STATE OF CALIFORNIA • DEPARTMENT OF TRANSPORTATION

TECHNICAL REPORT DOCUMENTATION PAGE

TR0003 (REV 10/98)

ADA Notice

For individuals with sensory disabilities, this document is available in alternate formats. For information call (916) 654-6410 or TDD (916) 654-3880 or write Records and Forms Management, 1120 N Street, MS-89, Sacramento, CA 95814.

1. REPORT NUMBER	2. GOVERNMENT ASSOCIATION NUMBER	3. RECIPIENT'S CATALOG NUMBER
CA11-2055		

4. TITLE AND SUBTITLE	5. REPORT DATE
Field Test Implementation of Queue Control	December 2011
	6. PERFORMING ORGANIZATION CODE

7. AUTHOR	8. PERFORMING ORGANIZATION REPORT NO.
R. Horowitz, P. Varaiya, R. Sanchez	

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. WORK UNIT NUMBER
California PATH 1357 South 46th Street, Bldg. 452 Richmond, CA 94804-4648	11. CONTRACT OR GRANT NUMBER 65A0320

12. SPONSORING AGENCY AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
California Department of Transportation Division of Research and Innovation, MS-83 1227 O Street, 5th Floor Sacramento CA 95814	Final Report 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

A major challenge for implementing queue control has been to accurately estimate on-ramp queue lengths, particularly during saturated on-ramp conditions, when the vehicle queue extends around or beyond the ramp entrance. The main outcome of the research was the development and field test verification of an accurate on-ramp queue length estimation algorithm based on vehicle re-identification, which performs reliably even under severe on-ramp saturated conditions. Four different queue estimation methodologies were studied using wireless magnetic sensors installed on a single lane loop on-ramp. Queue length estimation based on (i) occupancy measurements at ramp entrance, (ii) vehicle counts at on-ramp entrance and exit, (iii) speed measurements at ramp entrance, and (iv) vehicle re-identification were considered. These queue estimation methods were evaluated using available raw and processed sensor data, retrieved from the test site through mobile data communication and downloaded from a server. The accuracy and reliability of the queue estimation methods were studied using ground truth data obtained from video.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Field Test, Queue Control, Wireless Magnetic Sensors	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161

19. SECURITY CLASSIFICATION *(of this report)*	20. NUMBER OF PAGES	21. COST OF REPORT CHARGED
Unclassified	55	

Reproduction of completed page authorized.

FIELD TEST IMPLEMENTATION OF QUEUE CONTROL

TA-65A0320-15354

Final Report

Roberto Horowitz
Principal Investigator
Department of Mechanical Engineering
University of California, Berkeley, CA 94720
Tel: (510) 642-4675
horowitz@me.berkeley.edu

Pravin Varaiya
Principal Investigator
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, CA 94720
Tel: (510)642-5270
varaiya@eecs.berkeley.edu

Rene O. Sanchez
Graduate Student Researcher
Department of Mechanical Engineering
University of California, Berkeley, CA 94720
Tel: (510) 642-5109
r2sanche@me.berkeley.edu

December 2011

DISCLAIMER STATEMENT

This document is disseminated in the interest of information exchange. The contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California or the Federal Highway Administration. This publication does not constitute a standard, specification or regulation. This report does not constitute an endorsement by the Department of any product described herein.

For individuals with sensory disabilities, this document is available in Braille, large print, audiocassette, or compact disk. To obtain a copy of this document in one of these alternate formats, please contact: the Division of Research and Innovation, MS-83, California Department of Transportation, P.O. Box 942873, Sacramento, CA 94273-0001.

TABLE OF CONTENTS

ILLUSTRATIONS AND TABLES

ILLUSTRATIONS

TABLES

EXECUTIVE SUMMARY

A major challenge for implementing queue control has been to accurately estimate on-ramp queue lengths, particularly during saturated on-ramp conditions, when the vehicle queue extends around or beyond the ramp entrance. The main outcome of TA-65A0320-15354 was the development and field test verification of an accurate on-ramp queue length estimation algorithm based on vehicle re-identification, which performs reliably even under severe on-ramp saturated conditions.

Four different queue estimation methodologies were studied using wireless magnetic sensors installed on a single lane loop on-ramp. Queue length estimation based on (i) occupancy measurements at ramp entrance, (ii) vehicle counts at on-ramp entrance and exit, (iii) speed measurements at ramp entrance, and (iv) vehicle re-identification were considered. These queue estimation methods were evaluated using available raw and processed sensor data, retrieved from the test site through mobile data communication and downloaded from a server. The accuracy and reliability of the queue estimation methods were studied using ground truth data obtained from video.

The occupancy queue estimation method may be used to determine if the ramp is either empty or full, but it cannot be used to estimate the vehicle queue length accurately. This approach is highly dependent on the time calculation interval over which occupancy is calculated, and under on-ramp saturated conditions may yield misleading results due to vehicles' tendency to miss the sensor detection zone while at rest.

The queue length based on vehicle counts from wireless magnetic sensors, as it has been the case with loop detectors, is not a reliable or accurate method to estimate the queue length due to its inability to correct for errors due to detector miscounts and offsets resulting from initial conditions errors and drifts.

The speed based queue estimation method appears to be capable of instantaneously determine the mode of the on-ramp queue: unsaturated, saturated, or in transition. Nevertheless, based on the results obtained in this study, this queue estimation approach is not suitable to accurately estimate queue lengths.

The vehicle re-identification queue estimation method counts vehicles that enter and leave the on-ramp, while relying on vehicle re-identification to prevent drifts in the queue estimate. This queue estimation method performed better that the other three methods, but it initially under-performed in estimating queue lengths during saturated on-ramp conditions. This was due to the fact that the vehicle matching rate was very low during congestion. As a consequence of the original results obtained with this method, a modified vehicle re-identification algorithm was formulated and tested. With the new re-identification algorithm, the vehicle matching rate was significantly increased. A higher matching rate allows for a better and rapid correction of errors in the queue estimate. With the improved vehicle re-identification algorithm, the error remains bounded and the queue estimate is able to track the ground truth queue length with sufficient accuracy, even under congested on-ramp conditions.

Based on the observations at the on-ramp test site, it was also possible to point out some of the main factors that affect the performance of the different queue estimation methods considered for this project.

INTRODUCTION

The work presented in this report was conducted as part of a queue control field test. This project was proposed in order to implement queue control on the Hegenberger Rd. Loop on-ramp to 880 southbound in the Caltrans Bay Area District (FIGURE 1) to study its effect in minimizing queue and mainline density oscillations and enhancing performance. The queue regulation method to be used for the field test showed the potential to improve ramp storage utilization with respect to queue override, which is the method currently used in California freeways to regulate queues.

In the queue override scheme, occupancy measured with an inductive loop at the entrance of the on-ramp is used to regulate the queue and avoid its interference with arterial streets. This scheme is equivalent to an integral control with saturated integrating rate and resetting, which can be easily shown not to be asymptotically stable, given that the queue-length dynamics is a simple integrator. If the queue length could be measured, an asymptotically stable PI regulator could be designed and use to stabilize the close loop queue dynamics. However, the PI regulator needs the current queue length as its feedback, which unfortunately is not available in the field. Therefore, a suitable queue length estimator has to be designed using available measurements from the on-ramp, such as occupancy, vehicle speed measured by the queue detector, vehicle counts at the entrance and at the exit of the ramp, etc, in order to be able to implement this type of queue control.

The field test was designed to be completed by using a 2070 traffic controller running a modified Universal Ramp Metering Software (URMS), which is a recently developed program that allows the 2070 traffic controller to function as a ramp metering controller for use throughout California. Queue estimation and control as well as ramp metering algorithms were to be incorporated in the functionality of the URMS. These modifications would allow the use of queue regulation and control on any 2070 traffic controller running the URMS, and permit an actual field test of this control scheme on the Hegenberger on-ramp.

The field test proposed in TA-65A0320-15354 was designed to answer the following questions:

1. What is the best technique for estimating queue-length in an on-ramp?
2. Does the simple queue override produce oscillations in the length of the queue?
3. Are these oscillations attenuated by the Queue Controller?
4. Does the Queue Controller decrease the incidence of overspilling on-ramp queues?
5. Does the Queue Controller decrease the frequency of the override, and thus enhance coordination?
6. What data and hardware requirements are necessary to implement the queue regulator?

The criteria for the test site selection included:

1. the ramp is under local control,
2. it has functioning queue detectors,
3. it is subject to heavy demand and long on-ramp queues, leading to frequent queue overrides.

The project was designed to be completed in three stages. For the first stage of the project, which includes the first four tasks of TA-65A0320-15354, different queue estimation methods were to be evaluated. To complete the tasks in this stage it was not necessary to disrupt the current metering strategy at the Hegenberger on-ramp or any traffic operation conducted by District 4 at the ramp. For the second stage, which includes tasks 5 and 6 of TA-65A0320-15354, the queue-length estimator and regulator were to be tested at the on-ramp. These algorithms were going to be coded into the URMS, and used together with the current URMS traffic responsive ramp metering controller. Finally, the objective of the third stage of the project, which included task 7 and 8 of TA-65A0320-15354, was to implement a queue-length estimator and regulator together with ALINEA, a close loop local ramp metering strategy, at the Hegenberger on-ramp.

The original tasks of the project were not completed due to unforeseen circumstances. Out of the 3 project stages proposed in TA-65A0320-15354, only the first stage was finished. The results of the first stage of TA-65A0320-15354 are presented in this report.

SCOPE OF WORK

The scope of work consisted of the following tasks:

Tasks

1. Install the Sensys wireless vehicle detection system at the on-ramp test site.
   Status: Completed.

2. Develop a queue-length estimator based on car velocity measurements upstream of the on-ramp.
   Status: Completed.

3. Develop a queue-length estimator using vehicle re-identification algorithm.
   Status: Completed. This task took significantly more time than expected, since the original queue estimation results using this method were very inaccurate and needed improvements.

4. Test simultaneously the two queue-length estimators on the on-ramp test site.
   Status: Completed. For this task, queue estimation methods based on occupancy at the entrance of the ramp and based on counting vehicles entering and leaving the ramp were also included.

5. Modify the URMS software to incorporate the queue regulator with one of the two queue-length estimation methods.
   Status: Partially Completed. The URMS was studied, and it was identified which modifications were required to complete this task. However, due to the delays explain in the next section, we were not able to implement the modifications in the URMS.

6. Test the queue-length estimation and regulation algorithm with the current URMS traffic responsive ramp metering controller on the on-ramp test site.
   Status: Not Completed.

7. Make a theoretical stability analysis of the algorithm that combines ALINEA with the queue controller.

Status: Completed.

8. Test ALINEA with the queue-length estimation and regulation algorithm on the on-ramp test site.

Status: Not Completed.

9. Prepare workshop presentation and final report as well as the Prototype of the Modified URMS code that includes ALINEA and Queue Estimation and Regulation.

Status: Partially Completed.

REASONS FOR UNCOMPLETED TASKS

It was not possible to complete all the tasks proposed for TA-65A0320-15354 due to the following unanticipated delays that were beyond our control:

- Funding was delayed
- Original equipment installation required revision
- Encroachment permit processing time
- Extended University winter recess
- Weather conditions
- Queue estimation methods initially performed poorly and it was necessary to modify the queue estimation method based on vehicle re-identification
- The Universal Ramp Metering Software (URMS) was not approved for ramp metering field testing by Caltrans District 4

Though the project work was scheduled to begin on 6/26/09, the official award notice and full budget of 152,068.00 was not granted until 7/21/09. There were also logistical delays.

At a July 2nd 2009 meeting between Alan S. Chow, Sean Coughlin, Roberto Horowitz, Rene Sanchez and other Caltrans personnel that took place at the District 4 Office, the installation plan and field test implementation of queue control project were presented and discussed and it was determined that certain aspects of the sensor installation plan required further analysis and modifications. In the following weeks alternative options for the sensor installation were considered until agreement was reached on an installation plan that met District 4 regulations and guidelines. Though an encroachment permit application package was prepared and submitted to the District 4 office of permits on September 30, 2009, it took them a month to process it and it was not possible to start with any of the tasks that required access to the Hegenberger Rd. Loop on-ramp to S/B 880 until the permit was approved. After further minor modifications to the installation plan, the encroachment permit was approved on November 4, 2009. It was only then that Republic ITS, the company that conducted the sensor installation, was able to initiate their double encroachment permit application and were able to coordinate with Caltrans for on-ramp closure and traffic rerouting. Installation was further delayed by the intervening winter break during which the University business operations shut down for an

extended number of days due to furloughs. This delayed the processing of the necessary purchase orders. The installation finally took place on January 27 2010, but only after further rainy weather related delays and the rescheduling of road closure and traffic rerouting.

On the scientific side, data obtained from the Sensys vehicle detection system originally showed unrealistic queue lengths. There was a lack of correlation between queue length based on the vehicle re-identification method and velocity measurements at the entrance of the Hegenberger on-ramp. Ground truth data was collected with video cameras and used to determine the accuracy and reliability of the queue estimation methods. There was a lot of time invested in improving the queue estimation methods in order to be able to use one for the second and third stages of the project. This effort resulted in an improved vehicle re-identification algorithm that allow for accurate queue estimation under uncongested and congested on-ramp conditions.

The project proposal for TA-65A0320-15354 relied on the assumption that the URMS was going to be approved in Caltrans District 4 (D4) for deployment on the field around the time we were ready to start with the second stage of the project. The plan was to modify specific modules of the software to incorporate the queue regulation and ramp metering algorithms for testing in the Hegenberger on-ramp. However, the URMS approval process from District 4 took more time than expected, and it was not possible to get a version of the URMS approved for deployment in order to conduct the queue regulation field test. The second and third stages of the project relied on this approval process, which did not occur.

BACKGROUND

Ramp metering is an effective traffic control strategy for managing freeway congestion (*1*). It consists on restricting the flow of vehicles into the mainline with the objective to improve freeway efficiency, leading to an overall system performance improvement. At peak traffic times, ramp metering effectiveness directly relates to the amount of space available at freeway ramps to store vehicles that otherwise will enter the mainline, and contribute to further deteriorate traffic conditions. As a result, adequate implementation of ramp metering utilizes storage space available in the freeway system while avoiding on-ramp vehicle queue interference with arterial street traffic (*1*).

In freeways, ramp metering is usually employed along with queue control. This form of ramp metering regulates traffic conditions at the mainline while taking into account the vehicle queue formed at the on-ramps. The queue control part of the algorithm regulates the length of the queue as to maximize the utilization of storage available at the on-ramps during peak traffic hours while avoiding queues spilling into adjacent arterial streets. In order to implement queue control, it is necessary to have a reliable and accurate estimate of the vehicle queue length (*2*).

Queue control has been addressed in the literature and used extensively in traffic simulation studies. However, when ramp metering with queue control is implemented in the field, the performance of the queue control algorithm has been undermined by the queue estimation methods employed, which results in inaccurate regulation of the queue length that underutilizes on-ramp storage space (*2*).

The four methods that were considered for the field test are queue estimation (i) based on occupancy measured at the entrance of the on-ramp, (ii) counts of vehicles leaving and entering the on-ramp, (iii) queue estimation based on speed data at the entrance of the on-ramp, and (iv) queue estimation based on vehicle re-identification. The first method corresponds to the queue estimation method used with queue override. It consists of determining if the queue is at or beyond the queue limit by comparing the occupancy measured at the entrance of the ramp with an occupancy threshold. The second method counts vehicles entering and leaving the on-ramp. The difference between the counts gives the queue length, provided the initial number of vehicles at the ramp when the counting is initiated can be accounted for. This queue estimation method has been explored in detail, and has been known to introduce errors due to the inability of the method to correct for offset (*3*), (*4*), (*5*). Work has been done in order to improve the performance of this queue estimation approach by incorporating occupancy measurement from the ramp (*5*) (*6*) (*4*). The third method is a queue-length estimator based on a simplified model for the driving behavior of a vehicle that is approaching the end of the queue: the vehicle decelerates at a constant rate from its cruising speed, until it stops. This method was proposed in (*2*) and assumes it is possible to accurately calculate queue length from vehicle speed measurements at the entrance of the on-ramp. The fourth method estimates the length of the queue using a vehicle re-identification algorithm, as described in (*7*). This scheme is based on matching individual vehicle signatures obtained from wireless magnetic sensor arrays placed at the two ends of the on-ramp. It relies upon counting vehicles entering and leaving the on-ramp and corrects for errors using the vehicle re-identification algorithm. Vehicle re-identification has been validated on arterial streets and has yielded satisfactory results; nevertheless its performance has not been studied at on-ramps until this project.

The body of this report is organized as follows: the test site is described in Section 1. The vehicle detection system is presented in Section 2. Section 3 discusses the different types of data used in this study. The ground truth data collection method is explained in Section 4. Section 5 contains the results obtained with the different queue estimation methods while Section 6 focuses on the improved queue estimation method based on vehicle re-identification. Section 7 discusses factors affecting different queue estimation methods. Section 8 presents the work completed in relation to task 5, while Section 9 presents the analysis related to task 7.

FIGURE 1 Hegenberger on-ramp (a) aerial view with saturated queue, (b) side view with saturated queue and (c) arterial streets and intersections around the Hegenberger on-ramp.

BODY

1. TEST SITE

The Hegenberger on-ramp is suitable for the field test because it is under local control, has functioning queue detectors, and it is subject to heavy demand and long on-ramp queues that frequently go beyond the on-ramp capacity. Furthermore, it is a single lane ramp, which allows for the testing of the queue estimation methods without having to take into account multiple lanes on-ramp dynamics (e.g. lane changing).

Wireless magnetic sensors are commonly employed in arterial streets or freeway mainlines, in which lanes are straight and cars tend to travel in the middle of the lane. This is not the case with loop on-ramps like the one used in our study. The difference in height and the curvature of the ramp, as seen in FIGURE 1 and FIGURE 2, differs from the traditional lane geometries observed at arterial streets or at freeway vehicle detection stations, for which wireless magnetic sensors performance have been studied (*8*) (*9*) (*7*). The use of these sensors in loop on-ramps is something that has not been investigated in detail, which is another reason why the Hegenberger on-ramp was a convenient test site to conduct queue estimation methods analysis.

