	0.90
$p = 0.90$ MAE(P2)	1.10	0.74	0.97	0.70	1.21
MAE(P1)	1.40	1.14	1.45	1.20	1.97
MAE( $l_p$ )	2.62	1.39	1.39	1.04	0.95

Table 2.3: Results in the unit "number of vehicles" for queue on lane N, estimated vs SUMO queue length.

Results \ Scenarios	S1	S2	S3	S4	S5
Average SUMO Queue Length	4.79	3.39	4.20	2.89	3.84
Max SUMO Queue Length	15.22	9.71	14.11	8.61	12.28
$p = 0.05$ MAE(P2)	1.98	1.24	1.39	1.26	1.36
MAE(P1)	2.02	1.25	1.35	1.22	1.30
MAE( $l_p$ )	5.71	3.94	4.27	3.23	4.22
$p = 0.10$ MAE(P2)	1.89	1.17	1.30	1.25	1.37
MAE(P1)	2.02	1.25	1.37	1.22	1.30
MAE( $l_p$ )	2.81	2.53	1.93	1.88	1.93
$p = 0.15$ MAE(P2)	1.68	1.09	1.22	1.20	1.29
MAE(P1)	2.02	1.25	1.37	1.22	1.30
MAE( $l_p$ )	1.94	1.54	1.45	1.44	1.68
$p = 0.20$ MAE(P2)	1.55	1.05	1.19	1.12	1.28
MAE(P1)	2.02	1.25	1.37	1.22	1.30
MAE( $l_p$ )	1.54	1.25	1.33	1.20	1.58
$p = 0.50$ MAE(P2)	1.14	0.90	1.03	0.89	1.23
MAE(P1)	2.02	1.25	1.37	1.22	1.30
MAE( $l_p$ )	0.89	1.15	1.41	1.30	2.00
$p = 0.70$ MAE(P2)	1.23	0.82	0.96	0.77	1.14
MAE(P1)	2.02	1.25	1.37	1.22	1.30
MAE( $l_p$ )	0.96	1.13	1.44	1.40	2.16
$p = 0.90$ MAE(P2)	1.28	0.80	0.91	0.72	1.00
MAE(P1)	2.02	1.25	1.37	1.22	1.30
MAE( $l_p$ )	0.96	1.19	1.56	1.50	2.30

Table 2.4: Results in the unit "number of vehicles" for queue on lane M, estimated vs SUMO queue length.

which drivers choose the shortest queue. Moreover, the drivers do not have the same behavior, i.e. the choice is stochastic, i.e. the probability of taking the shortest queue increases with the threshold on the difference on the queue lengths. Another source of uncertainty and error is the variance of the arrival flows. Indeed, estimation of the queue length takes into account the average arrival flow  $\lambda$ , but in the simulation there is a variance of the time arrivals, which is then retrieved as an error of measurement.

We give for information the order of magnitude of the following indicator:  $\text{MAPE}(\text{P2,R}) :=$  the mean percentage absolute error between the estimated queue lengths as given by the estimator based on Proposition 2 and SUMO queue lengths on a subset of the data, at the end of the red time. Indeed, taking into account the MAPE at the beginning of the red time is meaningless (since the queues are not formed yet). This indicator rather emphasizes the difference between our assignment model  $\alpha^*$  and SUMO assignment model. This indicator varies between around 10% to 30% depending on the penetration ratio and the simulation scenarios. The assignment model of Proposition 4 is different from the SUMO assignment model. The estimations of the queue lengths given by Propositions 1-2 rely on the assignment model of Proposition 4. However, rather than using the assignment model of Proposition 4, the simulations were performed with the SUMO assignment model, which can then introduce some inconsistencies in the results.

In Fig. 2.7 and Fig. 2.8 we give the two lanes queue lengths in the scenario S4 where  $r_N = r_M$ . We estimate queue lengths for  $r_N > 0$  and  $r_M > 0$ . We notice that the estimation is more accurate as  $p$  gets higher.

### 2.5.3 Traffic light control and vehicles assignment onto the lanes

In this part, symmetric demand is assumed,  $\lambda_n = \lambda_m = \lambda_{nm} = 0.17$  vehicles/second. It is assumed that the traffic light cycle includes a phase of 8 seconds, where green light is given to lane  $M$ , while red light is given to lane  $N$ . Hence, the red duration on lanes  $N$  and  $M$  are different :  $r_N \neq r_M$ ,  $\bar{r} \neq 1$ . It is also assumed that  $\alpha = \alpha^*(\bar{r})$  given by Proposition 4.

On Fig. 2.9 we draw  $\bar{r}$  in blue and  $\alpha^*(\bar{r})$  in green depending on time. At the beginning of the cycle, the red is only for the lane  $N$ . The lane  $M$  is at this time with green light. Then, the two red lights are simply increasing as a line of slope 1. Starting from 20 seconds,  $\bar{r} \in I$  and  $\alpha^*(\bar{r}) = 1$ . This is because all the vehicles are assigned to lane  $M$  which is the shortest queue (we recall it was at green light until then). Then,  $\alpha^*(\bar{r})$  decreases slowly to reach approximately 0.6 which means that the two queue lengths are more equilibrated as the red durations on lanes  $N$  and  $M$  are getting less different. We notice that  $\alpha^*(\bar{r})$  would tend to  $1/2$  if the red time goes to infinity, because symmetric demand is assumed. Concerning the ratio of red lights  $\bar{r} = r_n/r_m$ , it is representing how the difference in the red

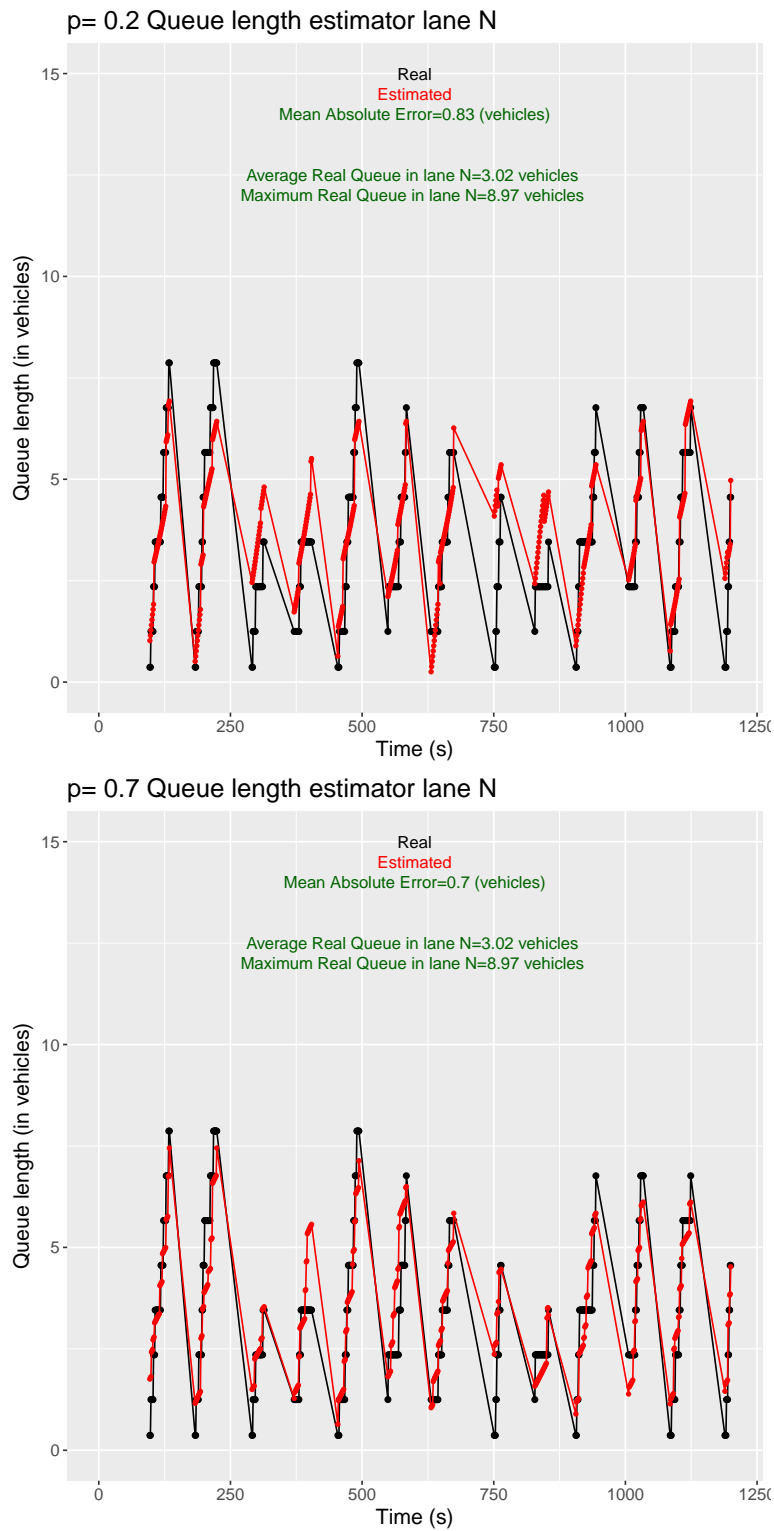


Figure 2.7: Queue lengths estimator as given by Proposition 2, for varying  $p = 0.2$ ,  $p = 0.7$  and lane N,  $\bar{r} = 1$

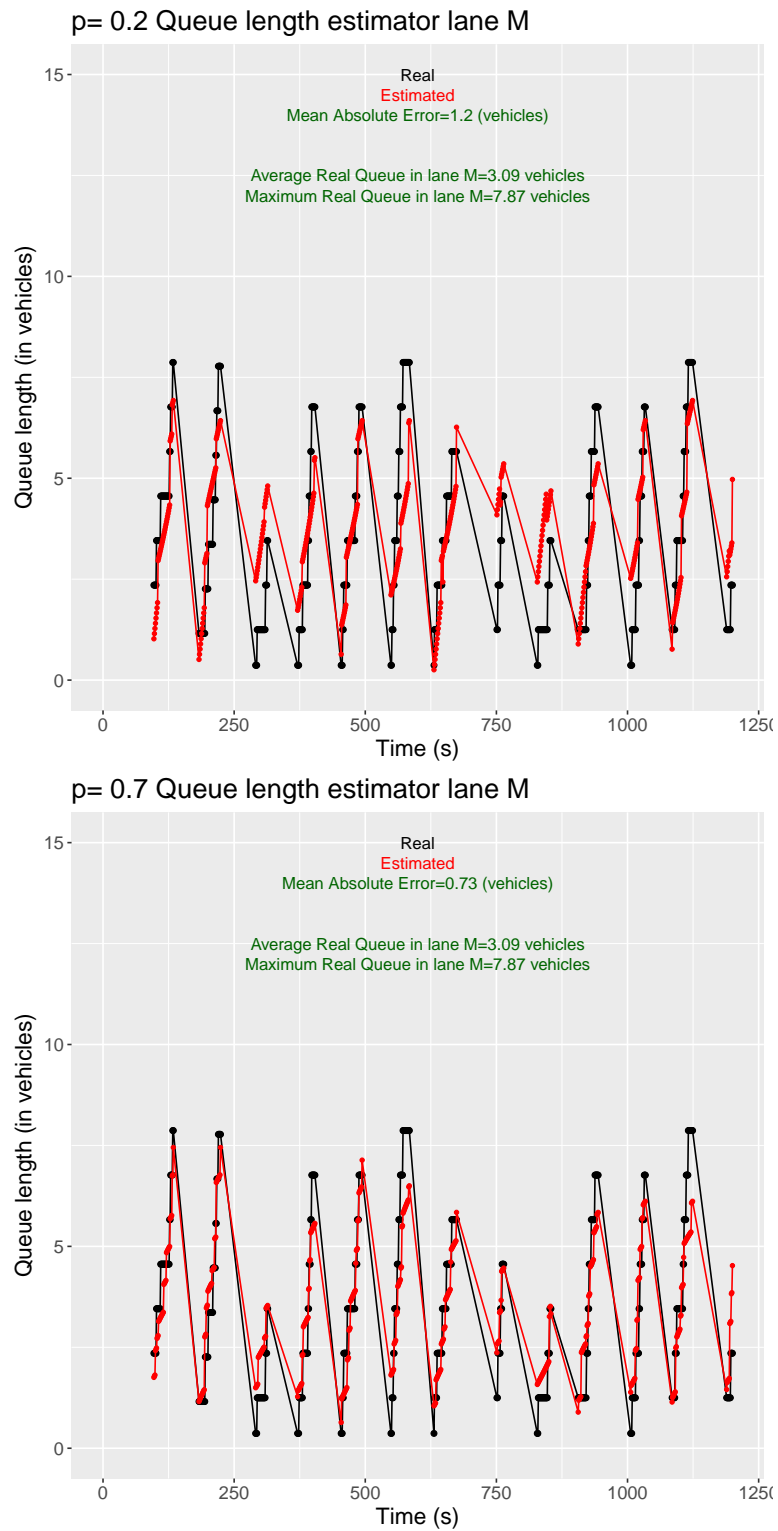


Figure 2.8: Queue lengths estimator as given by Proposition 2, for varying  $p = 0.2$ ,  $p = 0.7$  and lane M,  $\bar{r} = 1$

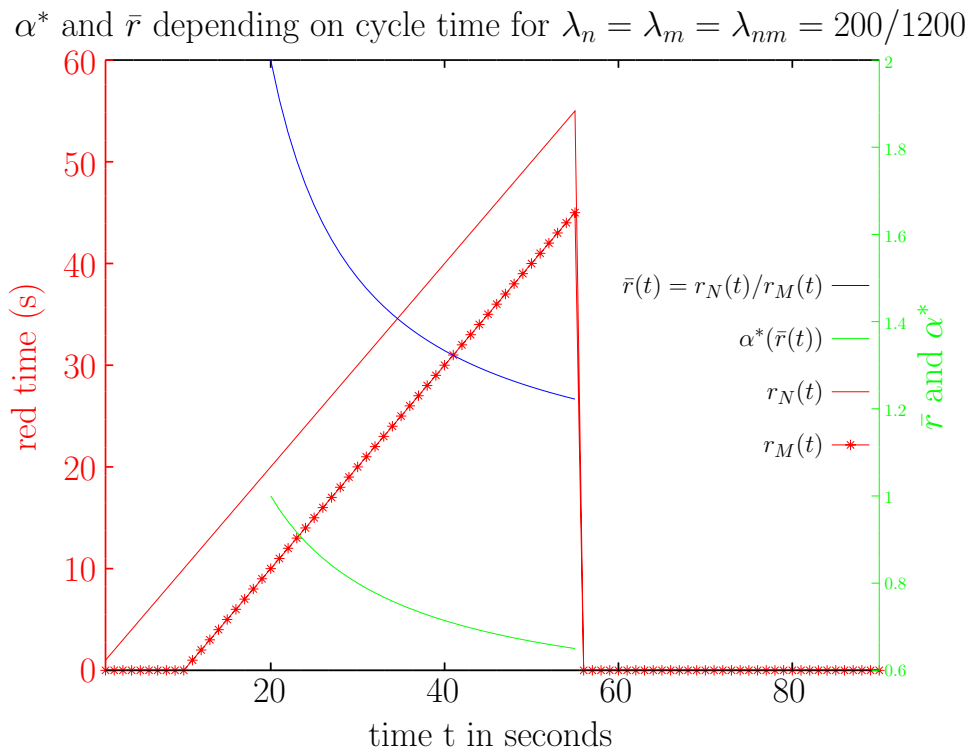


Figure 2.9:  $\alpha^*$  and  $\bar{r}$  depending on cycle time of 90s for  $\lambda_n = \lambda_m = \lambda_{nm} = \frac{200}{1200} = 0.17$  vehicles/second.



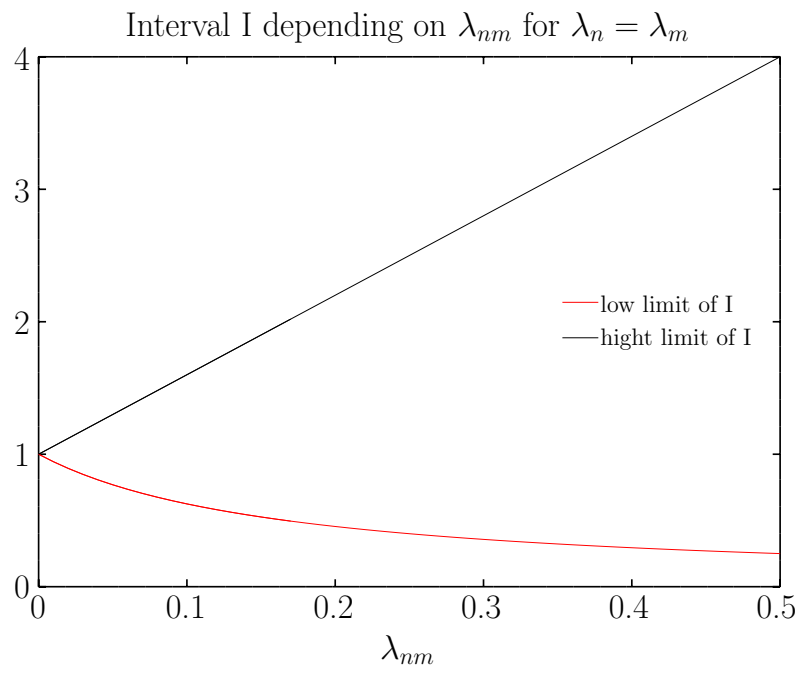


Figure 2.10:  $I = [\frac{\lambda_m}{\lambda_n + \lambda_{nm}}, \frac{\lambda_m + \lambda_{nm}}{\lambda_n}]$  depending on  $\lambda_{nm}$  for  $\lambda_n = \lambda_m = 1/6$

lights durations is decreasing as time is going on.  $\bar{r}$  decrease is due to a constant offset (corresponding to the duration where lane  $M$  is at green light while lane  $N$  is at red light) which becomes less significant as the red light durations are increasing. We notice that  $\bar{r}$  would tend to 1 if the red time goes to infinity.

The interval  $I = [\frac{\lambda_m}{\lambda_n + \lambda_{nm}}, \frac{\lambda_m + \lambda_{nm}}{\lambda_n}]$  as a function of  $\lambda_{nm}$  is represented on Fig. 2.10. We recall that if  $\bar{r} \in I$ , then  $0 < \alpha^*(\bar{r}) < 1$ , and in this case the common flow of vehicles can be splitted into two sub flows, i.e. the common flow is shared between the two lanes. We notice on Fig. 2.10 that as the common flow  $\lambda_{nm}$  gets higher, the interval  $I$  gets larger. Therefore, as the common flow  $\lambda_{nm}$  gets higher, there is more freedom to assign the vehicles onto a lane or another.

#### 2.5.4 Communication network performances

Some information on the communication network performances measured in simulation are given here. The communication performances are not disturbing the estimation of the queue lengths. As the order of magnitude of the end-to-end-delay <sup>2</sup> is very low (around 0.2 ms), no significant consequence on the queue length estimation application is expected, as it could happen in scenarios where more vehicles would communicate, and cause significant delays such as described in [115].

## 2.6 Conclusion and perspectives

In this paper, we have proposed a method for the estimation of urban traffic state. Estimations for the penetration ratio of probe vehicles and for the vehicles arrival rate, on any link of the road network, are given. Knowing the arrival rate of the incoming flow and its composition, the joint probability distribution of the queue lengths in the case of two lanes link is computed. For this purpose, a simple assignment model of vehicles onto the lanes has been proposed. In addition, the probability distribution of the queue lengths with the information provided by the probe vehicles is given. A control of the traffic light has been proposed in order to balance the queues of the two lanes. Moreover, a formula for computing the optimal assignment of the vehicles onto the lanes has been proposed. Numerical simulations have been conducted with Veins framework, and the work presented here has been evaluated. Road traffic control could benefit from the queue length estimations presented in the present paper, in order to improve travel conditions. The ideas given in this paper could be extended to a link of any number of lanes.

---

<sup>2</sup>The end-to-end delay is a communication indicator of performance that measures the delay from the time a message is sent from a communicating vehicle until the time it is received by the receiver (in our case the receiver is the RSU).

## 2.7 Appendix : Calculus of the bias of $p$ estimator

We consider queue lengths on lanes  $N$  and  $M$  respectively equal to  $n$  and  $m$ . We propose :

$$\hat{p} = \frac{c_p}{n + m} \quad (2.22)$$

By the way, in our case, the length  $n$  of queue  $N$  can be estimated with the number of arrivals on lane  $N$  during  $r_N(t)$  which is  $\mu_n$ . As  $\mu_n + \mu_m = \max(\mu_n, \mu_m) + \min(\mu_n, \mu_m)$  and by estimating  $\max(\mu_n, \mu_m) = l_p + i_c/p$ , where  $0 \leq i_c \leq 1$  and  $i_c/p$  represents the backlog of the queue (unequipped vehicles following  $l_p$ ). We want to determine  $0 \leq i_c \leq 1$  and compute  $\hat{p}$  such that the estimation of  $p$  is without bias. We can write :

$$\mu_n + \mu_m = \max(\mu_n, \mu_m) \left( 1 + \frac{\min(\mu_n, \mu_m)}{\max(\mu_n, \mu_m)} \right) \quad (2.23)$$

$$\mu_n + \mu_m = (l_p + i_c/p)(1 + \kappa) \quad (2.24)$$

We introduce  $c_\kappa = c_p/(1 + \kappa)$  and replace (2.24) in (2.22). Finally, we get the following equation :

$$\hat{p} = \frac{c_\kappa}{l_p + i_c/p} \quad (2.25)$$

$$\hat{p} = (c_\kappa - i_c)/l_p \quad (2.26)$$

We know from [27] how to compute the expectation of  $\hat{p}$  and we follow the same ideas below :

$$\begin{aligned} P(L_p^t = l_p, N_p^t + M_p^t = c_p) &= P(N_p^t + M_p^t = c_p | L_p^t = l_p) \\ &= \frac{P(L_p^t = l_p)}{P(L_p^t = l_p)} \end{aligned} \quad (2.27)$$

$$\begin{aligned} P(L_p^t = l_p, N_p^t + M_p^t = c_p) &= \binom{l_p - 1 + \min(l_p, n, m)}{c_p - 1} \\ &= p^{c_p - 1} (1 - p)^{l_p - 1 + \min(l_p, n, m)} P(L_p^t = l_p) \end{aligned} \quad (2.28)$$

$$\begin{aligned} \mathbb{E}\left(\frac{c_p/(1 + \kappa) - i_c}{l_p}\right) &= \sum_{l_p \geq 1} \sum_{c_p=1}^{l_p - 1 + \min(l_p, n, m)} \frac{c_p/(1 + \kappa) - i_c}{l_p} \\ &= \binom{l_p - 1 + \min(l_p, n, m)}{c_p - 1} \\ &= p^{c_p - 1} (1 - p)^{l_p - 1 + \min(l_p, n, m)} P(L_p^t = l_p) \end{aligned} \quad (2.29)$$

$$\begin{aligned}
 &= \sum_{l_p \geq 1} \frac{1}{l_p(1+\kappa)} \sum_{c_p=1}^{l_p-1+\min(l_p, n, m)} (c_p - 1 + 1 - i_c(1+\kappa)) \\
 &\quad \binom{l_p - 1 + \min(l_p, n, m)}{c_p - 1} \\
 &\quad p^{c_p-1} (1-p)^{l_p-1+\min(l_p, n, m)} P(L_p^t = l_p)
 \end{aligned} \tag{2.30}$$

To derive the next equation we use two arguments :

- the expectation of a binomial probability distribution law  $\mathbb{E}(\mathbb{B}(n_x, p)) = n_x p$  with  $n_x = l_p - 1 + \min(l_p, n, m)$  in our case.
- and the formula of Newton  $(a+b)^{m_x} = \sum_{k=0}^{m_x} \binom{m_x}{k} a^k b^{m_x-k}$ , with  $a = p$  and  $b = 1 - p$  in our case.

$$\begin{aligned}
 \mathbb{E}\left(\frac{c_p/(1+\kappa) - i_c}{l_p}\right) &= \sum_{l_p \geq 1} \frac{1}{l_p(1+\kappa)} \\
 &\quad (p(l_p - 1 + \min(l_p, n, m)) + 1 - i_c(1+\kappa)) \\
 &\quad P(L_p^t = l_p)
 \end{aligned} \tag{2.31}$$

We replace  $\min(l_p, n, m) = \kappa \max(l_p, n, m) = \kappa(l_p + i_c/p)$  :

$$\begin{aligned}
 \mathbb{E}\left(\frac{c_p/(1+\kappa) - i_c}{l_p}\right) &= \sum_{l_p \geq 1} \frac{1}{l_p(1+\kappa)} (p(l_p - 1 + \kappa(l_p + i_c/p)) + \\
 &\quad 1 - i_c(1+\kappa)) \\
 &\quad P(L_p^t = l_p)
 \end{aligned} \tag{2.32}$$

$$\begin{aligned}
 \mathbb{E}\left(\frac{c_p/(1+\kappa) - i_c}{l_p}\right) &= \frac{p}{1+\kappa} \mathbb{E}\left(\frac{l_p - 1}{l_p}\right) + \frac{p\kappa}{1+\kappa} + \\
 &\quad \frac{i_c\kappa}{1+\kappa} \mathbb{E}\left(\frac{1}{l_p}\right) + \frac{1 - i_c(1+\kappa)}{1+\kappa} \mathbb{E}\left(\frac{1}{l_p}\right)
 \end{aligned} \tag{2.33}$$

$$\mathbb{E}\left(\frac{c_p/(1+\kappa) - i_c}{l_p}\right) = p + \frac{\mathbb{E}\left(\frac{1}{l_p}\right)}{1+\kappa} (1 - p - i_c) \tag{2.34}$$

To get an estimator without bias, we write :

$$\mathbb{E}\left(\frac{c_p/(1+\kappa) - i_c}{l_p}\right) = p \tag{2.35}$$

Solving this equation gives :

$$i_c = 1 - p \quad (2.36)$$

Finally, we replace  $i_c$  in (2.26) :

$$\hat{p} = \frac{c_\kappa - (1 - p)}{l_p} \quad (2.37)$$

$$\hat{p} = \frac{c_\kappa - 1}{l_p - 1} \quad (2.38)$$

## Chapter 3

# Estimation of road traffic state at a multi-lanes controlled junction

# Estimation of road traffic state at a multi-lanes controlled junction

## 3.1 Abstract

We present in this paper a method for the estimation of traffic state at road junctions controlled with traffic lights. We assume mixed traffic where a proportion of vehicles are equipped with communication resources. The estimation of road traffic state uses information given by communicating vehicles. The method proposed is built upon a previously published method which was applied to estimate the traffic in the case where roads are composed of two lanes. In this paper, we consider the case where roads are composed of three lanes and we show that this solution can address the general case, where roads are composed of any number of lanes. The geometry of the road junction is assumed to be known, as well as its connections between incoming and outgoing lanes and roads. Using the location data provided by the communicating vehicles, first, we estimate some primary parameters including the penetration ratio of the probe vehicles, as well as the arrival rates of vehicles (equipped and non-equipped) per lane by introducing the assignment onto the lanes. Second, we give estimations of the queue length of the 3-lane road, without and with the additional information provided by the location of the communicating vehicles in the queue. The proposed model is discussed and illustrated with numerical simulations.

