DYNAMIC METERING IN CONNECTED URBAN STREET NETWORKS: IMPROVING MOBILITY

FINAL PROJECT REPORT

by

Ali Hajbabaie
North Carolina State University

Rasool Mohebifard
Washington State University

Sponsorship
University of Washington PacTrans
Washington State University

for
Pacific Northwest Transportation Consortium (PacTrans)
USDOT University Transportation Center for Federal Region 10
University of Washington
More Hall 112, Box 352700
Seattle, WA 98195-2700

In cooperation with the US Department of Transportation- Office of the Assistant Secretary for Research and Technology (OST-R)

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The Pacific Northwest Transportation Consortium, the U.S. Government and matching sponsor assume no liability for the contents or use thereof.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
	01701507	

4. Title and Subtitle	5. Report Date
Dynamic Metering in Connected Urban Street Networks: Improving Mobility	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Ali Hajbabaie, ORCID# 0000-0001-6757-1981; Rasool Mohebifard	2017-S-WSU-2

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
PacTrans Pacific Northwest Transportation Consortium University Transportation Center for Region 10 University of Washington More Hall 112 Seattle, WA 98195-2700	11. Contract or Grant No. 69A355174110

12. Sponsoring Organization Name and Address	13. Type of Report and Period Covered
United States of America Department of Transportation Research and Innovative Technology Administration	Research Final 14. Sponsoring Agency Code

15. Supplementary Notes

Report uploaded at www.pacTrans.org

16. Abstract

Traffic metering offers great potential for reducing congestion in oversaturated urban street networks. The available metering approaches rely on aggregated traffic flow dynamics to find metering rates that keep the number of vehicles inside congested areas within a predefined accumulation range. In such approaches, the aggregated traffic metering rates found for all gates surrounding the congested areas need to be divided among gates using a heuristic approach, which may compromise the quality of the suggested metering strategies.

We developed a methodology to dynamically optimize the metering rates for each metering signal in an urban-street network. We utilized the Benders decomposition technique and developed a solution technique that finds metering strategies within a tight optimality bound. We also proposed distributed optimization and coordination algorithms to be able to optimize traffic metering rates in real time. This approach estimates vehicle densities across the network by using connected vehicle location information and loop detector data at discretized time steps. The estimations are shared among several computational nodes across the network and each node optimizes traffic metering rates for a sub-set of gates for several time steps ahead of the current network time. Once the optimized rates have been found, the rates for only the current network time are applied to the network. This procedure continues until the analysis period has ended. Several numerical analyses showed that traffic metering is an effective strategy to improve traffic operations in congested urban networks.

17. Keywords	18. Distribution Statement
Traffic Metering, Perimeter Control, Connected Vehicles, Distributed and Coordinated Systems, Model Predictive Control	No restrictions.

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.		NA

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CTM: Cell transmission model

CV: Connected vehicles

DOCA: Distributed optimization and coordination algorithms

DTM: Dynamic traffic metering

LB: Lower bound

MFD: Macroscopic fundamental diagram

MINLP: Mixed-integer non-linear program

MPC: Model predictive control

OPT: Optimization

PR: Market penetration rate of connected vehicles

SIM: Simulation

UB: upper bound

Acknowledgments

The authors would like to thank the PacTrans Regional University Transportation Center for their financial support of this research project.

Executive Summary

Traffic metering or perimeter control is an effective means to mitigate congestion on urban-street networks. Bimodal traffic signals (i.e., green and red indications only) can be placed at the borders of congested areas, similar to the way on-ramps regulate the flow of vehicles. The objectives of this study was to develop a methodology for metering traffic on urban street networks and to study the effects of traffic metering on traffic operations on such networks.

This study developed an optimization program that aims to increase the number of completed trips on urban street networks by controlling traffic metering rates. We used the Benders decomposition technique and proposed a solution technique that could solve the program within an optimality bound. The case study results for a network of 20 intersections showed that traffic metering significantly improved traffic operations by reducing the travel time of vehicles inside the network by 30.8 percent and 34.2 percent in comparison to a no-metering strategy. Some vehicles were delayed at the metering signals; however, traffic metering reduced system-level travel times by 2.7 percent and 5.4 percent.

While the solution technique proved effective at providing metering strategies very close to the optimal ones, it did not scale well with the size of the problem, and its runtime increased. This issue was addressed by developing a distributed model predictive control approach. The distributed approach decomposed the network to several sub-networks and allocated computational resources to each sub-network. Each sub-network estimated vehicle density across all of its links at a time step, optimized the metering rate over a planning horizon in the future, shared information with other subnetworks and used the received information in its future optimizations, implemented the optimal metering rates only for the next time step, and then

repeated the process. In our tests, the distributed solution technique found solutions in real time that were at most 2.2 percent different that the optimal solutions.

Sensitivity analysis on the market penetration rate of connected vehicles in a microsimulation environment showed that network performance with traffic metering, in comparison to the no-metering strategy, improved significantly as the penetration rate increased from 0 percent to 30 percent. However, the improvements were negligible for penetration rates of more than 30 percent.

Chapter 1: Introduction

A study of 83 urban areas in the U.S. by the Harvard Center for Risk Analysis projected a $96 billion cost in wasted time due to congestion by 2030 (Levy et al., 2010). Traffic congestion is a major contributing factor to excessive travel delay, fuel consumption, and air pollution in urban areas. Researchers have studied several traffic congestion management strategies to mitigate traffic congestion in urban street networks, such as traffic signal timing (Han et al., 2016; Hosseini and Savla, 2016; Islam et al., 2020; Islam and Hajbabaie, 2017; Liang et al., 2020; Medina et al., 2011, 2010; Mehrabipour and Hajbabaie, 2017; Mohebifard and Hajbabaie, 2019; Nilsson et al., 2015; Stevanovic et al., 2007; Yan et al., 2019; Yang et al., 2016), variable speed limits and speed harmonization (Lee et al., 2006; Li et al., 2015; Tajalli et al., 2020; Tajalli and Hajbabaie, 2018a, 2018b; Zhu and Ukkusuri, 2014), and traffic assignment (Aziz and Ukkusuri, 2012; Jafari et al., 2017; Levin et al., 2016; Mehrabipour et al., 2019; Samaranayake et al., 2018). Traffic metering is another promising strategy for managing traffic congestion that can prevent queue spillovers and gridlocks that can occur in oversaturated conditions. (Mahmassani et al., 2013).

1.1. Background

Daganzo (2007) utilized the idea of macroscopic fundamental diagrams (MFDs) (Godfrey, 1900; Herman and Prigogine, 1979; Mahmassani et al., 1984) to explain the dynamics of vehicle accumulation inside an urban network and vehicle throughput. According to Daganzo (2007), the dynamics of an urban network vary between two extremes of single-link and single-ring systems. In a single-link system, the number of vehicles in a link and the number of exiting vehicles have a relationship similar to that shown in figure 1-1 (a). Therefore, the outflow increases with the density of the link (i.e., number of vehicles in the link) up to the density D_C corresponding the link capacity F_{max}. Figure 1-1 shows the declining part of the diagram with a dashed line because this

part represents congestion that occurs because of restrictions on outflows, such as downstream queue spillbacks. In other words, the declining part does not happen because of further increases of inflow. On the other hand, in the single-ring system of figure 1-1 (b), increases in inflow, and accordingly the density of the system, more than D_C create a decline. In the congested region of a single-ring system, the outflows can be reduced to zero by increasing the inflow and its density to the jam density D_J. Therefore, the declining part of the figure is created by the occurrence of gridlock.

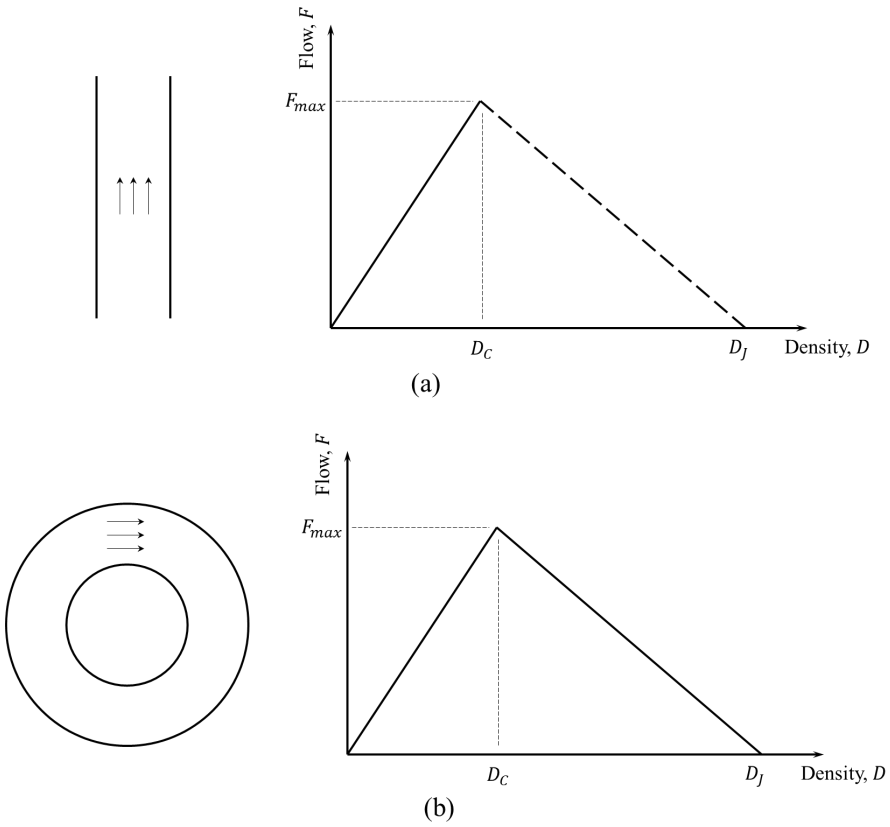


Figure 1-1. . (a) Single-link and (b) single-ring systems with their corresponding triangular fundamental diagrams (Daganzo, 2007)

The dynamics of real networks are somewhere between the dynamics of single-link and single-ring systems. Therefore, the diagram of real networks may have a generic shape similar to that in figure 1-2. In real networks, gridlock happens, but it does not last long because of the nature

of transportation networks. Therefore, the maximum observed density D_v inside the network can be somewhere between D_C and D_J (Daganzo, 2007).

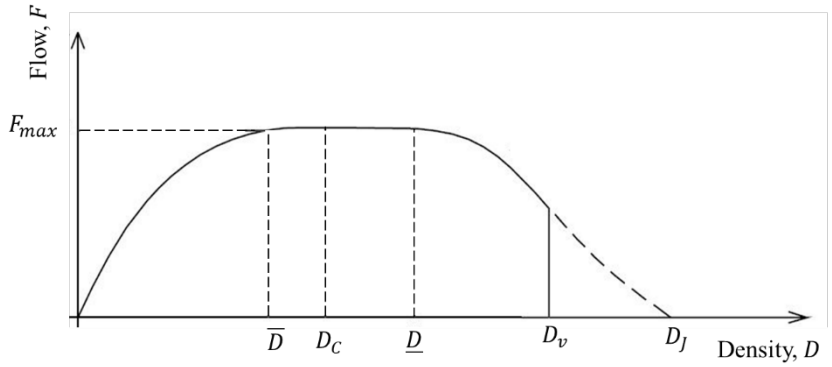


Figure 1-2. A generic fundamental diagram for real transportation networks (Daganzo, 2007)

Consequently, an opportunity exists in a transportation network to control its density and prevent outflow drops. For instance, keeping the density between $\underline{D}$ and $\overline{D}$ in a network with a generic fundament diagram like figure 1-2 ensures that the network can process the maximum number of vehicles without an outflow drop due to excessive congestion. Once the number of entry vehicles to a network has surpassed the optimal accumulation interval $[\underline{D}, \overline{D}]$, "metering signals" will be activated to reduce the excessive number of entry vehicles.

Figure 1-3 shows a portion of an urban network that is equipped with metering signals at the borders of a region that needs to be protected from oversaturated flow conditions. Bimodal traffic signals (with green and red indications) can be placed at the borders of congested areas to regulate the inflow of vehicles. The metering signals can also be traffic signals at intersections, whose timings can be modified to accomlish traffic metering through prolonged red signal indications. Note that the inbound approaches of such intersections should have enough capacity to hold possible queues of vehicles. Otherwise, queue spillovers from the metering signals may deteriorate upstream network performance.

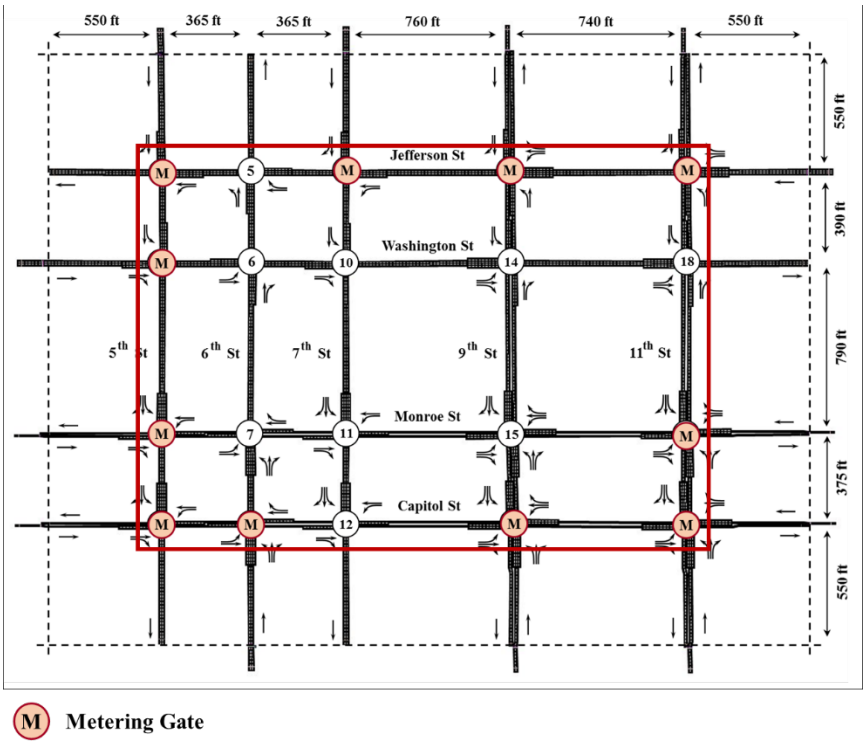


Figure 1-3.. Metering gates that are placed at the borders of a region that needs to be protected from oversaturated flow conditions

1.2. Research Objectives

The goal of traffic metering on urban street networks is to regulate the number of vehicles entering congested parts of the networks to achieve the highest throughput. Most of the available metering approaches directly use the control theory fundamental diagrams discussed above to maintain the number of vehicles within a predefined accumulation interval (e.g., Haddad and Geroliminis, 2012; Keyvan-Ekbatani et al., 2012). While such approaches are easy to implement for large-scale networks, their efficiency relies on the accuracy of optimal accumulation estimation. In other words, the available approaches assume that a well-defined macroscopic fundamental diagram (MFD) that provides an accurate estimate for the optimal vehicle accumulation can be derived for each urban network (or its sub-networks). However, deriving such diagrams for urban networks is not an easy task. Moreover, the complex dynamics of transportation networks cannot be accurately represented with a single accumulation interval (or

multiple intervals for networks with several sub-networks). Therefore, the objective of this study was to develop a methodology for traffic metering on urban street networks. The study was intended to

- Develop an optimization program with explicit representation of traffic dynamics to optimize metering rates at individual network metering signals
- Develop a solution technique that could solve the program with tight optimality bounds
- Develop a distributed model predictive control approach to find metering strategies in real time
- Evaluate the effects of traffic metering on traffic operations on urban street networks.

Accordingly, we developed an optimization program and two solution techniques to optimize traffic metering strategies by explicitly capturing traffic dynamics within the optimization program. The program employs the cell transmission model (CTM) (Daganzo, 1995, 1994) to capture traffic dynamics and maximizes network throughput by optimizing traffic metering rates at predefined gate locations. Because the program has large dimensions, it cannot be solved with regular optimization techniques. Therefore, we proposed a Benders decomposition-based approach (Benders, 1962; Geoffrion, 1972) to solve the program. The results of this approach are near-optimal traffic metering strategies for all metering signals over the entire analysis period, regardless of whether or not they can be found in real time.

We addressed the runtime limitation of the solution technique by proposing distributed optimization and coordination algorithms (DOCA) for dynamic traffic metering (DTM). The DOCA-DTM method decomposes a network into smaller sub-networks and allow each sub-network to optimize its own traffic metering rates and share information with other sub-networks to coordinate their decisions. In addition, the DOCA-DTM is implemented within a model

predictive control (MPC) structure to further reduce the problem complexity and find solutions in real time. Thus, the DOCA-DTM collects connected vehicle and loop detector data at discretized time steps, integrates them and estimates density over the network links, optimizes metering rates over a prediction period (several time steps), implements the optimized metering rates for the next time step, and continues the process over the entire study period. Like to other adaptive approaches (e.g., Hajibabai and Ouyang, 2016; Mirheli et al., 2020; Mirheli and Hajibabai, 2020) MPC allows the methodology to be responsive to dynamic changes in traffic demand and network state.

1.3. Report Organization

This report includes seven chapters. An extensive literature review on available traffic metering approaches, their strengths, and their limitations are discussed in Chapter 2. The mathematical formulation of the traffic metering optimization program is detailed in Chapter 3. A Benders decomposition-based solution technique to solve the traffic metering program centrally is developed in Chapter 4. Chapter 5 presents distributed optimization and coordination algorithms for traffic metering that allows metering strategies to be found in real time. Numerical experiments, including the effects of traffic metering on network performance, the optimality bounds of the central solution technique, the runtime and solution quality of the distributed approach, and the effects of the market penetration rate of connected vehicles on traffic metering, are presented in Chapter 6. Finally, Chapter 7 concludes this report with a summary of the key findings and trends for further research.

Chapter 2: Literature Review

This chapter presents a review of available traffic metering approaches. These approaches are categorized into scenario- and MFD-based studies. The details of each category follow.

2.1. Scenario-Based Traffic Metering

Scenario-based approaches evaluate improvements in network performance that can be achieved by using predefined metering scenarios. These scenarios can be fixed percentages of reductions in the traffic flows passing through the metering signals. Rathi and Lieberman (1989) simulated a portion of Manhattan, New York, and tested different entry volumes. The simulation results indicated that traffic metering had positive effects on improving network performance. Rathi and Lieberman (1989) did not consider the effects of traffic metering on the vehicles delayed at the metering signals and suggested further study for better evaluation of traffic metering benefits.

Hajbabaie and Benekohal (2011) evaluated several metering scenarios. They optimized traffic signals for several traffic metering scenarios so that the operational benefits of traffic metering could be better evaluated. They did not find optimal metering levels, but their results indicated that a metering level existed at which the network would have the best performance. A similar analysis by Medina et al. (2013) suggested that dynamic metering rates offered more efficient network performance than fixed metering levels.