The Hegenberger on-ramp length is 616 ft (187.8 m), which corresponds to a capacity ranging from 17 to 25 vehicles. The capacity varies due to vehicle type and inter-vehicle spacing or headway, which fluctuates throughout the day. Note that for this study, the analysis is concerned with queue estimation performance from the entrance to the exit of the ramp (see FIGURE 2(a)). When the ramp was saturated and vehicles were beyond the ramp entrance, the ground truth queue lengths presented in this analysis corresponds to the number of vehicles located between the exit and the entrance of the ramp.

2. VEHICLE DETECTION SYSTEM

The vehicle detection system deployed at the Hegenberger on-ramp was developed by Sensys Networks. This system consists of an Access Point (AP240-ESG), a repeater (RP240-B), and 18 wireless magnetic sensors (VDS240) installed as shown in FIGURE 2 (b) and (c). Two arrays of 7 sensors separated by 1 ft, and installed at the center of the lane, are located at the entrance and at the exit of the on-ramp. These entrance and exit arrays were necessary for the implementation of queue estimation based on vehicle re-identification, see (*7*) for more details. These arrays are also used for the analysis of the queue estimation methodology using vehicle counts, even though a single detector placed at the center of the lane at each end of the ramp would have been enough. Four sensors are installed at the entrance of the on-ramp and are arranged in a speed trap configuration. Two sensors were used for leading vehicle detection (SL1 and SL2), and another two for trailing vehicle detection (ST1 and ST2). This configuration increases the lateral detection zone to capture vehicles that may be traveling off-centered of the lane. The speed trap sensors were used to study the queue estimation methodology based on speed and occupancy and to compare the vehicle counting performance of a single sensor and a sensor array. See (*8*) for details on this vehicle detection system.

3. VEHICLE DETECTION SYSTEM DATA

For the analysis presented in section 5, two different types of data were used.

Processed Data

The processed data was obtained using the Sensys SNAPS server, which provides connectivity to the equipment installed at the on-ramp. This software computes counts, speed, occupancy, vehicle length, and other statistics over various time intervals, as well as real-time individual vehicle speed, length and headway using raw data from the sensors (*8*). It is also possible to obtain the queue length data based on vehicle re-identification in the same way.

The data processing application does not have the capability to Or-gate, i.e. automatically merge and treat as a single detector, the leading sensors (SL1 and SL2 shown in FIGURE 3) or the trailing sensors (ST1 and ST2 shown in FIGURE 3). This means that the data processing application treats Sensors SL1 and ST1 as one speed trap and sensors SL2 and ST2 as another one. Speed trap data processing is done independently for the SL1 and ST1 and for the SL2 and ST2 sensors.

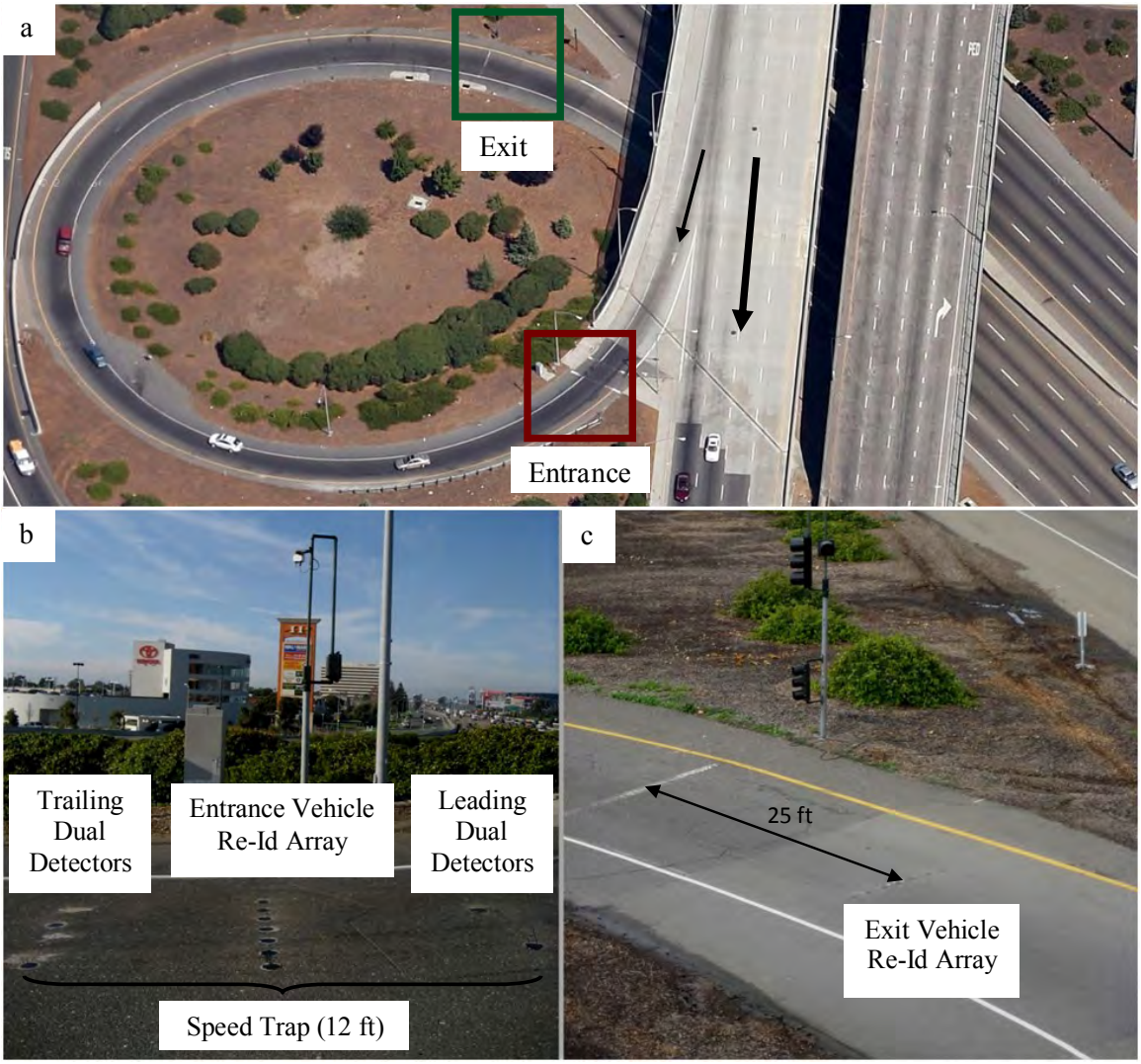


FIGURE 2 (a) Hegenberger on-ramp entrance and exit. (b) Vehicle detection system installation at the entrance of the ramp. (c) Sensor array installation at the exit of the ramp.

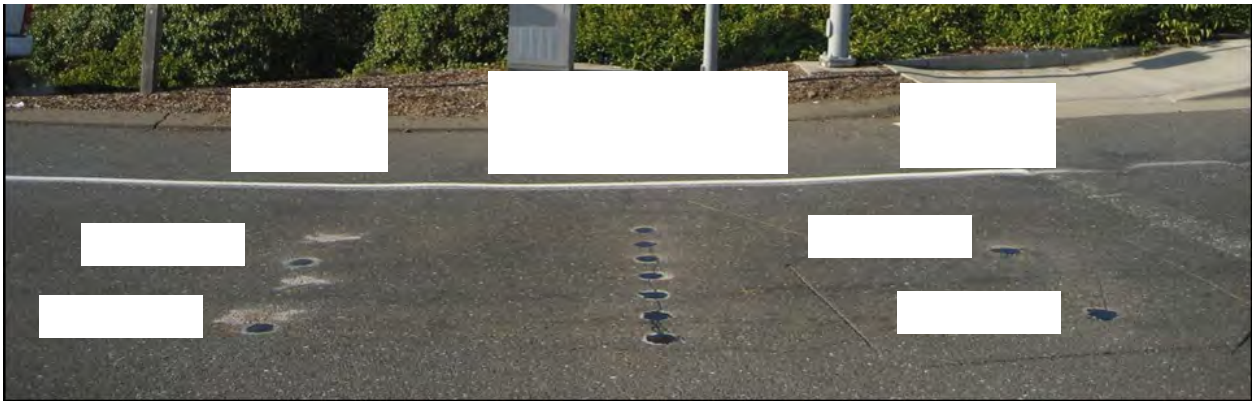


FIGURE 3 Wireless magnetic sensor configuration at the entrance of the ramp.

Raw Data

The raw data obtained from the sensors were used to analyze different sources of errors. The arrangement of the sensors in the field and the configuration mode of each of the sensors determine the type of raw data, e.g. speed trap (mode b) raw data or re-identification array (mode f) raw data. Data from any of these sensor modes can be used to derive counts and occupancy data. Accurate speed estimation is possible with the speed trap sensors' raw data, and vehicle re-identification is possible with the array sensors generating detailed raw data from which vehicle magnetic signatures are obtained.

Mode B sensor raw data

Each speed trap sensor (FIGURE 3) generates the following information every time a sensor event occurs.

```
<sensor_id> <unix epoch time> <event>
```

An `event` value of 0 is the transition from detect to undetect, a value of 1 is the transition from undetect to detect. Each of the sensors in the speed trap can be used independently to obtain vehicle counts and occupancy.

Mode F sensor raw data

The entrance and exit array raw data is generated every time a vehicle goes on top of the arrays. The data generated by the arrays and used for this analysis contains the following information:

```
<array_id> <undetect time> <duration over sensor> <queue(real time)>
```

`Undetect time` corresponds to the instance in which the vehicle is no longer detected by any of the sensors in the array. `Duration over sensor` indicates how much time at least one of the sensors of the array was active due to vehicle presence. `Queue(real time)` is the queue length calculated with the vehicle re-identification method in real time. Notice that the raw data is excluding the vehicle magnetic signature portion, which is also generated every time a vehicle goes on top of the arrays and used in the vehicle re-identification algorithm.

4. GROUND TRUTH DATA COLLECTION

In order to get ground truth data to validate the queue estimation methods, vehicles at the on-ramp were videotaped. Video was recorded on 3 different days, April 13, May 5, and May 11, 2010. Observations on those three times were important for the conclusions drawn for this analysis, however, the ground truth data that is presented in section 5 corresponds to the day of May 11, 2010 from 4:07 pm (16.1 hours) to 5:35 pm (17.6 hours), when ramp metering was active. The data collected on this day is adequate for the analysis presented in section 5 because different queue lengths (zero, intermediate sized as well as capacity) were observed during the recording time. This allows for the analysis of the queue estimation methods under saturated and unsaturated on-ramp conditions.

Three independent cameras were used to obtain the ground truth data. The first camera recorded vehicles waiting at the metering light and/or leaving the on-ramp and passing the exit sensor array (FIGURE 4 (a)). The second camera recorded vehicles entering the on-ramp and passing through the speed trap and the entrance sensor array (FIGURE 4 (b)). The third camera was not fixed at any specific location, and it was used to capture queue dynamics during the ground truth data collection period (FIGURE 4 (c)). The three cameras were synchronized with a common clock so that data extracted from different videos could be merged, or could be meaningfully compared with the vehicle detection system data.

Ground truth queue length was obtained with camera 1 and camera 2. From the video recorded with these cameras, it was possible to extract the time at which the frontal part of all the vehicles entering and leaving the ramp was aligned with the entrance and exit sensor array, respectively. Ground truth queue length was calculated by adding and subtracting unity to the queue length every time a vehicle was registered entering or leaving the on-ramp, respectively. The queue length was initially set to zero, since at 4:07 pm there were zero vehicles at the ramp. Motorcycles and vehicles bypassing the queue were recorded in the ground truth data set.

Ground truth data for queue estimation based on vehicle speed was obtained with the second video camera. The time instances of vehicle alignment with leading dual sensors (FIGURE 3, SL1 and SL2) and trailing dual sensors (FIGURE 3, ST1 and ST2) were extracted. The ground truth speed was calculated by dividing the distance between the leading and trailing sensors (12 ft (3.7m)) over the difference between the two time stamps extracted from the video for each vehicle. Camera 2 is a cassette video camera with a recording rate of 30 frames per second, which affects the resolution of the speed ground truth data. The largest discrepancies between ground truth data and the vehicle detection system data are expected at larger speed due to quantization.

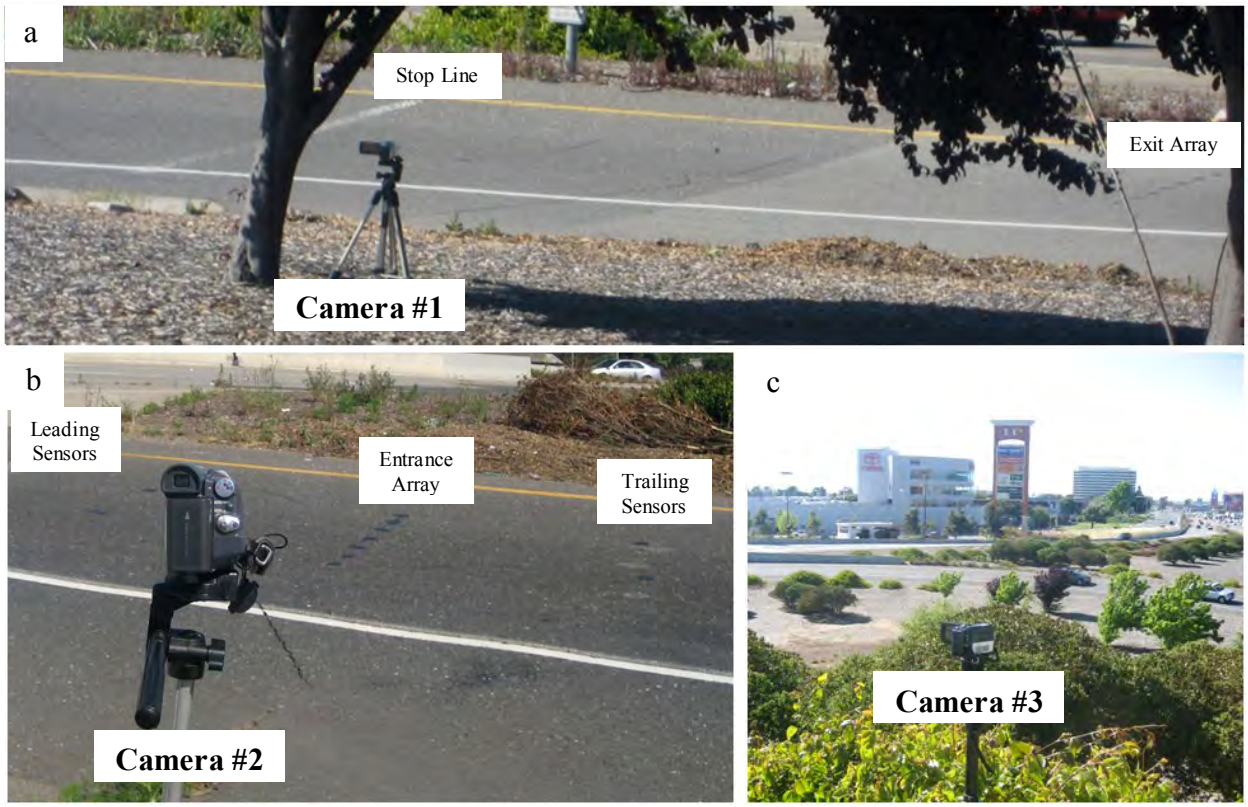


FIGURE 4 (a) Video camera recording vehicles leaving the on-ramp. (b) Cassette video camera recording vehicles entering the on-ramp. (c) Video camera to capture vehicle behavior and queue dynamics.

5. QUEUE ESTIMATION RESULTS

Occupancy based queue estimation method

Occupancy calculated with mode b speed trap sensors (sensor SL1 and sensor ST1) over a 30 second calculation interval is shown in FIGURE 5 (top). Before 17 hours, occupancy magnitude correlates with the ground truth queue length. Higher occupancies are observed when queue lengths are larger and vice versa. However, under on-ramp saturation conditions, after 17.2 hours, very low occupancies are observed even though the queue length was around or beyond the on-ramp entrance. This phenomenon, a so-called ZSZO (zero-speed zero-occupancy) situation (*6*), makes queue estimation based on short calculation interval occupancy unreliable.