## 3.2 Introduction and state of the art

### 3.2.1 Introduction

Limited capacities of roads and junctions, combined with traffic demand, determine the road traffic conditions experimented by the users in daily life. Road traffic can then be modeled by shared resources systems such as queuing systems, as it has been done for example in [140], where the max pressure algorithm adapted to road traffic is presented. In order to improve road traffic conditions experimented by the users, there is the possibility to control road traffic by guiding the users in the network, or by controlling the traffic lights to reduce the delays. However, controlling the road traffic needs an information on the state of the traffic. In particular, concerning the road traffic state we are interested in the queue lengths at the junctions.

Nowadays, road traffic can be probed from the inside with communicating vehicles equipped with localization capabilities. A probe vehicle is a vehicle which uses wireless communication to send information to another vehicle (vehicle to vehicle V2V), the infrastructure (vehicle to infrastructure V2I) or to any other device (V2X). The data provided by these mobile

sensors are quite different in their nature from the data provided by fixed sensors. With probe vehicles, we get data concerning individual sample vehicles trajectories, rather than global information on traffic state at a fixed location. This difference in the nature of the information furnished, raises the opportunity to develop new methods for queue length estimation. For example, in [27] a stochastic method has been proposed to evaluate road traffic parameters like the queue lengths with probe vehicles. In [111], we have presented a method to estimate road traffic state at controlled junctions for the two-lanes roads which extended the method in [27].

In the present paper, we aim at generalizing the method published in [111] to roads composed of any number of lanes. This includes estimating road traffic primary parameters such as the penetration ratio of equipped vehicles, the arrival rate of vehicles and the queue lengths per lane at the junction. These estimations are enabled by the introduction of the assignment of vehicles onto the lanes.

The outline of the paper is as follows : after the introduction 3.2.1, a brief state of the art concerning the estimation of road traffic in subsection 3.2.2 is given. In section 3.3 the problem and the notations used in the paper as well as the main estimation model, are introduced. In subsection 3.3.1 we introduce the assignment onto the lanes in order to balance the queues as much as possible and derive primary parameters such as the penetration ratio of equipped vehicles, the arrival rate of vehicles per lane, the total arrival rate of vehicles, the probability distribution of the queue lengths per lane. In subsection 3.3.2, we give three different probability distributions for the queue lengths on a three-lane road. By computing the expectations for these probability distributions, the queue lengths at the traffic light are estimated. The method presented in the present paper is general for any number of lanes. We show in section 3.3.3 that the estimation of the queue lengths of three-lane roads is sufficient to address the general n-lane roads case. In section 3.4, we perform some numerical simulations that we have conducted with Omnet++ [142], a discrete event simulator, and analyze the results. Section 3.5 concludes the article.

## 3.2.2 State of the art

### 3.2.2.1 Road traffic estimation with fixed sensors (historical approaches)

Historically, the estimation of road traffic state was done using sensors placed at fixed locations, such as magnetic loops, piezoelectric sensors or video cameras [90]. Among these approaches with fixed sensors, some estimations of the queue length and of the delay of vehicles at a traffic light have been given in [13] [148] [108] [110] [7]. These papers give analytic formulations for the under-saturated and over-saturated (i.e. when the arrival flow exceeds



the intersection capacity) cases. Among the input-output class of methods, we also cite [145] where a probabilistic model for the estimation of queue lengths at signalized junctions, which can capture spillback and gridlock phenomena, is presented. In 2009, the authors of [102] take advantage of shockwave traffic theory, combined with fixed detector and signal timings data input, in order to estimate queue lengths.

### 3.2.2.2 Road traffic estimation with probe vehicles

In this paragraph, we give an overview of the methods that have been developed upon the information given by these new mobile sensors, namely, probe vehicles. In [67], the authors provide a timely survey on traffic information collection and state estimation methods published in the last decade, which use the data provided by connected and automated vehicles (CAVs). They classify the different traffic observation methods which use CAVs data by distinguishing between deterministic and stochastic approaches. The same outline is used here.

Among the deterministic approaches, the authors of [118] expose a method to estimate queue lengths with probe vehicles as the single source of information. Position and instantaneous speed of probe vehicles are the input data of their method. The latter relaxes some common assumptions made in the literature, such as the knowledge of signal timings or arrival process distribution. The shockwave theory based on first order traffic models is the key model used in their queue length estimation method. In [10], Ban et al. use intersection travel times in order to estimate queue lengths and delays at junctions. These intersection travel times are measured when probe vehicles cross virtual trip lines (VTL), located upstream and downstream relatively to the intersection. The benefits of using intersection travel times are : respect the privacy of the users, the flexibility in defining the virtual trip lines (as they are virtual locations), and the pliancy which enables other sensor inputs such as Bluetooth Mac address matching, and other travel times collection systems.

Among the stochastic approaches, the authors of [71] define a vehicle index as “the position of vehicles in the departure process of the cycle”. That paper has proposed a method for estimating these vehicle indices which are described as a basic and primordial information that can be provided by probe vehicles. For example, knowing the index of a vehicle gives its position in the queue. With the intent of solving some privacy issues, their method relies only on intersection travel times as input data. They derive the intersection travel times from the arrival time and departure time of probe vehicles into virtual areas (Virtual Trip Lines, VTL) respectively upstream and downstream the intersection. They model the arrival process as a time-dependent Poisson process; and use a log-normal distribution to model the departure headways for every vehicle index. In this framework, the authors

use a Bayesian Network in order to estimate vehicle indices.

Vehicle indices are some basic information that can be used as input data for estimating queue length at junctions. Indeed in 2014, Hao et al. [70] have naturally pursued their work on vehicle indices with the estimation of queue lengths at intersections. Using as input data the intersection travel times and the vehicle indices, as determined in their former 2013 paper, they estimate queue lengths using a stochastic model based on Bayes theorem.

In 2009, Comert and Cetin [28] have proposed a method for the estimation of queue length using the data provided by probe vehicles. They have assumed that probe vehicles indices are available as input data. Assuming that the probability distribution of the queue length is given, they compute a conditional probability distribution of the queue length, knowing the locations of the probe vehicles in the queue. They show that for the 1-lane case, the location of the last probe vehicle in the queue is the only one needed. The authors of [28] have also derived the variance of the estimator. Furthermore, numerical analysis are performed, where the arrival processes models and arrival processes intensities are varied. This work “appears to be the first attempt to formulate the problem of estimating the queue length from probe vehicle data.”

In 2013, Comert [26] derives queue length, last probe location and queue joining times probability distributions, with or without overflow queue (residual queue at the end of the red time). Mean and variance for the queue length estimators are given. It has been shown that the estimators depend on probe proportion, red duration and arrival process properties.

In 2016, Comert [27] goes a step further by studying the cases with unknown probe proportions and unknown arrival rates. He gives analytical formulations for these primary parameters (proportions of probe vehicles and arrival rate), as well as various queue length estimators with or without overflow queue. Derivation of the estimation errors are also given, and numerical analysis performed with VISSIM microscopic simulator have been presented.

In 2017, Zheng and Liu [164] estimated traffic volumes for low penetration ratio of equipped vehicles. The method proposed uses as input data “vehicle trajectories approaching to an intersection as well as traffic signal status”. The trajectories of equipped vehicles are used to detect if a probe vehicle has stopped at the traffic light and its stopping position. With these information, the arrival rate is estimated and bounds for this arrival rate are given. Zheng and Liu have used for their estimation a time dependent Poisson arrival process and the Expectation Maximization (EM) algorithm. They have tested their method with data sets from an experiment where around 2800 probe vehicles were deployed in the city of Ann Arbor, and from data provided by commercial navigation service in China. Machine learning methods have also been used in order to predict traffic flow; for example, Lv et al. [105] used a data driven machine learning method : the stacked

autoencoders model is combined with the data provided by 15000 individual detectors deployed across California in order to predict traffic conditions. In 2021, Zhao et al. [163] have proposed a method based on maximum likelihood algorithm in order to estimate the penetration ratio of probe vehicles as well as the distribution of queue lengths. The expectation-maximization algorithm is used to solve the problem formulated as maximum likelihood. The accuracy of the estimation results shows the relevance of their method. We also refer to the work by Tang et al. published in 2021 [134] in which the authors propose a method for the estimation of queue lengths using license plate recognition detectors. The authors introduce in their paper a lane based estimation of queue lengths which gives promising simulation and empirical results. Although the work by Tang et al. [134] is not using probe vehicles location as input data, it is addressing the multi lanes case combined with shockwave or input output methods.

### 3.2.3 Assumptions and main contributions

For the roads composed of many lanes, it is not accurate to use the model proposed by Comert [28] [26] [27]. This is because the shortest lanes would be equated to the longest lanes, which obviously is false, in particular in the case of unbalanced demands on the various lanes. Controls of the traffic light which are lane based would be inaccurate by not differentiating between the lanes. In the present paper, we generalize the method presented in [111] which estimates road traffic state at controlled junctions for the two-lane roads, to the case where roads are composed of any number of lanes. These methods are based on existing works published in [28] [26] [27] which were addressing the roads composed of one lane. To our knowledge, the present work and our previous paper [111] are the first works to extend the papers by Comert and Cetin [28] [26] [27] to the multi lanes case. In addition to this generalization, we give methods to estimate the number of probe vehicles per lane as well as a new estimator for the penetration ratio of probe vehicles. Based on these estimates, we also give a new analytical formulation for estimating the queue length per lane (Proposition 3).

The main assumptions of this article are the following :

- a proportion  $p$  of vehicles are equipped with wireless communication and localization capabilities, namely probe vehicles. The probe vehicles send their GPS localization to a road side unit using wireless communication. The penetration ratio  $p$  of probe vehicles is considered as a variable in the present paper.
- the traffic demand is low or moderate such that we can assume that the arrivals are following a Poisson process. Hence it is supposed that there is no overflow queue : the queues are cleared at the beginning of the red time.

- it is assumed that in average the queues tend to balance among the different lanes under the constraints of the vehicle assignments onto the lanes, with respect to their destinations.
- the GPS localization of vehicles is not accurate enough to distinguish on which lane a probe vehicle is located. This assumption relies on the fact that the accuracy of GPS localization is such that we can assign a vehicle with 95% probability within a radius of  $7.8m$  [63] although the width of a lane is  $3.7m$  in the USA.

### 3.3 Road traffic state estimation

The notations used in the present paper are described in TABLE 3.1; see also Fig. 3.1. We consider a road junction composed of a number of incoming

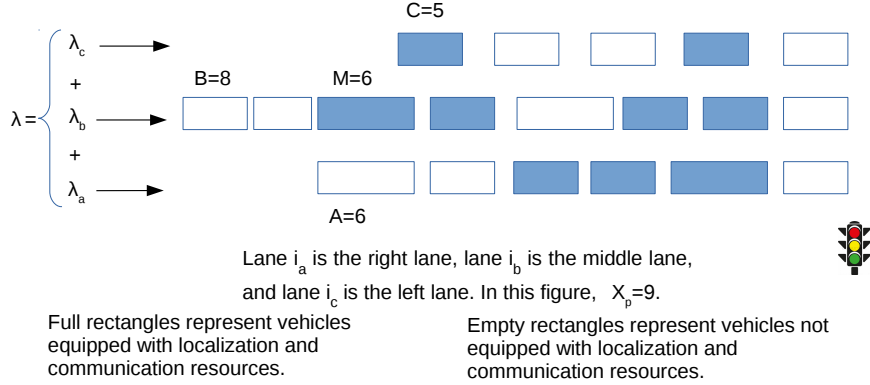


Figure 3.1: Vehicles queue at a controlled junction, in a three-lanes incoming road.

roads, one of them having three lanes, and  $d$  outgoing roads of any number of lanes, controlled by a traffic light. The vehicles come from an origin lane  $i \in O = \{i; i = i_a, i_b, i_c\}$  (where  $i_a, i_b, i_c$  denote the three origin lanes) of the 3-lanes incoming road; and go to a destination road  $j \in D := \{j; j = 1..d\}$ . We label the incoming vehicles by index  $k \in \mathbb{N}^*$  without taking into account their arrival order. Then, we introduce the two families  $O_k$  and  $D_k, k \in \mathbb{N}^*$  of random variables.  $O_k$  takes its values in  $O$ , such that  $O_k = i$  if vehicle  $k$  comes from origin lane  $i$ .  $D_k$  takes its values in  $D$ , such that  $D_k = j$  if vehicle  $k$  goes to destination road  $j$ . We assume that the probability that a vehicle  $k$  comes from incoming lane  $i \in O$  and goes to outgoing road  $j$ , is the same for all the vehicles  $k \in \mathbb{N}^*$ . Hence, we introduce the family of boolean random variables  $W_{ij}, i \in O, j \in D$  such that  $W_{ij} = 1$  if a vehicle  $k$  comes from the origin lane  $O_k = i$  and goes to destination road  $D_k = j$ ; and 0 otherwise. We will denote  $w_{ij} := P(W_{ij} = 1)$ , and  $W := (w_{ij})_{i \in O; j \in D}$  the matrix with

Name	Definition
$R$	the total red time in one cycle
$r_a(t), r_b(t), r_c(t)$	the time since the beginning of the red phase for respectively lane $i_a, i_b$ , or $i_c$ (it is 0 if we are not in red phase at time $t$ ), $0 \leq r_a(t), r_b(t), r_c(t) \leq R$ .
$\lambda_a, \lambda_b, \lambda_c$	the average arrival rate in vehicles/second respectively on lane $i_a, i_b$ and $i_c$ .
$\lambda = \lambda_a + \lambda_b + \lambda_c$	the total arrival rate for the incoming link in vehicles/second.
$X(t)$	the number of vehicles queuing on all the lanes, of the considered link at time $t$ .
$X_p(t)$	the number of communicating vehicles queuing on all the lanes of the considered link at time $t$ .
$Y(t)$	the number of vehicles (queuing and not queuing) of the considered link at time $t$ .
$Y_p(t)$	the number of communicating vehicles (queuing and not queuing) on all the lanes of the considered link at time $t$ .
$p, 0 \leq p \leq 1$	the penetration ratio of probe vehicles.
$A, B, C$	the number of vehicles in the queue at time $t$ and respectively lane $i_a, i_b$ , or $i_c$ . In this paper, $A, B, C$ are assumed to be random variables.
$(A_p, \hat{a}_p), (B_p, \hat{b}_p), (C_p, \hat{c}_p)$	the number of probe vehicles in the queue, and its estimate, at time $t$ and respectively at lane $i_a, i_b$ , or $i_c$ . In this paper, $A_p, B_p, C_p$ are assumed to be random variables.
$M$	the location (in number of vehicles) of the last probe in the queue, namely the last connected vehicle, at time $t$ . $M$ is assumed to be a random variable, taking value $m$ .
$L$	the incoming lane $i_a, i_b$ , or $i_c$ , of the last connected vehicle, at time $t$ . $L$ is assumed to be a random variable.
$O_k$	$O_k = i$ if a vehicle $k$ is located on an incoming lane $i = i_a, i_b, i_c$ . $O_k$ is considered a random variable.
$D_k$	$D_k = j$ if a vehicle $k$ is located on an outgoing road $j = 1, 2, \dots, d$ . $D_k$ is considered a random variable.
$W_{ij}$	$W_{ij}$ is a random variable. $W_{ij} = 1$ if a vehicle $k$ comes from the origin lane $i$ and goes to a destination road $j$ , such that $O_k = i$ and $D_k = j$ ; $W_{ij} = 0$ otherwise. We have the following notation : $P(W_{ij} = 1) = w_{ij}$ .
$W = (w_{ij})_{i=i_a, i_b, i_c; j=1..d}$	$W$ is the matrix with three lines and $d$ columns which represents the vehicles assignment from the incoming lanes to the outgoing roads.
$q_{sat}$	the saturation rate (exit rate) of roads outgoing from the junction
$t_e^{k,j}$	the exit time of the junction of a probe vehicle $k$ on an outgoing road $j$
$\pi(k, \mu_i) = \mu_i^k e^{-\mu_i} / k!$	the Poisson probability mass function of parameter $\mu_i$

Table 3.1: Notations

three lines and  $d$  columns which represents the vehicles assignment from the incoming lanes to the outgoing roads, and the assignment weight given to each couple (incoming lane, destination road).

### 3.3.1 Primary parameters

In this section, we determine the penetration ratio  $p$  of equipped vehicles, the total arrival rate  $\lambda$  of vehicles, the matrix  $W$ , and the arrival rates  $\lambda_i$  per incoming lane.

#### 3.3.1.1 Estimation model for the penetration ratio $p$

We will now estimate the penetration ratio  $p$  of probe vehicles. In [111], we have introduced the ratio  $\kappa = \min \mu_i / \max \mu_i$  and demonstrated that for the two lanes case and under the conditions of taking respectively the maximum and minimum queue lengths measurements as :  $l_p + 1/p - 1$ ,  $\kappa(l_p + 1/p - 1)$ , the estimator  $\hat{p} = \frac{X_p/(1+\kappa)-1}{m-1}$  is unbiased. However, these two queue lengths estimations are not given usually as measurements and we can not take  $\kappa$  as an average value because it will introduce some bias. This is why we propose another estimator for the penetration ratio  $p$  of probe vehicles which can be applied more easily in real life situations. The number of probe vehicles in an incoming road  $n$  is denoted by  $X_p(n)$ . Hence, the total number of probe vehicles in the queues of all incoming roads is  $\sum_n X_p(n)$  where the index  $n$  of the sum represents an incoming road. We denote  $t_e^{k,j}$  the exit time of a probe vehicle  $k$  on an outgoing road  $j$  during the green time. Considering one outgoing road  $j$ , the last probe vehicle going out to this road is located at  $q_{sat} \times \max_k \{t_e^{k,j}\}$  position (in number of vehicles) where  $q_{sat}$  is the saturation rate (exit rate) of vehicles per unit of time. Considering the vehicles waiting in the queue, we can virtually rearrange their order such that they are placed in the order they are going out of the junction. Notice that changing the order of vehicles does not change their total number. The following formula gives the estimation of  $p$  where the index of the sum  $n$  represents an incoming road and the index of the sum  $j$  represents an outgoing road.

$$\hat{p} = \frac{\sum_n X_p(n)}{q_{sat} \sum_j \max_k \{t_e^{k,j}\}} \quad (3.1)$$

The numerator of equation (3.1) is computed during the red time and the denominator during the green time.

#### 3.3.1.2 Estimation model for the total arrival rate $\lambda$

$\lambda$  can be estimated as follows :

$$\hat{\lambda} = \frac{Y_p(R) - Y_p(0)}{pR} \quad (3.2)$$

$Y_p$  represents the number of probe vehicles on the considered incoming road and  $R$  is the total red time.

### 3.3.1.3 Estimation model for the matrix $W$

The matrix  $W$  is used to calibrate the assignment model, by equilibrating the queue lengths among the three lanes  $i_a, i_b, i_c$ . The first step is to determine the matrix  $W$ . The probability that a vehicle  $k$  is coming from lane  $i$  is equal to  $w_i = \sum_j w_{ij}$ . We have  $0 \leq w_i \leq 1$  and  $\sum_i w_i = 1$ . The objective is to find the optimal assignment matrix  $W$  which equilibrates the ratios of inflows over the three lanes of the incoming road. The ideal case is :

$$\forall i, \sum_j w_{ij} = 1/3 \quad (3.3)$$

Therefore, in order to equilibrate the inflows over the three incoming lanes we propose to minimize the difference  $(\sum_j w_{ij} - 1/3)$  for all  $i = 1, 2, 3$ . We denote  $\bar{W} := (w_{11}, \dots, w_{1d}, w_{21}, \dots, w_{2d}, w_{31}, \dots, w_{3d})$  and  $v = (1/3, 1/3, 1/3)$ . We denote  $\bar{\lambda}_j$  the proportion of the arrival rate  $\lambda$  of vehicles going to the outgoing road  $j$ . The  $\tilde{w}_{ij} = 0$  represent the information given by the topology of the road junction, specifically the incoming lanes and outgoing roads which are not connected. The constraints are :

$$\sum_{ij} w_{ij} = 1 \quad (3.4)$$

$$\sum_i w_{ij} = \rho_j := \bar{\lambda}_j / \lambda, \forall j \quad (3.5)$$

$$\tilde{w}_{ij} = 0 \quad (3.6)$$

$$0 \leq w_{ij} \leq 1 \quad (3.7)$$

Note that, since the turn ratios  $\rho_j$  are fixed such that  $\sum_j \rho_j = 1$ , then the constraint (3.4) is automatically satisfied. The constraints (3.5) and (3.6) are linear and can be written as follows :

$$\bar{A}\bar{W} = \bar{b} \quad (3.8)$$

where

$$\bar{A} = \begin{pmatrix} I_d & I_d & I_d \\ & B & \end{pmatrix},$$

and  $\bar{b} = (\rho_1, \dots, \rho_d, 0, \dots, 0)$ , with  $I_d$  the  $d \times d$  identity matrix, and  $B$  the matrix satisfying  $B\bar{W} = \tilde{w}$ .

For the criterion, we define the  $3 \times (3d)$  matrix  $L$  such that :

$$(w_1, w_2, w_3) = \left( \sum_j w_{1j}, \sum_j w_{2j}, \sum_j w_{3j} \right) = L\bar{W}. \quad (3.9)$$

The objective function is :

$$\min_{\bar{W}} (L\bar{W} - v)'(L\bar{W} - v) \quad (3.10)$$

This can be written as :

$$\min_{\bar{W}} \bar{W}'Q\bar{W} - 2(v'L)\bar{W} \quad (3.11)$$

where  $Q = L'L$ . Finally, the minimization problem can be written as :

$$\begin{aligned} \min_{\bar{W}} \bar{W}'Q\bar{W} - 2(v'L)\bar{W} \\ \bar{A}\bar{W} = \bar{b} \\ 0 \leq w_{ij} \leq 1 \end{aligned} \quad (3.12)$$

The problem (3.12) is convex since the criterion is quadratic and the constraints are linear. Therefore, first order conditions of optimization are necessary and sufficient to solve this problem. Practically, an off-the-shelf optimization library provided with Octave software [40] is used to solve this problem. Then,  $\bar{W}$  (and then the matrix  $W$ ) is derived, given the turn ratios as input.

#### 3.3.1.4 Estimation of the arrival rates per lane $\lambda_i$

The arrival rates  $\lambda_i$  per incoming lane  $i$  in one cycle are given as follows :

$$\lambda_i = \lambda \sum_j w_{ij} = \lambda w_i \quad (3.13)$$

$(\lambda_i)_{i=i_a, i_b, i_c}$  represents the arrival rate of vehicles in the queue, respectively on lane  $i = i_a, i_b, i_c$ . This will enable to determine the probability distribution of queue lengths per lane in Proposition 5.

### 3.3.2 Traffic state estimation for roads composed of three lanes

In this section, we propose three probability distributions for the queue lengths. The expectation of these probability distributions are used as estimators. Proposition 5 gives the probability distribution without the information given by the probe vehicles but using the matrix  $W$ , especially the arrival rates per lane  $\lambda_i$ . Proposition 6 refines Proposition 5 by adding the information given by the probe vehicles. It is an extension of a previously published result [111]. In Proposition 7, we use the estimation of probe vehicles per lane to give another estimator.