2.2. MFD-Based Traffic Metering

The majority of traffic metering studies have been based on macroscopic fundamental diagrams (MFDs) for urban networks (Daganzo, 2007; Geroliminis and Daganzo, 2008). MFDs provide a relationship between the number of vehicles inside a network and its throughput (see Chapter 1 for more details). Therefore, traffic metering controllers can be designed to maintain the

number of vehicles in the network within the range that corresponds to the highest throughput. Moreover, accurate estimation of the optimal range requires a low scattered and well-defined MFD, which can be expected for networks with a homogenous distribution of traffic. Hence, the primary step in designing the controllers is to divide a network into homogenous sub-networks and estimate a well-defined MFD for each sub-network.

Geroliminis and Sun (2011) stated that despite several empirical studies on the existence of MFDs, these diagrams might not exist for all networks. Moreover, they stated that the shape and scatteredness of such diagrams were a function of a network and its control. Furthermore, they empirically and analytically showed that the spatial distribution of vehicle density was a key component that affected the shape and scatteredness of an MFD.

Zhang et al. (2013) used the stochastic cellular automaton model (De Gier et al., 2011) to compare the MFDs of a network under various conditions of adaptive signal settings and demand rates. Their simulation-based analysis showed that the shape of MFDs was influenced by the demand rate such that their case study network reached its capacity at higher densities when demand was more uniformly distributed in the network rather than when demand was non-uniformly distributed. They also observed that traffic signal settings that distributed density in the network uniformly resulted in an MFD with higher capacity and outflows than an adaptive signal control system.

Ji and Geroliminis (2012) proposed a clustering algorithm so that a heterogeneous network could be divided into several sub-networks, each with smaller link density variances than the original network. This approach allowed sub-networks with well-defined MFDs to be found so that each sub-network could be controlled with its own boundary gates.

Keyvan-Ekbatani et al. (2012) designed a feedback-based controller that maintained the number of vehicles in a network with the optimal range that could be found from the network MFD. They represented the network dynamics with a first-order, non-linear system that was linearized around the optimal accumulation range. The controller determined the total inflows (the number of vehicles that should be allowed to enter the network), and each gate processed a fraction of the total inflow proportional to its saturation flow rate.

Geroliminis et al. (2013) used MFDs to represent the dynamics of a two-region urban network. They formulated an optimal control problem for this system such that the number of completed trips was maximized by regulating the flow of vehicles at the shared boundaries of regions. Hence, the decision variables of the problem were the total number of vehicles that should be allowed to move between the regions. They solved the problem with model predictive control (MPC) so that the errors in the MFD and demand estimation could be considered in the perimeter control decisions. This approach was a high-level control problem that determined the overall metered flows rather than optimizing metering rates for each gate. Moreover, the application of this formulation to a real or simulated network required a detailed enough representation of flow dynamics that the effect of gate locations, metering rates at each gate, and the queues of vehicles at the gates could be captured to indicate network performance.

Haddad et al. (2013) considered perimeter and ramp metering problems on a mixed urban and freeway network. They considered a two-region urban network and one freeway facility that received or sent vehicles to each of those regions. They captured traffic dynamics in the urban regions with MFDs and used an asymmetric cell transmission model (Gomes and Horowitz, 2006) for the freeway facility. Although this study provided invaluable insights about controlling a large-scale and mixed traffic system, the proposed formulation and its MPC-based solution technique

could not capture detailed traffic dynamics and the interactions between the systems. For instance, the effect of queues at the freeway on-ramps was not captured for the dynamics of urban regions. Moreover, the perimeter control between the regions was represented by the total flow that should transfer between the regions, and therefore the effects of gate locations, queues, and traffic metering rates at each gate were not considered in this modeling approach.

Haddad and Shraiber (2014) stated that the previously discussed MFD-based controllers could regulate the flow of vehicles around a predefined accumulation setpoint, which was the accumulation that corresponded to the highest network throughput in a network's MFD. However, this study linearized the non-linear MFD-based system dynamics around a stochastic setpoint in such a way that uncertainties in the MFD-based dynamics could be considered in the controller design. The numerical analysis of this study showed that the performance of the designed controller was superior for different congestion levels, unlike previous controllers that were effective only once the congestion level had reached the setpoint. This study concluded that translating the optimized metered flows into signal indications at metering signals required further study.

Keyvan-Ekbatani et al. (2015) stated that different parts of a network might experience congestion at different times. Hence, they designed MFD-based controllers for concentric urban networks in which congestion spread from the innermost to outermost regions. Accordingly, different perimeter controllers would protect each region according to its own MFD and desired congestion level. The authors also proposed a policy based on queue lengths at the gates so that the total gated flows could be divided among gates by converting flows to signal indications. However, the proposed flow distribution approach might not be optimal because the controllers did not optimize the gate flows for each gate individually. Moreover, the gates could not enforce

the same gated flows as were optimized by the controllers because of the conversion of flows to signal indications.

Haddad and Mirkin (2017) stated that central controllers for multi-region networks were not robust in the face of data collection and communication failures. Therefore, they designed local adaptive perimeter controllers that could share upper-level information about the desired setpoints of regions with each other. This information could help with coordination amongthe regions and improve overall system performance. However, this coordination did not ensure the capacity restriction of neighboring regions. Therefore, congestion might spread from a downstream region to its upstream region, and the controllers would fail to maintain vehicle accumulation around the desired setpoints.

Kouvelas et al. (2017) said that the dynamics of multi-region networks were complex, and defining setpoints according to the MFD of each region might not be optimal for the entire network. In other words, some regions might need to held vehicle accumulation to a level more or less than their desired setpoint to favor other regions and to improve the overall operation of the network. Therefore, they proposed an adaptive approach based on network conditions to finding controller parameters such as setpoints, rather than using fixed values for the entire study period. Although their numerical analysis in a microsimulation environment showed the capability of this approach to improve network performance, the approach did not ensure the optimality of the parameters. Moreover, the optimized metering rates were distributed between the gates in proportion to their saturation flow rates.

Fu et al. (2017) stated that applying perimeter controls to networks with high demand that created heterogeneity in a network might not improve network performance. Therefore, they proposed a hierarchical perimeter control approach for a two-region network that allowed upper

and lower bounds to be found for the flow control variables in the first level to ensure system stability. Each of the two regions were further divided into sub-regions, and sub-regional-level perimeter controllers regulated the flow of vehicles with the objective of minimizing the heterogeneity among the sub-regions. The sub-regions were defined by a clustering algorithm and used a genetic algorithm to solve the non-linear control problem. However, the proposed bounds could not be ensured because the control problem optimized the total metered flows, but the flows had to be divided among gates and converted to traffic signals so that they could be implemented in a network. These steps created approximations that might violate the stability conditions.

A numerical analysis by Haddad (2017) of a two-region network with MFD-based dynamics showed that the capacity of gates to hold queued vehicles had a significant effect on perimeter control policies. He proposed a control model with explicit constraints on the aggregated gate capacities at the boundaries between regions. The limited capacity of the boundary gates could be captured explicitly in the perimeter control policies. However, the capacity restrictions were considered to be an aggregated measure in the model. Therefore, the model did not enforce the capacity restrictions of individual gates. In other words, gates could have different capacities, vehicle flows, and metering rates. Therefore, the aggregated capacity limits did not account for the locations and capacity of gates.

Ding et al. (2017) proposed an approach to simultaneously consider perimeter control and route guidance in multi-region networks using MFD-based dynamics. They captured the effect of route guidance by estimating the rate of drivers who would comply with suggested alternative routes that passed through less congested regions. The proposed algorithm had a hierarchical structure such that route guidance and perimeter control minimized the average delay of vehicles sequentially. The proposed approach required predefined homogenous sub-regions and gate

locations, while traffic guidance changed the distribution of traffic in a network and hence the shape of the MFD's, number of sub-regions, and gate locations.

Sirmatel and Geroliminis (2017) developed an MPC for perimeter control and route guidance using MFD-based network dynamics. They argued that keeping vehicle accumulation at predefined setpoints in multi-region networks might be infeasible for perimeter controllers. Moreover, the accumulation-based objective did not necessarily result in the best network performance. Hence, they minimized the total travel time in the MPC formulation. The proposed formulation in this study was a mixed-integer, non-linear program that was solved with an optimization engine. However, the quality of the solutions in terms of optimality bounds was not discussed in the study.

Yang et al. (2019) considered social welfare besides operational performance to be the objectives of perimeter control. They considered priority lanes that could be assigned dynamically to different traffic modes. The tolls for the priority lanes were optimized according to the lane choice information for connected vehicles. Analysis results showed that the proposed approach could distribute the total costs of delays and tolls more uniformly than a case without priority that showed improvements in social welfare.

2.3. Summary

The reviewed literature showed that scenario-based techniques do not provide dynamic metering rates, and their efficiency depends on the quality of the pre-defined scenarios. On the other hand, MFD-based techniques regulate the flow of vehicles dynamically and are suitable for large-scale transportation networks. However, they rely on well-defined MFDs for a network or its sub-networks. Although several algorithms have been developed to divide a network into several sub-networks, each with well-defined MFDs, these algorithms are rather heuristics, and

their accuracy depends on network geometry and spatial vehicle density distributions in the network. In addition, the MFD-based techniques find aggregated metering rates for all gates of a protected area. The aggregated rates require a post-processing step to be converted into the metering rates for each individual gate. The gate-level metering rates can be found by distributing the aggregated rates in proportional to the queue lengths of vehicles at the gates or to their saturation flow rates. Therefore, a gap exists in the traffic metering literature that can be filled with an approach that explicitly captures traffic dynamics instead of using MFDs and that optimizes traffic metering rates for each individual gate dynamically.

Chapter 3: Problem Formulation

The developed program for optimizing traffic metering rates is discussed in this chapter. The model is based on the cell transmission model (CTM) (Daganzo, 1995, 1994). The CTM discretizes both time and space and provides numerical solutions to the hydrodynamic traffic flow model presented by Lighthill and Whitham (1955) and Richards (1956). The discretization makes the CTM easy to implement, with the ability to consider different traffic flow regimes ranging from undersaturated to oversaturated and gridlock conditions in urban street networks (Mohebifard and Hajbabaie, 2018a). CTM divides each network link into homogenous segments that are called "cells." Figure 3-1 shows a simple network of four intersections with twelve one-way links that are represented by 52 cells according to the CTM.

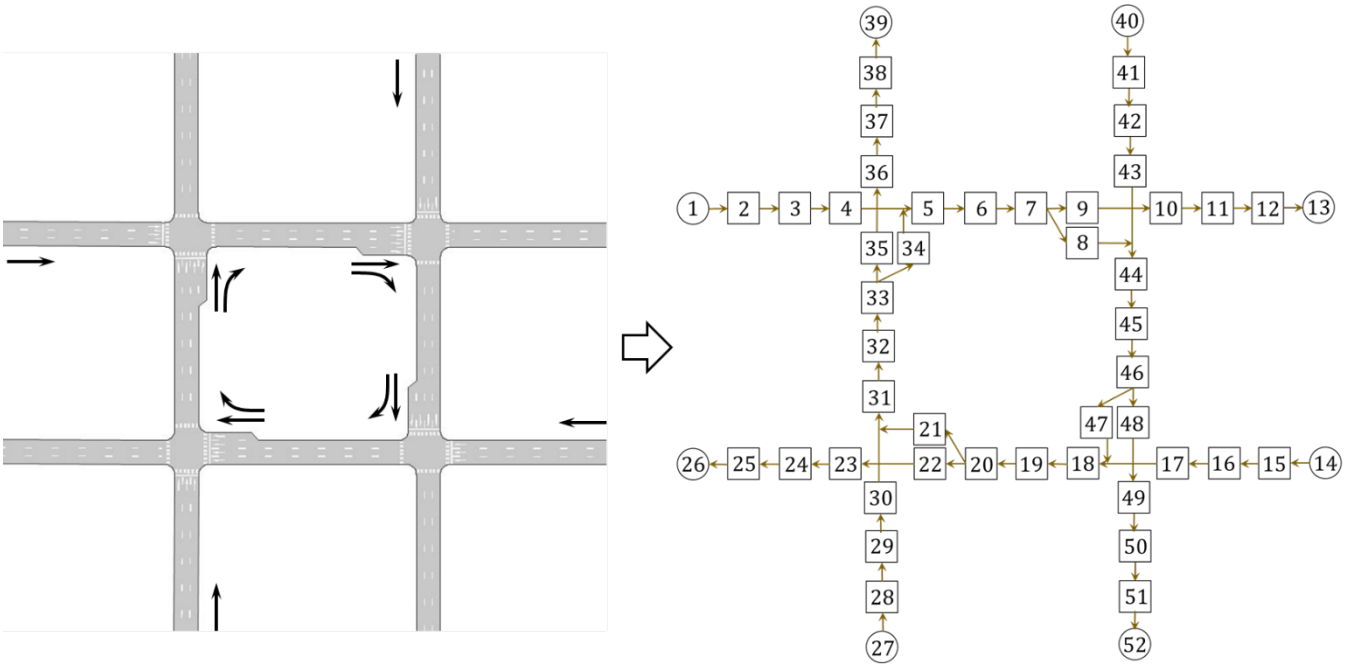


Figure 3-1. Cell representation of a sample network

We define five cell types, shown in figure 3-2. Gate cells represent the entry points to a network. The inflow to gates follows a predefined demand profile, while the outflow of the gates

will be defined on the basis of the optimized traffic metering rates. Sink cells absorb the flow of vehicles leaving the network. Merge cells receive vehicles from multiple immediate upstream cells, and diverge cells send vehicles to multiple immediate downstream cells. Ordinary cells are the cells that receive vehicles from and send vehicles to only one immediate cell. We also use intersection cells, which are like ordinary cells but with a variable saturation flow rate that represents the effect of signals. The notations for the set of all the cells are as follows: network C, gate C_G, sink C_S, merge C_M, diverge C_D, and intersection C_I. Therefore, the set of ordinary cells over all other cells is $C\backslash\{C_G, C_S, C_M, C_D, C_I\}$. In addition, the successor and predecessor cells of cell $i \in C$ are indicated by $S(i)$ and $P(i)$, respectively. The predecessor cells of cell $i \in C$ are those that are immediately upstream and successor cells are those that are immediately downstream.

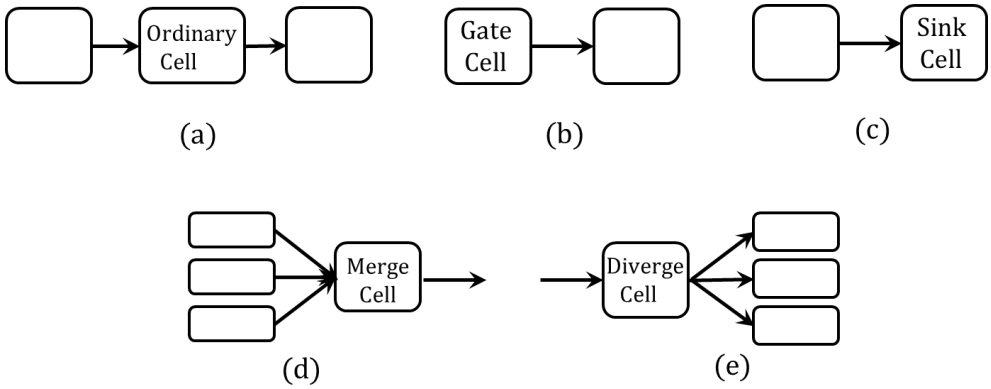


Figure 3-2. Types of cells according to the CTM

We explain a mixed-integer, non-linear program (MINLP) to optimize traffic metering rates at each metering signal. The decision variables of the program are the number of vehicles y_{ij}^t that should be allowed to leave gate cell $i \in C_G$ to its successor cell $j \in S(i)$ at time step $t \in T$. Constraints (3-1) through (3-6) ensure that the gate flows do not violate the flow feasibility conditions. Constraint (3-1) ensures that that gate flows y_{ij}^t do not exceed the available number of vehicles x_i^t in gate cell $i \in C_G$ at each time step $t \in T$. Moreover, the gate flows should not be

more than the available capacity $\delta\left(N_j - x_j^t\right)$ of receiving cell $j \in S(i)$, as shown in constraint (3-2). In these constraints, N_j is the capacity of a cell in terms of the number of vehicles it can hold, and δ is the ratio of backward to forward shockwave speeds (Daganzo, 1994). Constraints (3-3) and (3-4) guarantee that the gate flows are less than or equal to the saturated flow rates Q_i^t and Q_j^t of the sending and receiving cells, respectively. The number of vehicles x_i^t in gate cell $i \in C_G$ needs to be less than gate capacity N_i for holding a queue of vehicles (see constraint (3-5)). Constraint (3-6) shows the non-negativity requirements of the gate flows.

$$y_{ij}^t \leq x_i^t \qquad \forall i \in C_G, j \in S(i), t \in T \qquad (3\text{-}1)$$

$$y_{ij}^t \leq \delta\left(N_j - x_j^t\right) \qquad \forall i \in C_G, j \in S(i), t \in T \qquad (3\text{-}2)$$

$$y_{ij}^t \leq Q_i^t \qquad \forall i \in C_G\ , j \in S(i), t \in T \qquad (3\text{-}3)$$

$$y_{ij}^t \leq Q_j^t \qquad \forall i \in C_G\ , j \in S(i), t \in T \qquad (3\text{-}4)$$

$$x_i^t \leq N_i \qquad \forall i \in C_G, t \in T \qquad (3\text{-}5)$$

$$y_{ij}^t \geq 0 \qquad \forall i \in C_G, j \in S(i), t \in T \qquad (3\text{-}6)$$

The flow of vehicles between all other cells except the gate cells should follow the flow-density diagram of the CTM that is shown by flow feasibility (constraints (3-7) through (3-10). Constraint (3-7) ensures that the flow of vehicles in ordinary cells $i \in C \backslash \{C_G, C_S, C_D, C_I\}$ should be equal to the minimum of x_i^t, Q_i^t, Q_j^t, and $\delta\left(N_j - x_j^t\right)$. Note that cell $j \in S(i)$ is successor to cell i. The $\min(.)$ function guarantees that the flows follow the flow-density diagram of CTM (Daganzo, 1995), and the flow holding-back problem is eliminated (Lo, 1999; Mohebifard and Hajbabaie, 2019). Constraint (3-8) uses the variable saturation flow rates $g_i^t Q_i^t$ instead of Q_i^t to account for the effect of signal indications g_i^t on the saturation flow rate of intersection cells $i \in C_I$. The signal

indication g_i^t is a binary parameter that is 1.0 for green signals and 0.0 otherwise. Note that the signal indications are input to this constraint and should be defined on the basis of the signal settings of a case study network before the traffic metering problem is solved.

