Intersection management strategies to realise the benefits of new generation vehicles

Wu, Yuanyuan

2021

Wu, Y. (2021). Intersection management strategies to realise the benefits of new generation vehicles. Doctoral thesis, Nanyang Technological University, Singapore.
<https://hdl.handle.net/10356/152230>

<https://hdl.handle.net/10356/152230>

<https://doi.org/10.32657/10356/152230>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 23:08:58 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**INTERSECTION MANAGEMENT
STRATEGIES TO REALISE THE
BENEFITS OF NEW GENERATION
VEHICLES**

WU YUANYUAN

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING

2021

INTERSECTION MANAGEMENT STRATEGIES TO REALISE THE BENEFITS OF NEW GENERATION VEHICLES

WU YUANYUAN

School of Civil and Environmental Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

2021

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

08/07/2021

.....

Date

wy²

.....

WU YUANYUAN

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

08/07/2021

.....

Date

.....

Prof. Zhu Feng

Authorship Attribution Statement

This thesis contains material from [1] paper(s) published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

Chapter 2 is published as [Y. Wu, H. Chen, and F. Zhu. DCL-AIM: Decentralized coordination learning of autonomous intersection management for connected and automated vehicles. Transportation Research Part C: Emerging Technologies, 103, 246-260 \(2019\). DOI: 10.1016/j.trc.2019.04.012.](#)

The contributions of the co-authors are as follows:

- A/Prof Zhu provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts. The manuscript was revised by Dr Chen.
- I co-designed the study with A/Prof Zhu.
- I encoded the algorithms and conducted the training and simulations. I also analysed the data.

08/07/2021

.....

Date

wy²

.....

WU YUANYUAN

Acknowledgements

I would like to give my sincerest respect and greatest gratitude to my mentor, Asst Prof Zhu Feng. I really appreciate his impressive patience, continuous guidance, support and kindness to me. I would also like to thank Assoc Prof Wong Yiik Diew and Asst Prof Qian Shunzhi for taking the time to be my TAC members and their guidance of my work. This thesis would not have been possible without their continuous guidance and support throughout my PhD study at NTU.

My gratitude also goes to all my teammates, and the colleagues in N1-B1c-26. Thanks to them for their enlightening discussions, valuable suggestions, and caring encouragement.

I also want to thank my landlord and her dog Ruby, they have provided a lot of help and added a lot of joy to my life in Singapore. Thanks to coach Herman and my teammates, the experience of being a member of NTU football team will be unforgettable forever, the lore goal in the final against NUS will also be one of the precious memories of my life.

Special thanks to all my family members, for their unconditional support and love.

“Produce the Thought Which Abides Nowhere.”

—Vajra Sutra

To my dear family

Contents

Acknowledgements	ix
Table of Contents	xiii
List of Figures	xvii
List of Tables	xix
Abbreviations	xxi
Abstract	xxiii
1 Introduction	1
1.1 Background	1
1.2 Motivations	3
1.2.1 Intersection management optimisation	4
1.2.2 Trajectory organisation comparison	4
1.2.3 Mixed traffic exploration	5
1.3 Major Contributions	5
1.4 Organisation of the thesis	7
I Intersection Management Optimisation	9
2 Decentralised Coordination Learning of AIM for CAVs	11
2.1 Introduction	12
2.2 Related work	14
2.3 Problem statement	17
2.3.1 Problem description	17
2.3.2 A multi-agent formulation of AIM problem	18
2.4 Decentralised Coordination Learning	21
2.5 Simulations and results	26
2.5.1 A typical intersection	27
2.5.2 Vehicle generation	29

2.5.3	Learning performance	30
2.5.4	Effectiveness comparison	32
2.5.5	Results of discretisation level 2	36
2.5.6	Runtime comparison	36
2.6	Conclusions	37
3	Multiple Conflicting Objectives Optimisation for AIM	39
3.1	Introduction	39
3.2	Related work	41
3.3	Problem statement and modeling	44
3.3.1	A multi-objective formulation of AIM problem	44
3.3.2	State	44
3.3.3	Action	45
3.3.4	Rewards	45
3.4	Solving algorithm and process	46
3.4.1	Generalised Bellman equation for MORL	46
3.4.2	Pre-training	48
3.5	Case Study	50
3.5.1	Simulation set up	50
3.5.2	Implementation details	51
3.5.3	Comparison and discussion	52
3.6	Conclusions	57
II	Trajectory Organisation Comparison	59
4	Comparative Study on Roundabout and Intersection Management under CAV Environment	61
4.1	Introduction	62
4.1.1	Related work	63
4.2	Parameters and assumptions	65
4.3	Capacity	67
4.3.1	Intersection using FCFS policy (I-FCFS)	67
4.3.2	Roundabout using FCFS policy (R-FCFS)	69
4.3.3	Roundabout using major-minor policy (R-MM)	70
4.3.4	Capacity comparison	72
4.4	Delay	74
4.4.1	Delay formulation	75
4.4.2	Delay comparison	76
4.5	Conclusions	79

III	Mixed Traffic Exploration	81
5	Intersection Capacity under Mixed Traffic Environment	83
5.1	Introduction	83
5.2	Modelling	86
5.2.1	Capacity at intersections	86
5.2.2	Average occupation time	90
5.3	Analyses	94
5.3.1	Effect of platooning willingness	94
5.3.2	Effect of CAV penetration rate	94
5.3.3	Effect of platoon gaps	95
5.3.4	Effect of platoon size	96
5.4	Numerical analyses	99
5.5	Conclusions	103
6	Summary	105
	List of Author's Publications	109
	References	111

List of Figures

1.1	Wireless communication technologies of CAVs. <i>photo cr. Google</i> . . .	2
1.2	Cooperative adaptive cruise control of CAVs.	3
2.1	An illustration of cell occupancies	19
2.2	Procedure of coordination learning.	22
2.3	A coordination needed state and potential joint-actions involving two CAVs.	23
2.4	Different discretisation levels of intersection area	27
2.5	Illustration of a sequential decision process and the cell occupancies	28
2.6	Traffic in-flow distribution.	30
2.7	An example of convergence of DCL-AIM algorithm in two discretisation level experiments.	31
2.8	Convergence of individual rewards (left), coordination learning (right).	31
2.9	Uncongested intersection delay comparison.	32
2.10	Congested intersection delay comparison.	33
2.11	Fluctuation of intersection delay of 100 simulation runs.	33
2.12	Step delay comparison.	35
2.13	Intersection delay comparison of discretisation level 2.	36
2.14	Average runtime per decision step comparison	37
3.1	Examples of FCFS, optimisation-based and heuristic scheduling policies (Khayatian et al., 2020). In the right row, the value of unfairness obtained from this study is shown.	43
3.2	Training with travel time	49
3.3	Simulated intersection	50
3.4	Structure of the neural network used in the learning algorithm . . .	51
3.5	Performance comparison	53
3.6	An illustration of unfairness in FCFS-AIM	53
3.7	The comparison of the average delay	55
3.8	The relationship between average delay and emission	56
4.1	Intersectional and circular organisations at a road junction.	62
4.2	The situation analysed in this study.	66
4.3	Illustration of conflict types of a typical intersection or roundabout.	66
4.4	Stream structure of roundabout.	69
4.5	Delay comparison under different demand scenarios.	76

4.6	Delay comparison of simulation results.	77
5.1	Illustration of streams at intersections	87
5.2	Conflicts among three movements in Brilon and Wu (2001).	89
5.3	Underestimation of the capacity of the conflict area model.	90
5.4	State transition diagram.	91
5.5	An illustration of the different inter-departure time.	93
5.6	An illustration of the relationship between $t_{B,i}$ and p_c	95
5.7	A typical intersection	99
5.8	Effect of platooning willingness	100
5.9	Effect of intra-platoon gap.	101
5.10	Effect of penetration rate.	101
5.11	Comparison of modeling and simulation results	102

List of Tables

2.1	Action choices of vehicles	20
2.2	Parameter settings for DCL-AIM algorithm	30
2.3	Comparison of intersection delays	35
3.1	The average delay (s)	54
3.2	Comparison of defined fairness and delay	55
3.3	Comparison of average vehicle emission (g/veh)	56
4.1	Notations	65
4.2	The conflicting streams of a stream i within the intersection	69
4.3	The circulating stream of an entry stream e within the single-lane roundabout	70
4.4	The average delay (s) of simulations	78
5.1	Notations	86
5.2	The settings of variables in Fig.5.10	102

Abbreviations

AIM	Autonomous Intersection Management
CACC	Cooperative Adaptive Cruise Control
CAV	Connected and Automated Vehicle
DCL	Decentralised Coordination Learning
DQN	Deep Q-Network
FCFS	First Come First Served
HV	Human-driven Vehicle
LQF	Longer Queue First
MDP	Markov Decision Process
MAMDP	Multi Agent Markov Decision Process
MOMDP	Multi Objective Markov Decision Process
RL	Reinforcement Learning

Abstract

Automation and wireless communication technologies are bringing us the new-generation vehicles: Connected and Automated Vehicles (*CAVs*). This thesis aims to investigate the impacts and leverage of the advanced features of *CAVs* on traffic flow at road junctions, and propose corresponding intersection management to realise the promising benefits of *CAVs*.

Conventional intersection managements in terms of controlling traffic, such as signal control, may not necessarily be the optimal strategies when it comes to *CAVs* environment. The first part of the thesis is devoted to the intersection management optimisation for *CAVs*. This task is based on the framework of Autonomous Intersection Management (*AIM*), which is tailored for *CAVs* aiming at replacing the conventional traffic control strategies. In Chapter 2, using the communication and computation technologies of *CAVs*, the sequential movements of vehicles through intersections are modelled as multi-agent Markov decision processes, in which vehicle agents cooperate to minimise intersection delay with collision-free constraints. A decentralised coordination multi-agent learning approach is proposed to find the optimal decisions. The effectiveness of the proposed method is demonstrated under a variety of traffic conditions, and compared to three representative benchmarks. The single-objective problem is extended to multiple conflicting-objective problem in Chapter 3, in which traffic efficiency and fairness are comprehensively considered to customise the optimal control strategy of *AIM*. Deep multi-objective reinforcement learning algorithm is adopted to solve the problem. The outstanding performance of the proposed method in balancing traffic fairness and efficiency is achieved as expected.

The second part of the thesis discusses the organisation of *CAV* trajectories at road junctions. Using communication technologies, *CAVs* are able to mutually negotiate the right of way at road junctions, engendering the concept of signal-free management for *CAVs*. Roundabout, as a naturally signal-free strategy, is not receiving enough attention from researchers. The significant difference lies in the

organisation of vehicle trajectories. A key motivation of this part is to theoretically compare the performance of signal-free intersection management and roundabout under CAVs environment, and to demonstrate whether roundabout has comparable performance to signal-free intersection. The theoretical formulas of the capacity and average delay of each strategy are derived by applying M/G/1 queueing model. The safety time gaps for two vehicles consecutively and safely passing conflict points are the main variables, which are assumed to be generally distributed in the theoretical derivation, rather than following a specific distribution. To the best of our knowledge, it is the first time the performances of intersection and roundabout in a CAVs environment are directly and theoretically compared. Under the same circumstances, roundabout using first-come-first-served policy has the largest capacity and the least average delay among the three analysed signal-free strategies, revealing the optimal management of CAVs based on circulating trajectory organisation is worthy of in-depth study.

The full penetration of CAV technology requires further long-term development and evolution, which makes the mixed traffic composed of both CAVs and Human-driven vehicles (HVs) inevitable. In the third part, to understand how CAV platooning affects the mixed-traffic intersection capacity, an analytical formulation for mixed-traffic intersection capacity based on the Markov chain theory and conflict point technique is proposed, with the spatial distribution of heterogeneous vehicle headways among mixed traffic as the core of the modelling. In the model, not only the CAV penetration rate and the three core platooning parameters (*intra-platoon gap*, *inter-platoon gap* and *platoon size*) are captured, but also the individual willingness of CAVs to form platoons is involved to consider the behavioural aspect of the problem. The influences of platooning parameters and platooning willingness are in line with expectations, while the effect of CAVs penetration rate is not intuitive. It is not certain that higher penetration rate will lead to higher intersection capacity, but further depends on the other analysed parameters.

Chapter 1

Introduction

Automation and wireless communication technologies are bringing us the new-generation vehicles: connected vehicles, automated vehicles, and even autonomous vehicles (also known as self-driving vehicles or driverless vehicles). These new generation vehicles are denoted as *CAVs* in this thesis. Enabled by the advanced communication, automation, computing, sensing, and positioning technologies, CAVs have the potential to change the way we travel and bring about transformative safety, mobility, energy and environment benefits. The corresponding updates of traffic management strategies are worth studying to realise the promising benefits.

1.1 Background

CAVs are expected to revolutionise future transportation and quality of life. CAVs have the ability to communicate with each other (Vehicle-to-Vehicle, V2V), road infrastructures (Vehicle-to-Infrastructure, V2I), and even non-motorised road users including pedestrians (Vehicle-to-X, V2X), as shown in Fig.1.1. After reaching a certain level, CAVs can drive themselves without the help of human beings, relying only on the computer control systems they are equipped with.

Automated driving technology is booming around the world, and many auto companies are spending numerous efforts on the development of CAVs. For example, Ford has announced a partnership with American delivery company Postmates to develop autonomous vehicles in conjunction with Ford's Argo divisions. Ford has

FIGURE 1.1: Wireless communication technologies of CAVs. *photo cr. Google*

also expressed intentions to offer infrastructure services to cities, which would include a cloud platform, in order to build the network that will power connected mobility. General Motors (GM) has presented its pedal-less and wheel-less Cruise autonomous vehicle, which it claims to be the first production-ready vehicle that requires no human input whatsoever, and designed with ride-sharing in mind. Many other auto companies are also taking similar attempts. In September 2016, the 5G Automotive Association (5GAA), a global, cross-industry organisation of companies from automotive, technology, and telecommunications industries (ICT) was created, working together to develop end-to-end solutions for future mobility and transportation services ([5GAutomotiveAssociation, 2017](#)).

At the same time, governments are actively exploring the possibility of authorising autonomous vehicles on public roads, as well as related laws and policies. Nevada is the first State to authorise the operation of CAVs by law in USA in June 2011. Since then, 20 other States plus the District of Columbia have passed some sorts of legislation pertaining to CAVs, mostly encouraging research and development according to the National Conference of State Legislatures. The British Government is spending unprecedented amounts of money on encouraging the development of cars that drive themselves. Russian technology giant Yandex has conducted test drives with its self-driving taxis in a snowy scenario, ensuring that the autonomous vehicle would be able to keep passengers safe in wintry conditions. Singapore is testing the use of CAVs in mobility-on-demand (MOD) service, especially for the *first-and-last-mile* service ([Cai et al., 2019](#)).

Synchronised with manufacturing attempts in the industry, academic research on

CAVs is also in full swing. The major difference between CAVs and current human-driving vehicles (HVs) is the much less reaction and response time in CAVs thanks to the communication and automation technologies. The most popular control system of CAV is called cooperative adaptive cruise control (CACC), an extension of adaptive cruise control, in which vehicles can measure and keep a constant distance or headway to a predecessor and exchange information with surroundings by wireless communication, as shown in Fig.1.2.

FIGURE 1.2: Cooperative adaptive cruise control of CAVs.

Starting from the new travel behaviour of CAVs, with well defined control strategies, recent studies have demonstrated from several aspects how the emergence of CAVs will affect the way we travel. CAVs can make road traffic safer because their movements can be coordinated through connectivity. Moreover, due to the movements being coordinated, non-necessary stop-and-go behaviours will be reduced, thereby reducing vehicular emissions. Particularly, they can provide us with smart transport services, such as sharing the autonomous vehicles.

1.2 Motivations

It has been estimated that CAVs will fully enter the global market by 2040s ([IHS-Markit, 2017](#)). The rapid development of CAVs brings new challenges to the corresponding traffic flow management under CAVs environment. As the crux of traffic network, the management of intersections has a significant impact on the entire transportation network. Hence, updated intersection management for CAVs is an essential topic of study ([Chen and Englund, 2016](#)).

1.2.1 Intersection management optimisation

A lot of researches have been conducted on how to use the characteristics of CAVs to solve the crux of traffic — problems in road nodes management. Some researchers have proposed the concept of signal-free intersection management ([Malikopoulos et al., 2018](#)), which can be traced back to autonomous intersection management (AIM) proposed by [Dresner and Stone \(2008\)](#). Such signal-free control is a reservation based system, in which the intersection manager accepts or rejects reservations in accordance with its control policy.

First-come-first-served (FCFS) protocol is one of the most commonly used policy cooperating with reservations, in which the right of passing the intersection is assigned based on vehicle arrival time and reservation time. However, for a dynamic system at a road junction where the traffic environment changes over time, a predefined policy may become suboptimal or even inappropriate. Hence, the optimal policy based on environmental dynamics deserves more in-depth research. In this thesis, optimal AIM framework which fully incorporates the real-time traffic dynamics is the focus.

1.2.2 Trajectory organisation comparison

In existing literature, the vast majority of the studies on junction management for CAVs are based on the underlying assumption that vehicle trajectories are intersection organised, while related research based on roundabout organisation is far less.

With appropriate signal control, intersectional trajectory management can greatly improve the actual traffic throughput of the junction and perfectly separate conflicting streams, while roundabouts could reach capacity even with moderate traffic demands ([Yang et al., 2004](#)). It is an indisputable fact that for the same junction in the current situation where the traffic flow is made up of human-driving vehicles (HVs), the signalised intersection management has better performance in terms of throughput than a roundabout. But if it is in the era of CAVs, we have not yet known which is better. When it comes to the CAVs environment, capacity may no longer be a major concern in the use of roundabout management.

Therefore, investigating vehicle management at a road junction for CAVs, focusing on comparing and identifying the trajectory organisation and policy cooperation that enable better performance for CAVs, is another core topic of this thesis.

1.2.3 Mixed traffic exploration

The full penetration of CAV technology requires long-term development and evolution ahead, which makes the mixed traffic composed of both CAVs and HVs inevitable. Exploring the impact of CAVs on mixed traffic flow is meaningful for proposing more targeted management strategies. Existing research on intersection management of mixed CAV-HV traffic mostly focused on control and optimisation, and there is a lack of impact analysis on the features of CAVs. Therefore, this thesis preliminarily explores the influence of CAV characteristics on the capacity of intersections under mixed traffic conditions. Since the mixed CAV-HV traffic flow has not yet become a reality, more general modeling methods are needed.

1.3 Major Contributions

Taking the advanced features of CAVs into consideration, this thesis aims to further improve the performance of intersection management, compare the performance of different CAV trajectory organisations and explore the impacts of CAVs on intersection capacity under mixed traffic environment.

The first part of the thesis is devoted to the optimisation of autonomous intersection management (AIM) under pure CAVs environment. Specific contributions include:

- The single-objective optimisation for AIM problem is modelled as a multi-agent Markov decision process (MDP) in Chapter 2, in which CAVs cooperate to minimise intersection delay with collision-free constraints. The huge dimension scale incurred by the nature of multi-agent decision making problems is handled by a decentralised coordination learning approach.
- The single-objective problem is extended to multiple conflicting-objective optimisation problem in Chapter 3, in which two concrete challenges in finding

the optimal decisions are tackled: (1) the objectives are essentially conflicting (traffic fairness and efficiency); (2) the relative preferences of the multiple objectives are unknown. A multi-objective reinforcement learning algorithm is customised for optimisation.

- The effectiveness of the proposed methods is demonstrated under a variety of traffic conditions, and the best performances compared to other benchmarks are achieved.

The second part contributes to theoretically comparing the performance of signal-free intersection management and roundabout under CAVs environment. Specific contributions include:

- It is the first time the performances of intersection and roundabout in a CAVs environment are directly and theoretically compared.
- The derivation of the formulas is based on the general distribution of safety gaps rather than some specific distributions, which ensures the generality of the analysis results.
- An important point is revealed that the optimal management of CAVs based on circulating trajectory organisation could have at least comparable performance as the optimal policy based on the intersectional organisation.

The mixed traffic environment where CAVs share the traffic network with human-driving vehicles (HVs) is inevitable. Part III of the thesis moves on to the impact analysis of CAVs on the intersection capacity under mixed traffic environment. Specific contributions include:

- A new theoretical approach for the determination of intersection capacity under mixed CAV-HV traffic is proposed, whereby capacity is defined as the maximum sum of the throughput of each stream crossing the intersection, which is modelled as the objective function of linear programming (LP).
- Queuing theory is used to construct the constraints of LP problem, the conflict points within the intersection determine the constraints.
- Significant corollaries are obtained, which are corroborated by numerical analyses.

1.4 Organisation of the thesis

The overall thesis consists of three parts. Part I (Chapter 2 and Chapter 3) is on the optimisation of autonomous intersection management for CAVs. Specifically, Chapter 2 is focused on the single-objective optimisation for AIM in which vehicles cooperate to minimise intersection delay with collision-free constraints. Chapter 3 extends the single-objective optimisation to multiple conflicting-objective optimisation for AIM, in which traffic fairness is innovatively optimised. In Part II (Chapter 4), the performances of intersection and roundabout in a CAVs environment are directly and theoretically compared for the first time. Part III (Chapter 5) preliminarily explores the influence of CAV features on the capacity of intersections under mixed traffic conditions. Finally, Chapter 6 concludes the thesis and discusses some interesting future research directions.

Part I

Intersection Management Optimisation

Chapter 2

Decentralised Coordination Learning of AIM for CAVs

Conventional intersection managements, such as signalised intersections, may not necessarily be the optimal strategies when it comes to connected and automated vehicles (CAVs) environment. Autonomous intersection management (AIM) is tailored for CAVs aiming at replacing the conventional traffic control strategies. In this chapter, using the communication and computation technologies of CAVs, the sequential movements of vehicles through intersections are modelled as multi-agent Markov decision processes (MAMDPs) in which vehicle agents cooperate to minimise intersection delay with collision-free constraints. To handle the huge dimension scale incurred by the nature of multi-agent decision making problems, the state space of CAVs is decomposed into independent part and coordinated part by exploiting the structure properties of the AIM problem. A decentralised coordination multi-agent learning approach (DCL-AIM) is proposed to solve the problem efficiently by exploiting both global and localised agent coordination needs in AIM. The main feature of the proposed approach is to explicitly identify and dynamically adapt agent coordination needs during the learning process so that the curse of dimensionality and environment nonstationarity problems in multi-agent learning can be alleviated.

The effectiveness of the proposed method is demonstrated under a variety of traffic conditions. The comparison analysis is performed between the DCL-AIM and the First-Come-First-Served based AIM (FCFS-AIM), with Longer-Queue-First

(LQF-AIM) policy and the signal control based on the Webster’s method (Signal) as benchmarks. Experimental results show that the sequential decisions from DCL-AIM outperform the other control policies.

2.1 Introduction

Connected and automated vehicles (CAVs) are seen as the future of transportation (Fagnant and Kockelman, 2015). CAVs have the ability to communicate with each other (Vehicle-to-Vehicle, V2V), road infrastructures (Vehicle-to-Infrastructure, V2I), and even non-motorised road users including pedestrians (Vehicle-to-X, V2X). After reaching a certain level, CAVs can drive themselves without the input of human beings, relying only on the computer control systems they are equipped with. Currently, these communication technologies and control systems are still in the research and testing phase. However, the promising benefits of CAVs are foreseeable. On one aspect, CAVs can make the road traffic safer because they can coordinate their movements through connectivity. Moreover, as the movements are coordinated, unnecessary stop-and-go behaviors will be reduced, thereby reducing emission. Further, CAVs can also provide intelligent transport services, such as sharing the autonomous vehicles (Lokhandwala and Cai, 2018).

Once CAVs become a reality, further traffic flow improvements can be expected. CAVs will change the way people travel and reshape existing transportation system. The consequent updates of traffic control and management are necessary. As the crux of traffic network, updated intersection management for CAVs has been researched (Chen and Englund, 2016). Autonomous intersection management (AIM) is one of the updated intersection control strategies tailored for CAVs. AIM, proposed by Dresner and Stone (2004), divides intersections into a grid of space-time cells to manage the coordinated movement of CAVs. CAVs wishing to cross the intersection make requests to reserve cells from the Intersection Manager (IM). The IM accepts or rejects reservations in accordance with its control policy. The most commonly used policy is first-come-first-served (FCFS), in which reservations of CAVs are prioritised by their arriving time and reservation time. Based on FCFS policy, further studies assigned higher priority for emergency vehicles (Dresner and Stone, 2006) and determined the priority pattern by using auctions (Schepperle and Böhm, 2007). It has been demonstrated that FCFS based AIM

and its variants can reduce delay and emission compared to conventional signal control under certain traffic conditions ([Fajardo et al., 2011](#); [Li et al., 2013](#)).

However, for a dynamic system like AIM where the traffic environment changes with time, a predefined policy may become suboptimal or even inappropriate. [Levin et al. \(2016\)](#) demonstrated that FCFS based AIM may not outperform traffic signals on some realistic networks. This indicates that the optimal policy based on environmental dynamics deserves more in-depth research. In this chapter, we aim to propose an optimal AIM framework which fully incorporates the real-time traffic dynamics for CAVs.