We denote  $\mu_i$  the stock of vehicles waiting in the queue of lane  $i$  at time  $t$  :

$$\mu_i = \lambda_i r_i(t)$$



We denote  $\pi(k, \mu_i)$  the Poisson probability mass function of parameter  $\mu_i$ .

$$\pi(k, \mu_i) = \frac{\mu_i^k e^{-\mu_i}}{k!}, \forall i$$

Under the assumption of an arrival process of vehicles represented by a Poisson process of rate  $\lambda = \lambda_a + \lambda_b + \lambda_c$ , we have the following proposition :

**Proposition 5.**

$$P(A = a, B = b, C = c) = \pi(a, \mu_a)\pi(b, \mu_b)\pi(c, \mu_c)$$

*Proof.* The Poisson process of rate  $\lambda$  is subdivided into three independent Poisson processes of rate  $\lambda_a, \lambda_b, \lambda_c$  with probability respectively  $\lambda_a/\lambda, \lambda_b/\lambda, \lambda_c/\lambda$ . By subdividing the main Poisson process of parameter  $\lambda$  with these probabilities, we get three Poisson processes of parameter  $\lambda(\lambda_i/\lambda) = \lambda_i$  for each lane  $i$ . Furthermore, the three Poisson processes are independent. Concerning the subdividing of a Poisson process, see reference [57]. Because of the stationary increment property of a Poisson process, the expected number of vehicles queuing on lane  $i$  at time  $r_i(t)$  is  $\lambda_i r_i(t) = \mu_i$ .  $\square$

Knowing the location of the last probe vehicle into the queue and the total number of probe vehicles in the queue, the estimation of the queue lengths can be refined :

**Proposition 6.**

If  $m \leq \max(a, b, c)$  and  $x_p \leq a + b + c$ , then :

$$P(A = a, B = b, C = c | M = m, X_p = x_p) = \frac{\sigma_{a,b,c}(1-p)^{a+b+c}P(A = a, B = b, C = c)}{\sum_{\substack{j,k,l \geq 0 \\ \text{subject to} \\ m \leq \max(j,k,l) \\ x_p \leq j+k+l}} \sigma_{j,k,l}(1-p)^{j+k+l}P(A = j, B = k, C = l)}$$

Otherwise,

$$P(A = a, B = b, C = c | M = m, X_p = x_p) = 0.$$

where we define :

$$\sigma_{j,k,l} := \binom{m-1 + \min(m, j, k, l) + \text{mid}(j, k, l)}{x_p - 1}$$

and  $\text{mid}(j, k, l) := (j + k + l) - \min(j, k, l) - \max(j, k, l)$

*Proof.* The proof given in [111] for the Proposition 2 can be directly applied to the present case, with minor changes.  $\square$

We now propose to compute the estimation of the queues by adding the information of matrix  $W$  which summarizes the knowledge of the destinations of the probe vehicles at the road junction. Hence, we propose to compute for lane  $i_a$

$$p_A(a) = P(A = a | M = m, A_p = \hat{a}_p)$$

given the estimated number  $A_p = \hat{a}_p$  of probe vehicles on lane  $i_a$  and the location of the last probe in the queues. First, we will give estimations of the number of probe vehicles per lane  $A_p = \hat{a}_p$  in order to compute  $p_A(a)$ . We have :

$$A_p = \sum_k \mathbf{1}_{a_k},$$

where

$$\mathbf{1}_{a_k} := \begin{cases} 1 & \text{if vehicle } k \text{ comes from lane } i_a \\ 0 & \text{otherwise} \end{cases}$$

Then,

$$\begin{aligned} \mathbb{E}(A_p) &= \sum_k \mathbb{E}(\mathbf{1}_{a_k}) \\ &= \sum_k P(O_k = i_a) \\ &= \sum_{k,j} P(O_k = i_a | D_k = j) P(D_k = j) \end{aligned}$$

where  $D_k$  is observed. Then,  $P(D_k = j_k) = 1$  for the vehicle  $k$  after it has crossed the junction and has gone through an outgoing road  $j_k$  and  $P(D_k = j) = 0$  otherwise, for  $j \neq j_k$ .

We have :

$$\mathbb{E}(A_p) = \sum_k P(O_k = i_a | D_k = j_k) \quad (3.14)$$

$$\mathbb{E}(B_p) = \sum_k P(O_k = i_b | D_k = j_k)$$

$$\mathbb{E}(C_p) = \sum_k P(O_k = i_c | D_k = j_k)$$

We propose to estimate the number of probe vehicles on lanes  $i_a, i_b, i_c$  by respectively :  $\hat{a}_p = [\mathbb{E}(A_p)]$ ,  $\hat{b}_p = [\mathbb{E}(B_p)]$ ,  $\hat{c}_p = [\mathbb{E}(C_p)]$ , where  $[x]$  denotes rounding  $x$  to the nearest integer.

Now we propose to compute  $p_A(a)$  the queue length probability distribution on lane  $i_a$  with the information provided by the estimation of probe vehicles on lane  $i_a$  ( $\hat{a}_p$ ) and the location of the last probe vehicle on all the lanes ( $m$ ). The following proposition is also true for the other lanes by just inverting the lane  $i_a$  with  $i_b$  or  $i_c$ .

**Proposition 7.**

if  $a_p \geq 1, m \geq 1, a \geq a_p$ , then  $p_A(a) =$

$$\frac{\left(\lambda_a \binom{m-1}{\hat{a}_p-1} + \lambda_b S_{\mu_b}^{m, \hat{a}_p} \binom{a}{\hat{a}_p} + \lambda_c S_{\mu_c}^{m, \hat{a}_p} \binom{a}{\hat{a}_p}\right) (1-p)^a \pi(a, \mu_a)}{\sum_{n \geq a_p} \left(\lambda_a \binom{m-1}{\hat{a}_p-1} + \lambda_b S_{\mu_b}^{m, \hat{a}_p} \binom{n}{\hat{a}_p} + \lambda_c S_{\mu_c}^{m, \hat{a}_p} \binom{n}{\hat{a}_p}\right) (1-p)^n \pi(n, \mu_a)}$$

$$\text{if } a_p = 0, \text{ then } p_A(a) = \frac{(1-p)^a \pi(a, \mu_a)}{\sum_{n \geq 0} (1-p)^n \pi(n, \mu_a)}$$

if  $a < a_p$  or  $m < a_p$ , then  $p_A(a) = 0$

where we define :

$$S_{\mu}^{m, \nu} := \sum_{j \leq k, k \geq \max(m, \nu)} \binom{m-1}{j-1} p^j (1-p)^{k-j} \pi(k, \mu)$$

*Proof.* By Bayes' theorem, we have :

$$p_A(a) = \frac{P(A = a, M = m, A_p = \hat{a}_p)}{P(M = m, A_p = \hat{a}_p)} \quad (3.15)$$

The numerator can be written :

$$P(A = a, M = m, A_p = \hat{a}_p) = P(M = m | A_p = \hat{a}_p, A = a) P(A = a, A_p = \hat{a}_p) \quad (3.16)$$

Bayes theorem implies that :

$$P(A = a, A_p = \hat{a}_p) = P(A_p = \hat{a}_p | A = a) P(A = a) \quad (3.17)$$

So, we can write the second term of equation (3.16) as :

$$P(A = a, A_p = \hat{a}_p) = \binom{a}{\hat{a}_p} p^{\hat{a}_p} (1-p)^{a-\hat{a}_p} P(A = a) \quad (3.18)$$

Concerning the first term of the product in equation (3.16), we use the marginal distribution on the random variable  $L$ , and we can write :

$$\begin{aligned} P(M = m | A_p = \hat{a}_p, A = a) = & P(M = m, L = i_a | A_p = \hat{a}_p, A = a) + \\ & P(M = m, L = i_b | A_p = \hat{a}_p, A = a) + \\ & P(M = m, L = i_c | A_p = \hat{a}_p, A = a) \end{aligned} \quad (3.19)$$

Let us detail the first term of equation (3.19), in the case where  $i = i_a$ . Concerning  $P(M = m, L = i_a | A_p = \hat{a}_p, A = a)$ , we recall the arguments given in [28] which are : "The sample space for the experiment is the possible combinations of choosing  $\hat{a}_p$  probe vehicles from  $a$  vehicles, which is equal

to  $\binom{a}{\hat{a}_p}$ . The number of elements in the event space is equal to the number of possible placements of the remaining probe vehicles, other than the one at position  $m$ , into the previous slots since  $m$  is fixed.”, which is  $\binom{m-1}{\hat{a}_p-1}$ . As the last probe should be on lane  $i_a$  with probability  $\lambda_a/\lambda$ , we write :

$$P(M = m, L = i_a | A_p = \hat{a}_p, A = a) = \frac{\lambda_a}{\lambda} \frac{\binom{m-1}{\hat{a}_p-1}}{\binom{a}{\hat{a}_p}} \quad (3.20)$$

In addition, we have :

$$P(M = m, L = i_b | A_p = \hat{a}_p, A = a) = \frac{P(M = m, B \geq \max(m, \hat{a}_p), L = i_b | A_p = \hat{a}_p, A = a)}{P(B \geq \max(m, \hat{a}_p) | M = m, L = i_b, A_p = \hat{a}_p, A = a)} \quad (3.21)$$

Since  $P(B \geq \max(m, \hat{a}_p) | M = m, L = i_b, A_p = \hat{a}_p, A = a) = 1$ , we write :

$$P(M = m, L = i_b | A_p = \hat{a}_p, A = a) = P(M = m, B \geq \max(m, \hat{a}_p), L = i_b | A_p = \hat{a}_p, A = a) \quad (3.22)$$

We also have :

$$P(M = m, B \geq \max(m, \hat{a}_p), L = i_b) = \frac{P(M = m, B \geq \max(m, \hat{a}_p), L = i_b | A_p = \hat{a}_p, A = a)}{P(A_p = \hat{a}_p, A = a | M = m, B \geq \max(m, \hat{a}_p), L = i_b)} \times P(A_p = \hat{a}_p, A = a) \quad (3.23)$$

and because  $P(A_p = \hat{a}_p, A = a | M = m, B \geq \max(m, \hat{a}_p), L = i_b) = P(A_p = \hat{a}_p, A = a)$  we can write :

$$P(M = m, L = i_b | A_p = \hat{a}_p, A = a) = P(M = m, B \geq \max(m, \hat{a}_p), L = i_b) \quad (3.24)$$

By computing the marginal distribution probability on  $B_p$ , we have :

$$P(M = m, B \geq \max(m, \hat{a}_p), L = i_b) = \sum_{b_p, b \geq \max(m, \hat{a}_p)} P(M = m, L = i_b | B_p = \hat{b}_p, B = b) \times P(B_p = \hat{b}_p, B = b) \quad (3.25)$$

$$P(M = m, B \geq \max(m, \hat{a}_p), L = i_b) = \sum_{b_p \leq b, b \geq \max(m, \hat{a}_p)} \frac{\lambda_b}{\lambda} \frac{\binom{m-1}{\hat{b}_p-1}}{\binom{b}{\hat{b}_p}} \binom{b}{\hat{b}_p} p^{\hat{b}_p} (1-p)^{b-\hat{b}_p} P(B = b) \quad (3.26)$$

$$P(M = m, B \geq \max(m, \hat{a}_p), L = i_b) = \sum_{b_p \leq b, b \geq \max(m, \hat{a}_p)} \frac{\lambda_b}{\lambda} \binom{m-1}{\hat{b}_p - 1} p^{\hat{b}_p} (1-p)^{b-\hat{b}_p} P(B = b) \quad (3.27)$$

We define the variable  $S_\mu^{m,\nu}$  :

$$S_\mu^{m,\nu} := \sum_{j \leq k, k \geq \max(m,\nu)} \binom{m-1}{j-1} p^j (1-p)^{k-j} \pi(k, \mu)$$

Then, we can write :

$$P(M = m, B \geq \max(m, \hat{a}_p), L = i_b) = \frac{\lambda_b}{\lambda} S_{\mu_b}^{m, \hat{a}_p} \quad (3.28)$$

Similarly,

$$P(M = m, C \geq \max(m, \hat{a}_p), L = i_c) = \frac{\lambda_c}{\lambda} S_{\mu_c}^{m, \hat{a}_p} \quad (3.29)$$

Finally,

$$P(M = m | A_p = \hat{a}_p, A = a) = \frac{\lambda_a}{\lambda} \frac{\binom{m-1}{\hat{a}_p - 1}}{\binom{a}{\hat{a}_p}} + \frac{\lambda_b S_{\mu_b}^{m, \hat{a}_p}}{\lambda} + \frac{\lambda_c S_{\mu_c}^{m, \hat{a}_p}}{\lambda} \quad (3.30)$$

Hence, the numerator can be written as :

$$P(A = a, M = m, A_p = \hat{a}_p) = \left( \frac{\lambda_a}{\lambda} \frac{\binom{m-1}{\hat{a}_p - 1}}{\binom{a}{\hat{a}_p}} + \frac{\lambda_b S_{\mu_b}^{m, \hat{a}_p}}{\lambda} + \frac{\lambda_c S_{\mu_c}^{m, \hat{a}_p}}{\lambda} \right) \times \binom{a}{\hat{a}_p} p^{\hat{a}_p} (1-p)^{a-\hat{a}_p} P(A = a) \quad (3.31)$$

Using the marginal distribution and some simplifications, we can write in conclusion that :

$$p_A(a) = \frac{\left( \lambda_a \binom{m-1}{\hat{a}_p - 1} + \lambda_b S_{\mu_b}^{m, \hat{a}_p} \binom{a}{\hat{a}_p} + \lambda_c S_{\mu_c}^{m, \hat{a}_p} \binom{a}{\hat{a}_p} \right) (1-p)^a \pi(a, \mu_a)}{\sum_{n \geq \hat{a}_p} \left( \lambda_a \binom{m-1}{\hat{a}_p - 1} + \lambda_b S_{\mu_b}^{m, \hat{a}_p} \binom{n}{\hat{a}_p} + \lambda_c S_{\mu_c}^{m, \hat{a}_p} \binom{n}{\hat{a}_p} \right) (1-p)^n \pi(n, \mu_a)}$$

□

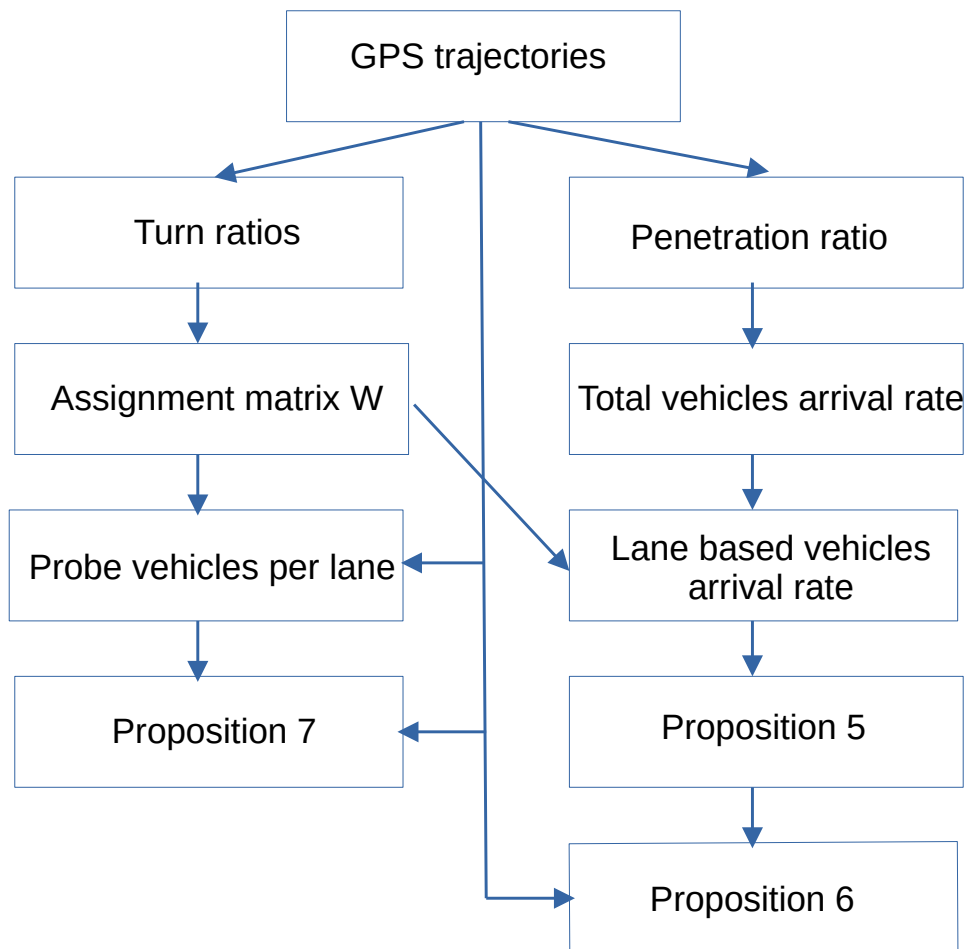


Figure 3.2: Summary of the estimation process

In Fig.3.2, the process of the road traffic estimation is summarized. Based on the trajectories of the probe vehicles, the turn ratios and the penetration ratio of probe vehicles are estimated. The assignment matrix  $W$  is derived as the solution of an optimization problem which tends to equilibrate the queues per lane as much as possible. The total vehicles arrival rate estimation combined with the assignment matrix  $W$  gives the vehicles arrival rate per lane. On another hand, the number of probe vehicles per lane is estimated with the assignment matrix  $W$  and the trajectories of the probe vehicles. Proposition 5 is derived from the vehicles arrival rate per lane. Proposition 6 is extending Proposition 5 while using the information given by the probe vehicles. Finally, Proposition 7 uses the estimated number of probe vehicles per lane combined with the information given by the probe vehicles, in order to estimate the queue length distribution probability per lane.

### 3.3.3 Application to roads composed of any number of lanes

The method proposed for the three propositions can be generalized easily to roads composed of an any number of lanes. However, we give in this section insights into the n-lane roads case and we demonstrate that the 3-lanes roads case is enough to address the general case. Indeed, with the accuracy of the GPS localization system nowadays, we can assign a vehicle with 95% probability within a radius of  $7.8m$ , as it is written in [63] : “the government commits to broadcasting the GPS signal in space with a global average user range error (URE) of  $\leq 7.8m$  (25.6 ft.), with 95% probability.”.

As the standard lane width in the United States is  $3.7m$ , any vehicle located on a lane can be assigned to a virtual three-lane road. All the virtual roads composed of successive three lanes are enumerated given the topology of the road junction. The virtual roads can overlap such that a lane can be in many virtual roads. The estimation of a queue length for a lane which is present only in a single virtual 3-lane road is straightforward with the method exposed. For the other lanes, which are represented in many virtual roads, we can take the average of the estimation done in each of the virtual roads. By taking the average of the queue length estimations for a given lane, on the set of all the virtual roads, we counterbalance the inaccuracy due to the estimation of  $A_p = \hat{a}_p$ ,  $B_p = \hat{b}_p$ ,  $C_p = \hat{c}_p$ . Indeed, we recall that these variables are function of the expected values :  $\hat{a}_p = [\mathbb{E}(A_p)]$ ,  $\hat{b}_p = [\mathbb{E}(B_p)]$ ,  $\hat{c}_p = [\mathbb{E}(C_p)]$ .

## 3.4 Numerical experiments

### 3.4.1 Estimation of the primary parameters

In this section, the model is discussed with simulation results. The simulation model is implemented with Omnet++ [142]. In the Omnet++ implementation, we represent each vehicle with one packet. A packet contains basic information such as the arrival and exit times, the destination, and whether the vehicle is a probe vehicle or not. Packets arrive as a Poisson arrival process of rate  $\lambda$ , i.e. the inter-arrival times between packets follow an exponential distribution probability law. The packets (vehicles) are assigned to the queue in accordance with matrix  $W$  and their destination, and are extracted from each queue by a traffic light at a given saturation rate during the green time. In [111], we have developed an implementation within Veins framework [131] which combines a road traffic simulator (SUMO [86]) and a communication simulator (Omnet++). We used to combine SUMO microscopic road traffic simulator to represent the dynamics of the vehicles and Omnet++ to simulate the communication between the vehicles and the infrastructure which is a Road Side Unit (RSU) in our case. The difference between the two implementations is the assignment model, which is not the same in SUMO and in our model. In the Omnet++ implementation, the assignment simulation model corresponds accurately with our theoretical assignment model. But in the SUMO implementation, the simulation assignment model is the SUMO lane changing model [44]. This is why in the present paper we use a simulation model using Omnet++ exclusively.

We consider a symmetric and an asymmetric scenarios for the simulation. For both scenarios, there are three outgoing roads. In both cases, the topology of the road junction is as follows : if a vehicle turns left, it must be on the left lane; if it turns right, it must be on the right lane; if it goes straight, it can be on any of the three lanes. So we have  $\tilde{w}_{21} = \tilde{w}_{31} = \tilde{w}_{13} = \tilde{w}_{23} = 0$ ; see section 3.3.1. The matrix  $W$  in the two scenarios is the result of the optimization problem given in (3.12) with the constraints (3.6) on the variable  $\tilde{w}$  given by the junction topology.

**Scenario S1** The turn ratios for the symmetric scenario S1 are  $(\rho_1, \rho_2, \rho_3) = (0.1, 0.8, 0.1)$ . The total arrival rate  $\lambda = 0.75$  vehicles/second. Then the optimization problem (3.12) gives the matrix  $W$  given in TABLE 3.2. We observe that this scenario S1 is symmetric, because the incoming lanes are balanced :  $w_{i_a} = w_{i_b} = w_{i_c} = 1/3$ .

**Scenario S2** The turn ratios for the asymmetric scenario S2 are  $(\rho_1, \rho_2, \rho_3) = (0.7, 0.15, 0.15)$ . The total arrival rate  $\lambda = 0.5$  vehicles/second. Then the optimization problem (3.12) gives the matrix  $W$  given in TABLE 3.3. The scenario S2 is asymmetric since  $w_{i_a} \neq w_{i_b}$  and  $w_{i_a} \neq w_{i_c}$ . In Fig. 3.3 is displayed the penetration ratio estimator proposed in section 3.3.1. The estimator  $\hat{p} = \frac{\sum_n X_p(n)}{q_{sat} \sum_j \max_k \{t_e^{k,j}\}}$  generalizes the estimators proposed in [27] to



Symmetric scenario S1	Destination			$w_i$
Incoming lane	1	2	3	
$i_a$	0.1	0.23	0	1/3
$i_b$	0	0.33	0	1/3
$i_c$	0	0.23	0.1	1/3
$\rho_j$	0.1	0.8	0.1	

Table 3.2: Matrix  $W$  for the symmetric scenario. Total arrival rate=0.75 vehicles/second.

Asymmetric scenario S2	Destination			$w_i$
Incoming lane	1	2	3	
$i_a$	0.7	0	0	0.7
$i_b$	0	0.15	0	0.15
$i_c$	0	0	0.15	0.15
$\rho_j$	0.7	0.15	0.15	

Table 3.3: Matrix  $W$  for the asymmetric scenario. Total arrival rate=0.5 vehicles/second.

the multi-lanes case. We observe on Fig. 3.3 that for  $p > 0.15$  the estimation error is very low compared to the case  $p < 0.15$ .

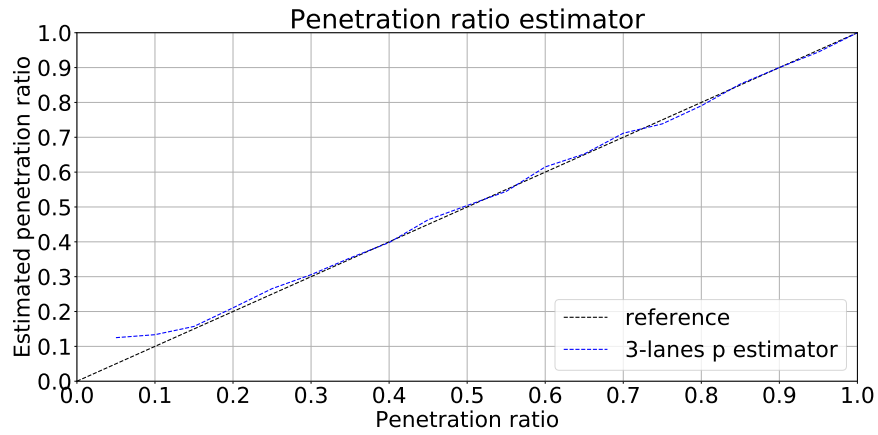


Figure 3.3: Penetration ratio estimator. Simulated time=25 hours.

In Fig. 3.4, the mean absolute error of various estimators for computing  $A_p$ ,  $B_p$ ,  $C_p$  are shown. The first estimator is the one which estimates the number of probes per incoming lane, based on the destination road of each probe vehicle. For example, with this estimator we will assume that all the vehicles that have turned right came from the right lane. We will denote

this estimator as  $E_0$ . With  $E_0$ , we write :

$$\hat{a}_p = \left[ \sum_k P(D_k = j_1) \right] \quad (3.32)$$

$$\hat{b}_p = \left[ \sum_k P(D_k = j_2) \right] \quad (3.33)$$

$$\hat{c}_p = \left[ \sum_k P(D_k = j_3) \right] \quad (3.34)$$

A more realistic estimator,  $E_1$  which is based on Equation (3.14) is recalled here :

$$\hat{a}_p = \left[ \sum_k P(O_k = i_a | D_k = j_k) \right] \quad (3.35)$$

$$\hat{b}_p = \left[ \sum_k P(O_k = i_b | D_k = j_k) \right] \quad (3.36)$$

$$\hat{c}_p = \left[ \sum_k P(O_k = i_c | D_k = j_k) \right] \quad (3.37)$$

We notice in Fig. 3.4 that the estimator  $E_1$  performs better than the estimator  $E_0$ . The estimation error for  $E_1$  is at maximum around 1 vehicle, although it can be around 3 vehicles for  $E_0$ . This is because with the estimator  $E_1$ , the matrix  $W$  is used to enhance the estimator  $E_0$  which estimates probe vehicles per lane based on the counting of probe vehicles per destination. We notice also that estimation error for  $E_0$  increases linearly with the penetration ratio  $p$  while the estimation error for  $E_1$  is clearly sub linear.

