

1. REPORT NUMBER CA10-2022	2. GOVERNMENT ASSOCIATION NUMBER	3. RECIPIENT'S CATALOG NUMBER
4. TITLE AND SUBTITLE Mobile Century Final Report for TO 1021 and TO 1029: A Traffic Sensing Field Experiment		5. REPORT DATE December 2010
7. AUTHOR Alexandre M. Bayen, Ph.D., Principal Investigator		6. PERFORMING ORGANIZATION CODE
9. PERFORMING ORGANIZATION NAME AND ADDRESS California Center for Innovative Transportation 2105 Bancroft Way, Berkeley, CA 94720 Phone: (510) 642-4522 Fax: (510) 642-0910		8. PERFORMING ORGANIZATION REPORT NO. UCB-ITS-CWP-2010-4
12. SPONSORING AGENCY AND ADDRESS California Department of Transportation Division of Research and Innovation, MS-83 1227 O Street Sacramento CA 95814		10. WORK UNIT NUMBER
		11. CONTRACT OR GRANT NUMBER Contract # 65A0212
		13. TYPE OF REPORT AND PERIOD COVERED
		14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

Mobile Century proved that data from GPS-enabled mobile phones alone were sufficient to infer traffic features, i.e., to construct an accurate velocity map over time and space. The methods employed were able to function properly during both congested and free flow traffic conditions, and to detect correctly a traffic incident that occurred during the deployment. Another important contribution from this work was that ground-truth travel times were recovered by re-identifying vehicles captured on videotape. Therefore all results in this report can be asserted with high confidence. We conclude that the quality of data obtainable from present-day smartphones is adequate for useful, real-time traffic applications, such as calculating travel times. The architecture of the traffic monitoring system was designed such that identity information is encrypted and handled separately from traffic information, with no single entity having access to both. The spatial sampling strategy is based on the use of virtual trip lines that can be re-configured on-the-fly. This feature builds-in guaranteed flexibility for future monitoring needs. The new paradigm demonstrated in Mobile Century yet requires substantial effort to bring to fruition. Any industrial-grade, real-time system will require partnerships between government, academia, and industry. Business cases for future deployment must address incentives for public participation. In conclusion, Mobile Century was the first to demonstrate the near-term potential for using velocity data from GPS cell phones to reconstruct traffic state with precision. This opens the door for further research in this area to scale up the solution and to deliver considerable value to Caltrans and the traveling public.

17. KEY WORDS GPS-enabled mobile phones, flow traffic, real-time traffic data, traffic information	18. DISTRIBUTION STATEMENT No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161
19. SECURITY CLASSIFICATION (of this report) Unclassified	20. NUMBER OF PAGES 443
	21. COST OF REPORT CHARGED

CALIFORNIA CENTER FOR INNOVATIVE TRANSPORTATION
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

Mobile Century Final Report for TO 1021 and TO 1029: A Traffic Sensing Field Experiment Using GPS Mobile Phones

Alexandre M. Bayen, Ph.D., Principal Investigator

**CCIT Research Report
UCB-ITS-CWP-2010-4**

This work was performed by the California Center for Innovative Transportation, a research group at the University of California, Berkeley, in cooperation with the State of California Business, Transportation, and Housing Agency's Department of Transportation, and the United States Department of Transportation's Federal Highway Administration.

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification, or regulation.

December 2010

Acknowledgements

The authors would first like to acknowledge the support of our sponsors at Caltrans, especially Larry Orcutt, who took a risk to support a truly innovative effort. We are also thankful for the the guidance and consideration of Greg Larson, Hassan Aboukhadijeh, Asfand Siddiqui, and Gurprit Hansra.

The authors wish to thank everyone who participated in Mobile Century, including those who contributed to the initial brainstorming, those who mobilized during the course of its evolution, and those who forged ahead to propel this project to heights unimaginable just three years ago. The success of Mobile Century would have been impossible without the leadership, dedication, and foresight of J.D. Margulici, Thomas West, and Joe Butler.

We are grateful to the entire staff of the California Center for Innovative Transportation for the logistical planning, and successful implementation, of Mobile Century. In particular, we thank Coralie Claudel, Marika Benko, Osama Elhamshary, Tia Dodson, Chris Flens-Batina, Manju Kumar, Jed Arnold, Benson Chiou, Lori Luddington, Xiaohong Pan, Erica Sherlock-Thomas, and Arthur Wiedmer.

We appreciate the hard work of all the execution officers, Emma Strong, Jason Wexler, Jean Parks, Jennifer Chang, Kristen Ray, Negin Aryaee, Timmy Siau, Anurag Sridharan, Madeline Ziser, Matthew Vaggione, Qingfang Wu, Tarek Rabbani, Josh Pilachowski, Kristen Parrish, Megan Smirti, Carl Misra, Christina Sedighi, Jessica Ariani, Elizabeth Kincaid, Sandy Do, Nick Semon, Trucy Phan, Timothy Racine, Alan Wang, Charlotte Wong, Irene Kwan, Karl David Cruz, Swe Shin Maung, Tyler Moser, Alexis Clinet, and Julie Percelay, who assisted in planning, preparation, and operations on the day of the 100-vehicle deployment.

We are also indebted to members of the Nokia team: John Paul Shen, Bob Ianucci, Dave Sutter, and John Loughney; members of the CITRIS team: Gary Baldwin, Lorie Mariano, Paul Wright, Aaron Walburg, and Khossrov Taherian; members of the ITS team: Margaret Chang, Norine Shima, Jillene Bohr, John Li, and Ann Guy; members of the CoE team: Shankar Sastry, Lisa Alvarez Cohen, and Barbara Blackford; the student team: Annalisa Schiaccioli, Saurabh Amin, and Dengfeng Sun; the viability team: Patrick Saint-Pierre, and Jean-Pierre Aubin; and Sarah Yang at the UC Berkeley Media Office.

The authors would like to apologize in advance to anyone who contributed to develop, build, and deploy the traffic monitoring system implemented as part of the Mobile Century experiment but whose name we neglected to mention through our own oversight. You have our gratitude.

During the compilation of this report, we thank the many people who provided crucial subject matter through personal interviews, Skype interviews, original project documents, briefing materials, personal notes, emails, and other sources that made it possible to provide the rich

narrative herein. Finally, we thank Sara Bagwell for her work to compile and format this document.

Executive Summary

Traffic monitoring is most commonly accomplished with government-deployed, dedicated equipment. Adopting new technology in this paradigm can be costly and slow. However, recent advances in the mobile internet, cell phone technology, and location-based services may be leveraged to transcend the old paradigm. Doing so will reduce costs, increase coverage and yield a wealth of new data that will empower the traveling public with real-time access to current traffic conditions. Furthermore, transportation operators will gain access to an unprecedented wealth of information to help them better manage road networks.

Nonetheless, significant technical barriers and privacy concerns may impede widespread acceptance of a new paradigm. To understand and overcome these barriers, the Mobile Century experiment was conceived as a proof-of-concept demonstration of a traffic monitoring system based on probe vehicles equipped with GPS-enabled mobile phones.

The sheer scale of the experiment required significant logistical effort. A base station was erected at Union Landing, to house a temporary control center that was linked to a secondary control center in Palo Alto. Over one hundred graduate students from UC Berkeley were employed to circulate in loops along Interstate 880 between Hayward and Fremont, California, for an entire day. During the experimental deployment, an average penetration rate of probe vehicles was sustained near 3% (a significant logistical feat), which is viewed as realistic in the near future considering the increasing penetration of GPS-enabled cellular devices.

Classical methods of traffic modeling operate in the vehicular density domain, and use data such as occupancies and flows from inductive loop detectors. Understanding how to use velocity measurements instead was a significant technical contribution. In this work, the classical model was converted to the velocity domain, and GPS-based measurements were directly fed into the model.