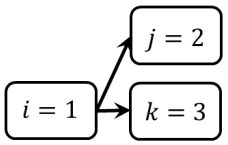
$$y_{ij}^t = \min\{x_i^t, Q_i^t, Q_j^t, \delta(N_j - x_j^t)\} \qquad \forall i \in C \backslash \{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \tag{3-7}$$

$$y_{ij}^t = \min\{x_i^t, g_i^t Q_i^t, Q_j^t, \delta(N_j - x_j^t)\} \qquad \forall i \in C_I, j \in S(i), t \in T \tag{3-8}$$

$$y_{ij}^t = \beta_j^t \min\left\{x_i^t, Q_i^t, \frac{Q_i^t}{\beta_j^t}, \frac{\delta(N_j - x_j^t)}{\beta_j^t}, \frac{Q_k^t}{\beta_k^t}, \frac{\delta(N_k - x_k^t)}{\beta_k^t}, \frac{Q_m^t}{\beta_m^t}, \frac{\delta(N_m - x_m^t)}{\beta_m^t}\right\} \qquad \forall i \in C_D, j \in S(i), k \in S(i), m \in S(i), j \neq k \neq m, t \in T \tag{3-9}$$

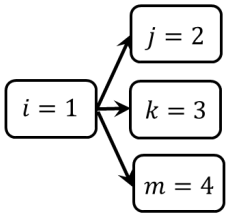
$$y_{ij}^t = \min\{x_i^t, Q_i^t\} \min\left\{1, \frac{\min\{Q_j^t, \delta(N_j - x_j^t)\}}{\sum_{k \in P(j)} \min\{x_k^t, Q_k^t\}}\right\} \qquad \forall j \in C_M, i \in P(j), t \in T \tag{3-10}$$

Constraint (3-9) shows the flow-feasibility conditions for divergent cells $i \in C_D$ based on dynamic turning ratios β_j^t, β_k^t, and β_m^t for distinct successor cells j, k, $m \in S(i)$. For divergent cells with two successor cells, the m index should be removed from these constraints. Figure 3-3 illustrates constraints for a divergent cell with two and three downstream cells. Dynamic turning ratios β_j^t, β_k^t, and β_m^t are time dependent and input to the constraint, and they show the fraction of flow that enters each of the downstream cells. Hence, the summation of the turning ratios should add up to 1.0, i.e., $\beta_j^t + \beta_k^t + \beta_m^t = 1$. Furthermore, constraint (3-9) ensures the first-in-first-out conditions for divergent cells. In other words, if cell $j \in S(i)$ is full and cannot accommodate any vehicles, then the flow of other downstream cells will be 0.0. This condition captures the capacity reductions at intersections due to queue spillovers or spillbacks.



$$y_{12}^t = \beta_2^t \min\left\{x_1^t, Q_1^t, \frac{Q_2^t}{\beta_2^t}, \frac{\delta(N_2 - x_2^t)}{\beta_2^t}, \frac{Q_3^t}{\beta_3^t}, \frac{\delta(N_3 - x_3^t)}{\beta_3^t}\right\}$$

$$y_{13}^t = \beta_3^t \min\left\{x_1^t, Q_1^t, \frac{Q_2^t}{\beta_2^t}, \frac{\rho(N_2 - x_2^t)}{\beta_2^t}, \frac{Q_3^t}{\beta_3^t}, \frac{\rho(N_3 - x_3^t)}{\beta_3^t}\right\}$$



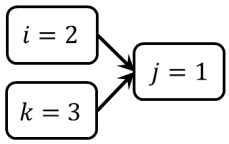
$$y_{12}^t = \beta_2^t \min\left\{x_1^t, Q_1^t, \frac{Q_2^t}{\beta_2^t}, \frac{\delta(N_2 - x_2^t)}{\beta_2^t}, \frac{Q_3^t}{\beta_3^t}, \frac{\delta(N_3 - x_3^t)}{\beta_3^t}, \frac{Q_4^t}{\beta_4^t}, \frac{\delta(N_4 - x_4^t)}{\beta_4^t}\right\}$$

$$y_{13}^t = \beta_3^t \min\left\{x_1^t, Q_1^t, \frac{Q_2^t}{\beta_2^t}, \frac{\delta(N_2 - x_2^t)}{\beta_2^t}, \frac{Q_3^t}{\beta_3^t}, \frac{\delta(N_3 - x_3^t)}{\beta_3^t}, \frac{Q_4^t}{\beta_4^t}, \frac{\delta(N_4 - x_4^t)}{\beta_4^t}\right\}$$

$$y_{14}^t = \beta_4^t \min\left\{x_1^t, Q_1^t, \frac{Q_2^t}{\beta_2^t}, \frac{\delta(N_2 - x_2^t)}{\beta_2^t}, \frac{Q_3^t}{\beta_3^t}, \frac{\delta(N_3 - x_3^t)}{\beta_3^t}, \frac{Q_4^t}{\beta_4^t}, \frac{\delta(N_4 - x_4^t)}{\beta_4^t}\right\}$$

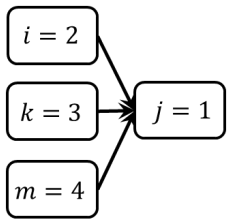
Figure 3-3. Illustration of the flow-feasibility constraints for two divergent cells with two and three downstream cells

The flow feasibility conditions of merge cells $j \in C_M$ are shown by constraint (3-10): the available capacity of the merge cell $\min\{Q_j^t, \delta(N_j - x_j^t)\}$ will be divided among its upstream cells $k \in P(j)$ in proportion to their flows $\min\{x_k^t, Q_k^t\}$. This constraint is further illustrated in figure 3-4.



$$y_{21}^t = \min\{x_2^t, Q_2^t\} \min\left\{1, \frac{\min\{Q_1^t, \delta(N_1 - x_1^t)\}}{\min\{Q_2^t, x_2^t\} + \min\{Q_3^t, x_3^t\}}\right\}$$

$$y_{31}^t = \min\{x_3^t, Q_3^t\} \min\left\{1, \frac{\min\{Q_1^t, \delta(N_1 - x_1^t)\}}{\min\{Q_2^t, x_2^t\} + \min\{Q_3^t, x_3^t\}}\right\}$$



$$y_{21}^t = \min\{x_2^t, Q_2^t\} \min\left\{1, \frac{\min\{Q_1^t, \delta(N_1 - x_1^t)\}}{\min\{Q_2^t, x_2^t\} + \min\{Q_3^t, x_3^t\} + \min\{Q_4^t, x_4^t\}}\right\}$$

$$y_{31}^t = \min\{x_3^t, Q_3^t\} \min\left\{1, \frac{\min\{Q_1^t, \delta(N_1 - x_1^t)\}}{\min\{Q_2^t, x_2^t\} + \min\{Q_3^t, x_3^t\} + \min\{Q_4^t, x_4^t\}}\right\}$$

$$y_{41}^t = \min\{x_4^t, Q_4^t\} \min\left\{1, \frac{\min\{Q_1^t, \delta(N_1 - x_1^t)\}}{\min\{Q_2^t, x_2^t\} + \min\{Q_3^t, x_3^t\} + \min\{Q_4^t, x_4^t\}}\right\}$$

Figure 3-4. Illustration of the flow-feasibility conditions for merge cells with two and three upstream cells

The flow conservation concept is shown by constraints (3-11) through (3-13) for different cell types. The number of vehicles x_i^{t+1} in gate cell $i \in C_G$ at time step $t+1 \in T$ is equal to the number of vehicles x_i^t that were available in the cell in the previous time step t plus the network demand D_i^t minus the number of vehicles $\sum_{j \in S(i)} y_{ij}^t$ that leave the gate cell to its downstream cells $j \in S(i)$. In this constraint, the number of vehicles entering the gate cells is determined by a predefined and time-dependent demand profile D_i^t, and the number of vehicles that leave the gate cells is indeed the optimized gate flow. Constraint (3-12) shows the flow conservation for sink cell $j \in C_S$. The sink cells are assumed to have infinite capacity and absorb all the vehicles that leave the network. With this representation, the summation of vehicles in all sink cells at each time step shows the accumulated number of completed trips up to that time step. Constraint (3-13) shows the flow conservation concept for ordinary cells.

$$x_i^{t+1} = x_i^t + D_i^t - \sum_{j \in S(i)} y_{ij}^t \qquad \forall i \in C_G, t \in T \qquad (3\text{-}11)$$

$$x_j^{t+1} = x_j^t + \sum_{i \in P(j)} y_{ij}^t \qquad \forall j \in C_S, t \in T \qquad (3\text{-}12)$$

$$x_i^{t+1} = x_i^t + \sum_{k \in P(i)} y_{ki}^t - \sum_{j \in S(i)} y_{ij}^t \qquad \forall i \in C \backslash \{C_S, C_G\}, t \in T \qquad (3\text{-}13)$$

The objective function of the problem is shown in (3-14); it maximizes the cumulative number of vehicles in all sink cells. This objective function maximizes the network throughput and is equivalent to the minimization of the total travel time. Previous research showed that network throughput maximization is a suitable objective function for oversaturated conditions (Hajbabaie, 2012; Hajbabaie et al., 2011, 2010; A Hajbabaie and Benekohal, 2011).

$$\text{maximize } Z = \sum_{\forall i \in C_S} \sum_{\forall t \in T} x_i^t \quad (3\text{-}14)$$

Given the discussed constraints and objective function, the program for optimizing the gate flows can be summarized as program (P1).

$$\text{(P1)} \quad \text{maximize } Z = \sum_{\forall i \in C_S} \sum_{\forall t \in T} x_i^t \text{ subject to (3-1)-(3-13)} \quad (3\text{-}15)$$

Chapter 4: Optimal Solutions of the Traffic Metering Problem

This section presents an approach that finds the optimal solutions of program (P1) within an optimality gap. Note that (P1) is a complex optimization program because of the excessive number of mixed-integer decision variables and nonlinear constraints. The decision variables of the program are continuous, but the representation of the min(.) operators in the flow-feasibility constraints (3-7)-(3-10) requires adding auxiliary dummy variables to the program. For instance, the standard representation of constraint (3-7) is equivalent to constraints (4-1) through (4-10) using the big-M technique. Constraints (4-1) through (4-4) ensure the flow feasibility conditions, while constraints (4-4) through (4-10) ensure that the flow of vehicles y_{ij}^t from each cell $i \in C\backslash\{C_G, C_S, C_D, C_I\}$ to cell $j \in S(i)$ at each time step $t \in T$ is exactly equal to either x_i^t, Q_i^t, Q_j^t or $\delta\left(N_j - x_j^t\right)$. In other words, the addition of constraints (4-4) through (4-10) with auxiliary binary variables θ_{ij}^t, ϑ_{ij}^t, ψ_{ij}^t, and χ_{ij}^t sets the flow of vehicles equal to one of the right-hand side values of constraints (4-1) through (4-4). In the latter constraints, M is an arbitrarily large number.

$$y_{ij}^t \le x_i^t \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}1)$$

$$y_{ij}^t \le Q_i^t \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}2)$$

$$y_{ij}^t \le Q_j^t \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}3)$$

$$y_{ij}^t \le \delta\left(N_j - x_j^t\right) \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}4)$$

$$y_{ij}^t \ge x_i^t - M\left(1 - \theta_{ij}^t\right) \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}5)$$

$$y_{ij}^t \ge Q_i^t - M\left(1 - \vartheta_{ij}^t\right) \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}6)$$

$$y_{ij}^t \ge Q_j^t - M\left(1 - \psi_{ij}^t\right) \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}7)$$

$$y_{ij}^t \ge \delta\left(N_j - x_j^t\right) - M\left(1 - \chi_{ij}^t\right) \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}8)$$

$$\theta_{ij}^t + \vartheta_{ij}^t + \psi_{ij}^t + \chi_{ij}^t = 1 \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}9)$$

$$\theta_{ij}^t, \vartheta_{ij}^t, \psi_{ij}^t, \chi_{ij}^t \in \{0,1\} \qquad \forall i \in C\backslash\{C_G, C_S, C_D, C_I\}, j \in S(i), t \in T \qquad (4\text{-}10)$$

Once the flow-feasibility constraints have been converted to standard constraints with the big-M technique, many auxiliary binary variables will be added to the program and will significantly increase the problem complexity. Hence, the conventional optimization techniques in commercial solvers such as CPLEX (CPLEX, 2009) cannot solve the program for medium-sized networks in a reasonable amount of time. We tackled this issue by developing a novel decomposition solution technique to solve (P1) more efficiently. Decomposition techniques have been used widely to solve complex transportation problems (e.g., Hajibabai et al., 2014; Hajibabai and Ouyang, 2013; Hajibabai and Saha, 2019).

The general idea of the solution technique is that once the gate flows are known, all other variables of the program can be determined with a CTM simulation without solving any optimization programs. Figure 4-1 shows the idea of the solution technique for a link with one gate. According to the figure, the gate can be separated from the rest of the link by relaxing the constraints that connect gate cell i to its successor cell $j \in S(i)$. Then, an optimization program for the gate can be solved to find gate flows y_{ij}^t and use them as input for the rest of the link to find all other variables. Then, by connecting the two steps and iterating between them, the optimal solution can be found.

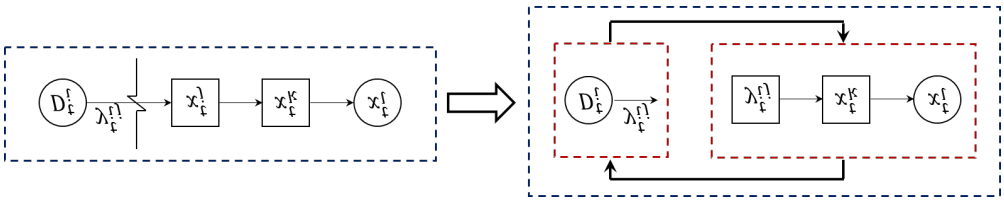


Figure 4-1. General decomposition idea for one link

We utilized the generalized Benders decomposition technique (Benders, 1962; Geoffrion, 1972) to implement the discussed ideas to ensure the convergence and optimality of the solution technique. The Benders decomposition technique was developed to address a class of complex optimization programs that can be solved efficiently once a set of decision variables has been temporarily set to predefined values. In other words, by fixing the values of the complicating decision variables, the remaining optimization program can be solved more easily than the original problem.

The Benders technique decomposes the original problem into a primal and a master problem. The primal problem is equivalent to the original problem in which the values of the complicating variables have been temporarily fixed. The solutions of the primal problem are the optimal values of the non-complicating variables, the optimized value of the objective function, and the dual values of the constraints that include both complicating and non-complicating variables. These values are used to construct the master problem. The master problem is the relaxed dual of the original problem, and its decision variables are the complicating variables that were fixed in the primal problem. Moreover, the master problem has the constraints of the original problem, which include only the complicating variables because those constraints were redundant in the primal problem because the values of all their variables were fixed. The master problem also includes several constraints that are added to the problem iteratively on the basis of the solutions of the primal problem. These constraints are called Benders cuts and reduce the feasible region of the master problem iteratively until the optimal solutions are found. The solution of the primal problem provides the lower bounds for the optimal solution of the original problem, ad the solution of the master problem provides the upper bounds. Hence, iterating between the problems continues until an acceptable gap has been achieved.

Chapter 5: Real-Time Solutions for the Traffic Metering Problem

This chapter describes the use of distributed optimization techniques and location information for connected vehicles to develop an approach to optimize traffic metering rates in real time. The approach has a model predictive control (MPC) structure that uses a simulated transportation network (or a real network) to collect the required information on the system state at discretized time steps (e.g., every six seconds), makes predictions about the system state over a prediction horizon (e.g., 120 seconds) that is several time steps ahead of its current time, optimizes traffic metering rates with a distributed architecture over the horizon, and implements the first optimized traffic metering rates in the next time step. This procedure continues until the study period is over (Mohebifard and Hajbabaie, 2018b). Different components of the proposed solution technique are shown in figure 5-1.

Figure 5-1 shows that the proposed solution technique collects location information from connected vehicles and vehicle counts from loop detectors to estimate the distribution of vehicles across the network links (Section 5.1). The state estimations are used in the distributed optimization component that optimizes traffic metering rates in real-time (Section 5-2). This task is performed by decomposing the central optimization program (P1) into several sub-network-level optimization programs. The sub-network level programs cooperatively optimizes traffic metering rates by sharing the required information among each other. Once the optimized traffic metering rates have been found, the metering rates are converted into green and red gate signal indications (Section 5.3). The signals are then transferred to the network for implementation. The details of each component follow.

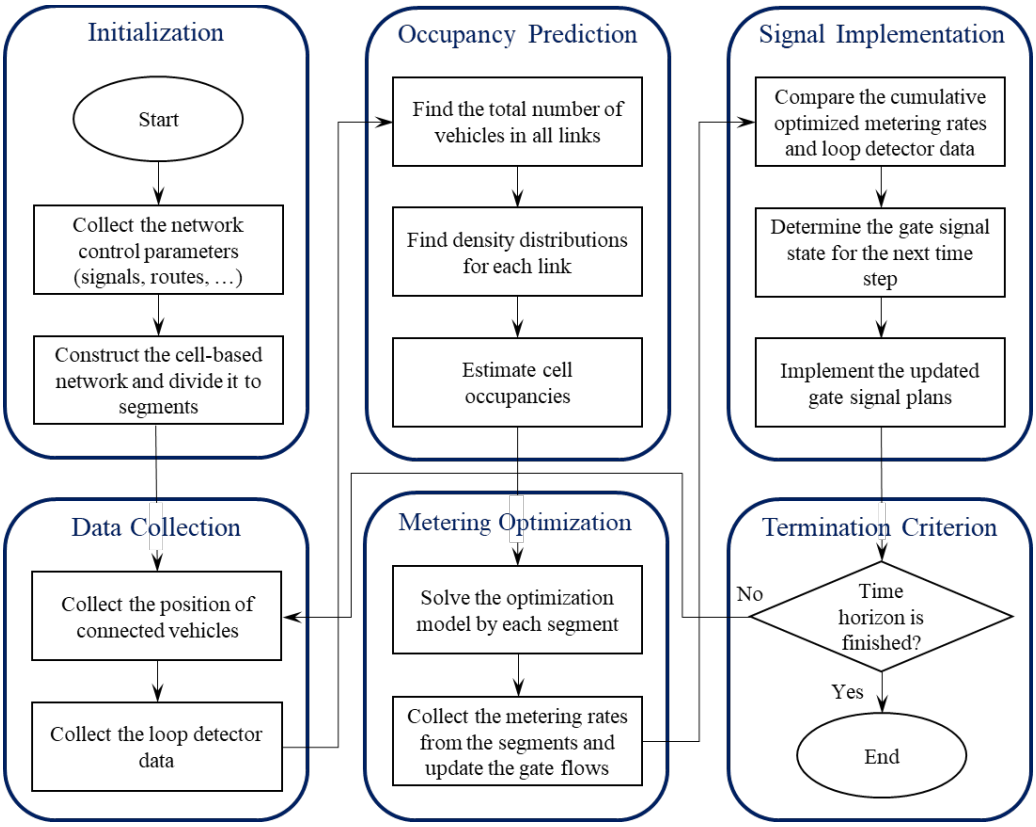


Figure 5-1. The proposed solution technique for real-time and distributed traffic metering rate optimization

5.1. System State Estimation

The initial system state in terms of cell occupancies x_i^0 for all $i \in C$ at time zero should be input to the optimization program. Hence, the location of vehicles in the network should be converted into cell occupancies. We used location information from connected vehicles (CV) and vehicle counts from loop detectors to estimate the cell occupancies.