### 3.4.2 Road traffic state estimation

In this section we evaluate the queue length estimations as given by the expectations of the probability distributions of Propositions 5,6 and 7. We notice that Proposition 6 is extending the work by Comert [28] to the multi lanes case.

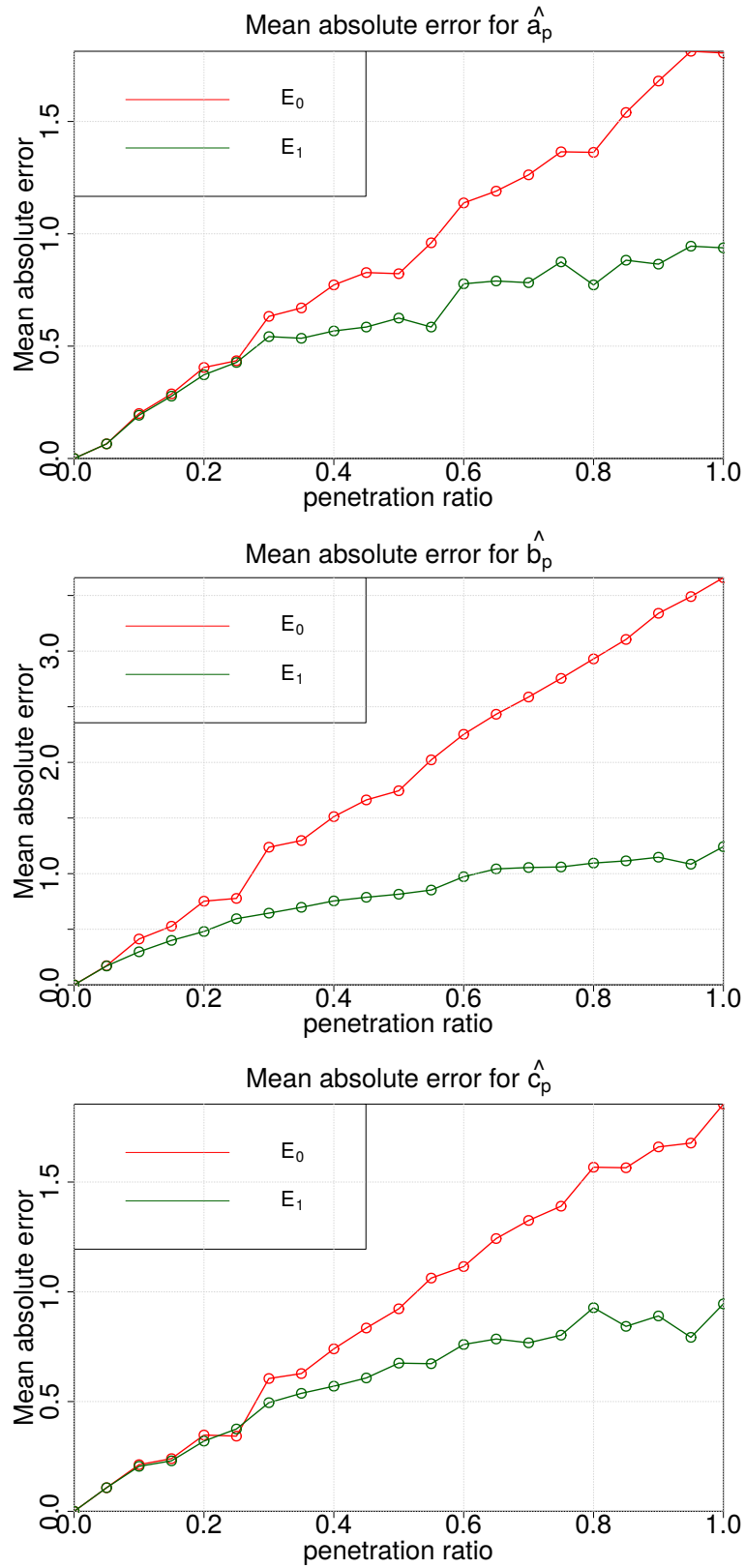


Figure 3.4: Mean absolute error for  $\hat{a}_p$ ,  $\hat{b}_p$ ,  $\hat{c}_p$  and estimators  $E_0$ ,  $E_1$ , averaged over 10 replications. Simulated time = 10 hours, asymmetric scenario.

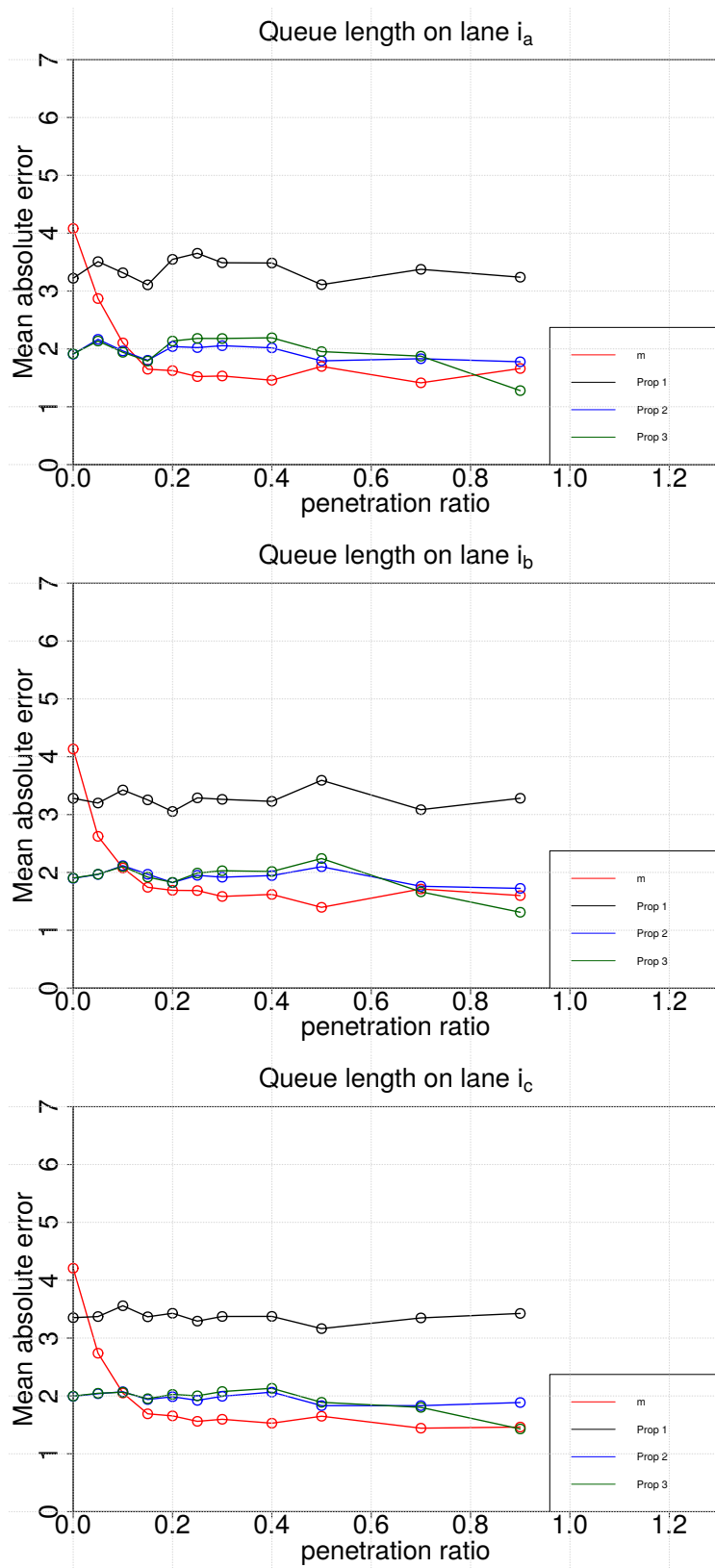


Figure 3.5: Mean absolute error for queue lengths on lanes  $i_a$ ,  $i_b$  and  $i_c$ , with Propositions 1, 2, 3.  $m$  means estimating the queue length with  $\hat{A} = \hat{B} = \hat{C} = m$  the last probe location. Simulated time = 2.5 hours and symmetric scenario.

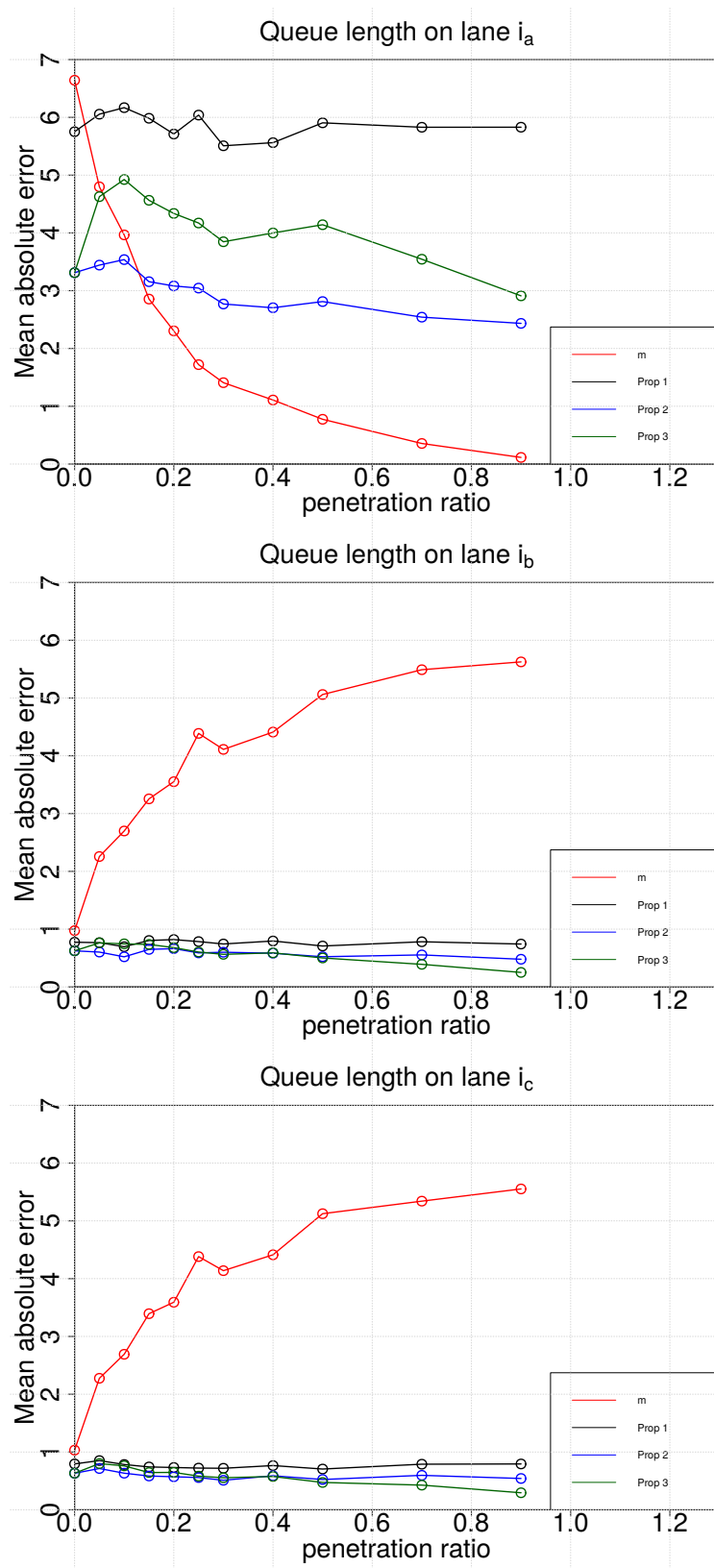


Figure 3.6: Mean absolute error for queue lengths on lanes  $i_a$ ,  $i_b$  and  $i_c$ , with Propositions 1, 2, 3.  $m$  means estimating the queue length with  $\hat{A} = \hat{B} = \hat{C} = m$  the last probe location. Simulated time = 2.5 hours and asymmetric scenario.

In Fig. 3.5 and Fig. 3.6, we compare the mean absolute error between the queue lengths and their estimated values with the different propositions respectively for scenario S1 and S2. We notice on Fig. 3.5 that the estimator  $m$  performs quite well for high penetration ratios, because we consider here the symmetric scenario. On the other hand, we notice on Fig. 3.5 that the estimator  $m$  is not accurate for low penetration ratios even in this symmetric scenario. We notice also that Proposition 6 and Proposition 7 perform better than Proposition 5 because we add some information to the estimations. In addition, we notice that all the estimators increase their accuracy as the penetration ratio of probe vehicles increases.

Let us now look at the simulation results of the asymmetric scenario. We notice on Fig. 3.6 that although the estimator  $m$  performs quite well on lane  $i_a$ , it is not true for lanes  $i_b$  and  $i_c$  because the demand is asymmetric. The estimator  $m$  is not an option for lanes  $i_b$  and  $i_c$ . On another hand, Proposition 6 and Proposition 7 perform better than Proposition 5 for all the lanes. Proposition 6 performs better than Proposition 7 for the longest lane  $i_a$  although for the lanes  $i_b$  and  $i_c$  Proposition 7 is more accurate than Proposition 6.

### 3.5 Conclusion and perspectives

A method to estimate the road traffic state at a multi-lanes road junction has been proposed. We have given an estimator for the penetration ratio of communicating vehicles as well as the arrival rate of vehicles. Based on the assumption that the queues tend to balance, an assignment matrix which gives the probabilities that a vehicle comes from an origin lane and goes to a destination road, has been derived. An existing method for road traffic state estimation on 2-lane roads has been extended to the case where roads are composed of any number of lanes. Three estimators for the queue lengths at the road junction have been proposed. The model was implemented with a discrete event simulator where vehicles are represented by packets. Numerical experiments allow to discuss the propositions and confirm that the model performs good especially for the asymmetric traffic demand scenarios. Concerning the future works regarding the present paper, we notice that the road traffic demand was assumed to be moderate or low such that the arrival process can be considered as a Poisson process. A future work could be to address the case of high traffic demand by taking into consideration the overflow queue. On another hand, it seems relevant to have assumed that the vehicles tend to choose the lane with the shortest queue although it lacks real assignment data to confirm this assumption. Finally, even if the GPS localization system becomes more accurate, there will be a need for robust models which can take into account inaccuracies on the localization of probe vehicles.



## Part III

# Control of urban road traffic





This part is built upon two conference papers [53] [139] from 2015 and 2017. It deals with control of urban road traffic with communicating vehicles. In chapter 4, a conference paper [53] extends a centralized approach (the Traffic Urban Control method [33, 35, 34]) by introducing some contention window where the vehicles communicate with each other to pass the junction. In chapter 5, a conference paper [139] details how such a contention window could be implemented with communicating vehicles.



## Chapter 4

# A semi-decentralized control strategy for urban traffic

# A semi-decentralized control strategy for urban traffic

## 4.1 Abstract

We present in this article a semi-decentralized approach for urban traffic control, based on the TUC (Traffic responsive Urban Control) strategy. We assume that the control is centralized as in the TUC strategy, but we introduce a contention time window inside the cycle time, where antagonistic stages alternate a priority rule. The priority rule is set by applying green colours for given stages and yellow colours for antagonistic ones, in such a way that the stages with green colour have priority over the ones with yellow colour. The idea of introducing this time window is to reduce the red time inside the cycle, and by that, increase the capacity of the network junctions. In practice, the priority rule could be applied using vehicle-to-vehicle (v2v) or vehicle-to-infrastructure (v2i) communications. The vehicles having the priority pass almost normally through the junction, while the others reduce their speed and yield the way. We propose a model for the dynamics and the control of such a system. The model is still formulated as a linear quadratic problem, for which the feedback control law is calculated off-line, and applied in real time. The model is implemented using the Simulation of Urban MObility (SUMO) tool in a small regular (American-like) network configuration. The results are presented and compared to the classical TUC strategy.

## 4.2 Introduction

Recent advances in information and communication technologies improve vehicular traffic in urban road networks by enabling the development of innovative urban traffic control strategies. While the traffic control in urban road networks is still done by setting traffic lights, intelligent transportation systems (ITS) are being tested in many cities. Various agents in the road network will be able to communicate from vehicle to vehicle (V2V) or from vehicle to infrastructure (V2I) for example. Big data sets, with different levels of information (microscopic, macroscopic) will be processed in real time and adaptive control strategies will be applied. The whole process of urban traffic control needs to be redefined in order to take into account this development.

Several levels of information are distinguished in the big amounts of data that are made available by ITS. The whole information cannot be optimally exploited with a unique centralized or distributed traffic control system. A multi-level control system needs to be developed in order to optimally use

each level of information for the corresponding control level. Macroscopic information could be transmitted to the centralized controller, while the microscopic one could be used by the local controller, which should operate in a short time horizon, compared to the high-level controller. Multi-level control schemes have been recently proposed; see for example [119, 141]. In [119], the control uses macroscopic fundamental diagrams (MFD) [59, 29, 49, 50, 51, 47, 48].

Using traffic lights, the main urban traffic parameters are: phase specification, split, cycle time, and offset. Fixed time urban traffic control (UTC) strategies appeared in the 1950s with coordination of signals that optimizes the offsets. These strategies use historic datasets, and therefore, are unable to adjust to changing conditions. The most well-developed and widely used UTC system is TRANSYT [122]. With advances in detection, communication, data processing, and control strategies, traffic responsive UTC systems appeared, where centralized and distributed systems are distinguished. Among the main centralized ones, we cite SCOOT [77, 18], SCATS [103], RHODES [74], MOTION [20], and TUC [33]. For distributed responsive UTC, we cite UTOPIA [39], PRODYN [46], OPAC [58]. Other UTC systems define an intermediate level of centralization.

Traffic responsive UTC systems use feedback controls on the state of the traffic and permit, by that, to meet traffic demand. Moreover, the control may be set in such a way to be robust, in the sense that it responds rapidly to disruptions. Furthermore, such controls are automatically adaptive to works and operations, and so installation and maintenance costs are reduced.

We propose in this article an extension of the traffic responsive urban control strategy TUC (Traffic Urban Control) [33, 35, 34]. Our extension introduces a kind of decentralization in the optimization of the right of way assignment. We introduce a contention time window inside the cycle time, where antagonistic stages alternate a priority rule. The priority rule is set by applying green colours for given stages and yellow colours for antagonistic ones, in such a way that the stages with green colour have priority over the ones with yellow colour. A TUC-based centralized control determines the optimal split of green, red and yellow lights at the level of every junction. A decentralized system manages the traffic of antagonistic stages during the yellow signal, taking into account the characteristics of each junction. By doing this, we aim to reduce the red time inside the cycle, increase the capacity of the network, and reduce users' delays. The traffic management during the yellow times would be realized based on vehicle to vehicle (v2v) and/or vehicle to infrastructure (v2i) communications.

We present in this article preliminary results of this semi-decentralization on a small American-like city. The results demonstrate the efficiency of this extension with respect to the classical TUC control. On a selected scenario of traffic demand, we show that the semi-decentralized TUC controls better the traffic, in the sense that it is able to respond efficiently and rapidly to

congestion.

### 4.3 A short review of TUC

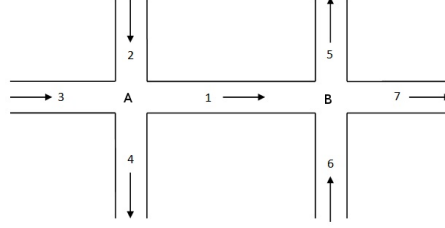


Figure 4.1: Academic example explaining the TUC strategy.

TUC [33, 35, 34] is a coordinated control strategy based on a store-and-forward approach. It can be implemented for large-scale networks, in real time, even under saturated traffic conditions. The split control part of TUC varies the green-stage durations of all stages at all the junctions of an urban network around given nominal values, and under a simplified traffic dynamics. The objective is to avoid oversaturations and spillbacks of link queues. In order to briefly explain the approach, let us consider the small network of Figure 4.1, with the following notations.

$c$	cycle time duration, in seconds.
$k$	discrete time integer index, corresponding to a duration of $kc$ sec.
$x_i(k)$	number of cars on link $i$ at discrete time $k$ .
$\bar{x}_i$	constant nominal number of cars on link $i$ .
$\Delta x_i(k)$	$= x_i(k) - \bar{x}_i$ .
$s_i$	saturation flow on link $i$ .
$g_i(k)$	green time duration for link $i$ during the $k$ th cycle.
$\bar{g}_i$	constant nominal green time duration for the stream coming from link $i$ .
$\Delta g_i(k)$	$= g_i(k) - \bar{g}_i$ .
$u_i(k)$	$= (g_i(k)/c)s_i$ average outflow from link $i$ during the $k$ th cycle.
$d_i(k)$	arrival demand flow to link $i$ at discrete time $k$ .
$\bar{d}_i$	constant nominal arrival demand flow to link $i$ .
$\Delta d_i(k)$	$= d_i(k) - \bar{d}_i$ .
$\alpha_{ij}$	turning movement ratio from link $i$ to link $j$

The definition of  $u_i(k)$  assumes sufficient demand on link  $i$ . Note that the oscillations of vehicle queues in the links due to green/red communications, and the effect of offset for consecutive junctions cannot be described by the model.

According to Figure 4.1, the number of cars on link 1 is updated as follows.

$$x_1(k+1) = x_1(k) + d_1(k) + \alpha_{21}s_2g_2(k) + \alpha_{31}s_3g_3(k) - s_1g_1(k). \quad (4.1)$$

Then, by introducing the nominal amounts, and by using vectorial notations, we get :

$$\Delta x(k+1) = \Delta x(k) + B\Delta g(k) + D\Delta d(k), \quad (4.2)$$

where  $B$  and  $D$  are matrices built basing on the dynamics (4.1) written on the whole network.

Assuming that the variations of the arrival demand flows on every link inside the cycle time sum to zero, we get the following linear system :

$$\Delta x(k+1) = \Delta x(k) + B\Delta g(k), \quad (4.3)$$

Bounds for minimum green times and maximum storage capacity of links must also be considered.

The criterion is the following, where  $\lambda$  is a discount factor, and where an infinite time horizon is considered.

$$J = \min_{\Delta g} \frac{1}{2} \sum_{k=0}^{+\infty} \frac{1}{(1+\lambda)^k} (\|\Delta x(k)\|_Q^2 + \|\Delta g(k)\|_R^2), \quad (4.4)$$

where  $Q$  and  $R$  are non-negative definite, diagonal weighting matrices. The first term on (4.4) aims to minimize the risk of oversaturation and the spill-back of link queues, while the second term is used to influence the magnitude of the control.

The control bounds are treated externally of the LQ problem solving. The solution for such problems consists in solving an algebraic Riccati equation, which then leads to the following optimum feedback control where  $L$  is the gain matrix :

$$g(k) = \bar{g} - Lx(k). \quad (4.5)$$

Considering the following linear system with  $A = I$  and criterion 4.4, we have :

$$\Delta x(k+1) = A\Delta x(k) + B\Delta g(k), \quad (4.6)$$

We refer to [33, 35, 34] [16] for taking into account the discounting factor  $\frac{1}{1+\lambda}$ .  $P$  is the solution of the following algebraic Riccati equation :

$$P = \bar{A}'P\bar{A} - \bar{A}'PB(R + B'PB)^{-1}B'P\bar{A} + Q \quad (4.7)$$

where  $\bar{A} = \frac{1}{\sqrt{1+\lambda}}A$ . The matrix  $L$  is given by :

$$L = (R + B'PB)^{-1}B'PA \quad (4.8)$$

$P$  is computed by solving the Riccati equation off line with a Python library. Matrix  $L$  is computed numerically in dependence on  $P$ .



## 4.4 Semi-decentralization

The model we present here is an extension of the classical model presented above. Instead of considering only green and red time durations in a cycle time (in addition to the lost time, which we consider implicit here and for which we assign the orange colour), we also consider yellow time durations. The objective here is to reduce the red time duration. To do that, we divide this duration into two time periods : red and yellow. By that, when a stage is assigned a red or a yellow time, the antagonistic stage is assigned a green time.

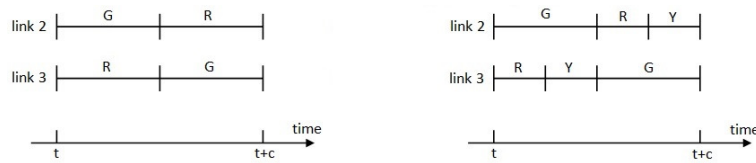


Figure 4.2: The cycle time in the classical model, and in the new model. G: green, R: red, Y: yellow.

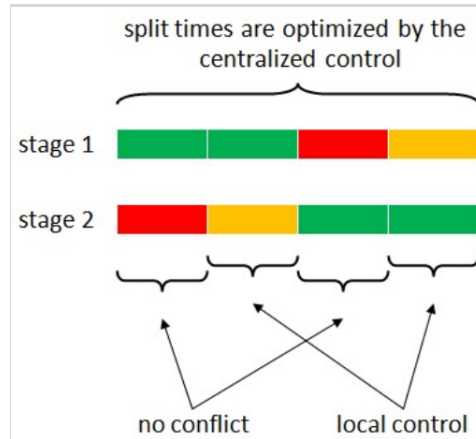


Figure 4.3: The cycle time in the new model.

