QUEUE LENGTH ESTIMATION AND PLATOON RECOGNITION USING CONNECTED VEHICLE TECHNOLOGY FOR ADAPTIVE SIGNAL CONTROL

A Dissertation

by

KAMONTHEP TIAPRASERT

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Yunlong Zhang
Committee Members,	Xiubin Bruce Wang
	Luca Quadrifoglio
	Kiavash Kianfar
Head of Department,	Robin Autenrieth

May 2016

Major Subject: Civil Engineering

ABSTRACT

This dissertation presents mathematical and analytical models for real-time queue length estimation and platoon recognition using the connected vehicle technology (CVT). Information on queue length and platoon is a crucial part of traffic signal control and is difficult to obtain accurately with traditional technologies such as loop detectors. The past studies are either limited to fixed-time signal control or lacked verification on the applicable range or evaluation of the performance of algorithms.

The proposed algorithms focused on estimating the queue length for adaptive signal control and platoon characteristics for signal coordination and adaptive signal control. For queue length detection, an algorithm was developed to determine the estimated value between the last stopped vehicle and the first moving vehicle for different market penetration ratios. Discrete wavelet transform is applied to the estimated queue lengths to improve accuracy and consistency.

The platoon recognition model is developed based on time headway so that the arrival times can be computed directly from the estimated platoon data. First, the detected platoon is identified by a modified critical time-headway. Then, platoon size and starting and ending times are estimated. Lastly, a filtering process for "qualified" detected platoon is proposed to optimize detectability. The results show that the proposed algorithms can estimate well in various traffic conditions and under both fixed-time and actuated signal control without relying on inputs that are hard to obtain in practice. Furthermore, an analytical model to estimate the platoon detection rate is

proposed and shown to be close to the numerical results. Therefore, Traffic engineers can use the analytical model to estimate the required market penetration ratio for the application without field experiments or microscopic simulation. Accordingly, the proposed algorithms can be an important part of adaptive signal control focusing on real-time coordination.

ACKNOWLEDGEMENTS

I am really grateful to my committee members for their guidance and insight that greatly helped improve the quality of my dissertation. During my PhD. years, I have the important persons; my advisor, Dr. Yunlong Zhang, and my close friend, Adel Khodakarami.

My advisor, Dr. Zhang, is my "G. H. Hardy" who was a mentor to Indian mathematician Srinivasa Ramanujan. While others were not confident in my research capability, only Dr.Zhang almost immediately recognized my worth and always stood for my defense. When I had a rough and slow start at early years, he was always patient with me and my flaws. With his support, I can focus on my mathematics, and this dissertation became complete before I had realized it. It is my greatest honor to work with him.

The relationship between me and Adel is more or less "the Friendship between Guan Zhong and Bao Shuya". To describe Adel, in my humble opinion, there is no better words than the quote "My parents gave birth to me, but it is Bao (Adel) who knows me best!"

This dissertation was intended to be a tribute for the "*first*" teacher in my life. The one who gave my heart and a world to live in was, without a doubt, him. I was taught the true essence of mathematics and the way of living an honest life.

Finally, thanks to my sister, mother and father for their encouragement.

NOMENCLATURE

CV	Connected vehicle
CVT	Connected vehicle technology

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW*

Introduction

Traffic congestion in urban area has been a critical problem in transportation for a long time (*1*). Traffic signal is very important for traffic management. A good traffic signal can reduce delay, travel time, gas emissions and incidents. Though traffic has similar daily patterns, it still varies from hour to hour. Pre-timed signal cannot handle such an unpredictable change so traffic signal that can adapt to real-time incoming demand certainly will provide more effective control. Thus, adaptive signal control has recently gained attentions from traffic engineers as one of the solutions with a great potential. Adaptive signal control can cope with traffic bursts by estimation and prediction from traffic sensors and select the best signal timing based on such information.

Aside from signal control, traffic detector is important as well. The most common type of detector is loop detector. Although it is a very accurate fix-location detector, expensive construction and maintenance cost prevents it from implementing at every necessary location. It is also costly to tracks queue length effectively, especially when queue size has very high variation. With the emerging of wireless technology, connected

vehicle technology (CVT) (*2-8*) is very promising as a probe vehicle. Probe vehicle can collect vehicle data at any location with potentially lower cost than the conventional loop detector. However, a new algorithm and adaptive signal logic are needed to develop for such a new and different detector. Moreover, a low penetration ratio of probe vehicles can lower the performance so the effect of penetration ratio needs to be studied before employing in the real world.

The current challenge of adaptive signal control using CV is to find the one which works best with low penetration ratios and maximize the potential of spatial information of CV. The other challenge is how to estimate traffic state, queue length and platoon as inputs for adaptive signal control without relying on other detectors or basic inputs which are hard to obtain in practice. Most queue length estimation relies on signal timing, duration of red time and assumed arrival distribution. However, such signal data and assumption is not practical especially for adaptive signal control at multiple intersections. Thus, the real-time queue length estimation without relying on signal and arrival data needs to be developed. For platoon recognition, only a few studies were conducted, and none of them had a full analysis on the impact of penetration ratio and performance consistency.

This research aims to develop queue estimation and platoon recognition algorithms for adaptive signal control. The two developed algorithms are verified and validated with microscopic simulation, VISSIM (*9; 10*), in various traffic conditions and signal settings.

Finally, the performance of queue length estimation and platoon recognition using CV is evaluated. The applicability range of the algorithm is identified from numerical results in chapter IV and V.

Literature Review

Studies similar to the dissertation and related topics have been discussed in this chapter. First, advantages and disadvantages of loop detectors and CV are compared. Second, adaptive signal control are reviewed. Lastly, queue length and platoon estimation methods from CV and their limitations are discussed.

Traffic Detectors

Traffic data are essential for signal control. Without accurate and sufficient traffic data, one cannot time traffic signal effectively. A loop detector is the most common detector in the US. It can provide an accurate traffic count in real time. However, the downsides are cost and fix-location detection. Construction and maintenance costs for loop detectors are expensive to be deployed in a large scale. Moreover, the loop detector can only collect traffic data in only one location per unit, taking the cost into account, making it impractical to get traffic data in all necessary locations. The location of detector can affect the performance of queue estimation and needs adjustment based on traffic condition. The detector near a stop line is not affected by measurement error from platoon dispersion effect but cannot identify a long queue. On the other hand, the detector far from a stop line can detect long queue but cannot effectively detector a short queue. It is also affected by platoon dispersion. There are two kinds of techniques to

estimate queue length from loop detectors with fixed-time signal, input–output models (*11-14*) and shockwave models (*15-19*).

To solve the aforementioned problems, many alternatives methods have been considered for a long time. Connected Vehicle has been getting attentions as the alternative traffic detector recently. Traffic data are collected through probe vehicles, vehicles equipped with GPS and wireless communication devices. Location and speed of probe vehicles can be gathered to estimate travel time, speed and queue length.

However, even though the penetration ratio of probe vehicles is expected to increase in the future, the penetration ratio is still relatively small. The penetration ratio heavily affects the estimation error of traffic state. How probe vehicle collects data is also different from a loop detector. For loop detector, all vehicles are detected only at specific locations where loop detectors are installed. On the other hand, CV can collect traffic data at every location by "sampling" from all vehicles. In other word, only probe vehicles are detected. Thus, traffic state algorithms of these two kinds of detectors are not the same. The new traffic estimation algorithms need to be developed which are discussed in chapter II.

Adaptive Signal Control

The examples of adaptive signal control systems using loop detector are Split, Cycle and Offset Optimization Technique (SCOOT) (*20*), Sydney Coordinated Adaptive Traffic System (SCATS) (*21*), Optimized Policies for Adaptive Control (OPAC) (*22*), Real-time Hierarchical Optimizing Distributed Effective System (RHODES) (*23*) and InSync (*24*). RHODES and InSync find the optimal signal by decomposing optimization

algorithm into sequences. For RHODES, they are “dynamic network loading”, “Network flow control” and “intersection control”. InSync uses two level optimization, Global and Local optimization to allow continuous movement for main arterial and allocate green time to minor streets respectively. Most of them have a problem when it is in oversaturate condition since the queue length is longer than the placement of loop detectors. Thus queue detections are failed to detect the actual demand and the systems cannot find a good signal setting for such a situation. The other problem is the loop detectors cannot distinguish between a normal queue and a queue due to spill back from the downstream intersection. As a result, it is ineffective to solve such a situation.

There are a few literatures related to adaptive signal control with CV. Cai *et al* (*25*) used travel time from CV. They stated that queue length is not suitable index for adaptive signal control as it misses some information. PAMSCOD, Platoon-based arterial multi-modal signal control with online data, is proposed by He *et al* (*26*). The other studies on adaptive signal control from CVT can be seen in (*27-29*)

Queue Length Estimation

For delay minimization-based signal control, queue length is one of the most common indicators that can be directly related to delay of an approach and overall delay. Not only that, queue length is also used to be an input for signal control for queue control or minimization. It is, therefore, essential to obtain queue information. For queue length estimation by only CV or probe vehicle, there are two major methods. The first one is proposed by Cormert *et al.* (*30-33*). By using the location of the last probe vehicle in the queue and the probability function of the queue, the expectation of queue length is

then obtained. The relationship and analytical model between penetration ratio and the performance of queue length estimation are exploited. However, the practical limitation of this method is: the probability function of total queue must be obtained. The total queue is an arrival queue plus overflow queue. The arrival queue is a queue from new arrivals during a red signal. Overflow queue is a remaining queue from the previous cycle. Both vary with traffic volume, flow rate and, most importantly, traffic signal. As mentioned by the authors, traffic signal must be pre-timed and the duration of red time must be known in order to apply the algorithm. In addition, only queue length during the red time can be estimated. Finally, the algorithm has been proposed for isolated intersections with undersaturated conditions only. Equation (1) shows the queue estimation proposed by Cormert *et al.*.

$$E(N \mid L_p = l_p) = \sum_{n=l_p}^{\infty} n \frac{\left[(1-p)\right]^n P(N=n)}{\sum_{k=l_p}^{\infty} \left[(1-p)\right]^k P(N=k)}, \quad \text{for } l_p \geq 1 \qquad (1)$$

Where N is the actual total queue length

L_p is the location of the last probe vehicle in the queue

p is penetration ratio

$P(N)$ is the probability function of the total actual queue length

The second method is probe trajectory (*34-41*). By using shockwave theory (*18; 19; 42*), queue length can be estimated directly from probe trajectory. Unlike Cormert's method, this method does not require the assumption of knowing the penetration ratio. In general, knowing signal timing and arrival rate are still required. In (*36; 38*), with

queuing delay pattern model, real-time queue length estimation was obtained from traveling time data from mobile phone. To account for stochastic characteristic of traffic flow, vehicle index estimation was proposed by Hao *et al.* in (*39*). The obtained vehicle index can be used to estimate queue length. In (*35; 37*), Cheng *et al.* attempted to overcome these limitations by defining and using critical points to detect signal timing. The red signal detection using critical points and queue estimation are tested in a two-intersection network in normal and coordinated mode with simulation program and GPS from field data. However, the aforementioned researches with the second method are still limited to pre-timed signals and the performance of queue estimation heavily relies on how accurate signal timing detection is. In large and complicate adaptive signal intersections, this method has not been tested yet.

Aside from two mentioned methods for queue, the maximum likelihood method is proposed by Neumann (*43*). However, it relies on historical traffic profile so it is only applicable for offline applications. Not only relying on CV, IntelliFusion utilized both loop detectors and CV (*44*). Like the second method, it used shockwave theory to estimate queue length. The other queue length estimation research based on fused data from loop detector and CV was proposed by Li *et al.* (*45*). Shockwave-theory-based algorithm was also developed to detect queue spillback on arterial road (*46*). Cormert also extended his works from using only CV to queue length estimation using both loop detectors at the stop line and CV to improve the accuracy (*30; 31*). He *et al.* proposed queue length estimation by using linear regression model to estimate all parameters necessary for queue estimation (*26*).

To be applicable for adaptive signal control, the desired properties of a queue length algorithm should

1. provide real-time estimation,
2. does not require signal timing, traffic volume and queue characteristic as basic inputs,
3. can apply to non pre-timed signal,
4. is robust to low penetration ratio and the availability of CV and gives consistent and accurate queue estimation, and
5. can be applied to not only isolated intersections but multiple coordinated intersections.

Thus, it was the goal of this research to develop the proposed queue estimation algorithm with the above properties or attributes.

Platoon Recognition

To realize the full potential of adaptive signal control, instead of being responsive and acting on input estimation, adaptive signal control should be proactive and include traffic state prediction as inputs into their control logic. One of these predicted inputs is the arrival of a platoon, which calls for methods of platoon recognition and prediction.

The most relevant work on the topic of platoon identification is published by (*26*). He *et al.* proposed a modified critical distance-headway to identify platoon from CVT. A probability that two adjacent two CVs are in a platoon is formulated as geometric distribution with a probability of success as the penetration ratio. The modified distance-headway is equal to a mean plus three times standard deviation of the geometric

distribution. After a detected platoon is formed by modified distance-headway, three parameters of the platoon, beginning and ending times and platoon size, are estimated from CV data by using linear regression equations. While the concept is interesting, the value and impact of modified distance-headway has not yet been fully studied. While the results in this study presented the estimation error in term of root-mean-square error (RMSE), detection and misdetection rates were left unmentioned. Thus, the consistency of platoon detection was not verified yet as RMSE' results rarely reflected fully the detectability of the algorithm.

The second approach, the car-following method, which is related to estimating platoons from CVT was studied by Goodall *et al.* (*47; 48*) and Feng *et al.* (*29*). The assumption is that most, if not all, vehicles behave according to some car-following model. Then an undetected vehicle is deducted from the microscopic behavior of any nearby CV. For any pair of adjacent CV, relative speed and acceleration and distance of the pair are factors to consider for inserting a non-CV in a gap between the pair. If the vehicle fits to the car-following model, then the vehicle is inserted into a gap between two CVs. As pointed out by Feng *et al.*, one major limitation of this method is: it cannot estimate non-CVs in free-flow region where an interaction between vehicles are almost non-existing (*29*). A boundary between each region may need to be adjusted in real-time according to the traffic signal control states and traffic conditions. In particular, a boundary between slow-down and free-flow regions is difficult to justify with only CVT. Moreover, this method emphasizes on the number of actual vehicles rather than platoon resolution in time and space such as a platoon's beginning and ending time. For a signal

coordination system, detecting the beginning time of a platoon is arguably one of the most important aspects. One other platoon recognition algorithm was proposed by Smith *et al*. by using k-mean clustering, but had technical difficulty removing outliers (*27*).

In summary, there are a few studies attempting to estimate platoons from CVT, however, they all had their limitations. The detection rate of a given actual platoon size and penetration ratio is often not provided. As the threshold of platoon size differs from one application to another, such knowledge could be used to identify the limitations of CVT for platoon estimation. In addition, all of the algorithms were based on distance headway. While a distance-based platoon can be physically visualized, an arrival time has to be converted from distance and speed, and the latter is varied from one arterial to others. The objective of the research is to develop the platoon recognition algorithm and analytical model that can

1. identify platoon from time headway
2. provide real-time estimation
3. estimate platoon based on CVT as the only input
4. be applicable not only to fixed time signal control.

This dissertation also proposes a method to measure the effectiveness of the platoon estimation algorithm. The proposed algorithm can serve as a foundation for platoon prediction in the future for adaptive signal control from CVT.

CHAPTER II

QUEUE LENGTH ESTIMATION FROM CONNECTED VEHICLE*

Methodology

Unlike the loop detector, Connected Vehicle technology can track trajectories of probe vehicles so it can be used to predict long queue length without additional installation cost. The proposed algorithm(*49*) assumes distance and instantaneous speed of equipped individual vehicles can be obtained from Connected Vehicle data. A penetration ratio is known and is an input though it might be estimated from CVT. The algorithm is also developed without cooperating data from other sources such as loop detectors, signal timing and vehicle arrival assumption. In the proposed algorithm, There are four assumptions.

1. The penetration ratio is known.
2. A vehicle is always operating at a speed faster than the stopping speed when it is not in the queue.
3. The probability of a vehicle being a connected vehicle follows Bernoulli distribution.
4. The individual location and speed of a connected vehicle can be collected.

The first assumption can be justified by historical data or approximation. This also allows the accuracy of queue estimation versus penetration ratio to be evaluated later. The second assumption might not be justifiable in case of incident or illogical driver. However, our work aims to apply to signalized intersections in normal circumstances. When a vehicle arrives towards the end of a queue, a stopping speed is used to determine whether the vehicle is slow enough to be considered a part of the queue or is still a moving vehicle. The third assumption is the same one proposed by Comert and Cetin in (*32*). Every vehicle is assumed to have equal probability of being detected or being a connected vehicle. In the last assumption, it is quite common for CV to have the location information available. In contrast, speed may or may not be available or accurate depending on specific connected vehicle technology. However, our proposed algorithm does not require speed data with very high precision since it is used only to determine whether the connected vehicle is stopping or still moving. The accuracy of speed measurement is, therefore, not an issue of concern in this work.

Based on the assumptions, the queue length estimation has three steps as follows:

1. collecting the data from connected vehicles,
2. determining whether a connected vehicle is a stopping or moving,
3. estimating queue length based on two main cases: no stopped vehicle is detected and stopped vehicle is detected.

A flowchart of the proposed queue estimation algorithm is shown in Figure 1.

In the second step, individual speed is used to determine whether the detected vehicle is stopping or moving as shown in Figure 2. Following the second assumption,

stopping speed is used to determine whether an incoming vehicle is slow enough to be considered to be in a queue. It depends on the definition of the queue. If the individual speed is greater than the stopping speed, then it is a moving vehicle; otherwise, it is a stopped vehicle. With the stopped and moving vehicle data from the second step, queue estimation is divided into two cases:

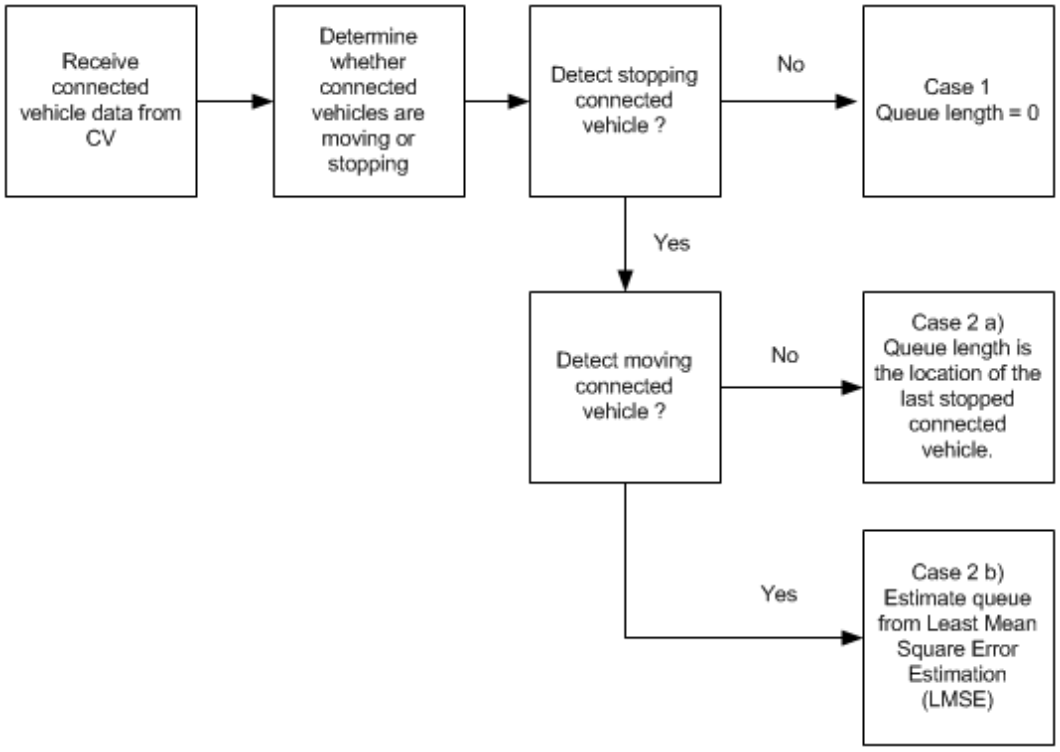

Figure 1 Flowchart of the proposed queue estimation algorithm

1. No stopped vehicle is detected at all.

In this case, the estimated queue is zero.

$$\hat{Q} = 0 \tag{2}$$

It is noteworthy that there is a chance that the queue does exist but cannot be detected since there is no connected vehicle in the queue. This problem is frequently found where the queue and penetration are small. It is problematic to the method of the last probe vehicle in the queue as well. To overcome it, other works attempt to incorporate other kinds of traffic detector such as a loop detector at the stop line (*30; 31*). However, this issue is beyond the scope of this dissertation; however, it is worthy of future work.

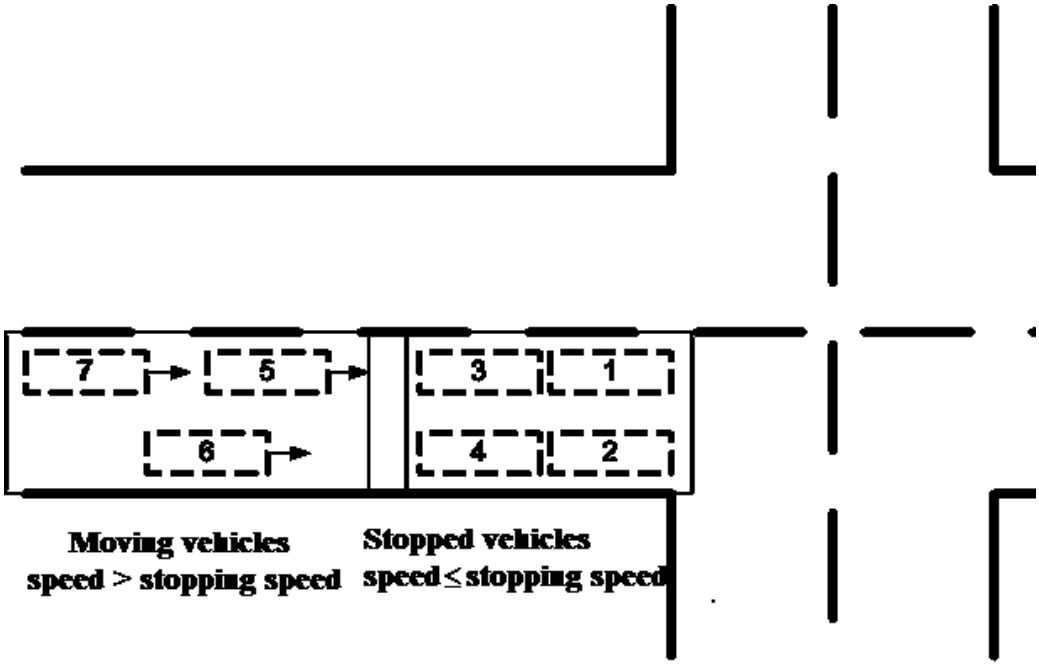

Figure 2 Illustration of moving and stopped vehicles

2. At least one stopped vehicle is detected.

   In this case, the minimum queue length can be calculated from

$$I = \max(i), v_i \le v_{stop} \tag{3}$$

$$\min(Q) = \left\lceil \frac{s_I}{l} \right\rceil \tag{4}$$

I is an index of the last or farthest stopped connected vehicle

(4) calculates the minimum queue length from the last stopped vehicle. By ordering distances and speeds of individual connected vehicle from closest to farthest from a stop line.

$s_1, s_2, s_3 ...$ are distances of connected vehicle from a stop line in the unit of vehicle length

$v_1, v_2, v_3, ...$ are speeds of connected vehicles corresponding to $s_1, s_2, s_3 ...$

v_{stop} is the stopping speed

The calculation is further divided into two subcases:

a) No moving vehicle is detected at all.

In this case, the estimated queue length equals to the minimum queue length

$$\hat{Q} = \min(Q) \tag{5}$$

b) At least one moving vehicle is detected.

In this case, the maximum queue length can be found by;

$$J = \min(j), v_j > v_{stop} \tag{6}$$

$$\max(Q) = \left\lfloor \frac{s_J}{l} \right\rfloor \tag{7}$$

$$\min(Q) \le \hat{Q} \le \max(Q) \tag{8}$$

where J is an index of the closest moving connected vehicle.

After the minimum and maximum queue length are determined, with the first and third assumptions, the calculation of estimated queue length begins with joint probability function of a total number of vehicles and connected vehicles hence

$$P(N = n, x = k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{9}$$

where $P(N = n, x = k)$ is joint probability function of total number of vehicles in the queue and connected vehicle in the queue equal to n and k, respectively;

n is a number of total vehicle in the queue, which equals to I multiplying with a number of lanes,

k is a number of connected vehicle in the queue, which equals to Q multiplying with a number of lanes, and

p is the penetration ratio of connected vehicles.

In (9), the probability function is binomial distribution due to the third assumption of probability of connected vehicle being Bernoulli distribution. Then the conditional probability function can be obtained as

$$\begin{aligned} P(N = n \mid x = k) &= \frac{P(N = n, x = k)}{\sum_{n=n_{\min}}^{n_{\max}} P(N = n, x = k)} \\ &= \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\sum_{n=n_{\min}}^{n_{\max}} \binom{n}{k} p^k (1-p)^{n-k}} \end{aligned} \tag{10}$$

The estimated total vehicles in the queue can be obtained by least-mean-square-error method

$$\begin{aligned} \hat{n}_{LMSE} &= \arg\min_{\hat{n}} \left\{ E\left[\left(\hat{n} - n \right)^2 \right] \right\} \\ &= \arg\min_{\hat{n}} \left\{ \frac{\sum_{n=n_{\min}}^{n_{\max}} \left[P(N=n, x=k) \left(\hat{n} - n \right)^2 \right]}{\sum_{n=n_{\min}}^{n_{\max}} P(N=n, x=k)} \right\} \end{aligned} \tag{11}$$

Since the denominator of (11) is constant, the estimation equation is reduced to

$$\hat{n}_{LMSE} = \arg\min_{\hat{n}} \sum_{n=n_{\min}}^{n_{\max}} \left[P(N=n, x=k) \left(\hat{n} - n \right)^2 \right] \tag{12}$$

Finally, the estimated queue length can be obtained from (13)

$$\hat{Q} = \left\lceil \frac{\hat{n}_{LMSE}}{m} \right\rceil \tag{13}$$

where m is the number of lanes

In (11)-(13), least-mean-square-error method is applied to find the estimated queue length because there is no probability function of a queue available. If the probability function is obtained, the estimated queue length can be obtained directly from (1). However, the probability function is difficult to obtain in case of non fixed-time signal control. The examples of how probability function is obtained for queue length estimation using connected vehicle data can be seen in the works done by Comert and Cetin (*30-33*).