In our research framework, the processes of how CAVs cross the intersection are modelled as multi-agent Markov decision processes (MAMDPs) which are extended from single-agent MDPs. The curse of dimensionality incurred by the large number of agents is the main problem in solving MAMDPs. Reinforcement learning (RL), especially multi-agent reinforcement learning (MARL), has been proposed to solve MAMDPs. Generally, in RL, agents learn the optimal policy by trial-and-error interaction with the dynamic environment which can be formally described by Markov decision processes (MDPs) ([Sutton and Barto, 1998](#)). In AIM, CAVs are the autonomous agents who perceive and interact with the real time traffic situations (environment) using their advanced communication technologies and sensors. More specifically, we customise an MARL algorithm, called DCL-AIM algorithm, to solve the decision-making problems of CAVs.

The remainder of this chapter is organised as follows. Section 2.2 briefly reviews previous work on AIM. Section 2.3 is the problem statement, detailing the multi-agent formulation of AIM problem. Section 2.4 introduces the main methods we used in the research, including how we model the optimisation problem of AIM into decentralised coordination learning problem and the algorithms proposed to solve the problem. Section 2.5 presents the details of simulation implementations, the performance of the solution methods and the analysis results of the effectiveness comparison. Section 2.6 concludes the chapter and identifies the direction of future work.

2.2 Related work

The next revolution in transportation is expected to be Connected and Autonomous Vehicles (CAVs). Conventional intersection managements, such as signalised intersections, are not necessarily the optimal strategies when it comes to CAVs environment. There are new alternative intersection managements customised for CAVs.

One such new intersection management that has attracted a lot of attention is the reservation based autonomous intersection management (AIM) proposed by [Dresner and Stone \(2004\)](#). AIM divides an intersection into a grid of space-time cells to manage the coordinated movement of CAVs. The vehicle agents attempt to reserve cells in the intersection, then the *intersection manager* decides whether to grant or reject requested reservations according to an *intersection control policy*, named FCFS, because of the *First-Come-First-Served* nature of the reservations. Based on FCFS, emergency vehicles are given higher priorities in ([Dresner and Stone, 2006](#)), and auctions are used to determine the priority pattern in ([Carlino et al., 2013](#)). The seminal work of Stone and his team ([Dresner and Stone, 2008](#); [Hausknecht et al., 2011](#)) have demonstrated that the FCFS protocol outperforms conventional intersection control protocols, such as signal control and all-stop control, under certain traffic conditions. The similar efficacy of auction-based control protocol also has been validated ([Schepperle and Böhm, 2007, 2008](#); [Vasirani and Ossowski, 2012](#); [Carlino et al., 2013](#)).

Although it has been demonstrated in the literature that AIM and its variants have the potential to mitigate traffic congestion in urban areas, their pre-determined control policies may result in the priority assignments being only feasible but not optimal.

Some further studies focused on the optimisation of AIM. [Lee et al. \(2013\)](#) proposed a cooperative vehicle intersection control (CVIC) algorithm to minimise total length of overlapped trajectories within the intersection zone. Their objective function is equal to delay or travel time minimisation, however, the constraints are nonlinear, which may result in the final solution set containing overlapped trajectories (collisions). In order to overcome the shortcomings of the nonlinear optimising algorithms, [Zhu and Ukkusuri \(2015\)](#) proposed a linear programming formulation of AIM, which minimises the total travel time by modelling the conflict points.

Levin et al. (2017) aggregated the conflict points model into conflict regions and derived a corresponding integer programming to further improve computational efficiency. Sun et al. (2017) proposed a multi-objective mixed-integer non-linear programming designed to maximise intersection capacity. Mirheli et al. (2018) proposed a signal-head-free intersection control logic (SICL) for AIM to maximise the intersection throughput under collision-free constraints. Further on, taking SICL as a benchmark, a distributed SICL (DC-SICL) is proposed (Mirheli et al., 2019), the vehicle-level mixed-integer non-linear programming is also applied with the objective of minimising travel time and speed variations of each CAV. There are also works on fuel consumption minimisation (Zhang et al., 2016), energy-optimal control (Malikopoulos et al., 2018), and multiple performance measures (delay, fuel consumption, and safety) (Guo et al., 2019), etc.

The optimisation approaches for AIM mentioned above are more or less faced with the computational intractability problem caused by traffic dynamics. More simple and effective optimisation control is worth studying.

The ability to learn and find out the optimal decisions through dynamic interaction with the environment enables RL to be competent for the adaptive control tasks of many intelligent transportation systems (Bazzan and Klügl, 2014), and RL is applied at most to the adaptive traffic signal control system. One of the first trials was conducted by Abdulhai et al. (2003). They demonstrated that the RL-based adaptive traffic signal has better performance than conventional traffic signal control through a case study of an isolated intersection. This isolated intersection case then was extended and applied to a signal network situation using multi-agent RL approaches (Arel et al., 2010; Jin et al., 2012). There is a detailed review of RL-based adaptive traffic signal control given by Mannion et al. (2016).

Particularly, RL was applied to the dynamic speed limit control of connected vehicles (Zhu and Ukkusuri, 2014), human-like autonomous car-following behavior learning (Zhu et al., 2018), dynamic tolling approach for traffic congestion alleviation (Chen et al., 2018), optimal fuel consumption control (Qi et al., 2019), etc.

However, the above mentioned frameworks are targeted at different problems, and there are relatively few studies using RL to solve AIM problems. Isele et al. (2018) proposed a single-agent RL approach to navigate one autonomous vehicle through

the intersection. They only focus on one agent, but the traffic system is essentially multi-agent. [Chen et al. \(2018\)](#) proposed an RL-based dynamic tolling approach for traffic congestion alleviation. However, their framework is targeted at macroscopic problems over an entire traffic network, and thus cannot be applied to our problem. In this study, multi-agent reinforcement learning (MARL) is used to find the optimal sequential actions for each CAV (agent). To the best of our knowledge, this study is one of the first trials to use MARL to solve AIM problems.

Most MARL algorithms are derived from a model-free single-agent RL algorithm called *Q-learning* ([Watkins and Dayan, 1992](#); [Peng and Williams, 1994](#); [Littman, 2001](#)). MARL methods can be mainly divided into two categories: centralised learning (CL) and decentralised learning (DL) ([Busoniu et al., 2008](#)). In CL, there is a central agent with powerful computing ability that learns the value and/or policy of the multi-agent joint-action in the system. The main challenge of CL is the scalability issue caused by the exponential growth of state space. In order to avoid the curse of dimensionality, in DL, each agent independently learns its own value of actions and/or its own policy, treating other agents as part of the environment ([Tan, 1993](#)). However, DL introduces a new challenge for MARL that the environment becomes non-stationary from the point of view of each agent, as it contains other agents who are themselves learning, rolling out any convergence guarantees. Fortunately, substantial empirical evidence has shown that DL approach are often effective in practice ([Matignon et al., 2012](#)).

Decentralised coordination learning (DCL) has the potential to solve the environment nonstationarity problem in decentralised MARL. The main idea of DCL is to capture local dependencies or coordination needs between the different agent in MAMDPs, and allow the *overall Q-function* to be decomposed into *local Q-functions* that can be more easily optimised ([Guestrin et al., 2002](#); [Kok and Vlassis, 2004](#)). DCL is more effective for systems where the interaction between agents is sparse or loosely coupled. [Melo and Veloso \(2011\)](#) modelled these systems as sparse-interaction multi-agent systems, Dec-SIMDPs. In addition to the theoretical contributions, they also proposed a solution method that utilises the particular structure of Dec-SIMDPs. Further on, instead of specifying the coordination needs, [Yu et al. \(2015\)](#) proposed an algorithm to enable agents to learn efficient coordinated behaviours by exploiting agent independence in loosely coupled multi-agent systems.

In AIM, the interaction between CAVs is also sparse and not coupled at every decision step but only in relatively infrequent situations — CAVs only need to coordinate with the vehicles on their conflicting lanes when their moving intentions have potential conflicts. Due to this property, DCL is applicable for the decision-making problem in AIM. However, a key difficulty of the AIM problem is that the number of agents keeps changing over time, which makes it infeasible to directly apply DCL approaches to our problem. Based on the characteristics of traffic flow, we propose several novel adaptations and customise a decentralised coordination learning algorithm for AIM, called DCL-AIM algorithm. Details are shown in the following sections.

2.3 Problem statement

2.3.1 Problem description

Introduced by [Dresner and Stone \(2004\)](#), AIM is a multi-agent system, which applies a reservation-based protocol. In the protocol, vehicle agents need to send a request to the intersection manager (IM) agent reserving the space-time occupancy of the intersection. The space of the intersection is divided into cells, and the reservation is about the desired occupancy of cells along its trajectory for a certain duration. There may be conflicts in some of the reservations, i.e. the same cells are requested at the same time. Then the IM agent needs to decide whether to grant or reject the requested reservations. Different intersection control policies are employed. At each simulated step, the employed policy determines cells the vehicle will occupy using the state of vehicle agents including the time and velocity of arrival, vehicle size, etc. Our goal is to find an optimal control policy that minimises intersection delay.

For the definition of delay of an intersection, we follow the same definition from the seminal work of [Dresner and Stone \(2004\)](#). The number of all vehicles passing through an intersection in a period of time (e.g. one minute) is denoted as C . Assume that there are no other vehicles on the road, vehicle v_i would pass the intersection with the maximally allowable speed in time OT_i . However, due to the presence of other vehicles and the need of avoiding potential conflicts, the actual

travel time of v_i is AT_i . Hence, the intersection delay of an intersection (DoI) is:

$$DoI = \sum_{v_i \in C} (AT_i - OT_i) \quad (2.1)$$

However, the centralised optimisation of reservation-based AIM is ultimately a dynamic traffic assignment problem and thus is extremely challenging in the sense of computational tractability (Levin et al., 2017). In order to avoid the curse of dimensionality in the centralised optimisation, we propose a multi-agent formulation of the AIM problem, where instead of sending reservations to the IM agent and relying on the optimal assignments of the IM agent, vehicle agents cooperate and learn a coordinated cell space-time occupancy policy to minimise the defined delay.

2.3.2 A multi-agent formulation of AIM problem

Upon entering the intersection area, without the control of IM agent, vehicle agents need to coordinate and negotiate with each other about the space-time occupancy of the cells, so as to avoid collisions as the primary condition and minimise the delay. It is a multi-agent decision making problem, which can be modelled as multi-agent Markov decision process (MAMDP): $\langle n, \mathcal{S}, A_1, A_2, \dots, A_n, \delta, r_1, r_2, \dots, r_n \rangle$, where n is the number of agents, $\mathcal{S}$ is the finite set of environment states describing the possible joint-states of all agents, A_i is the finite set of actions available to agents yielding the joint action set $\mathcal{A} = A_1 \times \dots \times A_n$, $\delta : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$ is the state transition function, and $r_i : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathbb{R}, i = 1, \dots, n$ is the reward function.

In the following, the elements of MAMDP are introduced in the context of AIM.

- *States*: states are the information used to determine what happens next. Vehicle agents need to make a decision based on the joint-states of each other. For each agent, the individual state is the combination of its *current position*, *speed*, *moving intention* and the *queue length* of its current lane.

The intersection is divided into cells, where each cell has its own fixed coordinates. The position of a vehicle can be represented by the cells it is occupying. A cell can only be occupied by one vehicle at the same time, and if any part of the cell is occupied, the state of the cell will be regarded as occupied (see Fig.2.1).

FIGURE 2.1: An illustration of cell occupancies

The speed of a vehicle is discretised into the number of cells per time step. The arc length of the turning movements is converted to the length of the cell. This discretisation is commonly used in the simulation of traffic flow using Cellular Automata model and has been demonstrated to capture the characteristics of traffic flow well (Nagel and Schreckenberg, 1992). The speed is related to the length of cell and the step size of the simulation.

Example 2.1. A speed of $20m/s$ is equal to 2 cell/step when the length of a cell is $5m$ and the step size is $0.5s$.

The concept of *desired cells* is proposed to describe the moving intention of a vehicle at each time step which is related to its destination, trajectory and current speed V . Vehicles desire to cross the intersection in the shortest amount of time, i.e. accelerate to or maintain the speed limit and move towards destinations (this behavior is referred to as ‘*keep going*’ in this study). However, for safety reasons, it has to maintain a safe stopping distance, i.e. safety buffer. Therefore, the number of desired cells (*DCs*) ahead of moving direction can be calculated as:

$$DCs = \begin{cases} \max(\lceil V^2/2a \rceil, \lceil V\Delta t + \frac{1}{2}a\Delta t^2 \rceil), & V < V_m \\ \max(\lceil V^2/2a \rceil, \lceil V\Delta t \rceil), & V = V_m \end{cases} \quad (2.2)$$

where a is the acceleration/deceleration value of the vehicle in unit of cell/step^2 , V_m is the speed limit, and $V^2/2a$ is the safe stopping distance needed to avoid

potential collisions. Step size Δt is set close to 0, and it takes at least one step for a vehicle to accelerate to speed limit. If a cell is occupied and desired by the same vehicle, the cell will be regarded as the desired cell of this vehicle.

Example 2.2. For a vehicle with a current speed $V_m = 4 \text{ cell/step}$, and acceleration $a = 1 \text{ cell/step}^2$, its $DCs = 8 \text{ cells}$. For a vehicle with a speed of 3 cell/step , its $DCs = \lceil 4.5 \text{ cells} \rceil = 5 \text{ cells}$.

Example 2.3. Let the intersection depicted in Fig.2.1 be represented as a 24×24 two-dimensional array S_e . Assume that the current speed of vehicle v_3 is $V_m = 4 \text{ cell/step}$ and acceleration $a = 2 \text{ cell/step}^2$, and there are 10 vehicles queueing on lane 11. The individual state (i.e. the position, speed, moving intention and queue length) of v_3 can be represented as:

$$\{S_e[3 : 8][17 : 20], 4 \text{ cell/step}, S_e[9 : 12][17 : 20], 10veh\}$$

- *Actions*: vehicle agents execute actions based on the states they observed. The actions of vehicles are their state of motion. Generally, the state of motion of a vehicle can be expressed in term of its speed ($V_0, \dots, V_m$) and acceleration/deceleration ($(\pm a, 0)$). The choice of an action is constrained by the relationship between speed and acceleration/deceleration and the range of speed: $[V_0, V_m]$.

TABLE 2.1: Action choices of vehicles

Current speed	V_0	$\{V_1, \dots, V_{m-1}\}$	V_m
Action choices	$\{+a, 0\}$	$\{\pm a, 0\}$	$\{-a, 0\}$

The global state and action (S, A) is an aggregation of the individual states and actions (s_i, a_i) .

- *Rewards*: scalar rewards are used to indicate how well agents are doing at time step t in achieving the goal. In order to discourage congestion, our objective is to minimise the intersection delay, so the negative of delay (Eq.2.1) is set as a reward. Assuming that the distance traveled by vehicle v_i at time step t is L_i , then the optimal time required to pass the distance is L_i/V_m , so the reward at time step t can be calculated as:

$$r(S, A) = - \sum_{v_i \in C} (\Delta t - L_i/V_m) \quad (2.3)$$

- *State transition*: the states of vehicles are transited according to the law of motion.

For the cooperative MAMDP of AIM, our objective is to find a joint-policy $\Pi^* : \mathcal{S} \rightarrow \mathcal{A}$ which maximises the total rewards of all agents. Π^* can maximise the long-term discounted reward for all agents, and it can be split into n component policies for each agent $\pi_i : \mathcal{S} \rightarrow A_i$ with $\Pi^* = (\pi_1, \dots, \pi_n)$. One way to find the optimal policy Π^* is by computing an optimal joint-action value function:

$$Q^*(S, A) = \max_{\Pi^*} Q(S, A) \quad (2.4)$$

Theoretically, Eq.(2.4) can be solved through the use of dynamic programming techniques. By iterating, evaluating and calculating value functions, it is possible to find the optimal policies. However, it is almost impossible to compute the joint-action value function for a system like AIM within acceptable execution time because the number of (S, A) pairs in AIM system is enormous. In multi-agent decision making problem, each agent adds its own variables to the joint state-action space which leads to the complexity of solving MAMDPs being exponential in the number of agents. However, since the interaction between CAVs in AIM is sparse and not coupled at every decision step, but only in relatively infrequent situations, with the new MAMDP formulation, we are able to decompose the original centralised optimisation into a distributed one. This leaves us the possibility of using a decentralised learning framework towards solving the AIM problem. In the following, we adapt the decentralised coordination learning (DCL) framework into solving our formulated multi-agent AIM problem.

2.4 Decentralised Coordination Learning

Instead of learning an optimal joint policy, agents in DCL learn and update the coordinated policy. DCL is driven by the fact that many multi-agent systems can exhibit a large amount of context-specific independence so that a general MAMDP problem can be decomposed into sub-problems which are easier to solve (Roth et al., 2007). One such context-specific independent state of a vehicle agent in AIM is that there are no vehicles on the conflicting lanes. The optimal action in this state is easier to determine, that is, the vehicle acts independently, accelerates to or maintains the maximum speed.

The primary task of coordination learning algorithms is to train the agents to learn where and how to coordinate (Yu et al., 2015). At each learning step, the main procedure of coordination learning is shown in Fig.2.2.

FIGURE 2.2: Procedure of coordination learning.

In DCL, the large-scale environment state space is divided into two parts: *coordinated states*, in which agents need to coordinate their actions so that the decision-making process can be improved; *independent states*, in which agents can act independently and coordination with other agents does not improve the decision-making process.

In AIM, for an agent i , its independent states are the states that its desired cells are not occupied and there is no overlap (even a point) with the occupying or desired cells of other vehicles. The rest are the states that need to be coordinated.

Example 2.4. Consider the state depicted in Fig.2.1, in the current state, vehicle v_3 is independent of other vehicles, thus no coordination is needed; vehicle v_1 and vehicle v_2 are in a coordinated state. The overlap of desired cells indicates that collisions may occur if vehicles v_1 and v_2 both take independent actions. Coordination is needed between vehicles v_1 and v_2 .

We define *coordinated joint-actions* as the available actions that can be taken in coordinated states, excluding the joint-actions that would cause any conflicts.

Example 2.5. For a coordination-needed state shown in Fig.2.3, the coordinated joint-actions are Actions 1, 2 and 3, while Action 4 is eliminated. If these two vehicles both choose ‘keep going’ action, a collision cannot be avoided at the next state, because the safe stopping distance cannot be guaranteed.

FIGURE 2.3: A coordination needed state and potential joint-actions involving two CAVs.

Algorithm 1: DCL-AIM

Establish Intersection Environment:

reset and update states

emit rewards and observations

Generate Vehicles: in-flow x for each lane l

Initialise Q – table: Q_i and Q^I

for $episode \leftarrow 1$ **to** $MaxEpisodes$ **do**

for $t \leftarrow 0$ **to** T **do**

 update the number of agents

 update the states

for *each agent* i **do**

 Algorithm 2

return Q – table for each agent i

Based on this idea, an algorithm DCL-AIM (Algorithms 1 and 2) is proposed to solve the decision-making problem of AIM by explicitly identifying and dynamically adapting agent coordination needs during the learning process.

In DCL-AIM algorithm, vehicles in the same direction from the same lane all inherit and contribute to learning and updating the same Q -function. Corresponding to the decomposition of state, Q -function of agent i also has two parts: a single-action value table $Q_i(s_i, a_i)$ and a joint-action value table $Q^I(s_K, a_K)$ for states which require coordination. At independent states, agents will act independently and choose actions based on the optimal Q -function Q_i^* . At coordinated states, CAVs will choose actions from joint optimal Q -function $Q^{I*}(s_K, a_K)$.

Algorithm 2: LEARNING ALGORITHM FOR AGENT i

Inherit $Q - table$: Q_i and Q^I

while *agent i is within the intersection* **do**

get observations

if s_i, s_K *not in* $Q - table.index$ **then**

 append the new state to $Q - table$

choose action (ε -greedy)

if *agent i is involved in coordination* **then**

 choose a coordinated action

else

 choose an independent action

obtain rewards

update $Q - table$

The decomposition relies on the assumption that in any given state, the agents are either independent or can share state information, as such agents can use state information from the other agents in the coordinated states to choose its actions. With the advanced sensors and the V2V communication technologies, each CAV is able to perceive unambiguously its own position in the environment. Also, when agents are simultaneously present in the coordination needed area, namely coordinated states, they are able to sense each other and thus can perceive their joint state clearly.

At time step t , the global reward corresponding to Eq.(2.3) is calculated in Eq.(2.5).

$$r(S, A) = - \sum_{i=1}^{N^t} wt_i \quad (2.5)$$

where, N^t is number of CAVs present at time t . wt_i is the delay of agent i at step t . Since the vehicles are always moving in and out, N^t changes over time.

Correspondingly, when an agent i is in an independent state at time t , its *individual reward* is:

$$r(s_i, a_i) = -wt_i \quad (2.6)$$

Otherwise, when agent i is in a coordinated state and is in subset $K = \{i_1, \dots, i_h\}$, its *joint-reward* is:

$$r(s_K, a_K) = - \sum_{i=1}^h wt_i \quad (2.7)$$

The learning progress begins with an initialisation of the action value functions and followed by the optimal policy learning process. In order to guarantee that the sequence Q_t converges to Q^* , the agents need to keep trying all actions in all states with nonzero probability. The trade-off between exploitation and exploration guarantees the convergence. ε - *greedy* exploration is a widely used approach. At each step, the agent can choose a random action with probability $\varepsilon \in (0, 1)$, and the greedy action with probability $(1 - \varepsilon)$.

Each episode in the algorithm simulates a possible process of the vehicles crossing the intersection, and ends when all CAVs exit the intersection, followed by a reset to the initial state. The DCL-AIM algorithm estimates all the possibilities of how CAVs cross the intersection, and then outputs the best decisions for the actual operation of CAVs.

The main tasks of DCL-AIM algorithm at every time step t are as follows:

- Update the environment.
 - The environment will be updated based on the actions of the current agents and the new agents just arrived. Function *env.update()* returns the new occupancy state of the intersection and *env.agents()* gives out the presence of agents including local states. In addition, the environment will be reset to the same initial state using the function *env.reset()* at the end of each episode.
 - The environment class also includes the functions of rewards: *env.rewards()* and observations: *env.observations()*, which are in charge of emitting rewards and observations to agents based on the environment states resulted from their actions.
- Update the Q-values.

At time step t , if agent i is not involved in any interaction, and moves from an independent state s_i to another independent state s'_i , the independent Q-values can be estimated by the standard Q-learning operation ([Watkins and](#)

Dayan, 1992), shown in Eq.(2.8).

$$Q_{e+1}(s, a) = Q_e(s, a) + \alpha[r(s, a) + \gamma \max_{a' \in A} Q_e(s', a') - Q_e(s, a)] \quad (2.8)$$

where α is the learning rate and specifies how far the current estimate $Q_e(s, a)$ at the e th update is adjusted toward the update target.

Otherwise, let K denote the subset of all h agents interacting with agent i at time step t .

◦ Case 1, agent i moves from a coordinated state s_K to an independent state s_i . The joint Q-values are updated by:

$$Q^I(s_K, a_K) = (1 - \alpha)Q^I(s_K, a_K) + \alpha \sum_{i=1}^h [r_i(s_K, a_K) + \gamma \max_{a'_i} Q_i(s_i, a'_i)] \quad (2.9)$$

◦ Case 2, conversely, agent i moves from an independent state s_i to a coordinated state s_K . The independent Q-values are updated by:

$$Q_i(s_i, a_i) = (1 - \alpha)Q_i(s_i, a_i) + \alpha[r_i(s_i, a_i) + \gamma \frac{1}{h} \max_{a'_K} Q^I(s_K, a'_K)] \quad (2.10)$$

That is, each agent is rewarded with the same fraction of the expected future discounted reward from the resulting coordinated state. This essentially implies that each agent contributes equally to the coordination.

◦ Case 3, agent i moves from a coordinated state s_K to another coordinated state s'_K . The update of Q-values of $Q^I(s_K, a_K)$ also follows the operation shown in Eq.(2.8) according to the joint-rewards.

$QUpdate()$ is used to compactly denote the Q-values update operation described above.

2.5 Simulations and results

DCL-AIM algorithm has three main components: intersection environment, vehicle generation and Q -table learning. The specific details of these three parts in the simulation implementation are described in the following subsections. The results of the simulations are then analysed and presented.

2.5.1 A typical intersection

All simulations are conducted using Python 3.6, and run on a computer with a i7-7700 CPU @3.60GHz, 8.00GB RAM and 64-bit Operating Windows System. A typical intersection with 4 approaches and three incoming lanes of each approach is simulated. The movements of CAVs are set to strictly follow the lane instruction, i.e. the trajectories of vehicles from different lanes are settled.

The control range of the intersection is set as $210m$ within the effective communication range of dedicated short range communication (DSRC) technology (Kenney, 2011). The width of a lane is $3.5m$, the speed limit is set as $7m/s$, the simulation step size is $0.5s$.

In AIM, the intersection area is discretised into cells, the length of the cell represents the level of discretisation, and two different levels of discretisation are simulated, as shown in Fig.2.4.

FIGURE 2.4: Different discretisation levels of intersection area

Definition 2.1 (Discretisation level 1). The intersection area is divided into 6×6 cells, each cell has the length/width of $3.5m$. The length of an CAV is set less than $3.5m$, so the occupancy of a vehicle is at least 1 cell, the specific occupancy is related to its trajectory, such as vehicle v_1 shown in Fig.2.4 (left) is occupying more than one cells. The maximum moving distance of each CAV in each decision step is $7m/s \times 0.5s = 3.5m$ (the length of one cell), the set of actions of each agent is represented as $A = \{0, 1\}$ accordingly. ‘0’ represents the actions of decelerating or

waiting, ‘1’ represents the action of accelerating or keeping going. The arc length of the turning movement is converted to the length of the cell. The allowable acceleration/deceleration is set as $\{\pm 0.5, 0\}(\text{cell}/\text{step}^2)$.