We notice here that our model is an extension of the classical TUC model, because it is sufficient to set the yellow times to zero to get the classical model.

In order to explain the model, let us consider junction  $A$  of example of Figure 4.1. Only two stages can be considered here, each of them with only one stream. One stage is associated to link 2 and the other to link 3. In this case, and in the classical TUC model, at every cycle  $k$ , we only have one independent control variable on that junction, which is the green or

red duration of any of the two streams. All the other time durations are dependent variables. We consider  $g_2(k)$ : the green time duration for link 2 as the independent control variable, then the dependent variables can be easily obtained as follows :

- $r_2(k) = c - g_2(k)$  : red duration for link 2
- $g_3(k) = r_2(k)$  : green duration for link 3
- $r_3(k) = g_2(k)$  : red duration for link 3

By considering yellow time durations, we need to choose three independent control variables, among six variables. For example the following three independent control variables can be considered.

- $g_2(k)$  : green time duration for link 2
- $y_2(k)$  : yellow time duration for link 2
- $y_3(k)$  : yellow time duration for link 3

The other three dependent control variables are given as follows (see figure 2):

- $r_2(k) = c - g_2(k) - y_2(k)$  : red time duration for link 2
- $r_3(k) = g_2(k) - y_3(k)$  : red time duration for link 3
- $g_3(k) = c - g_2(k)$  : green time duration for link 3

#### 4.4.1 The dynamics

Let us consider the following additional notations.

- $q_J^{\max}$  : capacity (maximum flow) of junction  $J$ .
- $Q_{ij}(k)$  : total flow going from link  $i$  to link  $j$  during the  $k$ th cycle.
- $Q_i^{\text{out}}$  : total flow exiting from link  $i$  during the  $k$ th cycle.
- $\gamma_J$  : friction coefficient on junction  $J$ , with  $0 \leq \gamma \leq 1$ .

We write the traffic dynamics on link 1 of Figure 4.1 with the new control model.

$$x_1(k+1) = x_1(k) + d_1(k) + Q_{21}(k) + Q_{31}(k) - Q_1^{\text{out}}(k),$$

with  $Q_{21}(k)$ ,  $Q_{31}(k)$  and  $Q_1^{\text{out}}(k)$  given in (4.9)-(4.11), where we introduce a new parameter  $\gamma_J$  (for junction  $J$ ) which we call here a friction coefficient, and which expresses the bother between vehicles entering into the junction from antagonistic stages during the contention time window. For example, in (4.9), the flows of vehicles going from link 2 to link 1 during different time durations of the  $k$ th cycle are given as follows.

- During  $r_3(k) = g_2(k) - y_3(k)$ , the flow is  $\alpha_{21}s_2(g_2(k) - y_3(k))$ , as usual.
- During  $y_3(k)$ , the flow is  $\alpha_{21}s_2y_3(k)$  as usual, but multiplied by the friction coefficient  $\gamma_A$  between the streams coming from link 2 (with green time) and link 3 (with yellow time), since the local control is activated with a priority rule setting. Link 2 has priority over link 3 during this time period.
- During  $r_2(k)$ , the flow is zero.
- During  $y_2(k)$ , the stream coming from link 3 has priority over the one coming from link 2. Therefore, the whole junction capacity  $q_A^{\max}y_2(k)$  is used by the stream of link 3, and the remaining capacity  $q_A^{\max}y_2(k) - s_3y_2(k)$  is used by link 2. This flow is also multiplied by the coefficient friction  $\gamma_A$  since the two streams pass through junction  $A$  during the same time period.

$$Q_{21}(k) = \alpha_{21}s_2(g_2(k) - y_3(k)) + \gamma_A\alpha_{21}s_2y_3(k) + \gamma_A(q_A^{\max}y_2(k) - s_3y_2(k)). \quad (4.9)$$

$$Q_{31}(k) = \alpha_{31}s_3(c - g_2(k) - y_2(k)) + \gamma_A\alpha_{31}s_3y_2(k) + \gamma_A(q_A^{\max}y_3(k) - s_2y_3(k)). \quad (4.10)$$

$$Q_1^{out}(k) = s_1(g_1(k) - y_6(k)) + \gamma_Bs_1y_6(k) + \gamma_B(q_B^{\max}y_1(k) - s_6y_1(k)). \quad (4.11)$$

The dynamics (4.9)-(4.11) are still linear on the variables  $x_i, g_i$  and  $y_i$ . We notice here that the dynamics are written with only independent controls. As it has already been explained above, on junction  $A$ , for example, the independent controls are  $g_2, y_2$  and  $y_3$ . As in the classical TUC model, we consider nominal demands  $\bar{d}_i$ , nominal numbers of cars  $\bar{x}_i$  and nominal independent controls  $\bar{g}_i$  and  $\bar{y}_i$ . The choices of  $\bar{x}$  and  $\bar{g}$  can be done by the same way as in the classical TUC model. One way to choose  $\bar{y}$  is to take  $\bar{y}_i = c - \bar{g}_i$ . This is equivalent to say that the nominal red time is zero. This choice can also be dependent on the junction design. Then it is very easy to derive a linear dynamics similar to (4.3). For the criterion we take exactly the one of (4.4), written with the new (independent) control variables  $\Delta g_i$ . Again, a linear quadratic problem is obtained, and the optimal control is derived by solving a Riccati equation as in the classical TUC model.

## 4.5 Numerical example

In this section, we apply the control model presented above, on a small regular (American-like) network of four horizontal and four vertical roads, with alternated directions, as shown in Figure 4.4.

For the saturation flow values, we take the recommended ones in urban networks ( $s_i = 1800 \text{ veh./h}, \forall i$ , as shown in Table 4.2), without corrective factors; see for example [24]. To compute the optimal cycle, we consider here a fixed cycle time that we approximate to 60 seconds, using the Webster Method [147]:  $c = (1.5T + 5)/(1 - Y)$ , where  $T$  is the total lost time per cycle,  $Y$  is the junction load. The cycle time is then projected onto the interval  $[40s, 90s]$ .

#### 4.5.1 Model implementation and Simulation Tools

We used SUMO, see for example [14], and its interface TRACI [149] to simulate and implement the model. The source code has been written in Python. The main tasks were :

- build the network topology and the demand using SUMO tools and original configuration files.
- design an algorithm and the source code architecture that enable the construction of the  $B$  matrix in equation (4.3).
- implement the contention time window and the associated priority rule.
- solve a Riccati equation, and at every cycle, measure the state, and apply the control on the traffic light signals.
- analyze the simulation data outputs, including state and control vectors, by rendering graphical results.

The time contention window is implemented as follows. On a given junction, and inside such contention time window, we consider first vehicles in incoming approaches. We compute the distances from those vehicles to the junction. In order to avoid conflicts, at a given time in the time window, if the distance to the junction, of the first vehicle on the link with yellow stage, is less than a given constant distance  $m$ , and if the distance to the junction, of the first vehicle on a link with green antagonistic stage, is less than a given constant distance  $M$ , we slow down the first vehicle on the link with yellow stage.

In general, the vehicles moving on an approach with a green stage (priority approach) pass through the junctions without checking for the antagonistic approaches. However, the vehicles moving on the approaches with yellow stages slow down at a distance  $m$  to the junction, to check if there is any vehicle coming from an antagonistic approach with green stage.

In the numerical example we consider in this article, we chose  $m = 15$  meters and  $M = 50$  meters. Our choice takes into account the reaction time of the drivers in SUMO, and also the simulation step length.

We plan to implement this conflict management using a communication simulator, for example the Network Simulator [106].

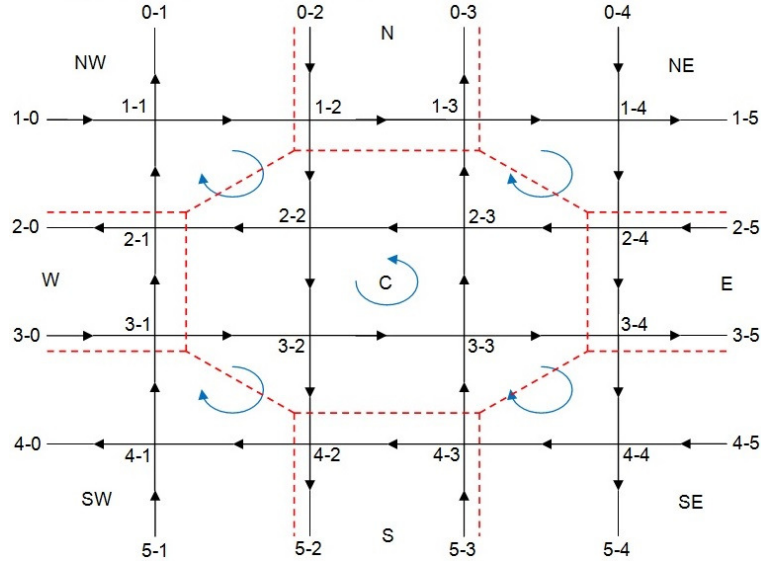


Figure 4.4: Regular network example.

#### 4.5.2 Network configuration

We discuss here, the configuration of the network of Figure 4.4. In this network, circuits are formed. We distinguish two types of circuits. The central circuit in which vehicles turn in the anticlockwise direction, and the other four circuits in which vehicles move in the clockwise direction. As already shown in [51, 48], the car-densities on the circuits of links are determinant in the stage transition of a vehicular network. Indeed, if a circuit is full of vehicles, then a deadlock occurs and spreads on the network.

In the network we consider here, the central circuit (which we call here the main circuit) is critical compared to the other four circuits, (which we call here the secondary circuits). Indeed, the secondary circuits have exits that are not constrained by any output supply, and they are closer to the borders.

In case of congestion, we need to clear out vehicles from the main circuit in order to improve the traffic, so that the number of vehicles we take out is bigger than the one we take into the circuit. Hence, for that circuit, the controller needs to favour the vehicles coming from the left side at the level of the four junctions around the main circuit. For example, if we take symmetric turn ratios, half of vehicles leaving the approaches are likely to leave the circuits, while the other half of vehicles are likely to remain in the

circuit. However, when the way is given to the vehicles coming from the right (with respect to the junction), half of those vehicles are likely to enter to the circuit, while the other half is likely to not enter to the circuit. For the secondary circuits, in case of congestion, the control shall favour vehicles coming from the right side links at the level of the junctions associated to those circuits, in order to clear them out.

The four junctions of the main circuit are shared with other secondary circuits. We think that the control needs to foster the evacuation of the main circuit with respect to the secondary circuits. Therefore, the control should favour the vehicles coming from the left side approaches to the main circuit.

### 4.5.3 Preliminary results

We present in this section the preliminary results we obtained. For the traffic demand, we took the scenario of Table 4.1 with symmetric turning ratios as input. The origins and destinations are predefined. In this scenario, we have some traffic demand inside the network. This permits us to attain saturated and congested stages. In the other side, the traffic demand from and towards the central zone is low comparing to that from and towards the boundary zones. This choice makes the states of the traffic controllable in the central zone of the network.

Table 4.1: The traffic demand.

	Central zone	Other zones
Central zone	0	40 ( <i>veh / h</i> )
Other zones	40 ( <i>veh / h</i> )	250 ( <i>veh / h</i> )

The other parameters are given in Table 4.2, where

- $r$  is a positive scalar such that  $Q = I$  and  $R = rI$ , with  $I$  the associated identity matrix,
- $g_{i-min}$  is the minimal green time duration on link  $i$ ,
- $l_i$  is the length of link  $i$ .

Table 4.2: The values of other parameters.

$r$	$\lambda$	$\bar{x}_i$	$s_i$	$\bar{g}_i$	$g_{i-min}$	$c$	$l_i$	$\alpha_{ij}$
0.5	0.1	10.5 veh	1800 veh/h	30 s	4 s	60 s	300 m	0.5

In Figure 4.5, we give the state of the traffic at the final time of simulation. The evolution over time of the running vehicles in the network is

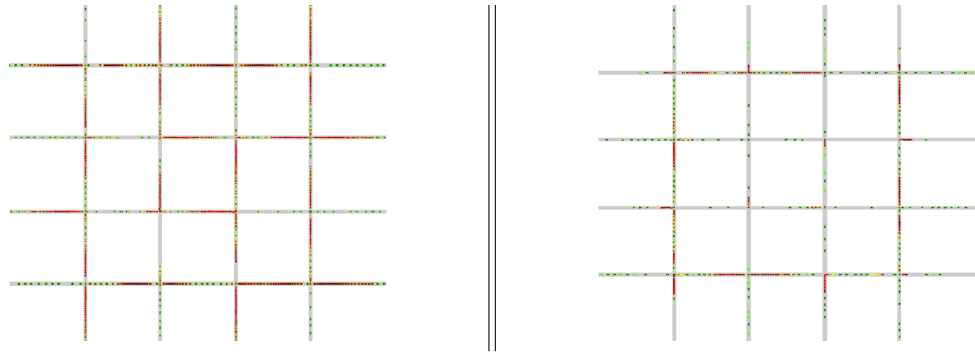


Figure 4.5: The state of the traffic at the end of simulation. The colours of vehicles correspond to their speed (green: high speed, red: low speed). On the left side: Centralized TUC. On the right side: semi-centralized TUC.

given on Figure 4.6, where we compare the classical TUC control with our semi-decentralized control by varying the value of the friction parameter  $\gamma$  in  $\{0.3, 0.5, 0.7\}$ . We see that with our semi-decentralized control, the car-density is limited in order to optimize the capacity of the network. The best result is obtained with  $\gamma = 0.3$ .

In Figure 4.6, we also compare the two controls in terms of the cumulated ended cars through the time, and in terms of the average travel time of cars in the network. We see clearly that our control improves the whole capacity of the network. Indeed a congestion appeared at a time around 1000 seconds. We observe that as long as the simulation runs, the two controls clear the congestion, but the semi-decentralized control do it very rapidly compared to the centralized one. We see clearly that the difference between the number of running vehicles decreases over time, but, even at the final time of simulation (which is 6 hours here), this difference is still important. Figure 4.5 tells clearly that the state of the traffic with the two controls is different (fluid with the decentralized control, and saturated with the centralized one). These results are confirmed by Figure 4.6, where we compare the running and the ended vehicles, as well as the average travel time of the cars through the network.

In Figure 4.7, we give the results of simulation for the semi-decentralized control. We show on the first row the time-average number of vehicles in the circuits of the network. On the second (resp. third) row of that figure, we show the control (in terms of durations of the green, yellow and red times) for the approaches coming from the left side (resp. right side) of the circuit junctions. The left side column of the figure corresponds to the main circuit (the circuit of the central zone), while the right side column corresponds to the secondary circuits (the circuits on the boundary of the network).

We observe on the first row of Figure 4.7 that the main circuit is more

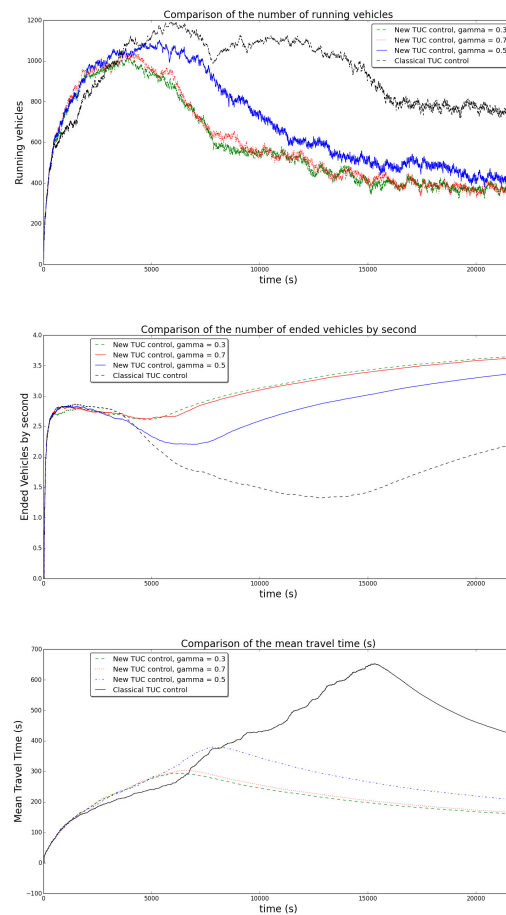


Figure 4.6: Comparison of the classical TUC with the semi-decentralized TUC in terms of the number of running vehicles on the network, the flow of ended vehicles, and the average travel time through the network, respectively, in function of time.



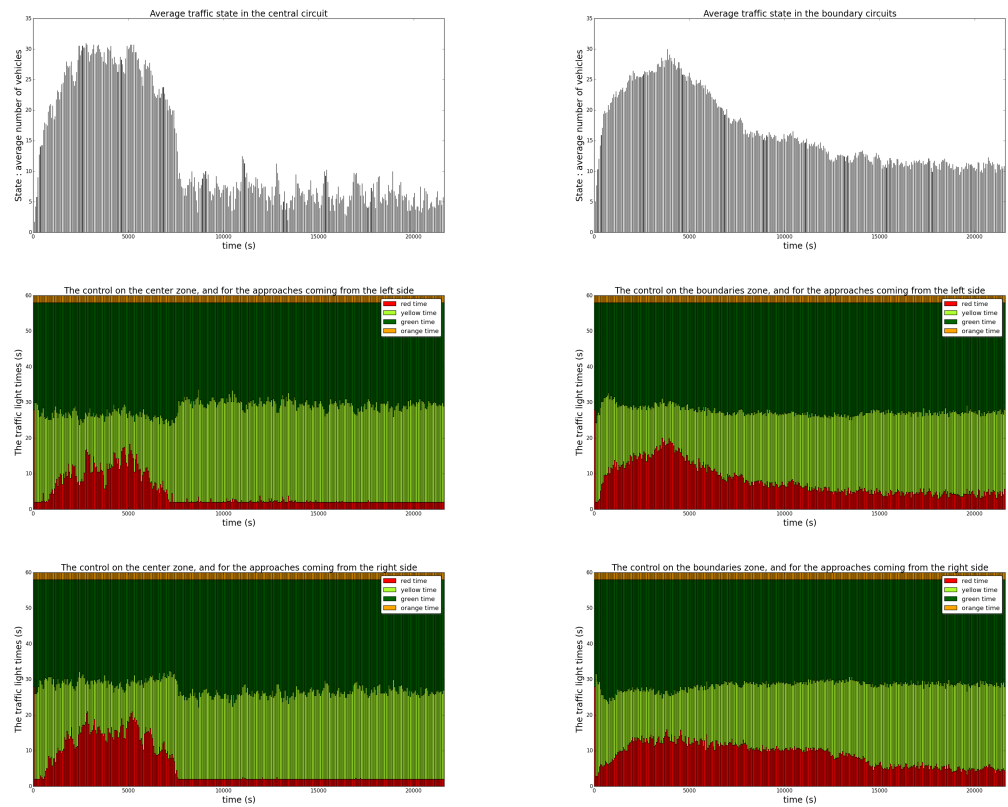


Figure 4.7: Semi-decentralized TUC. The control in terms of the traffic light times into the cycle time, through the simulation time, on different zones (center and boundaries), and for approaches coming from left and right sides.

cleared out than the secondary circuits. This observation confirms our intuition given above. We see in the second and third rows of Figure 4.7 that the control frees the approaches coming from the left side of the junctions' main circuit and limits the flow on the antagonistic approaches of the same circuit, while it does the opposite for the secondary circuits.

Figure 4.7 shows another important result, which is that the yellow time is almost fully used (i.e. the red time is almost zero) in the case of free traffic flow, while the red time appears with important values in case of congestion. This result is very important because it confirms the importance that the activation as well as the duration of the local control (the contention time window with yellow times) are both controlled by the centralized control, which optimizes them in function of the state of the traffic in the network.

## 4.6 Preliminary conclusions

We presented in this article a TUC-based approach for the control of urban traffic. By defining a time contention window inside the time cycle, we introduced a little of decentralization of the control. We have implemented and simulated the new control on a small American-like network. The traffic has been simulated using the Simulation of Urban MObility tool while the control has been implemented with Python. We are aware that we need more investigations in order to validate our assertions. For that we will improve the implementation of our control by better managing the contention time window, in particular using communication network simulators. On this small network, we showed that our approach is effective in terms of many parameters including the total network capacity as well as the average travel time. Another important result we obtained is the confirmation that the centralized control optimizes the activation as well as the duration of the decentralized control (the contention time window) in function of the state of the traffic in the network.



## Chapter 5

# A vehicle-to-infrastructure communication based algorithm for urban traffic control

# A vehicle-to-infrastructure communication based algorithm for urban traffic control

## 5.1 Abstract

We present in this paper a new algorithm for urban traffic light control with mixed traffic (communicating and non communicating vehicles) and mixed infrastructure (equipped and unequipped junctions). We call equipped junction here a junction with a traffic light signal (TLS) controlled by a road side unit (RSU). On such a junction, the RSU manifests its connectedness to equipped vehicles by broadcasting its communication address and geographical coordinates. The RSU builds a map of connected vehicles approaching and leaving the junction. The algorithm allows the RSU to select a traffic phase, based on the built map. The selected traffic phase is applied by the TLS; and both equipped and unequipped vehicles must respect it. The traffic management is in feedback on the traffic demand of communicating vehicles. We simulated the vehicular traffic as well as the communications. The two simulations are combined in a closed loop with visualization and monitoring interfaces. Several indicators on vehicular traffic (mean travel time, ended vehicles) and IEEE 802.11p communication performances (end-to-end delay, throughput) are derived and illustrated in three dimension maps. We then extended the traffic control to an urban road network where we also varied the number of equipped junctions. Other indicators are shown for road traffic performances in the road network case, where high gains are experienced in the simulation results.

## 5.2 Introduction

### 5.2.1 Introduction

Penetration rate of communicating vehicles is expected to increase in the next years. Compared to in-road detectors and video sensors, a wireless road side unit (RSU) can collect more detailed vehicle data such as location, speed and acceleration rate, more than once a second, on some hundred meters range and probably at lower cost [56]. This new amount of high resolution data provided by V2X communication enables new traffic signal controls. We present a new reactive algorithm based on V2I communications using WAVE/IEEE 802.11p protocol. Simulation for road traffic and communication networking has been conducted using VEINS framework [131]. This simulation framework led to a performance study of both road traffic and communication protocols. We show that the gain in road traffic performance is significant most of the time, especially in the case of a high penetration rate for vehicles and junctions.

### 5.2.2 State of the art

In the field of traffic signal control based on vehicular communication, several approaches have been developed in the few last years [62] : over-saturation algorithms which tend to avoid blockages by using V2I communication, gap-out algorithms which terminate the phase green if no vehicle is detected during a gap-out time, and platoon based algorithms which use vehicle clustering to provide acyclic timing plans. Some other approaches tend to minimize cumulative delays.

In [6], a dynamic gap-out algorithm has been presented. Total vehicular delays are minimized and the optimization determines “phase sequence, phase green times, and gap-out times (both dynamic and regular gap-outs).” In [88], a reactive control based on VANET communication is detailed. Different weights are assigned to vehicles depending on their distance to the junction. A timing plan is then computed and applied using these weights.

Some papers have finely evaluated performances of WAVE/IEEE 802.11.p protocols [146], some of them comparing pros and cons of WAVE and alternatives such as LTE [69]. Coupling road traffic and communication simulators have recently been achieved in VEINS [131]. We also report ITETRIS [87] and VSimRTI [126] that declare successful coupled simulation, even if we haven’t been in measure to evaluate these last two softwares in detail. However, in the case of road traffic control applications, we did not see communication performance studies with specialized communication simulators.

### 5.2.3 Paper organization

We aimed in this paper to propose a new traffic light control algorithm, based on V2I communication and evaluated with a fine grained and extended simulation tool, VEINS [131]. We modified VEINS in order to include TCP/IP support over IEEE 802.11p. We present some performance indicators of the WAVE protocol stack in the scenario of this new kind of road traffic control. This paper is organized in four parts. In part 5.2, it has been provided the global context and state of the art in the field of connected traffic light signal control. In part 5.3, a new algorithm is presented for road traffic control. Then simulation scenarios and results are shown for one junction and for a small American like road network in part 5.4. In part 5.5 we open perspectives to future works.

## 5.3 Connected traffic light signal control

### 5.3.1 Algorithm description

In this section we describe a new local control algorithm. This control makes some *hypothesis* on vehicles and infrastructure and is composed of

the following subtasks : building *a map*, *electing* a vehicle and *actuating* the TLS.

### 5.3.1.1 Assumptions

We use the terminology described in [56]. We assume that some junctions of the road network are equipped with traffic light signals (TLS) with communication capabilities. In our case the communication protocol is IEEE 802.11p coupled with the Internet Protocol version 4 (IPv4) and the Transmission Control Protocol (TCP). TCP adds transport services to IEEE 802.11p, such as a reliable and ordered delivery of byte streams [133]. It is used in conjunction with IP which provides network routing services. Hence, we suppose that some TLS are able to communicate with the TCP/IP protocols over IEEE 802.11p and we consider it as an Intersection Agent (IA). Similarly, we suppose that some cars are equipped with the same communication capabilities and are also able to localize themselves, for example with GPS modules which provide in addition global time synchronization. We call them equipped vehicles or Vehicle Agents (VA) [56].