Note that the estimated queue length in case a) tends to underestimate the actual queue length. If the maximum queue length can be implied (no blockage at the upstream intersection) and not too large then (6)-(13) can be used to find the estimated queue length instead of using (5)

Discrete Wavelet Transform and its Application in Queue Estimation Using CV

One of the major problems of CV is a low penetration ratio which leads to uncertain and inconsistent estimation. There are four major types of error when the queue exists.

1. no stopped connected vehicle,
2. have a stopped connected vehicle (s) but no moving connected vehicle,
3. have both stopped and moving connected vehicles but the distance between them is too far away,
4. have moving vehicles in a queue.

The shortcoming in the first type could be addressed by incorporating with an other kind of detector such as a loop detector at the stop line. However, as this work focuses on only CV as a detector, this kind of estimation error will not be the main focus of our work. The second type causes the proposed algorithm to underestimate the actual queue length. The third type makes overestimates. To solve the second and especially the third case, discrete wavelet transform (DWT) (*50*) can alleviate inconsistencies in such a case by assuming that the queue build or discharge is slower than the abrupt change from the second and third types. A random spike is considered to be a noise. In other words, when plotting in time series, most of the actual queue length is a low frequency component of the estimate queue length. On the other hand, most of the estimation error is a high

frequency component of the estimated queue length. DWT can be applied to a real-time application and its multi-resolution property (*51*) is useful in filtering noise at different penetration ratios which is shown in chapter IV. In the last type, some vehicles in middle of a long queue may move with "crawling speed" due to shockwave propagation. Then, it might contradict the second assumption and lead to underestimating of queue length. This issue could be solved by two methods. The first one is by setting stopping speed in the proposed algorithm to be higher than crawling speed. The second method is to use moving speed to determine if a vehicle leaves a queue. Moving speed is a minimum speed in which a vehicle is considered to leave a queue. Both stopping speed and moving speed are used to define queue definition in most simulation including VISSIM. In chapter IV, the last issue was observed in microscopic simulation, however, the proposed algorithm still estimated queue length reasonably well without calibration of stopping speed or applying moving speed. Nonetheless, an effect of stopping and moving speed to the proposed algorithm is still worthy of future work.

For denoising, DWT applies either a soft or hard threshold to obtain the thresholding wavelet coefficients. However, it needs a knowledge about the nature of the actual queue and the noise which may be affected by the type of traffic signal setting and penetration ratio. The scope is the dissertation is to be the first to demonstrate the contribution of DWT to queue estimation using CV. Therefore, a simple method, high-pass component complete elimination (detail component) was adopted in the numerical results. DWT is input with the estimated queue length obtained from the basic proposed algorithm to obtain the approximation coefficients Level 1-3 as shown in Figure 3. Then the denoised

estimated queue lengths Level 1-3 (DWT Level 1-3) are obtained by applying inverse discrete wavelet transform to the approximation coefficients Level 1-3 CA1, CA2, and CA3. Even though the chosen method is simple, the results in a chapter IV show that it can improve the accuracy and robustness of the proposed queue length estimation algorithm. To investigate the suitable strategy or threshold is beyond the scope of the dissertation but is worthy of future work. In addition, how the level of DWT and the penetration ratio affects the RMSE minimization is exploited; that is, the level of DWT, which performs the best is directly related to the penetration ratio.

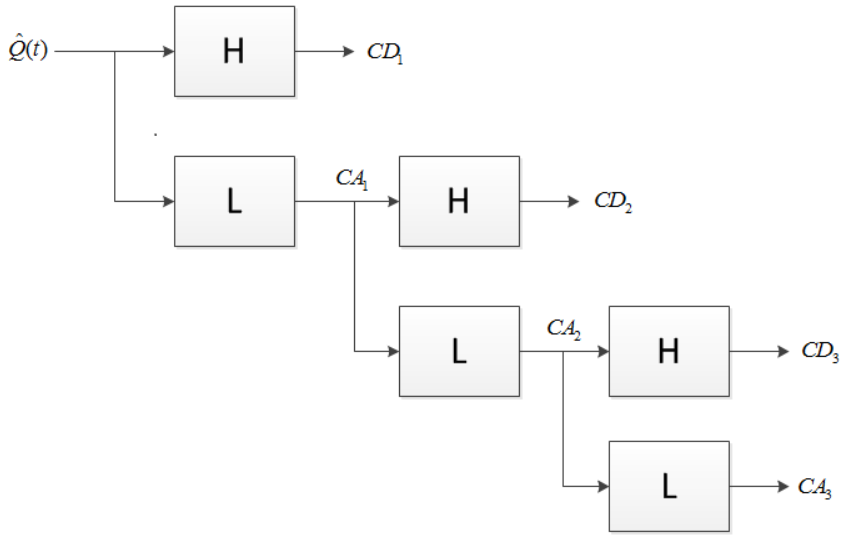

Figure 3 3-level decomposition using discrete wavelet transform in the proposed algorithm

CHAPTER III

PLATOON RECOGNITION FROM CONNECTED VEHICLE

In this chapter, the platoon is recognized by the time headways between two adjacent CVs at a chosen location on an arterial. When CV passes the location, the time stamp is collected. The differences of consecutive time stamps are the time headways between CVs, which are used to identify and then estimate the platoon parameters. The platoon recognition algorithm is divided into 4 steps of

1. platoon identification,
2. platoon size estimation,
3. platoon starting and ending time estimation, and
4. platoon filtering.

Platoon Identification

With CVT inputs alone, only CVs are observed, and non-CVs are undetectable. The lower the penetration ratio, the more missing vehicle data and the larger headway between two adjacent connected vehicles in the same platoon. Therefore, to identify a platoon using CV, a conventional platoon recognition method by critical distance or time-headway is rendered ineffective. A new algorithm is required to detect a platoon. In this research, follow (*26*), the platoon identification begins with formulating a modified critical time-headway by assuming a geometry distribution. A probability that two adjacent CVs are in a platoon is formulated as geometric distribution with the

penetration ratio as the probability of success. To avoid confusion, it is noteworthy to mention that a modified critical headway in (*26*) is a modified critical distance-headway. However, the modified critical headway proposed in our study is a modified critical headway in time domain.

Here, the modified critical time-headway with a penetration ratio of p , T_p , is treated as a calibrated parameter to better suit the application of platoon detection and traffic operation. The range of T_p is between $\frac{T_1}{p}$ and $\frac{T_1}{p}+\frac{3T_1\sqrt{1-p}}{p}$. Where T_1 denotes a critical time-headway with a 100% penetration ratio. The lower bound $\frac{T_1}{p}$ is obtained from the mean of geometric distribution with a probability of success p . The upper bound is obtained from the mean plus three times standard deviation of the geometric distribution.

He *et al.* proposed modified distance-headway to be equal to the upper bound regardless of platoon arrival, penetration ratio or traffic condition. A larger modified critical headway is better for capturing a small-sized platoon in a long term as the total number of vehicles converses to total number of CVs divided by p . On the other hand, a smaller modified critical headway is better at platoon clustering and starting and ending time identification of the platoon. The smaller modified critical headway tends to miss a small-sized platoon. Despite the flaw, some traffic operation, such as signal coordination in this case, does not give the right of way to or consider a small group of

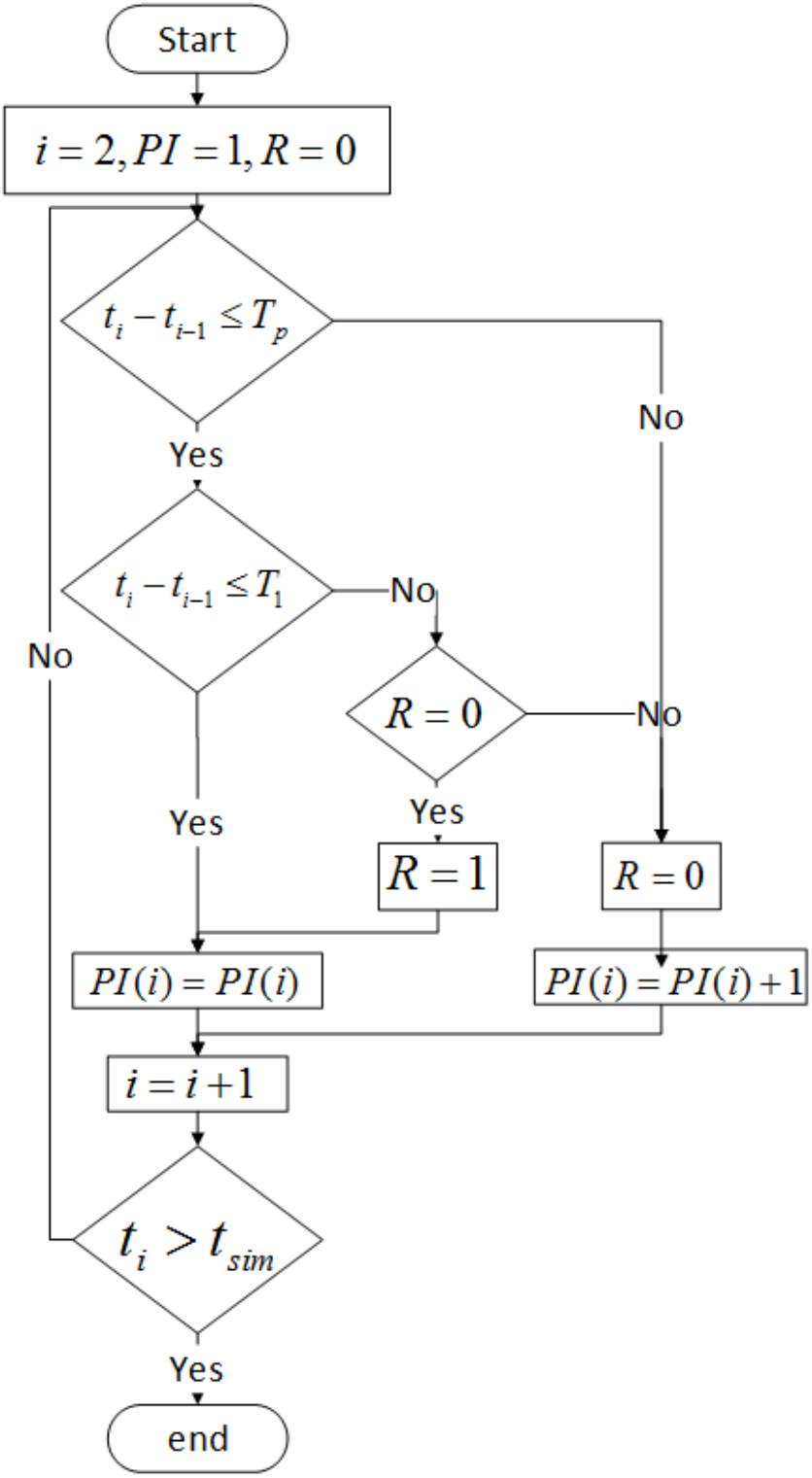

Figure 4 The flowchart of the proposed platoon identification algorithm

vehicles as a platoon, so undetectable small-sized platoon is less critical in this case. In addition, while a larger modified critical headway is more suitable for random arrival with time-independent arrival rate, a smaller modified critical headway is better for signalized platoon. Thus, T_p is set to be obtained through a calibration process.

After T_p is obtained, it is applied to detect a platoon from individual detected vehicle's time stamps. The methodology to identify a platoon is further enhanced from (*26*). Two adjacent vehicles are considered to be in the same platoon if their headway is less than T_1. If their headway is more than T_1 but less than T_p, they are in the same platoon, if and only if, there are no other CV with headway greater than T_1 before. The proposed platoon identifying algorithm is demonstrated in Figure 4 where i is CV index in chronological order, $PI(i)$ is platoon index of $i-$th CV, R is the number of times that T_P is used to form the platoon being considered.

Platoon Size Estimation

After a platoon is detected, maximum and minimum time stamps can be obtained from the last and the first CVs in the platoon as in Equation (14) and (15).

$$t_{\min} = min(t_i) \tag{14}$$

$$t_{\max} = \max(t_i) \tag{15}$$

Then the maximum and minimum platoon sizes can be calculated from Equation (16) and (17)

$$\min(n) = \max\left(\left\lceil \frac{t_{\max} - t_{\min}}{T_1} \right\rceil, x\right) \tag{16}$$

$$\max(n) = \left\lceil \frac{t_{\max} - t_{\min}}{f_1 T_1} \right\rceil \cdot m \tag{17}$$

where t_i is the time stamp of CV i in the platoon

$t_{\min}$ is the time stamp of the first CV in the platoon

$t_{\max}$ is the time stamp of the last CV in the platoon

x is the number of CVs in the platoon

n is the platoon size in vehicle unit

m is the number of lanes

f_1 is the minimum actual time-headway factor and is a calibrated parameter

$T_{\min}$ is the minimum time-headway for safety purpose.

The minimum platoon size in Equation (16) is obtained from assuming platoon vehicles in one lane with maximum time-headway, T_1. However, minimum platoon size must not be less than the number of CVs in platoon. For maximum platoon size, platoon vehicles are assumed to be as dense as possible. Thus, a minimum actual time-headway and vehicles in all lanes are considered in this case. $f_1 T_1$ is ranged between minimum time-headway for the safety purpose, $T_{\min}$, and T_1. If $f_1 = m$, then $\max(n) = \min(n)$, which is the highest possible value of f_1. As a result, the range of f_1 is $\frac{T_{\min}}{T_1} \le f_1 \le m$.

Then the estimated platoon size can be obtained by least-mean-square error method (LMSE). First, assume probability of a vehicle being detected as CV is the Bernoulli distribution. Unlike a queue, the platoon is sparse and have a density fluctuation, so a

probability of finding CV at any time is a function of the probability to have an actual vehicle at any given time and the penetration ratio. We make the second assumption that the probability to have an actual vehicle at any time between the start and ending time of platoon is time-invariant and independent from the penetration ratio. So, the probability of finding CV at any time is obtained as in Equation (18).

$$\bar{p} = f_2 \cdot p \tag{18}$$

where f_2 is a vehicle factor and is also a calibrated parameter and $\bar{p}$ is the probability to find an actual vehicle at a given location in the platoon. The range of f_2 is between 0 and 1. Then, the probability of the number of total vehicles in a platoon is derived as binomial distribution in Equation (19).

$$P(N = n, X = x) = \binom{n}{x} \bar{p}^x (1 - \bar{p})^{n-x} \tag{19}$$

A conditional probability of the total number of vehicles given the number of CVs in a platoon can be formulated as in Equation (20) and (21).

$$P(N = n \mid X = x) = \frac{P(N = n, X = x)}{\sum_{n=\min(n)}^{\max(n)} P(N = n, X = x)} \tag{20}$$

$$P(N = n \mid X = x) = \frac{\binom{n}{x} \bar{p}^x \left(1 - \bar{p}\right)^{n-x}}{\sum_{n=\min(n)}^{\max(n)} \binom{n}{x} \bar{p}^x \left(1 - \bar{p}\right)^{n-x}} \tag{21}$$

The formula for least-mean-square error calculation is shown in Equation (22). By substituting the condition probability function from Equation (21), Equation (23) is obtained

$$\hat{n}_{LMSE} = \arg\min_{\hat{n}} \sum_{n=\min(n)}^{\max(n)} \left[P\left(N = n \mid X = x\right)\left(\hat{n} - n\right)^2 \right] \tag{22}$$

$$\hat{n}_{LMSE} = \arg\min_{\hat{n}} \frac{\sum_{n=\min(n)}^{\max(n)} \left[\binom{n}{x} \bar{p}^x \left(1-\bar{p}\right)^{n-x} \left(\hat{n} - n\right)^2 \right]}{\sum_{n=\min(n)}^{\max(n)} \binom{n}{x} \bar{p}^x \left(1-\bar{p}\right)^{n-x}} \tag{23}$$

where $\hat{n}_{LMSE}$ is the estimated platoon size from LMSE. Since the denominator in Equation (23) is a constant term over $\hat{n}$, Equation (23) is reduced to Equation (24)

$$\hat{n}_{LMSE} = \arg\min_{\hat{n}} \sum_{n=\min(n)}^{\max(n)} \left[\binom{n}{x} \bar{p}^x \left(1-\bar{p}\right)^{n-x} \left(\hat{n} - n\right)^2 \right] \tag{24}$$

Platoon Starting and Ending Time Estimation

After the estimated platoon size is calculated by Equation (24), the starting time of the platoon is estimated. Due to the penetration ratio, a first CV may not be the head of the platoon. The starting time of the platoon is estimated to be between the time stamp of the first CV in the platoon and the last CV in the previous platoon. A gap between an estimated starting time and the ending time of the previous platoon should not be less than T_1, else two platoons are merged into one platoon. The main factor to consider whether to extend the starting time or not is the average platoon flow rate. The starting time estimated from the average platoon flow rate and the last time stamp in the platoon are shown in Equation (25);

$$\hat{t}_{Bg} = t_{\max} - \frac{\hat{n}_{LMSE}}{f_3} \tag{25}$$

where f_3 is the average platoon vehicle flow rate per second and $\hat{t}_{Bg}$ is the estimated starting time of the platoon. The physical meaning of f_3 is the average of total number of vehicles divided by total time span of vehicle in platoon, and is obtained by calibration. The lower and upper bounds of f_3 are $\frac{1}{T_1}$ and $\frac{m}{T_{\min}}$, respectively. Then if the estimated starting time in Equation(25) is greater than the starting time from the minimum time stamp, the later should be used instead, hence the estimated starting time from the minimum and maximum time stamps of connected vehicles and the estimated platoon size is shown as below;

$$\hat{t}_{Bg} = \min\left(t_{\min}, t_{\max} - \frac{\hat{n}_{LMSE}}{f_3} \right) \tag{26}$$

As mentioned before, the difference between the estimated starting and the ending time of the previous platoon should not be smaller than T_1. Finally, the estimation equation that takes all factors into account is shown in Equation (27);

$$\hat{t}_{Bg} = \max\left(\hat{t}_{E2} + T_1, \min\left(t_{\min}, t_{\max} - \frac{\hat{n}_{LMSE}}{f_3} \right) \right) \tag{27}$$

where $\hat{t}_{E2}$ is an estimated ending time of the last connected vehicle in the previous platoon. If the current platoon is the first platoon then $\hat{t}_{E2} = -T_1$.

After that, the ending time is obtained directly from the last CV in the platoon as in Equation (28)

$$\hat{t}_E = t_{\max} \tag{28}$$

where $\hat{t}_E =$ the estimated ending time of the platoon.

It should be noted that the starting time of the platoon is adjusted from the time stamp of the first CV in the platoon. Such an adjustment is important because a platoon that arrives earlier than predicted would likely have to stop and incur significant delay, hence deteriorating the quality of coordination. The ending time of the platoon on the other hand is not adjusted from the time stamp of the last CV in the platoon here, because the impact on coordination is not significant in general. Adjustments of both the starting and ending times at the same time is possible, but we choose to only adjust the starting time in this research as we focus on continuous movement of the arriving platoon.

Platoon Filtering

This part is optional for platoon recognition. In coordinated signal systems, signal control is set to give the right of way to the through movement on the major arterial when there is an arrival of a platoon with the number of vehicles more than a threshold. There is no clear cut number for the threshold and it depends on the application. This process of identifying a ‘’qualified’’ platoon is referred as platoon filtering in the dissertation. A platoon with lower number than a threshold is ignored.

Due to the uncertainty in the platoon size estimation, the threshold from the case with 100% penetration ratio cannot apply directly to the other cases with lower penetration ratio and needs to be modified. The modified threshold needs to be adjusted to maximize the detection rate while minimizing the false detection rate. The filtering

process relates the threshold in the case with actual vehicle or with 100% penetration ratio, y_1 to its counterpart, the modified threshold with the penetration ratio p, y_p. y_p is used to determine whether the detected platoon size is large enough to be considered a platoon. A detected platoon size which is smaller than $\hat{y}$ would not trigger a signal coordination and is filtered out from the platoon list.

y_p is set to be a calibration parameter with its lower bound corresponding to the penetration ratio. The upper bound is set to be y_1 regardless of the penetration ratio. However, when the penetration ratio is equal to or greater than 90%, a lower bound is set to be a constant with $0.8 \cdot y_1$ so the bound is not too tight. The benefit of setting the lower bound of y_p depending on the penetration ratio is to prevent the calibration parameters to fit so well with the training data that they cannot be used to estimate platoon with different platoon profiles or under traffic signal control settings. In summary, the lower and upper bounds of y_p are $y_1 \cdot \min(0.8, p)$ and y_1, respectively. The numerical results in the chapter V attempt to validate the performance of applying calibrated parameter values to other test cases without recalibrating.

Platoon Detection Rate Analysis

In this subchapter, the platoon detection rate is estimated by an analytical approach. First, we begin with a simplification that T_p is assumed to be large enough that a platoon of any size can be detected with only two CVs. Generally, to ensure that two CVs in the platoon are sufficient enough for platoon detection, T_p must be larger than the time-stamp difference between the first and the large vehicles in the platoon. For a large-sized

platoon, because of its large time span, it is not 100% possible as it depends on a number of lanes and T_p. However, with a proper calibration, T_p is assumed to be large enough to detect any platoon with at least two connected vehicles with a probability close to 100%. In addition, a chance of only two connected vehicles found in a large-sized platoon is significantly lower than a chance of finding more than two connected vehicles so such a case can be approximated as mentioned. The next assumption is the actual and detected platoons are matched one to one. In other word, we assume there is no merging or diverging issues, particularly under a low penetration ratio situation.

As the probability of a vehicle being CV follows a Bernoulli distribution, the conditional probability of having x CVs given a platoon size n is;

$$P(X = x \mid N = n) = \binom{n}{x} p^x (1-p)^{n-x} \tag{29}$$

and since the probability of having CVs is independent from a threshold of interest when $n \geq y_1$,

$$P(X = x \mid N = n, Y = \{y\_1\}) = P(X = x \mid N = n) \tag{30}$$

The probability of platoon detected given a platoon size n and the threshold y_1 is equal to a summation of probability of having connected vehicle equals or greater than 2 as below;

$$\begin{aligned} &\sum_{x=2}^{n} P(X = x \mid N = n, Y = y_1) = 1 - P(X = 0 \mid N = n) - P(X = 1 \mid N = n) \\ &= 1 - (1-p)^n - np(1-p)^{n-1} \end{aligned} \tag{31}$$

A marginal probability of platoon detected given the threshold y_1 is obtained by summation over platoon size n from y_1 to infinity as follow;

$$P(X \geq 2) = \frac{\sum_{n=y_1}^{\infty} \left(P(N=n) \sum_{x=2}^{n} P(X=x \mid N=n, Y=y_1) \right)}{\sum_{n=y_1}^{\infty} P(N=n)} \tag{32}$$

When a platoon profile is obtained, a detection rate can be estimated from Equation (32). However, the probability function of platoon size $P(N=n)$ may not be available. Hence, a lower bound of detection rate is derived further.

Since $\sum_{x=2}^{b} P(X=x \mid N=b, Y=y_1) \geq \sum_{x=2}^{a} P(X=x \mid N=a, Y=y_1)$ for all a, b, y_1 that $2 \leq y_1 \leq a \leq b$, Equation (33) is obtained from Equation (32)

$$\begin{aligned} P(X \geq 2) &\geq \sum_{x=2}^{y_1} P(X=x \mid N=y_1, Y=y_1) \\ P(X \geq 2) &\geq 1-(1-p)^{y_1} - y_1 p (1-p)^{y_1-1} \end{aligned} \tag{33}$$

Equation (33) is applied to find a lower bound of the probability to detect a platoon when the platoon profile is unavailable. On the other hand, if the platoon profile is obtained, Equation (32) can be applied to get a more accurate estimation.

Figure 5 shows the detection rates obtained from the lower bound calculated by Equation (33). The implication from the figure is, for instance, if the penetration ratio and the minimum required platoon size is 30% and 10 vehicles, a platoon detection rate is no less than 80%. The figure also implies that; the smaller the platoon size, the less chance to detect, due to a smaller sampling size.