Consider the link in figure 5-2. The link includes several CVs that are equipped with onboard units that can transmit different information, such as the location of CVs in relation to roadside units in a network. If the penetration rate (PR) of connected vehicles is 100 percent, then the state estimation is straightforward. The locations of vehicles will be mapped to their corresponding cells, and the cell occupancies will be the summation of vehicles in the cells.

However, the locations of unequipped vehicles must be estimated once the penetration rate is less than 100 percent.

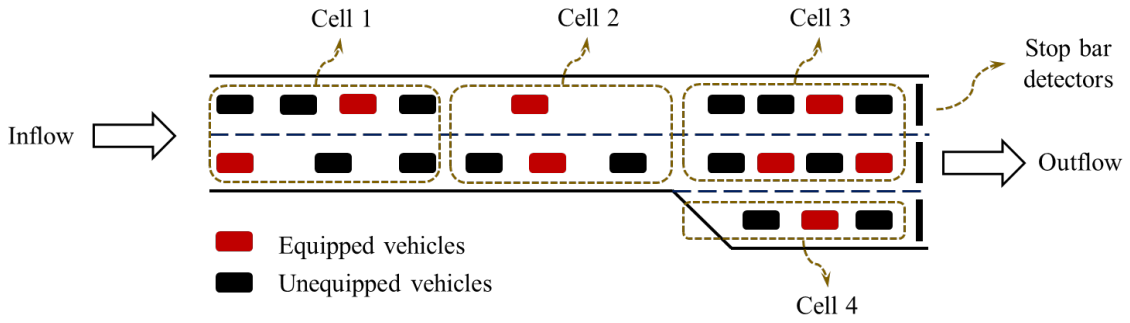


Figure 5-2. A link with the corresponding cells, including both equipped (CV) and unequipped (non-CV) vehicles

The proposed approach for system state estimation is the integration of two different estimation techniques (Mohebifard and Hajbabaie, 2018c). The first technique approximates the density distribution of vehicles in each network link with a vehicle sample that includes CVs. The second technique uses flow feasibility and conservation of CTM to estimate cell occupancies. These two estimations are averaged on the basis of the market penetration rate of CVs.

5.1.1. State Estimation Using Location Information from CVs

The distribution of vehicles in each network link can be estimated by using Equation (5-1). In the equation, xe_i^t is the number of CVs in cell $i \in C$ at time step $t \in T$. By dividing xe_i^t over the total number of CVs in the link $l \in L$ that contains cell i, the distribution of CVs in the link can be found. In this equation, L represents the set of all network links. By multiplying the estimated distribution of CVs by the total number of vehicles in the link V_l^t, the distribution of both CVs and non-CVs can be found. The total number of vehicles in each link can be estimated by tracking the inflow and outflow of each link such that V_l^t will be equal to the cumulative outflow minus the

cumulative inflow of each link up to time step $t \in T$, assuming that all links are equipped with stop-bar detectors.

$$x_i^{t,CV} = V_l^t \frac{xe_i^t}{\sum_{i \in l} xe_i^t} \qquad \forall i \in C, l \in L, t \in T \qquad (5\text{-}1)$$

Note that as the penetration rate of CVs increases, the distribution estimation error decreases because the sample size of CVs increases and better represents the distribution of vehicles in a link. However, the estimation error increases with low penetration rates. To address this issue, we used the CTM flow conservation and feasibility equations to adjust estimations for low penetration rates.

5.1.2. State Estimation Using CTM Flow Feasibility and Conservation Principles

With this technique, we used the flow feasibility and conservation equations (3-7) through (3-13) to track cell occupancies in each link. For more clarification, consider figure 5-2 again. The cell occupancies will be initialized to zero at the beginning of the study when the network is empty. As vehicles enter the network, the loop detectors can track vehicles entering and exiting each link. This information can be utilized to simulate each link with the demand rate equal to total vehicle counts of the link's upstream stop-bar detectors. If we assume the flow of vehicles follows the flow conservation and feasibility equations (3-7) through (3-13), we can estimate the cell occupancies $x_i^{t,CTM}$ for each cell $i \in C$ and time step $t \in T$. Note that this approach does not require any information from CVs and relies only on vehicle counts of loop detectors, assuming that all links are equipped with stop-bar detectors.

5.1.3. Combined CV and CTM State Estimations

Once the estimations of $x_i^{t,CV}$ and $x_i^{t,CTM}$ have been found with the proposed approaches in sections 5.1.2 and 5.1.3, then the estimations are combined by using Equation (5-2). In this

equation, the state estimations are adjusted by the penetration rate $0 \leq PR \leq 1$ of CVs. This adjustment gives higher weights to estimations based on CVs because those estimations are more accurate, but with low penetration rates, the estimations of the CTM-based approach will have more weight. Moreover, the estimated cell occupancies cannot increase the capacity of the cell N_i.

$$x_i^t = \min\left\{PR x_i^{t,CV} + (1 - PR) x_i^{t,CTM}, N_i\right\} \qquad \forall i \in C, l \in L, t \in T \tag{5-2}$$

5.2. Distributed Traffic Metering Optimization

Once the initial cell occupancies have been estimated, traffic metering rates for each gate can be optimized. This optimization can be achieved by solving (P1) that optimizes all metering rates centrally. However, the solutions might not be found in real time. Moreover, as the spatial scale of a network increases, finding the optimal solutions requires more computation time. Therefore, we developed the Distributed Optimization and Coordination Algorithms for Dynamic Traffic Metering (DOCA-DTM) method, which can optimize traffic metering rates in real time. The DOCA-DTM has two main components, distributed optimization and distributed coordination. In distributed optimization, a network is divided into smaller sub-networks. The central optimization program is also decomposed into several sub-network programs. Each sub-network program optimizes traffic metering rates for its corresponding sub-network. The distributed coordination component shares information about the predicted number of vehicles that enter the sub-networks and the available capacity for receiving vehicles among all sub-networks. This information is incorporated as several constraints into the sub-network programs so that they can coordinate their decisions and push their local solutions toward the system-wide optimal solutions.

5.2.1. Distributed Optimization

The first step in the distributed optimization technique is to define appropriate sub-networks that can serve as computational nodes, and then decompose the optimization model according to the defined sub-networks.

5.2.1.1. Model Decomposition

Defining appropriate computational nodes is a critical factor in the efficiency of the solution technique. Although intersection-level decomposition has been used widely in traffic operations (Mehrabipour, 2018), traffic metering decision variables are distributed at the boundary of a network. Therefore, we proposed to use arterial-level sub-networks that start from the gates and end at the network sinks. To clarify the point, figure 5-3 shows a sample network that is decomposed into six arterial-level sub-networks.

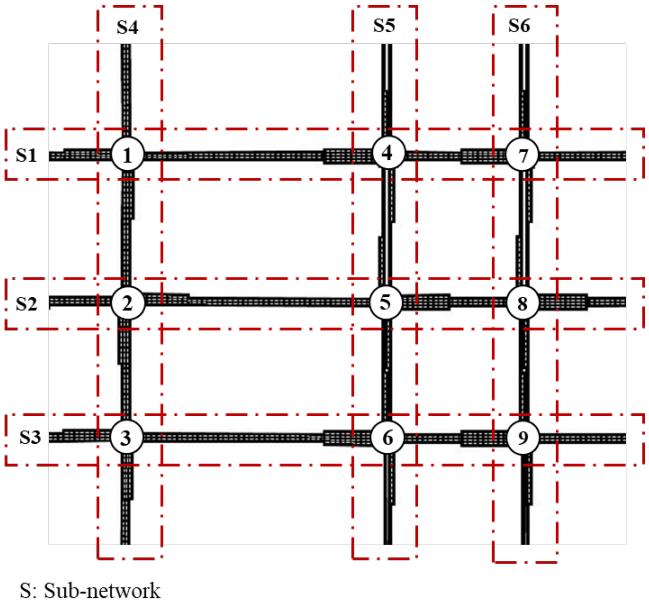


Figure 5-3. Spatial decomposition of a network into six sub-networks

Once the sub-networks have been defined, the central optimization program should be decomposed accordingly. The optimization model can be decomposed by relaxing flow feasibility constraints (3-7) through (3-10) on the links that connect cells $i \in C^s$ and $j \in C^p$ in such a way that

cells i and j belong to two neighboring sub-networks $s, p \in \mathcal{S}$. Thus, each sub-network will have a stand-alone optimization program that can be solved independently to optimize traffic metering rates for the gates within each sub-network. Note that solutions for the sub-network programs will be local optimal solutions, as the interactions of the sub-networks are ignored. We discuss later in this chapter how these relaxed constraints can be reenforced to coordinate the sub-networks and push the local solutions toward the global optimal solutions.

5.2.1.2. Linearization of Sub-Network Programs

Although spatial decomposition of the central program (P1) produces several sub-network programs that each have lower complexity than the original problem, they are still MINLPs that are hard to solve. However, the sub-networks have a simple cycle-free structure. Therefore, we can linearize the sub-problems by replacing non-linear flow-feasibility constraints (3-7) through (3-10) with linear constraints and modifying the objective function to eliminate the flow hold-back problem with the approach proposed by Zhu and Ukkusuri (2013). In other words, the min(.) operators in constraints (3-7) through (3-10) can be replaced by several linear "inequality" constraints and still prevent the flow-hold back problem, provided that the objective function of the program is modified. In this approach, the objective function is modified by assigning appropriate weights to the cells. The modified objective function of a sub-problem $s \in \mathcal{S}$ is shown in (5-3).

$$\text{maximize } Z^s = \sum_{\forall i \in C^s} \sum_{\forall t \in T} w_i x_i^t \tag{5-3}$$

The weights w_i should be defined in such a way that the weight of an upstream cell is lower than the downstream cell. For instance, figure 5-4 shows a sub-network with 14 cells whose weights are also shown. The weights are assigned in such a way that the sink of the sub-network

has a very high weight in comparison to those of the other cells. The objective function still maximizes the throughput of the sub-network (and hence the network). Moreover, the weight of the cells that send flow to the sub-network is one unit higher than the weight of all other cells except the sink cell that is also the network sink. In this way, a vehicle will be sent to the network if it can move to the sink cell without holding back; otherwise, it will stay in the gate cells that represent the traffic metering concept.

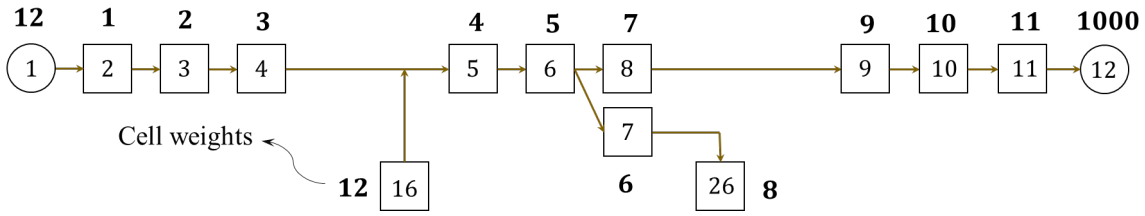


Figure 5-4. A sample sub-network with the weights assigned to each cell

Note that cell weights can be found for a sub-network without cycles according to the concept of topological ordering (Ahuja et al., 1993; Mohebifard et al., 2019). Figure 5-5 shows the proposed algorithm for finding cell weights according to the discussed characteristics of the weights. In this algorithm, M is an arbitrarily large number.

```
Algorithm segment cell weights;
Begin
    unmark all cells in the segment
    mark the gate cell;
    next: = 1;
    LIST: = gate cell;
    while LIST ≠ ∅ do
    begin
        select a cell i in LIST;
        if cell i is incident to an arc (i, j) and is admissible
        then
            begin
                mark cell j ;
                next: = next +1;
                w_j: = next;
                add cell j to LIST;
            end
            else delete cell i from LIST;
        end
    end
    for all gate and source cells
    begin
        select a cell i;
        w_i: = next+1;
    end
    for all sink cells
    begin
        select a cell i;
        w_i: = M;
    end
end
```

Figure 5-5.. Pseudocode for the topological ordering of cells in each sub-network

With the discussed modification, linearized sub-problem $s \in \mathcal{S}$ can be written as the optimization problem (P4).

$$\text{(P4)} \qquad \text{maximize } Z^s = \sum_{\forall i \in C^s} \sum_{\forall t \in T} w_i x_i^t \tag{5-3}$$

Subject to:

$$\sum_{j \in S(i)} y_{ij}^t \leq x_i^t \qquad \forall i \in C^s \backslash C_S^s, t \in T \tag{5-4}$$

$$\sum_{j \in S(i)} y_{ij}^t \leq Q_i \qquad \forall i \in C^s \backslash C_S^s, t \in T \tag{5-5}$$

$$\sum_{i \in P(j)} y_{ij}^t \leq Q_j \qquad \forall j \in C^s \backslash C_G^s, t \in T \qquad (5\text{-}6)$$

$$\sum_{i \in P(j)} y_{ij}^t \leq \delta\left(N_j - x_j^t\right) \qquad \forall j \in C^s \backslash C_G^s, t \in T \qquad (5\text{-}7)$$

$$x_i^t \leq N_i \qquad \forall i \in C_G^s, t \in T \qquad (5\text{-}8)$$

$$y_{ij}^t = \beta_j \sum_{k \in S(i)} y_{ik}^t \qquad \forall i \in C_D^s, j \in S(i), t \in T \qquad (5\text{-}9)$$

$$\sum_{j \in S(i)} y_{ij}^t \leq g_i^t Q_i^t \qquad \forall i \in C_I^s, t \in T \qquad (5\text{-}10)$$

$$x_i^{t+1} = x_i^t + D_i^t - \sum_{j \in S(i)} y_{ij}^t \qquad \forall i \in C_G^s, t \in T \qquad (5\text{-}11)$$

$$x_j^{t+1} = x_j^t + \sum_{i \in P(j)} y_{ik}^t \qquad \forall j \in C_S^s, t \in T \qquad (5\text{-}12)$$

$$x_i^{t+1} = x_i^t + \sum_{k \in P(i)} y_{ki}^t - \sum_{j \in S(i)} y_{ij}^t \qquad \forall i \in C \backslash \{C_S^s, C_G^s\}, t \in T \qquad (5\text{-}13)$$

$$y_{ij}^t \geq 0 \qquad \forall i \in C^s \backslash C_S^s, j \in S(i), t \in T \qquad (5\text{-}14)$$

Note that sub-problem (P4) is a linear program with much less computational complexity than the original problem. This sub-problem will be solved by each sub-network (computational nodes) to optimize traffic metering rates.

5.2.2. Distributed Coordination

We decompose the central formulation (P1) into several sub-network programs (P4) by relaxing the flow-feasibility constraints on the links between boundary cells of neighboring sub-networks. However, those constraints are critical for coordinating the sub-networks and avoiding sub-optimal operations. Therefore, we developed a distributed coordination algorithm that modifies the sub-network program (P4) by adding dummy source and sink cells and two additional constraints. The objective of this modification is to allow sub-networks to share information about the predicted number of vehicles that will enter a sub-network and the available capacity of a sub-network for receiving vehicles. To clarify the point, consider Sub-network 1 in figure 5-6. Sub-

network 1 has cells number 1 to 12, while this sub-network receives vehicles from cell 16 of sub-network 2 and sends vehicles to cell 26 of sub-network 3. In decomposing the central problem (P1), the flow feasibility constraints on links that connect cells 16 to 5 and cells 7 to 26 are relaxed. Therefore, dummy cell $1\acute{6}$ is added as a dummy source, and cell $2\acute{6}$ is added as a dummy sink cell to sub-network 1 to capture their effects on the traffic metering rate optimization.

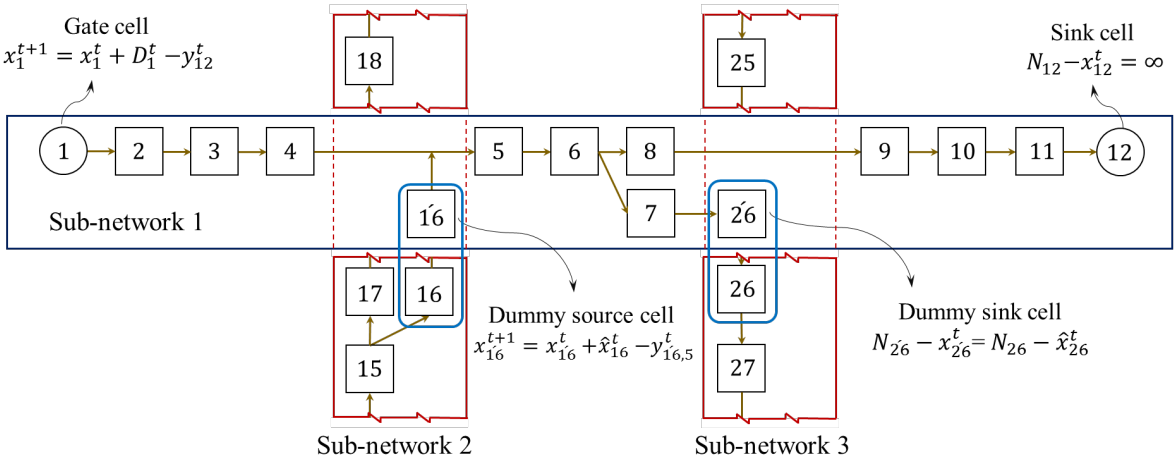


Figure 5-6. Illustration of dummy source and sink cells for sub-network 1

Accordingly, the number of vehicles that are predicted to enter sub-network 1 from sub-network 2 can be represented by adding constraint (5-15) to (P4), and the available capacity of sub-network 3 to receive vehicles from sub-network 1 can be represented by adding constraint (5-16) to (P4).

$$x_i^{t+1} = x_i^t + \hat{x}_i^{t+1} - \sum_{j \in S(i)} y_{ij}^t \qquad \forall i \in CD_G^s, t \in T \qquad (5\text{-}15)$$

$$\sum_{i \in P(j)} y_{ij}^t \le \delta\left(N_j - \hat{x}_j^t\right) \qquad \forall j \in CD_S^s, t \in T \qquad (5\text{-}16)$$

In Constraint (5-15), CD_G^s is the set of dummy gate cells that are added to a sub-network $s \in \mathcal{S}$, and $\hat{x}_i^{t+1}$ is the predicted number of vehicles in dummy gate cell $i \in CD_G^s$ at time step $t \in$

T. Similarly, CD_S^s is the set of added dummy sink cells to the sub-network $s \in \mathcal{S}$, and $\hat{x}_j^t$ is the predicted number of vehicles in the cell. The addition of these two constraints to (P1) allows each sub-network to coordinate its decisions on the basis of the available information from its neighboring sub-networks.