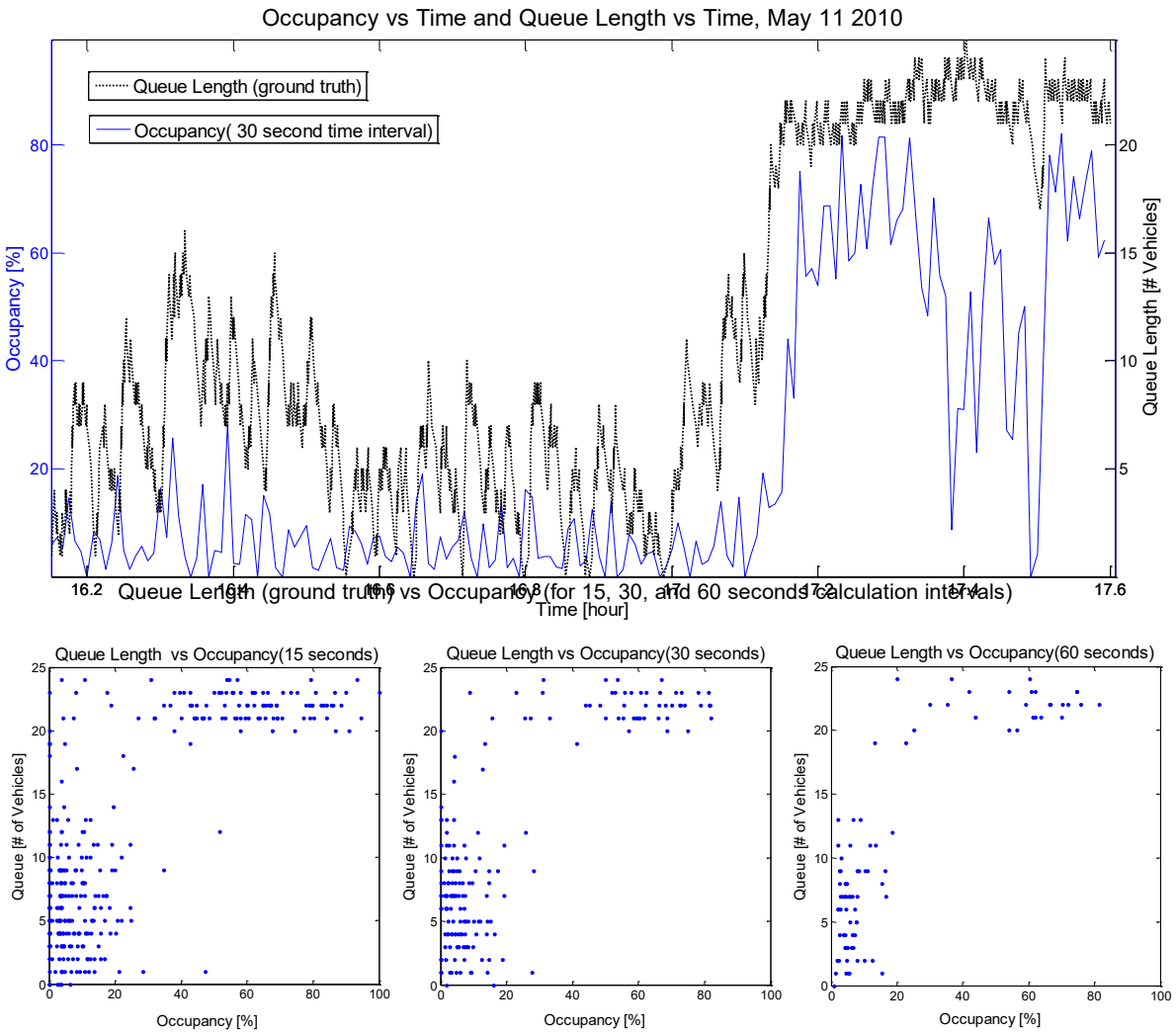


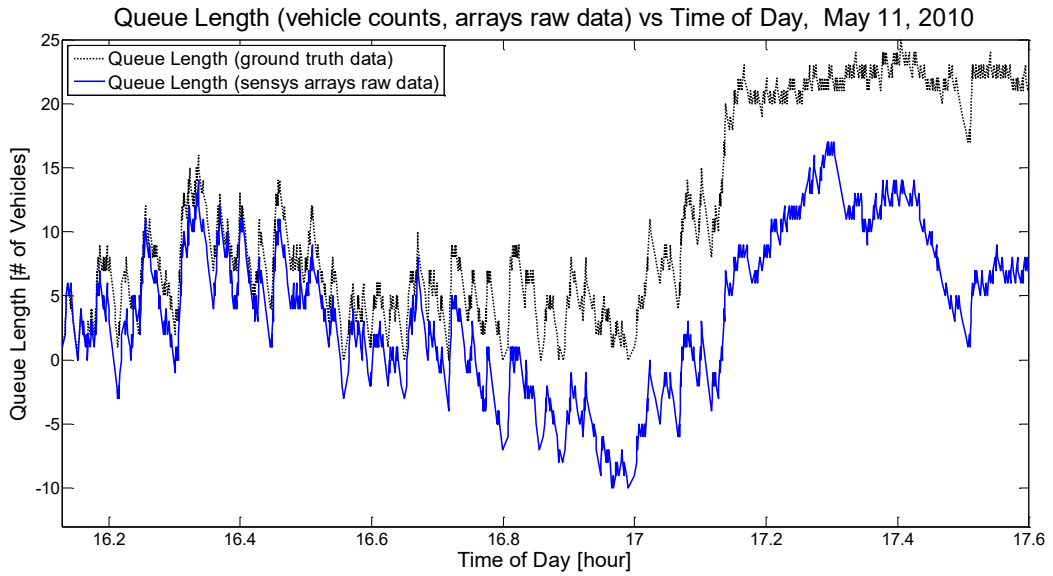
FIGURE 5 (top) Occupancy and queue length as a function of time of day. (bottom) Scatter plots of queue length versus occupancy for 15, 30 and 60 seconds time intervals.

FIGURE 5 (bottom) shows the scatter plots of queue length versus occupancy for multiple calculation intervals. In these scatter plots it is evident that occupancy measured at the entrance of the ramp may only be used to predict if the on-ramp is either saturated or unsaturated using an occupancy threshold as a reference. Based on this scatter plots, less data points are present in the upper left corner of the plot as the calculation interval is increased. This shows that the accuracy (or applicability) of the occupancy threshold method increases with larger calculation intervals (e.g. less chances of ZSZO-phenomenon to occur). However, increasing the calculation interval introduces a delay in detecting queue saturation (see FIGURE 5(top) between 17.1 and 17.2 hours), which is an undesirable queue estimation characteristic for accurate queue control.

Vehicle Counts based queue estimation method

Arrays vehicle counts

543 vehicles entered and 522 left the on-ramp during the ground truth data recording period. The raw data coming from the sensor arrays registered that 527 vehicles entered and 520 left the on-ramp during the same time period. It would seem that the accuracy of the counting method is adequate, with only 3% error for entering vehicles and less than 1% error for exiting vehicles. This conclusion would be erroneous; when the queue estimation is calculated as a function of time and compared with the ground truth, it is clear that there exists vehicle under and over-counting.



	ENTRANCE	*EXIT*
TOTAL VEHICLES (Ground Truth)	**543**	**522**
Trucks	21	20
Motorcycles	1	1
TOTAL VEHICLES (Arrays)	**527**	**520**
Repeated Detection	**21**	**15**
due to Trucks	9	5
Undetected Vehicles	**40**	**19**
due to merging into one detection event	11	0

FIGURE 6 (top) Queue length based on ground truth data and arrays' raw data. (bottom) Total vehicle counts for the ground truth and the arrays raw data between 16.13 hours and 17.6 hours.

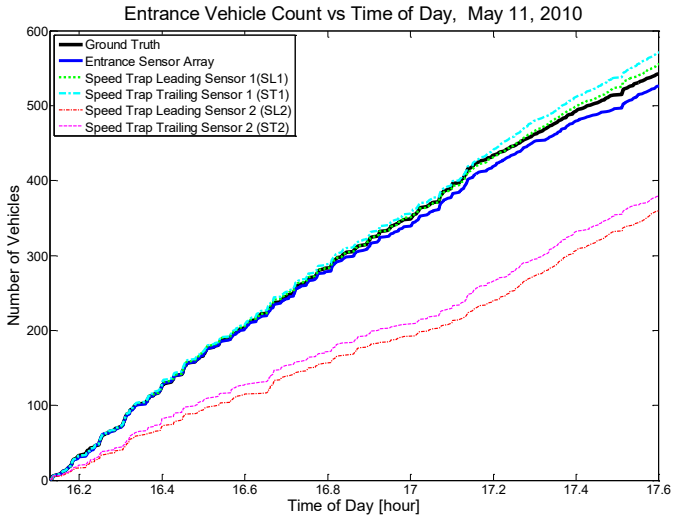
FIGURE 6 (top) shows that there exist large discrepancies between the estimated and ground truth queue lengths. At 17 hours, the estimated queue reaches a value around -10, which suggests considerable offset due to miscounting even before ramp saturation. When the ramp is

saturated, which occurs after 17.1 hours, the error between both queue lengths is more pronounced. When the raw data events are analyzed (FIGURE 6 (bottom)) it is evident that there are some vehicles that are not detected, while others are counted multiple times (e.g. trucks). Note that the entrance array has twice as much undetected vehicles as the exit array, from which 11 were undetected due to merging. During congestion, cars tend to travel slower and closer together as they move through the entrance array, which sometimes results in multiple cars being registered as a single vehicle. The exit array does not present any case in which multiple vehicles are merged into a single detection event, since there is significant spacing between vehicles as they exit the ramp because the exit array is located after the ramp metering light.

Entrance Vehicle Count Comparison

Mode b sensors were not installed at the exit of the on-ramp. They were only installed at the entrance of the ramp as part of the speed trap sensor arrangement (SL1, SL2, ST1 and ST2). As a result, it is not possible to calculate queue length using only mode b sensors. Nevertheless, it is possible to investigate the reliability of mode b sensors for vehicle counting by comparing vehicle counts gathered using the speed trap sensors to the entrance array and the ground truth vehicle counts.

The comparison between the entrance vehicle counts based on speed trap sensors, the entrance array, and ground truth is shown in FIGURE 7 (left). If cars traveled through the middle of the lane, a similar count would be expected from all the speed trap sensors. However, it was observed that vehicles tend to travel on the right side of the lane, which is reflected in FIGURE 7, since sensors SL2 and ST2, which are located on the left side of the lane, register significantly less vehicle counts over the complete time interval. The vehicle counts discrepancies for sensors SL1 and ST1 and the ground truth vehicle counts becomes evident only after the 17.1 hours, which corresponds to the time when the ramp goes into saturation mode. It seems that congestion affects the counting performance of both mode b and mode f sensors.



	Total Vehicle Counts
Entrance GT	543
Entrance Array	527
SL1	557
ST1	573
SL2	362
ST2	381

FIGURE 7 (left) Entrance vehicle counts based on ground truth, entrance array and SL1, SL2, ST1 and ST2 sensors. (right) Total entrance vehicle counts between 16.13 hour and 17.6 hour.

As with the sensor array data, single sensors sometimes register multiple detections for the same vehicle as well as a single detection event for multiple vehicles. Nevertheless, notice that sensors SL1 and ST1 total count exceeded the ground truth data, which is the opposite of the total entrance array vehicle count. This suggests that mode b sensors may be more likely to generate multiple detections for the same vehicle than mode f sensors.

For the most part, leading and trailing detectors on the same side of the lane were expected to have very similar vehicle counts; however, trailing sensors register higher total vehicle counts than the leading sensors. This suggests that sensor vehicle counting performance for mode b sensors is dependent on the lateral as well as the longitudinal location of the sensor in the ramp lane.

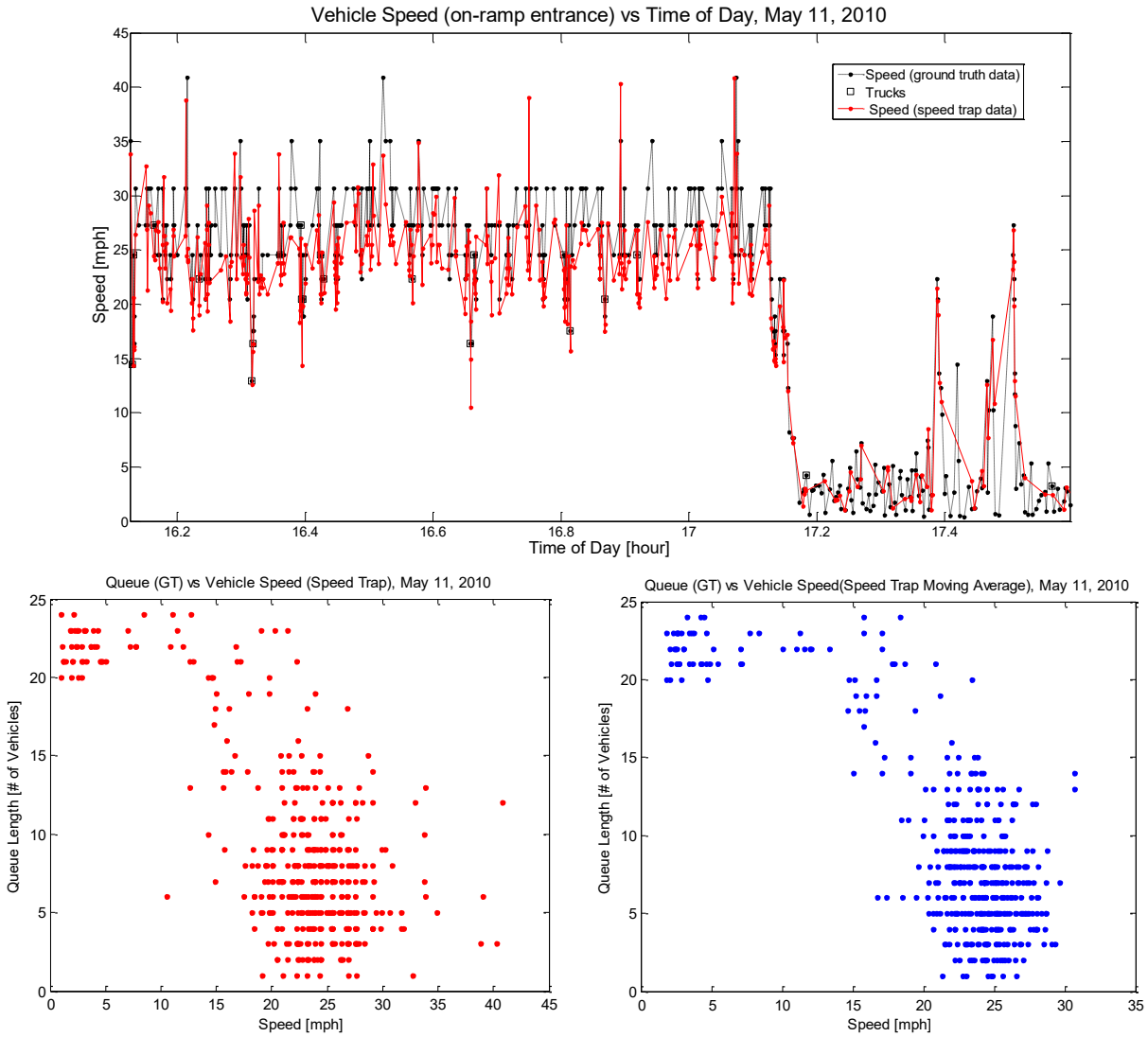


FIGURE 8 (top) Speed trap measurements and ground truth speed as a function of time of day. (bottom, left) Scatter plot of queue lengths versus speed at the entrance of the on-ramp. (bottom, right) Scatter plot of queue lengths versus 3 trailing point moving average of speed.

Queue Based on Speed

Before analyzing the queue estimation method based on vehicle speed measured at the entrance of the ramp, it was necessary to make sure that speed determined with the speed trap was accurate. FIGURE 8 (top) shows the comparison between the speed trap and the ground truth speed data sets. Even though both data sets correlate well, it was not possible to calculate a speed for all the vehicles that entered the on-ramp. The ground truth data set has 543 data points while the speed trap data set only has 425 data points, which means that it was not possible to determine the speed of 22 % of the vehicles that entered the on-ramp during the ground truth recording time. During on-ramp saturation, there are a larger number of consecutive vehicles for which speed was not calculated. FIGURE 8 (top) shows, at around 17.25 and 17.4 hours, clear examples of this problem. Inability to determine the speed of every entering vehicle affects the applicability of this method for queue estimation, since several minutes could pass before a vehicle speed can be calculated again to update the queue estimate.

Speed measurements correlate well with the on-ramp mode. When the queue is unsaturated, the speeds are larger, usually above 20 mph, and when the queue is large, speed tends to drop accordingly. However, there are multiple factors that result in speed drop at the entrance of the ramp, which are unrelated to the queue dynamics. This is the case for large trucks entering the on-ramp, which, independent of the queue length, usually slow down. FIGURE 8 (top) shows drops of speed at unsaturated queue conditions, before 17.1 hours, for most of the instances when trucks entered the ramp. Another factor that affects the speed of vehicles in a similar way, is the presence of pedestrians crossing the street and blocking the entrance of the ramp.

FIGURE 8 (bottom, left) is a scatter plot of queue length vs speed that shows that a wide range of entrance speeds correspond to a given queue length value. Taking a moving average of the speed data, as done on FIGURE 8 (bottom, right), decreases the range of speed that corresponds to a given queue length, but it does not allow to find a relationship between queue length and speed that could be used for precise queue control. It is not possible, based on these results, to obtain a relationship between queue length and speed as obtained in (*2*) using traffic simulation results. Nevertheless, using speed at the entrance of the ramp seems to be a good approach to determine if the queue is saturated, unsaturated or transitioning from one mode to the other in an almost instantaneous way, an improvement over using occupancy measurements at the entrance of the ramp.

Queue based on Vehicle Re-Identification

There were two queue estimates based on the vehicle re-identification method. The first one is provided in real time using sensor array raw data and the second one uses post processing offline.

FIGURE 9 (top) shows a comparison between the queue estimates calculated using vehicle re-identification in real-time and the ground truth queue length. During on-ramp uncongested conditions, both correlated properly, and even when some notorious errors appeared, e.g. around 16.43 hours, the method was able to compensate for the offsets. However, when on-ramp saturation occurred, the queue estimator performance dropped, and eventually collapsed around 17.2 hours. During congestion, vehicle counting is more inaccurate, as shown

earlier, and it becomes very difficult to make appropriate corrections when a vehicle is re-identified, since this method is relying on vehicle counting to correct for offsets. Despite its underperformance during saturation, this method seems to do better than the queue estimation method based only on counting cars, see FIGURE 6 (top).