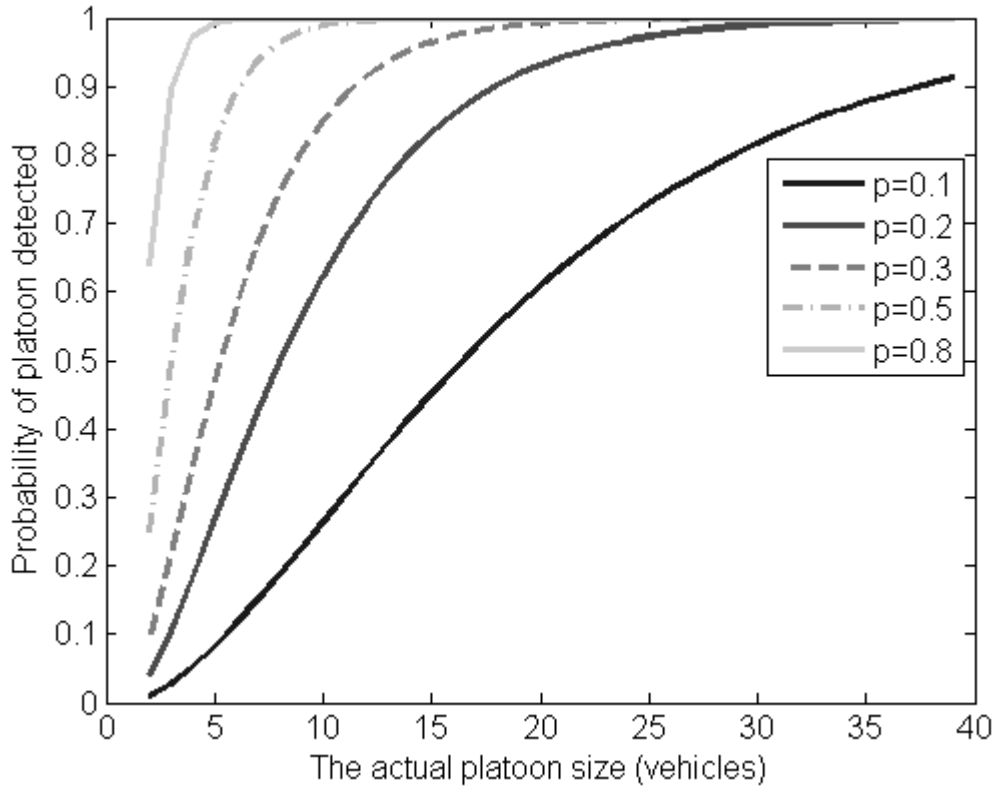

Figure 5 The estimated detection rate from analytical approach 2

Parameter Calibration

There are 5 unknown parameters that need to be calibrated before the proposed algorithm can accurately estimate a platoon. First, we define an objective function as a function of weight and error vectors as follows:

$$fitness = \sum_{i=1}^{I} \vec{W}^T \vec{Z}_i \tag{34}$$

where $\vec{W}$ and $\vec{Z}_i$ are weight and error vectors, respectively. The error vector contains five error measurements. $RMSE_{Size}(i), RMSE_{Bg}(i), RMSE_E(i), DR(i)$ and $MR(i)$ are root-

mean-square errors (RMSE) of platoon size, beginning and ending times, and detection and misdetection rates after a filtering process of iteration i, respectively. Weight vector is defined as a vector of weights to each error measurement. The higher the weight value, the more focus on minimizing the respective error measurement. I is the total number of iterations. For each iteration, a vehicle is randomly selected to be CV with a probability equals to p. RMSE calculation, therefore, computes with an actual and detected platoon after the filtering.

To compute RMSEs, one needs to identify the pair of estimated and actual platoon data. However, the proposed algorithm forms a different set of platoons depends on the penetration ratio and randomness in each iteration. Actual platoons can be either or both overlapped, merged or diverged to one or more detected platoons. As a result, a detected platoon cannot straightforwardly match one-to-one with an actual platoon. Before computing RMSEs, a matching matrix between actual and detected platoons has to be created.

A matching matrix is defined as $M \times N$ matrix, $A = [a_{ij}]$, whose members is either 1 or 0, where M and N are a number of detected and actual platoons respectively. $a_{ij} = 1$ if a detected platoon i is overlapped with an actual platoon j. The calculation of a matching matrix is shown in Equation (35).

$$a_{ij} = \begin{cases} 1 & \text{, if } \hat{t}_{Bg}(i) < t_E(j) \text{ and } \hat{t}_E(i) > t_{Bg}(j) \\ 0 & \text{, otherwise} \end{cases} \tag{35}$$

where $t_{Bg}(j)$ and $t_E(j)$ are beginning and ending times of the actual platoon j.

There are two issues after the matching matrix is found: merging and diverging of detected platoons. First, we solve merging issue, then solve the diverging issue. An merging issue occurs when there are multiple actual platoons overlapped with a detected platoon. A merged platoon can be observed from the matching matrix by the sum of the row, $\sum_{j=1}^{n} a_{ij}$. For any $i-$th row with $\sum_{j=1}^{n} a_{ij} > 1,$ $i-$th detected platoon has a merging issue. To solve this problem, the detected platoon is demerged to smaller detected platoons with the number equals to the number of associated actual platoons. The beginning and ending time of the new broken-down platoons are determined by a breaking point. A breaking point is defined as a time in the middle between two mentioned actual platoons. Mathematically speaking, breaking point is the average of the ending time of the preceded actual platoon and beginning of the following actual platoon. The size of each broken-down detected platoon is proportional to its time span.

After all merged detected platoons are broken down to one-to-one match with the actual platoons, the next issue is the diverged detected platoon. A diverged platoon is defined as a single actual platoon being paired to multiple detected platoons. It can be observed by the sum of the column, $\sum_{i=1}^{m} a_{ij}$. For any $j-$th column with $\sum_{i=1}^{m} a_{ij} > 1,$ $j-$th actual platoon is determined to be a diverging platoon. In this case, the detected platoons overlapped with the respective actual platoon are combined into one detected platoons for a comparison. Without loss of generality, let diverged detected platoon index is a member of set I . The platoon size, beginning and ending time for the merged detected platoon for matching with the actual platoon j are obtained as follow;

$$\tilde{N}_j = \sum_i \hat{N}_i, \forall i \in I \tag{36}$$

$$\tilde{t}_{Bg,j} = \min(\hat{t}_{Bg,i}), \forall i \in I \tag{37}$$

$$\tilde{t}_{E,j} = \max(\hat{t}_{E,i}), \forall i \in I \tag{38}$$

To give an example of how demerging and remerging rematching actual and detected platoon one-to-one, Figure 6 demonstrates two actual and three detected platoons which are overlapped with each other and have both merging and diverging issues. As shown in the figure, $t_1,...,t_{10}$ are beginning and ending times of actual and detected platoons. $t_1,...,t_4$ and $t_5,...,t_{10}$ are in chronological order. $s_1,...,s_5$ are sizes of actual platoons 1 and 2, and initial detected platoons 3-5. First, demerging detected platoon 3 into two smaller platoons according to the middle time between actual platoons 1 and 2. Then, remerging detected platoon 1,2 and 6 into a new platoon which can match one-to-one with the actual platoon 1. The beginning and ending times, and platoon sizes are shown in Figure 6

After transforming with merged and diverged platoons, RMSE can be computed one-to-one for any pair ij . Unpaired actual and detected platoon are excluded from RMSE calculation, but are taken into account for detection and misdetection rate, respectively. Let $k(u)$ and $q(u)$ be u -th pair index of actual platoon and its matched detected platoon. RMSEs, detection and misdetection rates for one-to-one pair are shown in the following equations;

$$RMSE_{Size} = \sum_{u \in U} \left(\hat{N}_{q(u)} - N_{k(u)} \right)^2 \tag{39}$$

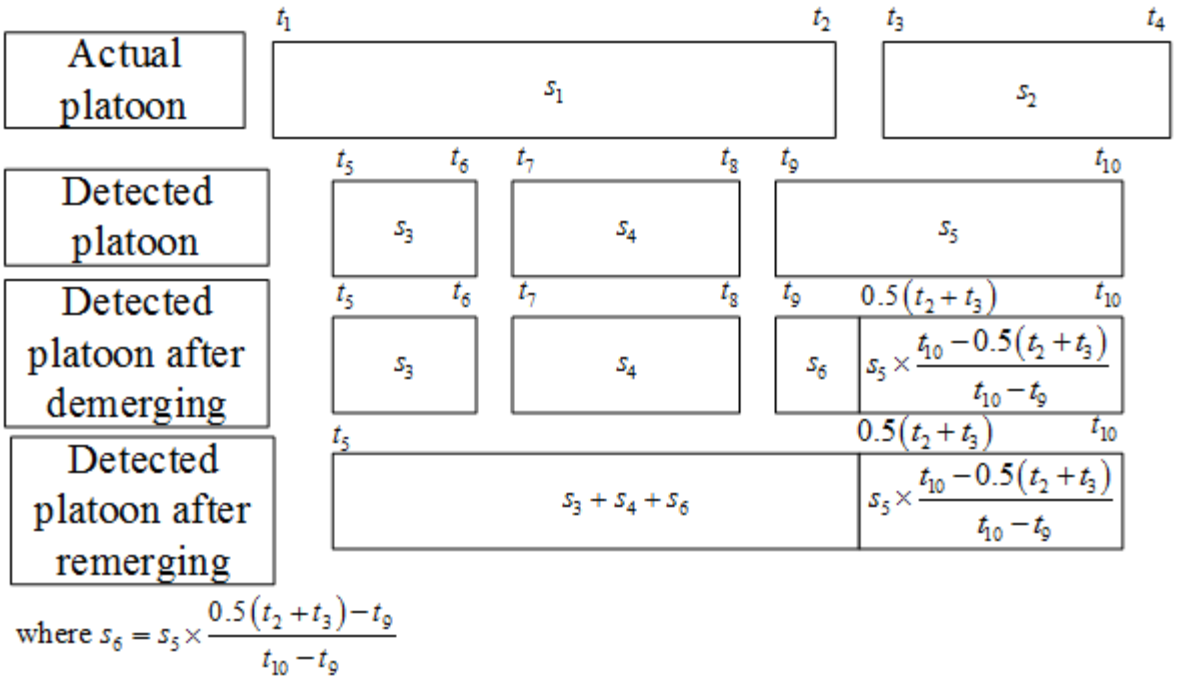

Figure 6 The example of rematheing detected platoons into one-to-one

$$RMSE_{Bg} = \sum_{u \in U} \left(\hat{t}_{Bg,q(u)} - t_{Bg,k(u)} \right)^2 \tag{40}$$

$$RMSE_{E} = \sum_{u \in U} \left(\hat{t}_{E,q(u)} - t_{E,k(u)} \right)^2 \tag{41}$$

$$DR = \frac{number\left(\sum_{i=1}^{m} a_{ij} > 0 \right)}{n} \times 100 \tag{42}$$

$$MR = \frac{number\left(\sum_{j=1}^{n} a_{ij} = 0 \right)}{m} \times 100 \tag{43}$$

Here, 5 performance indicators are obtained through the process mentioned above. RMSEs of platoon size, beginning and ending time indicate the accuracy of platoon estimation in case an actual platoon is being detected. When the actual platoon is

undetected, RMSE is not computed, but the case is considered in computing the detection rate. The more actual platoons are undetected, the lower the detection rate. Equation (42) shows the detection rate is equal to the number of columns with the sum of the row greater than 0 divided by the number of total columns. Misdetection rate is an indicator how often the proposed algorithm falsely detects a platoon. It is defined as the number of detected platoons that is not overlapped with any actual platoon at all divided by the total number of detected platoons. Equation (43) shows misdetection rate is equal to the number of rows with the sum of the column greater than 0 divided by the total number of rows.

CHAPTER IV

NUMERICAL RESULTS FOR QUEUE LENGTH ESTIMATION*

To test queue length estimation from Connected Vehicle data, microscopic simulation, VISSIM 5.40 (*10*), was selected as a test bed to generate individual vehicle's information and observe queue length. For the proposed algorithm, the basic inputs extracted from VISSIM are individual vehicle data, which are instantaneous speed and distance from the approaching intersection.

The two-second-average queue length data from VISSIM are collected to compare with the average queue length estimated from the proposed algorithm. Individual vehicle data are collected every 0.5 second. The estimated queue length is calculated when the next-step individual data are available, which is every 0.5 second. However, the average queue length data from VISSIM is set to be obtained every 2 seconds. As a result, the 2-second average estimated queue length is computed to compare with the 2-second average queue length from VISSIM. Stopping speed was set to 0 km/h in the calculation of the proposed algorithm.

To demonstrate a penetration ratio, a part of individual data was randomly removed from queue length estimation corresponding to the penetration ratio. For example, if the ratio were 30%, there was a 30% chance that individual vehicle data were used in the estimation algorithms; that is, the greater the penetration ratio, the less missing data.

To account for the effect of randomness, the availability of vehicle data was randomized 100 times; then the root-mean-square error from 100 randomizations was averaged as in the following equations:

$$RMSE(i) = \sqrt{\frac{\sum_{t=1}^{T} \left(q(t) - \hat{q}_i(t) \right)^2}{T}} \tag{44}$$

$$AVG\_RMSE = \frac{\sum_{i=1}^{I} RMSE(i)}{I} \tag{45}$$

where $q(t)$ is the actual average queue length from VISSIM at time step t;

$\hat{q}_i(t)$ is the estimated average queue length from the proposed algorithm at time step t and i-th randomization;

T is total number of time steps being computed

$RMSE(i)$ is root-mean-square error of i-th randomization;

AVG_RMSE is average root mean square error of total randomization, and

I is total number of randomization being computed.

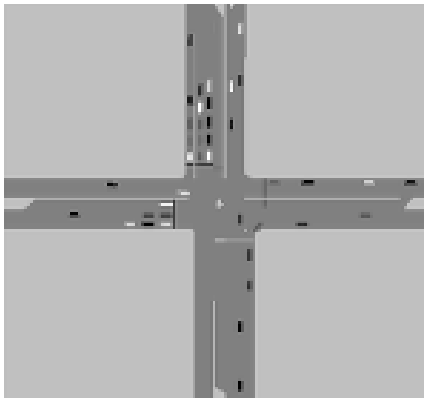

Figure 7 An isolated intersection used in the experiment

The test network is an isolated intersection as shown in Figure 7. East and west bound approaches have one left-turn lane and two through lanes. The right lane is shared between through and right-turn traffic. For north and south bound approaches, there is one left-turn lane, two through lanes and one right-turn lanes. The left-turn for all approaches is protected. Simulation time is 2,000 seconds. Each approach is a 270 meter-long road. The volume for each approach and the signal control are varied from case to case. The experiments are divided into 4 cases: pre-timed signal and actuated signal control under undersaturated and saturated conditions. For an undersaturated condition, the Poisson arrival rate assumption is valid. On the other hand, the Poisson assumption is no longer valid when traffic volume increases to saturated condition. The objective of having two traffic conditions is to compare the performance of both algorithms where the Poisson assumption is valid and invalid. The actuated control is included in the experiments to verify the applicability range of the proposed algorithm with non-fixed signal settings. In both a pre-timed signal control situation and actuated cases, the comparison between queue length from the proposed algorithm and VISSIM and RMSE are given.

Pre-timed Signal Control with Undersaturated Condition

In this case, the signal control is set to be pre-timed with a 100-second cycle length. The green time for all left turn and through movement is 15 and 35 seconds, respectively. The volume for all approaches for undersaturated condition is 1,000 vehicle per hour (vph). Figure 8- Figure 10 shows the results of queue length estimation for a penetration ratio equaling 10%, 50% and 80%.

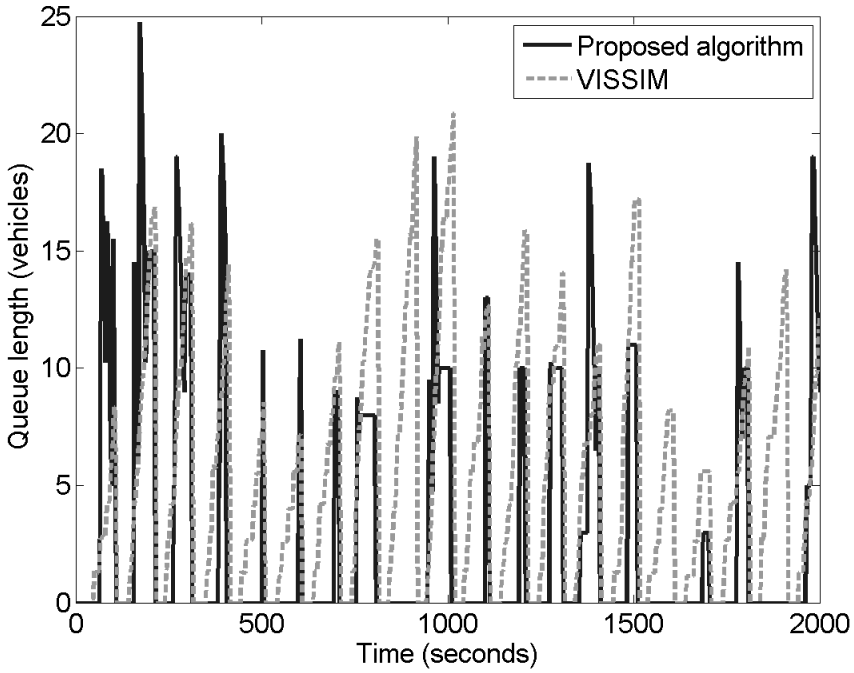

Figure 8 Case A: The estimated queue length when penetration ratio = 10%

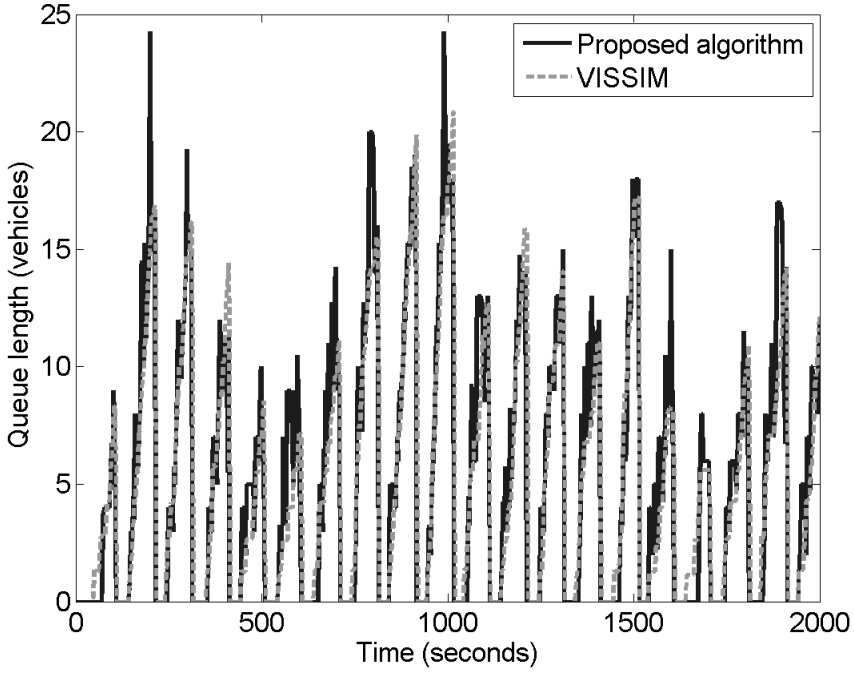

Figure 9 Case A: The estimated queue length when penetration ratio = 50%

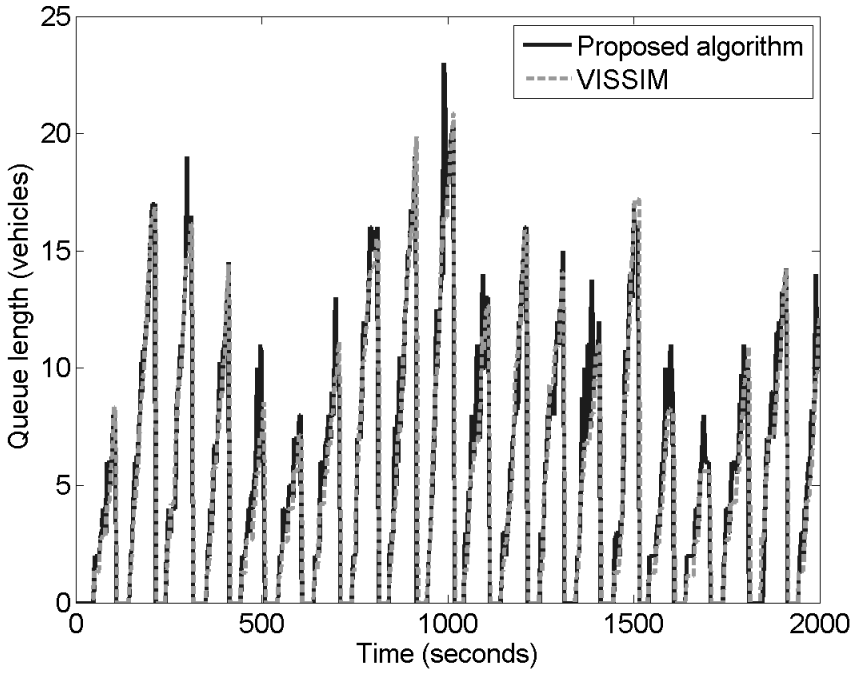

Figure 10 Case A: The estimated queue length when penetration ratio = 80%

From Figure 8, the maximum of 2-second average queue length and AVG_RMSE are 20.87 and 5.7 vehicles. Thus, the proposed algorithm is able to estimate the actual queue length, with RMSE less than 30% of maximum queue length. The RMSE of Case A-D can be seen in the last subchapter. When a detected moving vehicle is far away from the last stopped vehicle, peaks or random high spikes occur which cause high error and inconsistency in queue estimation. This problem is improved in the later subchapter with DWT. It is noteworthy that the proposed algorithm is unable to track the queue when there is no connected vehicle in queue, in particular, at the beginning of queue forming up. As a result, the proposed algorithm cannot detect at the beginning of the queue at some cycles, for instance, at time 900 and 1,600 seconds. However, such a

problem is less likely when the penetration ratio increases as shown in Figure 9 and Figure 10. The other solution is to have other kinds of detectors at the stop line to detect a small queue; however, it is out of the scope of the dissertation.

Pre-timed Signal Control with Saturated Condition

Everything is similar to Case A but with volume 1,400 vph for EB and WB approaches. NB and SB remain 1,000 vph as in case A. Assuming a 1900 vphpl saturation flow rate, the degree of saturation for the through movement is 1.0. Figure 11-Figure 13 show the results of queue length estimation for a penetration ratio equal to 10, 30 and 80%, respectively.

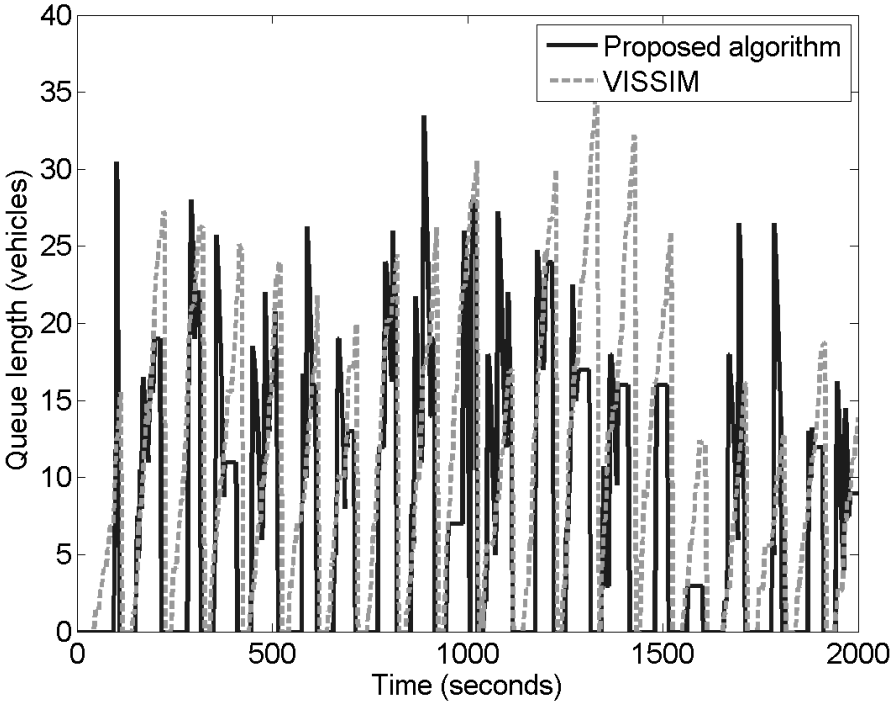

Figure 11 Case B: The estimated queue length when penetration ratio = 10%

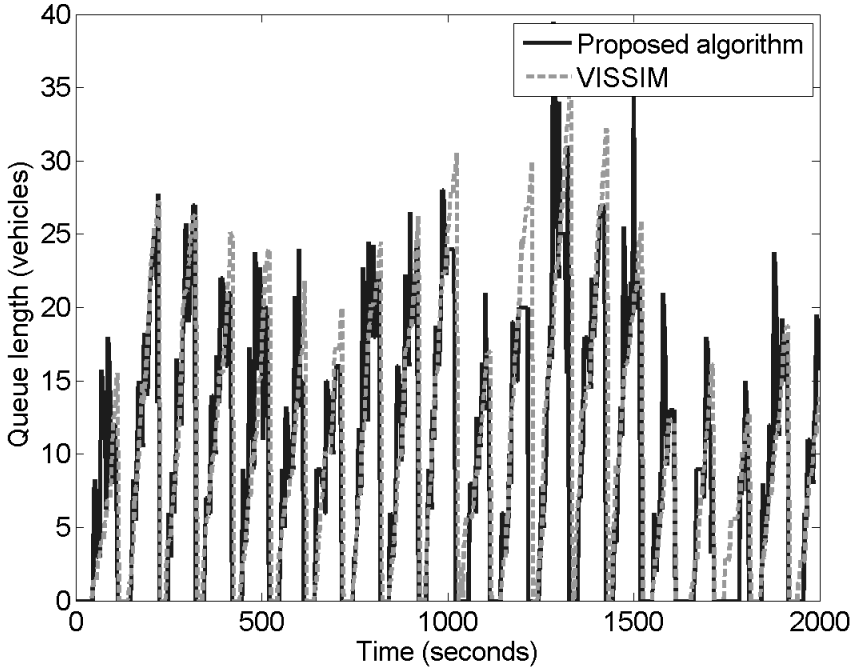

Figure 12 Case B: The estimated queue length when penetration ratio = 30%

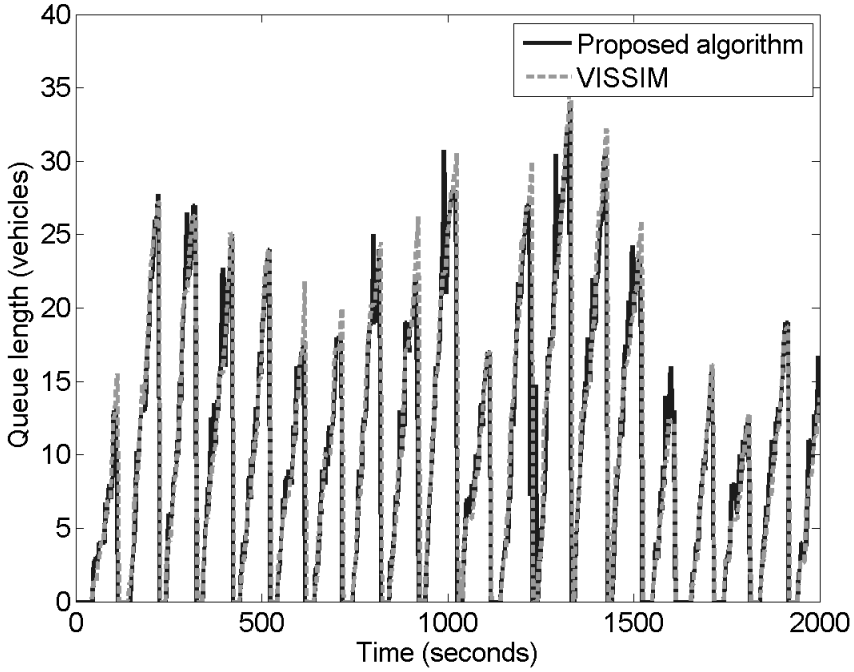

Figure 13 Case B: The estimated queue length when penetration ratio = 80%

With increasing volume from Case A, the Poisson arrival assumption is weaker; nonetheless, the proposed algorithm still reasonably estimates the actual queue length. It implies that the proposed algorithm is capable of estimating queue with non-Poisson arrivals and without arrival and overflow queue data as basic inputs. This property is crucial for multiple intersections and coordinate systems as volume may change from undersaturated to a congested condition, and vehicles may not arrive in Poisson distribution but in platoons. Like in Case A, there are random spikes and misdetection of queue estimation. However, since there are more vehicles and, therefore, more connected vehicles in the network, random spikes appearing in Case B is less compared to Case A.