Example 2.6. For the two vehicles in a coordinated state depicted in Fig.2.5, in which the cells are numbered to refer to cell coordinates. At the coordination-needed state $S(t)$, they choose the coordinated actions of $\{a_1 : 0 - \text{decelerating}, a_2 : 1 - \text{keep going}\}$, and move to the next state $S(t+1)$. At $t+1$, parts of cell 2 is occupied by vehicle v_1 ($V_{v_1}\Delta t - 0.5a\Delta t^2 = 0.75 \text{ cell}$), while the rest of cell 2 is desired by vehicle v_1 ($0.5a\Delta t^2 = 0.25 \text{ cell}$). In this case, we still use cell 1 to represent the current position of vehicle v_1 , and cell 2 is used to represent its moving intention. The state of vehicle v_1 at $t+1$ can be represented as $\{\text{cell } 1, 0.5 \text{ cell}/\text{step}, \text{cell } 2, q_1\}$. Similarly, the state of vehicle v_2 : $\{\text{cell } 6, 1 \text{ cell}/\text{step}, \text{cell } 3, q_2\}$.

FIGURE 2.5: Illustration of a sequential decision process and the cell occupancies

Definition 2.2 (Discretisation level 2). The intersection area is divided into 24×24 cells, each cell having the length/width of $0.875m$. The length of an CAV is set as $4.5m$ and width is $1.8m$, so the occupancy of a vehicle is at least 6×3 cells, and the specific occupancy is related to its trajectory, for example, a turning vehicle v_1 in Fig.2.4 (right) is occupying 6×5 cells. The speed limit is equal to $4 \text{ cell}/\text{step}$, then the speed of a vehicle $V \in \{0, 1, 2, 3, 4\}(\text{cell}/\text{step})$. In order to provide better passenger experience, the allowable acceleration/deceleration is set as $\{\pm 1, 0\}(\text{cell}/\text{step}^2)$.

The objective of the decision model of CAVs is to maximise the long-term rewards. During the learning process, the coordination needs are explicitly identified and

dynamically adapted, collision-free is guaranteed by the coordinated joint-actions, and the reward function minimises delay. As a result, the optimal policy given by the model equals the minimisation of the intersection delay and is collision-free.

2.5.2 Vehicle generation

In order to simulate the traffic flow more realistically, vehicles are generated randomly using Poisson distribution. Each lane has different vehicle arriving rate λ , and intersection traffic generating rate is the sum of each lane. A Poisson Process is a kind of Markov Chain process, and has unique characteristics:

- the number of arrivals in non-overlapping intervals are independent, i.e., for every integer $m = 2, 3, \dots$ and time points $t_0 = 0 < t_1 < t_2 < \dots < t_m$, the number of vehicle arrivals of the intervals are independent.
- for any time t and positive number Δt , the probability distribution of the number of vehicle arrivals occurring between t and $t + \Delta t$ depends only on the interval length Δt , and not on the time t .
- as $\Delta t \rightarrow 0^+$, the probability of at least one vehicle arrives, is proportional to the interval length Δt :

$$P\{N((t, t + \Delta t]) \geq 1\} = \lambda \Delta t + o(\Delta t), \Delta t \rightarrow 0^+ \quad (2.11)$$

- the probability of two or more vehicles arrive in an interval of length $\Delta t \rightarrow 0^+$ is negligible:

$$P\{N((t, t + \Delta t]) \geq 2\} = o(\Delta t), \Delta t \rightarrow 0^+ \quad (2.12)$$

With the characteristics described above, we use the probability of at least one vehicle arrives to generate vehicles at each simulation step ($\Delta t = 0.5s$). The generating rule is as following:

$$\text{if } P\{N((t, t + \Delta t]) \geq 1\} > \text{random}(0, 1), \text{ then } v_i \text{ generated} \quad (2.13)$$

where, $\text{random}(0, 1)$ is a pseudo-random number in $(0, 1)$, v_i is the i_{th} generated vehicle at time t .

Three examples of traffic in-flow distribution are shown in Fig.2.6. *ITGR* represents intersection traffic generating rate. *vph* represents the number of arriving vehicles per hour. In this way, the uniform distribution of traffic flow is avoided and more realistic traffic flows are generated.

FIGURE 2.6: Traffic in-flow distribution.

2.5.3 Learning performance

The setting of parameters affects the performance of the *Q-values* learning performance. The parameter setting in the experiment is shown in Table 2.2.

TABLE 2.2: Parameter settings for DCL-AIM algorithm

Greedy policy: epsilon ε	0.9
Learning rate: alpha α	0.05
Discount factor: gamma γ	0.9
Step size: Δt	0.5

The learning rate α affects how quickly the model can converge. However, higher learning rate means higher loss in the learning process, while lower learning rate means the longer convergence runtime. After comparing the performance of $\alpha = 0.01, 0.05, 0.1$, we set $\alpha = 0.05$. An example of the convergence of DCL-AIM algorithm in two discretisation level experiments is shown in Fig.2.7.

As can be seen from Fig.2.7, the episodic reward is well converged after approximately 5000 episodes for each discretisation level experiment, and it also converges

FIGURE 2.7: An example of convergence of DCL-AIM algorithm in two discretisation level experiments.

well in the case of heavy traffic. Note that the fluctuations in reward afterwards are caused by the ε – *greedy* policy.

FIGURE 2.8: Convergence of individual rewards (left), coordination learning (right).

The construction of reward function also has a significant influence on the performance and convergence of DCL-AIM algorithm. Fig.2.8 (left) shows a convergence example (discretisation level 1 experiment) in which the agents update their Q -values only depending on individual rewards in all states, while Fig.2.8 (right) shows an example in which Q -values are updated according to $QUpdate()$. When updating and learning with individual rewards in all states, the algorithm still converges, but it will converge to a poorer policy (less episodic rewards, which means higher delay) compared to the policy converged using $QUpdate()$. The advantage of DCL-AIM algorithm is that it can effectively decompose the difficult decision-making problem into sub-problems without losing too much global performance.

2.5.4 Effectiveness comparison

This section focuses on comparing the performance of FCFS-AIM and the proposed multi-agent coordination learning based AIM (DCL-AIM) after the Q -values are trained and learned, using the data of discretisation level 1 experiment. In addition, Longest-Queue-First policy (LQF-AIM) and the traditional signal control (labeled Signal in Fig.2.10) based on the Webster's method (Webster, 1958) are set as benchmarks.

For all control policies, for the sake of fairness, the traffic in-flow and time horizon are the same. The major road - minor road intersecting situations are simulated. The traffic flow on the major road is twice that of the minor road. And, the through-flow is twice the turning-flow on each approach.

FIGURE 2.9: Uncongested intersection delay comparison.

Different traffic scenarios with different traffic generating rate are simulated. The comparison of intersection delay is shown in Fig.2.9 and Fig.2.10, where *traffic generating rate* is the through-flow rate of the major road. Because the traffic flow in this paper is randomly generated, in order to eliminate the impact of the randomness on the results when comparing the effectiveness of the policies, for each traffic generating rate scenario, the simulations were conducted 100 times. The average delay (s/veh) with 95% confidence interval of the 100 simulation runs are also shown in Fig.2.9 and Fig.2.10.

FIGURE 2.10: Congested intersection delay comparison.

FIGURE 2.11: Fluctuation of intersection delay of 100 simulation runs.

To compare the stability of different methods, we show the results across different simulation runs, as shown in Fig.2.11. Under the same traffic generation rate scenarios, the intersection delay of each simulation is different. Traffic dynamics have the greatest impact on FCFS policy. It is because that the priority of FCFS is assigned based on the arrival time of vehicles. The amplitude of fluctuation shown in Fig.2.11 reflects that FCFS-AIM is not as good as DCL-AIM in dealing with uncertainty. DCL-AIM addresses the uncertainty problem through the trade-off between exploration and exploitation.

The experimental results show that DCL-AIM significantly outperforms FCFS-AIM under different random traffic conditions. *Safety buffer* is an important concept in FCFS-AIM (Dresner, 2009), which is used to enhance the safety of the control. The less buffer the less delay, so we set the extra time safety buffer as 0s in the comparison. In our approach, the *front and rear collisions* are eliminated by the concept of occupying cells and the desired cells, DCL-AIM system does not need additional safety enhancements, the coordinated actions in DCL-AIM algorithm eliminate the conflicting situations.

From the perspective of system optimisation, at a certain time point t , it seems better to give higher priority for the longer queue, which is why the longer queue first (LQF) policy is often applied to intersection management, especially in the adaptive signal control. However, in the long run, it may not be the best choice.

From Fig.2.10 and Fig.2.11, we can see that the Webster’s signal control policy is not dominant in the management of CAVs. The major reason for the higher delay of signal control (even when the traffic flow is relatively light) is that it cannot make good use of the space-time resources of the intersection. In the signal control, only up to two streams can be discharged simultaneously, while the spatiotemporal resources in other directions are “wasted”. Another characteristic of signal control can also be seen from the figures, that is, its standard deviation is relatively small, indicating that the performance of signal control is little affected by the randomness of traffic flow. This is why signal control may be superior to FCFS-AIM in some cases, which can also be seen in Fig.2.10. It is worth mentioning that the signal control compared in this study is not adaptive. The adaptive and optimal signal control for CAVs is another research topic that has been studied extensively in the literature but out of the scope of this study.

As the traffic flow approaches the intersection capacity, the intersection delay increases dramatically, and the effectiveness of all control policies begins to decline, especially FCFS-AIM policy, while DCL-AIM consistently performs the best under all the traffic generating rates. The main reason for the decline in effectiveness is the limited availability of the space-time cells of the intersection.

In order to show how the spatiotemporal resources affect the performance of the control policies, we have designed some scenarios specifically.

Scenario 1. Lane 2 is congested, with the traffic in-flow rate $3000veh/h$, while the arriving vehicles on lane 5, lane 8 and lane 11 are randomly generated as in-flow rate $1500veh/h$. No vehicles are generated on other lanes.

Scenario 2. Lane 2 and lane 8 are congested, with the traffic in-flow rate $3000veh/h$, while the arriving vehicles on other lanes are randomly generated as $1000veh/h$.

Scenario 3. All of Lane 2, lane 5, Lane 8 and Lane 11 are congested, with the traffic in-flow rate $3000veh/h$. No vehicles are generated on other lanes.

For each scenario, we run the simulation 20 times, the averaged intersection delay of the policies is shown in Table 2.3.

TABLE 2.3: Comparison of intersection delays

Scenario	Intersection delay (s)		
	FCFS	LQF	DCL
1	333.5	297.5	111.5
2	3589.6	2752.4	2582.2
3	3585.2	3036.4	2542.1

FIGURE 2.12: Step delay comparison.

For an asymmetrical traffic flow designed in Scenario 1, there are still spatiotemporal resources available in the other three directions, so all the three AIM control policies can manage the intersection well. While, for Scenario 2 and 3 where congestion is developing in all directions, there are relatively less spatiotemporal resources can be used, it is difficult for all three strategies to handle. However, it is promising to see that DCL-AIM algorithm still can produce better control policy.

Fig.2.12 shows the step delay comparison of the three control policies in a 2-minute simulation of Scenario 3. As can be seen, DCL-AIM maintains a better performance at every time step and needs less time to evacuate the congestion.

2.5.5 Results of discretisation level 2

The proposed solution technique is readily applicable to different levels of discretisation. The intersection delay comparison results of discretisation level 2 are shown in Fig.2.13. From the comparison results and trends shown in Fig.2.13, we can draw the same conclusions as the experiment of discretisation level 1, that the proposed DCL-AIM algorithm can learn a better policy and the FCFS-AIM policy is most affected by the randomness of traffic flow.

FIGURE 2.13: Intersection delay comparison of discretisation level 2.

2.5.6 Runtime comparison

After the policy is trained and learnt, the policy obtained by DCL-AIM algorithm can be read and executed in milliseconds for real-world scale problems, just as other rule-based policies (FCFS, LQF and Signal compared in this research). An example of runtime is shown in Fig.2.14. The computational time required for each policy basically increases with traffic flow, but is still much less than the step size.

This shows that all policies can be applied to real world scale problems. As can be seen from the figure, DCL-AIM policy needs a little bit more time to make and execute the decisions. The extra time is spent on the reading of the policy. This time is related to the data storage architecture and the way of reading, and can be further optimised, but it is not the focus of this thesis.

FIGURE 2.14: Average runtime per decision step comparison

It is worth pointing out that the computational effort is mainly required at the policy training period. The runtime of training depends not only on the computing power of the computer used, but also the convergence level desired. There is no training time required for the other policies since they are rule-based.

2.6 Conclusions

AIM is an alternative intersection control tailored for CAVs. With the wireless communication technologies and advanced computer control systems, CAVs have the ability to benefit the traffic system substantially. Although CAVs in AIM system can be programmed with behaviors designed in advance, for example FCFS, it is often necessary that they learn new behaviors online, such that the performance of the agent and the whole multi-agent traffic system gradually improves. In this study, a multi-agent reinforcement learning approach is proposed to learn the optimal policy from the real-time environment states.

We model the crossing processes of CAVs as multi-agent MDPs, in which vehicle agents cooperate to maximise rewards with collision-free constraints. The optimal sequential actions of vehicles can be given by the proposed DCL-AIM algorithm. The main feature of the proposed approach is to explicitly identify and dynamically adapt agent coordination needs during the learning process so that the curse of dimensionality and environment nonstationarity problems in multi-agent learning can be alleviated.

A variety of traffic conditions are simulated within a typical intersection, demonstrating the effectiveness of the proposed model. In order to more realistically simulate the traffic flow, vehicles are randomly generated rather than uniformly distributed. The comparative analysis is conducted between the DCL-AIM and FCFS-AIM, with LQF policy and the Webster's signal control as benchmarks. Results shows that the optimal sequential decisions from DCL-AIM outperform the other control policies.

In this chapter, the traffic efficiency is focused and optimised. Further on, multiple conflicting objectives optimisation will be investigated for AIM in next chapter.

Chapter 3

Multiple Conflicting Objectives Optimisation for AIM

In this chapter, multiple conflicting objectives are optimised for AIM. The problem is modelled as a multi-objective decision process, in which traffic safety is fundamentally guaranteed in each decision making step, and traffic efficiency, fairness, fuel consumption/emissions are taken as the major measurements. Optimised control policies are learnt over multiple conflicting objectives with unknown relative preferences. The standard Bellman optimality equation is extended to a general version to learn a parametric representation for the optimal policies over the space of all possible preferences. Pre-training process is conducted to perform the normalisation of the different reward signals (objectives). The effectiveness of the proposed method is evaluated in a simulated real-world intersection, and compared with three benchmark policies. Better performance of the trained neural network is demonstrated.

3.1 Introduction

With advanced on-board sensors, control systems, Vehicle-to-Everything (V2X) communication technologies, Connected and Automated Vehicles (CAVs) can perceive information with each other and surroundings, including other V2X device equipped road users and traffic control infrastructures ([Molina-Masegosa and Gozalvez, 2017](#)). CAVs have been considered as promising solutions to realise envisioned

autonomous traffic management systems in the future (Wang et al., 2020). Autonomous intersection management (AIM) is one of the envisaged systems tailored for CAVs. The original AIM concept was proposed by Dresner and Stone (2004), in which the traditional signal control is replaced by a reservation-based control system using the communication ability of CAVs. CAVs are seen as autonomous agents which would communicate with the intersection manager (IM) agent and reserve the spatial-temporal occupancy of the intersection. The IM agent accepts or rejects reservations in accordance with control policies, with *collision-free* as the primary principle. First-Come-First-Served (FCFS) is the most commonly used rule because of its straightforward and fair features (Dresner and Stone, 2008), referred as FCFS-AIM. However, although traffic fairness is guaranteed in FCFS-AIM, it is often unsatisfactory in terms of traffic efficiency and may cause more delays than signal control under higher traffic flow (Levin et al., 2016). In addition, as a predefined control policy, the performance of FCFS-AIM is seriously affected by the dynamics of traffic flow compared to other control policies, such as Longer-Queue-First (LQF) policy (Wu et al., 2019). In the follow-up development, other rule-based policies were proposed, such as platoon based AIM (Jin et al., 2013). In platoon-based control, represented as Platoon-AIM, CAV agents can proactively form groups and traverse the intersection using the platoon-based reservation scheme, that is, the IM assigns the right-of-way to platoons instead of individual vehicles as in FCFS-AIM. It was asserted that leveraging the platooning capability of CAVs in AIM could improve the efficiency of any policy at an intersection, in terms of average delay reduction (Bashiri and Fleming, 2017). However, Timmerman and Boon (2021) showed that platoon based policies with lower average delays sometimes are relatively unfair, indicating a potential need for balancing average delay and fairness.

A key motivation of this study is to provide optimal policies for IM agent, which fully incorporate the real-time traffic dynamics, and comprehensively consider the conflicting objectives of traffic fairness and efficiency. Traffic efficiency is measured by average travel time, emissions and fuel consumptions. Observing the stochastic real-time traffic states, the IM agent needs to decide which actions to take to optimise the multiple objectives. This is modelled as a Multi-Objective Markov Decision Process (MOMDP). We tackle two concrete challenges in finding the optimal decisions: (1) the objectives are essentially conflicting; (2) the relative preferences of the multiple objectives are unknown. An Multi-Objective Reinforcement

Learning (MORL) algorithm is customised for optimisation, in which the dynamic environment is formally described by the modelled MOMDP. The proposed algorithm is evaluated in a simulated real-world isolated intersection. Its performance in terms of traffic fairness and efficiency is compared with three benchmark policies.

The rest of the chapter is structured as below. Section 3.2 briefly reviews related work. Section 3.3 is the problem statement, detailing the multi-objective formulation of AIM problem. Section 3.4 introduces the solving algorithm and training process. Section 3.5 presents the details of simulation implementations, the performance of the solution methods and the analysis results of the effectiveness comparison. Lastly, section 3.6 concludes the chapter and identifies the direction of future work.

3.2 Related work

The management of CAVs at isolated intersections has experienced the development from reservation to optimisation based methods (Yu et al., 2019). The AIM mentioned hereafter in this chapter refers to the use of CAV features for intersection management in a general sense, rather than specifically referring to the framework proposed by Dresner and Stone (2004). For single-objective optimised AIM, maximising capacity/throughput (Sun et al., 2017; Mirheli et al., 2018), minimising average travel time/delay (Zhu and Ukkusuri, 2015; Wu et al., 2019), reducing fuel consumptions/emissions (Zhang et al., 2016) are the most common optimisation goals. Specifically, a decentralised energy-optimal control framework for CAVs was proposed by Malikopoulos et al. (2018), in which substantial dual benefits of saving energy and fuel consumption while improving travel time were illustrated through simulations. In Han et al. (2020), platoon trajectory was optimised with the objective of reducing CAVs' fuel consumption, and the same dual benefits of fuel saving and travel time reduction were demonstrated. It is showed that reducing travel time and minimising fuel consumption for CAVs at intersections are not conflicting.

As for the multi-objective optimisation of AIM, it is basically a combination of the above mentioned objectives. In Wang et al. (2018), guaranteeing safety, alleviating traffic congestion and improving the performance of fuel consumption were

optimised for the cooperative driving for CAVs at non-signalised intersections; Aiming at improving travel efficiency, fuel consumption and safety at signalised intersection, [Zhou et al. \(2019\)](#) proposed a reinforcement-learning-based real-time car following model for CAVs to obtain an appropriate driving behaviour. Further on, taking advantages of CAV capabilities, vehicle trajectories, signal timing and phasing were jointly optimised in [Guo et al. \(2019\)](#), in which combined dynamic programming and shooting heuristic approach are used. Sound performance with high efficiency and significant benefits in fuel consumption savings were illustrated by comparing to the adaptive signal control. [Zhao et al. \(2019\)](#) proposed a decentralised multi-objective control algorithm based on model predictive control for CAVs, in which performances of safety, traffic efficiency, fuel consumption and riding comfort were improved at intersection area. Specifically, passenger discomfort and energy consumption were jointly minimised at signal-free intersections in [Zhang and Cassandras \(2019\)](#), and the effectiveness of the proposed decentralised optimal control was illustrated by extensive simulations. In [Kumaravel et al. \(2020\)](#), the problem of coordinating platoons of CAVs at signal-free intersection was addressed by a decentralised, two-level optimal framework with the objectives to minimise travel delay and fuel consumption of every platoon. It is worth noting that platoon based optimisation is likely to compromise the traffic fairness.

The latest systematic review of intersection management of CAVs can be found in [Khayatian et al. \(2020\)](#), in which the existing intersection control policies for CAVs is grouped into three main categories: i) First-Come First-Served, ii) Heuristic, and iii) Optimisation-based. FCFS order, i.e. the crossing order of vehicles is scheduled by the approaching order, is commonly seen as the fairest policy ([Wierman, 2011](#)). Using this logic, if the crossing order of a vehicle co_v is greater than its approaching order ao_v , unfairness is resulted. Corresponding unfairness of the three categories are illustrated in Fig.3.1. From the figure, existing optimisation-based approaches try to optimise traffic efficiency related goals regardless of the approaching order, causing a potential disparity in the crossing and arriving order of vehicles, leading to unfairness as a result. It can be seen that traffic fairness is rarely regarded as an optimisation objective of AIM, and it may compromise traffic efficiency. In this study, the scenarios where FCFS-AIM strategy will cause congestion are focused, and we tackle the challenge of optimising multiple conflicting objectives for AIM. MORL framework is applied to solve the modelled problem.

FIGURE 3.1: Examples of FCFS, optimisation-based and heuristic scheduling policies (Khayatian et al., 2020). In the right row, the value of unfairness obtained from this study is shown.

MORL has relatively few applications in intersection management and is mainly used for signal optimisation (Yin et al., 2010; Khamis and Gomaa, 2014; Mannion et al., 2016). The latest related research is the adaptive traffic signal control optimisation in Gong et al. (2020), where traffic efficiency and safety are optimised simultaneously. The waiting time of vehicles measures the traffic efficiency, and a crash risk model calibrated based on the local historical crashes indicates the traffic safety. These two objectives are mutually independent, while one is microscopic, the other is macroscopic. Weighted sum approach is applied for the optimisation, and the weight (preference) of the two objectives is fixed at 0.5 each. Note that if preference is fixed to a single value, the multi-objective optimisation problem could be downgraded to a single-objective optimisation problem. We tackle another challenge in multi-objective optimisation: the relative preferences of objectives is unknown.

3.3 Problem statement and modeling

3.3.1 A multi-objective formulation of AIM problem

The management of road intersection is essentially the assignment of right of way, that is the IM agent needs to decide how to assign the crossing order of approaching vehicles to achieve the goals of optimising traffic efficiency and fairness. It is a multi-objective decision making problem, which can be modelled as a multi-objective Markov decision process (MOMDP): $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathbf{r}, \Omega, f_\Omega \rangle$, where $\mathcal{S}$ is the state space, $\mathcal{A}$ is the finite set of actions, $\mathcal{P}(s' | s, a)$ is the state transition function, $\mathbf{r}(s, a)$ is the vector reward function, Ω is the space of preferences, and preference functions $f_\omega(\mathbf{r})$ produces a scalar utility using preference $\omega \in \Omega$. The goal of a MOMDP is to recover a set of optimal policies $\Pi_{\mathcal{L}} : \mathcal{S} \rightarrow \mathcal{A}$ for the entire convex coverage set of the Pareto frontier, which depends on the *relative preferences* ($\omega \in \Omega$) of the objectives. In this study, linear preference function is considered:

$$f_\omega(\mathbf{r}(s, a)) = \omega^\top \mathbf{r}(s, a)$$

For each $\pi \in \Pi_{\mathcal{L}}$, there exists at least one linear preference ω such that no other policy π' generates higher utility under that ω :

$$\pi \in \Pi_{\mathcal{L}} \implies \exists \omega \in \Omega \text{ s.t. } \forall \pi' \in \Pi, \omega^\top v^\pi(s_0) \geq \omega^\top v^{\pi'}(s_0)$$

where s_0 is a fixed initial state, and v^π is the value function, i.e., $v^\pi(s) = \mathbb{E}_\pi[\hat{\mathbf{r}} | s_0 = s]$. Given any preference ω , $\Pi_{\mathcal{L}}(\omega)$ determines the optimal policy.

3.3.2 State

In this chapter, the states are the real-time traffic situations within the control range of intersection. For each route i , its state includes the number of queued vehicles q_i and the approaching order ($ao_i^n, n \in q_i$) of each vehicle. Worth mentioning that the approaching order of vehicles is sorted by all vehicles from all the routes in the intersection. For an intersection with L routes, that state at step k has the

following form:

$$S_k = \begin{bmatrix} q_1^k & ao_{1,k}^1 & ao_{1,k}^2 & \cdots & \cdots & ao_{1,k}^{q_1^k} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ q_i^k & ao_{i,k}^1 & ao_{i,k}^2 & \cdots & ao_{i,k}^{q_i^k} & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ q_L^k & ao_{L,k}^1 & ao_{L,k}^2 & \cdots & \cdots & \cdots & ao_{L,k}^{q_L^k} & 0 \end{bmatrix} \quad (3.1)$$

where, element 0 is used to ensure the same dimension. For an environment of CAVs, these states are easily accessible.