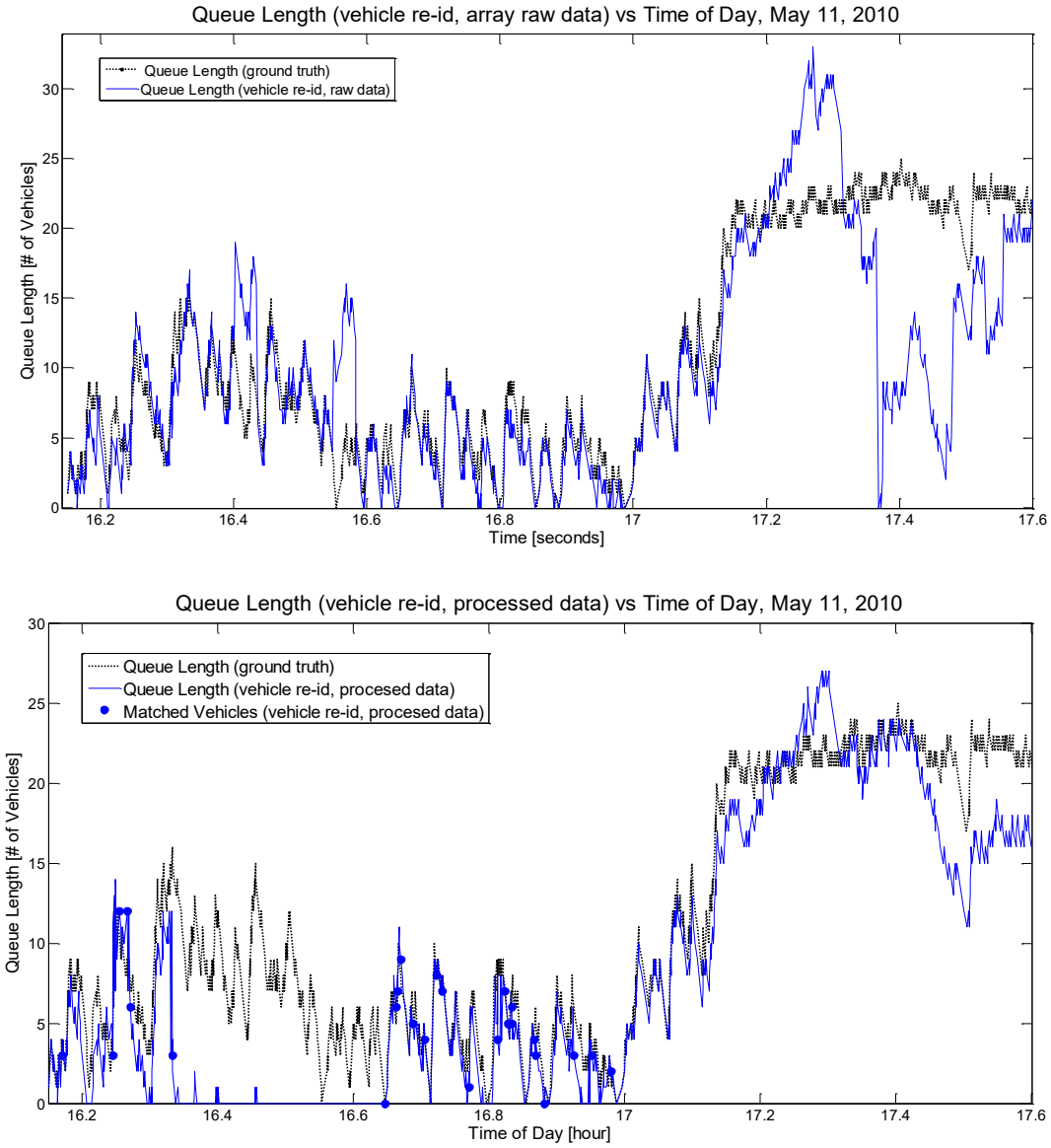


FIGURE 9 (top) Queue length based on real time vehicle re-identification versus time of day. (bottom) Queue length based on offline vehicle re-identification algorithm versus time of day.

FIGURE 9 (bottom) shows a comparison between the queue estimates based on vehicle re-identification (offline) and the ground truth queue length. For this approach, the performance seems to be better during on-ramp saturated conditions. However, there is a long period of time, between 16.3 and 16.6 hours, during which the queue length estimates were much smaller than the ground truth values. FIGURE 9 (bottom) shows the instances at which a vehicle was re-identified (matched). From the 543 vehicles that entered the on-ramp, only 27 were matched. The matches occurred only during unsaturated on-ramp conditions, which suggest that saturation

not only affects vehicle counting, but also vehicle re-identification matching rate. The matching rate obtained accounts for around 5% of the vehicles, which is significantly lower than the 70% that has been reported for arterials (*7*).

The vehicle re-id queue estimation method was implemented on the Hegenberger on-ramp as it has been implemented on arterial streets. The assumption that this method would perform similarly on arterial streets and loop on-ramps seems to be incorrect. Loop on-ramps have characteristics that affect the performance of this method, and need to be taken into account. In order to make this method reliable for queue estimation and regulation, it was necessary to work on the vehicle re-identification algorithm and take into account on-ramp specific factors like ramp curvature, slope, vehicle headway, location of sensors, etc. The improved queue estimation method based on vehicle re-identification is presented in section 6.

6. IMPROVED VEHICLE RE-IDENTIFICATION QUEUE ESTIMATION

The queue estimation method based on the vehicle re-identification algorithm used by Sensys Networks performed better than the other three methods, but it significantly under-performed in estimating queue lengths during saturated on-ramp conditions. This was due to the fact that the vehicle re-identification matching rate was very low, especially during congested on-ramp conditions, when vehicles stop and move forward at very low speeds as they go through the sensors. The poor performance of the vehicle matching algorithm in the queue estimation study called for an algorithm revision, improvement and performance analysis, which is presented in this section. The improvements in the vehicle re-identification algorithm help to improve the queue estimation method based on vehicle re-identification.

This part of the report is organized as follows: the vehicle re-identification method is summarized in the first section. The ground truth (GT) and vehicle detection system (VDS) data are explained in the second section. The revision details are presented in the third section. The fourth section discusses the modifications on the algorithm. The fifth section contains the performance analysis of the original and modified vehicle re-identification method. Section six shows a comparison between the original and the modified queue estimates. Finally, the last section shows the modified queue estimation results for the month of March 2011 at the Hegenberger on-ramp. Conclusions are collected in the last section.

Vehicle Re-Identification Method

The vehicle re-identification method summarized in this section is described in (*7*).

Vehicle Magnetic Signature

The magnetic vehicle signature consists of a collection of peak value sequences (local maxima and minima) extracted from the 'raw' magnetic signals measured by an array of sensors. Each sensor has a three-axis magnetometer that measures the x, y and z directions of the earth's magnetic field as a vehicle goes over it. Each extracted peak is paired with a local time stamp, which can be used to determine relative separation among peaks generated from the same sensor. Each sensor generates three peak sequences extracted from the x, y and z component signals,

which constitute a signature slice $X_i = (X_x^i, X_y^i, X_z^i)$. Seven slices constitute the vehicle's signature. FIGURE 12 (a) shows two slices of a vehicle's signature measured at the entrance and at the exit arrays plotted using the local time stamp component and the peak amplitudes before any processing.

The vehicle re-identification algorithm takes two signatures, $X_i = (X_1, \cdots, X_7)$ and $Y_j = (Y_1, \cdots, Y_7)$ where (X_q, Y_r) are slices, and computes a distance (a measure of dissimilarity) between each pair of slices componentwise. The distance $\delta(X_i, Y_j)$ is defined as the minimum of the distances between all pairs of slices (X_q, Y_r). This method uses a dynamic time warping approach to calculate the distance between signatures.

Vehicle Re-Identification Algorithm Summary

The vehicle re-identification is done in two steps:

1) Signal Processing Step: In this step, each pair (X_i, Y_j) of entrance and exit vehicle signatures is compared to produce a distance $d(i, j) = \delta(X_i, Y_j) \geq 0$ between them. The smaller $\delta(X_i, Y_j)$ the more likely it is that X_i, Y_j are signatures of the same vehicle. This step reduces the two signature arrays $X = \{X_i, i = 1, \cdots, N\}$ and $Y = \{Y_j, j = 1, \cdots, M\}$ to the $N \times M$ distance matrix $D = \{d(i, j) \mid 1 \leq i \leq N, 1 \leq j \leq M\}$.

2) Matching Step: In the second step a *matching function* assigns to each distance matrix D a matching $\boldsymbol{\mu}$: $\{1, \ldots, N\} \rightarrow \{1, \ldots, M, \tau\}$, with the following interpretation: $\boldsymbol{\mu}(i) = j$ means that the upstream vehicle i is declared to match (be the same as) downstream vehicle j; $\boldsymbol{\mu}(i) = \tau$ means i is declared not to match any downstream vehicle.

It is assumed that the distance matrix D is characterized by two probability density functions (pdf), f and g: f is the pdf of the distance $\delta(X_v, Y_v)$ between the signatures at the entrance and exit sensor arrays of the same randomly selected vehicle v, and g is the pdf of the distance $\delta(X_v, Y_w)$ between two different randomly selected vehicles $v \neq w$. The f and g pdfs are assumed Gaussian and their statistics, the mean μ and the standard deviation σ, are part of the algorithm parameters that must be determine beforehand.

The Probability of Turn (β) is another matching algorithm parameter. In an arterial implementation, β can be defined as the percentage of vehicles that went through the upstream array but turned before reaching the downstream array, and it can be determined from field observations or experience. But another consideration may govern the choice of β. The larger β is, the more stringent is the requirement of a match, and the lower is the probability of an incorrect match.

Ground Truth and Vehicle Detection System Data

Ground Truth Data

Ground truth data was obtained from the videos recorded on May 11, 2010 from 4:07 pm to 5:35 pm. Three independent cameras were used to obtain the ground truth data (see FIGURE

4) to study the vehicle re-identification algorithm. From the second camera (FIGURE 4 (b)) it was possible to obtain the time s_{GTk} when each entering vehicle k crossed the entrance array, where $s_{GT1} < s_{GT2} < \cdots s_{GTN_{GT}}$. From the first camera (FIGURE 4 (a)) it was possible to get the time t_{GTl} when each exiting vehicle l went through the exit array, where $t_{GT1} < t_{GT2} < \cdots t_{GTM_{GT}}$. The GT data for this analysis consists of two vectors $\{s_{GTk}, k = 1, \cdots, N_{GT} = 543\}$ and $\{t_{GTl}, l = 1, \cdots, M_{GT} = 534\}$. The GT matching of upstream to downstream vehicles $k \rightarrow l$ was done visually and resulted in 534 matches.

Vehicle Detection System Data

Consider the link formed by the entrance and exit arrays shown in FIGURE 2. During the video recording time interval, detection events indexed $i = 1, \cdots, N$ were registered by the entrance array at times $s_1 < s_2 < \cdots s_N$. Detection events indexed $j = 1, \cdots, M$ were registered by the exit array at times $t_1 < t_2 < \cdots t_M$. The upstream sensor measures a signature X_i each time there is a vehicle detection event i and the corresponding time s_i. The downstream sensor measures a signature Y_j each time there is detection event j and the corresponding time t_j . For this study, the vehicle detection system data consists of two arrays $\{(s_i, X_i), i = 1, \cdots, N = 522\}$ and $\{(t_j, Y_j), j = 1, \cdots, M = 527\}$. Note that due to the nature of the vehicle detection system, detection errors cannot be avoided and may create multiple signatures of the same vehicle at one location or may result on missing signatures due to undetected vehicles at the entrance and/or exit array. To be able to determine the number of vehicles that are correctly matched by the algorithm and the percentage of those vehicles that are mismatched, mappings of the form $k \rightarrow i$ and $k \rightarrow l \rightarrow j$ were obtained. With them it is possible to determine if a signature (X_i, Y_j) corresponds to the same vehicle or to different vehicles.

Vehicle Subsets

During the video recording period, the ramp presented two traffic modes: uncongested and congested. For this analysis, uncongested conditions correspond to the time interval from 4:07 pm to 5:07 pm, when the on-ramp queue was below on-ramp capacity and vehicles with index k for $1 \leq k \leq 399$ went through the on-ramp. The uncongested vehicle subset is constituted by these vehicles. The congested conditions time interval occurs from 5:07 pm to 5:35 pm, when the queue length was around or beyond on-ramp capacity, vehicles were stopping or going slowly over the entrance array, and vehicles with entering vehicle index k for $400 \leq k \leq 534$ went through the ramp. The congested vehicle subset is composed of these vehicles. Based on the vehicle detection system data, it is possible to achieve a maximum correct matching of 477 vehicles assuming a perfect matching algorithm, where 362 vehicles correspond to the uncongested vehicle subset and 115 to the congested one.

23 Chosen Vehicles

In order to be able to analyze the vehicle re-identification algorithm in detail, 23 vehicles were chosen from the 543 that entered the ramp, as shown in Table 1. The selection criteria were that their entrance and exiting signatures were available and unique and that they were traveling through the middle of the lane when going over the arrays.

Table 1 23 Chosen Vehicles

Uncongested Conditions				Congested Conditions			
Index k	Veh. Type	Org. Dist.	Mod. Dist.	Index k	Veh Type	Org. Dist.	Mod. Dist.
76	minivan	.25	.23	449	car	.50	.41
135	truck	.28	.23	459	SUV	.30	.35
140	truck	.20	.18	471	pickup	.42	.24
187	pickup	.29	.24	472	car	.28	.16
194	truck	.71	.70	490	SUV	.40	.38
222	car	.15	.18	492	car	.33	.29
236	minivan	.24	.14	516	car	.34	.35
259	SUV	.30	.30	517	car	.36	.25
282	pickup	.45	.32	519	car	.35	.18
305	car	.46	.28	527	car	.53	.15
380	car	.29	.25				
393	SUV	.48	.48				
419	van	.38	.29				

Vehicle Re-Identification Method Revision

Signal Processing Step Revision

The plots on FIGURE 10 are gray scale coding of the distance matrix of the 23 chosen vehicles signatures (left) and the distance matrix of the complete vehicle signatures data set (right) calculated using the original signal processing algorithm. Each square in the plots represents a distance between a X_i and Y_j vehicle signature combination; a darker color indicates shorter distance and a greater chance that X_i and Y_j come from the same vehicle. The gray scale used in FIGURE 10 and FIGURE 11 goes from $min(D)$ to $.75median(D)$, where D represents the distance matrix being plotted.

FIGURE 10 (left) shows multiple dark squares along the diagonal, where diagonal entries correspond to distances between signatures of the same vehicle. The distance values of the diagonal entries are listed in Table 1. Note that the diagonal entries corresponding to the congested vehicle subset are not as dark (i.e. small) as the uncongested ones. It seems that phenomena occurring specifically during congestion affect the f and g probability density functions (pdf). This is further corroborated by the plot of the distance matrix for the complete data set, shown on FIGURE 10 (right). For the uncongested vehicle subset, a diagonal line is present and is consistently darker and visually distinguishable from the off-diagonal D entries. However, dark diagonal entries vanish at the lower right portion of the plot, where the distances between vehicles signatures from the congested vehicle subset are plotted.

The original signal processing algorithm produces different f and g pdfs for different on-ramp modes. The f and g pdfs become similar during on-ramp congested conditions, which reduces the effectiveness of the matching algorithm. The signal processing method should be able to maintain the f and g pdfs invariant to traffic conditions.

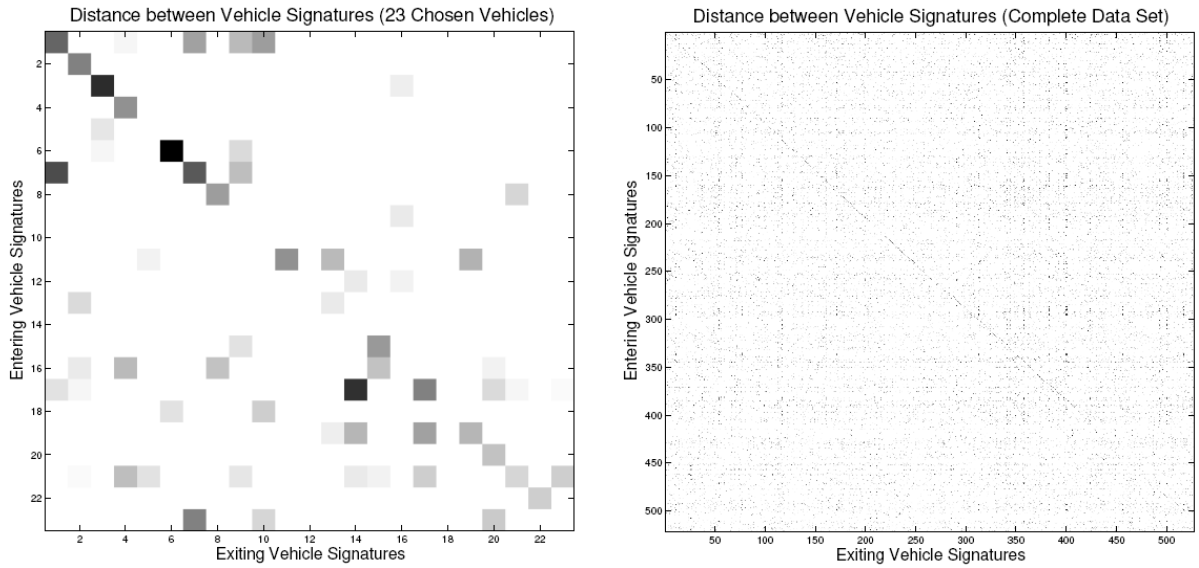


FIGURE 10 Original distance matrix from entrance to exit array for the (left) 23 chosen vehicles (right) complete vehicle data set

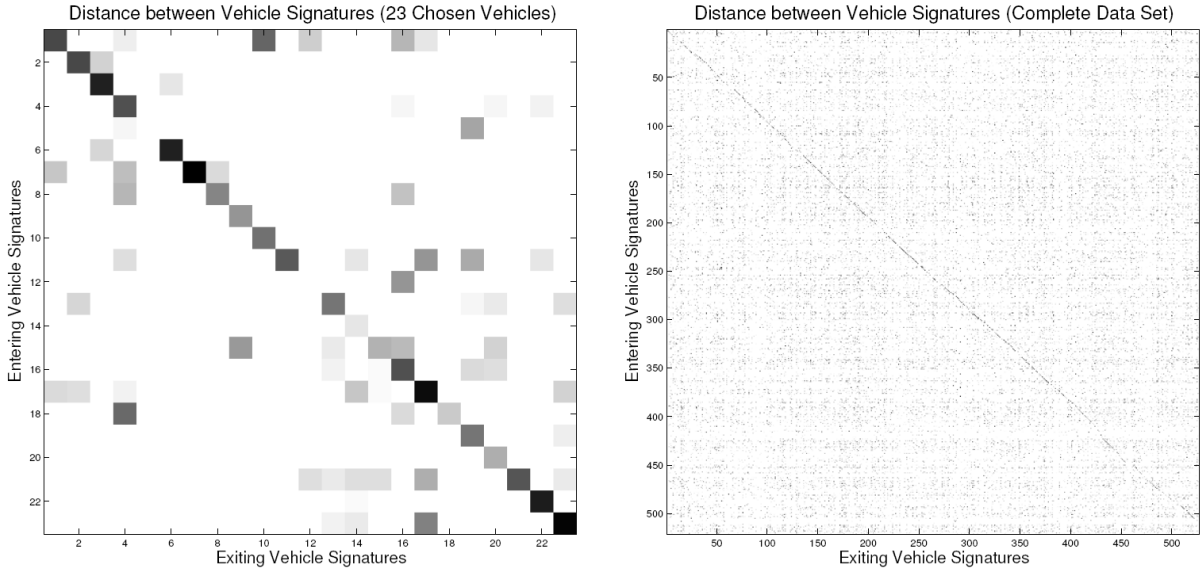


FIGURE 11 Modified distance matrix from entrance to exit array for the (left) 23 chosen vehicles (right) complete vehicle data set

Vehicle Signatures Revision

Entrance and exit signatures of the 23 chosen vehicles were studied in order to explain the different *f* and *g* pdfs observed for the uncongested and congested vehicle subsets.