Actuated Signal with Undersaturated Condition

The traffic signal is an actuated control in this case. Actuated control is used to demonstrate how the proposed algorithm can work with time-varying traffic signal without signal timing information. The volume is the same as in Case A. Actuated signal control has min and max green time 6 and 26 seconds for left turn and 10 and 36 seconds for through movements. All phrases have 2 seconds of vehicle extension. Figure 14-Figure 16 show the results of queue length estimation for penetration ratios equaling to 10, 30, and 50% respectively.

Since the signal timing becomes time-varying, the cycle length and green time are not constant. The trend of maximum queue length in each cycle becomes less monotonic than in Case A. Overall, the actual queue length in this case is shorter than in Case A

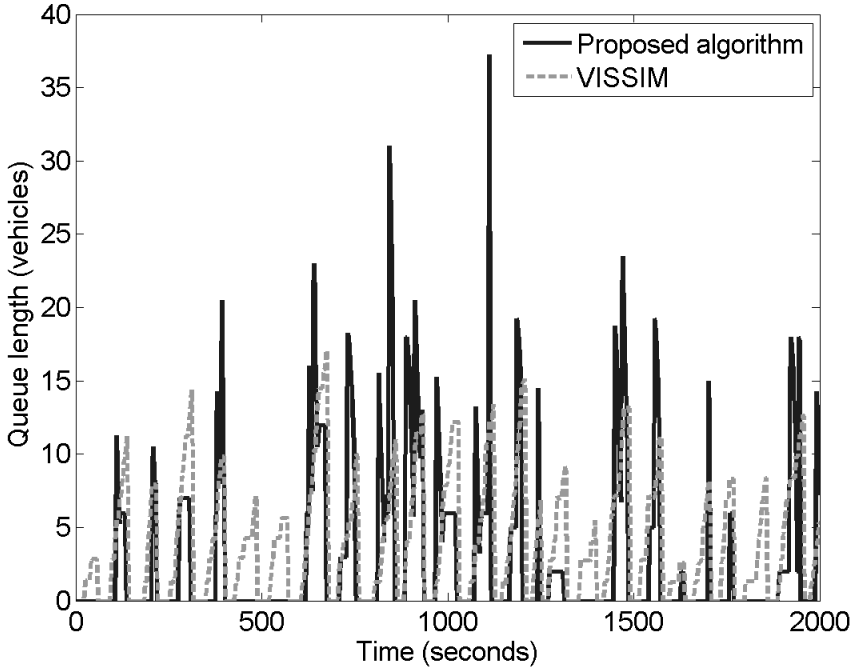

Figure 14 Case C: The estimated queue length when penetration ratio = 10%

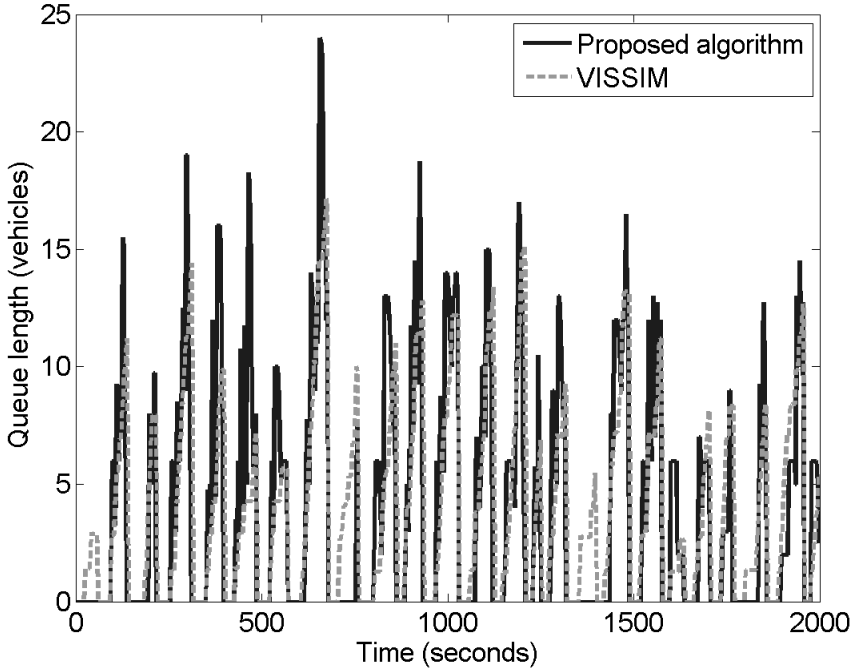

Figure 15 Case C: The estimated queue length when penetration ratio = 30%

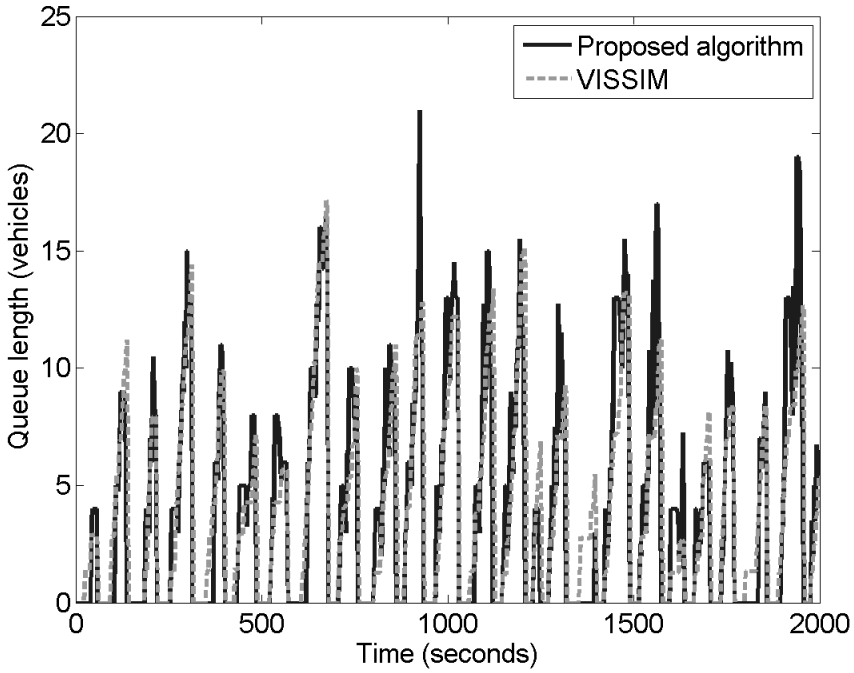

Figure 16 Case C: The estimated queue length when penetration ratio = 50%

since actuated control can adjust green time according to the presence of vehicles. The accuracy of estimated queue length, measured by AVG_RMSE, is about the same as in Case A. For penetration ratio = 10%, the maximum of 2-second average queue length and AVG_RMSE are 17.13 and 4.8 vehicles. Even though signal timing and duration of red stop time are not used as basic input in the proposed algorithm, the estimated queue length is able to estimate queue length with RMSE less than 30% of maximum queue length and is about the same as in Case A. The only remaining problem is inconsistency from random spikes.

Actuated Signal with Saturated Condition

In this case, traffic has an actuated signal of volume 1,400 vph. It is also approximately at the point of saturation. This was confirmed in VISSIM animation. Figure 17-Figure 19 shows the results of queue length estimation for penetration ratios of 30, 50, and 80%, respective

In this case, actuated control can manage traffic better than pre-timed signaling in Case B so the queues are generally shorter. Like in Case B, the misdetection of estimated queue length is almost nonexistence even with a low penetration ratio. It is a result of more vehicles in the road network. The performance of the proposed algorithm in case B and D are not much different.

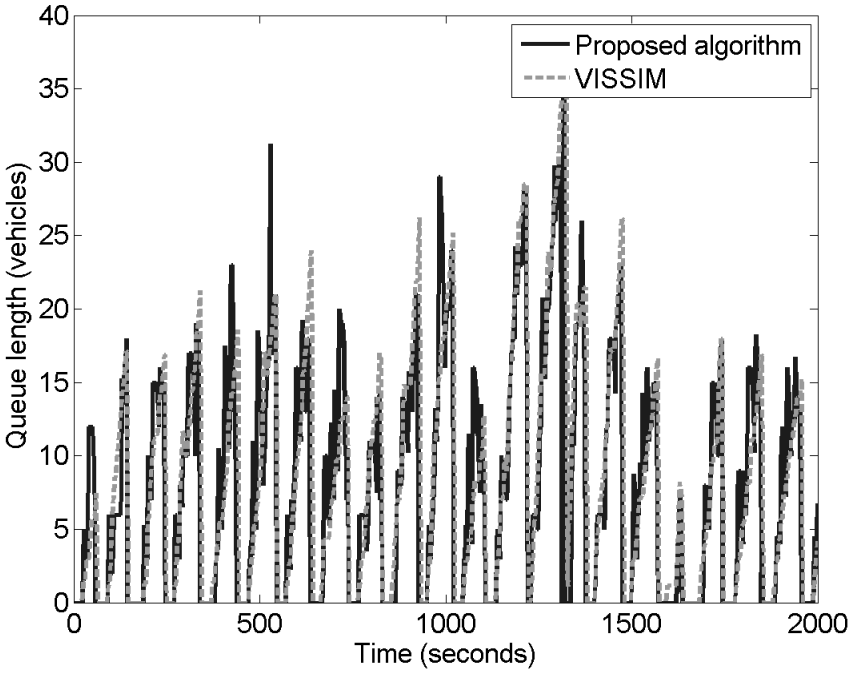

Figure 17 Case D: The estimated queue length when penetration ratio = 30%

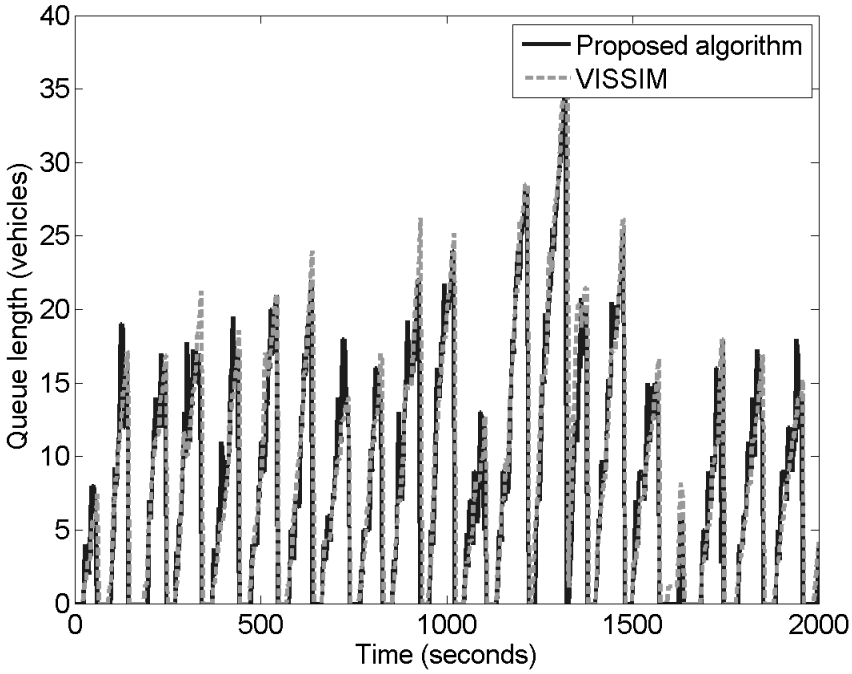

Figure 18 Case D: The estimated queue length when penetration ratio = 50%

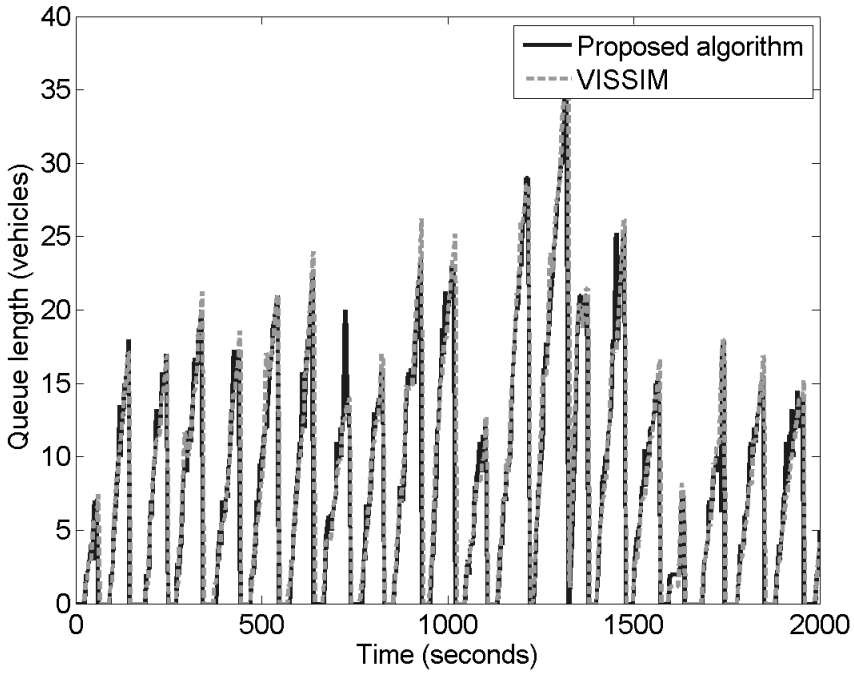

Figure 19 Case D: The estimated queue length when penetration ratio = 80%

To further verify the impact of stopping speed, extensive analysis and results on different threshold values for multiple scenarios of signal control and volume levels, such as 5km/h were carried on. The RMSE of the predicted queue length does not change much with the increase of the threshold value. As a result, the stopping speed does not change the overall performance of the prediction accuracy.

Enhanced Queue Estimation with Discrete Wavelet Transform

The numerical results from Case A to D in the previous subchapter show that the proposed algorithm can estimate queue length quite well. The only common problem is the fluctuation or random spikes which may cause instability to adaptive signal control when the queue is overestimated. This subchapter demonstrates the numerical results after applying discrete wavelet transform to all four cases.

Figure 20-Figure 23 show the root-mean-square error of estimated queue length before and after applying Haar DWT Level 1-3 to the proposed algorithm.

In all four cases, the RMSE decreases as the penetration ratio increases. The RMSE is also proportional to the total average queue length. In actuated control cases, the total average queue length is lower than the pre-timed control with the same traffic condition. As a result, Case B has the highest RMSE and Case C has the lowest RMSE among all cases. From RMSE plots, DWT can reduce RMSE, but the performance depends on penetration ratio. At the low penetration ratio, the proposed algorithm contains more error so DWT Level 3 can perform the best by filtering out most of high frequency component. On the other hand, DWT Level 1 can filter noise less than

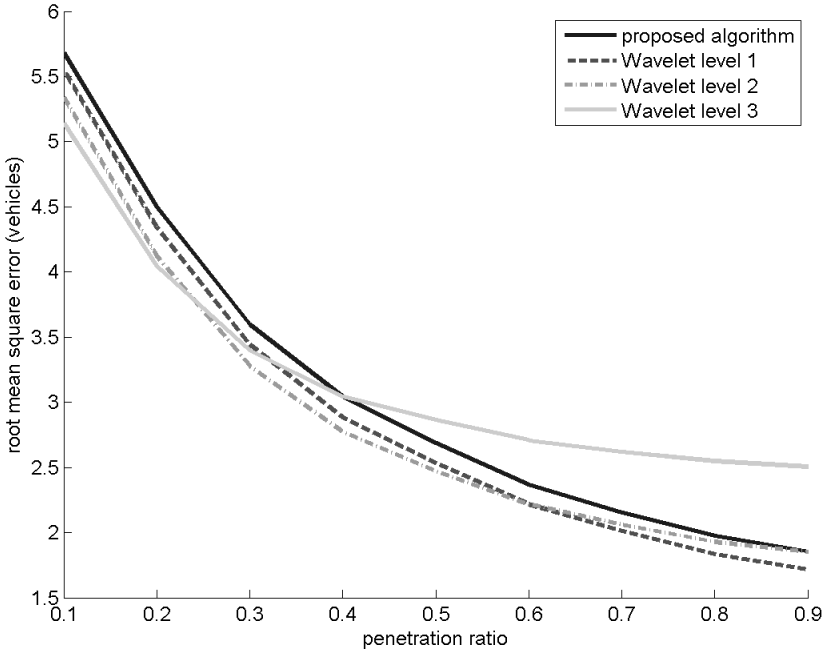

Figure 20 Root-mean-square error of queue length estimation in Case A

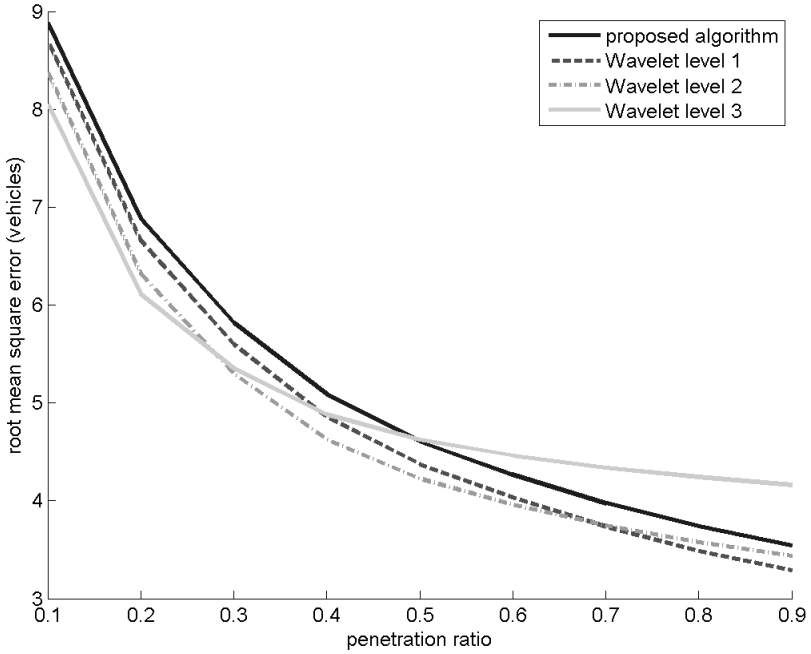

Figure 21 Root-mean-square error of queue length estimation in Case B

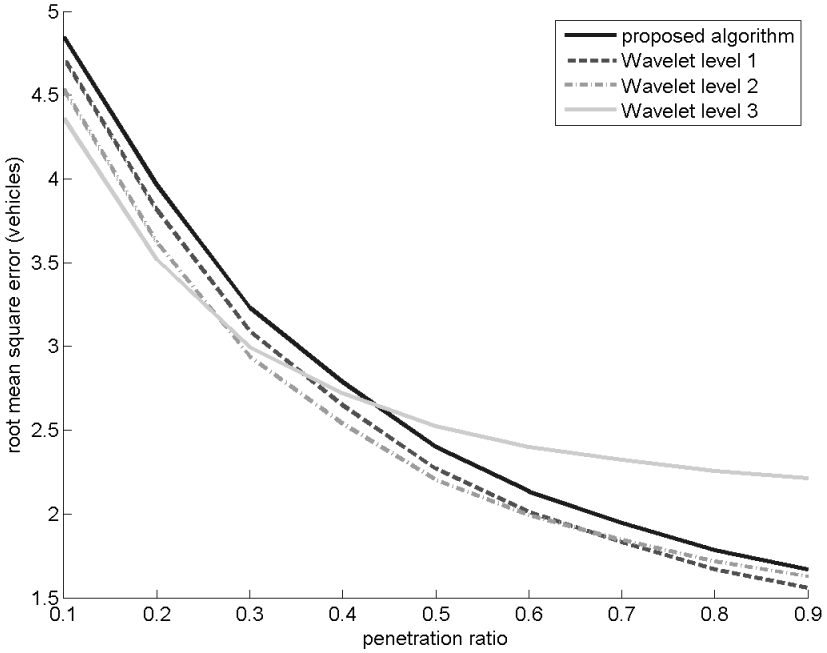

Figure 22 Root-mean-square error of queue length estimation in Case C

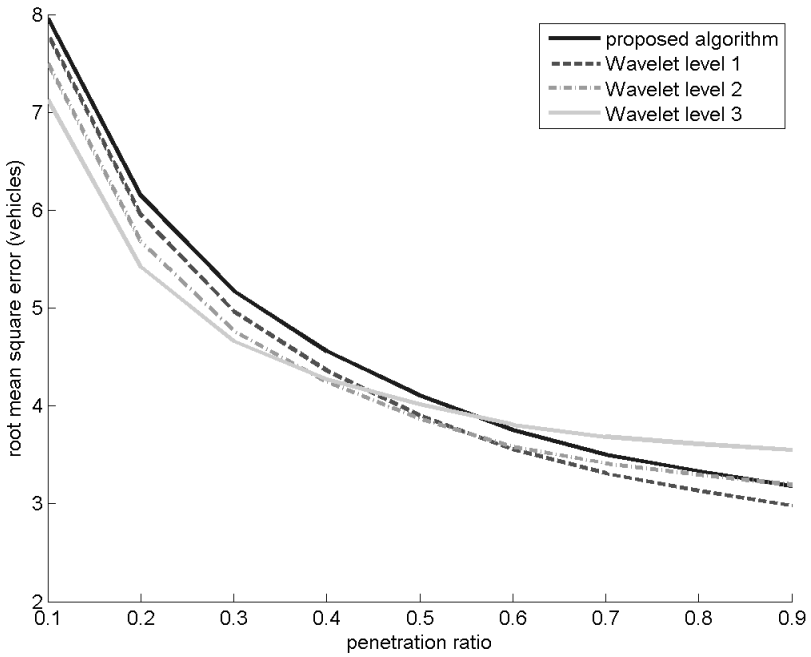

Figure 23 Root-mean-square error of queue length estimation in Case D

DWT Level 3 so the error reduction is less. At high penetration ratios, high-frequency component of the proposed algorithm contains less noise or error. Moreover, most of the high-frequency components are dynamic part of the queue length. Filtering out high-frequency components disables the estimated queue length to capture the dynamics of the actual queue length. Consequently, DWT Level 3 increases the RMSE of the estimation of queue length. On the other hand, DWT Level 1 can retain the dynamic component better than DWT Level 1 while eliminating part of errors so DWT Level 1 perform the best at high penetration ratio. For all ranges of penetration, DWT level 2 consistently reduce RMSE the best regardless of penetration ratio.