3.3.3 Action

The intersection manager needs to decide which action to take based on the current state S_k . An action is represented by a binary vector with L elements; the i th element (a_i^k) corresponds to route i , and is equal to 1 if route i is chosen to be discharged at time step k and to 0 otherwise.

$$a_k = [a_1^k \ a_2^k \ \cdots \ a_i^k \ \cdots \ a_L^k] \quad (3.2)$$

The actions can be chosen (the action space $\mathcal{A}$) are constrained as:

$$1 = a_i^k \perp a_j^k = 1, \forall i \in L, \forall j \in L_i^c \quad (3.3)$$

where L_i^c is the set of routes that have potential conflicts with route i . This complementarity constraint is to ensure the *collision-free* rule of intersection management, that is, routes with potential conflicts cannot be discharged at the same time. With this setting of the actions, the learned policy is not limited to fixed phase setting as in most cases of adaptive signal control, and is more accountable for traffic dynamics.

3.3.4 Rewards

Comprehensive performance of the IM is expected. Safety is guaranteed in action selection, and several other key goals are set as reward signals for learning. For

traffic efficiency, the goal is to minimise the travel time of vehicles.

$$r_{tt}^k = - \sum_{v \in V_k} T_v \quad (3.4)$$

For traffic fairness, the goal is to minimise the unfairness caused by the discharge policy of the IM:

$$r_{uf}^k = - \sum_{v \in V_k} Uf_v \quad (3.5)$$

where uf_v is the unfairness of vehicle v calculated from the difference between crossing order co_v and approaching order ao_v of vehicle v :

$$Uf_v = \begin{cases} co_v - ao_v, & \text{if } co_v > ao_v; \\ 0, & \text{otherwise.} \end{cases} \quad (3.6)$$

Because it is uncertain whether the power source of the CAVs is electric or fuel, the objective function of reducing fuel consumption and emissions is modified to minimise the number of idling vehicles in front of the intersection at each step, which is directly related to the reduction of fuel usage and emission:

$$r_{ef}^k = - \sum_{i \in L} H_i^k \quad (3.7)$$

where H_i^k is the number of halting vehicles on route i at step k .

3.4 Solving algorithm and process

3.4.1 Generalised Bellman equation for MORL

Learning policies over multiple preferences of MOMDP is challenging. A widely used strategy is to convert the multi-objective problem into a single-objective one through various techniques (Rădulescu et al., 2020). *Weighted sum approach* (Kim and De Weck, 2006) is one of the techniques, which linearly weighted sums the *Q-values* (the expected return of a state-action pair given the policy) for all the m

objectives to obtain a synthetic Q function:

$$Q_i(s, a) = \mathbb{E}\left\{\sum_{t=0}^{\infty} \gamma^t r_{k+t+1}^i \mid S_k = s, A_k = a\right\}$$

$$SQ(s, a) = \sum_{i=1}^m w_i Q_i(s, a)$$

With the synthetic Q function, the standard Q-Learning algorithm for single-objective RL (Watkins and Dayan, 1992) can be applied to learn the optimal action value function $Q^*(s, a)$, which satisfies the Bellman optimality equation (Bellman, 1966) :

$$Q^*(s, a) := r(s, a) + \gamma \mathbb{E}_{s' \sim \mathcal{P}(\cdot | s, a)} [\sup_{a' \in \mathcal{A}} Q(s', a')] \quad (3.8)$$

where $\sup_{a' \in \mathcal{A}} Q(s', a')$ is an optimality filter over the Q-values for the next state s' . γ is a discount factor. *Q-Learning* turns the Bellman optimality equation into an iterative approximation procedure and estimates Q^* from transitions consisting of *state-action-reward-next state* experienced in the learning process. However, this approach is only sufficient when the weights are known beforehand.

Let $\mathcal{Q} \subseteq (\Omega \rightarrow \mathbb{R}^m)^{\mathcal{S} \times \mathcal{A}}$ be the value space including all bounded functions ($Q(s, a, \omega)$), $Q^* \in \mathcal{Q}$ be the desired optimal value function in the value space, as shown below:

$$Q^*(s, a, \omega) = \arg_Q \sup_{\pi \in \Pi} \omega^\top \mathbb{E}_{\tau \sim (\mathcal{P}, \pi) | s_0=s, a_0=a} \left[\sum_{k=0}^{\infty} \gamma^k r(s_k, a_k) \right] \quad (3.9)$$

It can be proven that $Q^* = \mathcal{T}Q^*$, see Yang et al. (2019).

With the extended operator, the constructed MOMDP for AIM can be solved by some classic deep reinforcement learning techniques (Mnih et al., 2013, 2015). A deep neural network with s, ω as input and $|\mathcal{A}| \times m$ Q-values as output is used to approximate the desired value space $\mathcal{Q} \subseteq (\Omega \rightarrow \mathbb{R}^m)^{\mathcal{S} \times \mathcal{A}}$ by minimising the following loss function at each step k :

$$L^A(\theta) = \mathbb{E}_{s, a, \omega} [\| \mathbf{y} - Q(s, a, \omega; \theta) \|^2] \quad (3.10)$$

where $\mathbf{y} = \mathbb{E}_{s'} [r + \gamma \arg_Q \max_{a, \omega'} \omega^\top Q(s', a, \omega'; \theta_k)]$ is the target for iteration k . In practice, it is difficult to directly optimise L^A because there are many discrete

solutions in the optimal frontier, which makes the landscape of the loss function quite unsmooth. Homotopy optimisation method is applied to address this concern by combining L^A with an auxiliary loss function L^B (Watson and Haftka, 1989):

$$\begin{aligned} L^B(\theta) &= \mathbb{E}_{s,a,\omega} [|\omega^\top \mathbf{y} - \omega^\top \mathbf{Q}(s, a, \omega; \theta)|] \\ L(\theta) &= (1 - \lambda) \cdot L^A(\theta) + \lambda \cdot L^B(\theta) \end{aligned} \quad (3.11)$$

where λ is a weight to trade off between losses L^A and L_k^B . The value of λ is increased slowly from 0 to 1 during the training. The combined loss function is effective, since the prediction of $\mathbf{Q}$ is firstly ensured by L^A to be close to any real expected total reward (although it may not be optimal), and the direction with better utility is guided by L^B .

The network $\mathbf{Q}_\theta$ is trained and updated by descending its stochastic gradient according to the combined loss function:

$$\nabla_\theta L(\theta) = (1 - \lambda) \cdot \nabla_\theta L^A(\theta) + \lambda \cdot \nabla_\theta L^B(\theta) \quad (3.12)$$

3.4.2 Pre-training

The goal of this study is to comprehensively consider the safety, fairness, efficiency and emissions/fuel consumption for AIM, where safety is fundamentally ensured by action selection at each decision-making step, while the other objectives indicated by the rewards need to be achieved through iterative exploration and learning. However, the dimensions of the three rewards are different, scalarisation is needed. In order to perform the de-dimension, the *Min-Max normalisation* is applied (Patro and Sahu, 2015). Pre-training is designed to get the maximum and minimum values of each reward.

In the pre-training process, the preference ω is set as $[1.0, 0.0, 0.0]$ for travel time, fairness and emissions/fuel consumption. That is, a single-objective of minimising travel time is learnt for 3000 episodes with original rewards, and the three rewards $[r_{tt}, r_{uf}, r_{ef}]$ of each step are recorded to get the estimated range of the rewards values ($[r_{min} \ r_{max}]$).

In the formal training, normalised rewards (r_{nor}) are used:

$$r_{i,nor} = \begin{cases} 0, & \text{if } r_{i,original} = 0; \\ \frac{r_{i,original} - r_{i,min}}{r_{i,max} - r_{i,min}}, & \text{otherwise.} \end{cases}; \quad r_i \in [r_{tt}, r_{uf}, r_{ef}]$$

FIGURE 3.2: Training with travel time

In addition, the performance of the pre-training is shown in Fig. 3.2. It can be seen that although only travel time is used as a reward signal for learning, reward of fuel consumption/emissions has the same converging trend. The dual benefits of improving traffic efficiency and reducing fuel consumption/emissions are corroborated, as stated in literature review (Malikopoulos et al., 2018; Han et al., 2020). With this prior knowledge obtained from literature and verified by the pre-training, reward of emission/fuel consumption is no longer trained during the formal training to save learning time.

3.5 Case Study

The proposed method was tested in a simulated intersection. SUMO traffic simulator is used as the environment server, and TraCI interface is used to provide access to SUMO and connections with the algorithms (Lopez et al., 2018). Learning algorithm is implemented using Python. All the training and simulations are conducted on a computer with a i7-7700 CPU @3.60GHz, 8.00GB RAM and 64-bit Operating Windows System.

3.5.1 Simulation set up

The simulation scenario was built based on a real intersection of Pan Island Expressway (PIE) intersecting with Pioneer Rd North road in Singapore. We focus on the motorised traffic, so the intersection is simplified (left-turning routes are not considered as they have no conflicts with other motorised routes), shown in Fig.3.3. *Route_SN* is the route on which vehicles travel from south to north, *Route_SH* is the route where vehicles travel from south and merge into the highway, and so on in the same manner.

FIGURE 3.3: Simulated intersection

Cooperative adaptive cruise control (CACC) is adopted to simulate the car-following behavior of CAVs (Milanés and Shladover, 2014), in which the safety headway of CAVs $h_c = 1s$, and the driver imperfection $\sigma = 0$. The default values in SUMO are used for other parameters. Under this setting, the saturation flow rate is $3600veh/h$. In the simulation experiments, the traffic flow of the highway-exiting routes (*Route_HS* and *Route_HN*) is set to the 80% saturation flow level

(2880veh/h), while the traffic flow rate of the other routes is 1000veh/h, so that the scenario where FCFS-AIM loses its effectiveness in traffic efficiency can be generated.

3.5.2 Implementation details

We implement the multi-objective Q -networks with 2 convolutional layers and 2 fully connected layers. The convolutional layers are used for feature extraction from traffic state input. The extracted features are then concatenated with preferences, and fed into the two-layer fully connected networks for output. The diagram of the learning algorithm is shown in Fig.3.4.

FIGURE 3.4: Structure of the neural network used in the learning algorithm

In the training process, *double Q-learning* (Van Hasselt et al., 2016) technique is employed to yield more accurate value estimates, in which the *target network* has the same architecture as the function approximator *Q-network* but with frozen parameters; and for every C iterations, the parameters from the Q -network are copied to the target network. Target (y_i) for iteration i is provided by the target network to the loss function. C is set to 100 in the training process. *Prioritised experience replay* (Schaul et al., 2015) technique is employed to speed up the training process, which can increase learning speed with mini-batches. The replay memory

size is 20000 and the batch size is 32. An Adam optimiser is used for updating the parameters after at least one batch experiences is stored in reply buffer, with an initial learning rate $lr = 0.001$ (Kingma and Ba, 2014).

The control range of the intersection is set as $210m$ within the effective communication range of dedicated short range communication technology (Kenney, 2011). The vehicle standing space is $7.5m$ (vehicle length $5m + \text{minGap } 2.5m$ defaulted in simulation setting), hence the number of queued vehicles within the communication range cannot be greater than 29. Combining the definition of states in Eq.(3.1), the number of columns of traffic states is set to 30. The action size is 7, because there are 7 different binary vectors that satisfy the complementarity constraint in Eq.(3.3) for the simulated case:

$$\mathcal{A} = \begin{cases} a_1 : [1 \ 1 \ 0 \ 0 \ 0 \ 0]; & a_2 : [1 \ 1 \ 0 \ 0 \ 0 \ 0]; & a_3 : [1 \ 0 \ 0 \ 1 \ 0 \ 0]; \\ a_4 : [0 \ 1 \ 0 \ 0 \ 1 \ 0]; & a_5 : [0 \ 0 \ 1 \ 0 \ 0 \ 1]; & a_6 : [1 \ 0 \ 0 \ 1 \ 0 \ 0]; \\ a_7 : [0 \ 0 \ 0 \ 1 \ 0 \ 1]; & a_k : [a_{SN} \ a_{SH} \ a_{HN} \ a_{NS} \ a_{NH} \ a_{HS}]. \end{cases}$$

3.5.3 Comparison and discussion

Multiple conflicting objectives are optimised in this study, and we use *MCO-AIM* to represent the trained policies. After learning, the agent is provided with specified preferences of the objectives to output the total returns. Three sets of preferences of fairness (w_f) and travel time (w_t) are given to the agent: $[w_f \ w_t] = [0.2 \ 0.8], [0.5 \ 0.5]$ and $[0.8 \ 0.2]$, represented as the cases of $w_f = 0.2, 0.5, 0.8$, respectively. The performance of the trained agent is compared with the three previously introduced policies:

- **FCFS-Individual-Vehicle (F-AIM).** The passing rights are assigned by the individual approaching order of vehicles. Whoever comes first can get to cross the intersection first.
- **FCFS-Platoon-Vehicles (P-AIM).** At every decision making step, the leader of queueing vehicles with the earlier arrival time gets the passing priority, and the following vehicles in the same platoon also get the passing right.

- **Longer-Queue-First (LQF).** In this policy, the passing right is not related to approaching order, and the longest queue gets the priority.

FIGURE 3.5: Performance comparison

The performance comparison is shown in Fig.3.5.

FIGURE 3.6: An illustration of unfairness in FCFS-AIM

First, we use the following case to illustrate why there is unfairness in FCFS-AIM. At time t , there are three vehicles v_1, v_2, v_3 in front of an intersection, see Fig.3.6. Vehicles v_1 and v_3 are from two non-conflicting routes, while v_2 has potential conflicts with both of them. The approaching order of v_1, v_2, v_3 is 1, 2, 3. According to

FCFS policy, vehicle v_1 has the right to cross the intersection first, and vehicle v_3 can cross at the same time since there are no conflicts. Then the crossing order of v_1, v_2, v_3 is 1, 3, 2, unfairness is caused for vehicle v_2 . Nevertheless, in line with its essence, FCFS-AIM is still the fairest with the lowest unfairness index, as seen in Fig.3.5 (the table). However, it is evident from the table that the FCFS-AIM policy performs poorly in terms of traffic efficiency, as it has the longest average travel time. The performance of Platoon-AIM and LQF further indicates the trade-off between traffic fairness and travel time. From the three adapted cases, it is shown that the trained MCO-AIM can improve efficiency to a large extent while ensuring relative fairness, hence the goal of balancing efficiency and fairness is achieved. In addition, the convexity of the solutions is also reflected in Fig.3.5.

The performance is further analysed in terms of average delay, and the delays of vehicles on different routes are compared.

TABLE 3.1: The average delay (s)

	F-AIM	P-AIM	LQF	$w_f = 0.5$
<i>Route_SN</i>	204.98	163.08	177.26	174.40
<i>Route_SH</i>	206.22	188.81	178.68	177.71
<i>Route_HN</i>	123.94	120.03	72.33	83.37
<i>Route_NS</i>	662.74	192.33	234.83	128.47
<i>Route_NH</i>	364.49	183.22	198.62	122.92
<i>Route_HS</i>	123.57	136.63	78.41	76.17
Overall average	225.28	151.58	125.81	110.08

The radar chart in Fig.3.7 illustrates the pros and cons of each strategy. It can be seen that our trained agent performs best in terms of overall average delay, while FCFS-AIM performs the worst. For vehicles on the highway-exiting routes (*Route_HS* and *Route_HN*) where longer queues are generated, LQF is the best policy in terms of efficiency. It can be clearly seen that less delay on the major roads in LQF comes at the cost of a longer waiting time on the minor roads. In our trained agent, since traffic fairness is comprehensively considered, the passing right of vehicles on the major roads is transferred to the minor roads to a certain extent, so that the delay of vehicles on each route tends to be balanced. An outstanding performance of the proposed method in balancing fairness and efficiency is achieved. The specific values for delays in Fig.3.7 can be found in Table 3.1.

Fairness is an amorphous concept that is impossible to define in a universal way. In existing studies, traffic fairness is usually measured by parameters related to traffic

FIGURE 3.7: The comparison of the average delay

TABLE 3.2: Comparison of defined fairness and delay

	F-AIM	P-AIM	LQF	$w_f = 0.2$	$w_f = 0.5$	$w_f = 0.8$
Unfairness	8.82	11.14	18.12	12.41	10.41	9.84
Average of delay	225.28	151.57	125.81	107.34	110.08	113.89
Stdev* of delay	393.75	109.94	97.90	70.16	77.33	77.71

* standard deviation

efficiency, such as the standard deviation of trip times (Lu and Kim, 2016) and delays (Bashiri and Fleming, 2017) of all crossed vehicles: the lower the standard deviation, the higher the traffic fairness. The defined fairness and the standard deviation of vehicle delay is presented in Table 3.2. It can be seen that FCFS-AIM has the highest standard deviation of delay for the simulated case because

it cannot account for the real-time traffic dynamics. While the proposed MOC-AIM has better performance in terms of the defined fairness, delay and standard deviation of delay.

TABLE 3.3: Comparison of average vehicle emission (g/veh)

	F-AIM	P-AIM	LQF	$w_t^* = 0.8$	$w_t = 0.5$	$w_t = 0.2$
Average of delay	225.28	151.57	125.81	107.34	110.08	113.89
Average of emission	27.44	20.33	17.15	12.31	13.19	13.82

* preference of travel time

As explained, the rewards of emissions are omitted in the formal training because of the dual benefits of travel time and emission reduction revealed during the pre-training process and by some existing papers. To confirm the consistency of the final results, the average vehicle emission (carbon monoxide, CO) under each policy is compared, as shown in Table 3.3. Carbon monoxide (CO) is chosen to represent vehicle emissions by assuming CAVs are fuel powered.

FIGURE 3.8: The relationship between average delay and emission

It can be seen from Table 3.3 and Fig.3.8 that the relationship between average delay and emission is close to a linear correlation. Our omitting of emission rewards is reasonable. To a certain extent, this also shows that it is not necessary to consider delay reduction and emission reduction as two different optimisation goals.

3.6 Conclusions

Utilising the advanced features of CAVs to further improve the performance of intersection management (AIM) in multiple conflicting-objectives is the focus of this study. The control agent observes real-time traffic state and selects *collision-free* actions to guarantee traffic fairness and improve traffic efficiency. A multi-objective reinforcement learning framework using a generalised version of the Bellman equation is applied to learn the optimal policies over the space of all possible preferences of the conflicting objectives. Convolution neural networks are used for feature extraction from the traffic state input, prioritised experience reply and double Q-learning techniques are employed to speed up the training process.

Traffic fairness is typically defined as the difference between the approaching order and crossing order of vehicles, and FCFS is considered the fairest policy for AIM. To the best knowledge of the authors, it is the first time that the conflicting efficiency and fairness are comprehensively considered for AIM policies, with unknown preferences of the objectives. The effectiveness of the trained agent is compared to three benchmark policies, and better performance in balancing traffic fairness and efficiency is achieved as expected.

At current development, homogenous traffic flow is considered. Studying the right-of-way assignment problem of heterogeneous traffic is the focus of future work, including the mixing of CAVs with prioritised buses, with HVs, or even with non-motorised traffic flow. The proposed approach can easily be transplanted to the situation of heterogeneous traffic flow. The key lies in the acquisition of real-time traffic states, which involves the effects of different CAV market penetration rates. Using partial CAV information to infer the overall state of mixed traffic flow is also one of the future research directions. The impact of CAV features on intersection capacity under mixed CAV-HV traffic is preliminarily discussed in Chapter 5.

Part II

Trajectory Organisation Comparison

Chapter 4

Comparative Study on Roundabout and Intersection Management under CAV Environment

Connected and automated vehicles (CAVs) are envisioned to be the new generation vehicles. Using communication technologies, CAVs are able to mutually negotiate the right of way at road junctions. The concept of signal-free management for CAVs came into being, in which the conventional signal control is replaced by various right of way negotiation mechanisms. In most proposed strategies, despite the replacement of traffic lights, the organisation of vehicle trajectories remains the same as that of traffic lights. However, as a naturally signal-free strategy, roundabout do not get enough attention. The significant difference lies in the organisation of vehicle trajectories. A key motivation of this study is to theoretically compare the performance of signal-free intersection management and roundabout under CAVs environment, and to demonstrate whether roundabout has comparable performance to signal-free intersection.

4.1 Introduction

The advanced development of automation and wireless communication technologies enables the new generation vehicles: Connected and Automated Vehicles (CAVs). It is estimated that CAVs will fully enter the global market by 2040s (IHSMarkit, 2017). How to prepare for the opportunities and challenges brought by CAVs has become a popular research topic (Fagnant and Kockelman, 2015), including the new challenges of the corresponding traffic flow management under CAVs environment. As the crux of traffic network, the intersection management for CAVs has been researched (Chen and Englund, 2016). With the ability to communicate with each other and between infrastructures, CAVs can negotiate the right-of-way at intersections and coordinate their movements without the indications of signals. Based on these characteristics, Autonomous Intersection Management (AIM) is customised for CAVs. In AIM (Dresner and Stone, 2004), the intersection area is divided into cells, CAVs need to send requirement to reserve the spatial-temporal occupancy of cells from the Intersection Manager (IM). With *collision-free* as the primary principle, the reservations would be accepted or rejected in a First-Come-First-Serve manner (FCFS-AIM). Based on FCFS-AIM, variant *signal-free* intersection management approaches for CAVs have been proposed and studied.

However, the vast majority of the signal-free management for CAVs are based on the setting that the vehicle trajectories are intersection-organised, while research based on circular organisation, i.e. roundabout, is far less. The two organisations at a road junction are depicted in Fig.4.1.

FIGURE 4.1: Intersectional and circular organisations at a road junction.

As a typical signal-free management, roundabouts have high safety standards, economic and social benefits (Mauro and Cattani, 2012). For human-driving vehicles

(HVs), the respective applicable scenarios and conditions for roundabouts and intersections have been clearly defined (Kimber, 1980; Veith, 2010; HCM, 2016), and the state of the art calculation method of capacity can be found in (Mauro and Pompigna, 2020; Pompigna et al., 2020). However, the situation may be different when it comes to the era of CAVs. This study attempts to examine the performance of different traffic management strategies (with intersectional organisation and circular organisation) for CAVs under the same traffic demand at the same road junction.

4.1.1 Related work

How to leverage the advantages of CAVs to further improve the traffic management at road junctions has attracted extensive research attention. Traffic flow can be improved in terms of traffic stability, safety and highway capacity with more vehicles participating in CACC platoons (Van Arem et al., 2006).

For the management of CAVs at road junctions, some researchers have proposed the concept of *signal-free* intersection management (Malikopoulos et al., 2018), which can be traced back to Autonomous Intersection Management (AIM) proposed by Dresner and Stone (Dresner and Stone, 2008).

The effectiveness of AIM using FCFS policy in terms of delay reduction under certain traffic conditions has been demonstrated (Hausknecht et al., 2011; Fajardo et al., 2011). In order to eliminate collisions and ensure safety, the potential conflict points are the most critical consideration in any signal-free strategy. Conflict points are modelled as the core constraints in the total travel time minimisation formulation for AIM (Zhu and Ukkusuri, 2015), intersection capacity and throughput maximisation (Sun et al., 2017; Mirheli et al., 2018), energy consumption minimisation (Malikopoulos et al., 2018), etc. More conflicts imply more constraints and more computational efforts. The conflict points within an intersection are aggregated into conflict regions to improve computational efficiency in Levin et al. (2017). For a typical road junction with 4 approaches and 3-in-3-out lanes on each approach using intersectional organisation, there are at least 16 conflict points need to be carefully modelled to avoid the potential conflicts. However, if the junction is organised as roundabout, the number of conflict points reduces to 4, as shown in Fig. 4.1 (Diverging conflict points are not counted as the diverging conflicts can

be easily avoided). Fewer conflict points bring the circular organisation substantial safety benefits, which is why the *roundabout first* policy has been adopted in some States of the U.S. (NYSDOT, 2018). In addition, fewer conflict points require fewer modelling constraints and less computational efforts. From the perspective of fewer conflict points, intersectional organisation may not necessarily outperform roundabout organisation under CAVs environment.

However, the research of roundabouts under CAVs environment is rare, and limited to merging control. Such as, using model predictive control to merge CAVs at roundabouts (Nor and Namerikawa, 2018); optimising the decision-making process of accessing the potential merging gaps for CAVs at single-lane roundabouts (Debada and Gillet, 2018); applying a cooperative merging control for CAVs at roundabouts in a mixed-traffic environment (Zhao et al., 2018). Under the same traffic demand, the performance comparison of intersectional organisation and roundabout at the same road junction has not been investigated.

The reason why roundabouts are not getting enough attention in junction management research for CAVs may be that the setting of intersections is more common in urban roads. With the appropriate signal control, the intersectional management can greatly improve the actual traffic throughput of the junction and separate conflicting streams, while roundabouts could reach capacity even with moderate traffic demands (Yang et al., 2004). However, the situation may be different in the CAVs era. It has been demonstrated that the compact headway travel behavior of CAVs will double the intersection capacity (Lioris et al., 2017) and so do the roundabouts. It is very likely that with the special car-following behavior of CAVs, roundabout may no longer have a limitation in terms of capacity. The purpose of this study is to theoretically derive and analyse the performance of FCFS-AIM and roundabout in terms of capacity and delay under the same CAVs flow.

The remainder of this chapter is organised as follows. The parameters and assumptions are elaborated in Section 4.2. The detailed derivations of the capacity and delay of each control strategy (as well as the comparison analysis) are given separately in Section 4.3 and Section 4.4. Section 4.5 concludes the study and discusses the direction of future work.