FIGURE 12 (b) shows two slices of the entrance and exit signatures of vehicle $k = 527$. Note that the upstream slices show more peaks than the downstream ones. Vehicle 527 went slowly over the entrance array and at free flow speed over the exit array. The extra peak

phenomenon was observed in the signatures of vehicles going slowly or stopping over the entrance arrays. During congestion, vehicle signatures tend to have more noisy peaks, most of which are small compared to the dominant peaks of the signatures. A vehicle signature with noisy peaks leads to a larger μ_f and σ_f, since the difference on the number of peaks is penalized in the distance calculation independently of their magnitude.

Another finding was related to the importance of the *x*, *y* and *z* component of a vehicle signature slice when calculating distances. The original distance method assigned the same weight to the distance obtained between the *x*, *y*, and *z* components when calculating distances between signature slices. By changing these weights, it was observed that it is possible to increase the dissimilarity between the *f* and *g* pdfs during congestion.

Sometimes signature slice components (*x*, *y* and *z*) of different vehicles look very similar after the peak sequences have been normalized by the maximum absolute value of their elements as a preprocessing step before the distance calculation is performed. The distance between such slices is small and may leads to significant errors in the matching step. It was observed that when the raw amplitude of the peaks is considered in the distance calculation between such signatures, the distance increases for signatures from different vehicles and remains unchanged for signatures from the same vehicle.

Matching Step Revision

The matching algorithm was studied using the 23 chosen vehicles. It was concluded that the match rate and accuracy are directly related to μ_f , σ_f , μ_g, σ_g and β . Furthermore, it was observed that one of the reasons that explain the low matching rate during the queue estimation study, even during uncongested on-ramp conditions, was directly related to the *f* and *g* parameters used at the Hegenberger on-ramp. The *f* and *g* statistics used at the ramp where assumed to be equal to the ones satisfactorily used at many arterial installation sites (see Table 2, second column). This assumption was incorrect, since *f* and *g* statistics based on the 23 chosen vehicles distance matrix (see Table 2), show that the default *f* and *g* statistics do not model accurately the signature distances at the on-ramp.

Table 2 Different *f* and *g* statistics for the original and modified signal processing method

	Default	23 Veh.		Iterative	
	Both	*Org*	*Mod*	*Org*	*Mod*
μ_f	.125	.36	.27	.30	.28
σ_f	.058	.12	.09	.09	.08
μ_g	.67	.54	.60	.56	.62
σ_g	.23	.11	.17	.13	.16

It is important to calculate *f* and *g* parameters for each test site, since they are site dependent and influence the matching rate. These parameters can be obtained using an iterative method as the one suggested in (*7*), which does not require GT. Note that the iteratively obtained *f* and *g* parameter extracted from the complete vehicle data set are very similar to the ones extracted from the 23 chosen vehicles for both the modified and the original signal processing method, and very different to the default values.

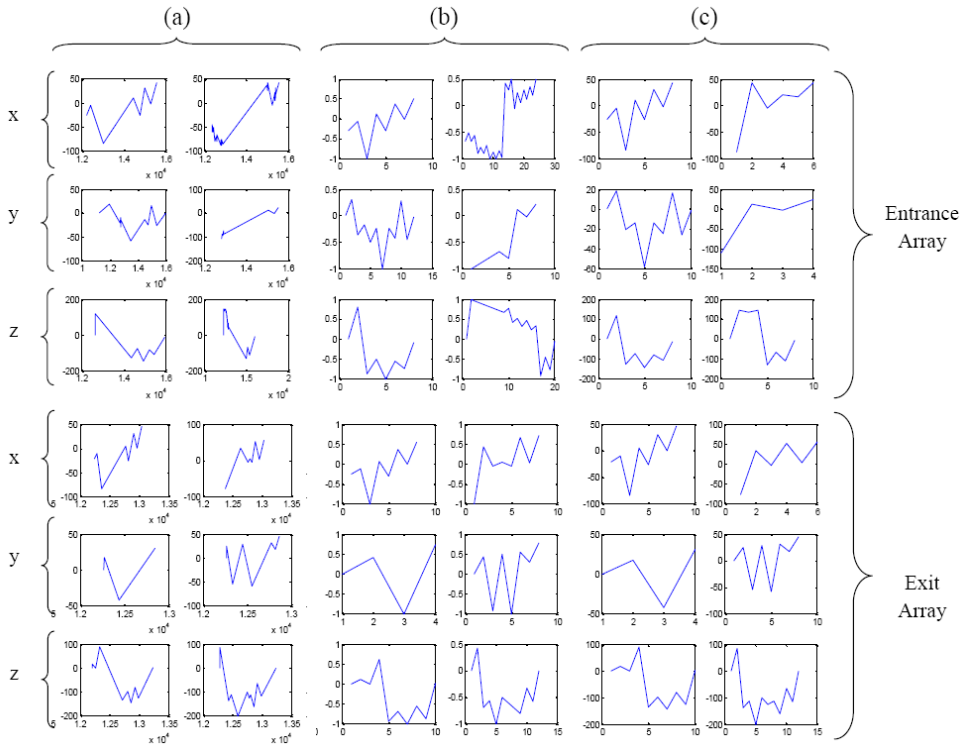


FIGURE 12 Vehicle k=527 two sensor signature slice at the entrance and exit arrays (a) Raw peak values (b) Processed for distance calculation (original method) (c) Processed for distance calculation (modified method)

Vehicle Re-Identification Method Modifications

Signal Processing Algorithm Modification

The following modifications were implemented in the original signal processing algorithm:

i) There were adjustments in the way vectors of different sizes are compared using dynamic time warping. This included modifying the way in which extra peaks are penalized when vectors being compared are of different sizes.

ii) A peak processing step was implemented before the distance calculation in order to remove noisy peaks resulting from vehicles traveling slowly or stopping on top of the arrays. This step uses the local time stamp component available for each signature peak described before. See FIGURE 12 for a comparison between the preprocessed signature slices for the original and the modified method.

iii) Different weights were assigned to the *x*, *y* and *z* components of the distance between two vehicle slices. The *x* component was assigned the larger weight and the *y* component the smaller one.

iv) The distance calculation is performed without normalizing the peak sequences. Once a distance is obtained between the components of two signature slices, a normalization step is performed.

Signal Processing Improvements

The plot on FIGURE 11 (left) is a gray scale coding of the distance matrix of the 23 chosen vehicles signatures using the modified signal processing algorithm. Note that in contrast to FIGURE 10 (left), dark squares are present along the entire diagonal, thus suggesting that f and g statistics are somehow invariant to on-ramp traffic conditions. In FIGURE 11 (right), where the distance matrix of the complete data set calculated using the modified distance method is plotted, it is possible to see a darker diagonal band at the bottom right corner of the plot. This diagonal band corresponds to distances generated from the same vehicle belonging to the congested vehicle subset. This distinction between diagonal and off-diagonal entries was not present in the original distance matrix shown in FIGURE 10 (right), which helps explain why the matching rate was particularly low during congestion.

From Table 1 it is possible to see that the modified distance among signatures from the same vehicle, $\delta(X_k,Y_k)$, are similar for the uncongested and congested vehicle subsets. FIGURE 13 shows for each of the 23 chosen vehicles, k, $D_k = \{d(k, l) = \delta(X_k,Y_l) \mid 1 \leq l \leq 23\}$. For each D_k, the data points given by $\delta(X_k,Y_l)$ for $l \neq k$ are plotted as black dots while the $\delta(X_k,Y_k)$ distance is plotted as a blue circle. The top plot contains the distances obtained using the original method. This plot shows that for vehicles from the congested vehicle set, $\delta(X_k,Y_k)$ is generally not the smallest entry of D_k, which is undesired and affects the performance of the matching algorithm. With the modified signal processing method, as displayed in FIGURE 13 (bottom), the $\delta(X_k,Y_k)$ data point, represented by a blue circle, is generally the lowest element of D_k for vehicles belonging to both uncongested and congested vehicle subsets. This improves the matching algorithm accuracy.

Finally, after comparing the f and g pdfs for the original and modified method, listed in Table 2, it is clear that the modified distance method f and g statistics will benefit the matching algorithm performance due to the larger difference between μ_f and μ_g and the smaller values of σ_f and σ_g.

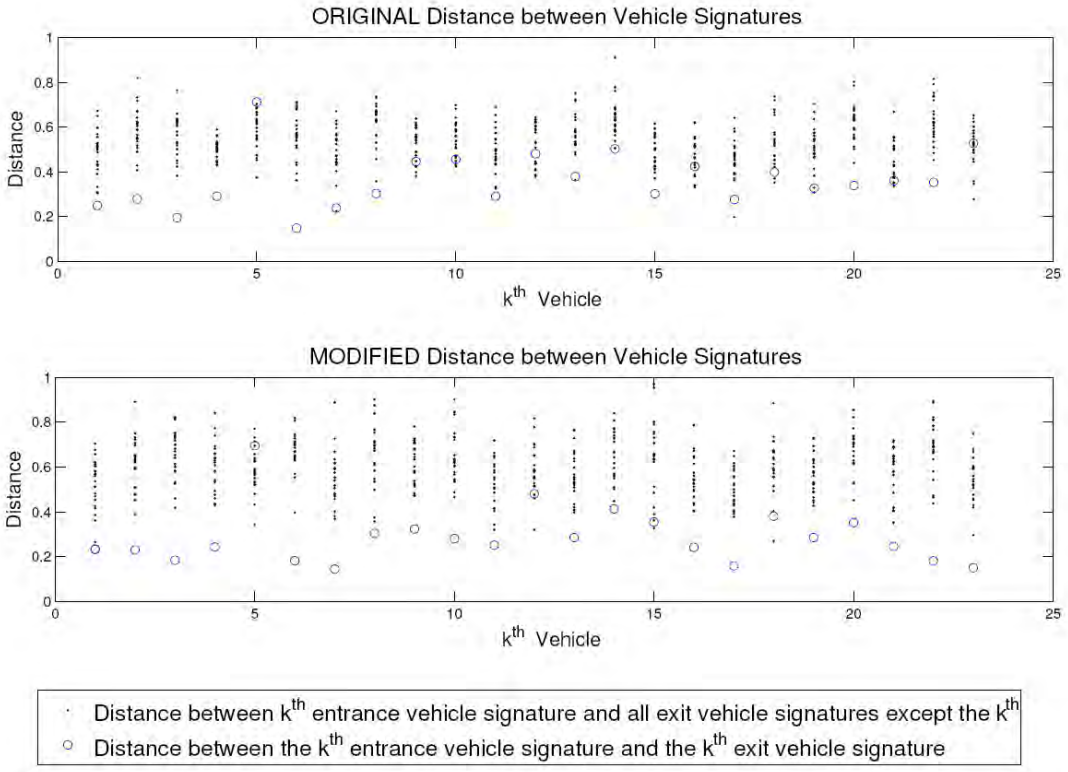


FIGURE 13 Comparison between the original distance method and the modified distance method

Matching Algorithm Modification

The matching algorithm was not modified. However, it was observed that the *f* and *g* statistics as well as β play an important role in the number of matched vehicles and the percentage of mismatched vehicles produced by the algorithm.

Vehicle Re-identification Results

Default vs Iterative μ_f, σ_f, μ_g, σ_g Results

In this section, two different sets of μ_f, σ_f, μ_g, σ_g were used to study the effect of *f* and *g* pdfs variations on the matching algorithm performance for the original and the modified signal processing methods. The default and the iteratively obtained set of parameters, listed in Table 2, were used.

Table 3 shows the results obtained with the matching algorithm for the complete vehicle data set, the uncongested vehicle subset and the congested vehicle subset. The modified method has a higher number of re-identified vehicles in comparison to the original method for both set of parameters. The number of matches over the original method results increased by 56% for the default case and by 13% for the iterative one. The increase in vehicle matches did not result in an increase of the percentage of incorrectly matched vehicles. For the original method 23% of

vehicles were incorrectly matched when using the default parameters and 16% when using the iterative parameters. For the modified method, from 166 matched vehicles obtained using the default parameters, 7% were incorrectly matched, while 7% out of 368 vehicles were misidentified using the iterative values. The modified distance method seems to increase the matching rate in comparison with the original distance method while keeping the percentage of incorrect matches low and constant for significant μ_f, σ_f, μ_g, σ_g variations. Note that the use of f and g statistics that properly model the distance between signatures (e.g. iteratively obtained parameters), makes a significant difference in the number of matched vehicle but does not impact the accuracy of the matches.

Table 3 Matching results using the default and iterative f and g statistics using the modified and original method

	Default		Iterative	
	Org	*Mod*	*Org*	*Mod*
Matched(total)	106	166	325	368
Incorrectly M.	24	11	53	24
Matched(uncong)	90	139	257	279
Incorrectly M.	14	8	32	14
Matched(cong)	16	27	68	89
Incorrectly M.	10	3	21	10

When the results of the matching algorithm are analyzed by vehicle subset, further differences were encountered. First, the percentage of mismatched vehicles is larger for the congested vehicle subset in comparison to the uncongested one, for the original and the modified signal processing methods. The percentage of incorrectly matched vehicles changes from 15% to 62.5% when using default parameter values and the original method. This value changes from 12% mismatched percentage to 31% when the iterative parameters are used instead. The modified distance method also shows differences in the accuracy obtained for the uncongested and congested vehicle subset results. The percentage of incorrect matches changes from 6% to 11% when using the default parameters. This number changes to 5% incorrect matches for the uncongested vehicles subset and 11 % for the congested subset when using the iterative parameters. The decrease in performance of the matching algorithm during congestion is observed for both the original and the modified signal processing method. Note that the original method accuracy is highly dependent on the values of μ_f, σ_f, μ_g, σ_g while the accuracy of the modified method seems to remain unchanged.

Iterative μ_f, σ_f, μ_g, σ_g with varying β Results

The matching algorithm results presented in this section were obtained using the iteratively obtained μ_f, σ_f, μ_g, σ_g parameters and using both the original and the modified vehicle re-identification method as β was varied.

From FIGURE 14 (left) it is observed that the modified method has higher matching rate for all β values considered for this analysis for the complete vehicle data set, the uncongested vehicle subset and the congested vehicle subset. FIGURE 14 (middle) shows the percentage of

incorrectly matched vehicles for the uncongested vehicle subset as function of β. Note that both methods have a mismatch percentage that remains somehow constant as β is varied. However, an advantage of the modified distance method is that the percentage of incorrect matches is around or below 8%, in comparison to the 15% obtained with the original method. FIGURE 14 (right) shows the percentage of incorrectly matched vehicles for the congested vehicle subset. The percentages of incorrect matches are larger for all β for the original and modified methods in comparison to the uncongested results. Observe that while the original method mismatch percentage increases from 20% at low β values to 60% for large ones, the modified distance method percentage of incorrectly matched vehicles remains around or below 14% for all β values.

The original algorithm matching rate and accuracy is affected by β variations. The modified distance method matching rate is affected by changes in β, but the accuracy of the matches seems to remain constant.

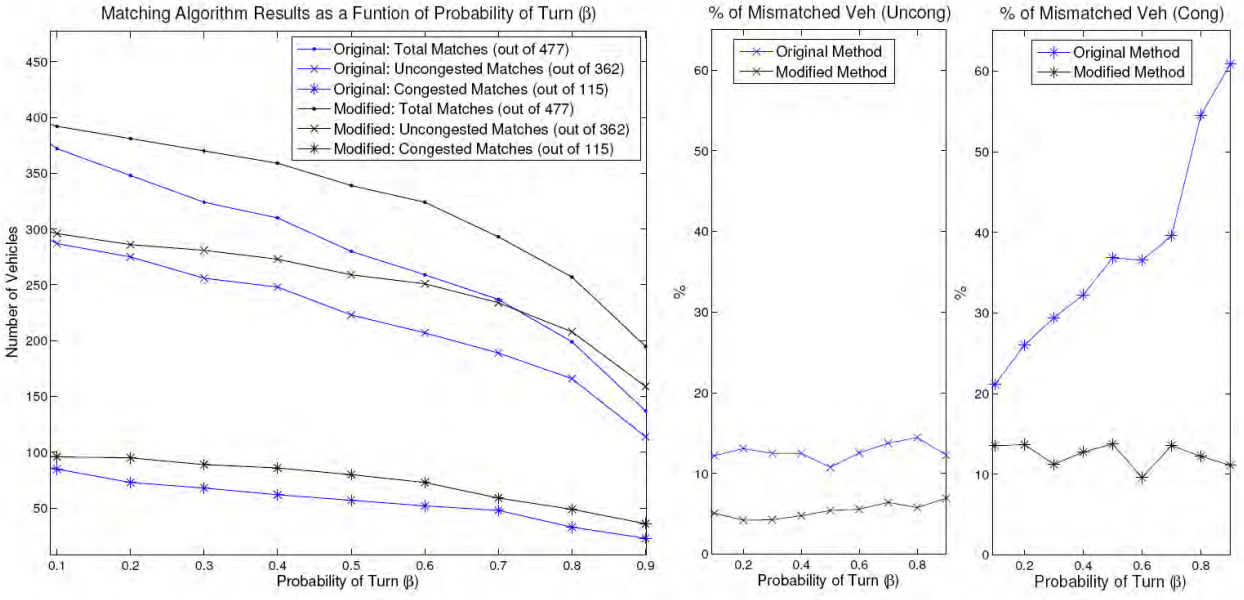


FIGURE 14 (left) Matched vehicles as a function of β (middle) % of uncongested mismatched vehicles as a function of β (right) % of congested mismatched vehicles as a function of β

Improved Queue Estimation based on Modified Vehicle Re-Identification Algorithm

With the modified vehicle re-identification algorithm, the vehicle matching rate during saturated ramp conditions was significantly improved, resulting in a more accurate and reliable queue estimate. A higher matching rate allows for a better and rapid correction of errors in the queue estimate. Mismatched vehicles cannot be completely avoided, but since they are around 6% of the total number of matched vehicles, they do not introduce significant errors in the queue estimation. During congestion, the queue estimation error increases in comparison to the uncongested case, but with the modified vehicle re-identification algorithm, the error remains bounded and the queue estimate is able to track the ground truth queue length with sufficient accuracy. A typical result is shown in FIGURE 15.