### 5.3.1.2 Dynamic Maps

With such capabilities, the IA can build a map of the connected vehicles coming and leaving the junction. Similarly each vehicle agent (VA) builds a map of the IAs approaching or leaving it in its communication range. To achieve this, we designed and programmed a map module in OMNET++ [142]. For the map of the vehicles, the coordinates system is relative to the earth. Instead, for the map of the IAs, the coordinates system is relative to the concerned vehicle position. This change of coordinates enables the use of the same module for IAs map and vehicles map. The maps are *dynamic*; they are updated periodically each time a message is received for the IA, and triggered on timer for the vehicles. Map's data that are older than a given time, named here *map\_module\_length*, are cleared.

An IA signals itself by broadcasting its IP address and coordinates via UDP protocol. Once the IA announced, the vehicles equipped with UDP client build a local map of all the IAs in their communication range. These vehicles then elect the closest IA approaching. So, the vehicles know their relative positions to the closest junction approaching, without need for communication with the IA anymore. Once close to the elected junction, the vehicles open a TCP connection with the IA.

After the TCP connection with the elected IA is established, all vehicles approaching the junction send periodically their position to the IA, each message being timestamped, with a time period named here *position\_send\_interval*. These vehicles data received by the IA enable the build of a map indexed by a unique vehicle identifier, described in Table 5.1. Among the different

fields of the map, we notice the state of the vehicle (approaching or leaving the junction) which is computed using the positions of vehicles and TLS.

Table 5.1: The IA and vehicle map module

unique vehicle identifier	
c	the IA_to_car_TCP_connection identifier
T	the trajectory which is an ordered map of couples (time, coordinates)
lst	the last time the vehicle data has been received by the IA
fst	the first time the vehicle data has been received by the IA
r	the radius is the distance the car is to the approaching junction
$\cos \theta$ $\sin \theta$	the car position is defined by its radius to the TLS and the angle this radius is from the (x) axis
s	the state of the car whether the car is coming or leaving the junction

### 5.3.1.3 Election

We say the road state and a given junction map are synchronized when all vehicles in the junction map have been detected in the last *map\_module\_timeout* seconds. Periodically, every *election\_interval* time and when the junction map is synchronized, the IA computes the lead vehicles on the approaching edges. In our case, we suppose that the junctions have only two incoming edges, each edge having one lane. So there can be two lead vehicles maximum in our case. See Fig. 5.1. The two incoming edge *priorities* alternate every *cycle\_duration/2*, where *cycle\_duration* is the duration of the TLS periodic program. A vehicle among lead vehicles from incoming edges is elected with the Algorithm 1. If the Algorithm 1 runs successfully, the IA sends a message to the elected vehicle. Otherwise the RSU will try to elect a vehicle after a *election\_interval* time.

### 5.3.1.4 Action

**The elected vehicle has now the power to set the TLS to a favorable state**, which is green light for the edge on which it is moving and red light for other edges. To do this, the elected vehicle sends a message to the IA with its established TCP connection. We set a minimum and a maximum duration for a given TLS state : *min\_state\_duration* and *max\_state\_duration*. If no



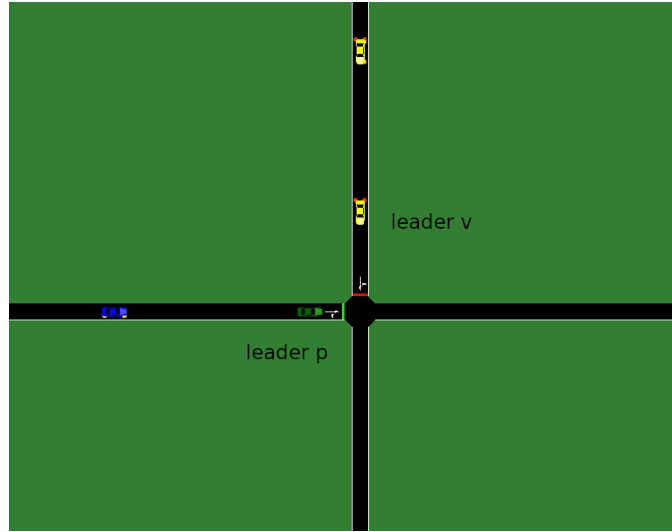


Figure 5.1: The lead vehicles

---

**Algorithm 1:** Vehicle Election
 

---

1 **function** *Elect* ( $p, v, d_p, d_v, d_{min}, \alpha$ )

**Input** :

- $p$  is the identifier of the lead vehicle on the prioritized edge, and it is *None* if no vehicle is detected on the prioritized edge
- $v$  is the identifier of the lead vehicle on the non prioritized edge, and it is *None* if no vehicle is detected on the non prioritized edge
- $d_p$  represents the distance  $p$  is to the junction, in case  $p \neq None$ ,
- $d_v$  represents the distance  $v$  is to the junction, in case  $v \neq None$ ,
- $d_{min} > 0$  is the minimum distance to consider a vehicle close to the junction,
- $\alpha > 1$  is a coefficient to ponderate the minimum distance.

**Output:**  $p$  or  $v$ .

2 **if** ( $p \neq None$  and  $v \neq None$  and  $d_p > \alpha d_{min}$  and  $d_v < d_{min}$ ) or  
 ( $p == None$  and  $v \neq None$  and  $d_v < d_{min}$ ) **then**

3 |  $electd = v$ ;

4 **else**

5 |  $electd = p$ ;

6 **return**  $electd$ ;

---

state switch has happened during *max\_state\_duration* time, the state of the TLS is automatically changed. Similarly, the state of the TLS must remain the same for at least *min\_state\_duration*. With the *min\_state\_duration* we ensure stability. With the *max\_state\_duration* we ensure dynamics of the states and avoid blockages of the TLS. As we set *max\_state\_duration* = *cycle\_duration*/2, if no vehicle is connected near a junction, then the associated TLS will follow an open loop cyclic program with *cycle\_duration* period. Once the connected vehicles know they are leaving the junction (with GPS and local map but not with communication means), they disconnect after they reach a given distance away from the junction.

The process starts again for the next junction and so on.

### 5.3.2 Properties of the algorithm

#### 5.3.2.1 Property 1

The local control is safe because the control is done by means of a TLS which never gives green light simultaneously to antagonistic phases.

#### 5.3.2.2 Property 2

It is not necessary for a vehicle to be equipped to pass the junction.

#### 5.3.2.3 Property 3

As the control tends to minimize delays for equipped vehicles, communicating equipments of vehicles are encouraged. Indeed, even if equipped vehicles are favoured, all vehicles benefit from the control. As the control presents gains for the road traffic, communicating equipments of junctions are encouraged.

#### 5.3.2.4 Property 4

When no vehicle is equipped near an equipped junction, the *max\_state\_duration* for a state induces that the TLS runs half time red and half time green light. It is equivalent to a simple open loop cyclic TLS program.

### 5.3.3 Implementation

We used VEINS Framework [131] which includes SUMO [86] as microscopic traffic simulator and OMNET++ [142] as communication network simulator. We modified and extended VEINS Framework in order to get TCP/IP support over IEEE 802.11p. To do this, “inet” models and Veins framework have been integrated and connected together.

Some application modules have been written : map, car, road side unit, TCP client and server, UDP client and server applications, which implement

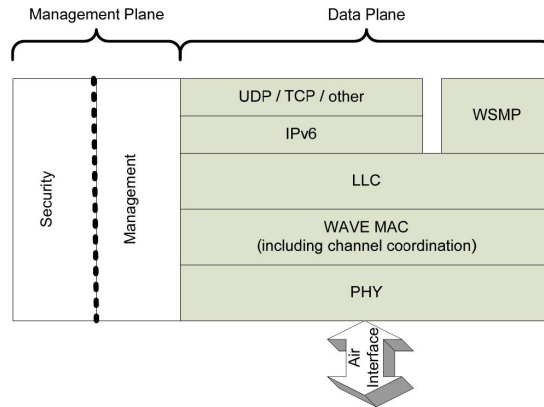


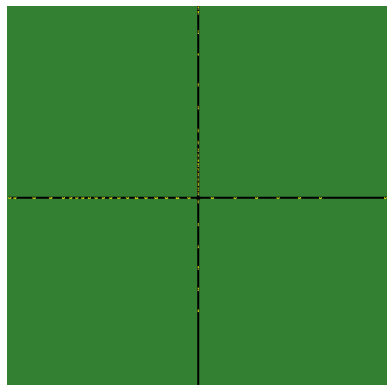
Figure 5.2: The IEEE WAVE protocols stack [3]

the algorithm described above. Commands to control the TLS states have been added. The MAC1609 module of VEINS framework module has been modified to connect TCP/IP to IEEE 802.11p layers.

## 5.4 Simulation results

We simulated a few runs with different seeds, each run being reproducible. We present statistical results for the road network case with 20 different simulations and preliminary results (one typical run) for the one junction case.

### 5.4.1 One Junction with connected TLS



#### 5.4.1.1 Scenario for one junction

We used the following road network, composed of one junction, with one lane incoming edges of 300 m length. We varied the demand which is the number of vehicles (uniformly inserted in time) per lane per hour. For each of this traffic demand, we varied the ratio of equipped vehicles with on board unit (OBU).

The total simulation time is 600 s. For the communication and the control algorithm, the main parameters are described in Table 5.2.

Table 5.2: Main parameters for the communication and road traffic control. Other parameters are VEINS defaults ones.

Parameter name	Parameter value
vehicle TCP <i>position_send_interval</i>	500 ms
UDP broadcasting interval	500 ms
IA <i>election_interval</i>	500 ms
<i>cycle_duration</i>	90 s
<i>max_state_duration</i>	45 s
<i>min_state_duration</i>	8 s
<i>map_module_timeout</i>	2 s
<i>map_module_length</i>	5 s
$d_{min}$	100 m
$\alpha$	2
MAC 1609 use service channel	true
MAC 1609 bitrate	27 Mbps
MAC 1609 carrier frequency	$5.890 \times 10^9$ Hz
transmit power	1 mW
application message payload	30 bytes
transceiver sensitivity	-89 dBm

#### 5.4.1.2 Simulation measurements

**For the communication** we have measured the mean TCP end-to-end delay, TCP throughput on RSU (Road Side Unit) and the amount of TCP application data sent divided by the total simulation time. We define the simulation indicator mean TCP end-to-end delay as the sum of all packet delays divided by the number of packets exchanged. The throughput on RSU is the sum of TCP application packet (successfully received) sizes divided by the simulation time. A given number of communicating hosts may be the result of different combinations of a demand multiplied by a ratio of equipped vehicles. For example, 100% equipped vehicles of 100 vehicles in total, gives the same number as 10% equipped vehicles of a total of 1000 vehicles.

In Figure 5.3 we can see that the amount of TCP application data sent by the nodes increases as the mean vehicle speed decreases. We assume that as the mean vehicle speed is low, the communicating vehicles remain connected longer, and then they send more messages.

In Figure 5.4, we see that the mean TCP end-to-end delay can be as high as 0.8s when vehicle speed is low (about  $8m/s$ ). We know from Figure 5.3 that there are more data sent by the nodes when mean speed is low. We suppose that as there are more messages sent in case of low speeds, the mean

TCP end-to-end delay will be higher.

The order of magnitudes of end-to-end delay is similar to the ones exposed in [69]. Clearly, in Figure 5.5, the throughput on RSU is increasing linearly with the number of communicating vehicles.

**TCP\_application\_data\_sent divided by the total simulation\_time (bit/s)**

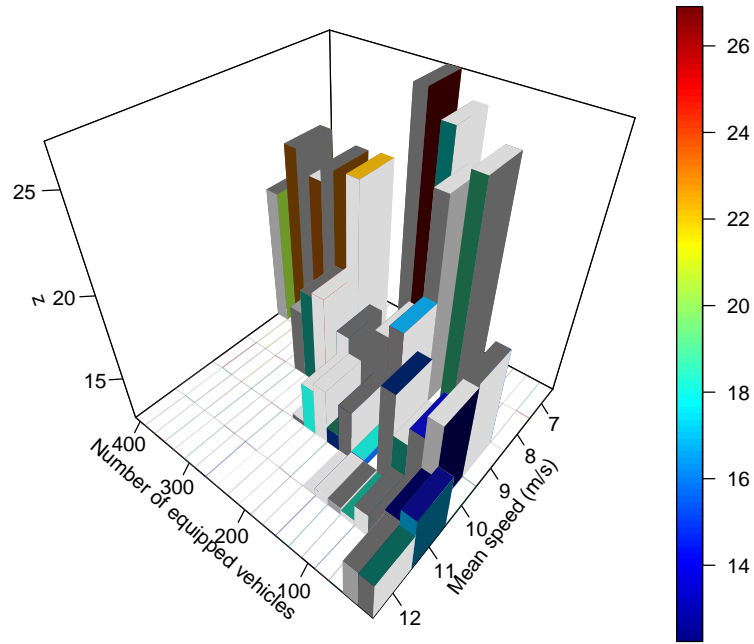


Figure 5.3: Amount of TCP application data sent divided by the total simulation time (bit/s)

**For the road traffic**, we measured the ratio of the ended vehicles by the inserted vehicles, and the mean travel time. We observe on Figure 5.6 that the ratio of the ended vehicles by the inserted vehicles increases when the demand decreases and the ratio of equipped vehicles increases. For a given ratio of equipped vehicles, as the road traffic demand increases, the ratio of the ended vehicles by the inserted vehicles decreases. For a given demand, this ratio increases with the number of equipped vehicles.

In Figure 5.7, for a given demand, the mean total travel time (for both equipped and non equipped vehicles) decreases as the number of equipped vehicles increases. This should encourage the spreading of vehicle communication capabilities. For a fixed demand, the difference in travel time between

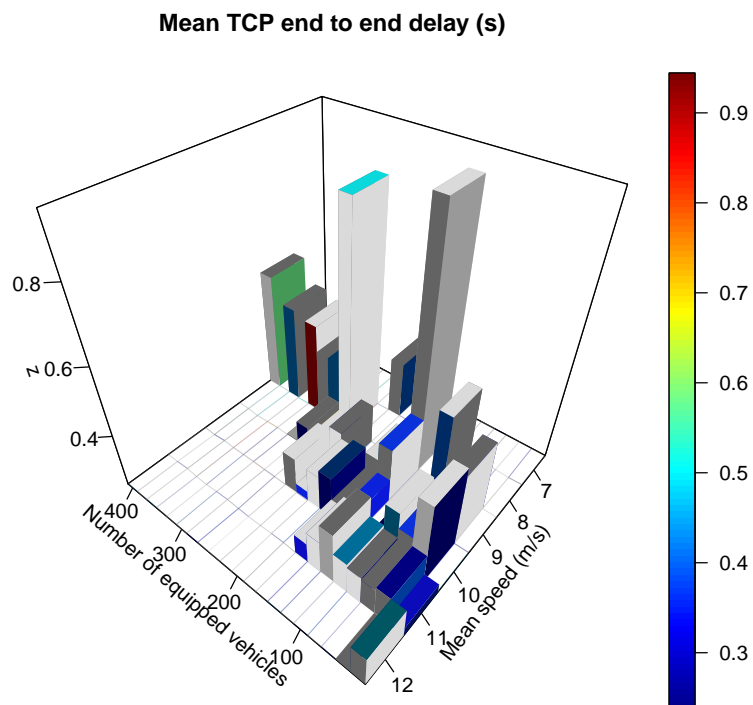


Figure 5.4: Mean TCP end-to-end delay (s)

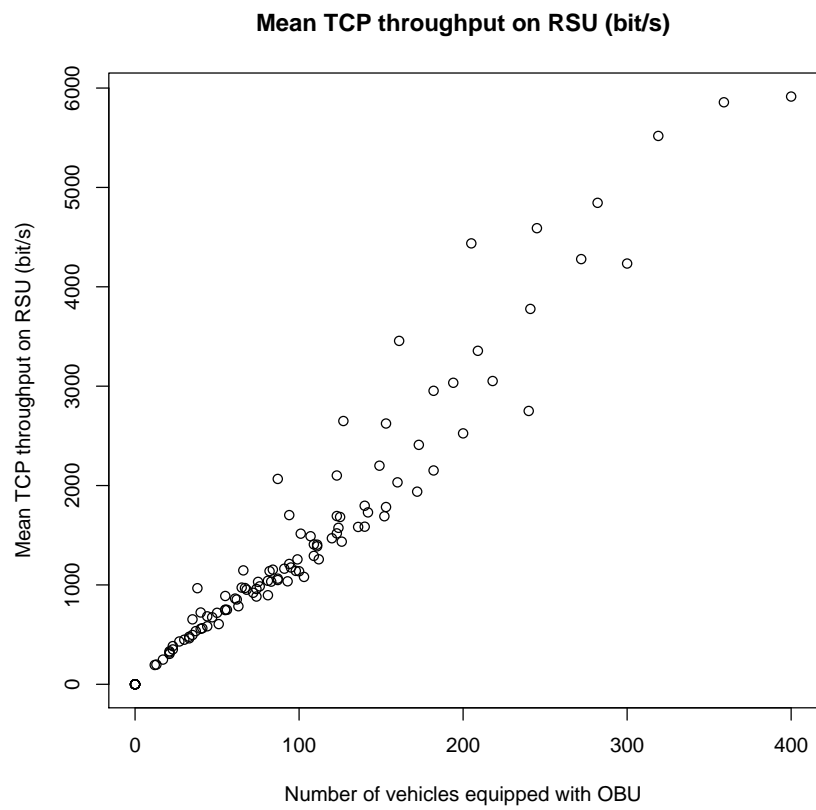


Figure 5.5: TCP application throughput on RSU (bit/s)

the worst and the best cases for a total distance of 600 m, may be as high as 20 seconds.

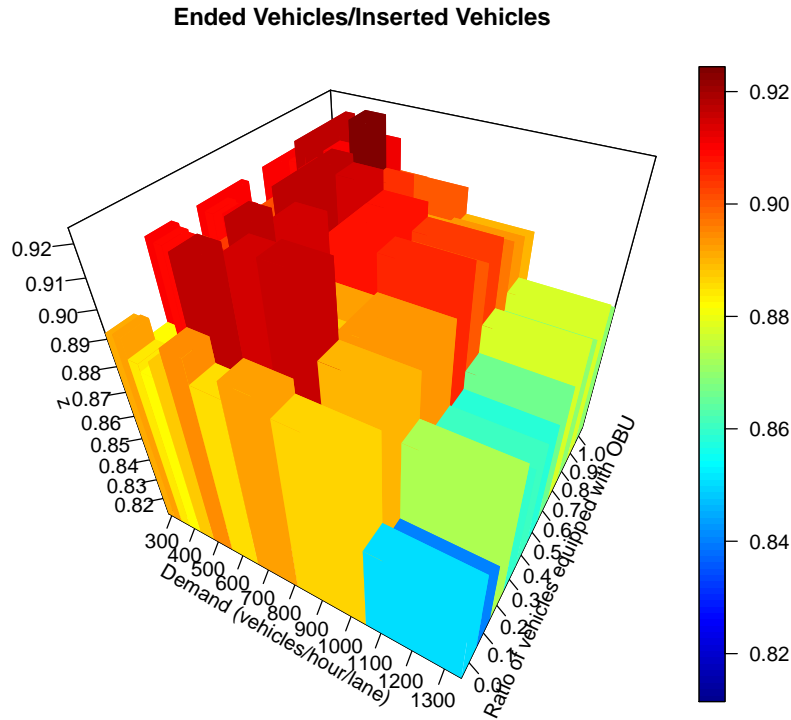


Figure 5.6: Ratio of ended vehicles by inserted vehicles at 600s simulation time

**For the actuator**, we define the *mean action interval* as the mean time interval between two consecutive changes of traffic lights state. We observe in Figure 5.8 that the traffic light state is stable for high traffic demand combined with a high equipped vehicle penetration rate. This minimizes the total yellow time of the traffic light, and then maximizes the junction capacity. For a low demand combined with a high equipped vehicle penetration rate, the control is more reactive.

## 5.4.2 Traffic control at the road network level

### 5.4.2.1 Scenario for the road network

We consider here an American like network with 16 junctions and 40 edges of length 500 m each. This is the same network considered in [53], where



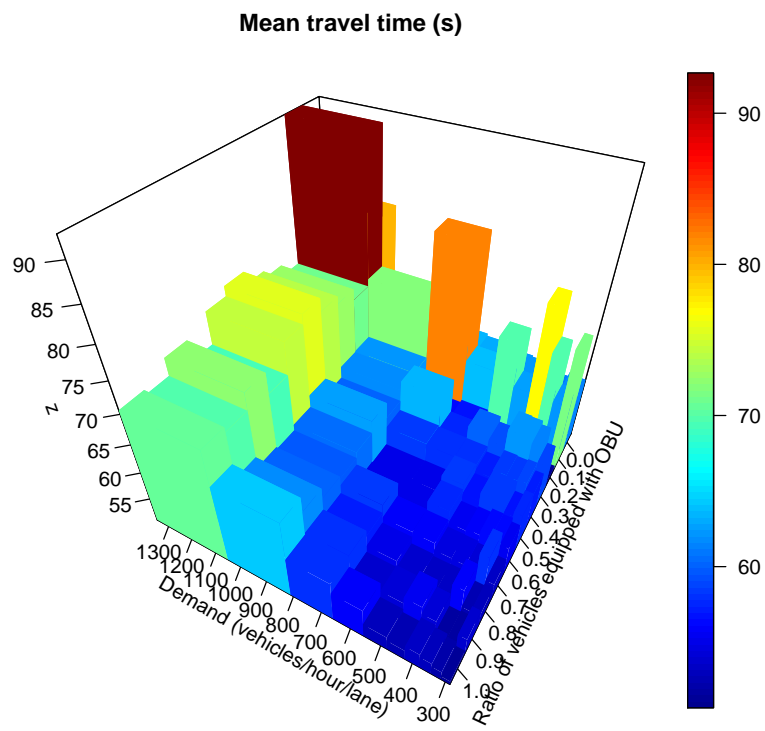


Figure 5.7: Mean Travel Time for equipped vehicles (s)

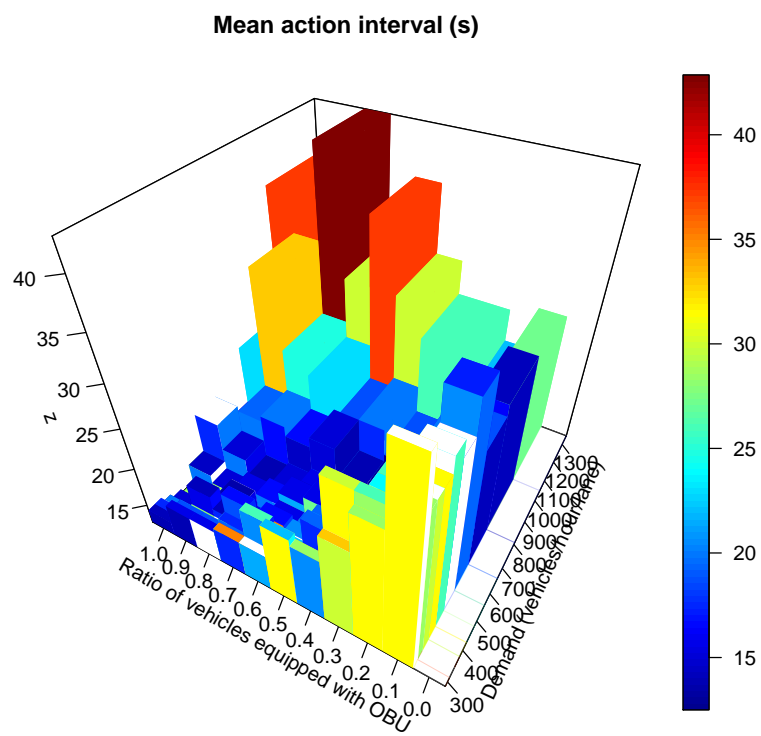


Figure 5.8: Mean action interval (s) for a cycle duration of 90 s.

centralized and decentralized road traffic controls have been combined. Each edge has one lane. We varied the number of equipped junctions : 25%, 50%, 100%; and the penetration rate of equipped vehicles : 20%, 50%, 80%. We defined nine zones in the network. The simulated time is 1800 s and the communication parameters are the same as in the “one junction” scenario.

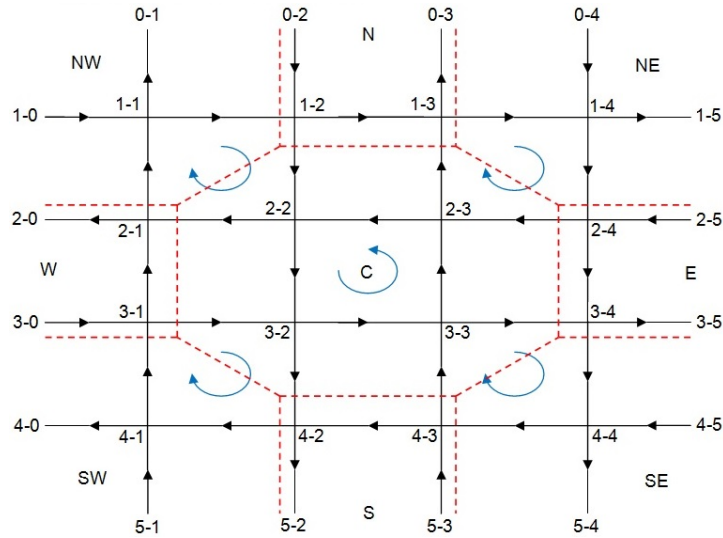


Figure 5.9: Regular network example.