Mobile Century proved that data from GPS-enabled mobile phones alone were sufficient to infer traffic features, i.e., to construct an accurate velocity map over time and space. The methods employed were able to function properly during both congested and free flow traffic conditions, and to detect correctly a traffic incident that occurred during the deployment.

Another important contribution from this work was that ground-truth travel times were recovered by re-identifying vehicles captured on videotape. Therefore all results in this report can be asserted with high confidence. We conclude that the quality of data obtainable from present-day smartphones is adequate for useful, real-time traffic applications, such as calculating travel times.

The architecture of the traffic monitoring system was designed such that identity information is encrypted and handled separately from traffic information, with no single entity having access

to both. The spatial sampling strategy is based on the use of virtual trip lines that can be re-configured on-the-fly. This feature builds-in guaranteed flexibility for future monitoring needs.

The new paradigm demonstrated in Mobile Century yet requires substantial effort to bring to fruition. Any industrial-grade, real-time system will require partnerships between government, academia, and industry. Business cases for future deployment must address incentives for public participation

In conclusion, Mobile Century was the first to demonstrate the near-term potential for using velocity data from GPS cell phones to reconstruct traffic state with precision. This opens the door for further research in this area to scale up the solution and to deliver considerable value to Caltrans and the traveling public.

Table of Contents

Acknowledgements.....	i
Executive Summary.....	iii
Table of Contents.....	v
List of Figures	xi
1 Introduction	1
1.1 Motivation.....	1
1.2 Historical Narrative	2
1.3 Scope	4
1.4 Summary of Findings.....	5
1.5 Organization of Report.....	6
2 Background	9
2.1 Traffic Monitoring with Dedicated Equipment (Road Infrastructure).....	9
2.2 Growth of the Mobile Internet	11
2.3 Barriers to a New Paradigm	12
3 Design Challenges	17
3.1 Preliminary Investigation	17
3.2 Scale of Experiment.....	23
3.3 Practical Considerations.....	25
3.4 Alternate Site.....	26
4 VTL Traffic Monitoring	29
4.1 Privacy Risks and Threat Model	29
4.2 Preserving Privacy with Virtual Trip Lines	31

4.3 Implementation..... 34

4.1 Experimental Deployment 38

4.2 Trip Line Placement..... 41

4.3 Estimating traffic 47

4.4 Discussion..... 53

5 Assimilating Lagrangian Data 55

5.1 Background..... 55

5.2 Description of Preliminary Concepts..... 57

5.3 Explanation of Proposed Methods..... 63

5.4 Assessment of the Methods..... 71

5.5 Numerical example of the NR nudging factor..... 81

6 Velocity-based Modeling 83

6.1 Related Work..... 83

6.2 Highway Traffic Flow Model..... 84

6.3 Speed Estimation..... 89

6.4 Implementation and Validation 90

6.5 Conclusion and Future Work..... 94

7 Video validation 95

7.1 Resolution Requirements and Camera Capabilities..... 95

7.2 Final Selection of Video Cameras..... 98

7.3 Practical Considerations for Camera Deployment..... 99

7.4 Trial Tests 99

7.5 Protocol for Camera Deployment 100

7.6 Narrative of Camera Deployment 103

7.7 Post-processing of Video Data 104

8 Probe Vehicle Deployment 107

8.1 Resources 107

8.2 Procedures 113

8.3 Collection of Probe Data 115

8.4 Narrative of Experiment..... 115

8.5 Data Pre-processing 118

9 Experimental results 121

9.1 Real-Time Traffic Monitoring Using Only Cell Phone Data 121

9.2 Trajectory data 122

9.3 Velocity Field 123

9.4 Ground-truth Travel Times..... 125

9.5 Reasons For Disparity Between Loop and VTL Data 129

9.6 Achieved Penetration Rate During Experiment 130

9.7 Inferring parameters from shockwave speed 132

10 Revisiting the Density Based Methods 135

10.1 Different Scenarios..... 135

10.1 Results 136

10.2 Conclusions..... 141

11 Discussion..... 145

11.1 Conclusions..... 146

11.2 Challenges 147

11.3 Future Goals 147

11.4 Steps Toward Deployment..... 148

References 151

Appendices..... 161

1 CPHS Protocol Narrative Form..... 163

2 Deployment Prototyping 203

 2.1 Strategy and Objectives 205

 2.2 Protocol for 20-Vehicle Deployment 211

 2.3 Instructions for Odd Drivers..... 219

 2.4 Instructions for Even Drivers 229

3 Contingency Plans..... 241

 3.1 Risk Management..... 243

 3.2 Emergency Response 255

 3.3 Directions for Phone Operators 259

4 Driver Briefings 271

 4.1 Red Team..... 273

 4.2 Yellow Team 281

 4.3 Orange Team 291

5 Driver Instructions 301

 5.1 Red AM route 303

 5.2 Red PM route 309

 5.3 Yellow AM route..... 315

 5.4 Yellow PM route..... 321

5.5 Orange AM route 327

5.6 Orange PM route..... 333

6 Field Experiment Protocol 339

7 Press Release Materials 405

7.1 Fact Sheet..... 407

7.2 Guest Program 411

8 Supplemental Tasks 417

8.1 AASHTO presentation..... 418

8.2 Mobile Millennium Planning, Design, and Server Development – Initial Stages..... 419

8.3 Mobile Millennium Arterial Modeling – Initial Stages 420

List of Figures

Figure 3.1: Vehicle trajectories from NGSIM. Shown in red, a flow fraction, θ , of trajectories are randomly designated as probes.....	18
Figure 3.2: One vehicle trajectory. Parameters are shown for real-time probe data reports.	19
Figure 3.3: Spatio-temporal dispersion of probe measurements for different combinations of penetration and sampling rates.....	20
Figure 3.4: Vehicle accumulations. Comparison of estimates using the Kalman filter method (top), and Nudging method (bottom). In both cases, the incorporation of Lagrangian data results in improved estimates over those using Eulerian data only.....	21
Figure 3.5: Vehicle accumulations. The higher the sampling rate, the higher the fidelity of the accumulation estimate.	22
Figure 3.6: Vehicle density estimated at the middle of the modeled expressway section.....	22
Figure 3.7: Travel time estimation for the modeled expressway section.	23
Figure 3.8: Alternate I-680 site.	27
Figure 4.1: Driving Patterns and Speed Variations in Highway Traffic.	31
Figure 4.2: Virtual Trip Line: Privacy-Preserving Traffic monitoring System Architecture. This system was implemented and ran for the entire duration of the 100-vehicle deployment of the Mobile Century experiment.....	33
Figure 4.3: Road networks extracted from Bay Area DLG files (Left) and Trip Lines per road segment in Palo Alto CA (Right).....	36
Figure 4.4: Comparison of the speed measurements recorded from the N95 (dots), the VTLs (boxes) and the vehicle speedometer (circles) as a function of time.	37
Figure 4.5: Satellite image of the first experiment site I-80 near Berkeley, CA. The red lines represent the locations of the VTLs, the blue squares show the speed recorded by the VTL, and the green squares represent the position and speed stored in the phone log. The brown circles represent the readings from the vehicle speedometer.	38
Figure 4.6: I880 Highway Segment for Twenty Car Experiment.....	40
Figure 4.7: Experimental Setup in a Car for Twenty Car Experiment.	40

Figure 4.8: Speed Measurements over Distance. 41

Figure 4.9: Speed Measurements over Time. 42

Figure 4.10: Linking prediction on a straight highway section. 43

Figure 4.11: Minimum Spacing Constraints for Straight Highway Section. 44

Figure 4.12: Linking attack near an on-ramp. 45

Figure 4.13: Actual travel times compared with an estimate given by the instantaneous method (30 second aggregation interval). 47

Figure 4.14: Travel time estimate errors by different sampling intervals using 15 VTLs. 49

Figure 4.15: Comparison between VTL-based spatial sampling and temporal periodic sampling against the same number of total anonymous samples. 49