Parameters $\hat{x}_i^{t+1}$ and $\hat{x}_j^t$ in constraints (5-15) and (5-16) need to be predicted before the optimization problem (P4) can be solved. We use a CTM simulation for this prediction. In the absence of any optimized traffic metering rates, the network will be simulated with the network demand. The result of this simulation is the number of vehicles in each cell at each time step. Then these cell occupancies are used for $\hat{x}_i^{t+1}$ and $\hat{x}_j^t$ in constraints (5-15) and (5-16). Once these values have been found, the sub-network-level optimization program is solved, and traffic metering rates are optimized. If optimized traffic metering rates are available, the CTM simulation uses the optimized number of entry vehicles to the network instead of network demand for flow prediction.

5.3. Gate Signal Control Structure

The optimization model determines the number of vehicles that should enter the network from each gate at each time step. We use signalized gates, similar to those implemented for on-ramps, to control the number of vehicles entering the network. By utilizing loop detectors placed at gate signal stop-bars, the number of vehicles that leave the gates can be determined. Accordingly, we track the number of vehicles that leave a gate and compare it to the number of vehicles that should enter the network according to the optimized rates. On the basis of this comparison, we determine whether the signal should be green or red. Figure 5-7 shows the control structure for the gate signals.

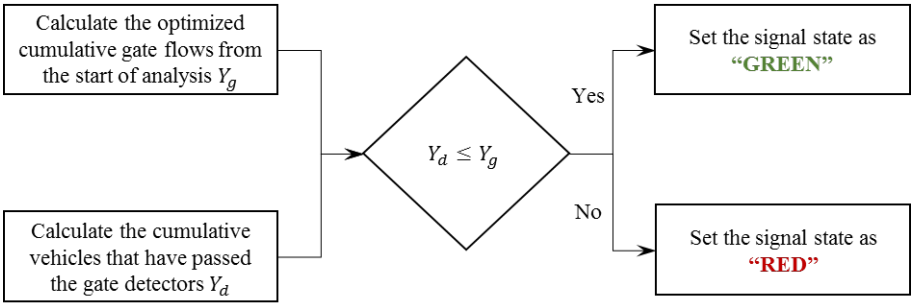


Figure 5-7. The control structure for each gate signal

On the basis of this simple control structure, a gate signal will remain green while the number of vehicles leaving a gate is less than the optimal number of vehicles that should enter the network. As soon as the number of vehicles that have passed a gate becomes greater than the optimal rate, the signal turns red. Notice that the optimization model updates the gate flows at each time step. Therefore, the state of the gate signals can change every time step on the basis of the updated solutions.

Chapter 6: Case Study and Numerical Results

6.1. Case Study Network

We used a portion of the urban street network in downtown Springfield, Illinois to evaluate the proposed traffic metering methodologies. The case study network is shown in figure 6-1. This network consisted of 20 intersections with a combination of one-way and two-way streets. The number of lanes in the streets varied between 1 to 3, as is shown in figure 6.1. We placed 13 gates at the boundary of this network to regulate the flow of vehicles entering the network according to the optimized metering rates. The network was analyzed with two demand profiles that are shown in figure 6.2. The total analysis period was 750 time steps (75 minutes), with 150 time steps for network loading. In addition, we optimized the signal timings of the intersections inside the network with the approach presented by Hajbabaie and Benekohal (2013, 2015) for each demand profile so that the effects of traffic metering on the network performance could be better evaluated. Other information on the case study network, such as cell characteristics, is summarized in table 6-1.

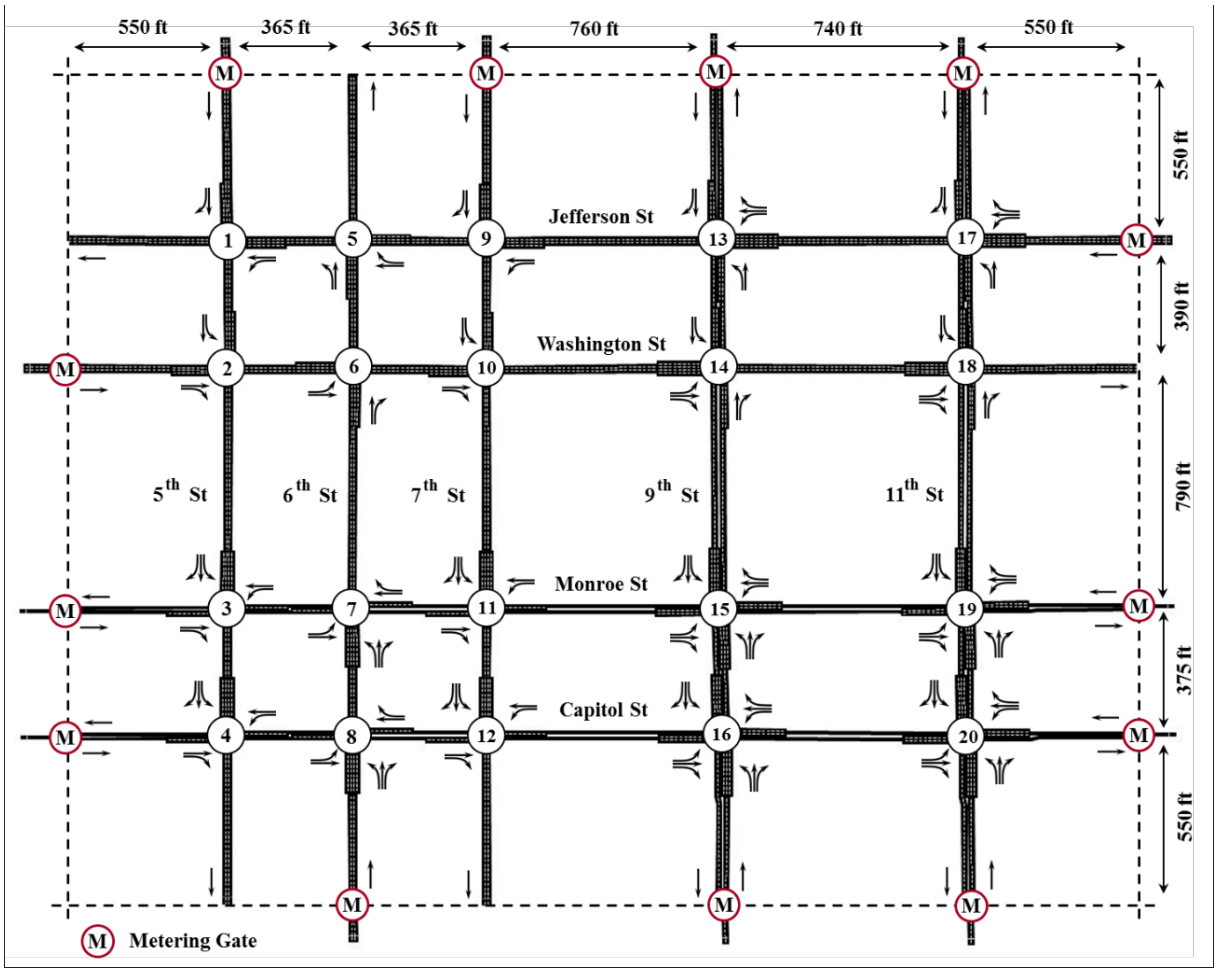


Figure 6-1. Downtown Springfield, Illinois, used as the case study network

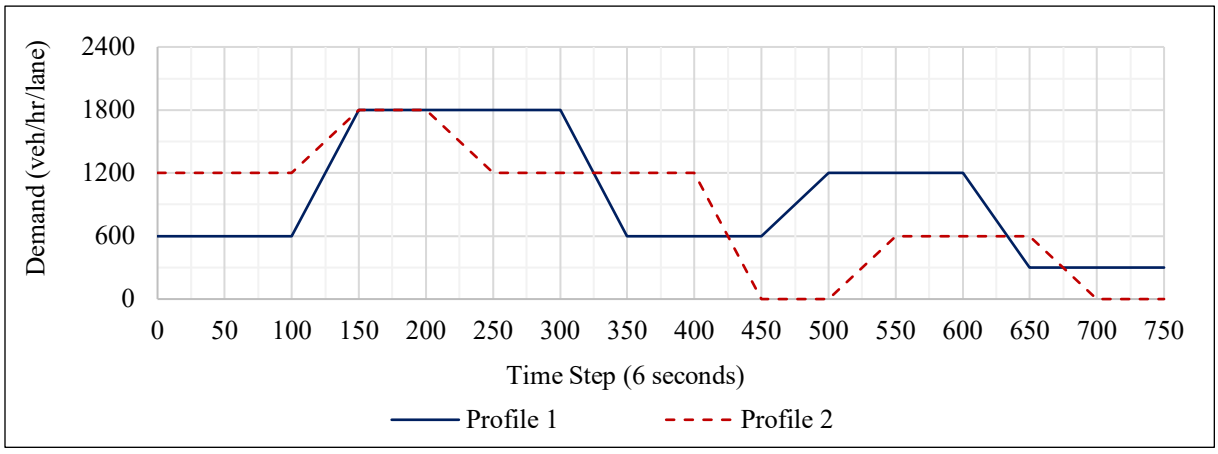


Figure 6-2. The demand profiles that were used for the analysis of the case study network

Table 6-1. Characteristics of the case study network

Element	Value
Free-flow speed (mph)	25
Saturation flow rate of links (veh/hr/lane)	1800
Time step duration (s)	6
Prediction horizon (min)	15
Total number of cells	316
Length of cells (ft)	225
Saturation flow rate of cells (veh/lane/time step)	3
Jam density of cells (veh/lane/cell)	12

6.2. Analysis Scenarios

For the presented analyses, the turning percentages at the intersections and penetration rate of connected vehicles were, respectively, set to 20 percent and 100 percent unless otherwise stated. We compared different performance measures under the following scenarios:

1. Simulation (SIM): This scenario showed network performance under the network's current conditions. In other words, we simulated the network without any traffic metering and calculated different performance measures. The results of this scenario were the do-nothing or no-metering performance measures in the case study.
2. Optimal Metering (OPT): We used the traffic metering approach proposed in Chapter 4 to find the optimal traffic metering rates for this scenario. This scenario showed network performance once the optimal traffic metering rates had been found centrally with the Benders decomposition-based optimization approach over the entire analysis period.
3. Distributed Optimization and Coordinated Algorithms for Dynamic Traffic Metering (DOCA-DTM): This scenario was designed to evaluate the efficiency of the distributed and coordination algorithms for traffic metering discussed in Chapter 5 once CTM had been

used in both the plant and optimization models. Hence, the results of this scenario could be directly compared to those of the OPT scenario to evaluate the optimality gaps between the real-time distributed solutions and the central optimal solutions.

6.3. Optimal Traffic Metering Results

We evaluated the performance of the case study network in the simulation (SIM) and optimal metering (OPT) scenarios for two demand profiles. The results in table 6-2 show that traffic metering in the OPT scenario increased the network throughput by 5.5 percent in Demand Profile 1 and by 3.4 percent in Demand Profile 2. In addition, the travel time of vehicles inside the network (excluding delayed vehicles at the gates) was reduced by 34.2 percent in Demand Profile 1 and by 30.8 percent in Demand Profile 2. Similarly, use of traffic metering decreased delay of vehicles inside the network by 42.3 percent and 37.2 percent, respectively, in comparison to the simulation scenarios in the two demand profiles. These results indicated that traffic metering could significantly improve operations of the case study network inside the region protected by gates.

The discussed improvements were achieved by delaying some vehicles at the gates. The delay of vehicles at the gates increased in the OPT by 2.3 percent for Demand Profile 1 and by 5.2 percent for Demand Profile 2 in comparison to the SIM scenario. However, if all vehicles are considered, including vehicles at the gates and inside the network, the overall system-level delay was reduced by 5.7 percent and 2.9 percent, respectively, as a result of traffic metering (see table 6-2).

Note that the benefits of traffic metering shown in table 6-2 are in fact the improvements that could be achieved in addition to signal timing optimization (see Section 6.1). In other words, the traffic signal timings of the network were optimized before the application of traffic metering.

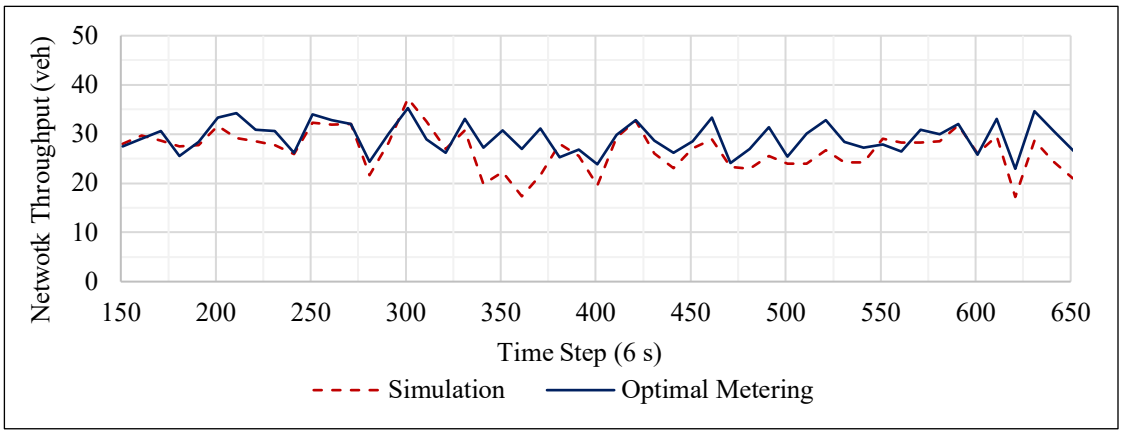
Hence, we could achieve further improvement in traffic operations with another level of control using traffic metering.

Table 6-2. Network performance measures for different scenarios

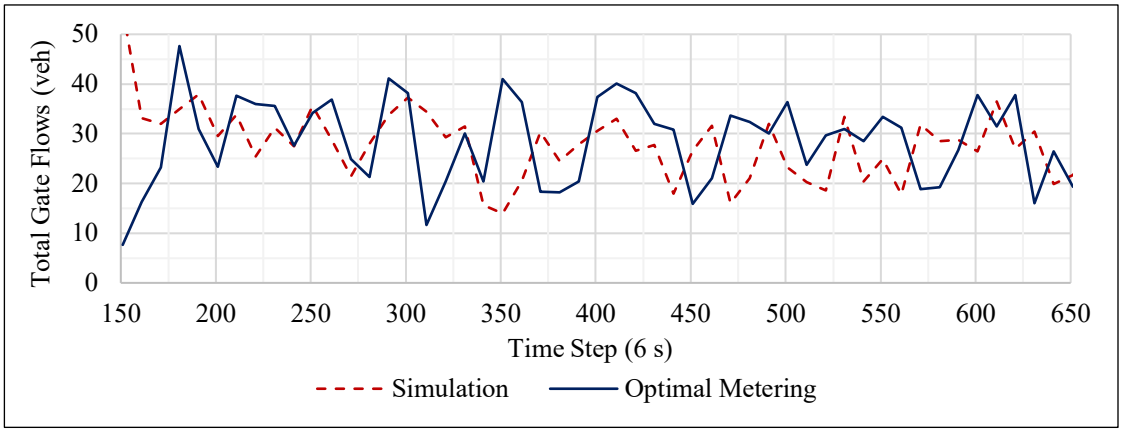
Criterion			
	Demand Profile 1		
Completed trips (veh)	18,681	19,708	5.5
Travel time of vehicles inside the network (min)	141,088	92,861	-34.2
Travel time of vehicles at the gates (min)	530,493	542,715	2.3
System-level travel time (min)	671,581	635,576	-5.4
Delay of vehicles inside the network (veh-min)	116,508	67,252	-42.3
Delay of vehicles at the gates (veh-min)	527,672	539,894	2.3
System-level delay (veh-min)	644,180	607,146	-5.7
System-level average speed (mph)	1.4	1.5	10.3
	Demand Profile 2		
Number of completed trips (veh)	19,265	19,926	3.4
Travel time of vehicles inside the network (min)	147,183	101,917	-30.8
Travel time of vehicles at the gates (min)	518,380	545,426	5.2
System-level travel time (min)	665,563	647,343	-2.7
Delay of vehicles inside the network (veh-min)	123,193	77,406	-37.2
Delay of vehicles at the gates (veh-min)	516,062	543,107	5.2
System-level delay (veh-min)	639,255	620,513	-2.9
System-level average speed (mph)	1.3	1.4	5.2

Figures 6-3 (a) and (c) show the network throughput, and figures 6-3 (b) and (d) show the gate flows for Demand Profiles 1 and 2 over time. Figures 6-3 (a) and (c) show that the Optimal Metering scenario could maintain higher network throughput than the Simulation scenario over

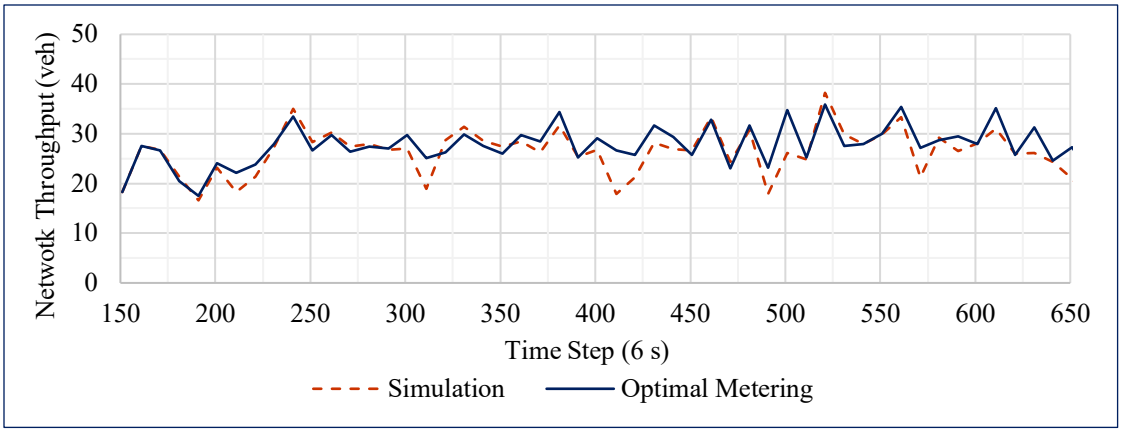
time. This improvement was achieved by reducing the number of vehicles inside the network once traffic metering had been applied at the beginning of the analysis at time step 150 (see figure 6-3s (b) and (d)). Note that the network was loaded from time step 0 to 150 without traffic metering, and the metering gates were activated from time step 150 until the end of the analysis period. After a significant reduction in the gate flows at time step 150, the gate flows in the Optimal Metering scenario fluctuated over a time that represented the traffic metering application.