Table 5.3: The traffic demand for the first 900 s.

Origins \ Destinations	Center zone	Each other zone
	Center zone	0
Each other zone	15 ( <i>veh</i> )	15 ( <i>veh</i> )

Table 5.4: The traffic demand for the last 900 s.

Origins \ Destinations	Center zone	Each other zone
	Center zone	0
Each other zone	20 ( <i>veh</i> )	20 ( <i>veh</i> )

We used SUMO “origin and destination edges instead of a complete list of edges. In this case the simulation performs fastest-path routing based on the traffic conditions found in the network at the time of departure/flow begin.” [1]. The road traffic demand is given in Tables 5.3 and 5.4.

### 5.4.2.2 Simulation results

Table 5.5: Ended vehicles in a scenario with the traffic demand of tables III and IV. Simulated time = 1800 s. Mean and standard deviation for 20 simulation runs.

Ended Equipped junctions	Penetration rate			
	0%	20%	50%	80%
25%	1373±19 (0±0)%	1470±33 (+7.1±2.5)%	1507±18 (+9.8±2.4)%	1484±19 (+8.1±1.9)%
50%	1373 ±19 (0±0) %	1499±49 (+9.2±3.4)%	1583±19 (+15.3±1.9)%	1571±20 (+14.5±2.3) %
100%	1373±19 (0±0)%	1281±151 (-6.7±11.2)%	1805±49 (+31.5±4.1)%	1877±29 (+36.7±2.8)%

Table 5.6: Running vehicles in a scenario with the traffic demand of tables III and IV. Simulated time = 1800 s. Mean and standard deviation for 20 simulation runs.

Running Equipped junctions	Penetration rate			
	0%	20%	50%	80%
25%	954±16 (0±0)%	842±32 (-11.8±3.5)%	817±18 (-14.4±2.7)%	848±21 (-11.2±2.4)%
50%	954±16 (0±0)%	835±36 (-12.4±3.8)%	764±17 (-19.9±2.1)%	778±21 (-18.4±2.5)%
100%	954±16 (0±0)%	962±80 (+0.9±8.7%)	583±43 (-38.9±4.7)%	517±27 (-45.8±2.9)%

Table 5.7: Mean Travel Time (s) in a scenario with the traffic demand of tables III and IV. Simulated time = 1800 s. Mean and standard deviation for 20 simulation runs.

MTT(s) Equipped junctions	Penetration rate			
	0%	20%	50%	80%
25%	413.9±1.8 (0±0)%	381.9±4.3 (-7.7±1.1)%	376.0±3.1 (-9.2±0.8)%	380.0±3.4 (-8.2±0.9)%
50%	413.9±1.8 (0±0)%	381.8±15.2 (-7.8±3.8)%	355.6±5.0 (-14.1±1.5)%	354.9±4.5 (-14.2±1.2)%
100%	413.9±1.8 (0±0)%	399.6±30.6 (-3.4±7.4)%	302.0±6.9 (-27.0±1.7)%	281.2±4.6 (-32.1±1.2)%

We present the simulation results in Tables 5.5 5.6 5.7. We compare our algorithm with an open loop fixed cycle TLS program, where the same cycle

time is considered in both cases. Mean and standard deviations are reported in each cell of Tables 5.5 5.6 5.7. In brackets, the mean and standard deviations are reported in percentage. This open loop cyclic control can also be achieved with zero communicating vehicles in our algorithm. The gain of our algorithm in terms of ended vehicles can be very high, as much as 30%. The gain in running vehicles can be as high as 40% and the gain in mean travel time can reach 30%.

We observe that when the penetration rate is 20% and when all junctions are equipped, the algorithm is not efficient. We suppose that as there are less vehicles communicating, less vehicles manage to connect in time (before having passed the junction). It then could be possible that the TLS map is not a good image of the real traffic. This could explain why the cycles of the TLS are not globally adequate. When all junctions are equipped, this can even be globally disturbing because of a lack of centralized coordination. By combining accurate estimation of the road traffic and V2I control, this phenomenon could be analyzed more deeply and potentially cleared.

## 5.5 Conclusions

We presented a new V2I based TLS control algorithm, with its design, implementation and performance study for communication and road traffic. Compared to an open loop cyclic TLS program, we showed that the presented algorithm features high gains in most of the configurations. However, in few cases, where low penetration rate for vehicles is combined with high ratio of equipped junctions, it seems that the algorithm is not efficient and produces some losses. A hypothesis for this phenomenon has been proposed. Future works could benefit from a mix of microscopic and macroscopic road traffic controls, based on vehicular communication networking.

**Part IV**

**Conclusion**



## 5.6 Conclusion

The stakes of road traffic range the economical, ecological, public health, and leisure domains. Intelligent transportation Systems (ITS) use the wireless connectivity of communicating and automated vehicles in order to improve road traffic conditions relatively to the stakes for the society and the individual. In this thesis, we have addressed the topic of estimation and regulation of urban road traffic. A literature review covering the main approaches used for road traffic estimation and regulation has been given. Based on the state of the art of this topic, we have proposed some models and algorithms for the estimation and control of urban road traffic in presence of communicating and/or automated vehicles. The thesis contributions on the estimation of urban road traffic are mainly based on the probability theory, while the thesis contributions on road traffic regulation are based on the control theory. To counterbalance the lack of experimental data for discussing the assumptions and models, simulation studies were performed by implementing the models in various communication and road traffic simulation frameworks.

Concerning the estimation of road traffic, two articles addressing the case of multi lanes roads at an urban controlled junction have been proposed. These papers are built upon an existing method which uses probabilistic models feeded by information provided by probe vehicles. This existing model is addressing the roads composed of one lane. However, in the case of unbalanced traffic demand over the multiple lanes, this existing model is inaccurate. This argument justifies the proposed work on the estimation of road traffic. We have considered a multi lanes road where vehicles arrive, form queues at the traffic light and then pass the junction. The purpose was to estimate the queue lengths on every lane, as well as other parameters such as the penetration ratio of equipped vehicles and the arrival rate of vehicles. For the estimation of road traffic, it is assumed low or moderate traffic demand, which ensures that the vehicle arrivals can be modeled as a Poisson process. It is also assumed that the queues at the traffic light are cleared at the beginning of the red time. Because of the inaccuracy of GPS systems, the probe vehicles cannot be assigned to a lane with confidence. Therefore, it was necessary to extend the existing method, which was the purpose of the thesis contribution. By letting the method robust to the GPS localization accuracy, the method adds new benefits to the existing one. As real traffic assignment data is not available, simulation studies have been performed. The simulation studies presented in this document use a discrete event simulator framework, sometimes coupled with a microscopic road traffic simulator. The coupling of communication and road traffic simulators (the VEINS framework) is enhancing a modern way to get experimental results on Vehicular Ad Hoc Networks. The results of the estimation methods show clearly the necessity to distinguish the various queues on a multi lanes road, especially when the demand among the



queues is not balanced. As a perspective to these methods, considering the case where there is a residual queue at the beginning of the red light, i.e. the moderate or high traffic demand case, would be interesting. Testing and refining the assignment model proposed could also be an interesting sequel to this thesis.

On another hand, based on the new potentials raised by communicating vehicles, two control strategies of the traffic lights in urban road networks have been proposed. The first strategy considers a semi-decentralized control which is extending a centralized approach (Urban Traffic Control algorithm) by introducing some contention window in the traffic light cycle, where vehicles communicate to pass the junction. Second, an algorithm which ensures the contention window efficiency has also been proposed to explicit the concurrency of passing through the junction. A mix traffic with communicating and non communicating vehicles is assumed. The penetration ratio of communicating vehicles is variable. The vehicles form queues at the traffic light and it was assumed that the queue lengths are inferior to the road length. Performances studies have been conducted for this algorithm, using the simulation framework VEINS which combines road traffic and communication simulators. The results show good performance of the control which can avoid congestion in some high demand simulated scenarios. The results depend on a parameter which represents the junction capacity loss when vehicles communicate to pass the junction. Further studies on this parameter could benefit to the method tuning. Other perspectives to these control methods could be to study their performances on larger networks and even in real applications. Validation of these methods on real road networks would make the application practical and useful compared to theoretical analyses.

In conclusion, this thesis proposes many contributions to the estimation and control of road traffic, with reasonable and explicit assumptions, which have been discussed through simulation studies. The development of these approaches is included in the general process of optimization of mix traffic (with communicating vehicles and traditional ones) which is an important phase of the near future, preceding a full automated road traffic. The simulation results show the relevance of the assumptions and of the methods developed. In particular, the refinement of the road traffic estimation for the multi lanes roads enables better regulation of urban road traffic. This type of methods will improve traffic estimation and optimization in urban road networks. In addition to improving the road traffic conditions that users experiment, the work presented in this thesis could benefit to the economical and ecological fields. It is important to keep in mind that the transportation systems are connected to domains that matter, on the mid and long terms.

## 5.7 Implications and consequences

The part of this thesis which deals with the estimation of road traffic has introduced an assignment model of vehicles onto the various lanes. This assignment model could be used in other road traffic fields such as lane assignment models. In our case, this assignment model is the basis for estimating the penetration ratio of communicating vehicles, the arrival rates of vehicles per lane and the queue lengths probability distributions per lane. This thesis has also proven that increasing the accuracy of the estimation of queue lengths in the cities, is particularly significant in the cases of imbalanced queues. We recall that this case of imbalance is coming from the combination of the junction topology and traffic demand properties. With this greater accuracy, the regulation of the traffic would lead to better performances which is valuable because the transition between a fluid and congested road traffic is very sensitive to both road traffic input data and road traffic control.

Concerning the road traffic control, the main contribution was the extension of the Traffic Urban Control (TUC) (centralized control) by the introduction of a concurrency time window where vehicles communicate in order to pass the junction (local control) which results in a semi decentralized control. The local control gives a practical solution for controlling traffic lights given the presence of communicating vehicles. Combined with the TUC control, it results in a ready made solution which enhances the state of the art strategies, in terms of performances, to avoid congestions in the cities. In order to apply this findings, it would be adequate to test the assumptions and evaluate the strategies both in simulation and in the field. This recalls us that transportation research matters for the society with its various and important stakes and application fields. By proposing contributions to the state of the art both in the estimation and control of urban road traffic, this thesis could contribute to enhance the conditions of transportation systems experienced by the individuals and the society on the mid and long terms.



# Bibliography

- [1] [http://sumo.dlr.de/wiki/Definition\\_of\\_Vehicles,\\_Vehicle\\_Types,\\_and\\_Routes](http://sumo.dlr.de/wiki/Definition_of_Vehicles,_Vehicle_Types,_and_Routes).
- [2] Ieee guide for wireless access in vehicular environments (wave) - architecture. *IEEE Std 1609.0-2013*, pages 1–78, March 2014.
- [3] Ieee guide for wireless access in vehicular environments (wave) - architecture. *IEEE Std 1609.0-2013*, pages 1–78, March 2014.
- [4] W F ADAMS. Road traffic considered as a random series. (includes plates). *Journal of the Institution of Civil Engineers*, 4(1):121–130, 1936.
- [5] National Highway Traffic Safety Administration.
- [6] Seli James Agbolosu-Amison, Ilsoo Yun, and Byungkyu Brian Park. Quantifying benefits of a dynamic gap-out feature at an actuated traffic signalized intersection under cooperative vehicle infrastructure system. *KSCE Journal of Civil Engineering*, 16(3):433–440, 2012.
- [7] Rahmi Akçelik. Time-dependent expressions for delay, stop rate and queue length at traffic signals.
- [8] Muhammad Alam, Joaquim Ferreira, and José Fonseca. *Introduction to Intelligent Transportation Systems*, pages 1–17. Springer International Publishing, Cham, 2016.
- [9] Kevin N. Balke, Conrad L. Dudek, and II Thomas Urbanik. Development and evaluation of intelligent bus priority concept. *Transportation Research Record*, 1727(1):12–19, 2000.
- [10] Xuegang (Jeff) Ban, Peng Hao, and Zhanbo Sun. Real time queue length estimation for signalized intersections using travel times from mobile sensors. *Transportation Research Part C: Emerging Technologies*, 19(6):1133 – 1156, 2011.

- [11] Maram Bani Younes and Azzedine Boukerche. Intelligent traffic light controlling algorithms using vehicular networks. *IEEE Transactions on Vehicular Technology*, 65(8):5887–5899, 2016.
- [12] Byungho Beak, K. Larry Head, and Yiheng Feng. Adaptive coordination based on connected vehicle technology. *Transportation Research Record*, 2619(1):1–12, 2017.
- [13] Martin J Beckmann, Charles B McGuire, and Christopher B Winsten. *Studies in the economics of transportation*. 1955.
- [14] Michael Behrisch, Laura Bieker, Jakob Erdmann, and Daniel Krajzewicz. Sumo-simulation of urban mobility-an overview. In *SIMUL 2011, The Third International Conference on Advances in System Simulation*, pages 55–60, 2011.
- [15] Francesco Bella and Roberta Russo. A collision warning system for rear-end collision: a driving simulator study. *Procedia-social and behavioral sciences*, 20:676–686, 2011.
- [16] Dimitri P Bertsekas et al. Dynamic programming and optimal control 3rd edition, volume ii. *Belmont, MA: Athena Scientific*, 2011.
- [17] Patrick Bonnel. *Prévision de la demande de transport*. PhD thesis, Université Lumière-Lyon II, 2002.
- [18] R.D. Bretherton, Wood K., and Bowen G.T. Scoot -version 4. In *Proceedings of the 9th international conference on Road Transport Information and Control*, 1998.
- [19] Walter Bronzi, Raphael Frank, German Castignani, and Thomas Engel. Bluetooth low energy performance and robustness analysis for inter-vehicular communications. *Ad Hoc Networks*, 37:76–86, 2016.
- [20] F. Busch. Traffic telematics in urban and regional environments. In *Proceedings of the Intertraffic Conference. Amsterdam*, 1996.
- [21] Shukai Chen and Daniel Jian Sun. An improved adaptive signal control method for isolated signalized intersection based on dynamic programming. *IEEE Intelligent Transportation Systems Magazine*, 8(4):4–14, 2016.
- [22] Yang Cheng, Xiao Qin, Jing Jin, Bin Ran, and Jason Anderson. Cycle-by-Cycle Queue Length Estimation for Signalized Intersections Using Sampled Trajectory Data. *Transportation Research Record*, 2257(1):87–94, January 2011. Publisher: SAGE Publications Inc.

- [23] Junil Choi, Vutha Va, Nuria Gonzalez-Prelcic, Robert Daniels, Chandra R Bhat, and Robert W Heath. Millimeter-wave vehicular communication to support massive automotive sensing. *IEEE Communications Magazine*, 54(12):160–167, 2016.
- [24] S. Cohen. *Ingénierie du trafic routier: Elément de théorie du trafic et applications*. Eyrolles, 1993.
- [25] G. Comert and M. Cetin. Analytical Evaluation of the Error in Queue Length Estimation at Traffic Signals From Probe Vehicle Data. *IEEE Transactions on Intelligent Transportation Systems*, 12(2):563–573, June 2011.
- [26] Gurcan Comert. Simple analytical models for estimating the queue lengths from probe vehicles at traffic signals. *Transportation Research Part B: Methodological*, 55:59 – 74, 2013.
- [27] Gurcan Comert. Queue length estimation from probe vehicles at isolated intersections: Estimators for primary parameters. *European Journal of Operational Research*, 252(2):502 – 521, 2016.
- [28] Gurcan Comert and Mecit Cetin. Queue length estimation from probe vehicle location and the impacts of sample size. *European Journal of Operational Research*, 197(1):196 – 202, 2009.
- [29] C. Daganzo and N. Geroliminis. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological*, 42(9), 2008.
- [30] Ruina Dang, Jieyun Ding, Bo Su, Qichang Yao, Yuanmu Tian, and Keqiang Li. A lane change warning system based on v2v communication. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1923–1928. IEEE, 2014.
- [31] Kashif Dar, Mohamed Bakhouya, Jaafar Gaber, Maxime Wack, and Pascal Lorenz. Wireless communication technologies for its applications [topics in automotive networking]. *IEEE Communications Magazine*, 48(5):156–162, 2010.
- [32] Christopher M. Day and Darcy M. Bullock. Detector-Free Signal Offset Optimization with Limited Connected Vehicle Market Penetration: Proof-of-Concept Study. *Transportation Research Record*, 2558(1):54–65, January 2016. Publisher: SAGE Publications Inc.
- [33] Christina Diakaki. *Integrated control of traffic flow in corridor networks*. PhD thesis, Ph. D. Thesis, Technical University of Crete, Department of Production Engineering and Management, Chania, Greece, 1999.

- [34] Christina Diakaki, Vaya Dinopoulou, Kostas Aboudolas, Markos Papageorgiou, Elia Ben-Shabat, Eran Seider, and Amit Leibov. Extensions and new applications of the traffic-responsive urban control strategy: Coordinated signal control for urban networks. *Transportation Research Record: Journal of the Transportation Research Board*, 1856(1):202–211, 2003.
- [35] Christina Diakaki, Markos Papageorgiou, and Kostas Aboudolas. A multivariable regulator approach to traffic-responsive network-wide signal control. *Control Engineering Practice*, 10(2):183 – 195, 2002.
- [36] George J. Dimitrakopoulos, Lorna Uden, and Iraklis Varlamis. *The Future of Intelligent Transport Systems*. Elsevier, March 2020.
- [37] Catalin Dimon. *Contributions à la modélisation et la commande des réseaux de trafic routier*. Theses, Ecole Centrale de Lille ; Universitatea politehnica (Bucarest), February 2012.
- [38] Felipe Domingos Da Cunha, Leandro Villas, Azzedine Boukerche, Guilherme Maia, Aline Carneiro Viana, Raquel A. F. Mini, and Antonio A. F. Loureiro. Data Communication in VANETs: Survey, Applications and Challenges. *Ad Hoc Networks*, 44(C):90–103, December 2016.
- [39] F. Donati, V. Mauro, G. Roncoloni, and M. Vallauri. A hierarchical decentralized traffic light control system - the first realisation. In *Progetto Torino Proceedings of the 9th World Congress Of the International Federation of Automotive Control. Budapest*, pages 2853–2858, 1984.
- [40] John W. Eaton, David Bateman, Søren Hauberg, and Rik Wehbring. *GNU Octave version 6.3.0 manual: a high-level interactive language for numerical computations*, 2021.
- [41] Oscar Egu-Festas. *Apports des données passives à la compréhension des comportements de mobilité?: Enjeux pour la planification et l'organisation des transports en commun*. PhD thesis, Université de Lyon, 2020.
- [42] Wael Ekeila, Tarek Sayed, and Mohamed El Esawey. Development of dynamic transit signal priority strategy. *Transportation Research Record*, 2111(1):1–9, 2009.
- [43] Oumaima El Joubari. *Mobility and traffic models for VANETs*. Theses, Université Paris-Saclay, March 2022.

- [44] Jakob Erdmann. Sumo's lane-changing model. In Michael Behrisch and Melanie Weber, editors, *Modeling Mobility with Open Data*, pages 105–123, Cham, 2015. Springer International Publishing.
- [45] ETSI. *Intelligent Transport Systems (ITS); ITS-G5 Access layer specification for Intelligent Transport Systems operating in the 5 GHz frequency band*, 2020.
- [46] J.L. Farges, K. Khoudour, and J.B. Lesort. Prodyn: on site evaluation. In *Third International Conference on Road Traffic Control. Institute of Electrical Engineers, London, UK,*, pages 62–66, 1990.
- [47] N. Farhi. Modeling and control of elementary 2d-traffic systems using petri nets and minplus algebra. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 2292–2297, Dec 2009.
- [48] N. Farhi, M. Goursat, and J.-P. Quadrat. The traffic phases of road networks. *Transportation Research Part C*, 19(1):85–102, 2011.
- [49] N. Farhi, M. Goursat, and J.P. Quadrat. Derivation of the fundamental traffic diagram for two circular roads and a crossing using minplus algebra and petri net modeling. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pages 2119–2124, Dec 2005.
- [50] N. Farhi, M. Goursat, and J.P. Quadrat. Fundamental traffic diagrams of elementary road networks. In *Control Conference (ECC), 2007 European*, pages 4944–4951, July 2007.
- [51] Nadir Farhi. *Modélisation Minplus et Commande du Trafic de Villes Régulières*. PhD thesis, Université Paris 1 Panthéon-Sorbonne, 2008.
- [52] Nadir Farhi. *Dynamic programming systems for modeling and control of the traffic in transportation networks*. Habilitation à diriger des recherches, Université Paris-Est Marne-la-Vallée, December 2018.
- [53] Nadir Farhi, Cyril Nguyen Van Phu, Mouna Amir, Habib Haj-Salem, and Jean-Patrick Lebacque. A semi-decentralized control strategy for urban traffic. *Transportation Research Procedia*, 10:41–50, 2015. 18th Euro Working Group on Transportation, EWGT 2015, 14-16 July 2015, Delft, The Netherlands.
- [54] Yiheng Feng, K. Larry Head, Shayan Khoshmaghham, and Mehdi Zamanipour. A real-time adaptive signal control in a connected vehicle environment. *Transportation Research Part C: Emerging Technologies*, 55:460–473, June 2015.



- [55] Tiago M Fernández-Caramés, Miguel González-López, and Luis Castedo. Mobile wimax for vehicular applications: Performance evaluation and comparison against ieee 802.11 p/a. *Computer Networks*, 55(16):3784–3795, 2011.
- [56] Ryan Florin and Stephan Olariu. A survey of vehicular communications for traffic signal optimization. *Veh. Commun.*, 2(2):70–79, April 2015.
- [57] R.G. Gallager. *Stochastic Processes: Theory for Applications*. Stochastic Processes: Theory for Applications. Cambridge University Press, 2013.
- [58] N.N. Gartner. Road traffic control: Demand responsive. In *In Papa-georgiou, M. (Eds), Concise Encyclopedia of Traffic and Transportation Systems*. Pergamon Press 37, pages 386–391, 1991.
- [59] N. Geroliminis and C. F. Daganzo. Macroscopic modeling of traffic in cities. *86th Transportation Research Board Annual Meeting*, (07-0413), 2007.
- [60] Shivani Godha. On-road obstacle detection system for driver assistance. *Asia Pacific Journal of Engineering Science and Technology*, 3(1):16–21, 2017.
- [61] Noah J. Goodall, Brian L. Smith, and Byungkyu (Brian) Park. Traffic signal control with connected vehicles. *Transportation Research Record*, 2381(1):65–72, 2013.
- [62] Noah J. Goodall, Brian L. Smith, and Byungkyu Brian Park. Traffic signal control with connected vehicles. *Transportation Research Record: Journal of the Transportation Research Board*, 2381:65–72, 2013.
- [63] Gps.gov. *GPS Accuracy*.
- [64] Victor Gradinescu, Cristian Gorgorin, Raluca Diaconescu, Valentin Cristea, and Liviu Iftode. Adaptive Traffic Lights Using Car-to-Car Communication. In *2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring*, pages 21–25, April 2007. ISSN: 1550-2252.
- [65] Jean Gregoire, Emilio Frazzoli, Arnaud de La Fortelle, and Tichakorn Wongpiromsarn. Back-pressure traffic signal control with unknown routing rates. *IFAC Proceedings Volumes*, 47(3):11332–11337, 2014.
- [66] Jean Gregoire, Emilio Frazzoli, Arnaud de La Fortelle, and Tichakorn Wongpiromsarn. Back-pressure traffic signal control with unknown routing rates. *IFAC Proceedings Volumes*, 47(3):11332–11337, 2014. 19th IFAC World Congress.

- [67] Qiangqiang Guo, Li Li, and Xuegang (Jeff) Ban. Urban traffic signal control with connected and automated vehicles: A survey. *Transportation Research Part C: Emerging Technologies*, 101:313 – 334, 2019.
- [68] Shereen Hamato, Sharifah Hafizah Syed Ariffin, and Norsheila Fisal. Overview of wireless access in vehicular environment (wave) protocols and standards. *Indian Journal of Science and Technology*, 7:07, 2013.
- [69] Zeeshan Hameed Mir and Fethi Filali. Lte and ieee 802.11p for vehicular networking: a performance evaluation. *EURASIP Journal on Wireless Communications and Networking*, 2014(1):89, 2014.
- [70] Peng Hao, Xuegang (Jeff) Ban, Dong Guo, and Qiang Ji. Cycle-by-cycle intersection queue length distribution estimation using sample travel times. *Transportation Research Part B: Methodological*, 68:185 – 204, 2014.
- [71] Peng Hao, Zhanbo Sun, Xuegang (Jeff) Ban, Dong Guo, and Qiang Ji. Vehicle index estimation for signalized intersections using sample travel times. *Transportation Research Part C: Emerging Technologies*, 36:513 – 529, 2013.
- [72] Qing He, K. Larry Head, and Jun Ding. PAMSCOD: Platoon-based arterial multi-modal signal control with online data. *Transportation Research Part C: Emerging Technologies*, 20(1):164–184, 2012.
- [73] Qing He, K. Larry Head, and Jun Ding. Multi-modal traffic signal control with priority, signal actuation and coordination. *Transportation Research Part C: Emerging Technologies*, 46:65–82, September 2014.
- [74] K Larry Head, Pitu B Mirchandani, and Dennis Sheppard. *Hierarchical framework for real-time traffic control*. Number 1360. 1992.
- [75] Dirk Heidemann. Queue length and delay distributions at traffic signals. *Transportation Research Part B: Methodological*, 28(5):377 – 389, 1994.
- [76] Jia Hu, Byungkyu Brian Park, and Young-Jae Lee. Coordinated transit signal priority supporting transit progression under Connected Vehicle Technology. *Transportation Research Part C: Emerging Technologies*, 55:393–408, June 2015.
- [77] P.B. Hunt, R.D. Robertson, R.D. Bretherton, and Winton R.I. Scoot - a traffic responsive method of co-ordinating signals. *Technical Report, Transport and Road Research Laboratory*, (1014), 1981.