Figure 4.16: Exclusion Area on Test Road Segment. Tracking starts from the point marked by star. 50

Figure 4.17: Spatial sampling and the benefit of an exclusion zone. 51

Figure 4.18: Travel time accuracy plotted vs. VTL spacing. 52

Figure 5.1 Highway US101 S used for the NGSIM dataset. 72

Figure 5.2 Schematic of the sampling strategy on equipped vehicle n. 73

Figure 5.3. Vehicle accumulation per cell for a) ground truth, and b) EDO case. 75

Figure 5.4 Vehicle accumulation (vehicles per cell) estimated using Newtonian relaxation method (left) and Kalman Filtering techniques (right) for scenarios 1 (top), 3 (middle), and 9 (bottom). 76

Figure 5.5 Total vehicle accumulation on the entire section. 77

Figure 5.6 Percentage of Improvement (PoI) in the RMSE as the number of Lagrangian measurements increases: (a) computing observed density using the fundamental diagram (Section 5.2.3), (b) using the actual density computed from vehicle trajectories as the observed density. Note that the two graphs are at different scales. 79

Figure 5.7 *True density* (computed from vehicle trajectories) versus *Observed density* (computed using the fundamental diagram as described in Section 5.2.3) for scenario 3 (left) and 9 (right). 80

Figure 6.1: Greenshields model. Left: Classical fundamental diagram (parabolic). Center: Linear relation between speed and density. Right: Flux function for the LWR-v PDE (6-4). The flux is parabolic with negative values. 86

Figure 6.2: Paramics velocity contours. Top: Ground truth velocity contour average across all vehicles. Bottom: Estimated velocity contour from the EnKF CTM-v algorithm (6-19) through (6-24) at 5% penetration rate. X-axis: position along highway in milepost; Y-axis: time of day.. 91

Figure 6.3: Error comparison of the EnKF CTM-v scheme, equations (6-19) through (6-24), (solid) and the averaging scheme (6-25), (dashed) using Paramics. Top: Relative error computed from (6-26) as a function of penetration rate. Bottom: Absolute error computed from (6-27) as a function of penetration rate. 93

Figure 7.1: Source image for resolution analysis..... 95

Figure 7.2: Simulated resolution for three camera types, assuming one camera per lane. 96

Figure 7.3: Simulated resolution for four camera types, assuming one camera for five lanes. .. 96

Figure 7.4: Image of license plate with blurring to simulate 2.0 cm resolution..... 97

Figure 7.5: Comparison of standard vs. HD camcorder image quality 98

Figure 7.6: Video frame with fence mesh between lanes. 100

Figure 7.7: HD Camcorder recording mode..... 100

Figure 7.8: Shutter priority switch. 101

Figure 7.9: Shutter control vs aperture control..... 101

Figure 7.10: Focus control..... 102

Figure 7.11: Custom image effects. 102

Figure 7.12: Controls for contrast and sharpness..... 103

Figure 7.13: Camera deployment during MC experiment. Stevenson and I880 (South bridge), Decoto and I880 (Central bridge), Winton and I880 (North bridge). 104

Figure 8.1: Traffic monitoring infrastructure built for field experiment. 109

Figure 8.2: Layout of Base Camp on the Union Landing Parking Lot..... 110

Figure 8.3: Tent Layout. 111

Figure 8.4: Study site and driver routes..... 112

Figure 8.5: Assigning drivers to routes and cars..... 113

Figure 8.6: Drivers logistics 117

Figure 9.1: Snapshot of the live traffic feed provided by the system in the present work (and from 511.org in the inset) at 10:52am on February 8, 2008. Traffic conditions after an incident on the northbound direction of I-880 are displayed. Numbers in circles correspond to speed in mph. 122

Figure 9.2: Vehicle trajectories in the northbound direction extracted from the data stored by 50% of the cell phones. The propagation of the shockwave from the accident can clearly be identified from this plot. The red lines in the close-up were drawn by hand by fitting a line through the points where trajectories change slope. 123

Figure 9.3: Loop detector locations along the northbound direction. Numbers indicating mileposts increase in the direction of traffic flow from left to right..... 124

Figure 9.4: Velocity fields in mph using: (a) 17 loop detector stations; (b) vehicle trajectories and Edie’s generalized definition; (c) 17 VTLs at the loop detector locations; and (d) 30 equally spaced VTLs..... 125

Figure 9.5: Travel time (in minutes) between Decoto Rd. and Winton Ave. The x-axis indicates arrival time at Decoto Rd. Dots correspond to individual vehicle travel times (4268 in total), collected manually using video. Black dash-dotted lines correspond to the standard deviation of travel times obtained from video cameras in 5-minute windows. 126

Figure 9.6: Loop detector vs. VTL velocity measurements (all locations). Dotted lines are the ± 5 mph thresholds.s 127

Figure 9.7: Loop detector and VTL velocity data collected at: (a) milepost 21.3, downstream of Decoto Rd.; (b) milepost 22.5, half-way between Alvarado Blvd. and Alvarado Niles Rd., (c) milepost 27.3, the most downstream detector near the Winton Ave. exit; and (d) milepost 24.0, downstream of Whipple Rd. Subfigure (e) shows the penetration rate at each of these four locations during the day. 129

Figure 9.8: Penetration rate map..... 131

Figure 9.9: (a) and (c): Average penetration rate over time at existing detector station locations during the morning and the afternoon. The range is one standard deviation below and over the mean. Traffic flows from left to right. (b) and (d): Histogram of the penetration rate including all the 17 locations during the morning routes and the afternoon routes, respectively. 132

Figure 10.1; Density field (in vpm) using 17 loop detector stations deployed in the section of interest (obtained through PeMS). 136

Figure 10.2: Density field (in vpm) using the Newtonian relaxation method (left) and the Kalman Filtering techniques (right) for scenario 1 (top), 2 (middle), and 3 (bottom). For each scenario, the boundary data is provided by loop detectors. 137

Figure 10.3: Flow comparison at mileposts (a) 21.3 (detector 1), (b) 24 (detector 7), and (c) 25.2 (detector 10). 139

Figure 10.4: Results of the quantitative analysis for the Newtonian relaxation method and the Kalman Filtering technique for scenario 1 (left), 2 (center), and 3 (right). Results obtained using loop detector data are also included for comparison. 140

A2.2 Figure 1: The 4 mile section to be used in the 20 cars experiment. 213

A2.2 Figure 2: Detail of the Alvarado-Niles interchange and the mall. 214

A2.2 Figure 3: Detail of the Tennyson Rd. interchange. 215

A2.2 Figure 4: Detail of the SR92 interchange. 215

A2.2 Figure 5: From I880S and I880N to the parking lot.. 216

A2.2 Figure 6: Lane numbering. 217

A2.3 Figure 1: From 1880S and I880M to the parking lot. 222

A2.3 Figure 2: From the parking lot onto the freeway. 223

A2.3 Figure 3: Lane numbering. 223

A2.3 Figure 4: Turn back at CA-92 (north end of the long loop). 224

A2.3 Figure 5: Turn back at Alvarado-Niles Rd. (south end of both the long and short loops). 225

A2.3 Figure 6: Turn back at Alvarado Blvd./Fremont Blvd. in case driver misses exit 23. 226

A2.3 Figure 7: Turn back at Winton Ave. in case driver misses off-ramp to CA-92W..... 227

A2.4 Figure 1: From I-880S and I-880M to the parking lot..... 232

A2.4 Figure 2: From the parking lot onto the freeway. 233

A2.4 Figure 3: Lane numbering..... 233

A2.4 Figure 4: Turn back at CA-92 (north end of the long loop). 234

A2.4 Figure 5: Turn back at W Tennyson Rd. (north end of the short loop). 235

A2.4 Figure 6: Turn back at Alvarado-Niles Rd. (south end of both the long and short loops). 236