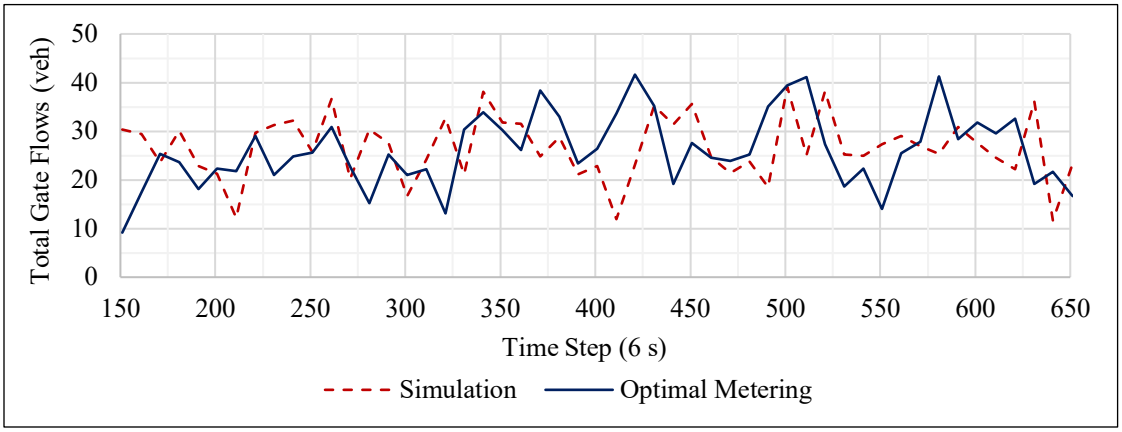
(a) Network throughput, Demand Profile 1



(b) Gate flows, Demand Profile 1



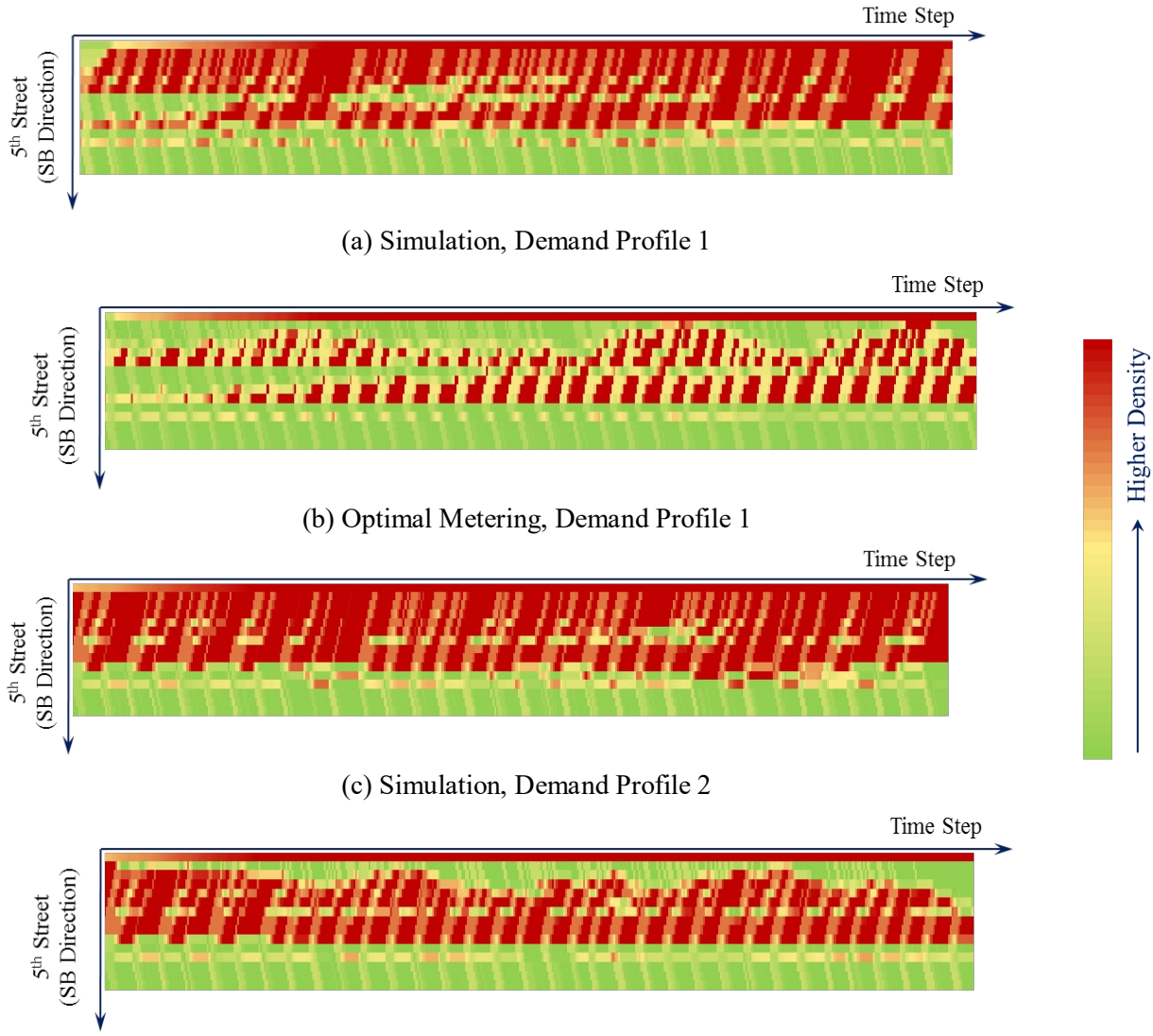
(c) Network throughput, Demand Profile 2



(d) Gate flows, Demand Profile 2

Figure 6-3. Network throughput and gate flows for two demand profiles

The flow reductions and traffic metering could be better observed by comparing the density profiles of a sample arterial street in the Simulation and Optimal Metering scenarios. Figure 6-4 shows the density profiles on 5th Street (figure 6-1) in the mentioned scenarios. The profiles in figure 6-4 show that the Simulation scenario resulted in high densities over an extended section of 5th Street. However, traffic metering reduced both spatial and temporal density profiles.



(a) Simulation, Demand Profile 1

(b) Optimal Metering, Demand Profile 1

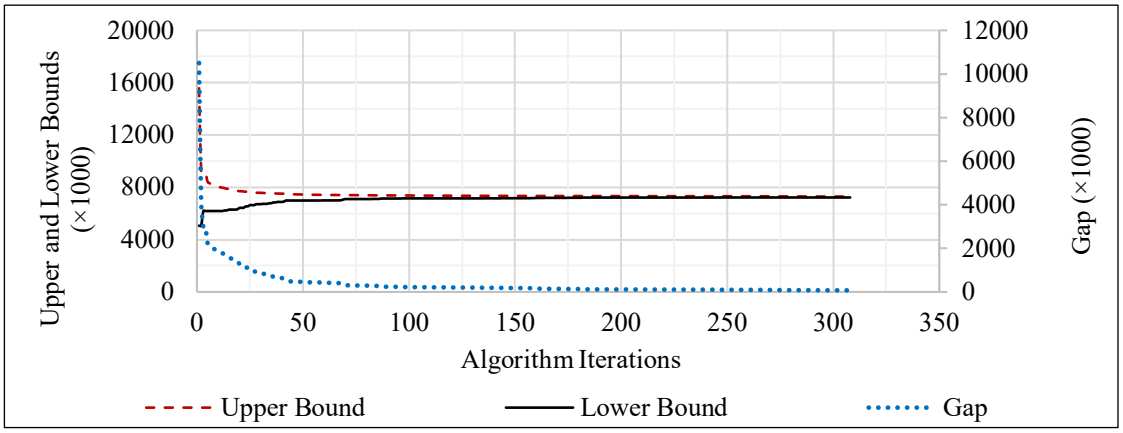
(c) Simulation, Demand Profile 2

(d) Optimal Metering, Demand Profile 2

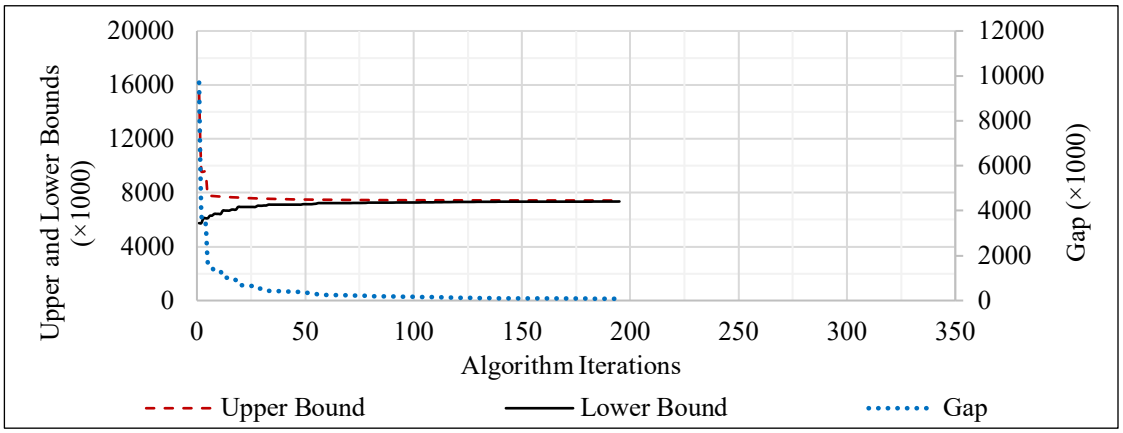
Figure 6-4. Density diagram for one arterial street

We set the convergence criterion of the optimal traffic metering algorithm to a 1 percent gap between the upper bound and lower bounds (see figure 6-3), and the algorithm was terminated once the criterion had been met. Figure 6-5 shows the bounds and their gap over the iterations. In the beginning, the gap between the bounds was relatively large, but the gap was reduced iteratively. The algorithm required 310 iterations to converge for Demand Profile 1 and 190 iterations for Demand Profile 2. Although the criterion was met after those iterations, the algorithm reduced the

gap significantly over the first 50 iterations. Therefore, we could terminate the algorithm after 50 iterations and find good solutions relatively quickly. The runtime for finding a solutions for Demand Profile 1 was 6.8 hours and for Demand Profile 2 was 4.7 hours on a quad-core PC with 18 GB of memory.



(a) Demand Profile 1



(b) Demand Profile 2

Figure 6-5. Convergence diagram of the algorithm

6.4. Real-Time Traffic Metering Results

While the optimal traffic metering solutions provided optimality bounds for the traffic metering solutions, they could not be found in real time (6.8 hours for Demand Profile 1 and 4.7

hours for Demand Profile 2 on a quad-core PC with 18 GB of memory). The Distributed Optimization and Coordination Algorithms (DOCA) approach compromises optimality to gain runtime efficiency. Therefore, it was critical to evaluate the quality of the DOCA-DTM solutions to ensure that the solutions were near-optimal. We designed different cases by changing the turning percentages in the case study to 5 percent, 10 percent, 15 percent, and 20 percent. We also used CTM as the plant model to make a fair comparison between the solutions of the DOCA-DTM and Optimal Metering apporaches. Table 6-3 shows the objective values of the DOCA-DTM and upper and lower bounds of the Optimal Metering solutions for different turning percentages and demand profiles. Table 6-3 shows that the solutions of the DOCA-DTM were less than 2.2 percent different from the upper bound solutions that could be found in the Optimal Metering scenario. Note that the upper bound solutions did not represent feasible solutions to the traffic metering problem. On the other hand, the gap between the solutions of the DOCA-DTM and the lower bound solutions (feasible solutions) was less than 1.2 percent. These results indicated that the DOCA-DTM found near-optimal solutions and performed well in the evaluated scenarios.

Table 6-3. Values of the DOCA-DTM and benchmark objective function

Percentage of Turning Movements		5%	10%	15%	20%
		Demand Profile 1			
	DOCA-DTM (veh)	7,601,215	7,434,969	7,045,405	7,230,630
Optimal Metering	Lower Bound (LB, veh)	7,673,203	7,507,528	7,130,123	7,221,560
	Upper Bound (UB, veh)	7,750,663	7,583,315	7,202,112	7,294,204
	Gap LB/UB	1.00	1.00	1.00	1.00
% Difference DOCA-DTM / LB		-0.94	-0.97	-1.19	0.13
% Difference DOCA-DTM / UB		-1.93	-1.96	-2.18	-0.87
		Demand Profile 2			
	DOCA-DTM (veh)	8,043,261	7,691,545	7,434,150	7,374,010

Percentage of Turning Movements		5%	10%	15%	20%
Optimal Metering	Lower Bound (LB, veh)	8,107,692	7,757,333	7,515,236	7,349,618
	Upper Bound (UB, veh)	8,186,066	7,835,667	7,591,137	7,423,570
	Gap LB/UB	0.96	1.00	1.00	1.00
% Difference DOCA-DTM / LB		-0.79	-0.85	-1.08	0.33
% Difference DOCA-DTM / UB		-1.74	-1.84	-2.07	-0.67

We further compared different performance measures resulting from the DOCA-DTM and the feasible solutions of the Optimal Metering scenario, shown in table 6-4. The table shows that in all evaluated scenarios, the number of completed trips was very close to that of the Optimal Metering solutions. Additionally, the total travel time and delay of vehicles in the DOCA-DTM cases were at most 1.5 percent higher than those of the Optimal Metering solutions (see table 6-4). The results of table 6-4 indicate that the DOCA-DTM found near-optimal solutions.

Table 6-4. Network-level performance measures in different scenarios

Performance Measure	Demand Profile 1 with 5% left turns			Demand Profile 1 with 10% left turns		
	Optimal Metering	DOCA-DTM	% Difference	Optimal Metering	DOCA-DTM	% Difference
Completed trips (veh)	21,400	21,144	-1.2	20,804	20,546	-1.2
Total travel time (min)	590,794	598,272	1.3	607,314	614,897	1.2
Total delay (veh-min)	560,090	568,087	1.4	577,283	585,417	1.4
Performance Measure	Demand Profile 1 with 15% left turns			Demand Profile 1 with 20% left turns		
	Optimal Metering	DOCA-DTM	% Difference	Optimal Metering	DOCA-DTM	% Difference
Completed trips (veh)	19,549	19,337	-1.1	19,708	19,777	0.4
Total travel time (min)	645,538	653,853	1.3	635,576	635,331	0.0

Total delay (veh-min)	617,277	626,038	1.4	607,146	606,687	-0.1

Performance Measure	Demand Profile 2 with 5% left turns			Demand Profile 2 with 10% left turns		
	Optimal Metering	DOCA-DTM	% Difference	Optimal Metering	DOCA-DTM	% Difference
Completed trips (veh)	22,030	21,768	-1.2	21,159	20,840	-1.5
Total travel time (min)	572,288	578,883	1.2	606,438	614,054	1.3
Total delay (veh-min)	542,625	549,620	1.3	577,704	586,129	1.5
Performance Measure	Demand Profile 2 with 15% left turns			Demand Profile 2 with 20% left turns		
	Optimal Metering	DOCA-DTM	% Difference	Optimal Metering	DOCA-DTM	% Difference
Completed trips (veh)	20,194	19,935	-1.3	19,926	20,073	0.7
Total travel time (min)	631,356	639,794	1.3	647,343	645,808	-0.2
Total delay (veh-min)	604,057	612,918	1.5	620,513	618,692	-0.3

Table 6-5 shows different network performance measures of the DOCA-DTM in comparison to those of the Simulation solutions once the turning percentages had been set to 15 percent and 20 percent. Like the improvement trends that were observed for the Optimal Metering solutions shown in table 6-2, the DOCA-DTM increased network throughput by 2.8 percent to 5.9 percent in comparison to throughput under the Simulation scenario in different cases. The system-level travel time also decreased by 2.3 percent to 5.4 percent in the DOCA-DTM in comparison to travel time under the Simulation scenario. The results in table 6-5 indicate that the DOCA-DTM could effectively improve traffic operations in the case study under various demand profiles and turning percentages.

Table 6-5. Network performance measures compared with No-Metering

Performance Measure	Demand Profile 1 with 15% left turns			Demand Profile 1 with 20% left turns		
	Simulation	DOCA-DTM	% Difference	Simulation	DOCA-DTM	% Difference
Completed trips (veh)	18,726	19,337	3.3	18,681	19,777	5.9
Travel time of vehicles inside the network (min)	136,965	58,429	-57.3	141,088	61,459	-56.4
Travel time of vehicles at the gates (min)	534,940	595,425	11.3	530,493	573,872	8.2
System-level travel time (min)	671,905	653,853	-2.7	671,581	635,331	-5.4
Delay of vehicles inside the network (veh-min)	112,407	33,435	-70.3	116,508	35,636	-69.4
Delay of vehicles at the gates (veh-min)	532,119	592,604	11.4	527,672	571,051	8.2
System-level delay (veh-min)	644,526	626,038	-2.9	644,180	606,687	-5.8
Performance Measure	Demand Profile 2 with 15% left turns			Demand Profile 2 with 20% left turns		
	Simulation	DOCA-DTM	% Difference	Simulation	DOCA-DTM	% Difference
Completed trips (veh)	19,388	19,935	2.8	19,265	20,073	4.2
Travel time of vehicles inside the network (min)	139,105	62,423	-55.1	147,183	63,481	-56.9
Travel time of vehicles at the gates (min)	515,894	577,371	11.9	518,380	582,327	12.3
System-level travel time (min)	654,999	639,794	-2.3	665,563	645,808	-3.0
Delay of vehicles inside the network (veh-min)	114,908	37,866	-67.0	123,193	38,684	-68.6
Delay of vehicles at the gates (veh-min)	513,575	575,052	12.0	516,062	580,008	12.4
System-level delay (veh-min)	628,483	612,918	-2.5	639,255	618,692	-3.2

The previous results were derived on a CTM-based plant with the same demand and turning percentages that were used in the optimization problems of the DOCA-DTM. Our purpose for using the same parameters was to be able to evaluate the solution quality of the DOCA-DTM. We

set the network demand to Demand Profile 1 and the turning percentages to 20 percent in the DOCA-DTM, but the network demand and the turning percentages were randomly (uniform distribution) changed, respectively, between 0.9 to 1.1 (10 percent change) and 0.95 to 1.05 (5 percent change) of the set parameters of the DOCA-DTM in the plant (i.e., the actual network). This case was designed to emphasize that we might not have the perfect estimations for actual network parameters (parameters of the plant), and we used only estimated parameters (demand rate and turning percentages) in the optimization programs of the DOCA-DTM. Thus, this case showed the performance of the DOCA-DTM once some errors existed in the estimated parameters.

Table 6-6 shows a comparison between the network performance measures of the DOCA-DTM and those of the Simulation scenarios. The results showed similar improvement trends in this condition, as such the DOCA-DTM increased the number of completed trips by 8.1 percent and decreased the system-level travel time by 4.6 percent. Moreover, this scenario indicated that even if an error existed between the parameters of the prediction model and the plant, which is the case for most real-world applications, the DOCA-DTM could effectively enhance network performance.