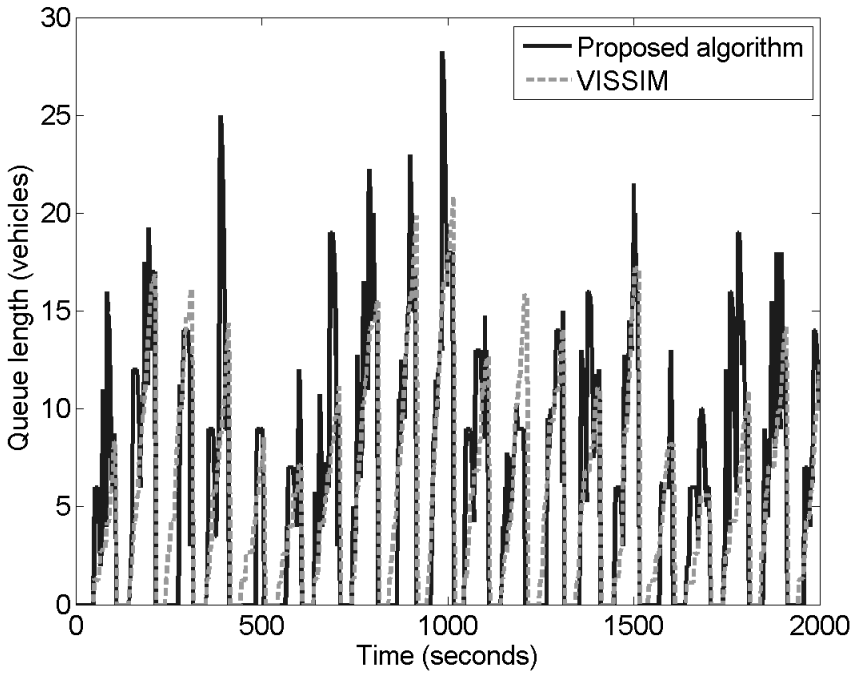

Figure 24 Case A: The estimated queue length when penetration ratio = 30%

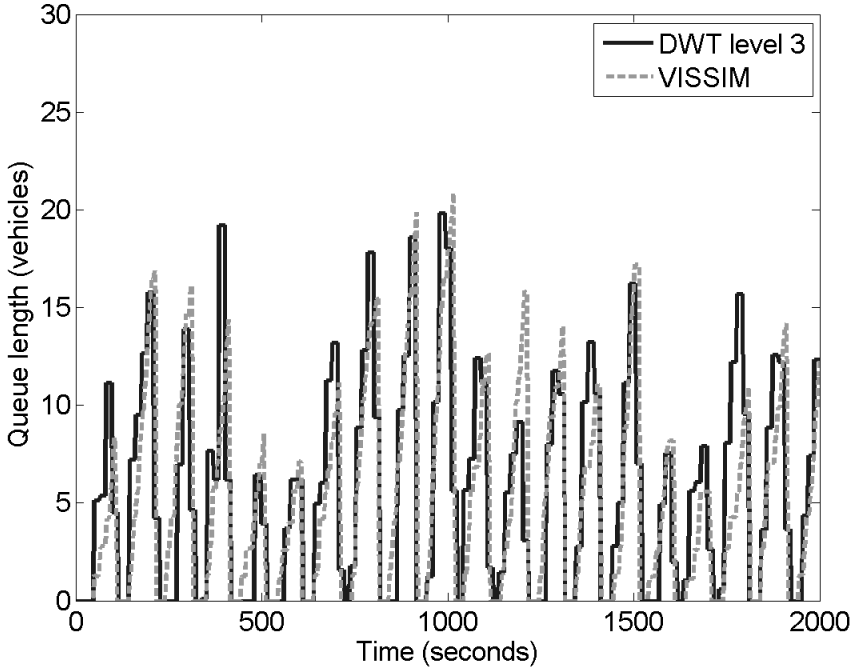

Figure 25 Case A: DWT level 3 when penetration ratio = 30%

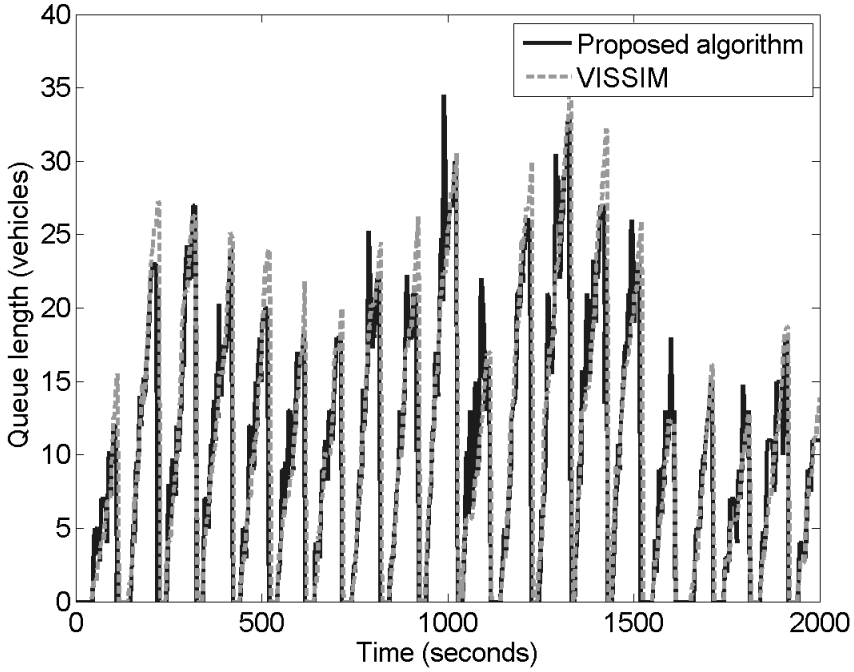

Figure 26 Case B: The estimated queue length when penetration ratio = 50%

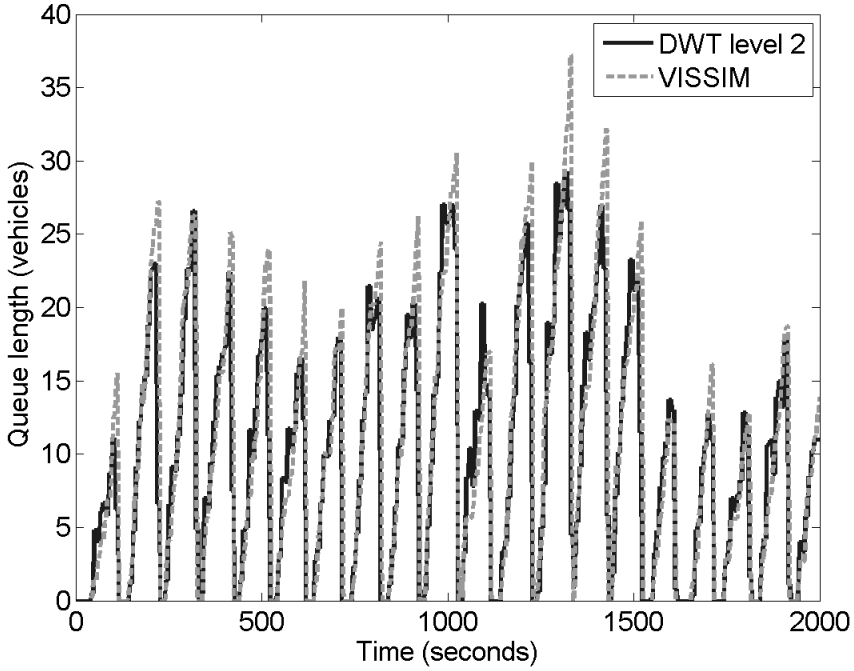

Figure 27 Case B: DWT Level 2 when penetration ratio = 50%

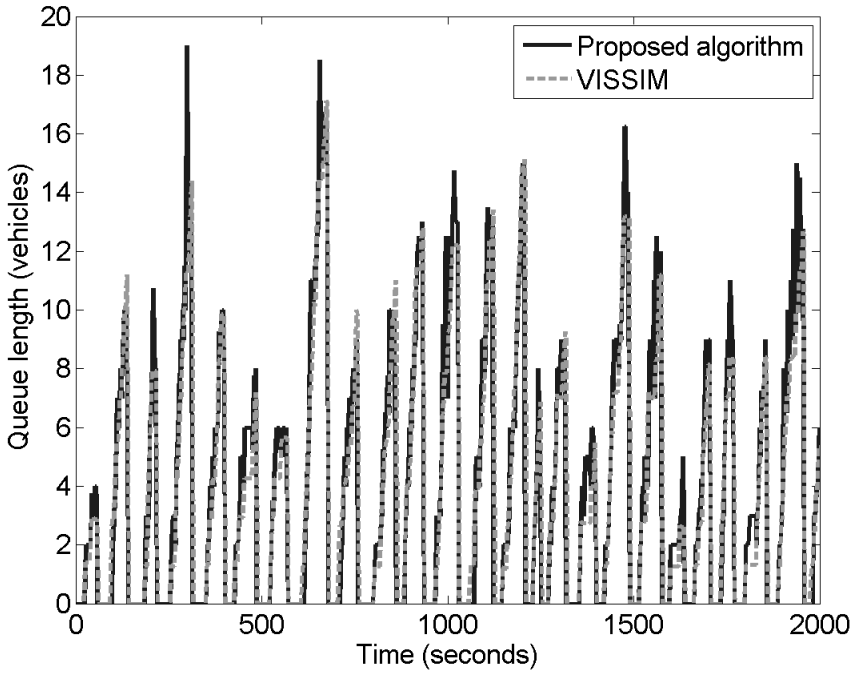

Figure 28 Case C: The estimated queue length when penetration ratio = 80%

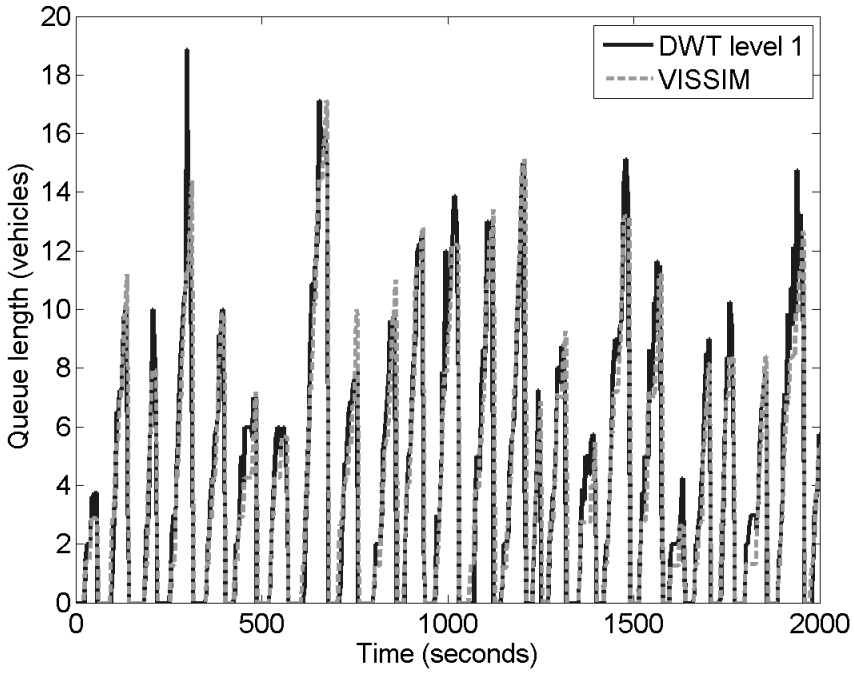

Figure 29 Case C: DWT Level 1 when penetration ratio = 80%

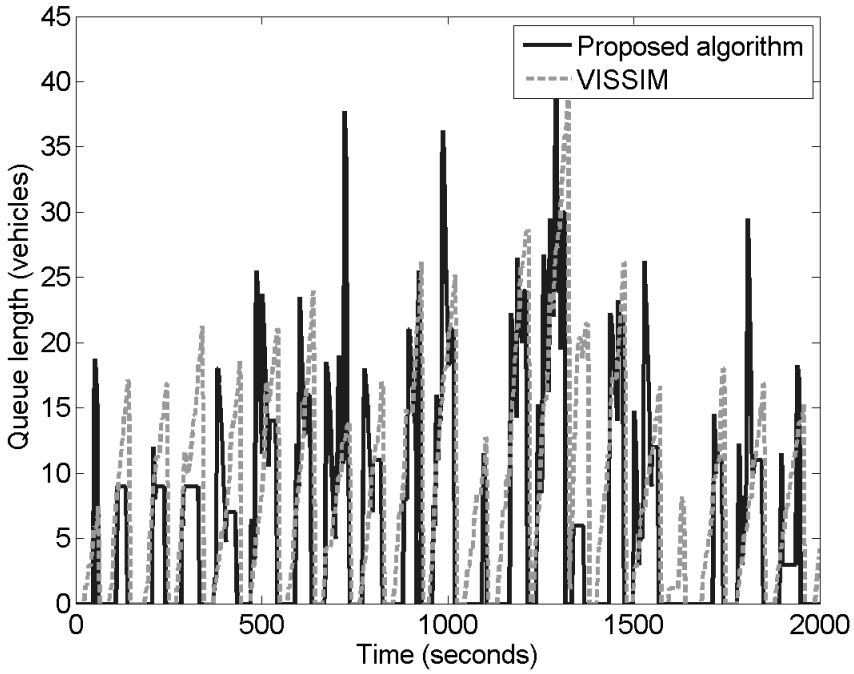

Figure 30 Case D: The estimated queue length when penetration ratio = 10%

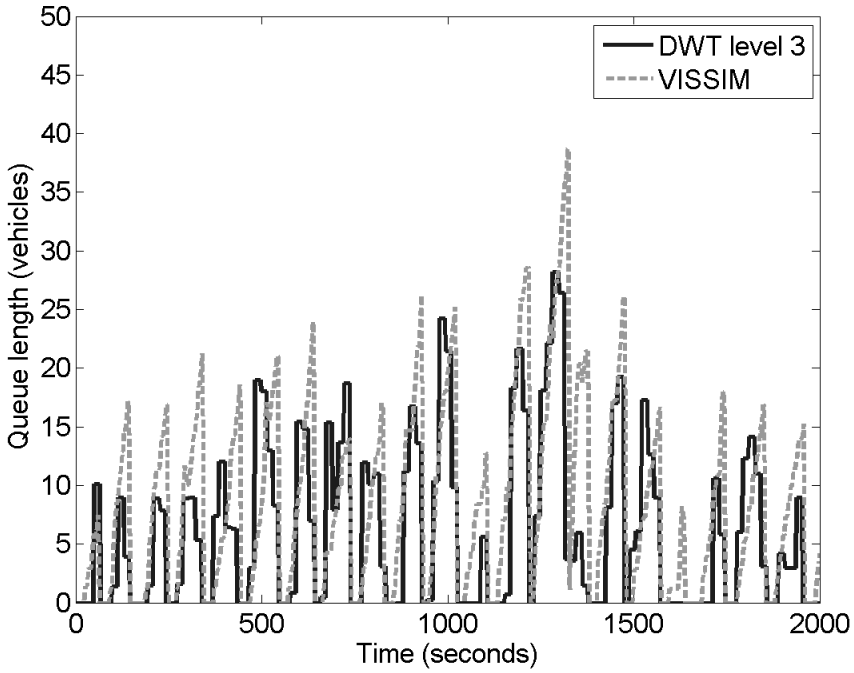

Figure 31 Case D: DWT Level 3 when penetration ratio = 10%

However, the most important contribution of applying DWT is to reduce random spikes and improve consistency of queue estimation. Figure 24-Figure 31 show the examples of time history of the proposed algorithm before and after applying with the best DWT at specific penetration ratios. The random spikes are reduced from before applying DWT in Figure 25, Figure 27, Figure 29, and Figure 31. The time series of queue estimation are smoothened by DWT.

Summary

Since most existing queue estimations from connected vehicle technology are only applied to pre-timed signals, a new algorithm aiming for adaptive signal control is proposed in this dissertation. To be applicable for adaptive signal control, the proposed algorithm can estimate queue length without the assumption of pre-timed signal, signal

interval, and specific arrival distribution. In addition, the volume, queue characteristic and signal timing data are not required as basic inputs. More importantly, the algorithm can work in both undersaturated and saturated conditions, which is necessary for adaptive signal control as it has to deal with a wide range of traffic condition. The proposed algorithm also focuses on the consistency of estimation when penetration ratio is low. Thus, discrete wavelet transform was originally applied to enhance the consistency of queue estimation in this work. Even though the simple frequency filter is used rather than using a soft or hard threshold, the results show that DWT can improve by both decreasing RMSE and making the estimation more consistent by reducing random spikes. Because the proposed queue estimation algorithm is able to estimate queue length under various traffic conditions and non pre-timed signals, it is expected that it will contribute significantly to adaptive signal control using connected vehicle technology. There are several ongoing efforts. The first one is to develop adaptive signal control strategies with queue estimation. The second one is to improve threshold scheme of DWT for better denoising performance.

CHAPTER V

NUMERICAL RESULTS FOR PLATOON RECOGNITION

In this chapter, there are two main parts. The first one is a calibration step to obtain calibrated parameters' value. After that, the obtained parameters are used to test the proposed algorithm with the same data as the training data. The second part is to use calibrated parameter values from the first part to test with other four data sets from different signal control, traffic volume, or data collection location. The second part intends to validate the applicability range of the proposed algorithm when traffic condition and/or signal control are changed without recalibrating the parameters.

A microscopic simulation, VISSIM 5.10 (*9*), is selected to be the test-bed for generating individual vehicle data. All testing is based on the network shown in Figure 32. The test network is an isolated intersection with two lanes for east and west bound and three lanes for north and south bound. All approaches are 270 meter-long with one left-turn bay. In the calibration process, traffic signal control is pre-timed signal with protected-left turn. The green time for left turn and through movement are 15 and 35 seconds for all approaches. Traffic volumes for EB and WB, and NB and SB are 1,400 and 1,000 vph, respectively. The platoons are observed from vehicles outbound from the intersection in the eastbound direction. Speed and distance of each vehicle are collected every 0.5 seconds at a location 30 meters from the intersection to calculate actual platoon data. The data collected at a distance 120 meters downstream in Figure 32 is used for validating the proposed algorithm for a different location without recalibration

in the second part of this chapter. The simulation time is 2,000 seconds. To demonstrate the penetration ratio affect, CVs are randomly selected corresponding to a penetration ratio value to obtain CV data.

Calibration Results

The calibration process is implemented to obtain calibrated parameter values via the minimization of the fitness function. The genetic algorithm (GA) is selected for executing the optimization process. The GA fitness function is defined as in Equation (34). In the experiment, $\vec{W}^T = [1 \quad 1 \quad 1 \quad 0.1 \quad 0.1]^T$. 100 randomization are carried out in this chapter to get reasonable steady values for fitness function, hence, $I = 100$. The GA population vector contains five calibration parameters, $\left[T_p \; f_1 \; f_2 \; f_2 \; y_p\right]^T$. The parameters, T_1, $T_{\min}$, m, and y_1, are set to be 5 and 2 seconds, 2 lanes, and 15 vehicles, respectively. The correspondent upper and lower bounds are shown in Table 1. The platoon data computed directly from VISSIM's vehicle data with penetration ratio 100%, no missing data, is considered as the ground truth. The error measurements are computed by comparing the estimated platoon data from the proposed algorithm from various penetration ratios with the ground truth (100% penetration ratio). With MATLAB's optimization toolbox, the GA setting is set to be the same as the default setting in MATLAB R2012A, except for the number of generations, which is set to be 50.

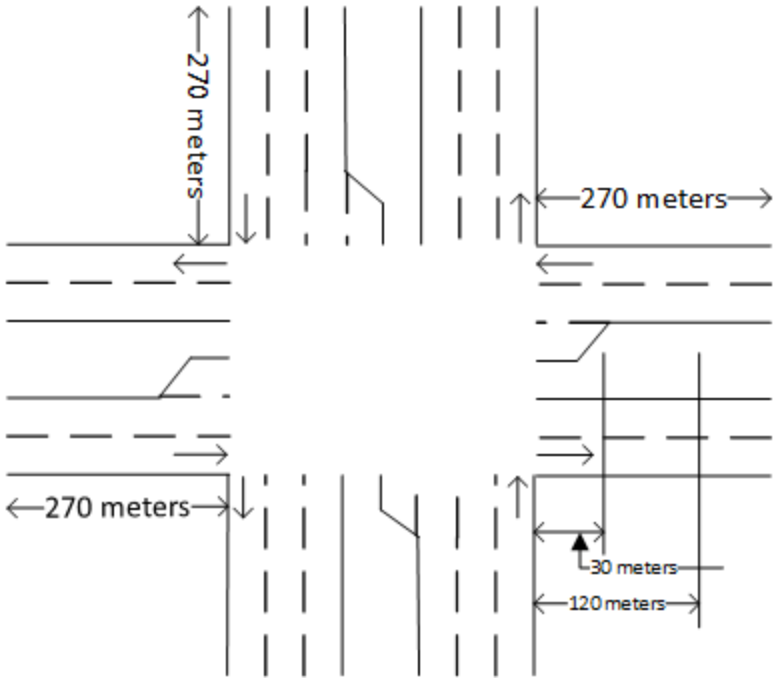

Figure 32 The topology of tested network

Table 2 shows resulting calibrated parameters. After obtaining calibrated parameters, the proposed algorithm then applies CV data and the calibrated parameter values to recognize platoons. The estimated platoon size, beginning and ending time are obtained, along with the estimation error measurements, RMSEs of the size and beginning and ending times, and the detection and false detection rates. The examples of estimated platoon history, before and after filtering, for penetration ratio 30%, 50% and 70% are shown in Figure 33-Figure 38.

Table 1 Lower and upper bounds of calibrated parameters

calibrated parameter	lower bound	upper bound
T_p	$\frac{5+15\sqrt{1-p}}{p}$	$\frac{5}{p}$
f_1	0.4	2
f_2	0	1
f_3	0.2	1
y_p	$15 \cdot \min(p, 0.8)$	15

Table 2 Resultant calibrated parameters

penetration ratio	T_p	f_1	f_2	f_3	y_p
10%	50.3	0.51	0.94	0.59	2
20%	26.8	0.43	0.22	0.75	3
30%	17.0	0.43	0.89	0.81	5
40%	13.1	0.43	0.98	0.95	6
50%	10.4	0.44	0.99	0.69	8
60%	8.4	0.41	0.99	0.75	10
70%	7.2	0.41	0.98	0.96	11
80%	6.4	0.40	0.97	0.70	12
90%	5.7	0.43	0.97	0.57	12