A2.4 Figure 7: Turn back at Alvarado Blvd./Fremont Blvd. in case driver misses exit 23. 237

A2.4 Figure 8: Turn back at Winton Ave. in case driver misses off-ramp to CA-92W..... 238

A3 Figure 1: Architecture of mobile probe data collection / travel time calculation system ... 252



Mobile Century

Final Report for TO 1021 & TO 1029: A Traffic Sensing Field Experiment Using GPS Mobile Phones

Prepared by:

Anthony D. Patire

Alexandre M. Bayen
Xuexang (Jeff) Ban
Christian Claudel
Marco Gruteser

Daniel B. Work
Quinn Jacobson
Ali Mortazavi
Murali Annavam

Juan C. Herrera
Olli-Pekka Tossavainen
Steve Andrews
Toch Iwuchukwu

Ryan Herring
Sebastien Blandin
Baik Hoh
Kenneth Tracton

For:

California Department of Transportation
Division of Research and Innovation

California Center for Innovative Transportation
University of California, Berkeley
2105 Bancroft Way, Suite 300, Berkeley, CA 94720-3830
Phone: (510) 642-4522 Fax: (510) 642-0910 <http://www.calccit.org>

1 Introduction

This final report for TO 1021 and TO 1029 documents the Mobile Century Project. At its heart, it is a narrative of painstaking preparation that ultimately culminated in an unprecedented deployment of 100-probe vehicles, and the subsequent analysis of the collected data. In a larger context, this is a story about how a singular experiment impacted the future paradigm of traffic monitoring.

This chapter begins with a discussion motivating the line of inquiry pursued in this work. A brief historical narrative follows, describing the initial actors, task orders, and subsequent amendments that shaped the outcome of Mobile Century. Within this context, the scope of this report is defined.

Key findings are summarized, and their significance is explained. In particular, the most promising directions for future research are identified, and near-term possibilities are explored. The chapter concludes with a description of the organization of this report.

1.1 Motivation

Expanding the scope and coverage of roadway Advanced Traveler Information Systems (ATIS) is a top-priority of Caltrans. Supporting statements for more and better traveler information across the state of California have come all the way from the Governor's office.

ATIS benefits the transportation system for at least two reasons. First, the availability of information enhances the service provided to travelers. Numerous studies reveal that commuters appreciate and value timely information, which reduces their uncertainty and their stress. Second, reliable information can arguably enable travelers to make educated choices about their itinerary, departure time or even transportation mode, with the result of bringing about system self-management. It remains to be established that system self-management can take place on a large scale and significantly impact network-level operations. However, at a more anecdotal level, information about an accident ahead or a scheduled ramp closure certainly influences driver decisions. An additional side benefit of ATIS is that it builds the awareness of the traveling public toward Intelligent Transportation Systems (ITS). Such awareness can translate into political support for ITS projects and enable more improvements in the long term.

One of the main pieces of ATIS content is undoubtedly travel time estimations. Travel times on selected itineraries represent information that is easy for the traveling public to understand and process. Travel times can be posted on freeway or arterial Changeable Message Signs (CMS) and reach a very large audience, as is currently done at dozens of locations in the San Francisco Bay Area and in Southern California. Estimating travel times, either at the present time or into the future, requires large amounts of good quality traffic data. Traditionally, traffic data is

collected by sensors such as inductive loops installed at fixed locations. While this method yields great results to estimate volume and occupancy, it does not provide accurate travel time information unless the sensor coverage is very dense. Traffic sensors also happen to be expensive to install and maintain. Therefore, except for some of the busiest corridors in the state, the data collected from fixed sensors is mostly inadequate for travel time estimation.

Appraising various methods of collecting traffic data and specifically travel times is a definite need for Caltrans. Besides providing the bulk of the content required for ATIS, travel times also represent precious data to Caltrans as a network operator. While travel times alone may not cover the full extent of the department's traffic data needs, accurate and reliable travel times can be used for both planning and operations purposes.

Over the past few years, a number of private industry vendors have approached Caltrans with solutions to collect travel time data on highways and city arterials. Solutions revolve around two basic concepts and trends. The first trend suggests leveraging new technologies that significantly lower the cost of fixed detection. Both in-pavement technologies such as wireless magnetometers from Sensys Networks, Inc. and off-pavement technologies such as radar-based sensors by Speedinfo, Inc. offer much more attractive price points than inductive loops and make it conceivable to augment detection to a level that would yield accurate traffic maps and travel time estimates. An alternative concept is to use so-called mobile traffic probes to measure travel times from actual trips. Mobile traffic probes are essentially vehicles that are tagged and tracked along a corridor. This concept can be implemented by toll collection tags and readers, or by automated license plate readers. In either of those two cases, travel times are collected for preset segments of roadways in-between readers. For instance, the San Francisco Bay Area 511 system relies for a large part on data collected from FasTrak readers.

In the past several years, cell-phone based technology has gained momentum as a promising avenue, although previous research and field tests have not been conclusive. This technology previously relied on positioning provided by cellular networks, which still has to overcome significant challenges. However, the introduction of GPS receiver chips into more and more handsets represents a new opportunity. The prospect of large numbers of GPS-equipped cell phones reporting position and speed with 10 meter / 3 mph accuracy at regular intervals represents a huge leap forward. Yet its implementation requires addressing key questions regarding individual privacy, data ownership, network load, and proper traffic flow estimation techniques. The emergence of new mediums to diffuse traffic information, such as in-vehicle telematics displays and GPS-equipped cell phones could bring about a shift in how travelers perceive and consume traffic information in years to come. As the State Department of Transportation, Caltrans needs to monitor and leverage this paradigm shift.

1.2 Historical Narrative

The present work originated from several distinct sources. One group was comprised of scientists at the University of California, Berkeley (UCB) who were investigating applications of mobile sensors. Another group from Nokia Research Center (NRC) in Palo Alto was interested

in social networking and location-based services that protect the privacy of participants. These two groups identified traffic monitoring as a potential area of overlap. Supported by a seed grant from the Center for Information Technology Research in the Interest of Society (CITRIS) they began to explore the rich milieu of research questions involving conflicting needs for both data collection and privacy preservation. Soon afterward, Nokia awarded the UCB group with another seed grant that was matched by the University of California under the MICRO program.

Independently from the aforementioned groups, another team at the California Center for Innovative Transportation (CCIT) wrote a proposal in December of 2006, entitled “Deployment of value-added mobile traffic probes.” This proposal was funded by Caltrans in January of 2007 under Task Order 1021. Although details of TO 1021 evolved over the years, the spirit remained unchanged.

Four crucial technologies for ATIS: (1) GPS, (2) GIS and digital maps, (3) internet networking, and (4) wireless data communications, are identified in the proposal. In addition, the proposal notes that these technologies have achieved levels of maturity and affordability that place the ATIS industry “on the verge of an unprecedented boom.”

Almost prescient in foresight, the proposal speculates that “research under this project may serve in creating a paradigm shift in traffic data collection, from fixed sensors paid for by the state to mobile probes deployed as part of a self-sustaining private business model.”

The proposal further notes that “in the past several years, cell-phone based technology has gained momentum as a promising avenue, although previous research and field tests are still not conclusive. Yet the salient point is that a 2-way communication device like a cell phone can provide data about the entire trip of a vehicle, rather than be limited to observations at specific locations. This potentially represents a considerable improvement over fixed readers in terms of both the resolution and the timeliness of the data being collected.”

Although recognizing the potential for cell phones as traffic probes, the technology choice for the original proposal was the Dash Navigation Unit. These units each included an accurate GPS receiver, a locally-stored comprehensive digital map, a complete historical traffic model for prediction, significant processing capability, and wireless communications built into a special purpose navigation appliance. Subsequent to the award, however, the provider of the navigation units withdrew from the project.