4.2 Parameters and assumptions

Main parameters and decision variables applied hereafter are summarised in Table 4.1.

TABLE 4.1: Notations

Set	Description
L	set of lanes (streams)
$\mathcal{G}_i$	set of non-conflicting streams of stream i
$\mathcal{C}_i$	set of conflicting streams of stream i
$\mathbf{E}$	set of entry approaches
Variable	Description
C	capacity of a control strategy
h_c	headway of CACC (s)
h_z	safety time gap between two vehicles from conflicting streams (s)
λ_i	vehicle arrival rate of stream i (veh/s)
λ_M	the sum of λ_i
μ	service rate of a queueing system
W	mean waiting time of a queueing system
D	average control delay

A typical urban road junction analysed in this study is depicted in Fig.4.2. The road junction consists of 8 traffic streams (left-turning streams are not considered as they have no conflicts with other motorised lanes). To be conservative, the single-lane type roundabout is analysed considering the limited geometric resources. Note that if there is enough space to set up a multi-lane roundabout, the performance should be better than the analysed single-lane roundabout.

All the possible conflict types in Fig.4.2 are elaborated in Fig.4.3. The safety time gap refers to the time required for vehicles to consecutively and safely pass a conflict point.

Since CAVs are not around the corner yet, it is difficult to know the specific distribution of the safety gaps, but they can be easily divided into two categories: the safety time gap (h_s, h_d, h_t, h_m) for two vehicles from conflicting streams and the safety time gap (h_c) for two vehicles from the non-conflicting streams. With this categorisation, the research in this study is based on the following assumptions:

FIGURE 4.2: The situation analysed in this study.

FIGURE 4.3: Illustration of conflict types of a typical intersection or roundabout.

- The communication capabilities of CAVs are perfectly operated, without any malfunctions.
- Lane changing is not allowed within the junction area, and vehicles strictly follow the traffic rules determined by the organisation of the junction; Wireless communication is perfectly guaranteed.
- The arrival process of vehicles on each stream $i \in L$ is a Poisson process with parameter λ_i and independent of each other.
- The difference between h_s, h_d, h_t, h_m is not distinguished in this study and a general distribution $f(h_z)$ with expectation $\mathbb{E}(h_z)$ for h_s, h_d, h_t, h_m is assumed. The distribution $f(h_c)$ of h_c is determined by the CACC policies with expectation $\mathbb{E}(h_c)$. Empirically, $\mathbb{E}(h_c) < \mathbb{E}(h_z)$.

4.3 Capacity

Capacity refers to the maximum sustainable hourly flow rate at which persons or vehicles reasonably can be expected to traverse a point or an uniform section of a lane or roadway during a given time period under predefined conditions. The capacity of the strategies are derived and compared in the following subsections.

4.3.1 Intersection using FCFS policy (I-FCFS)

For a vehicle v_i under *FCFS* policy, there are two situations for its preceding vehicle. One situation (denoted as event S_1) is that its preceding vehicle comes from the same stream or comes from the non-conflicting streams. These streams are grouped in the set $\mathcal{G}_i \subset L$. The service time of vehicle v_i is denoted by X_1 . Another situation (denoted as event S_2) is that its preceding vehicle comes from the conflicting streams, which are in the set $\mathcal{C}_i \subset L$. The service time of event S_2 is represented as X_2 .

Therefore, the service time X of a CAV under FCFS policy can be estimated as:

$$X_{fcfs}^I = \mathbb{P}(S_1^I)X_1 + \mathbb{P}(S_2^I)X_2 \quad (4.1)$$

where $\mathbb{P}(S_1)$ and $\mathbb{P}(S_2)$ denote the probabilities of occurrence of events S_1 and S_2 , respectively.

As the arriving process of vehicles are assumed to be the Poisson distribution (Heidemann and Wegmann, 1997), the time gap h_i between two consecutive vehicle in the stream i is exponentially distributed, i.e. $h_i \sim \exp(\lambda_i)$ and $\mathbb{E}(h_i) = 1/\lambda_i$. The probability that the n^{th} vehicle comes from stream i , can be calculated as

follows:

$$\begin{aligned}
\mathbb{P}_i^n &= \mathbb{P}(h_i < h_j, j \neq i; j \in L) \\
&= \int_0^\infty \mathbb{P}(h_i < h_j, j \neq i; j \in L \mid h_i = \Delta t) \lambda_i e^{-\lambda_i \Delta t} d(\Delta t) \\
&= \int_0^\infty (\prod_{j \neq i; j \in L} e^{-\lambda_j \Delta t}) \lambda_i e^{-\lambda_i \Delta t} d(\Delta t) \\
&= \int_0^\infty \lambda_i e^{-\sum_{j \in L} \lambda_j \Delta t} d(\Delta t) \\
&= \frac{\lambda_i}{\sum_{j \in L} \lambda_j} \\
&= \frac{\lambda_i}{\lambda_M}
\end{aligned} \tag{4.2}$$

With Eq.(4.2), the probabilities of occurrence of event S_2 can be calculated as:

$$\mathbb{P}(S_2) = \sum_{i \in L} \sum_{j \in \mathcal{C}_i} \mathbb{P}_i^n \mathbb{P}_j^{n-1} = \sum_{i \in L} \frac{\lambda_i}{\lambda_M} \frac{\sum_{j \in \mathcal{C}_i} \lambda_j}{\lambda_M} \tag{4.3}$$

For the sake of simplicity, we denote $\sum_{i \in L} \frac{\lambda_i}{\lambda_M} \frac{\sum_{j \in \mathcal{C}_i} \lambda_j}{\lambda_M}$ as $\mathcal{P}_{fcfs}^I$. Correspondingly,

$$\mathbb{P}(S_1) = 1 - \mathbb{P}(S_2) = 1 - \mathcal{P}_{fcfs}^I \tag{4.4}$$

The service time X_1 and X_2 are determined by the categorisation of safety gaps, corresponding to h_c and h_z , we have:

$$\mathbb{E}(X_1) = \int h_c f(h_c) d(h_c) = \mathbb{E}(h_c) \tag{4.5}$$

$$\mathbb{E}(X_2) = \int h_z f(h_z) d(h_z) = \mathbb{E}(h_z) \tag{4.6}$$

Capacity is equivalent to the service rate in queueing theory (Tanner, 1962). The intersection capacity under FCFS policy for CAVs can be estimated as:

$$C_{fcfs}^I = \frac{1}{\mathbb{E}(X_{fcfs}^I)} = \frac{1}{\mathbb{E}(h_c) + (\mathbb{E}(h_z) - \mathbb{E}(h_c)) \mathcal{P}_{fcfs}^I} \tag{4.7}$$

Table 4.2 shows the conflicting streams of a stream i within the intersection.

TABLE 4.2: The conflicting streams of a stream i within the intersection

stream i	conflicting stream set $\mathcal{C}_i$					stream i	conflicting stream set $\mathcal{C}_i$				
1	3	6	7	8		2	3	4	5	8	
3	1	2	5	8		4	2	5	6	7	
5	2	3	4	7		6	1	4	7	8	
7	1	4	5	6		8	1	2	3	6	

4.3.2 Roundabout using FCFS policy (R-FCFS)

The capacity of roundabout under FCFS policy can be determined by applying the same logic that derives C_{fcfs}^I . The difference lies in the potential collision mechanism due to the different organisation of the trajectories, i.e. the probabilities of occurrence of events S_1 and S_2 under circulating organisation are different.

At roundabout, an entry stream e only conflicts with the circulating stream c_e . The stream structure of roundabout is further elaborated in Fig.4.4. Take the eastward

FIGURE 4.4: Stream structure of roundabout.

approach as an example, the entry stream e_1 is made up of the vehicles on lane 1 and lane 2, conflicting with the circulating stream c_1 made up of the vehicles from lane 3, lane 4 and lane 6. The general arriving process of vehicles on each circulating flow is calculated as a superimposed Poisson process with parameter λ_{c_e} , Table 4.3 summarises all the superimpositions depicted in Fig.4.4. It is demonstrated that the simplification has little effect on the accuracy of the calculation (Wu,

1998). Note that the superimposition of k independent Poisson processes with rates $\lambda_1, \dots, \lambda_k$ is also a Poisson process with rate $\sum_i^k \lambda_i$.

TABLE 4.3: The circulating stream of an entry stream e within the single-lane roundabout

	entry stream		circulating stream		
superimposed of stream	1	2	3	4	6
superimposed of stream	3	4	5	6	8
superimposed of stream	5	6	2	7	8
superimposed of stream	7	8	1	2	4

With the distribution of the conflicting streams, the capacity of the roundabout under FCFS policy can be given by:

$$C_{fcfs}^R = \frac{1}{\mathbb{E}(X_{fcfs}^R)} = \frac{1}{\mathbb{E}(h_c) + (\mathbb{E}(h_z) - \mathbb{E}(h_c))\mathcal{P}_{fcfs}^R} \quad (4.8)$$

where,

$$\mathcal{P}_{fcfs}^R = \sum_{e \in \mathbf{E}} \sum_{c_e \in \mathcal{C}_e} \mathbb{P}_e^n \mathbb{P}_{c_e}^{n-1} = \sum_{e \in \mathbf{E}} \frac{\lambda_e}{\lambda_M} \frac{\sum_{c_e \in \mathcal{C}_e} \lambda_{c_e}}{\lambda_M} \quad (4.9)$$

4.3.3 Roundabout using major-minor policy (R-MM)

Typically, roundabout is controlled by the major-minor priority strategy, in which each entry stream e has the minor priority and has to find the proper gap between two consecutive circulating vehicles to merge into the circulating stream c_e . That is, the incoming streams need to overcome the impedance effect of the circulating flow. Under major-minor priority strategy, the expectation of service time of vehicles can be estimated as:

$$\mathbb{E}(X_{mm}^R) = \sum_{e \in \mathbf{E}} \frac{\lambda_e}{\lambda_M} \frac{\mathbb{E}[e^{\lambda_c^e h_z}] - 1}{\lambda_c^e} \quad (4.10)$$

Correspondingly,

$$C_{mm}^R = \frac{1}{\mathbb{E}(X_{mm}^R)} \quad (4.11)$$

where, $\lambda_c^e = \sum_{c_e \in \mathcal{C}_e} \lambda_{c_e}$.

Remark: The derivation of Eq.(4.10) is elaborated as follows (Boon et al., 2019).

Let $X_{mm}^{e,n}$ denote the inter departure time between the $(n-1)th$ and nth vehicles from stream e , and let Q_n^e denote the number of vehicles in the queue on the stream e when the nth vehicle from stream e merges into the stream c_e . The queue Q_n^e can be written as:

$$Q_n^e = \begin{cases} Q_{n-1}^e - 1 + A(X_{mm}^{e,n}), & Q_{n-1}^e \geq 1 \\ A(X_{mm}^{e,n}), & Q_{n-1}^e = 0 \end{cases} \quad (4.12)$$

where $A(\cdot)$ is the Poisson arrival process of stream e .

It can be verified that the process $\{Q_n^e, n = 1, 2, \dots\}$ has the dynamics of an $M/G/1$ queue length process at departures where $X_{mm}^{e,n}$ corresponds to the nth service time of the vehicles on stream e under major-minor priority policy.

Every CAV on the stream e needs a critical headway T for the mth attempt to merge into stream c_e ($m = 1, 2, \dots$). Every CAV sticks to the same critical headway, but different CAVs have different values of critical headway due to different control systems. The probability for the mth attempt of a CAV being successful is: $\mathbb{P}(\tau^c > T) = e^{-\lambda_c^e T}$, where τ^c is the gap of vehicles on the circulating stream, and is an exponential random variable with mean $1/\lambda_c^e$.

The distribution function of random variable $X_{mm}^{e,n}$ can be defined as:

$$B(x) \triangleq \left[\prod_{m=1}^k \mathbb{P}(\tau^c < T_m) \right] \mathbb{P}(T_{k+1} \leq \tau^c), k \in N \quad (4.13)$$

Since random variable $X_{mm}^{e,n}$ is generally distributed, we apply the Laplace transform of $B(x)$ as $\mathbb{E}[e^{-sX_{mm}^e}]$, which follows from:

$$\mathbb{E}[e^{-sX_{mm}^e}] = \sum_{k=0}^{\infty} \left\{ \prod_{m=1}^k \mathbb{E}[e^{-s\tau^c} \mathbb{P}(\tau^c < T_m)] \right\} \mathbb{E}[e^{-sT_{k+1}} \mathbb{P}(T_{k+1} \leq \tau^c)] \quad (4.14)$$

$$\begin{aligned} \mathbb{E}[e^{-s\tau^c} \mathbb{P}(\tau^c < T_m)] &= \mathbb{E}\left[\int_0^{T_m} e^{-s\tau^c} \lambda_c^e e^{-\lambda_c^e \tau^c} d\tau^c\right] \\ &= \mathbb{E}\left[\frac{\lambda_c^e}{s + \lambda_c^e} (1 - e^{-(s+\lambda_c^e)T_m})\right] \end{aligned} \quad (4.15)$$

$$\begin{aligned}\mathbb{E}[e^{-sT_{k+1}}\mathbb{P}(T_{k+1} \leq \tau^c)] &= \mathbb{E}[e^{-sT_{k+1}}\mathbb{P}(\tau^c \geq T_{k+1})] \\ &= \mathbb{E}[e^{-(s+\lambda_c^e)T_{k+1}}]\end{aligned}\quad (4.16)$$

Substituting Eq.(4.15) and Eq.(4.16) into Eq.(4.14), we get:

$$\mathbb{E}[e^{-sX_{mm}^e}] = \mathbb{E}\left[\sum_{k=0}^{\infty} \left(\frac{\lambda_c^e}{s + \lambda_c^e}\right)^k e^{-(s+\lambda_c^e)T_{k+1}} \prod_{m=1}^k (1 - e^{-(s+\lambda_c^e)T_m})\right] \quad (4.17)$$

For one CAV, $T_1 = T_2 = \dots = T_{k+1} = T$, Eq.(4.17) is the sum of a geometric sequence,

$$\mathbb{E}[e^{-sX_{mm}^e}] = \mathbb{E}\left[\frac{(s + \lambda_c^e)e^{-(s+\lambda_c^e)T}}{s + \lambda_c^e e^{-(s+\lambda_c^e)T}}\right] \quad (4.18)$$

Where T is distributed the same as h_z . Using the moment generating function, the mean service time of vehicles on stream e can be obtained:

$$\mathbb{E}[X_{mm}^e] = (-1) \frac{d\mathbb{E}[e^{-sX_{mm}^e}]}{ds} \Big|_{s=0} = \frac{\mathbb{E}[e^{\lambda_c^e h_z}] - 1}{\lambda_c^e} \quad (4.19)$$

Thus, Eq.(4.10) can be obtained.

4.3.4 Capacity comparison

By comparing Eq.(4.8) and Eq.(4.11), it can be easily concluded that R-FCFS policy outperforms major-minor policy in terms of capacity. The brief derivation is as follows:

$$\begin{aligned}
\mathbb{E}(X_{fcfs}^R) - \mathbb{E}(X_{mm}^R) &= \mathbb{E}(h_c) \times \mathbb{P}(S_1^R) + \mathbb{E}(h_z) \times \mathbb{P}(S_2^R) - \mathbb{E}(X_{mm}^R) \\
&= \sum_{e \in \mathbf{E}} \frac{\lambda_e}{\lambda_M} \left[\frac{\sum_{g_e \in \mathcal{G}_e} \lambda_{g_e}}{\lambda_M} \mathbb{E}(h_c) + \frac{\sum_{c_e \in \mathcal{C}_e} \lambda_{c_e}}{\lambda_M} \mathbb{E}(h_z) \right] - \mathbb{E}(X_{mm}^R) \\
&< \sum_{e \in \mathbf{E}} \frac{\lambda_e}{\lambda_M} \mathbb{E}(h_z) - \sum_{e \in \mathbf{E}} \frac{\lambda_e}{\lambda_M} \frac{\mathbb{E}[e^{\lambda_c^e h_z}] - 1}{\lambda_c^e} \\
&< \sum_{e \in \mathbf{E}} \frac{\lambda_e}{\lambda_M} \left\{ \frac{\lambda_c^e \mathbb{E}(h_z) - \mathbb{E}[e^{\lambda_c^e h_z}] + 1}{\lambda_c^e} \right\} \\
&\leq \sum_{e \in \mathbf{E}} \frac{\lambda_e}{\lambda_M} \left\{ \frac{\lambda_c^e \mathbb{E}(h_z) - e^{\lambda_c^e \mathbb{E}(h_z)} + 1}{\lambda_c^e} \right\} \\
&< \sum_{e \in \mathbf{E}} \frac{\lambda_e}{\lambda_M} \left\{ \frac{\lambda_c^e \mathbb{E}(h_z) - \lambda_c^e \mathbb{E}(h_z) - 1 + 1}{\lambda_c^e} \right\} = 0
\end{aligned} \tag{4.20}$$

In the above derivation, $\mathbb{E}[e^{\lambda_c^e h_z}] \geq e^{\lambda_c^e \mathbb{E}(h_z)} > 1 + \lambda_c^e \mathbb{E}(h_z)$ is applied. $\mathbb{E}(X_{fcfs}^R) < \mathbb{E}(X_{mm}^R)$ leads to $C_{fcfs}^R > C_{mm}^R$.

This result is in line with expectations. Using the communication technologies of CAVs, FCFS policy can assign the junction temporal-spatial resources more efficiently. If the major-minor priority pattern is still defaulted at roundabouts in the case of CAVs environment, the advantages of CAVs are wasted. The main purpose of this study is to explore whether the roundabout may replace the intersection settings at urban roads in the environment of CAVs. Since R-FCFS has been confirmed to be superior to R-MM, R-FCFS is used as the representative of roundabout management and is compared with FCFS-AIM. The performance of roundabout adopting major-minor priority pattern is no longer compared below.

The relationship between C_{fcfs}^I and C_{fcfs}^R depends on the traffic demand pattern, and the difference lies in $\mathcal{P}_{fcfs}^I$ and $\mathcal{P}_{fcfs}^R$:

$$\mathcal{P}_{fcfs}^I - \mathcal{P}_{fcfs}^R = \sum_{i \in L} \frac{\lambda_i}{\lambda_M} \frac{\sum_{j \in \mathcal{C}_i} \lambda_j}{\lambda_M} - \sum_{e \in \mathbf{E}} \frac{\lambda_e}{\lambda_M} \frac{\sum_{c_e \in \mathcal{C}_e} \lambda_{c_e}}{\lambda_M} \geq 0 \tag{4.21}$$

That is $\mathbb{E}(X_{fcfs}^I) \geq \mathbb{E}(X_{fcfs}^R)$, resulting in:

$$C_{fcfs}^R \geq C_{fcfs}^I \tag{4.22}$$

Specifically,

$$\begin{aligned}
\mathbb{E}(X_{fcfs}) &= \mathbb{P}(S_1) \times \mathbb{E}(h_c) + \mathbb{P}(S_2) \times \mathbb{E}(h_z) \\
&= (1 - \mathbb{P}(S_2)) \times \mathbb{E}(h_c) + \mathbb{P}(S_2) \times \mathbb{E}(h_z) \\
&< (1 - \mathbb{P}(S_2)) \times \mathbb{E}(h_z) + \mathbb{P}(S_2) \times \mathbb{E}(h_z) \\
&< \mathbb{E}(h_z)
\end{aligned} \tag{4.23}$$

And,

$$\begin{aligned}
\mathbb{E}(X_{fcfs}) &= \mathbb{P}(S_1) \times \mathbb{E}(h_c) + \mathbb{P}(S_2) \times \mathbb{E}(h_z) \\
&= (1 - \mathbb{P}(S_2)) \times \mathbb{E}(h_c) + \mathbb{P}(S_2) \times \mathbb{E}(h_z) \\
&> (1 - \mathbb{P}(S_2)) \times \mathbb{E}(h_c) + \mathbb{P}(S_2) \times \mathbb{E}(h_c) \\
&> \mathbb{E}(h_c)
\end{aligned} \tag{4.24}$$

It can be obtained that:

$$\frac{1}{\mathbb{E}(h_z)} < C_{fcfs} < \frac{1}{\mathbb{E}(h_c)} \tag{4.25}$$

In summary:

$$\frac{1}{\mathbb{E}(h_z)} < C_{fcfs}^I \leq C_{fcfs}^R < \frac{1}{\mathbb{E}(h_c)} \tag{4.26}$$

Essentially, the capacity of I-FCFS is relatively small because it has more impedance effects of more conflict points than the setting of roundabout.

4.4 Delay

In addition to capacity, efficiency is also an indicator that traffic management attaches great importance to. Therefore, average control delay of each strategy is used as another performance measurement and is derived and compared in this section.

4.4.1 Delay formulation

In queueing theory, by Little's law ([Little and Graves, 2008](#)), the mean waiting time in the system reads:

$$\mathbb{E}[W] = \frac{\mathbb{E}[N]}{\lambda} \quad (4.27)$$

where λ is the arriving rate, and $\mathbb{E}[N]$ is the mean number of waiting customers, i.e, the mean number of queueing vehicles in this study. The Little's law, of great generality, holds irrespective of the form of inter-arrival and service time distributions and discrepancy that may be within the system ([Medhi, 2002](#)), that is the relationship is not influenced by the arrival process distribution, the service distribution, or the service order ([Simchi-Levi and Trick, 2011](#)). It is widely accepted to calculate and analyse the performance of intersection traffic management ([Wu, 2001](#); [Boon et al., 2019](#)).

Derived from Little's law, for M/G/1 queue, the average waiting time by Pollaczek-Khinchin mean formula is:

$$\mathbb{E}[W] = \frac{\lambda \mathbb{E}(X^2)}{2(1 - \lambda \mathbb{E}(X))} \quad (4.28)$$

It should be noted that Eq.(4.28) is only applicable when the system is stable, meaning $1 - \lambda \mathbb{E}(X) > 0$. The predication of capacities and performance analysis under heavy flow conditions, especially for roundabouts with unbalanced entry flows can be found in [Akcelik et al. \(1996, 1998\)](#).

The service time X_{fcfs}^I and X_{FCFS}^R have the similar formulation (i.e., $X = \mathbb{P}(S_1)h_c + \mathbb{P}(S_2)h_z$).

Since h_c and h_z are independent variables, $\mathbb{E}(X^2)$ can be calculated as:

$$\begin{aligned} \mathbb{E}(X^2) &= \mathbb{D}(X) + \mathbb{E}^2(X) \\ &= \mathbb{P}^2(S_1)\mathbb{D}(h_c) + \mathbb{P}^2(S_2)\mathbb{D}(h_z) + (\mathbb{P}(S_1)\mathbb{E}(h_c) + \mathbb{P}(S_2)\mathbb{E}(h_z))^2 \end{aligned} \quad (4.29)$$

Calculating the corresponding values of $\mathbb{E}(X^2)$ and $\mathbb{E}(X)$ and substituting them into Eq.(4.28), the delay of I-FCFS and R-FCFS can be obtained:

$$D_{fcfs}^I = \frac{\lambda_M \mathbb{E}[(X_{fcfs}^I)^2]}{2\{1 - \lambda_M[\mathbb{E}(h_c) + (\mathbb{E}(h_z) - \mathbb{E}(h_c))\mathcal{P}_{fcfs}^I]\}} \quad (4.30)$$

$$D_{fcfs}^R = \frac{\lambda_M \mathbb{E}[(X_{fcfs}^R)^2]}{2\{1 - \lambda_M[\mathbb{E}(h_c) + (\mathbb{E}(h_z) - \mathbb{E}(h_c))\mathcal{P}_{fcfs}^R]\}} \quad (4.31)$$

4.4.2 Delay comparison

It is obtained that $\mathbb{E}(X_{fcfs}^I) \geq \mathbb{E}(X_{fcfs}^R)$ and $\mathcal{P}_{fcfs}^I \geq \mathcal{P}_{fcfs}^R$, with the same traffic arriving rate, comparing the theoretical formulas Eq.(4.30) and Eq.(4.31), getting:

$$D_{fcfs}^I \geq D_{fcfs}^R \quad (4.32)$$

In general, under CAVs environment, for the same road junction, we can see that roundabout using FCFS policy outperforms FCFS-AIM in capacity and efficiency, see Eq.(4.26) and Eq.(4.32). It is worth mentioning that this conclusion holds for any distribution of h_c and h_z , because they are assumed to be generally distributed in the derivation.

(a) Equal demand scenario

(b) Major-minor demand scenario

FIGURE 4.5: Delay comparison under different demand scenarios.

In literature, the headway h_c is normally assumed as uniformly distributed with the parameter in range of $[0.5s, 2.0s]$ (Ghiassi et al., 2017). In order to more intuitively reflect the performance of FCFS-AIM and FCFS-Roundabout, this uniform distribution of h_c is applied to calculate the values of delay under two representative traffic demand patterns according to the derived equations, as shown in Fig.4.5.

The random variable h_z is estimated as $2h_c$ in the calculation. In Fig.4.5, *Symmetrical traffic flow* scenario refers to the case where the traffic demand of each stream is the same, i.e. $\lambda_1 = \lambda_2 = \dots = \lambda_8 = \lambda$. Note that although the arrival rates are the same, the real traffic dynamics are not the same and independent of each other. The *asymmetrical traffic flow* scenario reflects the major-minor traffic pattern in which $\lambda_1 = 3\lambda$; $\lambda_2 = \dots = \lambda_8 = \lambda$. Better performance of R-FCFS is verified.