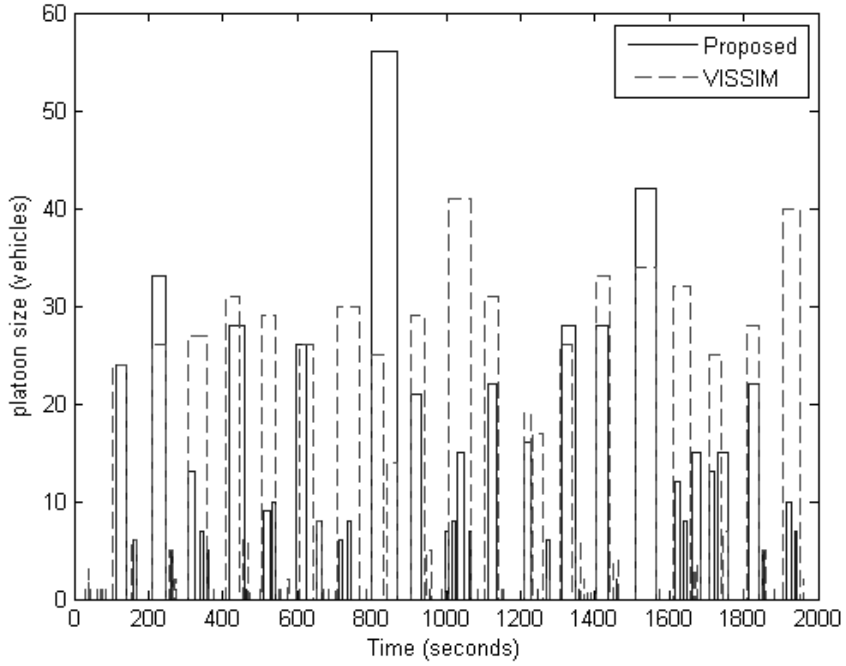

Figure 33 The pre-filtering estimated platoon when penetration ratio = 30%

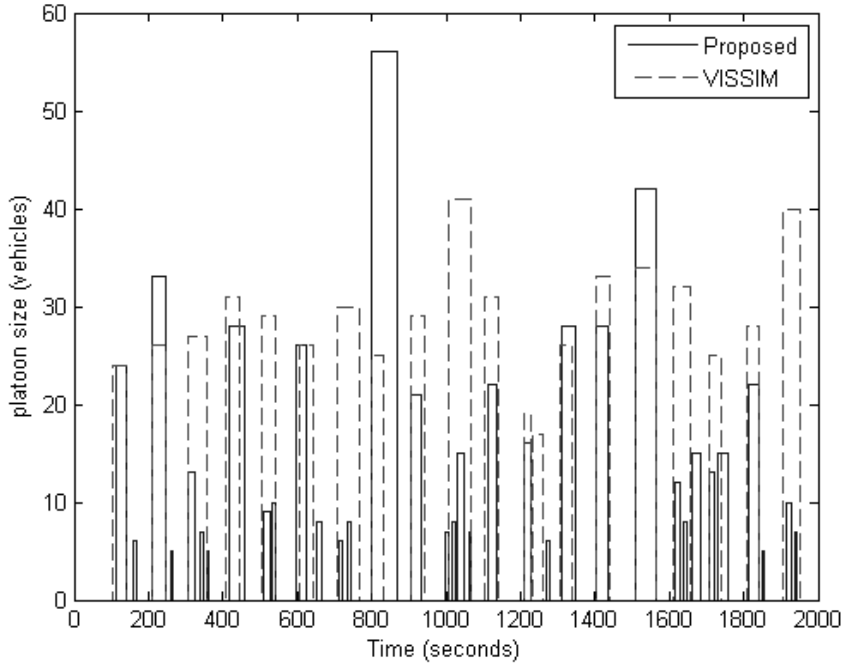

Figure 34 The post-filtering estimated platoon when penetration ratio = 30%

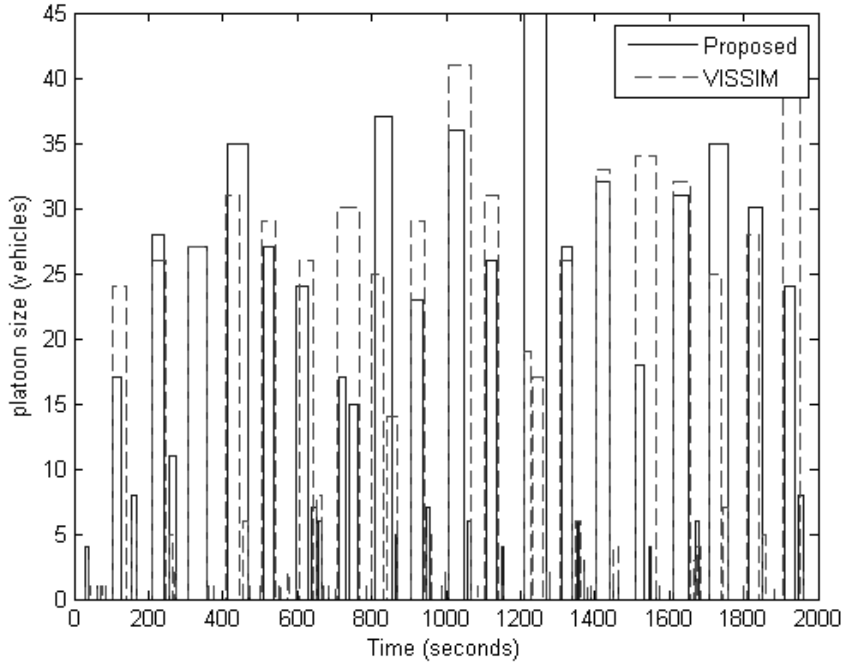

Figure 35 The pre-filtering estimated platoon when penetration ratio = 50%

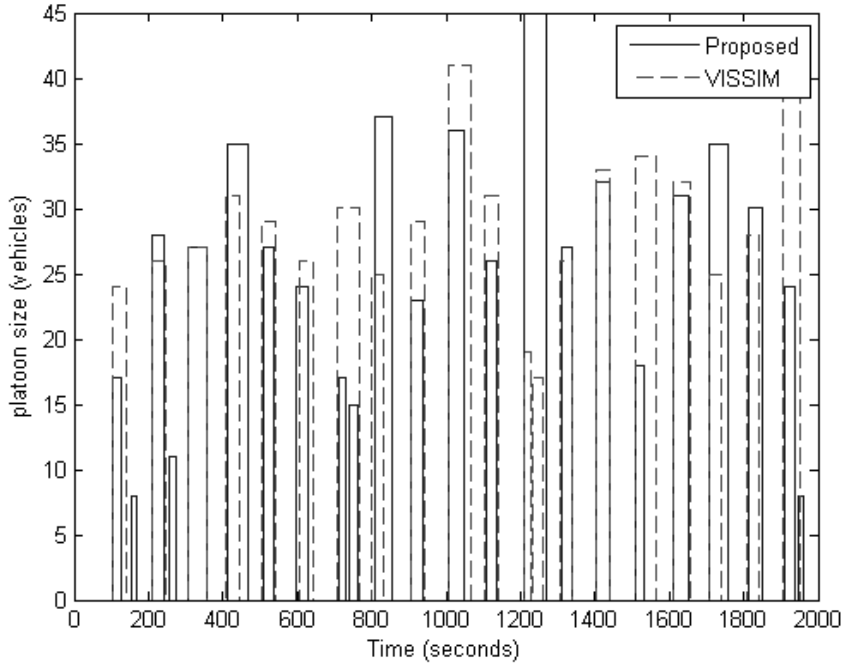

Figure 36 The post-filtering estimated platoon when penetration ratio = 50%

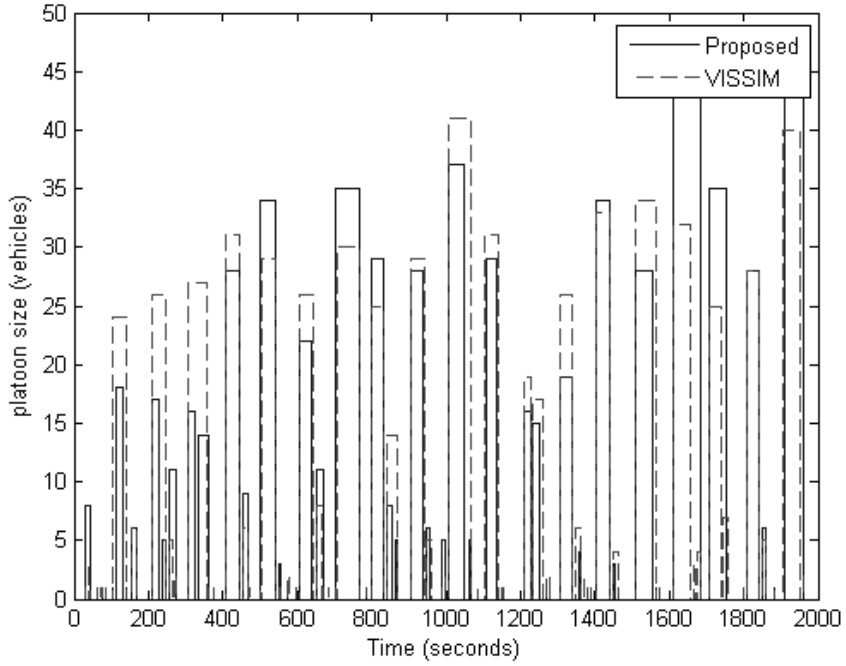

Figure 37 The pre-filtering estimated platoon when penetration ratio = 70%

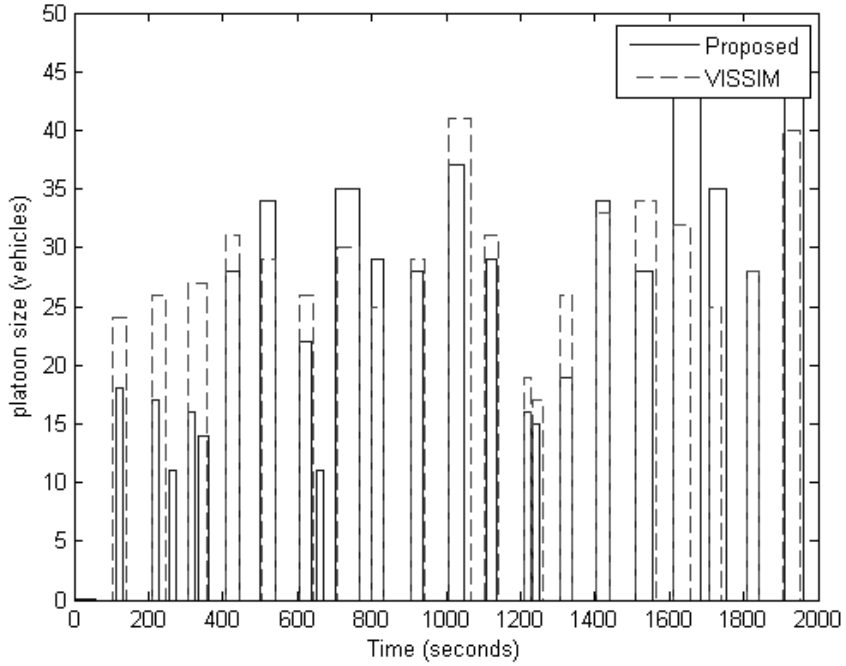

Figure 38 The post-filtering estimated platoon when penetration ratio = 70%

To avoid confusion, it should be noted that the platoon size is not equal to the area of Figure 33-Figure 38 but the y-axis value. In addition, the lower penetration ratio, the more frequent occurrence of merging and diverging. When merging occurs, two or more consecutive platoons are detected as one long platoon, for instance, at time between 400-500 seconds in Figure 35-Figure 36. The two actual-platoon sizes are 32 and 5 vehicles, while the detected-platoon size is 35 vehicles. The merged actual platoon then has 32+5 = 37 vehicles. The difference of platoon size is, therefore, 2 vehicles. Nonetheless, signal coordination yields a green time to the two actual platoons almost the same way it does for the single detected platoon because the time gap between the two actual platoon is too small to discard either one of them. Merging from platoons when their time gap(s) are too large rarely occurs because of the relative small T_p compared to the one used in (*26*). In addition, the platoon identification in Figure 4 is purposely developed to prevent such a case.

The diverging case can be observed in the same figures at time between 700-800 seconds. In this case, the diverged detected platoon occurs because of small T_p, as a tradeoff to prevent merging of platoons with large time gap(s). However, as in the merging case, signal coordination tends to allocate sufficient green time to closely spaced platoons without premature gap out. Hence, in most cases, even though merging and diverging can cause the estimation errors, the traffic operation may still perform well without significant problems due to consistent detectability, which is shown in the later subchapter. It is also noteworthy to mention the importance of y_p. The well fine-

tuned y_p is important to keep detected platoon in diverging cases. From the Figure 33-Figure 34, the diverged detected platoons are kept to allow to mimic the real platoon because y_p is low enough to let it pass through the filtering process. If y_1 is used in filtering process instead, most diverging detected platoons are excluded and signal coordination can neither detect or let a platoon move without stops. On the other hand, a too small y_p can increase false detection rate as ‘‘unqualified’’ platoons are not filtered out properly, resulting in an ineffective green time. Thus, a proper calibration of y_p can lessen the negative effects of merging and diverging, enhance detectability, and is varied with the penetration ratio and the platoon profile.

Validation Results

The calibrated parameters in Table 1 are further applied to estimate platoons with other four scenarios in this subchapter. The purpose is to validate the performance of the proposed algorithm when traffic condition and signal control change without recalibration. The base case in the first part is referred as case 1. The other four cases are

- case 2: pre-timed signal control with 1,000 vph on eastbound and westbound approaches
- case 3: actuated signal control with 1,400 vph on eastbound and westbound approaches
- case 4: actuated signal control with 1,000 vph on eastbound and westbound approaches

- case 5: same as the base case but the location of platoon detection is 120 meters away from the intersection

For both pre-timed and actuated signal controls, volumes 1,000 and 1,400 vph are approximately undersatured and saturated conditions, respectively. The actuated signal control in case 3 and 4 has min and max green time 6 and 26 seconds for left turn and 10 and 36 seconds for through movements. All phases have 2 seconds of vehicle extension. The platoon detecting location in case 5 is shown in Figure 32. The performance indicators for all 5 cases are shown in Figure 39-Figure 44.

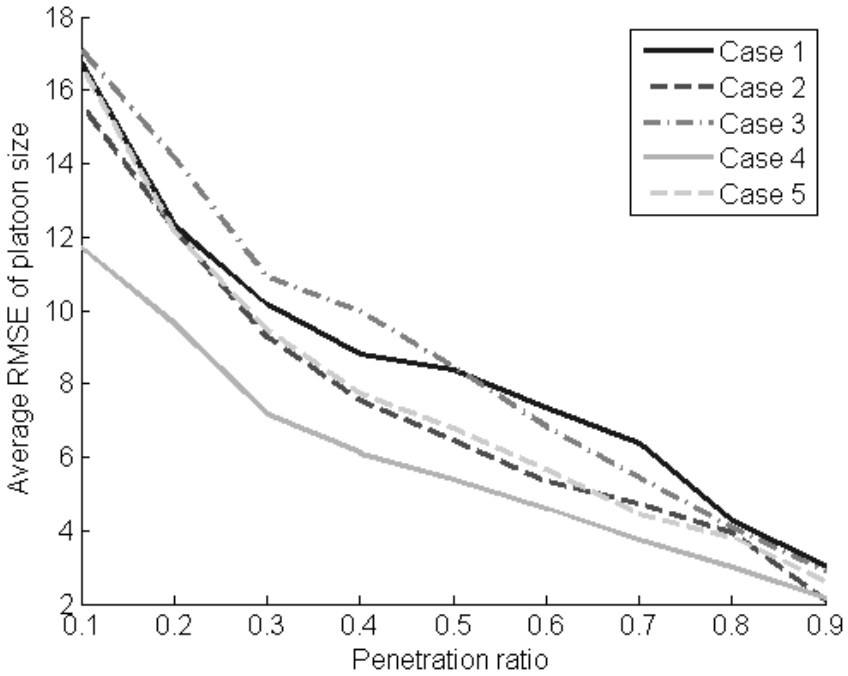

Figure 39 Average RMSE of platoon size

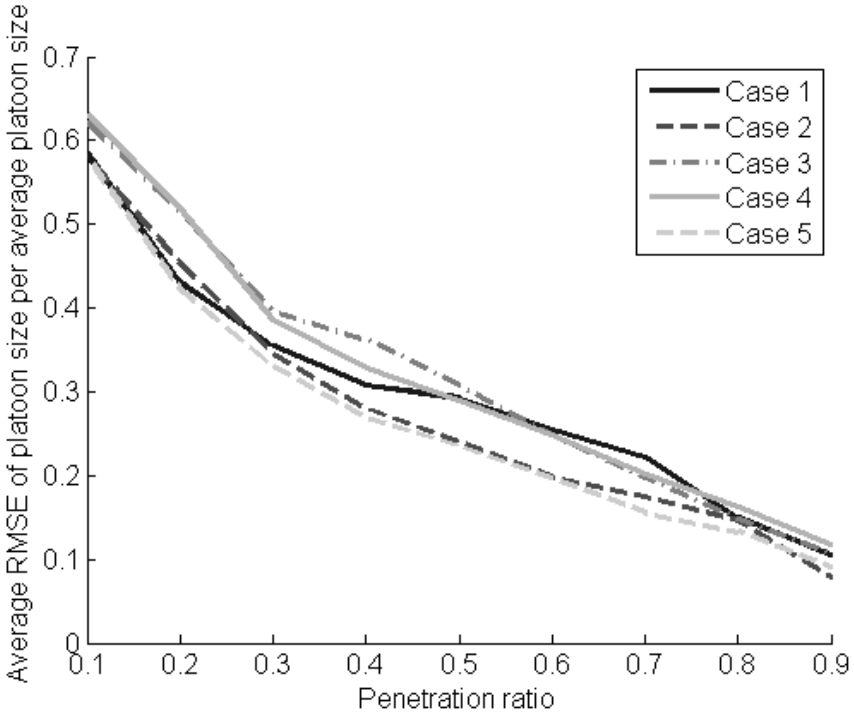

Figure 40 Average RMSE of platoon size per average actual platoon size

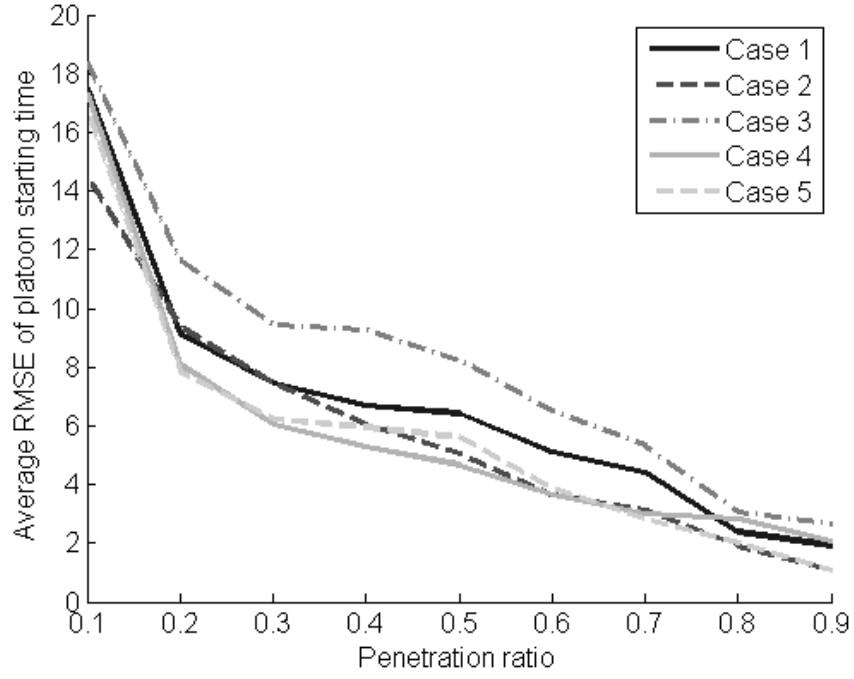

Figure 41 Average RMSE of beginning time

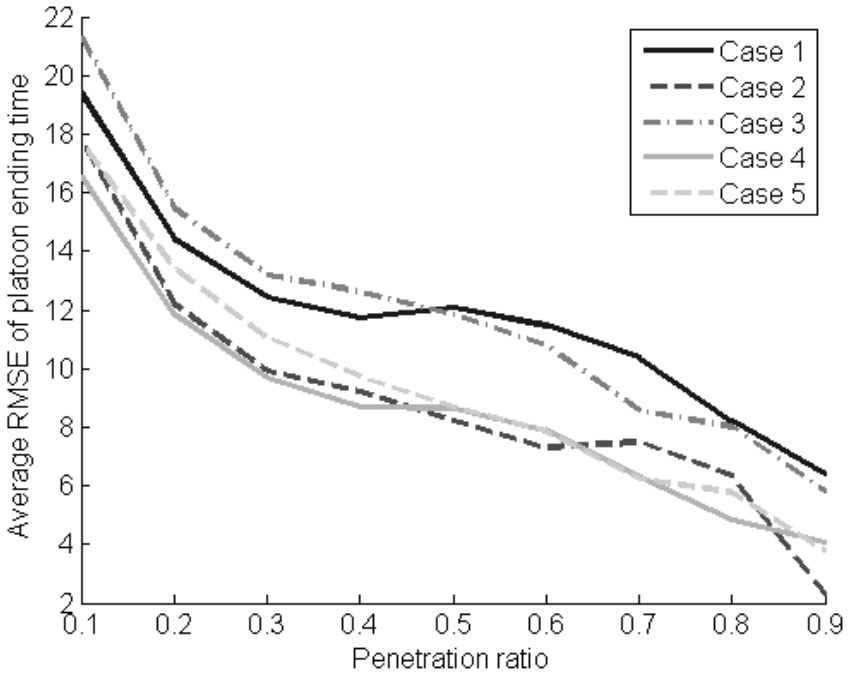

Figure 42 Average RMSE of ending time

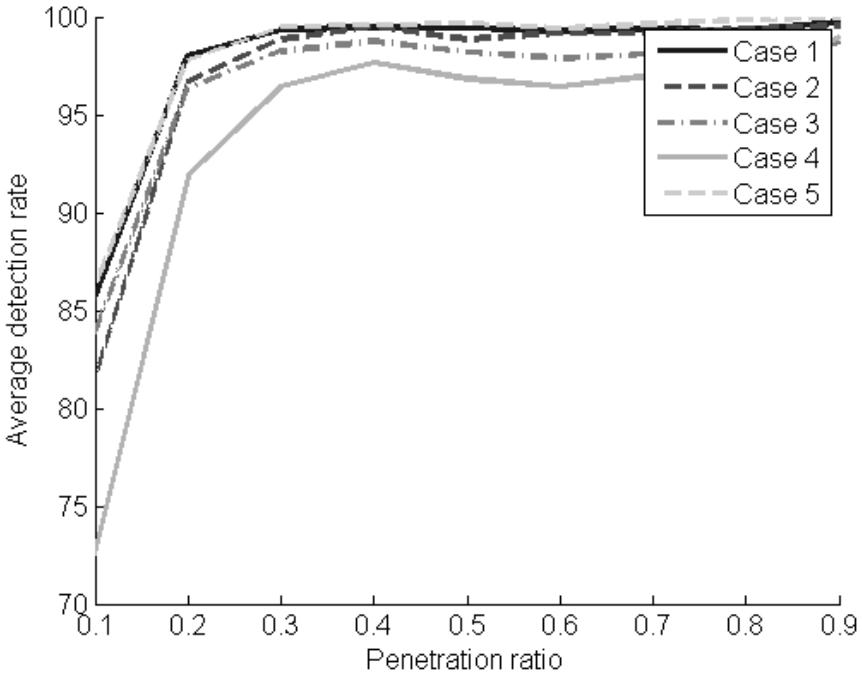

Figure 43 Average detection rate

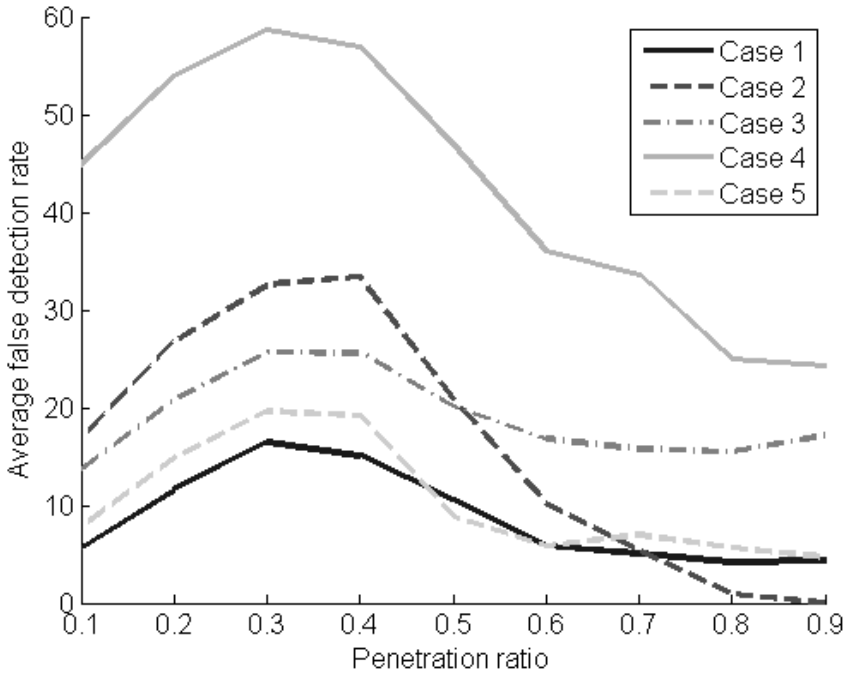

Figure 44 Average false detection rate

From Figure 39-Figure 42, all three RMSE decrease as the penetration ratio increases. It should be noted that, Figure 39 shows that while the RMSE for platoon size is smaller in cases 2 and 3, the average platoon sizes are significantly lower in both cases due to actuated signal control. In Figure 40, the ratio of RMSE of platoon size and the average platoon size is lower in the base case. The reason is RMSE of platoon size is proportional to average platoon size, and the selected value of calibration parameters fits better with the base case. Figure 41-Figure 42 show that the estimation of starting time is more accurate than the ending time by 2 seconds in average. These results imply that the adjustment applied to the estimation of starting time in Equation (27) can improve the accuracy of the estimation. Moreover, the improvement is obtained without the expense

of the other error measurements. Average false detection rate in Figure 44 decreases after the penetration ratio is lower than 30% because estimated platoons need to be merged to a larger one due to a rapid increase of T_p after the penetration ratio lower than 30%. Merging cases are not considered as misdetection or false detection in the simulation. However, they are considered in computing RMSEs of the size, and beginning and ending times as all three RMSEs increases significantly after the penetration ratio is lower than 30%.

In case 4, the false detection rate is abnormally high because of the platoon profile and the effect of the filtering. This is an actuated signal control with undersaturated condition that does not produce long platoons. In the experiment, y_1 is set to 15 vehicles. When an actual platoon with a size a bit smaller than 15 vehicles is detected, it can often lead to false detection. Because the actual platoon are filtered out by y_1, while the detected platoon is not filtered by y_p. Moreover, there are only 12 platoons out of 101 that have a platoon size greater than 15 vehicles. As a result, 89 actual platoons are filtered out and cause the detected platoons to be classified as false detection. The resultant false detection rate is, therefore, higher than it should be. It is noteworthy that a smaller platoon is harder to detect. As evidenced by Figure 43, the average detection rate is the lowest of 5 cases. Detecting more of them will increase false detection rate. By giving more weight to the error measurement of false detection in the fitness function, the calibration process can reduce false detection rate at the expense of the other error measurements.

A comparison of detection rate from simulation and two analytical approaches is shown in Figure 45-Figure 49. Two analytical approaches are given. The first one is when a platoon size profile is known and estimated by Equation (32). The second one is when a platoon size profile is unknown and a lower bound of detection rate is instead estimated by Equation (33). For a penetration ratio higher than 30%, analytical approach 1 and 2 tend to overestimate the detection rate because both analytical approaches do not take into account the effect of platoon filtering. A few detected platoons, that should match with the filtered actual platoon, are filtered out because their sizes are smaller than y_p, while the analytical approaches assume that all detected platoon are kept and matched with the actual platoons. In contrast, at a penetration lower than 30%, both analytical approaches tend to underestimate the detection rate because of the merging affect. As more detected platoons are merged at the lower penetration ratio, it can increase detection rate by covering more time span and detecting more platoons. The assumption is the analytical approaches do not take into account merging issue, so they ignore the increased detection rate from merged detected platoons. Nonetheless, analytical approach 1 still estimates the detection rate accurately with less than 5% error in most cases. As expected, analytical approach 2 is less accurate than the first one due to the lack of platoon profile as input. However, analytical approach 2 still yields the acceptable estimation when the penetration ratio is higher than 30%. It also does give a lower bound for detection rate at the penetration lower than 30%. The obtained lower bound is useful to roughly determine whether platoon recognition from CVT is applicable for the desire application.

Out of the 5 cases, the estimated detection rate from the analytical approach 2 is the most accurate in case 4. This is because the average platoon size post filtering is very close to y_1. As analytical approach 2 estimates the detection rate by assuming a filtered platoon size is equal to y_1, the smaller the average filtered platoon size, the more accurate the detection rate from analytical approach 2.