Under sponsorship from Caltrans, the team from CCIT joined forces with the UCB and NRC groups, who had already formed a strong partnership. The first readjustment of the research plan was to replace the Dash Navigation Unit with the Nokia N95 smartphone. These smartphones had lesser GPS capabilities and lacked the navigational aids made possible with locally-stored historical traffic, and digital mapping data. In contrast with the Dash Navigation Unit, the GPS-equipped N95 was more of a general purpose communication device. Note that the N95 was one of the first smartphones ever developed (before the iPhone) and is regarded as a precursor to today’s participatory sensing based traffic monitoring systems. At the time,

the limitations of the N95 platform forced the researchers to adopt a more focused and simple deployment plan.

The UCB and NRC groups contributed a substantially improved experimental protocol. For example, the underlying sampling scheme was made privacy-aware, as is further explained in Chapter 4. The privacy studies were performed in collaboration with a team from Rutgers University, with expertise in the field of privacy research, and funded by Nokia. For now, consider that this early consideration focused technology development in a direction more appropriate for widespread adoption. In addition, the details of the mobile probe deployment were improved to enable a scientific analysis of results. Rather than using a remote site, a well-studied section of I-880 was chosen to enable rigorous evaluation of the data quality from the GPS-enabled smartphones.

The outstanding success of the 100-vehicle deployment precipitated further expansions to the direction of research. Caltrans allocated a supplemental award (TO 1029), marshaling resources to process the enormous amount of data that were collected during the 100-vehicle deployment. In addition, efforts were refocused toward a second major deployment of probe vehicles at a scale of at least one order of magnitude larger than before. New partnerships (including subcontracts with Covaluate, an Information Technology Service Provider, and Rensselaer Technology Institute) were forged in accordance with the new directives, which will be the focus of the report on Mobile Millennium, a follow up to Mobile Century launched soon afterwards.

1.3 Scope

This is the final report for Task Order 1021 and Task Order 1029. The bulk of this work is embodied in what has become known as Mobile Century. In terms of task orders, this refers to the entirety of the work plan for TO 1029 (except for the small section entitled Mobile Millennium Traffic Server Development), and Task 2 as written in Amendment A (replacing the original task order, TO 1021).

All other tasks in TO 1021, TO 1029 and subsequent amendments fall in to one of three categories:

- (1) AASHTO presentation
- (2) Mobile Millennium Planning, Design, and Server Development
- (3) Mobile Millennium Arterial Modeling

We note here that the above tasks (2) and (3) related to Mobile Millennium were enormous in scope. The monies allocated from TO 1021 and TO 1029 toward these tasks amounted to a small fraction of the total required to bring these tasks to fruition. Herein we only document initial stages of work toward these tasks funded by TO 1021 and TO 1029. The ultimate success of tasks (2) and (3) lie outside the scope of Mobile Century, and will be addressed in the

forthcoming Final Report for Mobile Millennium (Agreement 65A0301). Tasks (1), (2), and (3) as they are relevant to TO 1021 and TO 1029 are reported in Appendix 8.

1.4 Summary of Findings

Exploring a new paradigm. Advances in the mobile internet, alongside current trends in cell phone technology and location-based services, place the field of traffic monitoring at the cusp of a new era in data collection. A new paradigm in which GPS-enabled smartphones supply the bulk of raw traffic monitoring data is very promising. Compared to the status quo, the costs of collecting traffic information would be drastically reduced, and coverage could be extended far beyond what is currently feasible with fixed detectors alone. The opportunity for government agencies is significant: the availability of data will empower the traveling public with real-time access to current traffic conditions, while transportation operators will gain access to an unprecedented wealth of information to help them better manage road networks.

A successful proof-of-concept. The successful 100-vehicle deployment presented in this report was conceived as a proof-of-concept for a traffic monitoring system based on GPS-enabled mobile phones. During the experimental deployment, an average penetration rate of equipped vehicles was sustained near 3%, which at the time of the experiment was representative of the 18-month growth forecast for the GPS fleet in the smartphone market.

Traffic reconstructed from smartphone data. Raw data from the GPS-enabled smartphones alone were sufficient to infer traffic features, i.e., to construct an accurate velocity map over time and space. Therefore, probe vehicles deployed during the Mobile Century experiment were evaluated as providing substantial added value. Since ground-truth travel times were recovered by re-identifying vehicles captured on videotape, these results can be asserted with high confidence. We conclude that the quality of data obtainable from present-day smartphones is adequate for useful, real-time traffic applications, such as calculating travel times.

VTL-based monitoring. As will be explained in Chapter 4, a data sampling approach using Virtual Trip Lines (VTLs) was designed and implemented. The VTL approach combined with a sustained 3% penetration rate of probes provided better data for travel time prediction than that of the PeMS loop detectors spaced at an average distance of 0.35 mi. Furthermore, the use of VTLs provides enough data for traffic monitoring purposes while protecting the privacy of participants. In addition to the privacy benefits, another key advantage of virtual trip lines over physical traffic sensors is the flexibility with which they can be deployed.

Challenges. As a business model, significant challenges yet remain. For example, participation of the traveling public is crucial for success. In order to create and maintain the desired service quality, a large number of participants must be recruited and sustained. To achieve this, the right incentives for participation are needed. Premature deployment would be counterproductive. One can imagine a worst case scenario in which a deployment fails for lack of public interest and participation.

Future work. The next iteration of this program includes efforts to extend Mobile Century in a number of ways. First, better methods are required for incorporating data from both static (loop detectors) and mobile sensors (GPS-enabled mobile phones). Inverse modeling and data assimilation algorithms aimed at identifying and circumventing potential deficiencies in available data are also necessary. Finally, the monitoring of arterials brings additional challenges that also require much future work. These issues will be explored in a follow up report for Mobile Millennium.

Deployment. Stated simply, any future deployment will require substantial research and development. We recommend an evolutionary progression of field operational tests, so that lessons learned during any particular iteration may be incorporated into subsequent efforts as the scope of the system is continually scaled up. To make this possible, technology infrastructure must be implemented to support the computational modeling that will be required. In addition, any future deployment effort must include a strong industry component, and proceed in a way that complements existing trends in mobile computing.

Future applications. Future work as a part of this program has direct application toward the strategic goals of Caltrans in the area of operational data collection. Data sharing modalities should be explored between industrial companies, Caltrans, and local public agencies. Although traffic data collected as part of the next iteration will be served back to the mobile phone users who originally generated the data, future iterations may disseminate information much more broadly (broadcast media, internet websites, personal navigation devices, and roadside changeable message signs).

1.5 Organization of Report

The limitations of the status quo, the possibilities enabled by the mobile internet, and the grand-scheme challenges to the new paradigm are introduced in Chapter 2. A substantial literature review is furnished in the broader context of these issues; this discussion applies beyond Mobile Century.

In contrast, Chapter 3 focuses specifically on how to build a medium-scale, one-day deployment of a proof-of-concept system. Efforts to determine the parameters of a workable experiment, and initial back-of-the-envelope calculations are described, thus setting the stage for everything that follows in this report.

Chapter 4 explores the challenge of building a traffic monitoring system that addresses the goals of (1) acquiring quality real-time probe data from GPS-enabled cell phones, and (2) protecting participants from privacy threats by design. Initial prototyping was performed to assess the software and hardware systems that were implemented and the quality of data obtained from the smartphones.

Assuming that real-time probe data is available, Chapters 5 and 6 describe algorithms to reconstruct the traffic state from that data. Assimilating Lagrangian data in the density domain

is the subject of Chapter 5. An alternate scheme is presented in Chapter 6, in which Lagrangian data is fed directly into a velocity domain model. This velocity domain model was ultimately the one adopted for the real-time, 100-vehicle deployment of February 8, 2008.

Chapter 7 stands alone as a narrative of the video validation effort. The selection of video camcorders, trial tests, deployment protocol, and post-processing procedures are described. This crucial aspect of Mobile Century is what made possible an objective comparison of the state-of-the-art, status quo monitoring system (based on inductive loop detectors) with the proof-of-concept implementation (based entirely on mobile probes).