As we emphasised, the derived delay function is only applicable when the capacity of each strategy is not reached. Therefore, traffic simulations are conducted to corroborate the theoretical conclusions and to demonstrate the delay of each strategy when the junction traffic arriving rate exceeds the capacity.

SUMO traffic simulator is used to build the road junction, and TraCI interface is used to customise the FCFS policy to SUMO (Lopez et al., 2018). Cooperative adaptive cruise control (CACC) is adopted to simulate the car-following behavior of CAVs (Milanés and Shladover, 2014), in which the platoon headway $h_c = 1s$, the critical inter-arrival between consecutive platoons is set as $2.5s$ (Jiang et al., 2006), and the driver imperfection is set to $\sigma = 0$ as perfect CAV behavior is expected. The situation shown in Fig.4.2 under equal traffic demand scenario is simulated. The comparison of the average vehicle delay caused by each control strategy is shown in Fig.4.6. For each simulated scenario, each control has the

FIGURE 4.6: Delay comparison of simulation results.

same incoming traffic flow, thus the difference in delay only comes from the different junction organisations and control strategies. As can be seen from Fig.4.6, under the same traffic flow situation, roundabout adopting FCFS policy always has the best performance among the three strategies. In addition, when the traffic flow is approaching the capacity, the control delay of each strategy increases greatly, which is very close to the theoretically derived value. The results of theoretical derivation are supported. The specific values for delays in Fig.4.6 can be found in Table 4.4.

TABLE 4.4: The average delay (s) of simulations

Junction traffic arriving rate	I-FCFS	R-FCFS	R-MM
800 <i>veh/h</i>	8.55	5.13	7.16
1200 <i>veh/h</i>	11.53	8.75	11.32
1600 <i>veh/h</i>	18.01	8.94	13.54
2000 <i>veh/h</i>	28.34	16.86	30.9
2400 <i>veh/h</i>	59.70	50.29	62.21
2800 <i>veh/h</i>	82.59	51.84	70.12
3200 <i>veh/h</i>	83.11	78.87	93.17
3600 <i>veh/h</i>	132.01	128.06	130.33
4000 <i>veh/h</i>	250.92	201.00	208.92

It is worth noting that the three strategies comparatively analysed in this study, especially I-FCFS and R-FCFS, are very sensitive to the real-time traffic dynamics, since the assignment of road rights in FCFS is directly related to the arrival-time of vehicles. In other words, the randomness of traffic flow has a great influence on the performance of these strategies (Wu et al., 2019). As a predetermined control policy, FCFS has its limitations. This is the reason why many optimisation approaches are proposed for AIM on the basis of FCFS in order to further improve traffic (Yu et al., 2019). However, as mentioned in the literature review, most of the research on signal-free junction management for CAVs, including the optimisation of AIM, are based on the intersectional organisation, circular organisation is rarely considered. Through the comparison in this study, under the same circumstances, FCFS-Roundabout has a better performance than FCFS-AIM, which makes it reasonable to believe that the optimisation based on FCFS-Roundabout may perform better than the optimisation based on FCFS-AIM.

4.5 Conclusions

This chapter sets out to compare the performance of FCFS-AIM and roundabout in terms of capacity and delay in the CAVs environment. The same in-coming traffic demand is guaranteed. Three management strategies are analysed: intersectional organisation using FCFS policy (I-FCFS), roundabout adopting FCFS policy (R-FCFS) and roundabout following the major-minor priority pattern (R-MM). Accounting for the characteristics of CAVs, the theoretical formulas to estimate the capacity and average control delay of different junction managements are derived using queueing theory. The derivation of the formulas is based on the general distribution of safety gaps rather than some specific distributions, which makes the formulas more applicable. The capacity of each strategy is inversely related to safety gaps: the smaller critical safety gaps, the larger the capacity. Under the same circumstances, R-FCFS has the largest capacity and the least average control delay among the three signal-free strategies. Traffic simulations are carried out to corroborate the theoretical conclusions.

The comparison results reveal an important point that the optimal management of CAVs based on circulating trajectory organisation could have at least the comparable performance as the optimal policy based on the intersectional organisation, which indicates that roundabout should be given more attention in the signal-free management for CAVs. For the same incoming traffic demand under CAVs environment, with the same optimisation goals, comparing the performance of intersection and roundabout is one of our future work. In addition, in order to facilitate the derivation of the formulas and the comparison of performances, some necessary assumptions are used. In the future, when the exact distributions of safety gaps are known, using the real data of junction traffic demand, the applicable scenarios and conditions of each strategy can be further distinguished using the proposed methods. This study stems from the curiosity about whether roundabout has comparable performance to FCFS-AIM under the pure CAVs environment. However, CAVs will coexist with conventional vehicles for a foreseeable transitional stage, as the full penetration of CAV technology still requires long-term development and evolution, which makes the mixed traffic composed of CAVs and conventional vehicles inevitable. The analysis of mixed traffic is of great significance, and is preliminarily investigated in next chapter.

Part III

Mixed Traffic Exploration

Chapter 5

Intersection Capacity under Mixed Traffic Environment

With the emergence of connected and automated vehicles (CAVs), the mixed traffic environment where CAVs share the traffic network with human-driven vehicles (HVs) is impending. Exploring the impact of CAVs on mixed traffic flow is significant to realise these potentials. In this chapter, a new theoretical approach for the determination of intersection capacity under mixed traffic is presented.

5.1 Introduction

Connected and Automated Vehicles (CAVs) are burgeoning. With the advanced control systems and wireless communication technologies, CAVs have the potential to progressively revolutionise intelligent transportation system (ITS) with increased efficiency, enhanced safety, reduced emissions, saved energies and improved environmental effects ([Vahidi and Sciarretta, 2018](#); [Gawron et al., 2018](#); [Tu et al., 2019](#)). These potential benefits are confirmed by many exploratory studies in which the characteristics of CAVs are modeled or simulated, assuming that CAVs have become a reality and are allowed to drive on public roads. These advanced discussions have a lot of practical significance and can prepare for the opportunities and challenges of the new era of CAVs ([Fagnant and Kockelman, 2015](#)).

Compared to Human-driven Vehicles (HVs), the most distinctive characteristic of CAVs is that they can exchange information with each other and surroundings through Vehicle-to-Everything (V2X) communication technologies (Molina-Masegosa and Gozalvez, 2017), resulting in the much less reaction and response time. That is, CAVs are able to maintain a smaller homogenous headway to a predecessor and move safely at a harmonised speed. This car-following behaviour is often referred as *platooning* and is realised by Cooperative Adaptive Cruise Control (CACC) (Milanés et al., 2013; Shladover et al., 2015). There are three critical parameter settings for platooning: intra-platoon gap, inter-platoon gap and platoon size (Amoozadeh et al., 2015). Intra-platoon gap corresponds to CACC time-headway setting, represented as τ_c in this study. Inter-platoon gap (τ_p) is the separation between two different platoons, which is assumed to be larger than intra-platoon spacing ($\tau_p > \tau_c$). Platoon size is the number of vehicles in the platoon. The most notable impacts of platooning on traffic flow are the increment of road capacity (Shladover et al., 2012), the improvement of string stability (Feng et al., 2019a), and the enhancement of road safety (Papadoulis et al., 2019).

Platoon based intersection control came into being, such as: platoon trajectory is optimised at a fixed-timing signalised intersection in Liu et al. (2019); an optimal speed advisory concept during green signal is proposed for CAV platoon in Feng et al. (2019b). Analytically, it has been demonstrated that CAV platoon can double the intersection throughput in a pure CAVs environment (Lioris et al., 2017).

However, the full penetration of CAV technology still requires long-term development and evolution, which makes the mixed traffic composed of both CAVs and HVs inevitable. There are many studies on the control and optimisation of mixed traffic flow, such as: a cooperative platoon control for a mixed traffic flow is proposed in Gong and Du (2018), in which both transient traffic smoothness and asymptotic stability of the mixed flow platoon is ensured by leveraging the communication and computation technologies equipped on CAVs; a speed harmonisation algorithm is proposed in Ghiasi et al. (2019) for mixed freeway traffic, yielding much smoother trajectories for not only CAVs but also HVs; in Li et al. (2020), an effective CAV controller is designed for the mixed-stream traffic stabilisation using empirical trajectory data. In terms of theoretical analysis, a Markov chain method is proposed to model the mixed traffic lane capacity in Ghiasi et al. (2017), in which the sufficient and necessary conditions for the mixed traffic capacity to

increase with CAV penetration rate and platooning intensity are identified. The fundamental traffic diagram for CAV-HV mixed traffic is developed wherein the capacity and jam density of the road is determined as a function of the penetration rate and the heterogeneous headways in [Shahri et al. \(2020\)](#). In [Zhou and Zhu \(2020\)](#), platooning intensity is considered in the modeling of the mixed traffic fundamental diagram.

Above analyses are for the lane-level mixed-traffic. We consider the more complex situations: intersections, where multiple streams from different directions converge to share the same intersection capacity. [Zhong et al. \(2020\)](#) point out that the single-lane freeway cases are prone to overestimate the mobility benefits brought by CAVs. It is also demonstrated that although platooned CAVs have smaller inter-vehicle spacing, uncoordinated and randomly arriving CAV platoons may act as *moving obstacles* and can induce local congestion, especially at bottlenecks ([Jin et al., 2020](#)). In this Chapter, we set out to analytically model the maximum throughput that an intersection can reach under mixed CAVs-HVs traffic flow conditions, and inspect the influence of CAVs platooning configuration on it. The specific control method to achieve the maximum throughput is not the focus of this research.

Furthermore, the CAV platooning evaluation has been focusing on the benefits that CAV could potentially bring to the transportation system from the perspective of the system operator, defaulting that the adjacent CAVs will form into platoons definitely. However, it is pointed out in [Sun and Yin \(2019\)](#) that from the perspective of energy saving, due to at different positions in a platoon save differently from the platooning, some drivers or owners of CAVs may not be willing to join or stay in the platoon even if they are advised to do so, yielding a behaviourally unstable platoon. It is alluded that such behavioural instability of platooning could have adverse impact on the traffic flow efficiency ([Kokkinogenis et al., 2019](#)), but has not yet been analysed as a variable in the modeling. This study intends to fill this gap and provide a detailed analysis, in which the platooning willingness is taken as an influencing factor along with the three critical parameters for platooning.

The main contribution of this chapter is to proposed a more practical procedure to determine the intersection capacity under mixed traffic without having to know the specific management of the intersection. The method can take into account the number of lanes at the approaches, the interaction between the different streams,

the distribution of different types of vehicles, and the different CAV platooning parameters.

The remainder of this chapter is organised as follows. Section 5.2 elaborates on the theoretical basis and preliminaries of the modeling of mixed traffic intersection capacity, including the Markov chain process and the conflict point technique. Analyses are conducted in Section 5.3, and a numerical case is given in Section 5.4. Section 5.5 concludes the chapter with a summary of the analyses and a brief discussion of future research directions.

5.2 Modelling

Main parameters and decision variables applied hereafter are summarised in Table 5.1.

TABLE 5.1: Notations

Variable	Description
$\mathbf{P}$	state transition probability matrix
$\mathbf{Y}$	vector of steady-state probabilities
C	capacity of a control strategy
Q_i	throughput of stream i
p_c	penetration rate of CAVs
p_h	penetration rate of HVs
w	willingness of CAVs to form into platoons
τ_c	intra-platoon gap/CACC gap setting
τ_m	critical departure headway of a CAV following a HV
$\tau_{m'}$	critical departure headway of a HV following a CAV
τ_h	critical departure headway of two consecutive HVs
τ_p	inter-platoon gap
$t_{B,i}$	average service time of vehicles on stream i

5.2.1 Capacity at intersections

At intersections, the vehicles in different streams have to pass the same conflict area alternately controlled by some rules. The maximum value of the sum of the

throughput (Q_i) of each stream corresponds to the capacity (C) of the intersection:

$$C = \max(\sum Q_i) \quad (5.1)$$

FIGURE 5.1: Illustration of streams at intersections

To derive the capacity of intersections, the conflict points are critical. The vehicles in different streams have to pass the same conflict area alternatively to avoid potential conflicts. Every vehicle of the streams occupies the conflict point by a time of t_B seconds. At a conflict point, there are at least two intersecting streams, as shown in Fig.5.1.a). The throughput (Q_i) of the two streams has the following relationship:

$$Q_1 \times t_{B,1} + Q_2 \times t_{B,2} \leq 3600 \quad (5.2)$$

where queueing theory is applied, each vehicle from movement i takes an average service time of $t_{B,i}$ passing the conflict point. 3600 is the total time (s/hour) available for vehicles from all the movements conflicting at the same point.

To better illustrate, we take the four conflict points case shown in Fig.5.1.d) as another example. The objective is to obtain the capacity:

$$C_d = \max(Q_1 + Q_2 + Q_3 + Q_4) \quad (5.3)$$

with the following constraints:

$$\begin{aligned} Q_1 \times t_{B,1} + Q_2 \times t_{B,2} &\leq 3600 \\ Q_2 \times t_{B,2} + Q_4 \times t_{B,4} &\leq 3600 \\ Q_1 \times t_{B,1} + Q_3 \times t_{B,3} &\leq 3600 \\ Q_3 \times t_{B,3} + Q_4 \times t_{B,4} &\leq 3600 \end{aligned} \quad (5.4)$$

Obviously, this is a linear programming problem, and can be solved when the values of occupation times $t_{B,i}$ are specified.

With the above examples, we give a general formulation for an arbitrary intersection. Suppose there are m streams, resulting in n conflict points within the intersection. For each of the conflict points $c \in n$, we have the following inequality:

$$k_1^c \times Q_1 \times t_{B,1} + k_2^c \times Q_2 \times t_{B,2} + \cdots + k_m^c \times Q_m \times t_{B,m} \leq 3600 \quad (5.5)$$

where,

$$k_i^c = \begin{cases} 1, & \text{if stream } i \text{ passes conflict point } c; \\ 0, & \text{otherwise.} \end{cases} \quad (5.6)$$

The problem is to maximise the sum of the throughput of the m streams,

$$C = \max_{i \in m} \left(\sum Q_i \right) \quad (5.7)$$

subjects to:

$$\begin{bmatrix} k_1^1 t_{B,1} & k_2^1 t_{B,2} & \cdots & k_m^1 t_{B,m} \\ \vdots & \vdots & \vdots & \vdots \\ k_1^c t_{B,1} & k_2^c t_{B,2} & \cdots & k_m^c t_{B,m} \\ \vdots & \vdots & \vdots & \vdots \\ k_1^n t_{B,1} & k_2^n t_{B,2} & \cdots & k_m^n t_{B,m} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \\ \vdots \\ Q_m \end{bmatrix} \leq \begin{bmatrix} 3600_1 \\ 3600_2 \\ 3600_3 \\ \vdots \\ 3600_n \end{bmatrix} \quad (5.8)$$

and,

$$Q_i \geq 0, \quad i \in m \quad (5.9)$$

There are obviously feasible solutions to the above linear programming problem. On this basis, applying the simplex method, an optimal solution can be found in a finite number of iterations ([Vanderbei, 2020](#)). Accordingly, we give the solution expression as:

$$C = \sum_{v \in V} \frac{3600}{\kappa_v t_{B,i}^v} \quad (5.10)$$

where $v \in V$ is the number of iterations. Q_i^v is the selected basic variable at the v th iteration. $\kappa_v t_{B,i}^v$ denotes the corresponding terms of the v th iteration, which depends on the constraints determined by the intersection conflict mechanism.

The modelling in this chapter is a general extension of the method proposed in [Brilon and Wu \(2001\)](#). The analysis in [Brilon and Wu \(2001\)](#) is based on a concept

FIGURE 5.2: Conflicts among three movements in Brilon and Wu (2001).

of *conflict group* proposed for unsignalised intersections. Taking the case shown in Fig.5.1.c) as an example. In Brilon and Wu (2001), the two situations shown in Fig.5.2 are seen as the same, the three adjacent conflict points are considered as a conflict group or conflict area, and share the capacity of 3600 as $C_1 \times t_{B,1} + C_2 \times t_{B,2} + C_3 \times t_{B,3} = 3600$. This means that the entire area only can be occupied by one vehicle at a time. For pure HVs traffic flow, in order to better guarantee safety, this method of modeling conflict group is feasible, and is demonstrated to be much easier to handle than the traditional gap-acceptance theory by avoiding many of the theoretical complications (Wu, 2000; Brilon and Wu, 2001).

For mixed CAVs-HVs traffic, this method of grouping conflict points into conflict area underestimates the traffic capacity of the intersection. Since these conflict points are separate, vehicles on different streams can occupy different conflict points at the same time, as long as it is ensured that they do not conflict with subsequent vehicles when they reach the next conflict point. In other words, if vehicles on different streams can coordinate the departure time well, more than one vehicle can be served at the same time. The communication capabilities of CAVs will facilitate such coordination. For the traffic flow involving CAVs, our proposed method of modelling each conflict point is more accurate and general.

According to Eq.(5.8), for the three conflict points case shown in Fig.5.1.c), we have the following equations:

$$\begin{aligned}
 Q_1 \times t_{B,1} + Q_2 \times t_{B,2} &\leq 3600 \\
 Q_2 \times t_{B,2} + Q_3 \times t_{B,3} &\leq 3600 \\
 Q_1 \times t_{B,1} + Q_3 \times t_{B,3} &\leq 3600 \\
 &\Downarrow \\
 Q_1 \times t_{B,1} + Q_2 \times t_{B,2} + Q_3 \times t_{B,3} &\leq 5400
 \end{aligned} \tag{5.11}$$

Applying the simplex method, the solution of this three points case is that:

$$C_c = \frac{1800}{t_{B,1}} + \frac{1800}{t_{B,2}} + \frac{1800}{t_{B,3}} \quad (5.12)$$

FIGURE 5.3: Underestimation of the capacity of the conflict area model.

An illustration example of the discussed underestimation is shown in Fig. 5.3, where $\Delta t_{C,i}$ is the travel time of vehicles on stream i from a upstream conflict point to the next downstream conflict point (depicted in Fig. 5.1), $V_i - j$ represents the j th vehicle from stream i , $\rightarrow$ indicates the movement of the vehicles. It can be seen that with the primary collision-free condition being guaranteed, within the same time, the maximum throughput of the three separated conflict points model is larger than the throughput of the conflict area model.

The occupation times $t_{B,i}$ are the decisive factors for the solutions. In Brilon and Wu (2001), $t_{B,i}$ values are roughly calibrated by the classical gap-acceptance theory established in HCM (HCM, 2000). In this study, $t_{B,i}$ values are considered in-depth to reflect the impact of CAV platooning on the capacity of intersections under mixed traffic situations, detailed in the following subsection.

5.2.2 Average occupation time

As modeled above, the average occupation time $t_{B,i}$ values constitute the parameters of the model. The inter-departure time between the $(n-1)$ th and n th vehicles is need to measure the $t_{B,i}$ values. Let,

- V_n^c : the n th vehicle is a CAV;
- V_n^h : the n th vehicle is a HV;
- V_n^{c,w^+} : the n th vehicle is a CAV and willing to platoon with other CAVs;
- V_n^{c,w^-} : the n th vehicle is a CAV but not willing to form into platoon.

FIGURE 5.4: State transition diagram.

Intuitively, the states ($\mathcal{S}$) of two adjacent vehicles can be enumerated as follows:

$$\mathcal{S} = \begin{cases} S_1 : \mathbf{c} \rightarrow \mathbf{c}, & V_n^{c,w^+} \text{ following } V_{n-1}^{c,w^+} \\ S_2 : \mathbf{c} \rightarrow \mathbf{h}, & V_n^c \text{ following } V_{n-1}^h \\ S_3 : \mathbf{h} \rightarrow \mathbf{c}, & V_n^h \text{ following } V_{n-1}^c \\ S_4 : \mathbf{h} \rightarrow \mathbf{h}, & V_n^h \text{ following } V_{n-1}^h \\ S_5 : \mathbf{p} \rightarrow \mathbf{p}, & V_n^{c,w^-} \text{ following } V_{n-1}^c \end{cases} \quad (5.13)$$

Denote the penetration rate of CAV as p_c , the willingness of CAVs to platoon with adjacent CAVs as w . Then the probability of HV is $1 - p_c$, represented as p_h . Corresponding to the stochastic arriving process of vehicles, we can have the state transition diagram as in Fig. 5.4, and the state transition probability matrix

($\mathbf{P} = [p_{S_i S_j}(1)]$):

$$\mathbf{P} = \begin{matrix} & \begin{matrix} S_1 & S_2 & S_3 & S_4 & S_5 \end{matrix} \\ \begin{matrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{matrix} & \begin{bmatrix} p_c w & 0 & p_h & 0 & p_c(1-w) \\ p_c w & 0 & p_h & 0 & p_c(1-w) \\ 0 & p_c & 0 & p_h & 0 \\ 0 & p_c & 0 & p_h & 0 \\ p_c w & 0 & p_h & 0 & p_c(1-w) \end{bmatrix} \end{matrix} \quad (5.14)$$

Let $T = \{t_0, t_1, t_2, \dots\}$, $X(t_n)$ be denoted by $X_n, n = 0, 1, 2, \dots$. The state space $S = \{1, 2, 3, 4, 5\}$. $X_m = k$ represents the state of the stochastic process at time t_m is S_k . X_0 is the initial state of the system. Denote the state probabilities as:

$$\begin{aligned} p_j(n) &= P(X_n = j) \\ &= \sum_{i \in S} P(X_n = j, X_0 = i) \\ &= \sum_{i \in S} P(X_n = j | X_0 = i) P(X_0 = i) \\ &= \sum_{i \in S} p_{ij}(n) p_i(0), \quad j \in S \end{aligned} \quad (5.15)$$

where, $p_{ij}(n)$ is the n -step transition probability (from state S_i to state S_j). Let $\pi(n) = [p_1(n) \ p_2(n) \ p_3(n) \ p_4(n) \ p_5(n)]$, Eq.(5.15) gives:

$$\pi(n) = \pi(0)\mathbf{P}(n) = \pi(0)\mathbf{P}^n \quad (5.16)$$

Let,

$$y_j = \lim_{n \rightarrow \infty} p_j(n), \quad j \in S \quad (5.17)$$

It is clear that,

$$\sum_j y_j = 1, \quad y_j \geq 0 \quad (5.18)$$

Denote,

$$\mathbf{Y} = \lim_{n \rightarrow \infty} \pi(n) = [y_1 \ y_2 \ y_3 \ y_4 \ y_5] \quad (5.19)$$

Normally, vector $\mathbf{Y}$ is called as the steady state of a Markov chain, which is the long-term probabilities of the system being in each state and is the equilibrium distribution of states, regardless of initial state that is approached by successive

iterations from any starting distributions (Viswanadham and Narahari, 2015). We have that:

$$\mathbf{Y} = \lim_{n \rightarrow \infty} \pi(n) = \lim_{n \rightarrow \infty} \pi(n-1)\mathbf{P} = \mathbf{Y}\mathbf{P} \quad (5.20)$$

Known the state transition probability matrix $\mathbf{P}$ in Eq.(5.14), the steady-state vector $\mathbf{Y}$ can be computed using Eq.(5.18) and Eq.(5.20):

$$\mathbf{Y} = [wp_c^2 \quad p_cp_h \quad p_cp_h \quad p_h^2 \quad (1-w)p_c^2] \quad (5.21)$$

FIGURE 5.5: An illustration of the different inter-departure time.

As illustrated in Fig.5.5, the inter-departure headway (τ), i.e. the occupation time, is different at each state. Because CAVs have not yet become a reality, leading to the specific distribution of headways is not knowable, the headway in each state is assumed to be an independent random variable with a general distribution. The inter-departure headways corresponding to S_1, S_2, S_3, S_4 and S_5 are denoted as $\tau_c, \tau_m, \tau_{m'}, \tau_h$ and τ_p , respectively.

With the probabilities of the states, the average occupation time can be expressed as:

$$\begin{aligned} t_{B,i} &= P(S_1)\mathbb{E}(\tau_c) + P(S_2)\mathbb{E}(\tau_m) + P(S_3)\mathbb{E}(\tau_{m'}) + P(S_4)\mathbb{E}(\tau_h) + P(S_5)\mathbb{E}(\tau_p) \\ &= wp_c^2\mathbb{E}(\tau_c) + p_cp_h[\mathbb{E}(\tau_m) + \mathbb{E}(\tau_{m'})] + p_h^2\mathbb{E}(\tau_h) + (1-w)p_c^2\mathbb{E}(\tau_p) \end{aligned} \quad (5.22)$$

For simplicity, denote $\mathbb{E}(\tau_c), \mathbb{E}(\tau_m), \mathbb{E}(\tau_{m'}), \mathbb{E}(\tau_h)$ and $\mathbb{E}(\tau_p)$ as $t_c, t_m, t_{m'}, t_h$ and t_p , respectively. Then,

$$t_{B,i} = wp_c^2t_c + p_cp_h(t_m + t_{m'}) + p_h^2t_h + (1-w)p_c^2t_p \quad (5.23)$$

With Eq.(5.23), the linear programming problem can be solved, that is, the capacity is determined.

5.3 Analyses

It has been stated that capacity of intersections under mixed traffic flow is an inverse function of the average occupation time $t_{B,i}$. An increase in average occupation time will result in a reduction in capacity, and conversely a decrease in the average occupation time will lead to a increase in capacity. On this basis, the influences of critical platooning parameters are analysed.