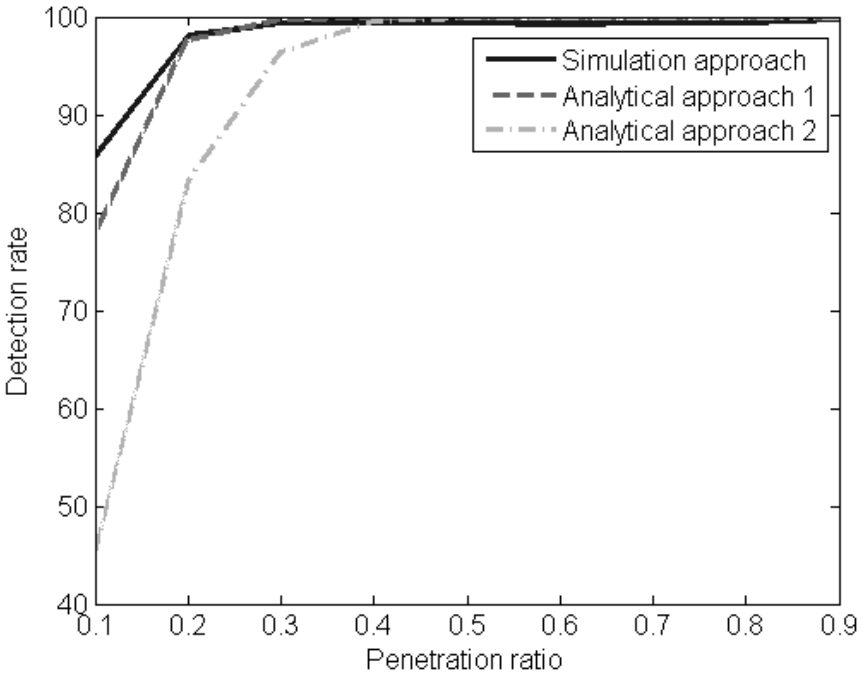

Figure 45 Average detection rate for case 1

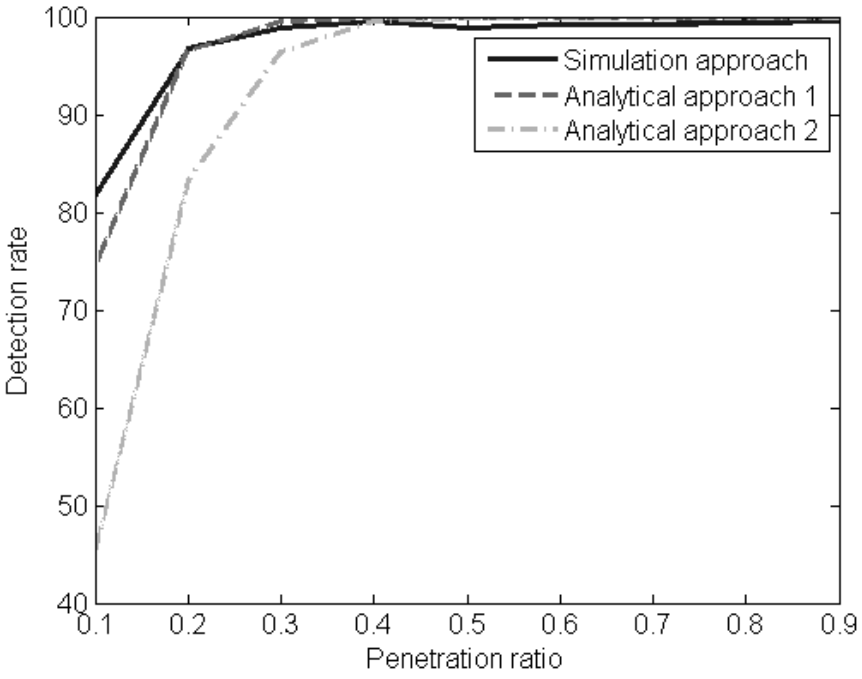

Figure 46 Average detection rate for case 2

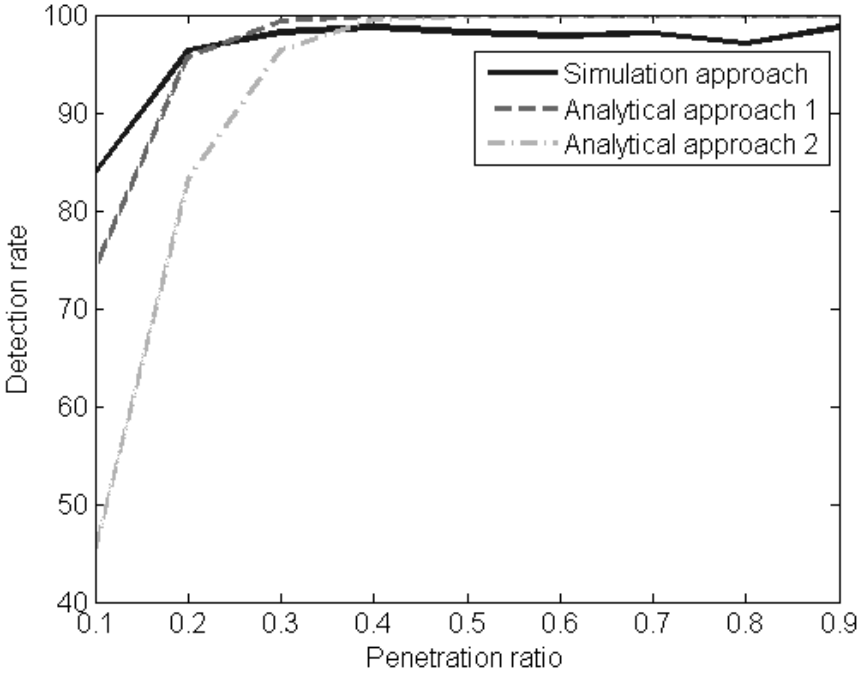

Figure 47 Average detection rate for case 3

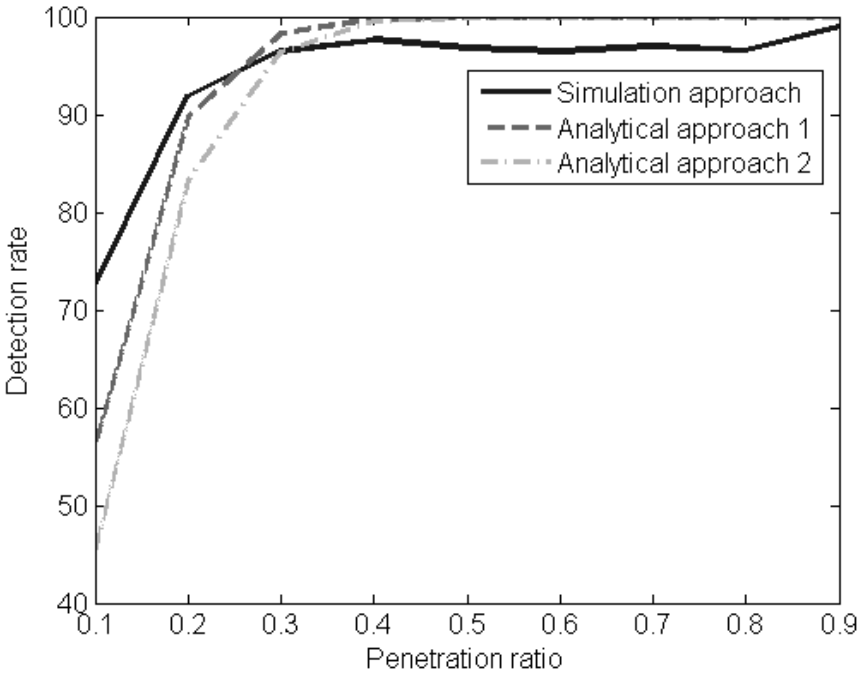

Figure 48 Average detection rate for case 4

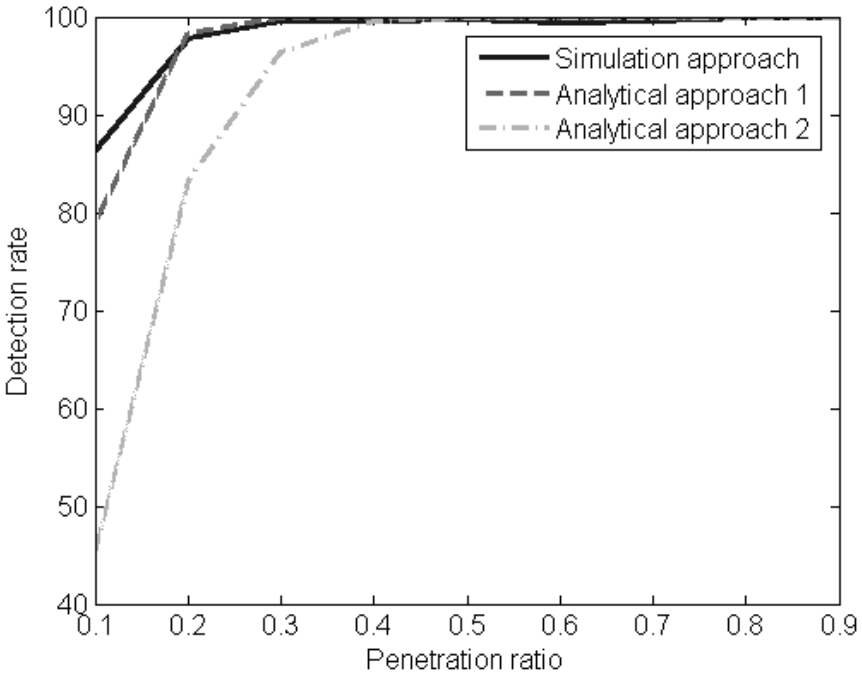

Figure 49 Average detection rate for case 5

Summary

There are a few existing platoon recognition algorithms using connected vehicle technology (CVT). However, none of them gave a thorough investigation with analysis on the performance, and the applicable range. The dissertation proposed a platoon recognition algorithm from CVT. The detected platoon is identified from time-stamp difference between two consecutive connected vehicles (CV) with a modified critical time-headway. In addition, the analytical approaches for estimating detection rate are proposed to estimate the required penetration ratio to obtain the detection rate based on the threshold and platoon profile. The new evaluation method is proposed to obtain performance separately: estimation error when platoon is detected, and detection and false detection rates when the platoon is undetected. Numerical results from VISSIM show that the proposed algorithm can detect platoons effectively in various traffic conditions, especially at a penetration higher than 30%, even without recalibration. The detection rate obtained from the analytical approaches corresponds to the detection rate from the numerical results, so it can be useful to determine the required penetration ratio for any application from CVT. The platoons are recognized without basic inputs that are hard to obtain from CVT, such as traffic volume and signal timing. Therefore, the proposed algorithm is expected to be well applicable to adaptive signal control. The future work is to develop platoon prediction from the platoon recognition and adaptive signal control logic.

CHAPTER VI

CONCLUSIONS AND FUTURE WORK

Contributions

This dissertation proposed queue length estimation and platoon recognition using connected vehicle technology. New algorithms were developed so that they can work with adaptive signal control. Both algorithms were evaluated in both pre-timed and actuated signal controls, and undersaturated and saturated traffic conditions. The contributions from this research are:

1. While there were some existing studies on queue length estimation, most of them were evaluated with pre-timed signal control only. This research is among the first to propose and verify queue length estimation from CVT with actuated signal control.
2. A few, if any, of existing queue length estimation considered consistency of the estimation results, for instance, random spikes in the estimation. Discrete wavelet transform was first purposed to enhance the accuracy and, more importantly, the consistency of the queue length estimation algorithm in this research.
3. Most queue length estimation algorithms relied on inputs that are hard to obtain in practice with CVT alone. The proposed algorithm can estimate queue length without relying on other detectors and inputs that are hard to

obtain in practice such as signal timing and traffic volume with reasonable accuracy and consistency.

4. To the best of our knowledge, we are the first to propose and full analysis on platoon recognition using CVT. The mathematic model as well as analysis on the impact of its parameters such as a modified critical time-headway was fully investigated and evaluated in separate scenarios, rather than a compound scenario.
5. We also believe to be the first to propose the analytical model on platoon detection rate using CVT. The evaluation of platoon recognition algorithm by two distinct measurements; the estimation error when platoons are detected and detection and false detection rates when platoon are undetected or falsely detected was included to give better insight into the numerical results. With the estimated detection rate and the new evaluation method, the applicable range of the platoon recognition can be perceived in more details by traffic engineers.

Future Research

Even though the numerical results verified and validated queue length estimation and platoon recognition algorithms in various cases, there are some investigations that is worth investigation in the future. The further extension to obtain inputs for adaptive signal control can be carried out in the future as new research topics.

1. In the queue length estimation, discrete wavelet transform with simple method, high-pass component complete elimination (detail component) was

chosen. In the future, a new threshold scheme to obtain wavelet coefficients should be developed for the proposed algorithm as it can improve the algorithm further.

2. Platoon recognition algorithm was verified with signalized platoons only. Verification with a random platoon arrival should be done.
3. The proposed queue length estimation and platoon recognition algorithms can be a foundation for the prediction algorithms for queue length and platoon arrival. As adaptive signal control should be proactive and include traffic prediction in a decision making, the two prediction algorithms are important for true adaptive signal control.
4. The adaptive signal control logic based on the estimated queue length and platoon recognition should be developed to verify the applicable range of CVT as traffic detector for adaptive signal control.

REFERENCES

1 Schrank, D., B. Eisele, and T. Lomax. TTI's 2012 urban mobility report. *Texas A&M Transportation Institute. The Texas A&M University System*, 2012.

2 Briesemeister, L., L. Schäfers, and G. Hommel. Disseminating messages among highly mobile hosts based on inter-vehicle communication.In *Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE*, IEEE, 2000. pp. 522-527.

3 Ott, J., and D. Kutscher. Drive-thru Internet: IEEE 802.11 b for" automobile" users.In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies, No. 1*, IEEE, 2004.

4 Consortium, C. V. S. C. Vehicle safety communications project: task 3 final report: identify intelligent vehicle safety applications enabled by DSRC. *National Highway Traffic Safety Administration, US Department of Transportation, Washington DC*, 2005.

5 Costa, P., D. Frey, M. Migliavacca, and L. Mottola. Towards lightweight information dissemination in inter-vehicular networks.In *Proceedings of the 3rd international workshop on Vehicular ad hoc networks*, ACM, 2006. pp. 20-29.

6 Andrews, S., and M. Cops. Final report: Vehicle infrastructure integration proof of concept executive summary–Vehicle. *US DOT, IntelliDrive (sm) Report FHWA-JPO-09-003*, 2009.

7 Morgan, Y. L. Managing DSRC and WAVE standards operations in a V2V scenario. *International Journal of Vehicular Technology*, Vol. 2010, 2010.

8 Zhang, J., F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen. Data-driven intelligent transportation systems: A survey. *Intelligent Transportation Systems, IEEE Transactions on*, Vol. 12, No. 4, 2011, pp. 1624-1639.

9 Vissim, P. 5.10 User Manual. *PTV, Karlsruhe, Germany*, 2008.

10 PTV, A. VISSIM 5.40 User Manual. *Karlsruhe, Germany*, 2011.

11 Vigos, G., M. Papageorgiou, and Y. Wang. Real-time estimation of vehicle-count within signalized links. *Transportation Research Part C: Emerging Technologies*, Vol. 16, No. 1, 2008, pp. 18-35.

12 Sharma, A., D. Bullock, and J. Bonneson. Input-output and hybrid techniques for real-time prediction of delay and maximum queue length at signalized intersections.

Transportation Research Record: Journal of the Transportation Research Board, No. 2035, 2007, pp. 69-80.

13 Akçelik, R. A queue model for HCM 2000. *ARRB Transportation Research Ltd., Vermont South, Australia,* 1999.

14 Webster, F. V. Traffic signal settings.In, 1958.

15 Liu, H. X., X. Wu, W. Ma, and H. Hu. Real-time queue length estimation for congested signalized intersections. *Transportation Research Part C: Emerging Technologies,* Vol. 17, No. 4, 2009, pp. 412-427.

16 Stephanopoulos, G., P. G. Michalopoulos, and G. Stephanopoulos. Modelling and analysis of traffic queue dynamics at signalized intersections. *Transportation Research Part A: General,* Vol. 13, No. 5, 1979, pp. 295-307.

17 Richards, P. I. Shock waves on the highway. *Operations research,* Vol. 4, No. 1, 1956, pp. 42-51.

18 Lighthill, M. J., and G. B. Whitham. On kinematic waves. II. A theory of traffic flow on long crowded roads.In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, No. 229*, The Royal Society, 1955. pp. 317-345.

19 Lighthill, M., and G. Whitham. On kinematic waves. I. Flood movement in long rivers.In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, No. 229*, The Royal Society, 1955. pp. 281-316.

20 Martin, P. T., and S. L. Hockaday. SCOOT--AN UPDATE. *ITE Journal,* Vol. 65, No. 1, 1995.

21 Sims, A. G., and K. W. Dobinson. The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits. *Vehicular Technology, IEEE Transactions on,* Vol. 29, No. 2, 1980, pp. 130-137.

22 Gartner, N. H. *OPAC: A demand-responsive strategy for traffic signal control*. 1983.

23 Head, K. L., P. B. Mirchandanai, and S. Shelby. *The RHODES prototype: a description and some results*. Transportation Research Board, 1998.

24 Siromaskul, S., and M. Selinger. InSync: The Next Generation of Adaptive Signal Systems.In *ITE annual meeting (westernite. org/annualmeetings)*, 2010.

25 Cai, C., Y. Wang, and G. Geers. Adaptive traffic signal control using wireless communications.In *Transportation Research Board 91st Annual Meeting*, 2012.

26 He, Q., K. L. Head, and J. Ding. PAMSCOD: Platoon-based arterial multi-modal signal control with online data. *Transportation Research Part C: Emerging Technologies,* Vol. 20, No. 1, 2012, pp. 164-184.

27 Smith, B. L., R. Venkatanarayana, H. Park, N. Goodall, J. Datesh, and C. Skerrit. IntelliDriveSM Traffic Signal Control Algorithms. *University of Virginia*, 2010.

28 Jin, Q., G. Wu, K. Boriboonsomsin, and M. Barth. Platoon-based multi-agent intersection management for connected vehicle.In *Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on*, IEEE, 2013. pp. 1462-1467.

29 Feng, Y., K. L. Head, S. Khoshmagham, and M. Zamanipour. A real-time adaptive signal control in a connected vehicle environment. *Transportation Research Part C: Emerging Technologies,* Vol. 55, 2015, pp. 460-473.

30 Comert, G. Effect of stop line detection in queue length estimation at traffic signals from probe vehicles data. *European Journal of Operational Research,* Vol. 226, No. 1, 2013, pp. 67-76.

31 Comert, G. Simple analytical models for estimating the queue lengths from probe vehicles at traffic signals. *Transportation Research Part B: Methodological,* Vol. 55, 2013, pp. 59-74.

32 Comert, G., and M. Cetin. Queue length estimation from probe vehicle location and the impacts of sample size. *European Journal of Operational Research,* Vol. 197, No. 1, 2009, pp. 196-202.

33 Comert, G., and M. Cetin. Analytical evaluation of the error in queue length estimation at traffic signals from probe vehicle data. *Intelligent Transportation Systems, IEEE Transactions on,* Vol. 12, No. 2, 2011, pp. 563-573.

34 Izadpanah, P., B. Hellinga, and L. Fu. Automatic traffic shockwave identification using vehicles' trajectories.In *Proceedings of the 88th Annual Meeting of the Transportation Research Board (CD-ROM)*, 2009.

35 Cheng, Y., X. Qin, J. Jin, and B. Ran. An exploratory shockwave approach for signalized intersection performance measurements using probe trajectories.In *89th Annual Meeting of Transportation Research Board*, 2010.

36 Ban, X. J., P. Hao, and Z. Sun. Real time queue length estimation for signalized intersections using travel times from mobile sensors. *Transportation Research Part C: Emerging Technologies,* Vol. 19, No. 6, 2011, pp. 1133-1156.

37 Cheng, Y., X. Qin, J. Jin, B. Ran, and J. Anderson. Cycle-by-cycle queue length estimation for signalized intersections using sampled trajectory data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2257, 2011, pp. 87-94.

38 Hao, P., and X. Ban. Vehicle queue location estimation for signalized intersections using sample travel times from mobile sensors.In *Transportation Research Board 90th Annual Meeting*, 2011.

39 Hao, P., Z. Sun, X. J. Ban, D. Guo, and Q. Ji. Vehicle index estimation for signalized intersections using sample travel times. *Transportation Research Part C: Emerging Technologies,* Vol. 36, 2013, pp. 513-529.

40 Sun, Z., and X. J. Ban. Vehicle trajectory reconstruction for signalized intersections using mobile traffic sensors. *Transportation Research Part C: Emerging Technologies,* Vol. 36, 2013, pp. 268-283.

41 Lee, S., S. Wong, and Y. Li. Real-time estimation of lane-based queue lengths at isolated signalized junctions. *Transportation Research Part C: Emerging Technologies,* Vol. 56, 2015, pp. 1-17.

42 May, A. D. *Traffic flow fundamentals*. 1990.

43 Neumann, T. A cost-effective Method for the Detection of Queue Lengths at Traffic Lights.In *Traffic Data Collection and its Standardization*, Springer, 2010. pp. 151-160.

44 Badillo, B. E., H. Rakha, T. W. Rioux, and M. Abrams. Queue length estimation using conventional vehicle detector and probe vehicle data.In *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*, IEEE, 2012. pp. 1674-1681.

45 Li, J.-Q., K. Zhou, S. Shladover, and A. Skabardonis. Estimating Queue Length Under Connected Vehicle Technology: Using Probe Vehicle, Loop Detector, and Fused Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2356, 2013, pp. 17-22.

46 Christofa, E., J. Argote, and A. Skabardonis. Arterial queue spillback detection and signal control based on connected vehicle technology. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2356, 2013, pp. 61-70.

47 Goodall, N. J., B. Park, and B. L. Smith. Microscopic Estimation of Arterial Vehicle Positions in a Low-Penetration-Rate Connected Vehicle Environment. *Journal of Transportation Engineering,* Vol. 140, No. 10, 2014, p. 04014047.

48 Goodall, N. J., B. L. Smith, and B. B. Park. Microscopic estimation of freeway vehicle positions from the behavior of connected vehicles. *Journal of Intelligent Transportation Systems*, No. ahead-of-print, 2014, pp. 1-10.

49 Tiaprasert, K., Z. Yunlong, X. B. Wang, and Z. Xiaosi. Queue Length Estimation Using Connected Vehicle Technology for Adaptive Signal Control. *Intelligent Transportation Systems, IEEE Transactions on,* Vol. 16, No. 4, 2015, pp. 2129-2140.

50 Chui, C. K. *An introduction to wavelets*. Academic press, 2014.

51 Akansu, A. N., and R. A. Haddad. *Multiresolution signal decomposition: transforms, subbands, and wavelets*. Academic Press, 2001.

APPENDIX

This part contains additional numerical results for queue length estimation that are not included in chapter IV.

The RMSE of Pre-timed signal with oversaturated condition is shown in Figure 50. The examples of time history results and the results with the best fit DWT levels are shown in Figure 51-Figure 56.