- [78] IEEE. *802.11p-2010 - IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments*, 2010.
- [79] S. Ilgin Guler, Monica Menendez, and Linus Meier. Using connected vehicle technology to improve the efficiency of intersections. *Transportation Research Part C: Emerging Technologies*, 46:121–131, September 2014.
- [80] S. M. A. Bin Al Islam, Ali Hajbabaie, and H. M. Abdul Aziz. A real-time network-level traffic signal control methodology with partial connected vehicle information. *Transportation Research Part C: Emerging Technologies*, 121:102830, December 2020.
- [81] S.M.A. Bin Al Islam and Ali Hajbabaie. Distributed coordinated signal timing optimization in connected transportation networks. *Transportation Research Part C: Emerging Technologies*, 80:272–285, 2017.
- [82] Janis Jansons, Ernests Petersons, and Nikolajs Bogdanovs. Wifi for vehicular communication systems. In *2013 27th International Conference on Advanced Information Networking and Applications Workshops*, pages 425–430. IEEE, 2013.
- [83] David Kari, Guoyuan Wu, and Matthew J. Barth. Development of an agent-based online adaptive signal control strategy using connected vehicle technology. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1802–1807, 2014.
- [84] John P Kasik. Keep calm: The communications air interface long and medium range (calm) effort is designed to create a standard for vehicle-to-vehicle and vehicle-to-infrastructure communication requirements-a vital project for vii. *Traffic Technology International*, 2006.
- [85] Roozbeh Kianfar, Bruno Augusto, Alireza Ebadighajari, Usman Ha-keem, Josef Nilsson, Ali Raza, Reza S Tabar, Naga VishnuKanth Irukulapati, Cristofer Englund, Paolo Falcone, et al. Design and experimental validation of a cooperative driving system in the grand cooperative driving challenge. *IEEE transactions on intelligent transportation systems*, 13(3):994–1007, 2012.
- [86] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. Recent development and applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3&4):128–138, December 2012.

- [87] V. Kumar, L. Lin, D. Krajzewicz, F. Hrizi, O. Martinez, J. Gozalvez, and R. Bauza. itetris: Adaptation of its technologies for large scale integrated simulation. In *2010 IEEE 71st Vehicular Technology Conference*, pages 1–5, May 2010.
- [88] S. Kwatirayo, J. Almhana, and Z. Liu. Optimizing intersection traffic flow using vanet. In *2013 IEEE International Conference on Sensing, Communications and Networking (SECON)*, pages 260–262, June 2013.
- [89] E Lalitha and D Jayachandran. Multimedia content downloading in vanet with density estimation. *International Journal of Engineering Research and General Science*, 2(2), 2014.
- [90] Guillaume Leduc et al. Road traffic data: Collection methods and applications. 2008.
- [91] J. Lee, B. Park, and I. Yun. Cumulative travel-time responsive real-time intersection control algorithm in the connected vehicle environment. *Journal of Transportation Engineering*, 139(10):1020–1029, 2013.
- [92] Seunghyeon Lee, S.C. Wong, and Y.C. Li. Real-time estimation of lane-based queue lengths at isolated signalized junctions. *Transportation Research Part C: Emerging Technologies*, 56:1–17, 2015.
- [93] Fuliang Li, Keshuang Tang, Jiarong Yao, and Keping Li. Real-Time Queue Length Estimation for Signalized Intersections Using Vehicle Trajectory Data. *Transportation Research Record*, 2623(1):49–59, 2017. eprint: <https://doi.org/10.3141/2623-06>.
- [94] Wan Li and Xuegang Ban. Connected vehicles based traffic signal timing optimization. *IEEE Transactions on Intelligent Transportation Systems*, 20(12):4354–4366, 2019.
- [95] Xiao (Joyce) Liang, S. Ilgin Guler, and Vikash V. Gayah. Signal timing optimization with connected vehicle technology: Platooning to improve computational efficiency. *Transportation Research Record*, 2672(18):81–92, 2018.
- [96] Xiao (Joyce) Liang, S. Ilgin Guler, and Vikash V. Gayah. A heuristic method to optimize generic signal phasing and timing plans at signalized intersections using Connected Vehicle technology. *Transportation Research Part C: Emerging Technologies*, 111:156–170, February 2020.
- [97] Chen-Fu Liao and Gary A. Davis. Simulation study of bus signal priority strategy: Taking advantage of global positioning system, au-

- tomated vehicle location system, and wireless communications. *Transportation Research Record*, 2034(1):82–91, 2007.
- [98] M J. Lighthill and J B. Whitham. On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 229(1178):317–345, 1955.
- [99] Iris Lin, Rong He, and Alain L. Kornhauser. Estimating Nationwide Link Speed Distribution Using Probe Position Data. *Journal of Intelligent Transportation Systems*, 12(1):29–37, 2008.
- [100] Jennie Lioris, Ramtin Pedarsani, Fatma Yildiz Tascikaraoglu, and Pravin Varaiya. Platoons of connected vehicles can double throughput in urban roads. *Transportation Research Part C: Emerging Technologies*, 77:292–305, April 2017.
- [101] Arthur D Little. The future of urban mobility 2.0. *Arthur D. Little Future lab study for UITP International association of public transport* <http://www.uitp.org/sites/default/files/members/140124%20Arthur%20D.%20Little>, 20:26, 2014.
- [102] Henry X. Liu, Xinkai Wu, Wenteng Ma, and Heng Hu. Real-time queue length estimation for congested signalized intersections. *Transportation Research Part C: Emerging Technologies*, 17(4):412 – 427, 2009.
- [103] P.R. Lowrie. The sydney coordinated adaptive traffic system. In *Proceedings of Institute of Electrical Engineers International Conference on Road Road Traffic Signalling*. London, pages 67–70, 1982.
- [104] Po-Wen Lu and Rongshun Chen. Infrared-based vehicular positioning with the automatic radiation-strength control. *IET Intelligent Transport Systems*, 8(3):273–285, 2014.
- [105] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Wang. Traffic flow prediction with big data: A deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, 16(2):865–873, April 2015.
- [106] Steven McCanne, Sally Floyd, Kevin Fall, Kannan Varadhan, et al. Network simulator ns-2, 1997.
- [107] Michael G McNally. The four-step model. In *Handbook of transport modelling*. Emerald Group Publishing Limited, 2007.
- [108] Donald R McNeil. A solution to the fixed-cycle traffic light problem for compound poisson arrivals. *Journal of Applied Probability*, 5(3):624–635, 1968.

- [109] Alan J Miller. Settings for fixed-cycle traffic signals. *Journal of the Operational Research Society*, 14(4):373–386, 1963.
- [110] Alan J Miller. The capacity of signalized intersections in australia. 1968.
- [111] Cyril Nguyen Van Phu and Nadir Farhi. Estimation of urban traffic state with probe vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 22(5):2797–2808, 2021.
- [112] Cyril Nguyen Van Phu and Nadir Farhi. Estimation of road traffic state at a multilanes controlled junction. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–0, 2022.
- [113] U.S. Department of Transportation.
- [114] Kartik Pandit, Dipak Ghosal, H. Michael Zhang, and Chen-Nee Chuah. Adaptive traffic signal control with vehicular ad hoc networks. *IEEE Transactions on Vehicular Technology*, 62(4):1459–1471, 2013.
- [115] C. N. Van Phu, N. Farhi, H. Haj-Salem, and J. P. Lebacque. A vehicle-to-infrastructure communication based algorithm for urban traffic control. In *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 651–656, June 2017.
- [116] Craig B. Rafter, Bani Anvari, Simon Box, and Tom Cherrett. Augmenting Traffic Signal Control Systems for Urban Road Networks With Connected Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 21(4):1728–1740, April 2020. Conference Name: IEEE Transactions on Intelligent Transportation Systems.
- [117] M. Ramezani and N. Geroliminis. Exploiting probe data to estimate the queue profile in urban networks. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pages 1817–1822, Oct 2013.
- [118] Mohsen Ramezani and Nikolas Geroliminis. Queue profile estimation in congested urban networks with probe data. *Computer-Aided Civil and Infrastructure Engineering*, 30(6):414–432, 2015.
- [119] Mohsen Ramezani, Jack Haddad, and Nikolas Geroliminis. Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control. *Transportation Research Part B: Methodological*, 74(0):1 – 19, 2015.
- [120] Paul I. Richards. Shock waves on the highway. *Operations Research*, 4(1):42–51, 1956.

- [121] George F. Riley and Thomas R. Henderson. The ns-3 Network Simulator. In Klaus Wehrle, Mesut Güneş, and James Gross, editors, *Modeling and Tools for Network Simulation*, pages 15–34. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [122] D. I. Robertson. "transyt" method for area traffic control. *Traffic Eng. Control*, 10:276–281, 1969.
- [123] Samuel Y. R. Rompis, Mecit Cetin, and Filmon Habtemichael. Probe vehicle lane identification for queue length estimation at intersections. *Journal of Intelligent Transportation Systems*, 22(1):10–25, 2018.
- [124] Michele Rondinone, Julen Maneros, Daniel Krajzewicz, Ramon Bauza, Pasquale Cataldi, Fatma Hrizi, Javier Gozalvez, Vineet Kumar, Matthias Röckl, Lan Lin, Oscar Lazaro, Jérémie Leguay, Jérôme Härrri, Sendoa Vaz, Yoann Lopez, Miguel Sepulcre, Michelle Wetterwald, Robbin Blokpoel, and Fabio Cartolano. iTETRIS: A modular simulation platform for the large scale evaluation of cooperative ITS applications. *Simulation Modelling Practice and Theory*, 34:99–125, 2013.
- [125] David Schrank, Luke Albert, Bill Eisele, and Tim Lomax. 2021 Urban Mobility Report. June 2021.
- [126] Björn Schünemann. V2x simulation runtime infrastructure vsimrti: An assessment tool to design smart traffic management systems. *Comput. Netw.*, 55(14):3189–3198, October 2011.
- [127] Marcin Sereczynski and Djamel Khadraoui. Complementing transit signal priority with speed and dwell time extension advisories. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1009–1014, 2014.
- [128] Majid Rostami Shahrababaki, Ali Akbar Safavi, Markos Papageorgiou, and Ioannis Papamichail. A data fusion approach for real-time traffic state estimation in urban signalized links. *Transportation Research Part C: Emerging Technologies*, 92:525 – 548, 2018.
- [129] Neha Shinde, Shobha Bhor, Trupti Wable, and Sayli Dubule. Zigbee based vehicular adhoc networking. *International Journal of Infinite Innovations in Technology*, 2015.
- [130] Emilia Simeonova, Janet Currie, Peter Nilsson, and Reed Walker. Congestion pricing, air pollution, and children's health. *Journal of Human Resources*, 56(4):971–996, 2021.

- [131] Christoph Sommer, Reinhard German, and Falko Dressler. Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis. *IEEE Transactions on Mobile Computing*, 10(1):3–15, January 2011.
- [132] Weili Sun, Jianfeng Zheng, and Henry X. Liu. A capacity maximization scheme for intersection management with automated vehicles. *Transportation Research Procedia*, 23:121–136, January 2017.
- [133] A.S. Tanenbaum and D. Wetherall. *Computer Networks*. Pearson Prentice Hall, 2011.
- [134] Keshuang Tang, Hao Wu, Jiarong Yao, Chaopeng Tan, and Yangbeibei Ji. Lane-based queue length estimation at signalized intersections using single-section license plate recognition data. *Transportmetrica B: Transport Dynamics*, 0(0):1–19, October 2021. Publisher: Taylor & Francis \_eprint: <https://doi.org/10.1080/21680566.2021.1991504>.
- [135] K. Tiaprasert, Y. Zhang, X. B. Wang, and X. Zeng. Queue Length Estimation Using Connected Vehicle Technology for Adaptive Signal Control. *IEEE Transactions on Intelligent Transportation Systems*, 16(4):2129–2140, Aug 2015.
- [136] Kamonthep Tiaprasert, Yunlong Zhang, Xiubin Bruce Wang, and Xiaosi Zeng. Queue Length Estimation Using Connected Vehicle Technology for Adaptive Signal Control. *IEEE Transactions on Intelligent Transportation Systems*, 16(4):2129–2140, 2015.
- [137] Lai Tu and Chung-Ming Huang. Forwards: A map-free intersection collision-warning system for all road patterns. *IEEE Transactions on Vehicular Technology*, 59(7):3233–3248, 2010.
- [138] F van Diggelen and P Enge. The world’s first gps mooc and worldwide laboratory using smartphones. In *Proceedings of the 28th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2015)*, pages 361–369, 2015.
- [139] Cyril Nguyen Van Phu, Nadir Farhi, Habib Haj-Salem, and Jean-Patrick Lebacque. A vehicle-to-infrastructure communication based algorithm for urban traffic control. In *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 651–656, 2017.
- [140] Pravin Varaiya. Max pressure control of a network of signalized intersections. *Transportation Research Part C: Emerging Technologies*, 36:177–195, 2013.



- [141] Pravin Varaiya. The max-pressure controller for arbitrary networks of signalized intersections. In Satish V. Ukkusuri and Kaan Ozbay, editors, *Advances in Dynamic Network Modeling in Complex Transportation Systems*, volume 2 of *Complex Networks and Dynamic Systems*, pages 27–66. Springer New York, 2013.
- [142] András Varga. The omnet++ discrete event simulation system. In *In ESM'01*, 2001.
- [143] Georgios Vigos, Markos Papageorgiou, and Yibing Wang. Real-time estimation of vehicle-count within signalized links. *Transportation Research Part C: Emerging Technologies*, 16(1):18–35, February 2008.
- [144] Francesco Viti and Henk van Zuylen. Modeling queues at signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board*, (1883):68–77, 2004.
- [145] Francesco Viti and Henk J. van Zuylen. Probabilistic models for queues at fixed control signals. *Transportation Research Part B: Methodological*, 44(1):120–135, January 2010.
- [146] Yi Wang, A. Ahmed, B. Krishnamachari, and K. Psounis. Ieee 802.11p performance evaluation and protocol enhancement. In *2008 IEEE International Conference on Vehicular Electronics and Safety*, pages 317–322, Sept 2008.
- [147] F. V. Webster. Traffic signal settings. *Traffic Eng. Control*, (39), 1958.
- [148] Fo Vo Webster. Traffic signal settings. Technical report, 1958.
- [149] A. Wegener, H. Hellbruck, C. Wewetzer, and A. Lubke. Vanet simulation environment with feedback loop and its application to traffic light assistance. In *GLOBECOM Workshops, 2008 IEEE*, pages 1–7, Nov 2008.
- [150] Tichakorn Wongpiromsarn, Tawit Uthaicharoenpong, Yu Wang, Emilio Frazzoli, and Danwei Wang. Distributed traffic signal control for maximum network throughput. In *2012 15th International IEEE Conference on Intelligent Transportation Systems*, pages 588–595, 2012.
- [151] Wei Wu, Wanjing Ma, Kejun Long, and Yinhai Wang. Integrated optimization of bus priority operations in connected vehicle environment. *Journal of Advanced Transportation*, 50(8):1853–1869, 2016. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/atr.1433>.
- [152] Weigang Wu, Jiebin Zhang, Aoxue Luo, and Jiannong Cao. Distributed mutual exclusion algorithms for intersection traffic control.

- IEEE Transactions on Parallel and Distributed Systems*, 26(1):65–74, 2015.
- [153] Pei Xiaofei, Liu Zhaodu, MA Guocheng, and YE Yang. Safe distance model and obstacle detection algorithms for a collision warning and collision avoidance system. *Journal of Automotive Safety and Energy*, 3(1):26, 2012.
- [154] Xiao-Feng Xie, Gregory J. Barlow, Stephen F. Smith, and Zachary B. Rubinstein. Platoon-based self-scheduling for real-time traffic signal control. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 879–884, 2011.
- [155] Biao Xu, Xuegang Jeff Ban, Yougang Bian, Jianqiang Wang, and Keqiang Li. V2i based cooperation between traffic signal and approaching automated vehicles. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pages 1658–1664, 2017.
- [156] Kaidi Yang, S. Ilgin Guler, and Monica Menendez. Isolated intersection control for various levels of vehicle technology: Conventional, connected, and automated vehicles. *Transportation Research Part C: Emerging Technologies*, 72:109–129, 2016.
- [157] Kaidi Yang, Monica Menendez, and S. Ilgin Guler. Implementing transit signal priority in a connected vehicle environment with and without bus stops. *Transportmetrica B: Transport Dynamics*, 7(1):423–445, 2019.
- [158] Zhihong Yao, Yangsheng Jiang, Bin Zhao, Xiaoling Luo, and Bo Peng. A dynamic optimization method for adaptive signal control in a connected vehicle environment. *Journal of Intelligent Transportation Systems*, 24(2):184–200, March 2020. Publisher: Taylor & Francis eprint: <https://doi.org/10.1080/15472450.2019.1643723>.
- [159] Juyuan Yin, Jian Sun, and Keshuang Tang. A Kalman Filter-Based Queue Length Estimation Method with Low-Penetration Mobile Sensor Data at Signalized Intersections. *Transportation Research Record*, 2672(45):253–264, December 2018. Publisher: SAGE Publications Inc.
- [160] Chunhui Yu, Yiheng Feng, Henry X. Liu, Wanjing Ma, and Xiaoguang Yang. Integrated optimization of traffic signals and vehicle trajectories at isolated urban intersections. *Transportation Research Part B: Methodological*, 112:89–112, June 2018.
- [161] Mahdi ZARGAYOUNA. *Multi-Agent Approaches for Dynamic Transportation Problems*. Habilitation à diriger des recherches, Université Paris Dauphine, January 2019.

- [162] Xiaosi Zeng, Xin Sun, Yunlong Zhang, and Luca Quadrioglio. Person-based adaptive priority signal control with connected-vehicle information. *Transportation Research Record*, 2487(1):78–87, 2015.
- [163] Yan Zhao, Wai Wong, Jianfeng Zheng, and Henry X. Liu. Maximum Likelihood Estimation of Probe Vehicle Penetration Rates and Queue Length Distributions From Probe Vehicle Data. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–9, 2021.
- [164] Jianfeng Zheng and Henry X. Liu. Estimating traffic volumes for signalized intersections using connected vehicle data. *Transportation Research Part C: Emerging Technologies*, 79:347 – 362, 2017.
- [165] Milan Zlatkovic, Aleksandar Stevanovic, and Peter T. Martin. Development and evaluation of algorithm for resolution of conflicting transit signal priority requests. *Transportation Research Record*, 2311(1):167–175, 2012.

# List of Figures

2.1	A signalized junction of the road network. . . . .	57
2.2	Queues in 2-lanes incoming link. Vehicles that can choose both lanes are assigned onto lane $M$ with probability $\alpha$ and onto lane $N$ with probability $(1 - \alpha)$ . . . . .	58
2.3	Estimated penetration ratio $p$ for a one lane incoming link, depending on penetration ratio for various demand scenarios. The arrival demand levels ( $\lambda$ ) are given in (vehicles/s). Simulated time = 40 min. . . . .	68
2.4	Estimated penetration ratio $p$ for a two lanes incoming link, depending on penetration ratio for various demand scenarios. The arrival demand levels ( $\lambda$ ) are given in (vehicles/s). Simulated time = 40 min. $\kappa = 0.5$ . . . . .	69
2.5	$(\lambda_n + \lambda_m + \lambda_{nm})$ in vehicles/second for a two lanes incoming link, depending on penetration ratio for various demand scenarios. Simulated time = 40 min. $\kappa = 0.5$ . . . . .	70
2.6	On the top : probability law of Proposition 1. On the bottom : probability law of Proposition 2. Scenario with $\lambda_n = 1/6, \lambda_m = 1/12, \lambda_{nm} = 1/24$ (veh/s)(thus $\alpha^*(1) = 1$ ), $r(t) = 41$ s, $p = 0.55$ , $c_p = 8$ , $l_p = 9$ , at time $t = 760$ s. Red dot is the expectation of the probability distribution. Simulation time=1200s. . . . .	72
2.7	Queue lengths estimator as given by Proposition 2, for varying $p = 0.2, p = 0.7$ and lane N, $\bar{r} = 1$ . . . . .	77
2.8	Queue lengths estimator as given by Proposition 2, for varying $p = 0.2, p = 0.7$ and lane M, $\bar{r} = 1$ . . . . .	78
2.9	$\alpha^*$ and $\bar{r}$ depending on cycle time of 90s for $\lambda_n = \lambda_m = \lambda_{nm} = \frac{200}{1200} = 0.17$ vehicles/second. . . . .	79
2.10	$I = [\frac{\lambda_m}{\lambda_n + \lambda_{nm}}, \frac{\lambda_m + \lambda_{nm}}{\lambda_n}]$ depending on $\lambda_{nm}$ for $\lambda_n = \lambda_m = 1/6$ . . . . .	80
3.1	Vehicles queue at a controlled junction, in a three-lanes incoming road. . . . .	91
3.2	Summary of the estimation process . . . . .	101
3.3	Penetration ratio estimator. Simulated time=25 hours. . . . .	104

3.4	Mean absolute error for $\hat{a}_p$ , $\hat{b}_p$ , $\hat{c}_p$ and estimators $E_0$ , $E_1$ , averaged over 10 replications. Simulated time = 10 hours, asymmetric scenario. . . . .	106
3.5	Mean absolute error for queue lengths on lanes $i_a$ , $i_b$ and $i_c$ , with Propositions 1, 2, 3. $m$ means estimating the queue length with $\hat{A} = \hat{B} = \hat{C} = m$ the last probe location. Simulated time = 2.5 hours and symmetric scenario. . . . .	107
3.6	Mean absolute error for queue lengths on lanes $i_a$ , $i_b$ and $i_c$ , with Propositions 1, 2, 3. $m$ means estimating the queue length with $\hat{A} = \hat{B} = \hat{C} = m$ the last probe location. Simulated time = 2.5 hours and asymmetric scenario. . . . .	108
4.1	Academic example explaining the TUC strategy. . . . .	118
4.2	The cycle time in the classical model, and in the new model. G: green, R: red, Y: yellow. . . . .	120
4.3	The cycle time in the new model. . . . .	120
4.4	Regular network example. . . . .	124
4.5	The state of the traffic at the end of simulation. The colours of vehicles correspond to their speed (green: high speed, red: low speed). On the left side: Centralized TUC. On the right side: semi-centralized TUC. . . . .	126
4.6	Comparison of the classical TUC with the semi-decentralized TUC in terms of the number of running vehicles on the network, the flow of ended vehicles, and the average travel time through the network, respectively, in function of time. . . . .	127
4.7	Semi-decentralized TUC. The control in terms of the traffic light times into the cycle time, through the simulation time, on different zones (center and boundaries), and for approaches coming from left and right sides. . . . .	128
5.1	The lead vehicles . . . . .	136
5.2	The IEEE WAVE protocols stack [3] . . . . .	138
5.3	Amount of TCP application data sent divided by the total simulation time (bit/s) . . . . .	140
5.4	Mean TCP end-to-end delay (s) . . . . .	141
5.5	TCP application throughput on RSU (bit/s) . . . . .	142
5.6	Ratio of ended vehicles by inserted vehicles at 600s simulation time . . . . .	143
5.7	Mean Travel Time for equipped vehicles (s) . . . . .	144
5.8	Mean action interval (s) for a cycle duration of 90 s. . . . .	145
5.9	Regular network example. . . . .	146

# List of Tables

2.1	Notations . . . . .	60
2.2	Demand for different scenarios (simulated time=1200 s) and for $r_N = r_M$ . . . . .	71
2.3	Results in the unit "number of vehicles" for queue on lane N, estimated vs SUMO queue length. . . . .	74
2.4	Results in the unit "number of vehicles" for queue on lane M, estimated vs SUMO queue length. . . . .	75
3.1	Notations . . . . .	92
3.2	Matrix $W$ for the symmetric scenario. Total arrival rate=0.75 vehicles/second. . . . .	104
3.3	Matrix $W$ for the asymmetric scenario. Total arrival rate=0.5 vehicles/second. . . . .	104
4.1	The traffic demand. . . . .	125
4.2	The values of other parameters. . . . .	125
5.1	The IA and vehicle map module . . . . .	135
5.2	Main parameters for the communication and road traffic control. Other parameters are VEINS defaults ones. . . . .	139
5.3	The traffic demand for the first 900 s. . . . .	146
5.4	The traffic demand for the last 900 s. . . . .	146
5.5	Ended vehicles in a scenario with the traffic demand of tables III and IV. Simulated time = 1800 s. Mean and standard deviation for 20 simulation runs. . . . .	147
5.6	Running vehicles in a scenario with the traffic demand of tables III and IV. Simulated time = 1800 s. Mean and standard deviation for 20 simulation runs. . . . .	147
5.7	Mean Travel Time (s) in a scenario with the traffic demand of tables III and IV. Simulated time = 1800 s. Mean and standard deviation for 20 simulation runs. . . . .	147