5.3.1 Effect of platooning willingness

Corollary 5.1. *Capacity is an increasing function of platooning willingness.*

Proof. To consider the influence of platooning willingness w , Eq.(5.23) is expressed as:

$$t_{B,i} = M - p_c^2(t_p - t_c)w \quad (5.24)$$

where M represents the part unrelated to w . We have $t_p > t_c$, then $t_{B,i}$ is a decreasing function of w , corollary 5.1 can be concluded. $\square$

5.3.2 Effect of CAV penetration rate

Let,

$$A = wt_c + t_h + (1 - w)t_p - (t_m + t_{m'}), B = 2t_h - (t_m + t_{m'}), \quad (5.25)$$

then, the average occupation time (Eq.(5.23)) can be expressed as:

$$t_{B,i} = Ap_c^2 - Bp_c + t_h \quad (5.26)$$

Two corollaries can be arrived, and corresponding illustration is given in Fig.5.6.

Corollary 5.2. *If $A \leq 0$, $t_{B,i}$ is a decreasing function of p_c , indicating that capacity increases as penetration rate of CAVs increases.*

FIGURE 5.6: An illustration of the relationship between $t_{B,i}$ and p_c .

Corollary 5.3. *If $A > 0$, let $p_m = \frac{B}{2A}$. Then, if $p_m > 1$, capacity increases as penetration rate of CAVs increases; if $0 < p_m < 1$, $t_{B,i}$ has a minimum value at $p_c = p_m$, while capacity has a maximum value at $p_c = p_m$.*

The above two corollaries indicate that it is not certain that higher penetration rate p_c will lead to higher capacity, but further depends on platooning willingness and the heterogenous headways. For the five different headways, τ_h and $\tau_{m'}$ are related to HVs, and the other three are related to CAVs. It is reasonable to assume that τ_h and $\tau_{m'}$ have the same average value ($t_h = t_{m'}$), because they both depend on the the following behaviour of HVs. Further, A can be expressed by only the three CAV related headways as $A = t_p - t_m - w(t_p - t_c)$. Let $\Delta t_{pm} = t_p - t_m$, $\Delta t_{pc} = t_p - t_c$ and $\alpha = \Delta t_{pm} / \Delta t_{pc}$. If there are some control strategies that enable,

$$\alpha \leq w \leq 1 \quad (5.27)$$

leading to $A \leq 0$. Then by Corollary 5.2, the positive effect of p_c on the intersection capacity is deterministic and no longer dependent on the other HV-related headways, suggesting that CAV penetration rate p_c and platooning willingness w need to be considered jointly when updating intersection management for mixed CAV-HV traffic flows.

5.3.3 Effect of platoon gaps

From Eq.(5.23), $t_{B,i}$ is an increasing function of the expectation of τ_c and τ_p . It can be concluded that:

Corollary 5.4. *Capacity increases as intra-platoon gap and inter-platoon gap decrease.*

Intra-platoon gap τ_c and inter-platoon gap τ_p are the core parameters of platooning: the smaller the gaps, the higher capacity. In other words, the more mature the technologies makes the safety gap requirements smaller, which in turn will be more beneficial to the performance of traffic flow management.

Unlike the influence of penetration rate, Corollary 5.4 is not affected by the other factors, nor is it affected by other specific control of traffic flow.

5.3.4 Effect of platoon size

The size or length of platoon is an important parameter for the management and control of CAV platooning. In the case of mixed traffic flow, the distribution of vehicles can be modeled as follows. In accordance with the stochastic nature of traffic flow, the probability that the $n - th$ vehicle is a HV:

$$\mathbb{P}\{V_n = HV\} = p_h = 1 - p_c \quad (5.28)$$

Similarly, the probability that the $n - th$ vehicle is a CAV, and there are $k - 1$ number of CAVs platooning and following with this $n - th$ vehicle:

$$\mathbb{P}\{V_n = CAV^k\} = p_c[p_h + p_c(1 - w)](p_cw)^{k-1} = p_c(1 - p_cw)(p_cw)^{k-1} \quad (5.29)$$

It can be easily verified by Eq.(5.29) that the probability that the $n - th$ vehicle is a CAV is equal to p_c :

$$\begin{aligned} \mathbb{P}\{V_n = CAV\} &= \sum_{k=1}^{\infty} \mathbb{P}\{V_n = CAV^k\} \\ &= p_c(1 - p_cw) \frac{1 - (p_cw)^k}{1 - p_cw} \Big|_{k \rightarrow \infty} \\ &= p_c \end{aligned} \quad (5.30)$$

With the distribution of vehicles in Eq.(5.28) and Eq.(5.29), when the platoon size is not controlled, the average occupation time can be estimated as (t_n denotes the

occupation time of the n th vehicle):

$$\begin{aligned}
 t'_{B,i} &= \mathbb{P}\{V_n = HV\}\mathbb{E}(t_n|V_n = HV) + \sum_{k=1}^{\infty} \mathbb{P}\{V_n = CAV^k\}\mathbb{E}(t_n|V_n = CAV^k) \\
 &= p_h(p_c t_{m'} + p_h t_h) + \sum_{k=1}^{\infty} p_c(1 - p_c w)(p_c w)^{k-1} [p_c w t_c + p_h t_m + p_c(1 - w)t_p] \\
 &= w p_c^2 t_c + p_c p_h(t_m + t_{m'}) + p_h^2 t_h + (1 - w)p_c^2 t_p
 \end{aligned} \tag{5.31}$$

which is exactly the same as the derived formula in Eq.(5.23). It is worth mentioning that with the distribution of vehicles, $t'_{B,i}$ can be simply formulated as $t'_{B,i} = \mathbb{P}\{V_n = HV\}\mathbb{E}(t_n|V_n = HV) + \mathbb{P}\{V_n = CAV\}\mathbb{E}(t_n|V_n = CAV)$, the reason for the formulation in Eq.(5.31) is to be consistent with the subsequent analysis of platoon size control. If the platoon size is controlled not to exceed a maximum value k_m , then:

$$\mathbb{P}\{V_n = CAV^k\} = \begin{cases} p_c(1 - p_c w)(p_c w)^{k-1}, & k < k_m \\ p_c(p_c w)^{k_m-1}, & k = k_m \end{cases} \tag{5.32}$$

When $k < k_m$, the n th CAV may be in any position of a platoon, t_n has the following cases:

$$t_n = \begin{cases} t_m, & V_n^c \text{ following } V_{n-1}^h \\ t_c, & V_n^{c,w^+} \text{ following } V_{n-1}^c \\ t_p, & V_n^{c,w^-} \text{ following } V_{n-1}^c. \end{cases}$$

When $k = k_m$, the n th CAV is a leading vehicle of a platoon with full size. Regardless of the platooning willingness of the n th CAV, due to platoon size restriction, an inter-platoon headway should be maintained if V_n^c is following V_{n-1}^c :

$$t_n = \begin{cases} t_m, & V_n^c \text{ following } V_{n-1}^h \\ t_p, & V_n^c \text{ following } V_{n-1}^c. \end{cases}$$

Denote $t''_{B,i}$ as the average occupation time under the control of platoon size:

$$\begin{aligned}
t''_{B,i} &= \mathbb{P}\{V_n = HV\} \mathbb{E}(t_n | V_n = HV) + \sum_{k=1}^{k_m} \mathbb{P}\{V_n = CAV^k\} \mathbb{E}(t_n | V_n = CAV^k) \\
&= p_h(p_c t_{m'} + p_h t_h) + p_c(p_c w)^{k_m-1}(p_h t_m + p_c t_p) \\
&\quad + \sum_{k=1}^{k_m-1} p_c[p_h + p_c(1-w)](p_c w)^{k-1}[p_c w t_c + p_h t_m + p_c(1-w)t_p] \\
&= w p_c^2 t_c + p_c p_h(t_m + t_{m'}) + p_h^2 t_h + (1-w)p_c^2 t_p + p_c(p_c w)^{k_m}(t_p - t_c)
\end{aligned} \tag{5.33}$$

From Eq.(5.33), $t''_{B,i}$ is a decreasing function of k_m , then corollary 5.5 can be concluded.

Corollary 5.5. *If the size of CAV platoon is controlled not to exceed a maximum value k_m , capacity is an increasing function of k_m .*

Further on,

$$\begin{aligned}
t''_{B,i} &= t_{B,i} + p_c(p_c w)^{k_m}(t_p - t_c) \\
&= t_{B,i} + t_{ex}
\end{aligned} \tag{5.34}$$

We have:

$$t''_{B,i} > t_{B,i} \implies C'' < C \tag{5.35}$$

It can be seen that the extra part $t_{ex} = p_c(p_c w)^{k_m}(t_p - t_c)$ is caused by the restriction on the platoon size. Under the size control, even if CAVs can and are willing to form a platoon with a size larger than k_m , they would have to be divided to form multiple platoons with sizes not exceeding k_m . That is, the probability of state $S_1 : \mathbf{c} \rightarrow c$ is no more $w p_c^2$, but the probability of the event that a $V^{c,w+}$ following a full size platoon should be subtracted: $w p_c^2 - p_c(p_c w)^{k_m-1} p_c w$. The subtracted part is restored to state S_5 . According to this logic, without the distribution of CAV platoons, we can also easily get that: $t''_{B,i} = t_{B,i} + p_c(p_c w)^{k_m}(t_p - t_c)$.

That is to say, if no measures are taken, the system is *idle* for $(\tau_p - \tau_c)$ time during each extra inter-platoon service period. Taking advantages of the advanced CAV technology, it is possible to reduce such system losses and even improve intersection

capacity through rational use of the inter-platoon gaps. However, as previously explained, specific control methods are not the focus of this Chapter.

5.4 Numerical analyses

FIGURE 5.7: A typical intersection

In order to more intuitively illustrate the above analytical results, a typical urban road intersection is analysed as a numerical case, and is depicted in Fig.5.7. The intersection consists of 8 traffic streams and 16 potential conflict points (left-turning streams are not considered as they have no conflicts with other motorised lanes). It is worth pointing out that the modeling method in this study can be easily extended to consider non-motorised traffic, by adding the conflict points between crossing walk lanes and the motor vehicle lanes to the modelling.

As modelled in Eq.(5.1) and Eq.(5.8), the objective is to maximise the total sum of the throughput:

$$\max_i \sum_i Q_i, i \in L = \{1, 2, \dots, 8\} \quad (5.36)$$

with 16 constraints for the conflict points. Take conflict point C_1 as an example, its constraint is:

$$Q_1 \times t_{B,1} + Q_3 \times t_{B,3} \leq 3600 \quad (5.37)$$

With specific values of $t_{B,i}$, the problem of Eq.(5.36) can be solved accordingly. It should be noted that the composition of traffic flow on each stream is assumed to be the same hereafter, i.e., $t_{B,i}$ values are identical for all movements in the calculation.

Fig.5.8 shows the effect of platooning willingness when the penetration rate of CAVs is *low*, *medium* and *high* ($p_c = 0.1, 0.5, 0.9$ respectively). The setting of

FIGURE 5.8: Effect of platooning willingness

other values are also given in the figure. As shown, at the same penetration rate, the higher platooning willingness, the higher capacity, and the influence of w is more obvious in the case of higher penetration rate p_c . Corollary 5.1 is supported. In addition, the analytical analysis shows that higher penetration rate p_c does not necessarily correspond to higher capacity. And this point can also be seen from the cases of $w = 0.0, 0.1, 0.2$ in Fig.5.8, where the capacity with $p_c = 0.9$ is less than the capacity when $p_c = 0.1$ and $p_c = 0.5$.

In order to demonstrate the effect of platoon gaps, the other four parameters are taken as fixed in the calculation ($t_{m'} = 2.0s$; $t_h = 2.5s$; $w = 1$; $p_c = 0.5$), and CAV-related headways are assumed to have the following relationship: $t_m = t_c + 0.5$, $t_p = t_c + 1.5$, $t_c \in [0.5, t_h]$. $t_c = t_h = 2.5s$ is considered the baseline, which assumes

that CAVs adopt a non-conservative car following behaviour, so that the required safe headway is less than that of HVs (Hussain et al., 2016). Results are illustrated in Fig.5.9. As can be seen, the capacity decreases with the increase of platoon

FIGURE 5.9: Effect of intra-platoon gap.

gaps, which is consistent with Corollary 5.4, again showing that the advances in CAV technologies benefit the improvement of traffic capacity.

FIGURE 5.10: Effect of penetration rate.

TABLE 5.2: The settings of variables in Fig.5.10

	w	t_c	t_m	$t_{m'}$	t_h	t_p
$A = 0$	1.0	1.0	1.5	2.0	2.5	3.0
$A < 0$	1.0	1.0	2.0	2.5	2.5	3.0
$A > 0, p_m > 1$	0.5	1.0	1.5	2.0	2.5	2.0
$A > 0, 0 < p_m < 1$	0.2	1.5	2.0	2.5	2.5	3.0

Furthermore, the relationship between capacity and p_c is shown in Fig.5.10, in which capacity is calculated with the value of p_c from 0 to 1 with a step size of 0.01. Other parameters are fixed, set to make A and p_m meet the corresponding conditions (see Table 5.2). Noted that CAV technologies should have little effect on the macroscopic distribution of τ_h , so the value of t_h is set to be the same in each case as 2.5s. Comparing Fig.5.10 and Fig.5.6, the analytical results (Corollary 5.2 and 5.3) are supported.

FIGURE 5.11: Comparison of modeling and simulation results

The expectations of the heterogeneous headways are used in the above analysis. To further verify the analytical results, Monte Carlo (MC) simulations are carried out for the numerical case. Departure headway in front of an intersection, rather than headways dedicated for uninterrupted traffic flow, is needed to measure the capacity. For HVs, it is widely accepted that the departure headways roughly follow log-normal distributions (Jin et al., 2009; Tan et al., 2013; Li and Chen, 2017).

Therefore, $\tau_h \sim \text{Lognormal}(\ln 2.5, 0.02^2)$ and $\tau_{m'} \sim \text{Lognormal}(\ln 2.0, 0.02^2)$ are assumed in the simulations. Since the headways related to CAVs (τ_c, τ_m, τ_p) are only affected by the settings of the control systems, rather than complicated human factors, it is assumed that they are uniformly distributed: $\tau_c \sim U(0.5, 1.5)$, $\tau_m \sim U(1.0, 2.0)$, $\tau_p \sim U(2.0, 3.0)$ (Ghiasi et al., 2017). The comparison of modeling and simulation results is shown in Fig.5.11, in which *part* of the MC simulation data is scattered. It can be seen that the approximated capacity by MC (i.e. the expected value of MC simulations) is very similar to the result calculated by the proposed modeling. The correctness of the analytical results is further supported.

5.5 Conclusions

In this Chapter, intersection capacity under mixed traffic is defined as the maximum sum of the throughput of each stream passing through the intersection, and is modelled as the objective of a linear programming (LP) problem under the by *collision-free* constraints. Conflict points are essential for determining the constraints. The average conflict point occupation time t_B constitutes the parameters of the solution of the LP problem. The mathematical basis of formulating t_B is Markov chain theory. The proposed method can take into account not only the layout of the intersection and the interaction between the different streams, but also the different platooning parameters and the distribution of heterogeneous headways ($\tau_c, \tau_m, \tau_{m'}, \tau_h, \tau_p$).

The penetration rate (p_c) of CAVs is analysed to reflect the impact of CAVs on the macro level, and platooning willingness (w) of CAVs is used as a micro factor to consider the behavioural side of the problem. The derivation and analysis reveal that capacity is an increasing function of platooning willingness w and an decreasing function of platoon gaps. This is in line with expectations, the mature of CAV technologies will be beneficial to the performance of traffic management. However, it is not certain that higher p_c will definitely lead to higher intersection capacity, it also depends on the relationship between the heterogenous headways. It is also showed that if there are some control strategies that enable the platooning willingness to a certain range, the positive effect of p_c could be deterministic and no longer dependent on the HV-related headways. Moreover, if the platoon size is controlled not to exceed a maximum value k_m , capacity is also an increasing

function of k_m . Numerical analyses are carried out to corroborate the analytical conclusions.

Based on the above conclusions on how different variables affect the capacity, targeted intersection management (such as, setting up CAV dedicated lanes) under mixed traffic situations is one of the future tasks. In addition, with the goal of improving intersection capacity, how to use the advanced features of CAVs to coordinate their interaction with non-motorised vehicles is also one of the research directions.

Chapter 6

Summary

This thesis revolves around the management of connected and automated vehicles (CAVs) at intersections. Utilising the advanced features of CAVs to improve the performance of intersection management, comparing the performance of different CAV trajectory organisations and exploring the impacts of CAVs on intersection capacity under mixed traffic environment are the focuses of this thesis.

Part I of the thesis is devoted to the optimisation of autonomous intersection management (AIM) under pure CAVs environment. The advanced features of the new-generation vehicles have the potential to prescribe for the cruxes of intersection management. Wireless communication technology of CAVs not only enables the observation and acquisition of real-time traffic states, but also makes it possible to optimise the coordination among vehicles. Chapter 2 proposes a multi-agent formulation of AIM, in which CAVs cooperate to minimise intersection delay with collision-free constraints. The huge dimension scale incurred by the nature of multi-agent decision making problems is handled by a decentralised coordination learning approach. The main feature of the proposed approach is being able to explicitly identify and dynamically adapt CAV coordination needs during the learning process so that the curse of dimensionality and environment nonstationarity problems in multi-agent learning can be alleviated. The effectiveness of the proposed method is demonstrated under a variety of traffic conditions, and the best performances compared to the other three benchmarks are illustrated. The single-objective optimisation problem solved in Chapter 2 is extended to the multiple conflicting-objective optimisation problem in Chapter 3, in which two concrete

challenges in finding the optimal decisions are tackled: (1) the objectives are essentially conflicting (traffic fairness and efficiency); (2) the relative preferences of the multiple objectives are unknown. An multi-objective reinforcement learning algorithm is customised for optimisation. The proposed approach is evaluated in a simulated real-world isolated intersection. Expected performance is achieved.

Part II of the thesis contributes to theoretically comparing the performance of signal-free intersection management and roundabout under CAV environment. Roundabout adopting first-come-first-serve policy (R-FCFS) and roundabout using the major-minor priority pattern (R-MM) are studied and compared to the most representative signal-free control protocol, the autonomous intersection management (AIM) using FCFS policy (I-FCFS). The theoretical formulas of the capacity and average delay of each strategy are derived by applying M/G/1 queueing model. The safety time gaps for two vehicles consecutively and safely passing conflict points are the main variables, which are assumed to be generally distributed in the theoretical derivation, rather than following a specific distribution. Comparing the derived formulas, R-FCFS has the highest capacity and the least average control delay among the three strategies under the same circumstance. The comparison results reveal an important point that the optimal management of CAVs based on circulating trajectory organisation could have at least comparable performance as the optimal policy based on intersectional organisation.

The mixed traffic environment where CAVs share the traffic network with human-driven vehicles (HVs) is inevitable. Part III of the thesis moves on to the impact analysis of CAVs on the intersection capacity under mixed traffic environment. Capacity is defined as the maximum sum of the throughput of each stream crossing the intersection, which is modelled as the objective function of linear programming under the *collision-free* constraints. The conflict points within the intersection determine the constraints, and the average conflict point occupation time t_B is the key to the optimal solution. Markov chain theory is applied to specify the value of t_B , and the specification relies on the spatial distribution of heterogeneous vehicle headways among mixed traffic. In the model, not only the CAV penetration rate and the three core platooning parameters (*intra-platoon gap*, *inter-platoon gap* and *platoon size*) are captured, but also the individual willingness of CAVs to form platoons is involved to consider the behavioural side of the problem.

Along the line of this thesis, there are a number of interesting research directions for future research.

- In part I, the optimisations are implemented at isolated intersections. It is worthwhile to investigate the coordination of multiple intersections. The framework of autonomous intersection management can be extended to network level, aiming at alleviating the congestion of traffic system. Furthermore, only homogenous CAV traffic flow is considered. Studying the right-of-way assignment problem of heterogeneous traffic can be the focus of future work, including the mixing of CAVs with prioritised vehicles (such as buses), with HVs, or even with non-motorised traffic flow. The proposed approach can easily be transplanted to the situation of heterogeneous traffic flow. The key lies in the acquisition of real-time traffic states, which involves the effects of different CAV market penetration rates. Using partial CAV information to infer the overall state of mixed traffic flow will be an interesting topic.
- In part II, the comparison results indicate that roundabout should be given more attention in the signal-free management for CAVs. For the same incoming traffic demand under CAVs environment, with the same optimisation goals, comparing the performance of intersection and roundabout is one of our future work.
- In part III, the effects of different CAV features on the intersection capacity under mixed traffic environment are analysed. Based on the concluded corollaries, targeted intersection management, such as CAV dedicated lanes, under mixed traffic situations is worthy of investigating. In addition, as only motorised traffic are focused in this thesis, how to use the advanced features of CAVs to coordinate the interaction with non-motorised traffic is also one of the research directions.

List of Author's Publications

Journal Articles

- **Wu, Yuanyuan**, Haipeng Chen, and Feng Zhu, “DCL-AIM: Decentralized coordination learning of autonomous intersection management for connected and automated vehicles,” *Transportation Research Part C: Emerging Technologies* 103 (2019): 246-260.
- **Wu, Yuanyuan**, Feng Zhu, “Intersection capacity under mixed traffic environment with connected and automated vehicles.” (submitted to *Transportation Research Part C: Emerging Technologies*).

Working Papers

- **Wu, Yuanyuan**, Feng Zhu, “Comparative study on roundabout and intersection management under connected and automated vehicles environment.”
- **Wu, Yuanyuan**, Feng Zhu, “Optimization of multiple conflicting objectives in intersection management for connected and automated vehicles.”

References

- 5GAutomotiveAssociation (2017), “An assessment of LTE-V2X (PC5) and 802.11p direct communications technologies for improved road safety in the EU”, 5G Automot. Assoc., Munich, Germany, 5GAA White Paper. [2](#)
- Abdulhai, B., Pringle, R. and Karakoulas, G. J. (2003), “Reinforcement learning for true adaptive traffic signal control”, Journal of Transportation Engineering, vol. 129, no. 3, 278–285. [15](#)
- Akcelik, R., Chung, E. and Besley, M. (1998), Roundabouts: Capacity and performance analysis, ARR 321. [75](#)
- Akcelik, R., Chung, E., Besley, M. et al. (1996), “Performance of roundabouts under heavy demand conditions”, Road and Transport Research, vol. 5, 36–57. [75](#)
- Amoozadeh, M., Deng, H., Chuah, C.-N., Zhang, H. M. and Ghosal, D. (2015), “Platoon management with cooperative adaptive cruise control enabled by VANET”, Vehicular communications, vol. 2, no. 2, 110–123. [84](#)
- Arel, I., Liu, C., Urbanik, T. and Kohls, A. (2010), “Reinforcement learning-based multi-agent system for network traffic signal control”, IET Intelligent Transport Systems, vol. 4, no. 2, 128–135. [15](#)
- Bashiri, M. and Fleming, C. H. (2017), “A platoon-based intersection management system for autonomous vehicles”, 2017 IEEE Intelligent Vehicles Symposium (IV), IEEE, pp. 667–672. [40](#), [55](#)
- Bazzan, A. L. and Klügl, F. (2014), “A review on agent-based technology for traffic and transportation”, The Knowledge Engineering Review, vol. 29, no. 3, 375–403. [15](#)
- Bellman, R. (1966), “Dynamic programming”, Science, vol. 153, no. 3731, 34–37. [47](#)
- Boon, M. A., Mandjes, M., Núñez-Queija, R. et al. (2019), “Congestion analysis of unsignalized intersections: The impact of impatience and markov platooning”, European Journal of Operational Research, vol. 273, no. 3, 1026–1035. [70](#), [75](#)

- Brilon, W. and Wu, N. (2001), “Capacity at unsignalized intersections derived by conflict technique”, Transportation Research Record, vol. 1776, no. 1, 82–90. [xviii](#), [88](#), [89](#), [90](#)
- Busoniu, L., Babuska, R. and De Schutter, B. (2008), “A comprehensive survey of multiagent reinforcement learning”, IEEE Trans. Systems, Man, and Cybernetics, Part C, vol. 38, no. 2, 156–172. [16](#)
- Cai, Y., Wang, H., Ong, G. P., Meng, Q. and Lee, D.-H. (2019), “Investigating user perception on autonomous vehicle (AV) based mobility-on-demand (MOD) services in singapore using the logit kernel approach”, Transportation, vol. 46, no. 6, 2063–2080. [2](#)
- Carlino, D., Boyles, S. D. and Stone, P. (2013), “Auction-based autonomous intersection management”, 2013 16th International IEEE Conference on Intelligent Transportation Systems-(ITSC), IEEE, pp. 529–534. [14](#)
- Chen, H., An, B., Sharon, G., Hanna, J. P., Stone, P., Miao, C. and Soh, Y. C. (2018), “DyETC: Dynamic electronic toll collection for traffic congestion alleviation”, Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-18). [15](#), [16](#)
- Chen, L. and Englund, C. (2016), “Cooperative intersection management: A survey”, IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 2, 570–586. [3](#), [12](#), [62](#)
- Debada, E. G. and Gillet, D. (2018), “Virtual vehicle-based cooperative maneuver planning for connected automated vehicles at single-lane roundabouts”, IEEE Intelligent Transportation Systems Magazine, vol. 10, no. 4, 35–46. [64](#)
- Dresner, K. and Stone, P. (2004), “Multiagent traffic management: A reservation-based intersection control mechanism”, Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2, IEEE Computer Society, pp. 530–537. [12](#), [14](#), [17](#), [40](#), [41](#), [62](#)
- Dresner, K. and Stone, P. (2006), “Human-usable and emergency vehicle-aware control policies for autonomous intersection management”, Fourth International Workshop on Agents in Traffic and Transportation (ATT), Hakodate, Japan. [12](#), [14](#)
- Dresner, K. and Stone, P. (2008), “A multiagent approach to autonomous intersection management”, Journal of Artificial Intelligence Research, vol. 31, 591–656. [4](#), [14](#), [40](#), [63](#)
- Dresner, K. M. (2009), “Autonomous intersection management”, Tech. rep., University of Texas at Austin Austin United States. [34](#)
- Fagnant, D. J. and Kockelman, K. (2015), “Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations”, Transportation Research Part A: Policy and Practice, vol. 77, 167–181. [12](#), [62](#), [83](#)