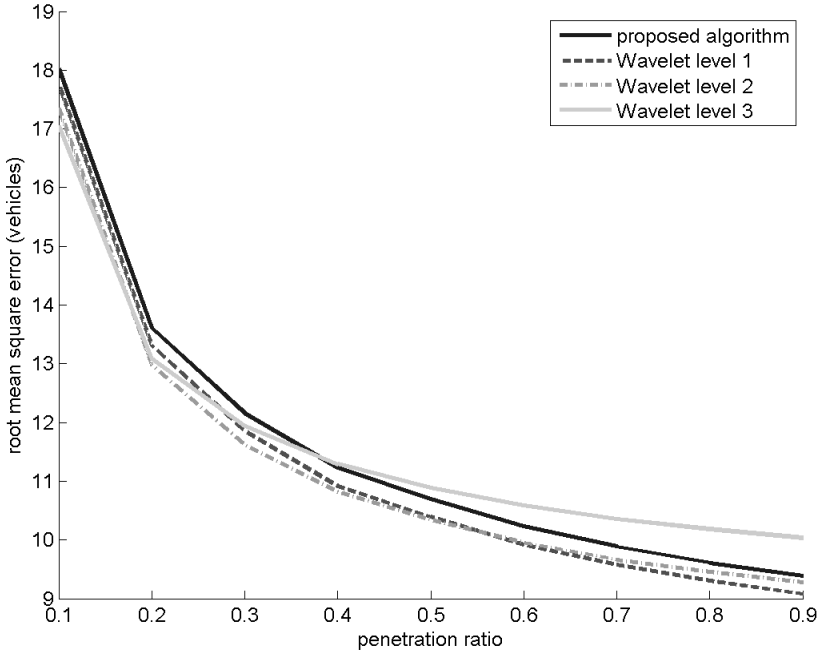

Figure 50 Root-mean-square error of queue length estimation in pre-timed with oversaturated condition case

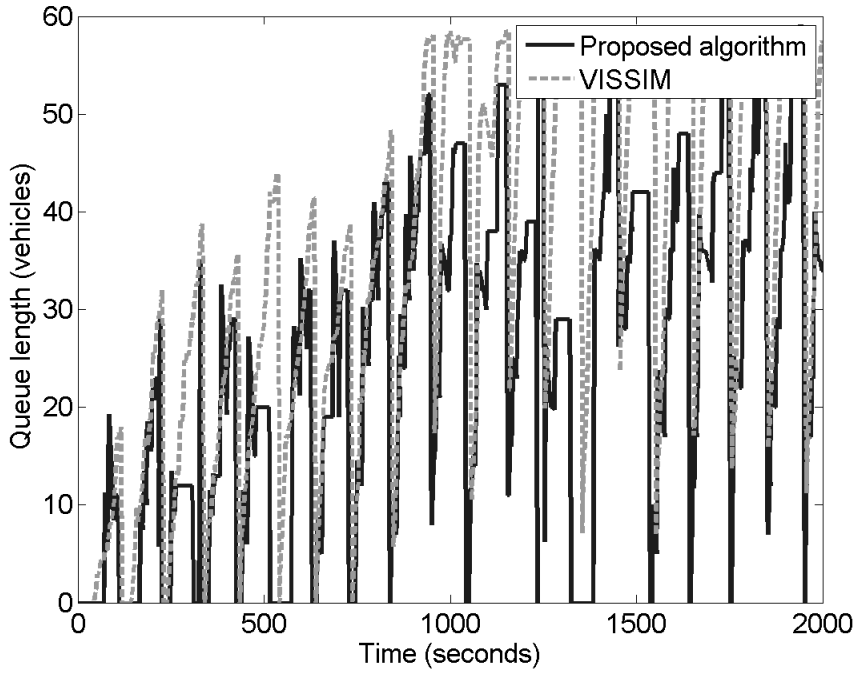

Figure 51 The estimated queue length when penetration ratio = 10%

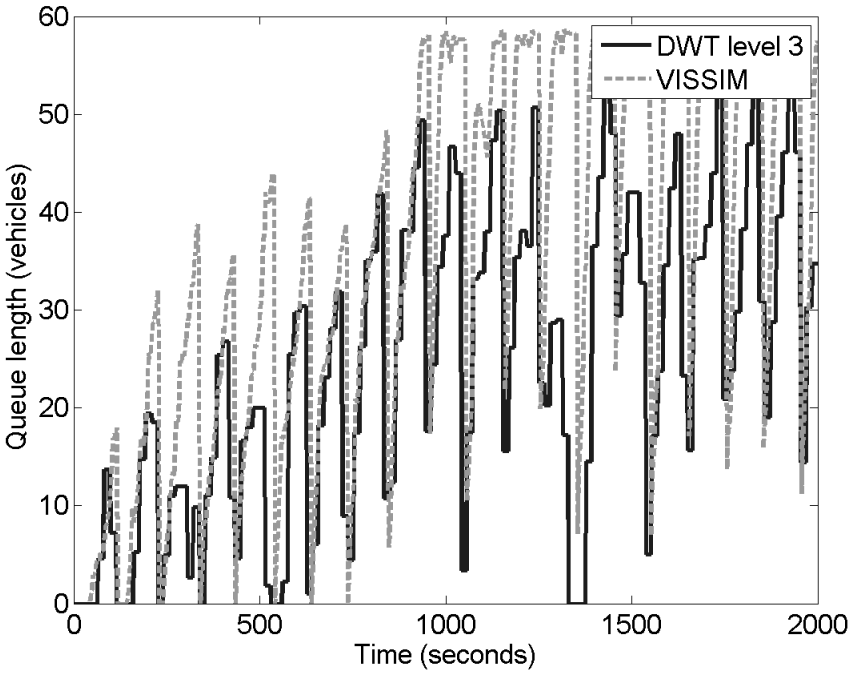

Figure 52 DWT level 3 when penetration ratio = 10%

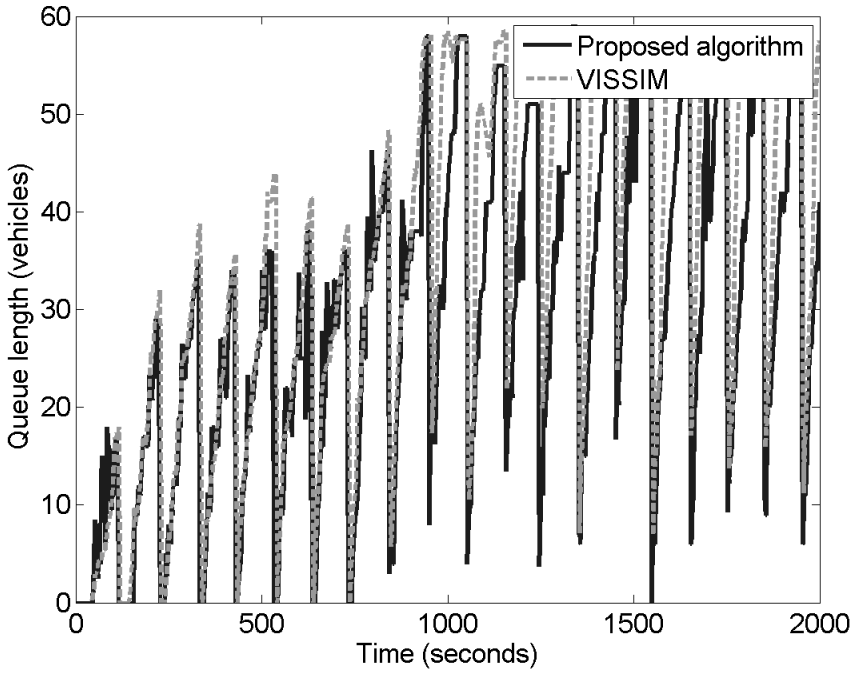

Figure 53 The estimated queue length when penetration ratio = 30%

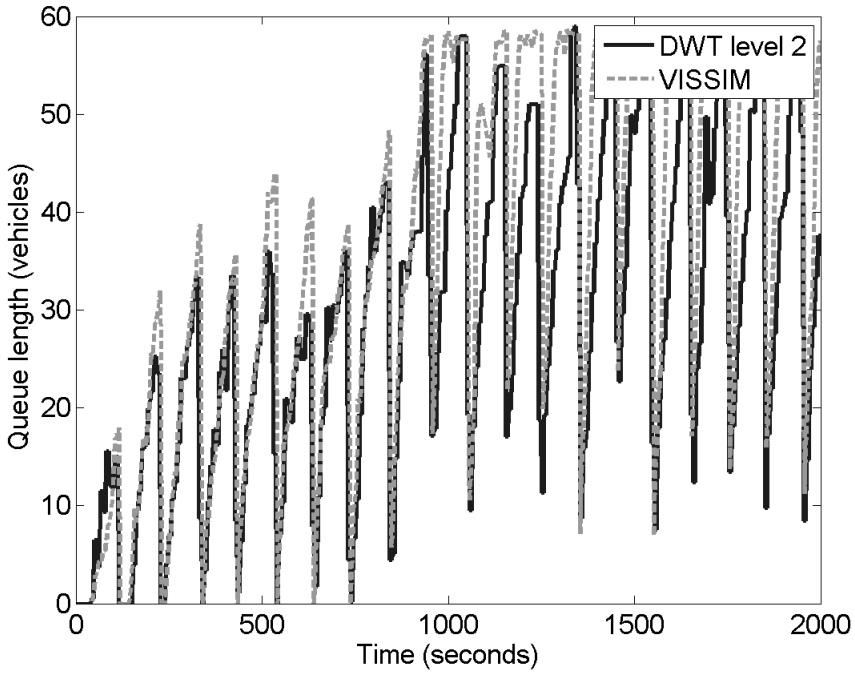

Figure 54 DWT level 2 when penetration ratio = 30%

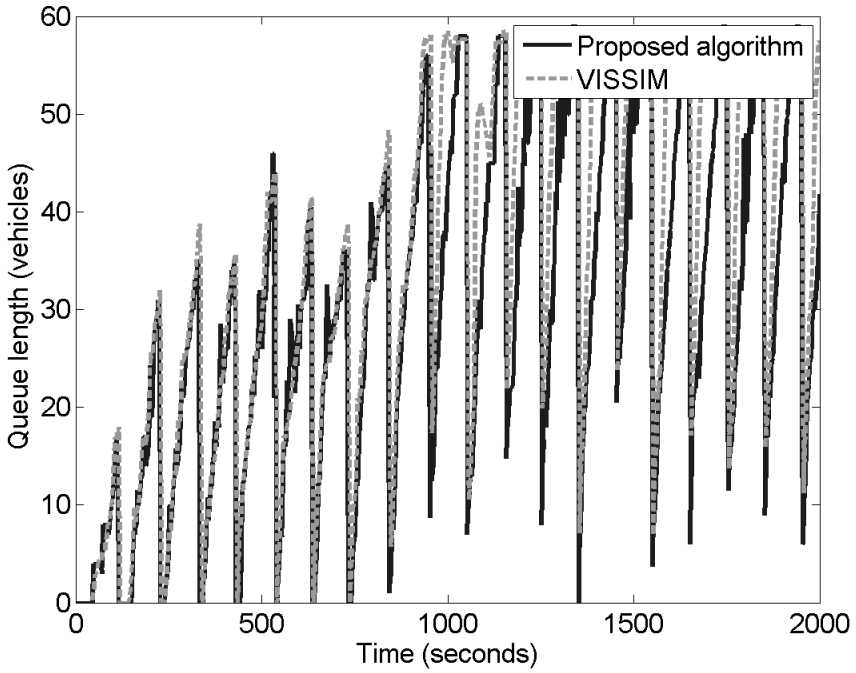

Figure 55 The estimated queue length when penetration ratio = 50%

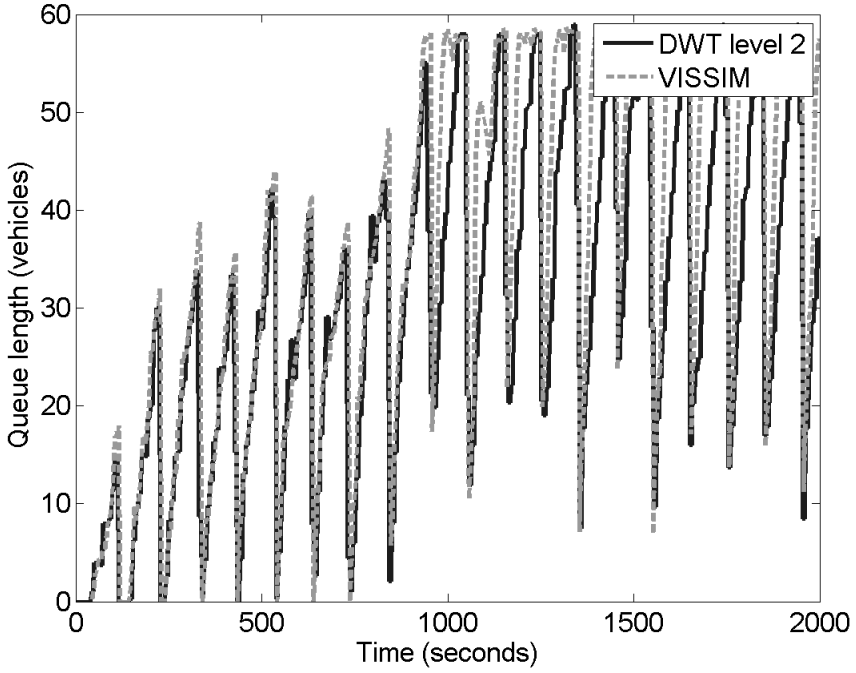

Figure 56 DWT level 2 when penetration ratio = 50%

The RMSE of Actuated signal with oversaturated condition is shown in Figure 57. The examples of time history results and the results with the best fit DWT levels are shown in Figure 58-Figure 64.

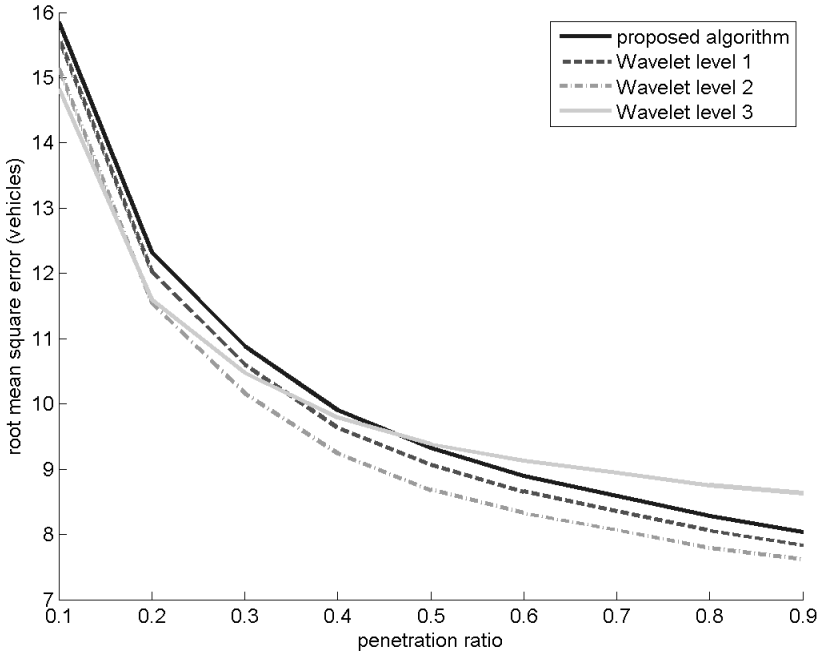

Figure 57 Root-mean-square error of queue length estimation in pre-timed with oversaturated condition case

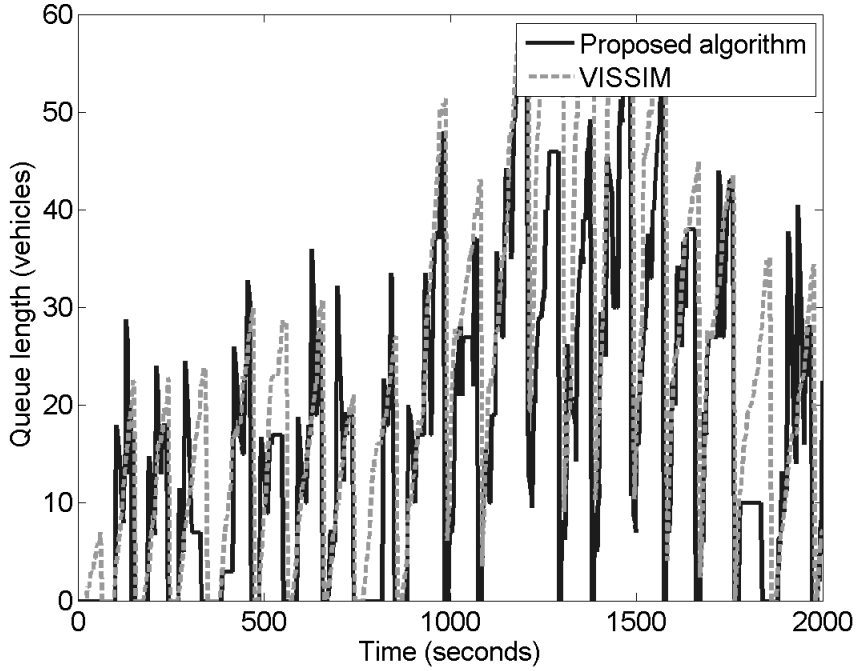

Figure 58 The estimated queue length when penetration ratio = 10%

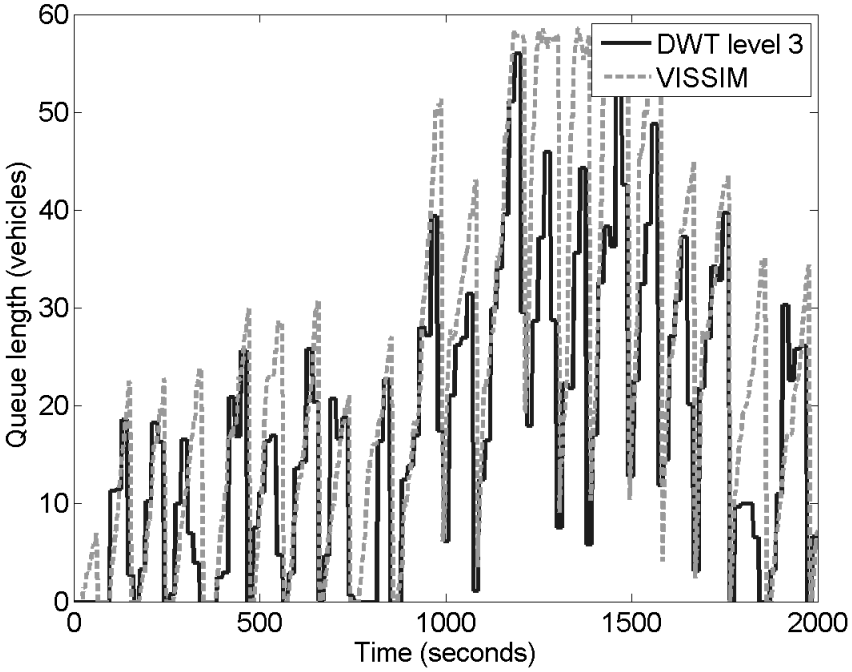

Figure 59 DWT level 3 when penetration ratio = 10%

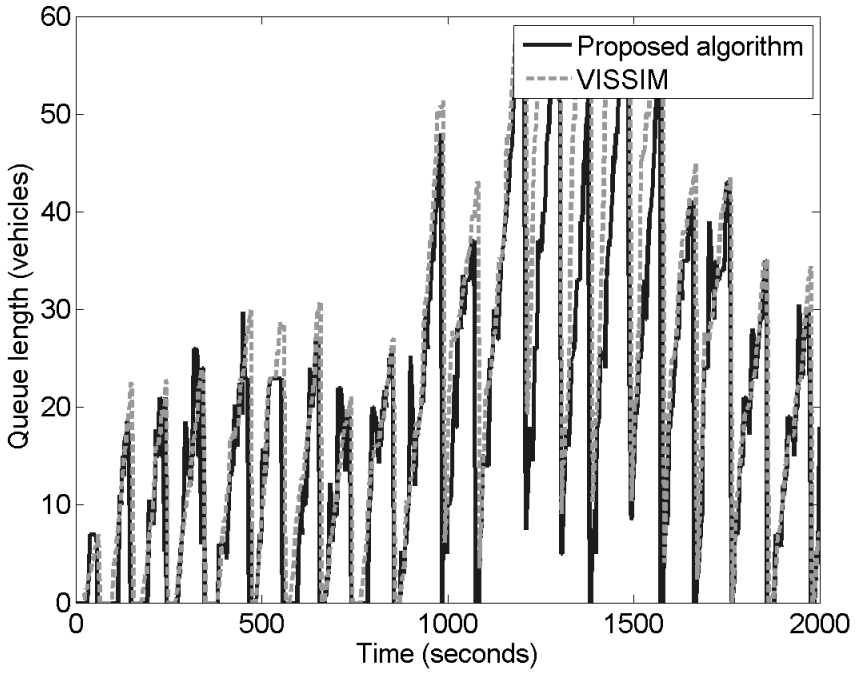

Figure 60 The estimated queue length when penetration ratio = 30%

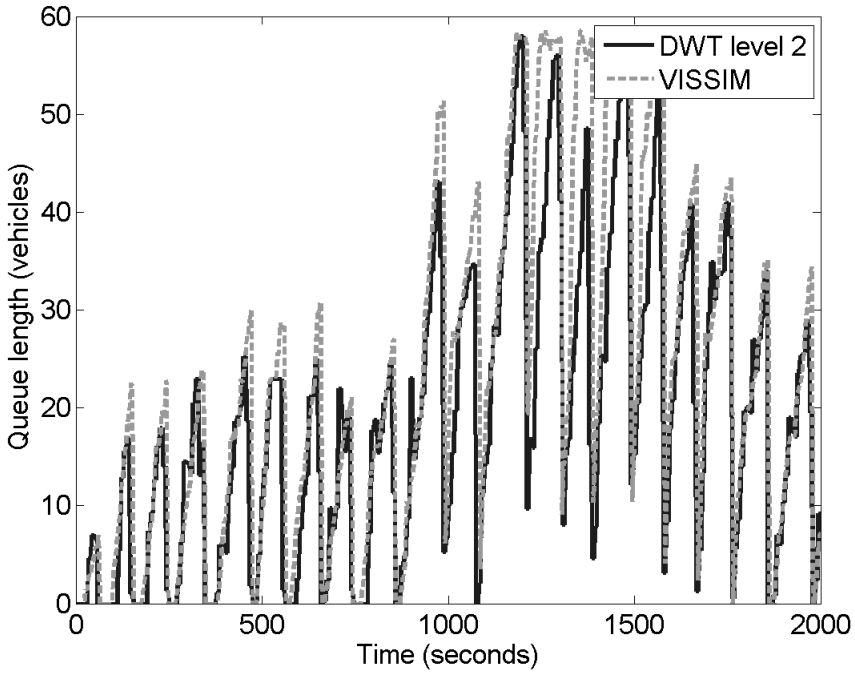

Figure 61 DWT level 2 when penetration ratio = 30%

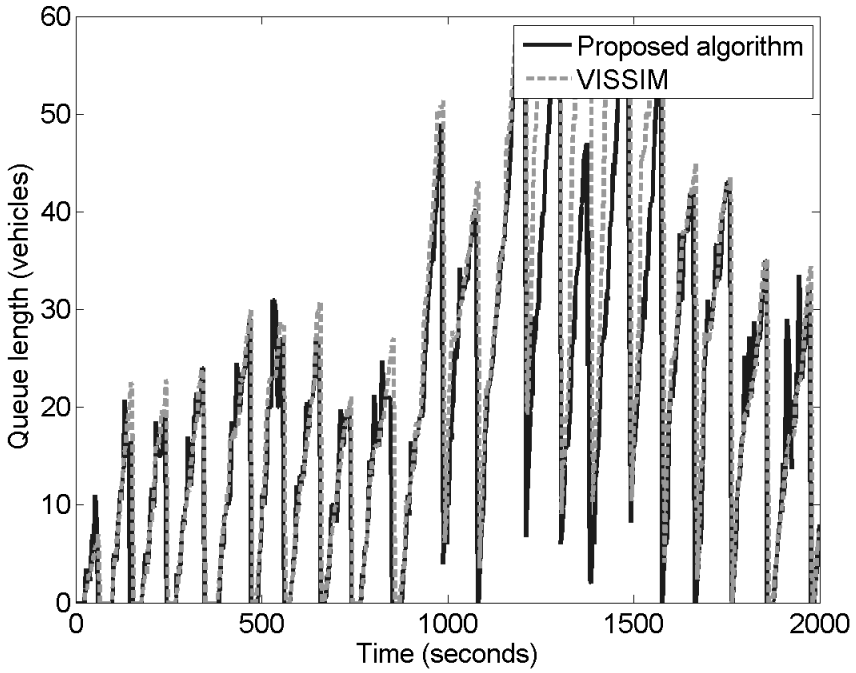

Figure 62 The estimated queue length when penetration ratio = 50%

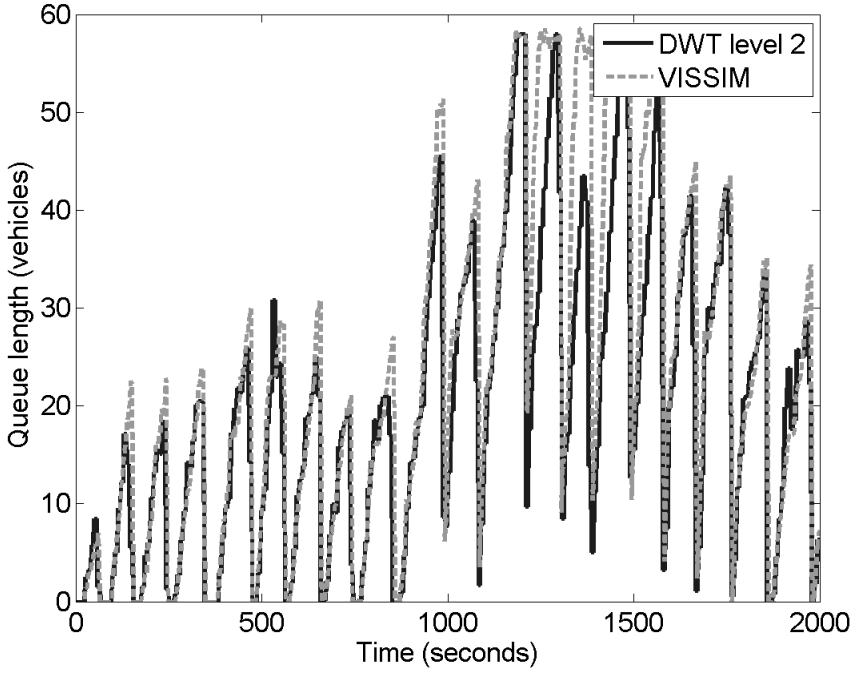

Figure 63 DWT level 2 when penetration ratio = 50%

The numerical results for Case C with two different stopping speed implemented by both VISSIM and the proposed algorithm are shown below

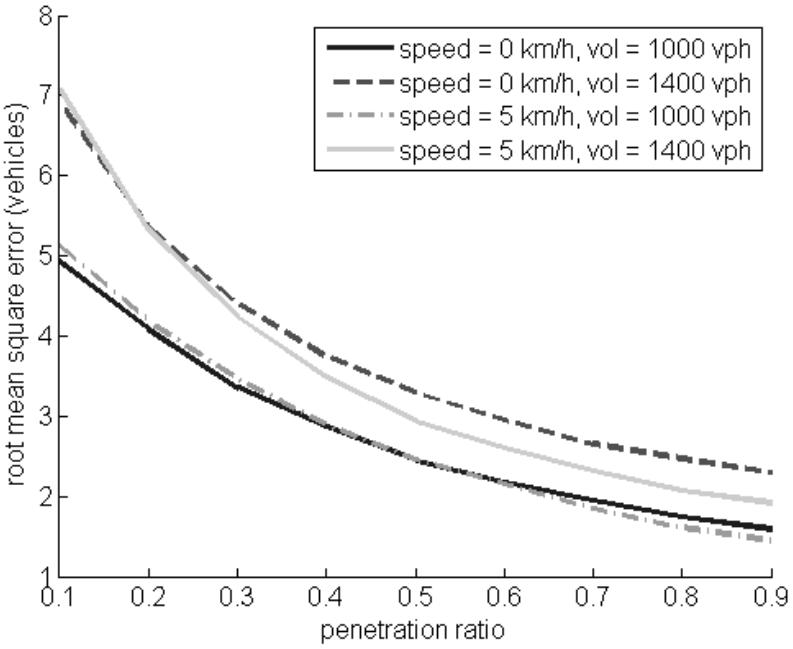

Figure 64 RMSE for Case C with two different stopping speeds

Average queue length for speed 0 km/h and 1000 vph, speed 0 km/h and 1400 vph, speed 5 km/h and 1000 vph and speed 5 km/h and 1400 vph are 3.7258, 6.5740, 3.9817 and 7.2057 vehicles. The performance of the algorithm did not change much with the change of the threshold. The RMSE of the predicted queue length does not change much with the increase of the threshold value.