An overview of the experimental protocol for the 100-vehicle deployment is described in Chapter 8. Explained are the employed resources, established procedures, gathered data, and post-processing efforts. Supplementary material, including more detailed logistics, schedule of execution, and emergency procedures are furnished in the Appendices.

Chapter 9 presents the experimental results of the 100-vehicle deployment, the cornerstone of Mobile Century. The trajectory data and reconstructed velocity fields are compared with the ground-truth supplied by the video cameras. Travel time estimated from loop detectors is compared with that estimated from probe data.

Chapter 10 revisits the density domain data assimilation methods that were introduced in Chapter 5. Additional evaluation of these methods is performed using the data from the 100-vehicle deployment.

This report concludes in Chapter 11 with an evaluation of the project and recommendations for future deployment of the new paradigm for traffic monitoring.

2 Background

This chapter begins with a discussion of the status quo in Section 2.1: traffic monitoring with dedicated equipment and sensing infrastructure. Experience shows that deploying, operating, and maintaining new technology in this paradigm is costly and slow. As explained in Section 2.2, advances in the mobile internet bring forth potential to leverage current trends in cell phone technology and location-based services to transcend the old paradigm. Discussed in Section 2.3 are the barriers that impede adoption of this new paradigm. In particular, technical barriers, and social acceptance issues related to privacy concerns are addressed.

2.1 Traffic Monitoring with Dedicated Equipment (Road Infrastructure)

Traffic monitoring with inductive loop detector (ILD) systems. ILD systems are the most common highway traffic monitoring tool, and have been in use for decades. The current highway monitoring system consists of wire inductive loops placed directly in the top layer of the pavement. When a vehicle passes over the sensor, it is recorded by a roadside controller. In the case of travel time (the most important performance metric to the driving public), these sensors suffer from some fundamental drawbacks.

ILD velocity estimation is inaccurate. ILDs are accurate sensors for flows (vehicle counts), but they often generate inaccurate velocity measurements. California's freeways are equipped with about 23,000 ILDs embedded in the pavement, accounting for roughly 8000 detector stations. Several of these stations feature a single inductive loop per lane, which cannot measure vehicle speed directly. Practitioners have attempted to create aggregate velocity estimates using the average length of a vehicle on the highway and the percentage of time the sensor is occupied. Even when the sensor is working properly, these estimates are particularly noisy (with estimates ranging from 20 mph to 120mph) for traffic flowing at greater than 50 mph [24]. This has lead researchers to develop algorithms to improve these single loop estimates [24, 49, 64, 83, 93]. In contrast, dual loops (composed of two successive inductive loops) compute velocity by matching the respective occupancy patterns. In practice, they also have been found to produce significant errors [24].

Loop detector stations are expensive to deploy and maintain. The cost of an ILD is roughly \$900-\$2000 depending on the type of the loop. More importantly, the direct and indirect costs of deployment are significant (staff to install the sensors, and corresponding impact on traffic). According to the PeMS system [91], only 65% of the detectors in California are working properly; the main causes of malfunction are problems with the controller. In [90] malfunction rates of loop detectors and their causes were studied using data obtained from loops on the same stretch of I-880 examined in the present work. The average malfunction rate was 21%, despite significant efforts to maintain system operations during the study.

RFID transponders for travel time measurements. Radio-frequency identification (RFID) transponders are often deployed to collect automatic toll payment, such as FasTrak in California

or E-ZPass in some states on the East coast. These transponders can also be used to obtain individual travel times based on vehicle re-identification [10, 116]. Readers located on the side of the road keep a record of the time the transponder (i.e., the vehicle) crosses that location. Measurements from the same vehicle are matched between consecutive readers to obtain travel time. This technology is successful only when drivers have an incentive to carry the transponder (such as sorter toll booth queues), and can only provide travel times between segments where the readers have been deployed.

Travel time measurement through LPR technology. License plate readers consist of high speed cameras that record the license plates of vehicles on the highway. As a vehicle passes multiple cameras, the travel time between the readers is computed. Although LPRs avoid the need for in-vehicle equipment, these systems are complicated to install, and require an additional camera for each lane of traffic to be monitored. The relatively high cost of the readers (in the \$10,000 range plus installation costs), have limited their widespread implementation. Example deployments include Traffic Master's passive target flow management (PTFM) on trunk roads in the United Kingdom [112], and Oregon DOT's Frontier Travel Time project [16].

Traffic monitoring with dedicated probe vehicles. Dedicated probe vehicles equipped with a Global Positioning System (GPS) device are capable of collecting information such as position, speed, and travel time. The work in [100] addressed some of the key issues of a traffic monitoring system based on probe vehicle reports, and concluded that they constitute a feasible source of traffic data. The work in [123] also investigated the use of GPS devices as a source of data for traffic monitoring. Two tests were performed to evaluate the accuracy of GPS as a source of velocity and acceleration data. The accuracy level was found to be good, despite limitations of the selective availability¹ feature that was imposed at the time of the study [92].

Deployment of probe vehicles. HICOMP [52] is an example of one small-scale deployment of dedicated probe vehicles using GPS devices to monitor traffic for some freeways and major highways in California. Unfortunately, dedicated probe vehicles equipped with a GPS device represent added cost that cannot be applied at a global scale. As pointed out by [74], the penetration of HICOMP is low and the collected travel times are not as reliable as other systems such as PeMS. Other approaches have investigated the possibility of using dedicated fleets of vehicles equipped with GPS or automatic vehicle location (AVL) technology to monitor traffic [17, 85, 102], such as FedEx, UPS trucks, taxis, buses or other dedicated vehicles. While industry models have been successful at gathering substantial amounts of historical data using this strategy, for example Inrix, the use of dedicated fleets always poses issues of coverage, penetration, bias due to operational constraints and specific travel patterns. Nevertheless, it appears to be a viable source of data, particularly in large cities.

¹ Selective availability is the intentional inclusion of positioning error in civilian GPS receivers. It was introduced by the Department of Defense of the U.S. to prevent these devices from being used in a military attack on the U.S. This feature was turned off on May 1, 2000.

Deployment of dedicated communication systems is slow. One policy intended to enable dedicated communication systems for the transportation network has achieved only limited deployment. On October 21, 1999, the Federal Communications Commission allocated 75MHz of spectrum as part of the US Department of Transportation's (DOT) Intelligent Transportation Systems (ITS) US-wide program, with mostly traveler safety, fuel efficiency and pollution in mind. The first industry-government supported standard followed on August 24, 2001, when ASTM's E17.51 Standards Committee voted 20-2 to base Dedicated Short Range Communication (DSRC) on a modification of the IEEE 802.11a specification, now named IEEE 802.11p. At the same time, the US DOT launched a plan that included the deployment of around 250,000 roadside DSRC radios, but only led to around 100 radios deployed for the entire US as of 2008 (mostly in Michigan and California). This example highlights the difficulty of creating a dedicated communication system for the transportation network.

2.2 Growth of the Mobile Internet

Smartphones as sensors of the built environment. The convergence of communication and sensing on multimedia platforms such as smartphones provides the engineering community with unprecedented monitoring capabilities. Smartphones include a video camera, numerous sensors (accelerometers, magnetometers, light sensors, GPS, microphones), wireless communication outlets (GSM, GPRS, WiFi, Bluetooth, infrared), computational power and memory. With the rise of the Android and the iPhone, this trend has now greatly accelerated. These phones can be used to listen to the radio, to watch digital TV, to browse the internet, to do video conferencing, to scan barcodes, to read PDFs, and the list is endless. The rapid penetration of GPS in smartphones is enabling device geopositioning and context awareness, which in turn is causing an explosion of Location Based Services (heavily relying on mapping) on the devices. For example, Nokia Maps displays theaters and museums near the phone, Google Mobile provides driving directions from the phone location, and the iPhone Travelocity shows hotels near the phone. Due to their portability, computation, and communication capabilities, smartphones are becoming useful for numerous applications in which they act as sensors moving with humans embedded in the built infrastructure. Large scale applications include everything from population migration tracking and traffic flow estimation to physical activity monitoring for assisted living.