Table 6-6. Network performance measures when different parameters were used in the plant and DOCA-DTM

Performance Measure	Simulation	DOCA-DTM	% Difference
Completed trips (veh)	16,288	17,615	8.1
Travel time of vehicles inside the network (min)	140,733	53,751	-61.8
Travel time of vehicles at the gates (min)	592,280	645,713	9.0
System-level travel time (min)	733,012	699,463	-4.6
Delay of vehicles inside the network (veh-min)	119,950	31,398	-73.8
Delay of vehicles at the gates (veh-min)	589,456	642,889	9.1
System-level delay (veh-min)	709,406	674,288	-5.0

Figure 6-6 shows the runtimes for solving the optimization programs in the DOCA-DTM for two of the sub-networks (Monroe Street and 9th Street). The runtimes were found by solving the DOCA-DTM on a PC with a quad-core processor and 16 GB of memory. The runtimes fluctuated between 1.00 and 2.00 seconds. Although we did not consider communication times, given that the time step was 6 seconds (see table 6-1), there was still a reasonable margin for all other processes that might be involved to ensure the real-time performance of the algorithm. In other words, the runtimes in figure 6-6 suggest that the DOCA-DTM was able to find solutions in real time.

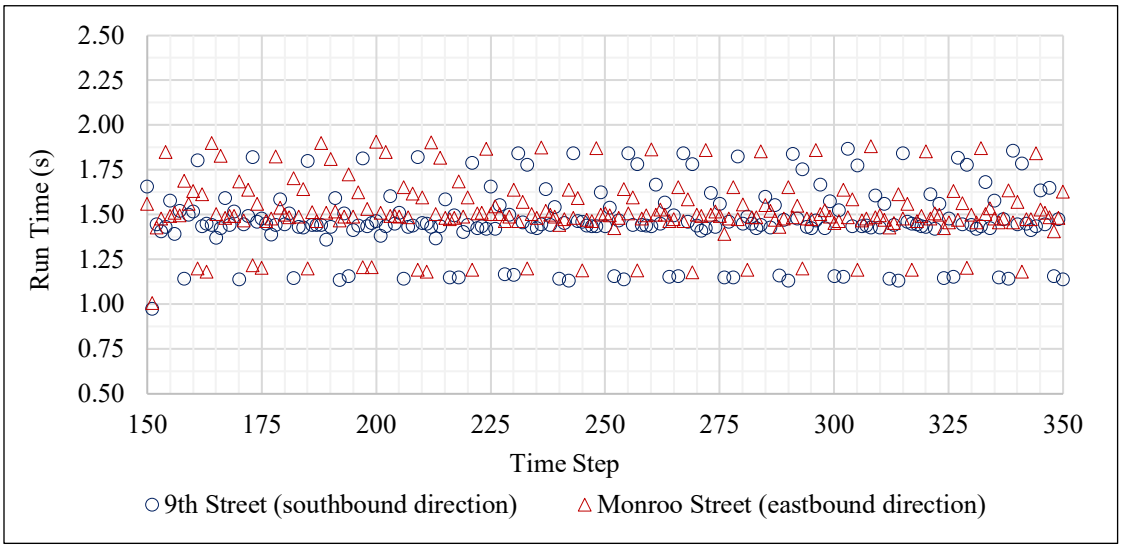
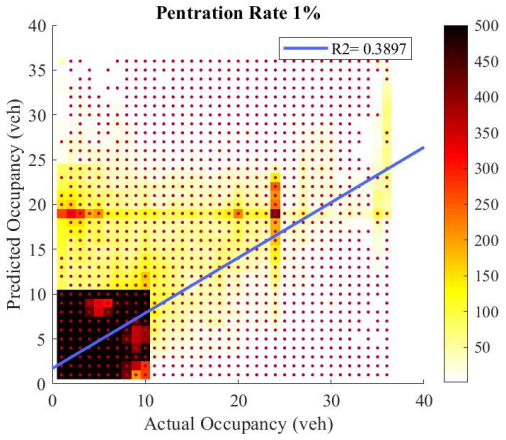


Figure 6-6. The runtime for solving the segment optimization model for two sub-networks

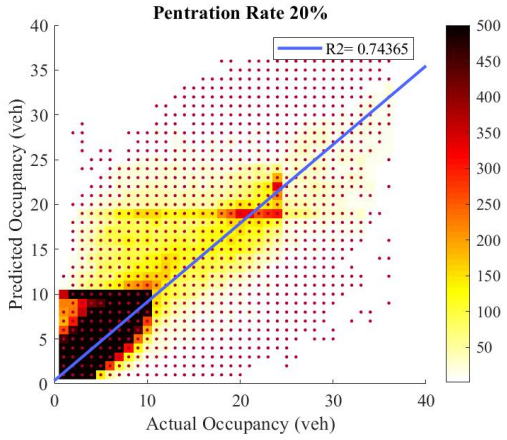
6.5. Effects of Connected Vehicles Penetration Rate on Traffic Metering

The results discussed in section 6.4 were for a 100 percent connected vehiclemarket share in a CTM plant model. We also evaluated the effects of different market penetration rates on traffic metering. We used the Vissim microscopic traffic simulator (PTV Group, 2013) as the plant and solved the DOCA-DTM for market penetration rates of 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 percent.

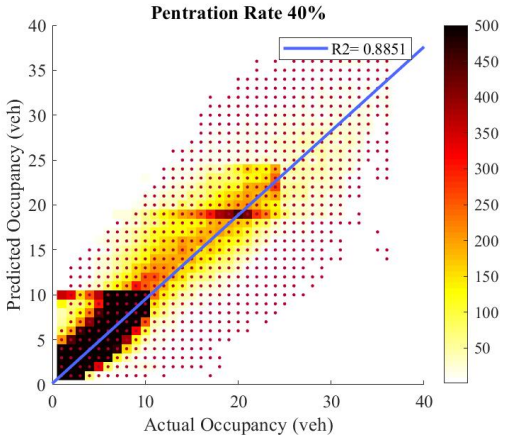
The penetration rate of CVs affected the accuracy of the estimated cell occupancies at the onset of solving the sub-network-level optimization programs because the programs in the DOCA-DTM required initial cell occupancies, the predicted number of vehicles entering the dummy gate cells, and the available capacity of the dummy sink cells (see Chapter 5). Hence, the market penetration rate of CVs had direct effects on these estimations. Figure 6-7 shows the actual cell occupancies versus the DOCA-DTM estimations for four penetration rates and two demand profiles. The colors of the figures show the density of points in different regions of the figures: a darker color shows a higher density of points. Figures 6-7 (a) and (e) show that once the penetration rate was almost zero, the predicted R-squared of the fitted regression line on the data had low values of 0.39, meaning that occupancy prediction in the absence of CV information was not accurate enough. However, as the penetration rate increased to 20 percent, the R-squared values significantly increased to 0.74 in Demand Profile 1 and 0.71 in Demand Profile 2 (see figure 6-7 (b) and (f)). Hence, only a 20 percent increase in the penetration rated had a significant effect on accuracy prediction. Once the penetration rate was 60 percent, the R-squared values were more than 0.9. Figures 6-7 (a) to (h) show that increasing the market penetration rate had a positive effect on occupancy prediction for the traffic metering application.



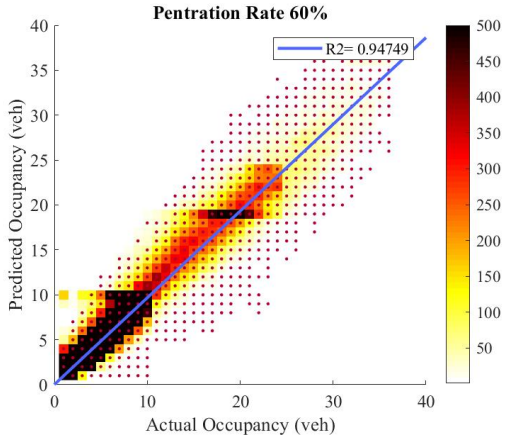
(a) Demand Profile 1, 1% penetration rate



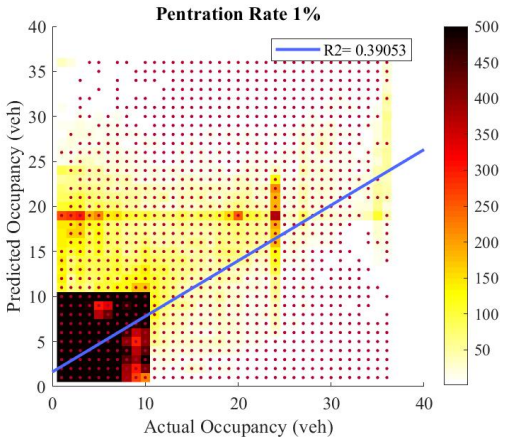
(b) Demand Profile 1, 20% penetration rate



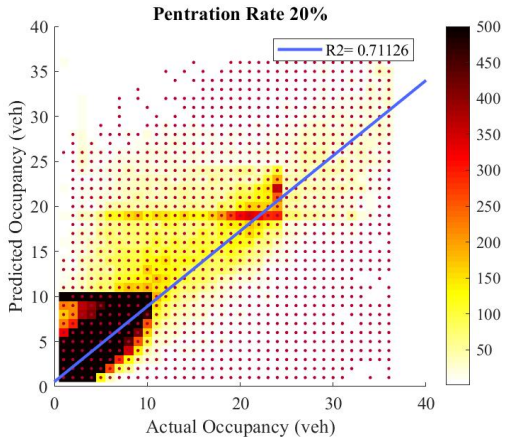
(c) Demand Profile 1, 40% penetration rate



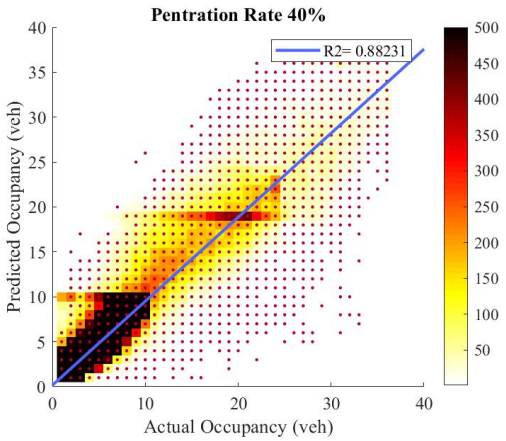
(d) Demand Profile 1, 60% penetration rate



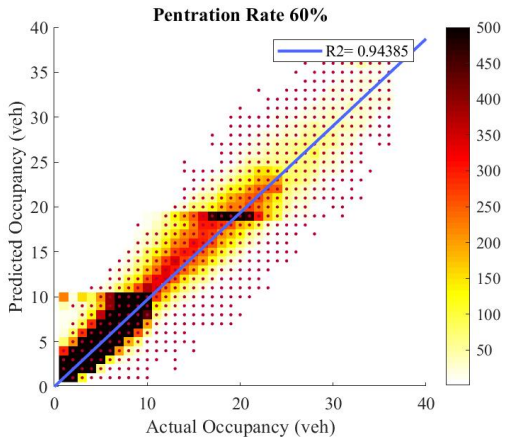
(e) Demand Profile 2, 1% penetration rate



(f) Demand Profile 2, 20% penetration rate



(g) Demand Profile 2, 40% penetration rate

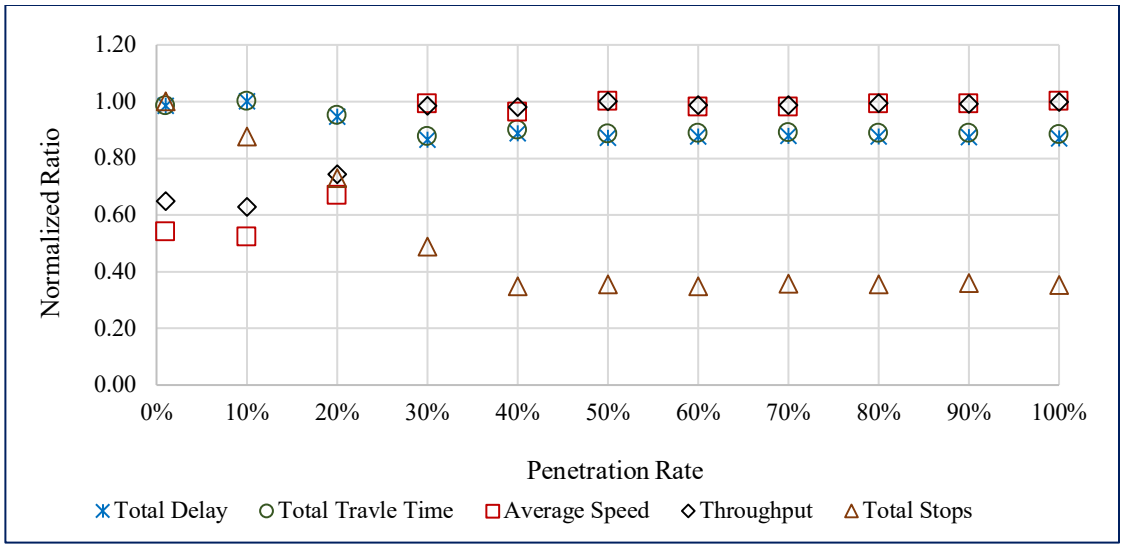


(h) Demand Profile 2, 60% penetration rate

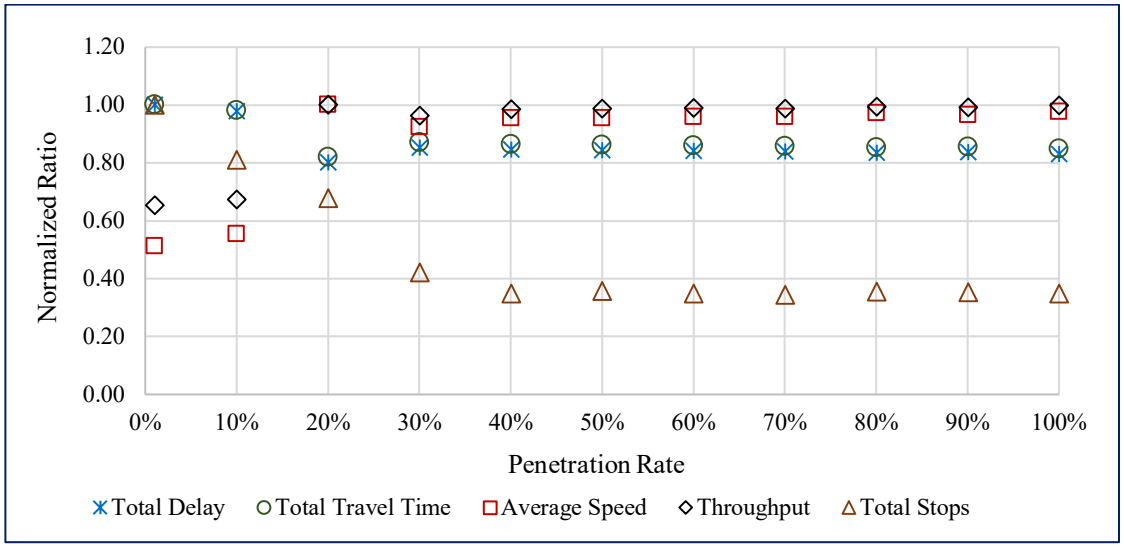
Figure 6-7. Prediction accuracy with respect to different penetration rates of CVs

We evaluated the effect of occupancy estimation accuracy on the solutions of the DOCA-DTM (figure 6-8). As shown in the figure, we normalized the values of different performance measures to be between 0.0 and 1.0. The normalization allowed us to show total delay, total travel time, average speed, throughput, and the total number of stops with unit-less values on the same vertical axis in figures 6-8 (a) and (b). The performance of the network improved with an increasing penetration rate of CVs: total delay and travel time decreased, and average speed and network throughput increased.

The results of figure 6-8 also show that increasing the market penetration rate from 1 percent to 30 percent significantly improved network performance, while the improvements beyond 30 percent were not significant. Hence, we can conclude that in our case studies, a 30 percent market penetration rate was critical for traffic.



(a) Demand Profile 1



(b) Demand Profile 2

Figure 6-8. Normalized performance measures of the case study for different penetration rates of connected vehicles

Chapter 7: Summary and Conclusions

Traffic congestion is one of the factors that contribute to excessive air pollution, traveler delay, and fuel consumption in urban street networks. As traffic congestion increases, the network capability to process vehicles decreases because of queue spillovers and gridlocks. Therefore, regulating the flow of vehicles entering congested areas allows the network to operate at its optimum level.

Traffic metering or perimeter control is one of the traffic congestion management strategies that improves traffic operations by preventing gridlock and queue spillbacks. Bimodal traffic signals can be placed at the borders of congested areas to regulate the flow of vehicles.

Thus far, the available traffic metering approaches have either failed to provide dynamic optimal metering rates or have required well-defined MFDs that cannot be easily derived for general transportation networks. This study filled this gap by developing a traffic metering formulation and two solution techniques to dynamically optimize traffic metering rates for predefined gate locations.

In this study, we developed a mixed-integer, non-linear optimization program that maximized the number of completed trips by optimizing traffic metering rates. The program used the cell transmission model to capture traffic flow dynamics. The non-linear flow conservation and feasibility constraints of the program made solving the program such a very challenging task that commercial optimization solvers could not solve the program for a reasonable-sized network.

Therefore, we employed the Benders decomposition technique, utilized the special structure of the problem, and proposed a solution technique that could solve the program within an optimality bound. The objective of this solution technique was to provide the capability of solving the program to a certain optimality level regardless of its runtime efficiency. The results

of applying this approach on a case study network showed that traffic metering significantly improved traffic operations by reducing the travel time of vehicles inside the protected area by 30.8 percent and 34.2 percent in comparison to a no-metering strategy. Some vehicles were delayed at the network gates because of traffic metering, but traffic metering reduced the system-level travel times, including vehicles both inside and at the gates of the network by 2.7 percent and 5.4 percent. These results indicated that traffic metering improved traffic operations by regulating the flow of vehicles at the gates. The solution technique reached an optimality gap of 1 percent for different tested demand profiles. These solutions were found in several hours of computation time. However, traffic operations require control strategies that can be implemented in real time.

To address the runtime issue, we presented a distributed model predictive control approach. The approach distributed the network-level traffic metering problem to several sub-network level problems with an allocated computational node and coordinated their decisions. The density of vehicles across network links was estimated by using connected vehicles and loop detector data at discretized time steps. The estimated densities were shared with the computational nodes, and each node optimized traffic metering rates for its assigned gates. Then, the first optimized values were implemented at the gates as red or green gate signal indications. This solution technique significantly reduced runtime. We performed several analyses and compared the results of this solution technique with the optimal solutions: the results showed that the solutions of this approach were only 2.2 percent different from the optimal solutions while they were found in realtime.

Furthermore, sensitivity analysis on the market penetration rate of connected vehicles in a microscopic simulation environment showed that the density estimation accuracy increased with the connected vehicle penetration rate. The network performance also improved with the market penetration rate. The increase in performance was significant from a 0 percent to 30 percent CV

market share. However, improvements were negligible for penetration rates of more than 30%. Therefore, the proposed traffic metering approach required at least a 30 percent penetration rate of connected vehicles in our case studies.

Connected automated vehicles are expected to bring significant changes to transportation systems, as studies show considerable improvements in intersection operations (Mirheli et al., 2019, 2018; Niroumand et al., 2020). It is important and interesting to study traffic metering effects on traffic operations in the presence of connected automated vehicles.

References

Ahuja, R.K., Magnanti, T.L., Orlin, J.B., 1993. Network flows: theory, algorithms, and applications.