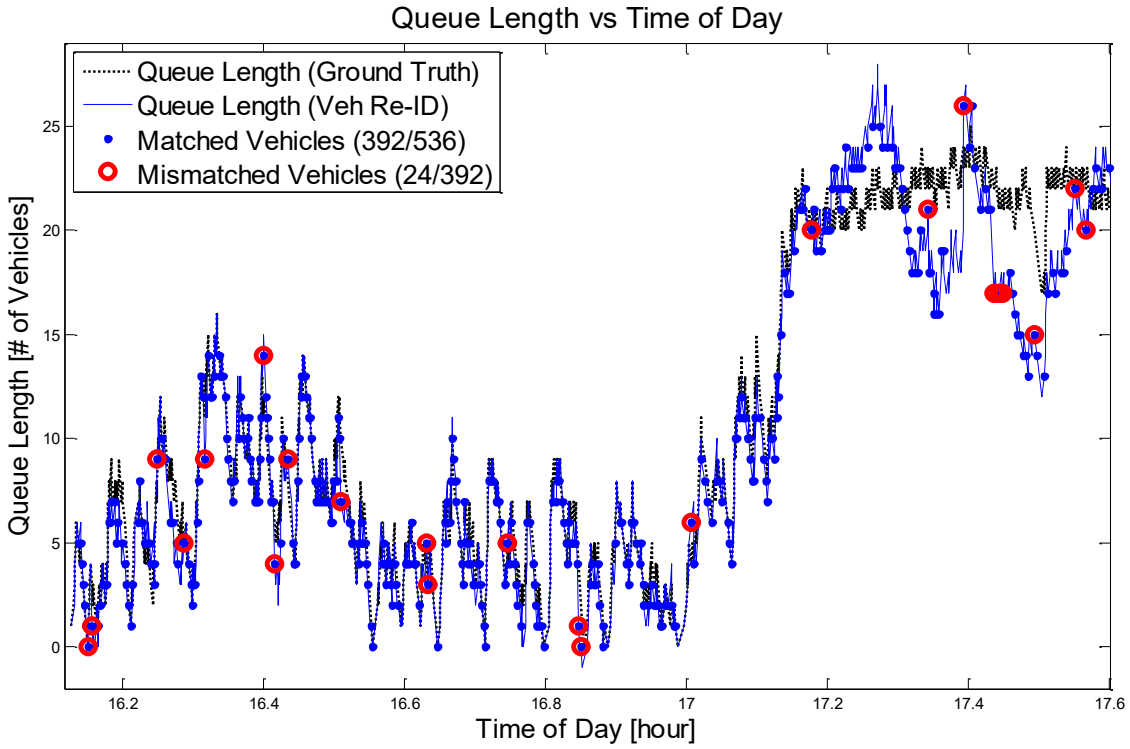


FIGURE 15 Results using the Improved Vehicle Re-Identification Queue Estimation Method

For this particular analysis, we used the Sensys vehicle detection system and re-identification algorithm to estimate the queue length at the on-ramp. However, a queue estimation method based on vehicle re-identification is not exclusive to Sensys. The vehicle queues at on-ramps can be accurately estimated using vehicle counts and a correction mechanism based on re-identifying vehicles using other detection systems. Queue estimation based on counting and vehicle re-identification could be implemented with the use of video cameras, loop detectors, Bluetooth technology, among others, provided that these systems are capable of accurate vehicle counting and reliable vehicle re-identification, and they are able to perform the re-identification in real time. The required matching rate for accurate queue estimation of a vehicle detection system would be a function of the accuracy of the system to measure vehicle flow.

Queue and Travel Time Estimation based on the Modified Vehicle Re-Identification Method for March 2011

The vehicle re-identification system installed at the Hegenberger on-ramp can be used to calculate queue lengths as well as travel times. With this system it is possible to calculate the number of vehicles in between the entrance and exit arrays as well as the travel time between the arrays for every vehicle that is re-identified. This information is useful, since queue length information may not be sufficient to evaluate traffic conditions at the ramp.

In this section we present the queue length and travel time estimates calculated using the modified vehicle re-identification system during the month of March 2011. We show how queue lengths and travel times vary as a function of the day of the week (i.e. weekday vs. weekend) and as a result of events taking place around the Hegenberger on-ramp (e.g. games at the Oakland Coliseum). Furthermore, we compare queue length and travel time estimates when ramp metering is active and when it is inactive. Finally, we show that by knowing travel time and queue lengths it is possible to get a better assessment of the driving conditions on the on-ramp.

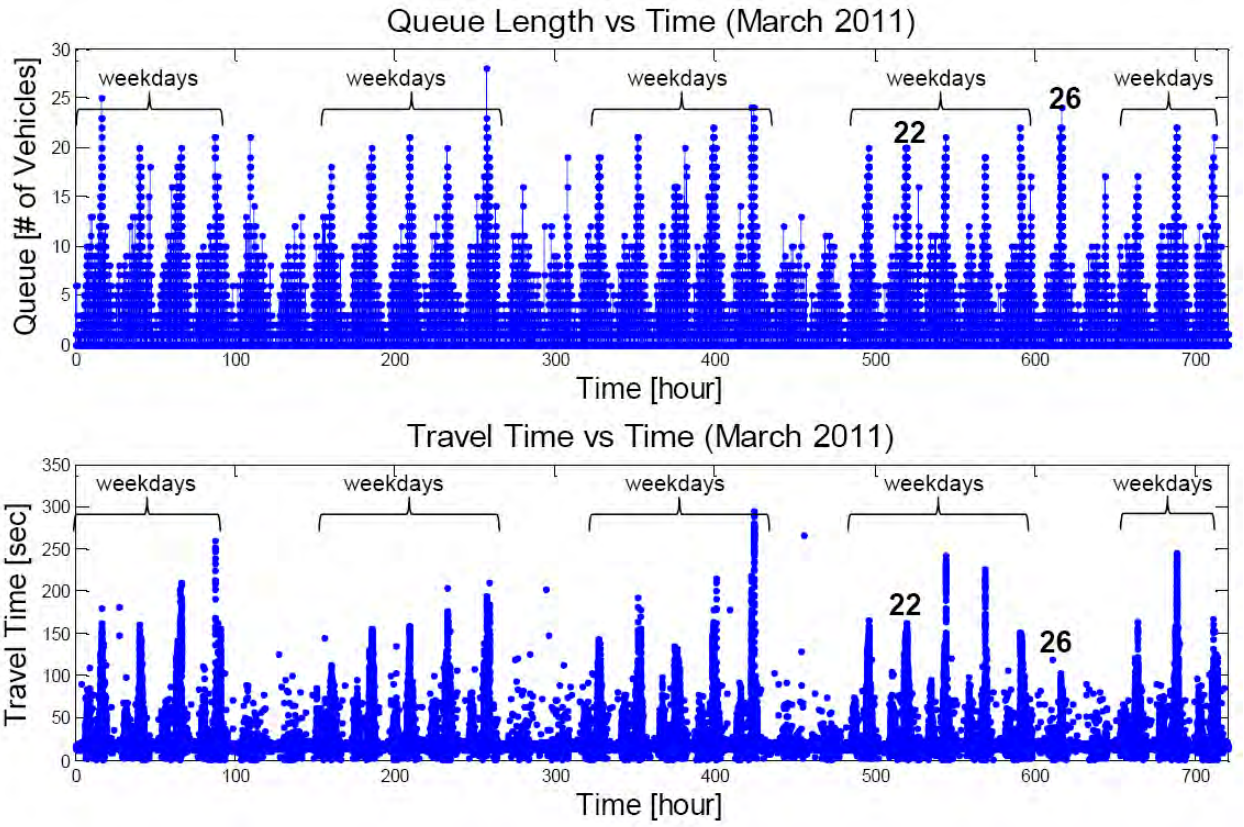


FIGURE 16 March 2011 Hegenberger on-ramp (top) queue estimates (bottom) travel time estimates

FIGURE 16 shows the queue length and travel time estimates for the month of March 2011. In the figure it is possible to distinguish estimates generated during weekdays from those obtained based on weekend data. For standard weekdays, the queue length reaches values of around 22 vehicles, and travel times of up to 300 seconds. During standard weekend days, a different pattern is observed. Queue lengths tend to be smaller, reaching up to around 10 vehicles, and travel times generally do not exceed 100 seconds. For the month of March 2011 it is also possible to observe non-standard days, in which queue lengths and travel times do not resemble the patterns for standard weekdays or weekend days. Non-standard days usually occur due to events happening at the Oakland Coliseum or because of incidents occurring around the Hegenberger ramp. In order to illustrate the difference between standard and non-standard days, as well as the effect of active ramp metering on the queue and travel time estimates, two days of the month of March were chosen: March 22, 2011, which corresponds to a standard weekday,

and March 26 2011, which corresponds to a non-standard weekend day. The data for these days are marked in FIGURE 16.

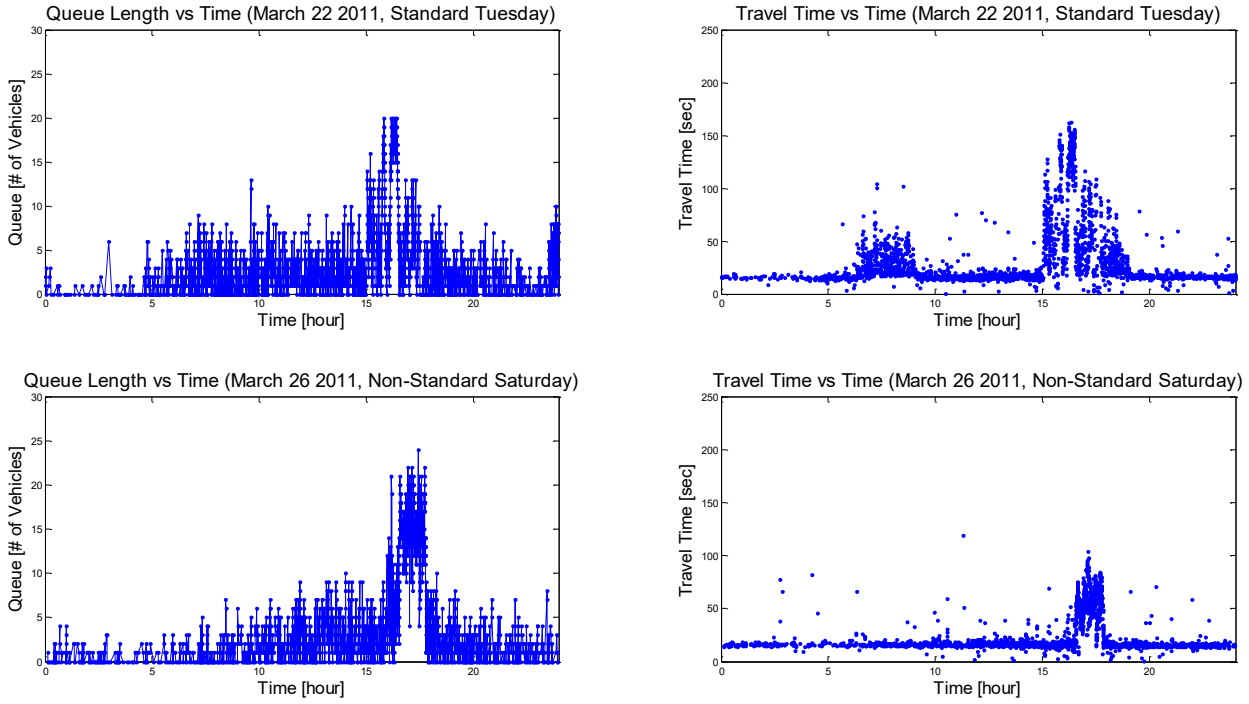


FIGURE 17 (top) Queue length and travel time estimates for March 22 2011 (bottom) Queue length and travel time estimates for March 26 2011

FIGURE 17 (top) shows the 24 hour queue length and travel time estimates for the standard weekday corresponding to March 22 2011. In the queue estimate plot (top, left), it is possible to observe a time interval for which the queue length is higher than during the rest of the day. This time period corresponds to the afternoon peak hour, when ramp metering is activated, demand is higher than during the rest of the day, and queues are frequently formed. During weekdays, ramp metering is also active during the morning peak hour. However, the queue length plot in FIGURE 17 (top, left) does not suggest that ramp metering is active and that long queues are forming during the morning. This is expected because at the Hegenberger on-ramp, demand is not high during the morning peak hour in comparison to the afternoon rush hour. Nevertheless, even when the vehicle queues do not reach values larger than 15 vehicles, it is possible to observe that during a standard weekday morning with short queues, drivers going through the on-ramp experience a significant delay as a result of active ramp metering, as observed in FIGURE 17 (top, right) between the 6 and 9 hours. This is a specific example where knowing travel time on top of queue length estimates at an on-ramp may be helpful to have a better assessment of the traffic conditions.

From FIGURE 17 (top) it is evident that during the afternoon peak traffic time, queue lengths as well as travel times significantly increase in comparison to the rest of the day. FIGURE 18 (top) shows the queue and travel time estimates for March 22 2011 during the afternoon peak hour. In these plots it is possible to see that queue lengths and travel times are higher than during the rest of the day, but the estimates oscillate even when ramp metering is

active and the mainline is congested. The presence of ramp metering induces the formation of vehicle queues at the ramp, but sometimes there are periods when demand drops allowing the vehicle queue to completely dissipate before forming again.

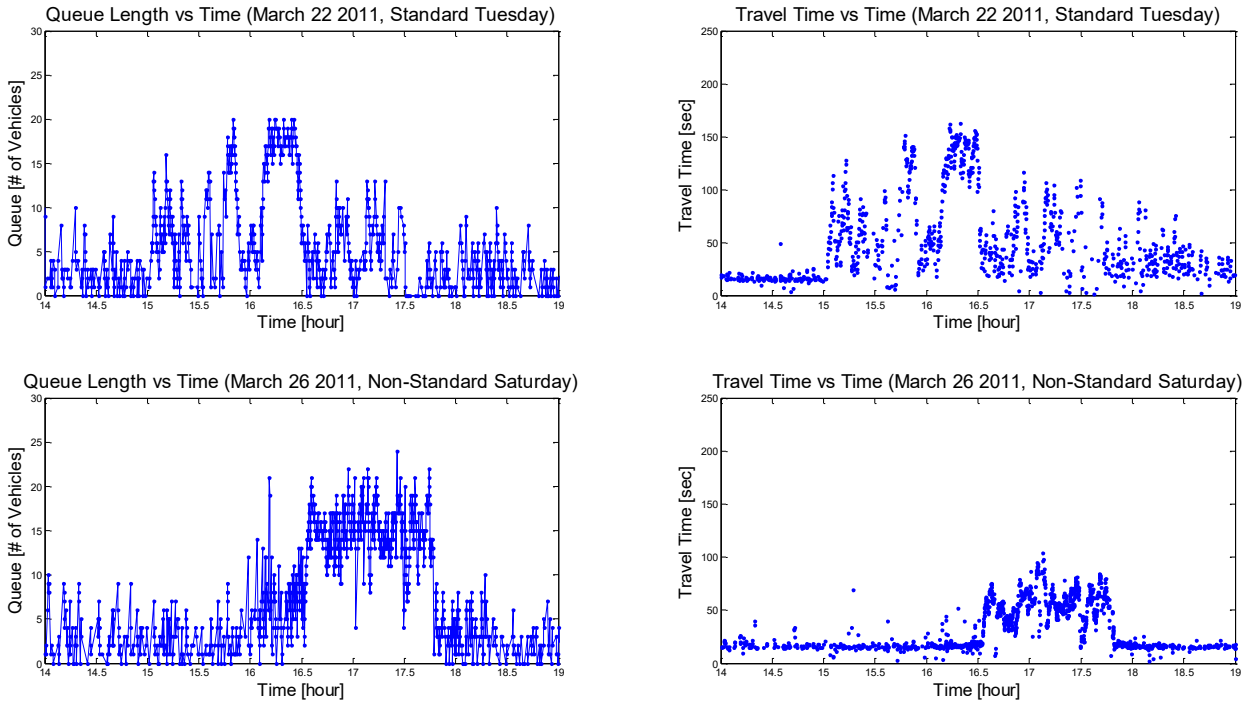


FIGURE 18 Queue length and travel time estimates for March 22 2011 between 14 and 19 hours (bottom) Queue length and travel time estimates for March 26 2011 between 14 and 19 hours

FIGURE 17 (bottom) shows the 24 hour queue length and travel time estimates for March 26 2011, a non-standard weekend day. For these plots it is possible to observe a time interval in the afternoon for which the estimates are significantly higher than during the rest of the day. In this particular day, there was a soccer game at the Oakland Coliseum that ended at around 16 hours, and that induced a large demand at the on-ramp. If this had been a standard weekend day, the queue lengths would have been shorter than 10 vehicles and the travel times would have been around 25 seconds. However, the event at the Oakland Coliseum significantly increased the Hegenberger on-ramp vehicle demand right after the soccer game as a result of many people exiting the Coliseum and trying to get into the freeway. Since the demand was high and constant, the queue length was not oscillating as observed for the afternoon peak hour on March 22 2011. Even when the number of vehicles at the on-ramp was around capacity from around 16.5 hours to 18 hours, as observed in figure FIGURE 18 (bottom), the travel time experience by the drivers during this time period was around 75 seconds, significantly less than the travel time observed during the weekday rush hour when vehicle queues had similar lengths and ramp metering was active. This is another scenario that illustrates that knowing travel time information besides queue length allows for a better understanding of driving conditions at the on-ramp in comparison to just knowing queue lengths. Note that for this particular day, the freeway segment adjacent to the Hegenberger on-ramp was heavily congested after the soccer

game, a condition that could have been avoided or at least improved by using proper ramp metering.