- Fajardo, D., Au, T.-C., Waller, S., Stone, P. and Yang, D. (2011), “Automated intersection control: Performance of future innovation versus current traffic signal control”, Transportation Research Record: Journal of the Transportation Research Board, , no. 2259, 223–232. [13](#), [63](#)
- Feng, S., Zhang, Y., Li, S. E., Cao, Z., Liu, H. X. and Li, L. (2019a), “String stability for vehicular platoon control: Definitions and analysis methods”, Annual Reviews in Control. [84](#)
- Feng, Y., He, D. and Guan, Y. (2019b), “Composite platoon trajectory planning strategy for intersection throughput maximization”, IEEE Transactions on Vehicular Technology. [84](#)
- Gawron, J. H., Keoleian, G. A., De Kleine, R. D., Wallington, T. J. and Kim, H. C. (2018), “Life cycle assessment of connected and automated vehicles: Sensing and computing subsystem and vehicle level effects”, Environmental science & technology, vol. 52, no. 5, 3249–3256. [83](#)
- Ghiasi, A., Hussain, O., Qian, Z. S. and Li, X. (2017), “A mixed traffic capacity analysis and lane management model for connected automated vehicles: A markov chain method”, Transportation Research Part B: Methodological, vol. 106, 266–292. [76](#), [84](#), [103](#)
- Ghiasi, A., Li, X. and Ma, J. (2019), “A mixed traffic speed harmonization model with connected autonomous vehicles”, Transportation Research Part C: Emerging Technologies, vol. 104, 210–233. [84](#)
- Gong, S. and Du, L. (2018), “Cooperative platoon control for a mixed traffic flow including human drive vehicles and connected and autonomous vehicles”, Transportation research part B: methodological, vol. 116, 25–61. [84](#)
- Gong, Y., Abdel-Aty, M., Yuan, J. and Cai, Q. (2020), “Multi-objective reinforcement learning approach for improving safety at intersections with adaptive traffic signal control”, Accident Analysis & Prevention, vol. 144, 105655. [43](#)
- Guestrin, C., Venkataraman, S. and Koller, D. (2002), “Context-specific multiagent coordination and planning with factored MDPs”, AAAI/IAAI, pp. 253–259. [16](#)
- Guo, Y., Ma, J., Xiong, C., Li, X., Zhou, F. and Hao, W. (2019), “Joint optimization of vehicle trajectories and intersection controllers with connected automated vehicles: Combined dynamic programming and shooting heuristic approach”, Transportation Research Part C: Emerging Technologies, vol. 98, 54–72. [15](#), [42](#)
- Han, X., Ma, R. and Zhang, H. M. (2020), “Energy-aware trajectory optimization of CAV platoons through a signalized intersection”, Transportation Research Part C: Emerging Technologies, vol. 118, 102652. [41](#), [49](#)
- Hausknecht, M., Au, T.-C. and Stone, P. (2011), “Autonomous intersection management: Multi-intersection optimization”, Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on, IEEE, pp. 4581–4586. [14](#), [63](#)

- HCM (2000), Highway Capacity Manual, Transportation Research Board, Washington, D.C. 90
- HCM (2016), Highway Capacity Manual 6th edition, Transportation Research Board, Washington, D.C. 63
- Heidemann, D. and Wegmann, H. (1997), “Queueing at unsignalized intersections”, Transportation research part B: methodological, vol. 31, no. 3, 239–263. 67
- Hussain, O., Ghiasi, A. and Li, X. (2016), “Freeway lane management approach in mixed traffic environment with connected autonomous vehicles”, arXiv preprint arXiv:1609.02946. 101
- IHSMarkit (2017), “The future of cars 2040”, <https://news.ihsmarket.com/press-release/energy-power-media/future-cars-2040-miles-traveled-will-soar-while-sales-new-vehicles->. 3, 62
- Isele, D., Rahimi, R., Cosgun, A., Subramanian, K. and Fujimura, K. (2018), “Navigating occluded intersections with autonomous vehicles using deep reinforcement learning”, 2018 IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp. 2034–2039. 15
- Jiang, Y., Li, S. and Shamo, D. E. (2006), “A platoon-based traffic signal timing algorithm for major–minor intersection types”, Transportation Research Part B: Methodological, vol. 40, no. 7, 543–562. 77
- Jin, L., Cicic, M., Johansson, K. H. and Amin, S. (2020), “Analysis and design of vehicle platooning operations on mixed-traffic highways”, IEEE Transactions on Automatic Control. 85
- Jin, Q., Wu, G., Boriboonsomsin, K. and Barth, M. (2012), “Multi-agent intersection management for connected vehicles using an optimal scheduling approach”, Connected Vehicles and Expo (ICCVE), 2012 International Conference on, IEEE, pp. 185–190. 15
- Jin, Q., Wu, G., Boriboonsomsin, K. and Barth, M. (2013), “Platoon-based multi-agent intersection management for connected vehicle”, 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), IEEE, pp. 1462–1467. 40
- Jin, X., Zhang, Y., Wang, F., Li, L., Yao, D., Su, Y. and Wei, Z. (2009), “Departure headways at signalized intersections: A log-normal distribution model approach”, Transportation research part C: emerging technologies, vol. 17, no. 3, 318–327. 102
- Kenney, J. B. (2011), “Dedicated short-range communications (DSRC) standards in the united states”, Proceedings of the IEEE, vol. 99, no. 7, 1162–1182. 27, 52

- Khamis, M. A. and Gomaa, W. (2014), “Adaptive multi-objective reinforcement learning with hybrid exploration for traffic signal control based on cooperative multi-agent framework”, Engineering Applications of Artificial Intelligence, vol. 29, 134–151. [43](#)
- Khayatian, M., Mehrabian, M., Andert, E., Dedinsky, R., Choudhary, S., Lou, Y. and Shirvastava, A. (2020), “A survey on intersection management of connected autonomous vehicles”, ACM Transactions on Cyber-Physical Systems, vol. 4, no. 4, 1–27. [xvii](#), [42](#), [43](#)
- Kim, I. Y. and De Weck, O. (2006), “Adaptive weighted sum method for multi-objective optimization: A new method for pareto front generation”, Structural and Multidisciplinary Optimization, vol. 31, no. 2, 105–116. [46](#)
- Kimber, R. (1980), “the traffic capacity of roundabouts, trrl laboratory report 942”, Transport and Road Research Laboratory, Crowthorne, United Kingdom. [63](#)
- Kingma, D. P. and Ba, J. (2014), “Adam: A method for stochastic optimization”, arXiv preprint arXiv:1412.6980. [52](#)
- Kok, J. R. and Vlassis, N. (2004), “Sparse cooperative Q-learning”, Proceedings of the Twenty-first International Conference on Machine Learning, ACM, p. 61. [16](#)
- Kokkinogenis, Z., Teixeira, M., d’Orey, P. M. and Rossetti, R. J. (2019), “Tactical level decision-making for platoons of autonomous vehicles using auction mechanisms”, 2019 IEEE Intelligent Vehicles Symposium (IV), IEEE, pp. 1632–1638. [85](#)
- Kumaravel, S. D., Malikopoulos, A. A. and Ayyagari, R. (2020), “Optimal coordination of platoons of connected and automated vehicles at signal-free intersections”, arXiv preprint arXiv:2001.04866. [42](#)
- Lee, J., Park, B. B., Malakorn, K. and So, J. J. (2013), “Sustainability assessments of cooperative vehicle intersection control at an urban corridor”, Transportation Research Part C: Emerging Technologies, vol. 32, 193–206. [14](#)
- Levin, M. W., Boyles, S. D. and Patel, R. (2016), “Paradoxes of reservation-based intersection controls in traffic networks”, Transportation Research Part A: Policy and Practice, vol. 90, 14–25. [13](#), [40](#)
- Levin, M. W., Fritz, H. and Boyles, S. D. (2017), “On optimizing reservation-based intersection controls”, IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 3, 505–515. [15](#), [18](#), [63](#)
- Li, L. and Chen, X. M. (2017), “Vehicle headway modeling and its inferences in macroscopic/microscopic traffic flow theory: A survey”, Transportation Research Part C: Emerging Technologies, vol. 76, 170–188. [102](#)

- Li, Y., Chen, S., Du, R., Ha, P. Y. J., Dong, J. and Labi, S. (2020), “Using empirical trajectory data to design connected autonomous vehicle controllers for traffic stabilization”, arXiv preprint arXiv:2010.05440. 84
- Li, Z., Chitturi, M., Zheng, D., Bill, A. and Noyce, D. (2013), “Modeling reservation-based autonomous intersection control in Vissim”, Transportation Research Record: Journal of the Transportation Research Board, , no. 2381, 81–90. 13
- Lioris, J., Pedarsani, R., Tascikaraoglu, F. Y. and Varaiya, P. (2017), “Platoons of connected vehicles can double throughput in urban roads”, Transportation Research Part C: Emerging Technologies, vol. 77, 292–305. 64, 84
- Little, J. D. and Graves, S. C. (2008), “Little’s law”, Building intuition, Springer, pp. 81–100. 75
- Littman, M. L. (2001), “Value-function reinforcement learning in Markov games”, Cognitive Systems Research, vol. 2, no. 1, 55–66. 16
- Liu, M., Wang, M. and Hoogendoorn, S. (2019), “Optimal platoon trajectory planning approach at arterials”, Transportation Research Record, 0361198119847474. 84
- Lokhandwala, M. and Cai, H. (2018), “Dynamic ride sharing using traditional taxis and shared autonomous taxis: A case study of NYC”, Transportation Research Part C: Emerging Technologies, vol. 97, 45–60. 12
- Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P. and Wießner, E. (2018), “Microscopic traffic simulation using sumo”, The 21st IEEE International Conference on Intelligent Transportation Systems, IEEE, URL <<https://elib.dlr.de/124092/>>. 50, 77
- Lu, Q. and Kim, K.-D. (2016), “Intelligent intersection management of autonomous traffic using discrete-time occupancies trajectory”, Journal of Traffic and Logistics Engineering Vol, vol. 4, no. 1, 1–6. 55
- Malikopoulos, A. A., Cassandras, C. G. and Zhang, Y. J. (2018), “A decentralized energy-optimal control framework for connected automated vehicles at signal-free intersections”, Automatica, vol. 93, 244–256. 4, 15, 41, 49, 63
- Mannion, P., Duggan, J. and Howley, E. (2016), “An experimental review of reinforcement learning algorithms for adaptive traffic signal control”, Autonomic Road Transport Support Systems, Springer, pp. 47–66. 15, 43
- Matignon, L., Laurent, G. J. and Le Fort-Piat, N. (2012), “Independent reinforcement learners in cooperative markov games: A survey regarding coordination problems”, The Knowledge Engineering Review, vol. 27, no. 1, 1–31. 16

- Mauro, R. and Cattani, M. (2012), “Functional and economic evaluations for choosing road intersection layout”, PROMET-Traffic&Transportation, vol. 24, no. 5, 441–448. [62](#)
- Mauro, R. and Pompigna, A. (2020), “State of the art and computational aspects of time-dependent waiting models for non-signalised intersections”, Journal of Traffic and Transportation Engineering (English Edition). [63](#)
- Medhi, J. (2002), Stochastic models in queueing theory, Elsevier. [75](#)
- Melo, F. S. and Veloso, M. (2011), “Decentralized MDPs with sparse interactions”, Artificial Intelligence, vol. 175, no. 11, 1757–1789. [16](#)
- Milanés, V. and Shladover, S. E. (2014), “Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data”, Transportation Research Part C: Emerging Technologies, vol. 48, 285–300. [50](#), [77](#)
- Milanés, V., Shladover, S. E., Spring, J., Nowakowski, C., Kawazoe, H. and Nakamura, M. (2013), “Cooperative adaptive cruise control in real traffic situations”, IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 1, 296–305. [84](#)
- Mirheli, A., Hajibabai, L. and Hajbabaie, A. (2018), “Development of a signal-head-free intersection control logic in a fully connected and autonomous vehicle environment”, Transportation Research Part C: Emerging Technologies, vol. 92, 412–425. [15](#), [41](#), [63](#)
- Mirheli, A., Tajalli, M., Hajibabai, L. and Hajbabaie, A. (2019), “A consensus-based distributed trajectory control in a signal-free intersection”, Transportation Research Part C: Emerging Technologies, vol. 100, 161–176. [15](#)
- Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M. (2013), “Playing atari with deep reinforcement learning”, arXiv preprint arXiv:1312.5602. [47](#)
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G. et al. (2015), “Human-level control through deep reinforcement learning”, Nature, vol. 518, no. 7540, 529. [47](#)
- Molina-Masegosa, R. and Gozalvez, J. (2017), “LTE-V for sidelink 5G V2X vehicular communications: A new 5G technology for short-range vehicle-to-everything communications”, IEEE Vehicular Technology Magazine, vol. 12, no. 4, 30–39. [39](#), [84](#)
- Nagel, K. and Schreckenberg, M. (1992), “A cellular automaton model for freeway traffic”, Journal de physique I, vol. 2, no. 12, 2221–2229. [19](#)

- Nor, M. H. B. M. and Namerikawa, T. (2018), “Merging of connected and automated vehicles at roundabout using model predictive control”, 2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), IEEE, pp. 272–277. [64](#)
- NYSDOT (2018), “NYSDOT’s web site dedicated to modern roundabouts”, <https://www.dot.ny.gov/main/roundabouts>. [64](#)
- Papadoulis, A., Quddus, M. and Imprialou, M. (2019), “Evaluating the safety impact of connected and autonomous vehicles on motorways”, Accident Analysis & Prevention, vol. 124, 12–22. [84](#)
- Patro, S. and Sahu, K. K. (2015), “Normalization: A preprocessing stage”, arXiv preprint arXiv:1503.06462. [48](#)
- Peng, J. and Williams, R. J. (1994), “Incremental multi-step Q-learning”, Machine Learning Proceedings 1994, Elsevier, pp. 226–232. [16](#)
- Pompigna, A., Guerrieri, M. and Mauro, R. (2020), “New extensions and applications of the modified chumanov model for calculating entry capacity of single-lane roundabouts”, Sustainability, vol. 12, no. 15, 6122. [63](#)
- Qi, X., Luo, Y., Wu, G., Boriboonsomsin, K. and Barth, M. (2019), “Deep reinforcement learning enabled self-learning control for energy efficient driving”, Transportation Research Part C: Emerging Technologies, vol. 99, 67–81. [15](#)
- Rădulescu, R., Mannion, P., Roijers, D. M. and Nowé, A. (2020), “Multi-objective multi-agent decision making: A utility-based analysis and survey”, Autonomous Agents and Multi-Agent Systems, vol. 34, no. 1, 10. [46](#)
- Roth, M., Simmons, R. and Veloso, M. (2007), “Exploiting factored representations for decentralized execution in multiagent teams”, Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multiagent Systems, ACM, p. 72. [21](#)
- Schaul, T., Quan, J., Antonoglou, I. and Silver, D. (2015), “Prioritized experience replay”, arXiv preprint arXiv:1511.05952. [51](#)
- Schepperle, H. and Böhm, K. (2007), “Agent-based traffic control using auctions”, International Workshop on Cooperative Information Agents, Springer, pp. 119–133. [12](#), [14](#)
- Schepperle, H. and Böhm, K. (2008), “Auction-based traffic management: Towards effective concurrent utilization of road intersections”, 2008 10th IEEE Conference on E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing, IEEE, pp. 105–112. [14](#)
- Shahri, P. K., Ghasemi, A. H. and Izadi, V. (2020), “Optimal lane management in heterogeneous traffic network using extremum seeking approach”, Tech. rep., SAE Technical Paper. [85](#)

- Shladover, S. E., Nowakowski, C., Lu, X.-Y. and Ferlis, R. (2015), “Cooperative adaptive cruise control: Definitions and operating concepts”, Transportation Research Record, vol. 2489, no. 1, 145–152. [84](#)
- Shladover, S. E., Su, D. and Lu, X.-Y. (2012), “Impacts of cooperative adaptive cruise control on freeway traffic flow”, Transportation Research Record, vol. 2324, no. 1, 63–70. [84](#)
- Simchi-Levi, D. and Trick, M. A. (2011), “Introduction to ‘little’s law as viewed on its 50th anniversary’”, Operations Research, vol. 59, no. 3, 535–535. [75](#)
- Sun, W., Zheng, J. and Liu, H. X. (2017), “A capacity maximization scheme for intersection management with automated vehicles”, Transportation Research Procedia, vol. 23, 121–136. [15](#), [41](#), [63](#)
- Sun, X. and Yin, Y. (2019), “Behaviorally stable vehicle platooning for energy savings”, Transportation Research Part C: Emerging Technologies, vol. 99, 37–52. [85](#)
- Sutton, R. S. and Barto, A. G. (1998), Reinforcement learning: An introduction, vol. 1, MIT Press Cambridge. [13](#)
- Tan, J., Li, L., Li, Z. and Zhang, Y. (2013), “Distribution models for start-up lost time and effective departure flow rate”, Transportation Research Part A: Policy and Practice, vol. 51, 1–11. [102](#)
- Tan, M. (1993), “Multi-agent reinforcement learning: Independent vs. cooperative agents”, Proceedings of the Tenth International Conference on Machine Learning, pp. 330–337. [16](#)
- Tanner, J. (1962), “A theoretical analysis of delays at an uncontrolled intersection”, Biometrika, vol. 49, no. 1/2, 163–170. [68](#)
- Timmerman, R. and Boon, M. A. (2021), “Platoon forming algorithms for intelligent street intersections”, Transportmetrica A: Transport Science, vol. 17, no. 3, 278–307. [40](#)
- Tu, R., Alfaseeh, L., Djavadian, S., Farooq, B. and Hatzopoulou, M. (2019), “Quantifying the impacts of dynamic control in connected and automated vehicles on greenhouse gas emissions and urban no2 concentrations”, Transportation Research Part D: Transport and Environment, vol. 73, 142–151. [83](#)
- Vahidi, A. and Sciarretta, A. (2018), “Energy saving potentials of connected and automated vehicles”, Transportation Research Part C: Emerging Technologies, vol. 95, 822–843. [83](#)
- Van Arem, B., Van Driel, C. J. and Visser, R. (2006), “The impact of cooperative adaptive cruise control on traffic-flow characteristics”, IEEE Transactions on Intelligent Transportation Systems, vol. 7, no. 4, 429–436. [63](#)

- Van Hasselt, H., Guez, A. and Silver, D. (2016), “Deep reinforcement learning with double q-learning”, Proceedings of the AAAI conference on artificial intelligence, vol. 30. [51](#)
- Vanderbei, R. J. (2020), Linear programming: Foundations and extensions, vol. 285, Springer Nature. [88](#)
- Vasirani, M. and Ossowski, S. (2012), “A market-inspired approach for intersection management in urban road traffic networks”, Journal of Artificial Intelligence Research, vol. 43, 621–659. [14](#)
- Veith, G. (2010), “Guide to road design, part 4A: Unsignalised and signalised intersections”, Tech. rep. [63](#)
- Viswanadham, N. and Narahari, Y. (2015), Performance modeling of automated systems, PHI Learning Pvt. Ltd. [93](#)
- Wang, J., Zhao, X. and Yin, G. (2018), “Multi-objective optimal cooperative driving for connected and automated vehicles at non-signalised intersection”, IET Intelligent Transport Systems, vol. 13, no. 1, 79–89. [41](#)
- Wang, Y., Cai, P. and Lu, G. (2020), “Cooperative autonomous traffic organization method for connected automated vehicles in multi-intersection road networks”, Transportation research part C: emerging technologies, vol. 111, 458–476. [40](#)
- Watkins, C. J. and Dayan, P. (1992), “Q-learning”, Machine Learning, vol. 8, no. 3-4, 279–292. [16](#), [25](#), [47](#)
- Watson, L. T. and Haftka, R. T. (1989), “Modern homotopy methods in optimization”, Computer Methods in Applied Mechanics and Engineering, vol. 74, no. 3, 289–305. [48](#)
- Webster, F. V. (1958), “Traffic signal settings”, Tech. rep. [32](#)
- Wierman, A. (2011), “Fairness and scheduling in single server queues”, Surveys in Operations Research and Management Science, vol. 16, no. 1, 39–48. [42](#)
- Wu, N. (1998), “Impedance effects for streams of higher ranks at unsignalized intersections”, Proceedings of the Third International Symposium on Highway Capacity. [69](#)
- Wu, N. (2000), “Capacity at all-way stop-controlled and first-in-first-out intersections”, Proceedings of the 4th International Symposium on Highway Capacity, Hawaii, Transportation Research Circular E-C018. [89](#)
- Wu, N. (2001), “A universal procedure for capacity determination at unsignalized (priority-controlled) intersections”, Transportation Research Part B: Methodological, vol. 35, no. 6, 593–623. [75](#)

- Wu, Y., Chen, H. and Zhu, F. (2019), “DCL-AIM: Decentralized coordination learning of autonomous intersection management for connected and automated vehicles”, Transportation Research Part C: Emerging Technologies, vol. 103, 246–260. [40](#), [41](#), [78](#)
- Yang, R., Sun, X. and Narasimhan, K. (2019), “A generalized algorithm for multi-objective reinforcement learning and policy adaptation”, Advances in Neural Information Processing Systems, pp. 14636–14647. [47](#)
- Yang, X., Li, X. and Xue, K. (2004), “A new traffic-signal control for modern roundabouts: Method and application”, IEEE Transactions on Intelligent Transportation Systems, vol. 5, no. 4, 282–287. [4](#), [64](#)
- Yin, S., Duan, H., Li, Z. and Zhang, Y. (2010), “Multi-objective reinforcement learning for traffic signal coordinate control”, 1th World Conference on Transport Research. [43](#)
- Yu, C., Sun, W., Liu, H. X. and Yang, X. (2019), “Managing connected and automated vehicles at isolated intersections: From reservation-to optimization-based methods”, Transportation research part B: methodological, vol. 122, 416–435. [41](#), [78](#)
- Yu, C., Zhang, M., Ren, F. and Tan, G. (2015), “Multiagent learning of coordination in loosely coupled multiagent systems”, IEEE Transactions on Cybernetics, vol. 45, no. 12, 2853–2867. [16](#), [22](#)
- Zhang, Y. and Cassandras, C. G. (2019), “Decentralized optimal control of connected automated vehicles at signal-free intersections including comfort-constrained turns and safety guarantees”, Automatica, vol. 109, 108563. [42](#)
- Zhang, Y. J., Malikopoulos, A. A. and Cassandras, C. G. (2016), “Optimal control and coordination of connected and automated vehicles at urban traffic intersections”, American Control Conference (ACC), 2016, IEEE, pp. 6227–6232. [15](#), [41](#)
- Zhao, L., Malikopoulos, A. and Rios-Torres, J. (2018), “Optimal control of connected and automated vehicles at roundabouts: An investigation in a mixed-traffic environment”, IFAC-PapersOnLine, vol. 51, no. 9, 73–78. [64](#)
- Zhao, X., Wang, J., Yin, G. and Zhang, K. (2019), “Cooperative driving for connected and automated vehicles at non-signalized intersection based on model predictive control”, 2019 IEEE Intelligent Transportation Systems Conference (ITSC), IEEE, pp. 2121–2126. [42](#)
- Zhong, Z., Lee, E. E., Nejad, M. and Lee, J. (2020), “Influence of cav clustering strategies on mixed traffic flow characteristics: An analysis of vehicle trajectory data”, Transportation Research Part C: Emerging Technologies, vol. 115, 102611. [85](#)

- Zhou, J. and Zhu, F. (2020), “Modeling the fundamental diagram of mixed human-driven and connected automated vehicles”, Transportation Research Part C: Emerging Technologies, vol. 115, 102614. [85](#)
- Zhou, M., Yu, Y. and Qu, X. (2019), “Development of an efficient driving strategy for connected and automated vehicles at signalized intersections: A reinforcement learning approach”, IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 1, 433–443. [42](#)
- Zhu, F. and Ukkusuri, S. V. (2014), “Accounting for dynamic speed limit control in a stochastic traffic environment: A reinforcement learning approach”, Transportation Research Part C: Emerging Technologies, vol. 41, 30–47. [15](#)
- Zhu, F. and Ukkusuri, S. V. (2015), “A linear programming formulation for autonomous intersection control within a dynamic traffic assignment and connected vehicle environment”, Transportation Research Part C: Emerging Technologies, vol. 55, 363–378. [14](#), [41](#), [63](#)
- Zhu, M., Wang, X. and Wang, Y. (2018), “Human-like autonomous car-following model with deep reinforcement learning”, Transportation Research Part C: Emerging Technologies, vol. 97, 348–368. [15](#)