The competition for probe traffic data collection as a proxy for the larger war to conquer the mobile internet. There has been a trend of increased levels of competition between cell phone manufacturers, network providers, internet service providers, computer and software manufacturers, and mapping companies. Following the transition from desktops to laptops to smaller and more portable devices, top companies in these industries are redefining themselves to remain relevant as the internet goes mobile. In the context of traffic monitoring, the examples below show the importance of information technology for transportation systems. In late 2007, Google made a move toward the phone industry with the launch of the Open Handset Alliance and the Linux-based Android platform (leading to the T-Mobile G1 Google phone). In part because of the pressure to use open platforms enhanced by the Google OS, Nokia, who manufactures 40% of the cell phones in the world, purchased Symbian, which

licenses the operating system running on more than half of the smartphones in the world. Nokia then established the Symbian Foundation, with the intention of unifying the platform and making it open-source (Apple also partially opened its iPhone OS to software developers with the release of a software development kit). To strengthen its own mapping capabilities, Nokia also bought Navteq, which is the largest mapping company in the world, following personal navigation device manufacturer TomTom's purchase of Tele Atlas, Navteq's chief competitor. Navteq in turn owns Traffic.com, one of the leading traffic data collection and broadcast companies. Its competitors include Inrix, which provides traffic data to Microsoft's web, desktop, and mobile applications.

Smartphones: a transformation from dedicated infrastructure to market-driven technology.

The scale at which cell phones are produced, and the rate at which they integrate new technology, is dramatic. The total number of cell phones worldwide exceeded three billion at the time of this project. Some European countries have a penetration rate of more than 150% (150 cell phones for 100 people), and forecast 1 billion smartphones by 2012. Nokia alone produces more than 13 phones a second; with the increasing penetration of GPS in the cellular phone fleet, cell phones will soon constitute one of the major traffic information sources available to the public. In North America and Europe, the overwhelming majority of commuters have a cell phone, potentially populating the entire arterial network with probe traffic sensors. Obviously, the use of cellular devices as traffic sensors has numerous benefits: (1) It is possible to leverage the market driven communication infrastructure already in place; (2) The spatio-temporal penetration of cell phones in the transportation network is increasing at an extremely fast pace; (3) The use of cell phones as traffic probes is device and carrier agnostic, leading to faster penetration; and, (4) Major car manufacturing companies already have cradles and interfaces with cell phones (for example BMW and the iPhone) in their new cars, so the sensing information gathered by modern cars can also be sent to such monitoring systems.

2.3 Barriers to a New Paradigm

New paradigm. The concept of using cell phones as sensors has the potential to usher in a new paradigm for traffic monitoring. Such systems promise to significantly improve coverage and timeliness of traffic information [5, 56, 58]. Near-term solutions will require GPS measurements to be fused with traditional sources of traffic information such as loop detectors, camera, and human reports. However, with sufficient penetration, this approach could potentially enable the collection of real-time traffic information over the complete road network, including arterials, at minimal cost for transportation agencies.

Barriers. Several studies have demonstrated the feasibility of probe based traffic estimation through analysis, simulations, and experiments [29, 37, 111, 120]. Yet many challenges, both technical and societal must be addressed.

2.3.1 Technical Barriers

Non-GPS based localization of cellphones is problematic. Multiple technological solutions exist to overcome the localization problem using cell phones. Historically, the seminal approach chosen for monitoring vehicle motion using cell phones (prior to the rapid penetration of GPS in cellular devices) uses cell tower signal information to identify a handset's location. This technique usually relies on triangulation, trilateration, tower hand-offs, or a combination of these. Several studies have investigated the use of mobile phones for traffic monitoring using this approach [13, 38, 81, 115, 117]. The fundamental challenge in using cell tower information for estimating position and motion of vehicles is the inherent inaccuracy of the method, which poses significant difficulties to the computation of speed. Several solutions have been implemented to circumvent this difficulty, in particular by the company Airsage, which historically developed its traffic monitoring infrastructure based on cell tower information [80, 104]. Based on the time difference between two positions, average link travel time and speed can be estimated. [119] conducted a field experiment to compare the performance of cell phones and GPS devices for traffic monitoring. The study concluded that GPS technology is more accurate than cell tower signals for tracking purposes. In addition, the low positioning accuracy of non-GPS based methods prevents its massive use for monitoring purposes, especially in places with complex road geometries. Also, while travel times for large spatio-temporal scales can be obtained from such methods, other traffic variables of interest, such as instantaneous velocity are more challenging to obtain accurately.

GPS based localization provides high quality data. Increasing numbers of smartphones or PDAs come with GPS as a standard feature. This technology can provide more accurate location information and thus more accurate traffic data such as speeds and/or travel times. Additional quantities can potentially be obtained from these devices, such as instantaneous velocity, acceleration, and direction of travel. In [38], cell phones are used for traffic monitoring purposes, and the need for GPS-level accuracy position information to compute reasonable estimates of travel time and speed is discussed. Furthermore, [118] and [119] concluded that if GPS-equipped cell phones are widely used, they will become a more attractive and realistic alternative for traffic monitoring. GPS-enabled mobile phones can potentially provide exhaustive spatial and temporal coverage of the transportation network when there is traffic, with the high positioning accuracy achieved by a GPS receiver.

Lagrangian vs. Eulerian information. While cellular phones provide an ideal bridge between the physical world (vehicle flows and dynamics on the road) and the information world (software systems monitoring the network), there is one major difference between the data collected by cell phones and traditional data, commonly used to estimate traffic in real time: the data collected by phones in cars is Lagrangian, i.e. gathered along cars trajectories, and not Eulerian, i.e., control volume based. This poses major challenges in building an information system for a cyberphysical infrastructure such as the transportation network. While a static loop detector or a camera (both Eulerian) can easily capture all vehicles going through the space monitored by the sensor, and therefore infer aggregate quantities (flows, counts, local speed), a Lagrangian

sensor can only monitor quantities following the vehicle, without direct access to flows, counts, etc.

Distributed models for the transportation network. Because GPS enabled phones measure velocity, or travel time between two consecutive GPS readings, constitutive models used to describe the evolution of the system need to incorporate these reading and bypass quantities which cannot be measured (density, flows, counts). The development of such flow models, for highways and arterials is still at its infancy. Techniques used for this include partial differential equations, queuing systems, and hybrid system models of flow equations.

Machine learning models to circumvent lack of geographical infrastructure information. Knowledge of signage, traffic light presence, and cycle information is difficult to procure. The presence of stop signs, lights, and their effect on traffic is not available from databases on a US-wide scale. Furthermore, they change too often to be incorporated into flow models. This difficulty has to be circumvented by machine learning algorithms capable of learning the flow features without knowledge of the detailed infrastructure, using techniques such as clustering analysis.

Inverse modeling and data assimilation. In the age of massive data collection, one of the most fundamental theoretical challenges associated with the reconstruction of traffic using mobile data will be the proper use of techniques to incorporate data into flow models or statistical models. The development of these techniques in fields such as oceanography or meteorology is relatively mature. For large-scale infrastructure systems, the state of modeling, model inversion and computation is still at its infancy, but promises significant breakthroughs in the near future.

Considerations for initially low penetration of equipped vehicles. As suggested in the literature [72, 94, 117, 118] field tests are needed to assess the potential of new technologies such as GPS-enabled mobile phones. Test deployments to assess the potential of traffic monitoring using cell phones go back to the advent of GPS on phones. In particular, the study of [30] investigates the deployment of 200 vehicles for an extended period of three months and the potential data that can be gathered from it. In light of that study, one of the main issues in experiments or pilot tests is the problem of penetration, i.e. percentage of vehicles equipped vs. total number of vehicles on the road.