7. SOURCES OF ERROR

Even though the number of motorcycles using the on-ramp is very low in comparison to the total number of vehicles, they can introduce important queue estimation errors over time. Motorcycles are more likely to be undetected due to their smaller size and because they tend to miss the vehicle detection station at the exit of the ramp as they have a tendency to bypass the queue and exit the on-ramp off-centered. FIGURE 19 (a) shows a motorcycle entering the on-ramp, bypassing the queue, and finally missing the exit sensor array. For this particular motorcycle, the entrance and exit sensor array did not register its presence.

There have been multiple observations of cars bypassing the queue. This affects the queue estimation methods because vehicles bypassing the queue tend to miss the detection sensors at the exit of the on-ramp. They usually travel off-centered of the lane, as shown in FIGURE 19 (b).

Trucks are consistently using the on-ramp, even though they account for a small percentage of the vehicles going through it. When they appear, they can affect the queue dynamics considerably since they are longer, slower and have to maneuver to make the turn properly. FIGURE 19(c) shows an example of a large truck maneuvering at the entrance of the on-ramp. It was observed with the raw sensor array data and speed trap sensor raw data that trucks like this one tend to introduce counting error of one or two vehicles.

Loop on-ramps have wide lanes that allow drivers to maneuver as they go through. The extra lateral space results on vehicles traveling highly off-centered. FIGURE 19 (d) shows multiple instances when vehicles were at different positions with respect to the center of the lane. Sometimes vehicles are off-centered and they end up completely missing the vehicle detection station. This phenomenon is an important source of error on the queue estimation methods that rely on counting vehicles.

As it was shown in section 5, queue estimation methods have their worst performance during saturated on-ramp conditions. One of the reasons for this phenomenon is that two adjacent cars close to each other are likely to stop on top or very close to the detection zone of the sensors. This creates undercounting problems, since two cars close to each other may be registered as one. FIGURE 19(e) shows examples of adjacent cars resting on top of the sensor arrangements at the entrance of the on-ramp. This results on errors for queue estimation methods relying on vehicle counting. It may also occur that vehicles are completely stopped at the on-ramp and none of them are within the sensor detection zone. When waiting time is large in comparison to the occupancy calculation interval, occupancy measurements may not reflect on-ramp queue conditions (ZSZO-phenomenon). This situation was observed in FIGURE 5 (top) around 17.4 hours.

Note that the sources of error mentioned here also apply to vehicle detection systems based on inductive loops and some of them have been addressed on (*6*).

FIGURE 19 (a) Motorcycle bypassing queue. (b) Vehicle bypassing queue. (c) Large truck entering the on-ramp. (d) Off-centered vehicles with respect to the lane. (e) Adjacent vehicles simultaneously stopped on top of the leading and trailing speed trap sensors.

8. TASK 5

As mentioned previously, the queue control algorithm was not fully incorporated into the URMS. In order to complete this task it was necessary that Caltrans D4 approve the URMS for ramp metering testing and provide us with the specific version of URMS that it had approved. Unfortunately, to the best of our knowledge, Caltrans D4 approval of URMS for ramp metering testing has not yet taken place.

In addition, it was necessary that Sensys Networks develop a new communication protocol between the Sensys Access Point (AP) and the URMS to transmit queue lengths. Although Sensys Networks agreed to help us develop such a protocol under the condition that the field test would actually take place, the protocol was not fully developed because it became clear that Caltrans D4 was not going to approve the URMS for ramp-metering testing in time for this project to be completed.

At the early stages of the project the URMS source code and data structure were analyzed in order to determine what modifications would be required in order to incorporate the queue control algorithm into the URMS code. It was determined that the incorporation of the queue control code into the URMS would have been a relatively straight forward programming exercise, since the queue control algorithm involves a relatively simple proportional plus integral (PI) controller and additional logic, as explained further in Section 8 and 9, and its placement within the ramp metering code was determined, as explained further in this section. However, the most challenging task of the queue control implementation would have been to establish the communication protocol between the Sensys AP and the URMS. Unfortunately, Sensys Networks was unwilling to devote precious engineering resources to complete this task once it became apparent that Caltrans D4 was not going to approve the URMS for ramp metering testing in time for the completion of this project. It should be emphasized that Sensys Networks remains committed to help us establish such a communication protocol when and if it is determined that the queue control field test implementation can be conducted.

In the remaining part of this section we describe the setup that was installed on the Hegenberger on-ramp controller cabinet to allow the Sensys AP to transmit the queue length estimates to the URMS. We provide a summary of the communication protocols that were considered for this purpose. Finally, we also describe the revisions that should have been done to the URMS code, in order to incorporate the queue control algorithm. Unfortunately, because of restrictions imposed on us by the NDA, we cannot show actual URMS code in this report nor provide detailed descriptions of its structure.

Controller Cabinet Set Up

The Sensys AP needed to be connected to the traffic controller cabinet for the following reasons:

1) To draw up to 6 Watts of power from the 12V or 24V lines of the traffic controller backplane to power itself. This connection was done using the Sensys master (CC) contact closure card and the Sensys access box, as shown in FIGURE 20.

2) To interface Sensys sensors signals to the 2070 traffic controller through the contact closure cards. Note that these cards have the capability to transmit detection events in the same way that regular loop detector cards transmit data to the URMS.

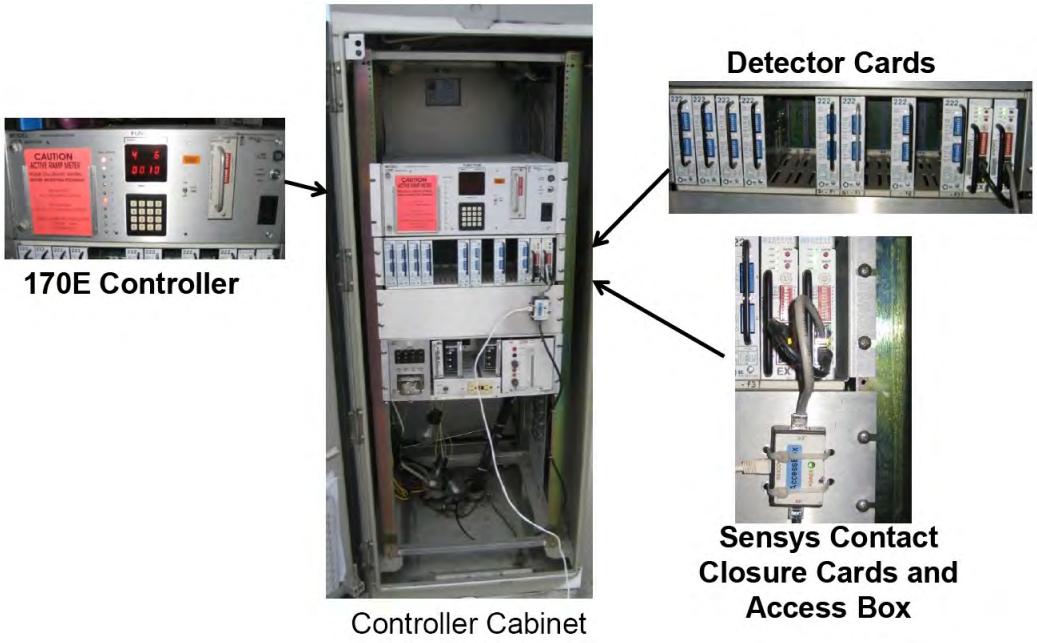


FIGURE 20 Hegenberger on-ramp controller cabinet

FIGURE 20 shows the Hegenberger on-ramp controller cabinet after the Sensys vehicle detection system installation was completed at the Hegenberger on-ramp. There was the need to install two Sensys contact closure cards and a Sensys access box. The controller cabinet was ready so that the 170 traffic controller could be replaced by a 2070 traffic controller running the modified URMS.

Communication Protocol between Sensys Access Point and URMS

The Sensys AP can communicate volume, occupancy, speed (using a speed trap), and queue length (as an occupancy value) through the contact closure cards to the URMS. When the Sensys vehicle detection system installation was completed at the Hegenberger on-ramp, the controller cabinet was modified so that task 6 and task 8 could be completed with any of the queue estimation methods studied. If queue estimates based on speeds would have been acceptable for stage 2 and stage 3 of the project, then the speed trap detection events would have been transmitted through channel 1 and channel 2 of the EX contact closure card (see FIGURE 21). If counting vehicles without any error correction mechanism would have been acceptable to calculate queue lengths, then channel 1 and channel 2 of the CC contact closure card could have been used to transmit volume measured with the Sensys arrays at the entrance and at the exit of the on-ramp (see FIGURE 21). It turned out that the only reliable method to estimate the queue, as discussed earlier, was the one based on vehicle re-identification. This method requires the transmission of a number, which usually goes from 0 to 25, to indicate the queue length estimate. The original approach involved communicating this number as an occupancy value, as explained

in FIGURE 22, but later on it was decided that a cleaner way to communicate this value would be trough the 2070 controller Ethernet port, as shown in FIGURE 23. This new approach was considered on April 7 2011, when Sensys Networks, Caltrans D3 People and David J. Wells met to discuss a direct interface between AP and the URMS that would eliminate the need to have contact closure cards. The idea was to include sensor detection events or queue lengths in a communication message sent directly from the AP to the URMS through the 2070 controller Ethernet port. The communication protocol to transmit queue length was never completed for the reasons stated earlier.

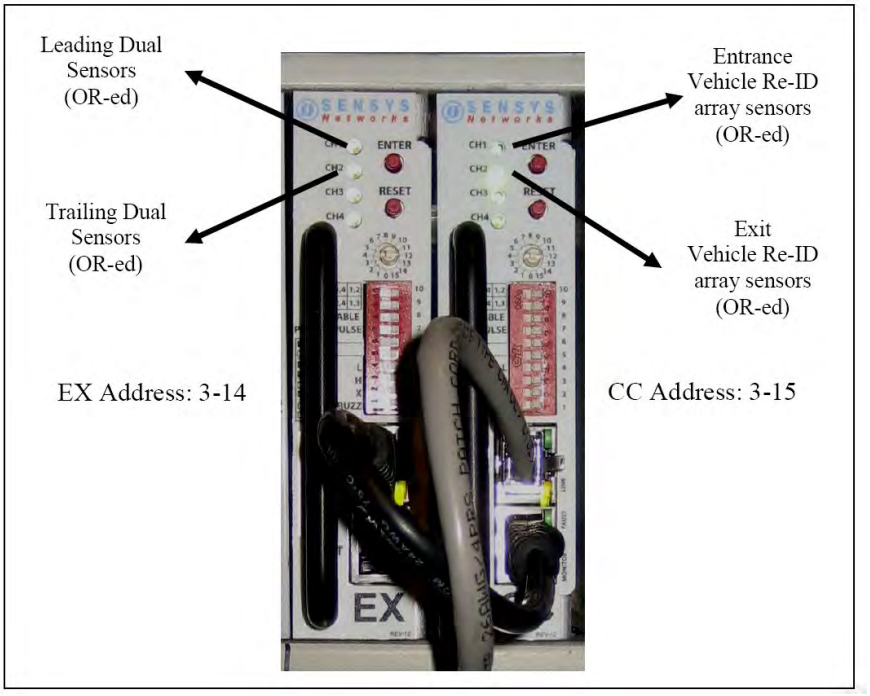


FIGURE 21 Sensys contact closure cards channels to sensors mapping

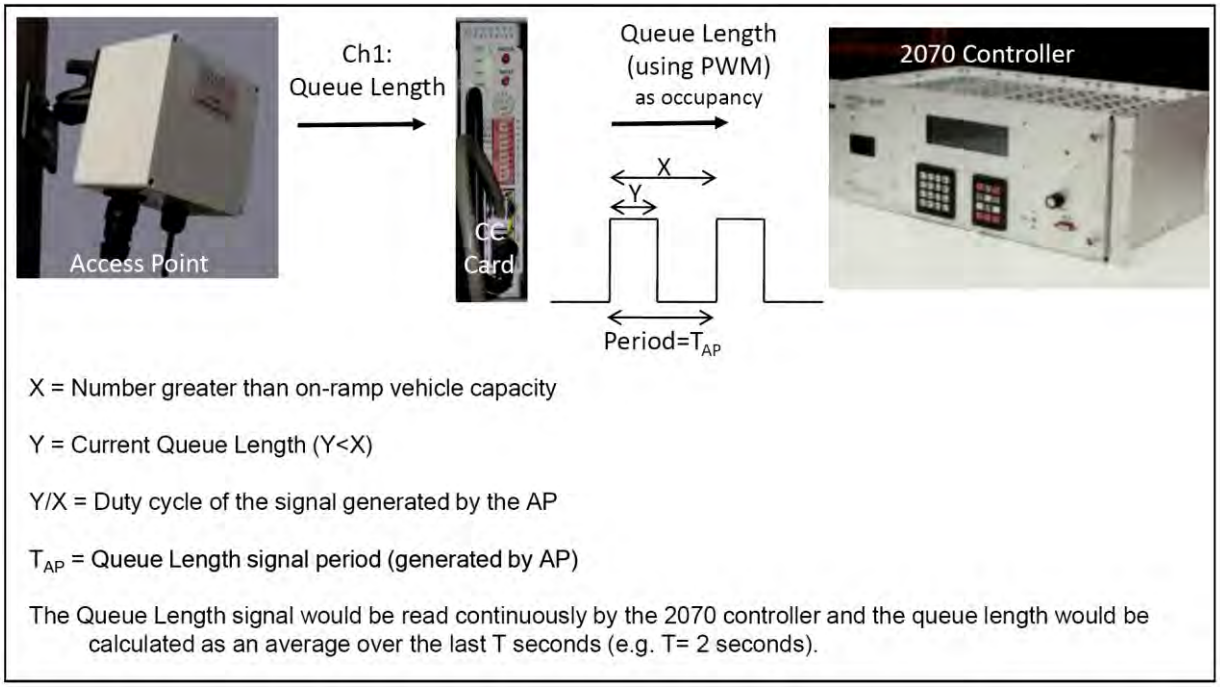


FIGURE 22 AP/Controller interface: old approach

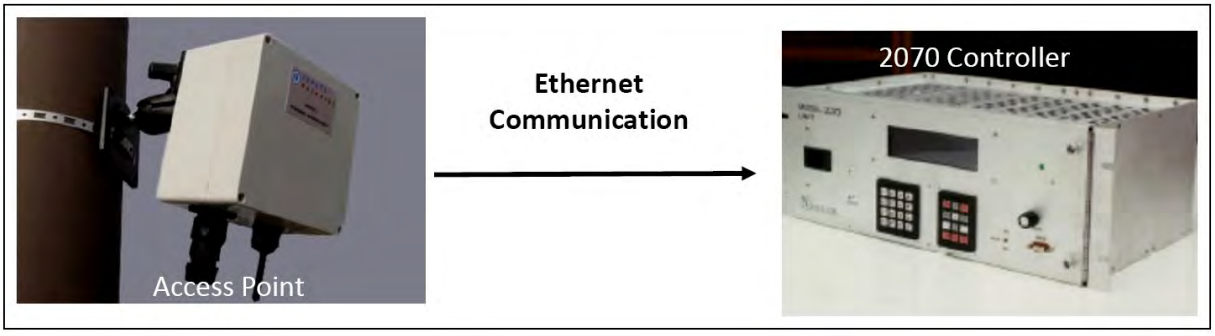


FIGURE 23 AP/Controller interface: new approach

URMS Source Code Analysis

The field test was designed to be completed by replacing the 170E traffic controller in the Hegenberger on-ramp by a 2070 controller running a modified version of the URMS. Due to the structure of the software, the timing module, one of the many modules of the URMS, needed most of the modifications for the successful implementation of the queue regulation algorithms. The timing module handles traffic calculations and meter timing and includes the implementation of queue override. David J. Wells released to us the source code of the URMS under a confidentiality and non-disclosure agreement so that we could study it, modify it and use it to complete the project.

Part of the analysis completed for this task, was the detailed study of the timing module. In this report, specific details of the URMS source code will not be disclosed, since it was released under a confidentiality and non-disclosure agreement. We will only illustrate that due to

the structure of the timing module, implementing the proportional plus integral (PI) queue control algorithm would not require any major modification to the URMS source code.

The PI queue regulator algorithm can be considered as a function that takes as its inputs the previous metering rate determined by the controller as well as the current and previous queue lengths at the on-ramp. The output of this function is a metering rate based on the proportional plus integral controller explained in Section 9. In relation to measurements from the field, the difference between this algorithm and queue override is that instead of using the occupancy measured with the queue detector, the PI queue regulator needs the queue length at the on-ramp to update the metering rate. In order to implement this algorithm, queue length would need to be communicated from the Sensys AP to the URMS, as occupancy is communicated to the URMS through the loop detector cards installed in the controller cabinet. In order for this function to be successfully included in the URMS as another method of queue regulation, it would have been necessary to establish the communication protocol between the AP and the 2070 controller and to include the queue length estimate in the data structure used in the URMS so that the information could be passed to the PI queue regulator function.