Aziz, H.M.A., Ukkusuri, S. V., 2012. Integration of Environmental Objectives in a System Optimal Dynamic Traffic Assignment Model. Comput. Civ. Infrastruct. Eng. 27, 494–511. https://doi.org/10.1111/j.1467-8667.2012.00756.x

Benders, J.F., 1962. Partitioning procedures for solving mixed-variables programming problems. Numer. Math. 4, 238–252.

CPLEX, I.B.M.I., 2009. V12. 1: User's Manual for CPLEX. Int. Bus. Mach. Corp. 46, 157.

Daganzo, C.F., 2007. Urban gridlock: Macroscopic modeling and mitigation approaches. Transp. Res. Part B Methodol. 41, 49–62.

Daganzo, C.F., 1995. The cell transmission model, part II: Network traffic. Transp. Res. Part B Methodol. 29, 79–93.

Daganzo, C.F., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transp. Res. Part B Methodol. 28, 269–287.

De Gier, J., Garoni, T.M., Rojas, O., 2011. Traffic flow on realistic road networks with adaptive traffic lights. J. Stat. Mech. Theory Exp. 2011, P04008.

Ding, H., Guo, F., Zheng, X., Zhang, W., 2017. Traffic guidance--perimeter control coupled method for the congestion in a macro network. Transp. Res. Part C Emerg. Technol. 81, 300–316.

Fu, H., Liu, N., Hu, G., 2017. Hierarchical perimeter control with guaranteed stability for dynamically coupled heterogeneous urban traffic. Transp. Res. Part C Emerg. Technol. 83, 18–38.

Geoffrion, A.M., 1972. Generalized Benders decomposition. J. Optim. Theory Appl. 10, 237–260.

Geroliminis, N., Daganzo, C.F., 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. Transp. Res. Part B Methodol. 42, 759–770.

Geroliminis, N., Haddad, J., Ramezani, M., 2013. Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach. IEEE Trans. Intell. Transp. Syst. 14, 348–359.

Geroliminis, N., Sun, J., 2011. Properties of a well-defined macroscopic fundamental diagram for urban traffic. Transp. Res. Part B Methodol. 45, 605–617.

Godfrey, J.W., 1900. The mechanism of a road network. Traffic Eng. Control 8.

Gomes, G., Horowitz, R., 2006. Optimal freeway ramp metering using the asymmetric cell transmission model. Transp. Res. Part C Emerg. Technol. 14, 244–262. https://doi.org/10.1016/j.trc.2006.08.001

Haddad, J., 2017. Optimal perimeter control synthesis for two urban regions with aggregate boundary queue dynamics. Transp. Res. Part B Methodol. 96, 1–25.

Haddad, J., Geroliminis, N., 2012. On the stability of traffic perimeter control in two-region urban cities. Transp. Res. Part B Methodol. 46, 1159–1176.

Haddad, J., Mirkin, B., 2017. Coordinated distributed adaptive perimeter control for large-scale urban road networks. Transp. Res. Part C Emerg. Technol. 77, 495–515.

Haddad, J., Ramezani, M., Geroliminis, N., 2013. Cooperative traffic control of a mixed network with two urban regions and a freeway. Transp. Res. Part B Methodol. 54, 17–36.

Haddad, J., Shraiber, A., 2014. Robust perimeter control design for an urban region. Transp. Res. Part B Methodol. 68, 315–332.

Hajbabaie, A., 2012. Intelligent dynamic signal timing optimization program. University of Illinois at Urbana-Champaign.

Hajbabaie, A., Benekohal, R., 2013. Traffic signal timing optimization: Choosing the objective function. Transp. Res. Rec. J. Transp. Res. Board 10--19.

Hajbabaie, A., Benekohal, R.F., 2015. A Program for Simultaneous Network Signal Timing Optimization and Traffic Assignment. IEEE Trans. Intell. Transp. Syst. 16, 2573–2586. https://doi.org/10.1109/TITS.2015.2413360

Hajbabaie, Ali, Benekohal, R.F., 2011. Does traffic metering improve network performance efficiency?, in: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC. pp. 1114–1119.

Hajbabaie, A, Benekohal, R.F., 2011. Which policy works better for signal coordination? Common, or variable cycle length, in: Proceedings of the 1st ASCE T&DI Congress. pp. 13–16.

Hajbabaie, A., Medina, J.C., Benekohal, R.F., 2010. Effects of ITS-Based Left Turn Policies on Network Performance, in: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC. pp. 80–84.

Hajbabaie, A., Medina, J.C., Benekohal, R.F., NEXTRANS, C., 2011. Traffic signal coordination and queue management in oversaturated intersections. Purdue Univ. Discov. Park.

Hajibabai, L., Nourbakhsh, S., Ouyang, Y., Peng, F., 2014. Network Routing of Snowplow Trucks with Resource Replenishment and Plowing Priorities: Formulation, Algorithm, and Application. Transp. Res. Rec. J. Transp. Res. Board 16–25.

Hajibabai, L., Ouyang, Y., 2016. Dynamic Snow Plow Fleet Management Under Uncertain Demand and Service Disruption. IEEE Trans. Intell. Transp. Syst. 17, 2574–2582.

Hajibabai, L., Ouyang, Y., 2013. Integrated planning of supply chain networks and multimodal transportation infrastructure expansion: model development and application to the biofuel industry. Comput. Civ. Infrastruct. Eng. 28, 247–259.

Hajibabai, L., Saha, D., 2019. Patrol route planning for incident response vehicles under dispatching station scenarios. Comput. Civ. Infrastruct. Eng. 34, 58–70.

Han, K., Liu, H., Gayah, V. V, Friesz, T.L., Yao, T., 2016. A robust optimization approach for dynamic traffic signal control with emission considerations. Transp. Res. Part C Emerg. Technol. 70, 3–26.

Herman, R., Prigogine, I., 1979. A two-fluid approach to town traffic. Science (80-.). 204, 148–151.

Hosseini, P., Savla, K., 2016. A comparison study between proportionally fair and max pressure controllers for signalized arterial networks.

Islam, S.M.A. bin Al, Aziz, H.M.A., Hajbabaie, A., 2020. Stochastic Gradient-Based Optimal Signal Control With Energy Consumption Bounds. IEEE Trans. Intell. Transp. Syst. 1–14.

Islam, S.M.A.B. Al, Hajbabaie, A., 2017. Distributed coordinated signal timing optimization in connected transportation networks. Transp. Res. Part C Emerg. Technol. 80, 272–285. https://doi.org/10.1016/j.trc.2017.04.017

Jafari, E., Pandey, V., Boyles, S.D., 2017. A decomposition approach to the static traffic assignment problem. Transp. Res. Part B Methodol. 105, 270–296.

Ji, Y., Geroliminis, N., 2012. On the spatial partitioning of urban transportation networks. Transp. Res. Part B Methodol. 46, 1639–1656.

Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I., Papageorgiou, M., 2012. Exploiting the fundamental diagram of urban networks for feedback-based gating. Transp. Res. Part B Methodol. 46, 1393–1403.

Keyvan-Ekbatani, M., Yildirimoglu, M., Geroliminis, N., Papageorgiou, M., 2015. Multiple concentric gating traffic control in large-scale urban networks. IEEE Trans. Intell. Transp. Syst. 16, 2141–2154.

Kouvelas, A., Saeedmanesh, M., Geroliminis, N., 2017. Enhancing model-based feedback perimeter control with data-driven online adaptive optimization. Transp. Res. Part B Methodol. 96, 26–45.

Lee, C., Hellinga, B., Saccomanno, F., 2006. Evaluation of variable speed limits to improve traffic safety. Transp. Res. Part C Emerg. Technol. 14, 213–228. https://doi.org/10.1016/j.trc.2006.06.002

Levin, M.W., Boyles, S.D., Duthie, J., Pool, C.M., 2016. Demand profiling for dynamic traffic assignment by integrating departure time choice and trip distribution. Comput. Civ. Infrastruct. Eng. 31, 86–99.

Levy, J.I., Buonocore, J.J., Von Stackelberg, K., 2010. Evaluation of the public health impacts of traffic congestion: a health risk assessment. Environ. Heal. 9, 65.

Li, Z., Liu, P., Xu, C., Wang, W., 2015. Optimal Mainline Variable Speed Limit Control to Improve Safety on Large-Scale Freeway Segments. Comput. Civ. Infrastruct. Eng.

Liang, X.J., Guler, S.I., Gayah, V. V, 2020. An equitable traffic signal control scheme at isolated signalized intersections using Connected Vehicle technology. Transp. Res. Part C Emerg. Technol. 110, 81–97.

Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads. Proc. R. Soc. A Math. Phys. Eng. Sci. 229, 317–345.

Lo, H.K., 1999. A novel traffic signal control formulation. Transp. Res. Part A Policy Pract. 33, 433–448.

Mahmassani, H.S., Saberi, M., Zockaie, A., 2013. Urban network gridlock: Theory, characteristics, and dynamics. Transp. Res. Part C Emerg. Technol. 36, 480–497.

Mahmassani, H.S., Williams, J.C., Herman, R., 1984. Investigation of network-level traffic flow relationships: some simulation results. Transp. Res. Rec. 971, 121–130.

Medina, J., Hajbabaie, A., Benekohal, R., 2013. Effects of metered entry volume on an oversaturated network with dynamic signal timing. Transp. Res. Rec. J. Transp. Res. Board 53--60.

Medina, J.C., Hajbabaie, A., Benekohal, R.F., 2011. A comparison of approximate dynamic programming and simple genetic algorithm for traffic control in oversaturated conditions - Case study of a simple symmetric network, in: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC. pp. 1815–1820.

Medina, J.C., Hajbabaie, A., Benekohal, R.F., 2010. Arterial traffic control using reinforcement learning agents and information from adjacent intersections in the state and reward structure, in: 13th International IEEE Conference on Intelligent Transportation Systems. pp. 525–530.

Mehrabipour, M., 2018. Real-time Network-level Signal Timing Optimization. Washington State University.

Mehrabipour, M., Hajbabaie, A., 2017. A cell based distributed-coordinated approach for network-level signal timing optimization. Comput. Civ. Infrastruct. Eng. 32, 599–616. https://doi.org/10.1111/mice.12272

Mehrabipour, M., Hajibabai, L., Hajbabaie, A., 2019. A decomposition scheme for parallelization of system optimal dynamic traffic assignment on urban networks with multiple origins and destinations. Comput. Civ. Infrastruct. Eng.

Mirheli, A., Hajibabai, L., 2020. Utilization Management and Pricing of Parking Facilities Under Uncertain Demand and User Decisions. IEEE Trans. Intell. Transp. Syst. 21, 2167–2179.

Mirheli, A., Hajibabai, L., Hajbabaie, A., 2018. Development of a Signal-head-free Intersection Control Logic in a Fully Connected and Autonomous Vehicle Environment. Transp. Res. Part C Emerg. Technol.

Mirheli, A., Tajalli, M., Hajibabai, L., Hajbabaie, A., 2019. A consensus-based distributed trajectory control in a signal-free intersection. Transp. Res. Part C Emerg. Technol. 100, 161–176. https://doi.org/10.1016/j.trc.2019.01.004

Mirheli, A., Tajalli, M., Mohebifard, R., Hajibabai, L., Hajbabaie, A., 2020. Utilization Management of Highway Operations Equipment. Transp. Res. Rec. 0361198120927400.

Mohebifard, R., Hajbabaie, A., 2019. Optimal network-level traffic signal control: A benders decomposition-based solution algorithm. Transp. Res. Part B Methodol. 121, 252–274. https://doi.org/10.1016/j.trb.2019.01.012

Mohebifard, R., Hajbabaie, A., 2018a. Dynamic traffic metering in urban street networks: Formulation and solution algorithm. Transp. Res. Part C Emerg. Technol. 93, 161–178. https://doi.org/10.1016/j.trc.2018.04.027

Mohebifard, R., Hajbabaie, A., 2018b. Distributed Optimization and Coordination Algorithms for Dynamic Traffic Metering in Urban Street Networks. IEEE Trans. Intell. Transp. Syst. 1–12. https://doi.org/10.1109/TITS.2018.2848246

Mohebifard, R., Hajbabaie, A., 2018c. Real-Time Adaptive Traffic Metering in a Connected Urban Street Network, in: Transportation Research Board 97th Annual Meeting.

Mohebifard, R., Islam, S.M.A.B. Al, Hajbabaie, A., 2019. Cooperative traffic signal and perimeter control in semi-connected urban-street networks. Transp. Res. Part C Emerg. Technol. 104, 408–427. https://doi.org/10.1016/j.trc.2019.05.023

Nilsson, G., Hosseini, P., Como, G., Savla, K., 2015. Entropy-like Lyapunov functions for the stability analysis of adaptive traffic signal controls, in: 2015 54th IEEE Conference on Decision and Control (CDC). pp. 2193–2198.

Niroumand, R., Tajalli, M., Hajibabai, L., Hajbabaie, A., 2020. Joint optimization of vehicle-group trajectory and signal timing: Introducing the white phase for mixed-autonomy traffic stream. Transp. Res. Part C Emerg. Technol. 116, 102659.

PTV Group, 2013. PTV Vissim 7 User Manual. PTV AG.

Rathi, A.K., Lieberman, E.B., 1989. Effectiveness of traffic restraint for a congested urban network: A simulation study.

Richards, P.I., 1956. Shock waves on the highway. Oper. Res.

Samaranayake, S., Krichene, W., Reilly, J., Monache, M.L.D., Goatin, P., Bayen, A., 2018. Discrete-time system optimal dynamic traffic assignment (SO-DTA) with partial control for physical queuing networks. Transp. Sci. 52, 982–1001.

Sirmatel, I.I., Geroliminis, N., 2017. Economic Model Predictive Control of Large-Scale Urban Road Networks via Perimeter Control and Regional Route Guidance. IEEE Trans. Intell. Transp. Syst.

Stevanovic, A., Martin, P., Stevanovic, J., 2007. VisSim-based genetic algorithm optimization of signal timings. Transp. Res. Rec. J. Transp. Res. Board 59–68.

Tajalli, M., Hajbabaie, A., 2018a. Dynamic Speed Harmonization in Connected Urban Street Networks. Comput. Civ. Infrastruct. Eng. 33, 510–523. https://doi.org/10.1111/mice.12360

Tajalli, M., Hajbabaie, A., 2018b. Distributed optimization and coordination algorithms for dynamic speed optimization of connected and autonomous vehicles in urban street networks. Transp. Res. Part C Emerg. Technol. 95, 497–515. https://doi.org/10.1016/j.trc.2018.07.012

Tajalli, M., Mehrabipour, M., Hajbabaie, A., 2020. Network-Level Coordinated Speed Optimization and Traffic Light Control for Connected and Automated Vehicles. IEEE Trans. Intell. Transp. Syst. 1–12. https://doi.org/10.1109/TITS.2020.2994468

Yan, H., He, F., Lin, X., Yu, J., Li, M., Wang, Y., 2019. Network-level multiband signal coordination scheme based on vehicle trajectory data. Transp. Res. Part C Emerg. Technol. 107, 266–286.

Yang, K., Guler, S.I., Menendez, M., 2016. Isolated intersection control for various levels of vehicle technology: Conventional, connected, and automated vehicles. Transp. Res. Part C Emerg. Technol. 72, 109–129.

Yang, K., Menendez, M., Zheng, N., 2019. Heterogeneity aware urban traffic control in a connected vehicle environment: A joint framework for congestion pricing and perimeter control. Transp. Res. Part C Emerg. Technol. 105, 439–455.

Zhang, L., Garoni, T.M., de Gier, J., 2013. A comparative study of macroscopic fundamental diagrams of arterial road networks governed by adaptive traffic signal systems. Transp. Res. Part B Methodol. 49, 1–23.

Zhu, F., Ukkusuri, S. V., 2014. Accounting for dynamic speed limit control in a stochastic traffic environment: A reinforcement learning approach. Transp. Res. Part C Emerg. Technol. 41, 30–47. https://doi.org/10.1016/j.trc.2014.01.014

Zhu, F., Ukkusuri, S. V, 2013. A cell based dynamic system optimum model with non-holding back flows. Transp. Res. Part C Emerg. Technol. 36, 367–380.

Appendix A. List of Notations

Table A-1 List of notations

Sets	
T	set of all time steps
C	set of all network cells
C_G	set of all gate cells
C_S	set of all sink cells
C_I	set of all intersection cells
C_D	set of all diverge cells
C_M	set of all merge cells
C_A	set of cells immediately downstream of gate cells
CD_G	set of all dummy resource cells
CD_S	set of all dummy sink cells
L	set of all network links
$S(i)$	set of all successor cells of cell $i \in C$
$P(i)$	set of all predecessor cells of cell $i \in C$
$\mathcal{S}$	set of all sub-networks
Decision (control) variables	
y_{ij}^t	number of vehicles advancing from gate cell $i \in C_G$ to cell $j \in S(i)$ at time step $t \in T$
Variables	
x_i^t	state variables; number of vehicles in cell $i \in C$ at time step $t \in T$
y_{ij}^t	number of vehicles advancing from cell $i \in C \backslash C_G$ to cell $j \in S(i)$ at time step $t \in T$
Z	objective value of the traffic metering program
μ	objective value of the master problem
$\theta_{ij}^t, \vartheta_{ij}^t, \psi_{ij}^t, \chi_{ij}^t$	binary variables
Parameters	
D_i^t	demand of gate cell $i \in C_G$ at time step $t \in T$
Q_i^t	saturation flow rate in cell $i \in C$ at time step $t \in T$

N_i	capacity of cell $i \in C$ in terms of the number of vehicles it can hold
β_i^t	portion of flow entering intersection cell $i \in C_I$ at time step $t \in T$ from the total flow leaving its upstream diverge cell
g_i^t	signal state of cell $i \in C_I$ at time step $t \in T$ that is equal to one for green and zero for red indications
$\hat{y}_{ij}^t$	temporarily fixed values if gate flows leaving gate $i \in C_G$ to cell $j \in S(i)$ at time step $t \in T$
x^{*t}_i	optimized number of vehicles in cell $i \in C$ at time step $t \in T$
y^{*t}_{ij}	optimized number of vehicles advancing from cell $i \in C$ to cell $j \in S(i)$ at time step $t \in T$
λ^{*t}_i	dual values of constraint i at time step $t \in T$
PR	penetration rate of connected vehicles
xe_i^t	number of equipped vehicles with onboard units (connected vehicles) in cell $i \in C$ at time step $t \in T$
V_l^t	total number of vehicles in each link $l \in L$ at time step $t \in T$
$x_i^{t,CV}, x_i^{t,CTM}$	estimated occupancy based on connected vehicle location information and CTM flow conservation and feasibility conditions, respectively, for cell $i \in C$ at time step $t \in T$
F, D_J	flow rate and jam density
v, w	speed of respectively forward and backward shockwave speeds
δ	ratio of backward shockwave speed to the free flow speed, $\delta = w/v$
w_i	weight factor for cell $i \in C$
M	the big-M
ρ	penalty parameter