The PI queue regulator algorithm was to be included as another option for queue regulation, queue override being the standard method already implemented in the URMS. The timing module would handle the ramp metering rate determined by the PI regulator exactly in the same way as it currently processes the metering rate output by the queue override algorithm. In this sense, the modifications to the timing module are very specific and do not require any major change to the source code. Nevertheless, the actual implementation of this queue regulation algorithm was very dependent on the queue estimate to be used for the later stages of the project, since that would determine additional modifications to the timing module and possibly other modules of the URMS. Irrespective of the queue estimator chosen, a knowledgeable C programmer that is familiar with the URMS source code would need to spend a few hours making the required modifications, since the code would only require the inclusion of a new option for queue regulation and the handling of the queue length estimate. On the other hand, someone unfamiliar with the URMS source code may need to invest considerable amount of time trying to understand the structure of the code, making the modifications, and compiling the source code to generate modules that can be loaded to the 2070 traffic controller.

9. TASK 7

The purpose of this task was to understand how the length of the queue would be affected by the interaction of the PI queue regulator algorithm with ALINEA, which would be different than the interaction that would exist between ALINEA with queue override. In particular, we wanted to show the advantages of using a PI queue regulator in comparison with queue override, which is currently accepted in California as a reliable method to avoid queue interference with arterial streets. With the following analysis, it would be shown that a PI queue regulator gives advantages in terms of utilization of on-ramp storage capacity when it is required, e.g. when the freeway system is at capacity or under congestion, while avoiding queue interference with arterial street traffic.

Queue Override

In California, queue override is the process used to adjust the metering rate determined based on freeway conditions, e.g. time of day or traffic responsive, in order to shorten the queue length at the on-ramp. Queue override is used in California, and implemented in the URMS, to prevent vehicles in the metered lanes from backing up and interfere with arterial streets. This queue regulation method uses the occupancy measured with the queue detectors, installed at the entrance of on-ramps, to monitor the state of the queue. In the queue override method, the occupancy measured by the queue detector is compared to a threshold previously specified. When the occupancy is greater than the threshold, it is assumed that the queue is long and should be shortened. As a result, the metering rate determined by the ramp metering algorithm is increased by a fixed amount every metering rate update time interval (e.g. 30 seconds). This process is repeated until the occupancy drops below another occupancy threshold (which can be the same as the first threshold) or until the metering rate saturates to the maximum allowed metering rate. In either case, once the vehicle queue dissipates and the occupancy drops, the metering rate becomes the rate determined by the ramp metering algorithm in place. As it is discussed in (*2*), queue override has been noted to lead to oscillatory behavior and underutilization of on-ramp storage capacities, and for this reason they introduce a proportional-integral (PI) queue controller as the solution to eliminate the oscillatory behavior observed. A detailed analysis was performed using simulations, and it was one of the objectives of this project to validate some of their results in an actual freeway on-ramp.

PI Queue Regulator

In order to understand the PI queue regulator, it is important to describe the queue dynamics, and illustrate how the queue controller helps regulate the queue under different mainline/on-ramp conditions. FIGURE 24 shows a diagram that helps to understand queue dynamics and how the design of the PI queue regulator comes into play. The queue at every calculation interval can be described as

$$l(k+1) = l(k) + T\left[d(k) - r(k)\right],$$

subject to $l \geq 0$, where $l(k)$ is the queue length at time k, T is the update time period, $d(k)$ is the demand (flow entering the queue), $r(k)$ is the on-ramp flow entering the mainline (the flow leaving the queue), and l_{ref} is the target queue length, which generally is set close to the capacity of the on-ramp.

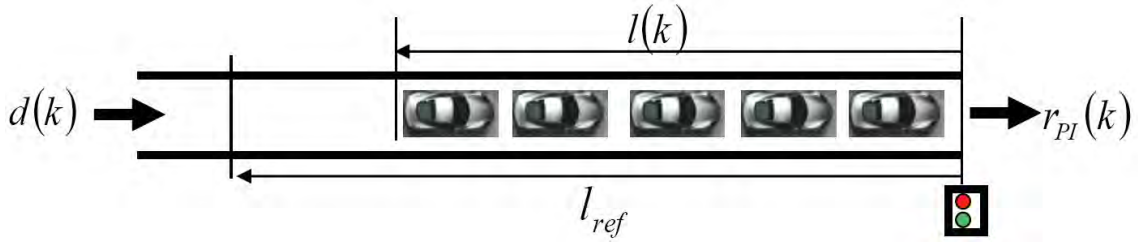


FIGURE 24 Queue length diagram

In order to implement a PI queue controller, it is necessary to choose *r(k)* with the following transfer function:

$$r_{PI}(z) = \left(k_p + \frac{k_I}{z-1} \right) \tilde{l}(z)$$

where $\tilde{l}(z)$ is the queue length error given by

$$\tilde{l}(z) = l_{ref} - l(z).$$

This algorithm can be easily implemented using the formula

$$r_{PI}(k) = k_p \left(l_{ref} - l(k) \right) + r_I(k),$$

where *k* represents the index of the metering rate update time interval and $r_I(k)$ is given by

$$r_I(k) = k_I \left(l_{ref} - l(k-1) \right) + r_I(k-1).$$

For the implementation of the queue regulator it would be necessary to use anti-windup and saturation mechanisms as it is suggested in (2). In the design of this regulator, *d(k)* is treated as an external disturbance and *r(k)* as the control variable. The closed-loop sensitivity function from the disturbance to the queue length error is

$$\frac{\tilde{l}(z)}{d(z)} = \frac{T(z-1)}{(z-1)^2 - Tk_p z + (k_I - k_p)T}.$$

As discussed in (2), proper PI gains k_P and k_I can be selected using the root-locus method.

ALINEA

ALINEA is a local traffic responsive ramp metering algorithm. The metering rate, r_A, is calculated by the following formula:

$$r_A(k) = r(k - 1) + K\left(\hat{o} - o(k) \right)$$

where $\hat{o}$ is the desired mainline occupancy, generally set to the critical occupancy, K>0 is a control gain, and $r(k)$ is the measured on-ramp flow merging into the mainline at time *k*. Note that the performance of ALINEA is not very sensitive to the value of the control gain K. The ALINEA equation shown above has an anti-windup mechanism that prevents unnecessary saturation of the metering rate, especially when on-ramp demand is low and it is not possible to make the occupancy at the mainline converge to the critical occupancy, the occupancy at which congestion is about to occur.

Ramp Metering with Queue Control

In order to implement ramp metering using ALINEA with a PI queue controller it is necessary to calculate two independent metering rates. The metering rate based on freeway traffic conditions is determined with the ALINEA algorithm, summarized in (*1*) and denoted in this report by r_A. On the other hand, the PI queue regulator determines a metering rate based on the traffic conditions at the on-ramp, and it is denoted in this report by r_{PI}. For every metering rate update time interval, the metering rate to be implemented in the field based on ALINEA with PI queue regulation, denoted in this report by r_{A+PI}, is the less restrictive of the two metering rates:

$$r_{A+PI} = \max\left[r_A, r_{PI}\right]$$

This is the same formula used in (*2*) to combine ALINEA with queue control. The metering rate to be implemented in the field, which for this report is represented by r, must belong to the range $[r_{min}, r_{max}]$, which represents the constraints imposed on the metering rate by Caltrans. If $r_{A+PI} < r_{min}$, then the metering rate is set to r_{min}, if $r_{min} \leq r_{A+PI} \leq r_{max}$, then the metering rate is set to r_{A+PI}, otherwise the metering rate is set to r_{max}. Note that the metering rate restrictions reduce the effectiveness of ALINEA and the PI regulator to accomplish their objectives. However, ALINEA with PI queue regulation has the potential to improve traffic conditions even under constrained ramp metering.

Analysis

Table 4 shows the different metering rates that would result with ALINEA, ALINEA with queue override, the PI queue controller, and ALINEA with PI queue regulation under different mainline/on-ramp conditions. This table shows that for the most part, ALINEA with PI queue regulation behaves in the same way as ALINEA with queue override. However, under two mainline/on-ramp traffic conditions highlighted in gray in Table 4, the metering rate determined by the algorithms is likely to be very different.

When the mainline is in free flow at capacity and the length of the queue is around capacity, the occupancy at the queue detector would most likely be above the threshold that triggers the queue override mechanism. For this reason, when the queue is around capacity, even if there is no interference with arterial street traffic, queue override would be activated and the metering rate would tend to saturate to r_{max}. Note that when the mainline is in free flow conditions at capacity, the successful application of the metering rate set by ALINEA is crucial to be able to maintain the mainline performing efficiently. Under this condition, it is essential to have the possibility to store as many vehicles as possible at the on-ramp if required. Nevertheless, the queue override mechanism is not able to maintain the queue length at capacity, and instead tends to discharge the queue into the mainline. When queue override is active, the ALINEA metering rate is ignored and instead a spike in the flow of vehicles into the mainline is induced. This spike in the inflow of vehicles to the mainline can be enough to trigger a change in the mainline conditions from free flow into congestion, which is undesirable and undermines the purpose of having ramp metering in the first place. On the other hand, the PI regulator does not create a spike on the flow of vehicles going into the mainline. Instead, this algorithm tries to

maintain the queue length at l_{ref}, which for this analysis we set to the capacity of the ramp. The PI regulator will try to maintain the length of the queue at capacity, which would result in the metering rate, $r_{PI,}$ convergence to the demand, *d(k)*. With this approach, the metering rate determined from ALINEA is more likely to be implemented at the on-ramp, as long as it satisfies $r_{min} \leq r_A \leq r_{max}$ and assuming that the on-ramp demand is low. Under this mainline/on-ramp condition, the metering rate in the field would be given by

$$r_{A+PI} = \max[r_A, d].$$

If the demand on the ramp increases and leads to a significant increase in the queue length, the PI queue regulator would dominate the metering rate and would converge towards r_{max} until the queue length is stabilized around l_{ref}. In this mainline/on-ramp mode, it is clear that using ALINEA with PI regulation is more advantageous than using ALINEA with queue override, since storage is utilized if required by the ALINEA algorithm while assuring that the queue stays around capacity and does not interfere with arterial street traffic.

When the mainline is congested and the queue is around capacity, both algorithms behave differently. In the ALINEA with queue override algorithm, queue override, as explained earlier, would be active and would saturate the metering rate to r_{max}. The queue override would dominate the behavior of the metering rate used in the field and ALINEA would be disregarded, since under congested mainline conditions, the metering rate set by ALINEA would tend to converge towards r_{min} in order restrict the inflow of vehicles into the mainline to allow the system to transition from a congested to free flow regime. Under the same mainline/on-ramp conditions, PI regulation offers an advantage over queue override because it minimizes the amount of vehicles that are allowed into the mainline. The metering rate would be set to *d(k)* instead of r_{max}, while making sure that the queue length does not interfere with arterial street traffic. Note that if the demand, *d(k)*, becomes larger than r_{max} , then the PI regulator would behave exactly like queue override.

Table 4 ALINEA with Queue Override vs ALINEA with PI Queue Controller

Freeway Conditions	***Queue Length***	***ALINEA (r_A)***	***Queue Override***	r_{A+O}		***PI Queue Regulator (r_{PI})***	r_{A+PI} ($max[r_A, r_{PI}]$)
Free Flow	**Below Capacity**	$\rightarrow r_{max}$	Inactive	r_A		$\rightarrow r_{min}$	r_A
	Around Capacity		Active	r_A		$\rightarrow$d	r_A
	Beyond Capacity		Active	r_A		$\rightarrow r_{max}$	r_A
Free Flow (Capacity)	**Below Capacity**	$r_{min} \leq r_A \leq r_{max}$	Inactive	r_A		$\rightarrow r_{min}$	r_A
	Around Capacity		Active	$\rightarrow r_{max}$		$\rightarrow$d	max $[r_A, r_{PI}]$
	Beyond Capacity		Active	$\rightarrow r_{max}$		$\rightarrow r_{max}$	$\rightarrow r_{max}$
Congested	**Below Capacity**	$\rightarrow r_{min}$	Inactive	r_A		$\rightarrow r_{min}$	r_A
	Around Capacity		Active	$\rightarrow r_{max}$		$\rightarrow$d	r_{PI}
	Beyond Capacity		Active	$\rightarrow r_{max}$		$\rightarrow r_{max}$	$\rightarrow r_{max}$

Extensive simulation studies have been conducted using ALINEA with queue regulation. The scope of this task was not to repeat these simulation efforts, but to show that by using the PI regulator in the field, it is possible to avoid interference between on-ramp queues and arterial streets like the queue override would if used with any ramp metering scheme. The main

advantage of PI queue regulation over queue override is that oscillations due to complete vehicle queue discharge and on-ramp storage underutilization are avoided.

CONCLUSIONS AND RECOMMENDATIONS

On ramp queue estimation based on vehicle re-identification performed significantly better than the other methods after the vehicle re-identification algorithm was modified. With the modified algorithm it was possible to achieve higher matching rates, which allows for a better and rapid correction of errors in the queue estimate. With the improved vehicle re-identification queue estimation method, the error remains bounded and the queue estimate is able to track the ground truth queue length accurately under uncongested and congested on-ramp conditions. It should be noted that the performance of the queue estimate drops during saturated on-ramp conditions, but even then the estimate is acceptable for implementations with queue control algorithms. Queue estimation based on vehicle re-identification, as discussed in this report, could be successfully implemented with different types of vehicle detection systems including loop detectors, provided the detector can count vehicles and can perform the re-identification of vehicles reliably and in real-time.

Occupancy queue estimation method may be used to determine if the ramp is either empty or full, but it cannot be used to estimate the queue length accurately. This approach is highly dependent on the time calculation interval over which occupancy is calculated, and under on-ramp saturation conditions may yield misleading results due to vehicles' tendency to miss sensor detection zone while at rest.

Queue length based on vehicle counts from sensors, as it has been the case with loop detectors, is not an acceptable method to estimate the queue due to its inability to correct for errors like detector miscounts and offsets resulting from initial conditions.

The speed based queue estimation method seems to be capable of instantaneously determine the mode of the ramp, either unsaturated, saturated, or in transition. The results obtained for this queue estimation approach does not match those results obtained based on traffic simulations.

There were several tasks in this project that could not be completed due to circumstances that were partly beyond our control. Most notably, the queue control algorithm was not fully incorporated into the URMS and, as a consequence, a field test of ramp metering using queue control could not be accomplished.

In order to incorporate the queue control algorithm into the URMS and to conduct a ramp metering field test with queue control, it was necessary that Caltrans D4 approve the URMS for ramp metering testing and provide us with the specific version of URMS that it had approved. Unfortunately, to the best of our knowledge, Caltrans D4 approval of URMS for ramp metering testing has not yet taken place.

It should be acknowledged that Alan S. Chow, the Office Chief of Traffic Systems in Caltrans District 4 and other Caltrans D4 personnel, including Sean Coughlin, were very

enthusiastic about this project and provided us with invaluable assistance in setting up the test site. We also received invaluable assistant in learning how to use and modify the URMS from David J. Wells in Caltrans Headquarters. Finally, Sensys Networks had express its commitment to develop, in collaboration with David Wells, the communication protocol to make the queue length estimation data available to the 2070 controller. Although the URMS source code was made available to us by Caltrans headquarters (after we signed an NDA that explicitly prevented us from providing the code to anyone else), the source code was not available to Caltrans D4 personnel, and this seems to have slowed down the URMS ramp metering certification and approval process. Perhaps we should have been more proactive and persistent in communicating to both Alan Chow at Caltrans D4 and our project manager, Nathan Loebs, the difficulties and causes of delays that we were experiencing with this project.

In order to prevent delays or incomplete tasks in future projects that involve field tests that rely on the use of a 2070 controller with the URMS, we recommend that either the field test be conducted in a district were the URMS is already approved for deployment or testing of ramp metering or that the district were the field test will take place provides an explicit commitment to certify the URMS with a firm date.

ACKNOWLEDGEMENTS

This work was supported by California PATH through Caltrans grant TA-65A0320-15354. In addition, we thank Adrian Levy, Alan S. Chow, David J. Wells, Lester S. Lee, Mehran Lajevardi, Melissa Clark, Nathan Loebs, Sean Coughlin, Steve Hague and Zhongren Wang for their participation and help with the project, as well as their useful comments and suggestions.

REFERENCES

1. Papageorgiou, M., and A. Kotsialos. Freeway Ramp Metering: An Overview. *IEEE Transaction on Intelligent Transportation Systems*, Vol. 3, no. 4, 2002, pp. 271-281.

2. Sun, X., and R. Horowitz. Set of New Traffic-Responsive Ramp-Metering Algorithms and Microscopic Simulation Results. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1959, 2006, pp. 9-18.

3. Wu, J., X. Jin, and A. J. Horowitz. Experiment to Improve Estimation of Vehicle Queue Length at Metered On-Ramps. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2099, 2009, pp. 30-38.

4. Wu, J., X. Jin, and A. J. Horowitz. Methodologies for Estimating Vehicle Queue Length at Metered On-Ramps. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2047, 2008, pp. 75-82.

5. Vigos, G., M. Papageorgiou, and Y. Wang. Real-time Estimation of Vehicle-Count Within Signalized Links. *Transportation Research Part C*, Vol. 16, no. 1, 2008, pp.18-35.

6. Vigos, G., and M. Papageorgiou. Relating Time Occupancy Measurements to Space-Occupancy, Density, and Link Vehicle-Count. *Tranportation Research Part C,* Vol. 16. no. 1, 2008, pp 1-17.

7. Kwong, K., R. Kavaler, R. Rajagopal, and P. Varaiya. Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors. *Transportation Research Part C: Emerging Technologies*, Vol. 17, no. 6, 2009, pp. 586-606.

8. Haoui, A., R. Kavaler, and P. Varaiya. Wireless Magnetic Sensors for Traffic Surveillance. *Transportation Research Part C: Emerging Technologies*, Vol. 16, no. 3, 2008, pp. 294-306.

9. Cheung, S. Y., S. Coleri, B. Dundar, S. Ganesh, C.-W. Tan, and P. Varaiya. Traffic Measurement and Vehicle Classification with Single Magnetic Sensor. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1917, 2005, pp. 173-181.