Real-time, online and robust availability. Unlike the more permanent Eulerian detectors, to which data quality, reliability and performance indices can be easily attributed, the penetration of cell phones at a given location and time is highly variable. Before this type of monitoring becomes the standard, the participation of the public will be spatially and temporally unpredictable. This means that the algorithms used for estimating traffic must be robust to variability in penetration.

2.3.2 Privacy Issues and Societal Acceptance

Privacy concerns with a new paradigm. Traffic monitoring through GPS-equipped vehicles raises significant privacy concerns, because the external traffic monitoring entity acquires fine-grained movement traces of the probe vehicle drivers. These location traces might reveal sensitive places that drivers have visited, from which, for example, medical conditions, political affiliations, speeding, or potential involvement in traffic accidents could be inferred. Furthermore, the correlation of this data with existing records poses specific threats to the preservation of privacy.

Example of data granularity. A variety of sampling techniques can be used to collect data from GPS enabled mobile devices. In the case of the Nokia N95, the embedded GPS chip-set is capable of producing a time-stamped geo-position (latitude, longitude, altitude) once every three seconds. From this time and position data, the instantaneous velocity is produced by the phone at the same frequency. Over time, this vehicle trajectory and velocity information produces a rich history of the dynamics of the vehicle and the velocity field through which it evolves.

Risk of unintended re-identification. While this level of detail is particularly useful for traffic estimation, it can be privacy invasive, since the device is ultimately carried by a single user. Even if personally identifiable information from the data is replaced with a randomly chosen ID through a process known as pseudo-anonymization, it is still possible to reidentify individuals from trajectory data. For example, pseudo-anonymous trajectories have been combined with free, publicly available data sets to determine the addresses of participant's homes [54].

Data value: sensitivity vs. utility. The transmission of high frequency data without regard to location also wastes resources throughout the system, which can pose scalability problems. In addition to disclosing sensitive information, the trajectory information on small roadways near users homes are of lower value to the general commuting public than major thoroughfares such as interstates. Thus, collection of low utility and highly sensitive data should be avoided when sampling using mobile devices.

Spatially aware sampling and privacy. At the heart of such a system, privacy-by-design sampling techniques must be used to prevent privacy invasion. In addition to proper anonymous data collection and encryption, sampling the vehicles at locations which are privacy safe is key to ensuring the ongoing participation of the public that is needed for such a system.

Disincentives for participation. Realistically, future users will have the option to choose the terms under which they share location information. Without providing tangible benefit, or safeguards to insure that an acceptable level of privacy can be guaranteed, adoption will not be widespread among the traveling public.

Current studies regarding privacy concerns are inadequate. Traffic monitoring applications based on a large number of probe vehicles have recently received much attention [21, 57, 120].

None of these works have addressed location privacy concerns in such systems. Since most traffic monitoring applications do not depend on the specific identification information about probe vehicles, the anonymization of sensing information has been a solution in practical deployments [58, 59, 110]. Not surprisingly, recent analyses of GPS traces [47, 71, 73], have shown that naive anonymization by simply omitting identifiers from a location dataset does not guarantee anonymity. Unique parts of GPS traces may be exploited to re-identify individuals using multi-target tracking, k-means clustering, or fingerprinting approaches to identify computer systems.

Centralized architectures for privacy protection. Therefore, several stronger protection mechanisms have been investigated. The k-anonymity concept [99, 108] provides a guaranteed level of anonymity for a database, although some recent studies [69, 82] have identified weaknesses. For location services, the k-anonymity concept has led to the development of centralized architectures that temporally and spatially cloak location-based queries [41, 46, 84]. This present work, in comparison, concentrates on providing privacy without requiring a single trustworthy entity.

Other best effort approaches require a trustworthy server. There are many best effort approaches [15, 65] that degrade information in a controlled way before releasing it. These approaches can be implemented in a centralized architecture or a decentralized approach. Many best effort approaches successfully preserve the privacy of users in high density areas, but they do not guarantee the privacy regardless of user density and user behavior pattern. [55] proposes the uncertainty-aware path cloaking algorithm to provide guaranteed privacy regardless of user density, but this again requires the existence of a trustworthy privacy server.

Perturbation and access control. Anonymous communication systems (e.g., onion routing [31, 43]) use a similar approach of distributing knowledge over several mixes. Random perturbation approaches for privacy-aware data mining [6, 7], which perturbs the collected inputs from users to preserve privacy of data subjects while maintaining the quality of data, are not applicable for time-series location data since noise with large variance does not preserve sufficient data accuracy, while noise with small variance may be filtered by tracking algorithms due to the spatio-temporal nature of the data [70]. Access control methods [39, 121] restrict access to data to permitted users. However, these techniques do not fully address the dishonest insider challenge. Further, they are not applicable to business models where the aggregated data is transferred to third party.

3 Design Challenges

This chapter furnishes an overview of how practical requirements shaped the one-day, experimental deployment of probe vehicles. Logical steps are retraced, and initial back-of-the-envelope calculations are described. The goal is to design an experiment to capture the essence of what a new traffic monitoring paradigm might look like. How might one address the key barriers outlined in Chapter 2? What constitutes a proof-of-concept? How might one build a miniature version of a near-term deployment? Projections suggested that in five to ten years, a substantial fraction of cell phones will be equipped with GPS receivers. In such a world, how accurately can traffic conditions be reconstructed in real-time with only GPS-enabled cell phone data? The experiment was designed to answer these types of questions.

Achieve desired penetration rate by design. A survey of previous work revealed that one key issue of smartphone based systems is the penetration rate. The penetration rate is defined as the fraction of vehicles (by flow) that act as probes to provide traffic data. The traffic state cannot be estimated to any useful precision when data is too sparse; i.e., the penetration rate is too low. From a traffic monitoring perspective, any successful experiment must maintain a penetration rate above some threshold. Previous studies reported that data coming from about 3% to 5% of the total flow are sufficient to obtain accurate estimates of the travel time [100, 115, 119]. This penetration rate is also realistic as a near-term future possibility. The present work distinguishes itself from previous studies in that a sufficiently high penetration rate was achieved by the design approach described in this chapter.

Addressing barriers to the new paradigm. Questions motivated by the discussion of Chapter 2 are reiterated here. The quality of data acquired from a privacy-aware sampling scheme is to be investigated. Is the information content of this data appropriate for reconstructing traffic states in practice? Are the algorithms employed able to reconstruct traffic states with adequate precision? How does the experimental system compare with a state-of-the-art system using ILDs? Specifically, how do both systems compare with ground-truth travel times, and how might one acquire that ground truth?

3.1 Preliminary Investigation

Confirm necessary penetration rate with NGSIM data. Trajectories from the NGSIM² data set provide accurate ground truth for all vehicles traveling along a 2000 ft stretch of expressway for a duration of 45 minutes. Measures such as vehicle accumulation and exact travel times can be calculated directly from the ground truth. The problem to be solved is to sample only a limited amount of information from the original data set, reconstruct the traffic flow, and to estimate vehicle accumulations and travel times based on the reconstruction. The estimated accumulations and travel times are compared with the ground truth, and errors are quantified.

² <http://ngsim.camsys.com/>

Consider Kalman Filtering and Newtonian relaxation algorithms. The accuracy of the estimates is limited by two factors, the quantity of information, and the algorithms for traffic reconstruction. At this early planning stage, sampling was assumed to occur at a fixed rate in time (temporal based sampling), and modeling was assumed to be performed in the density domain. Two algorithms were considered, Kalman Filtering and Newtonian relaxation (the latter is also called the “nudging method,” borrowed from oceanography). These two methods are described in detail in Chapter 5, and further evaluated in Chapter 10. This chapter furnishes a description of strictly preliminary findings used during the initial design stages of the experiment.

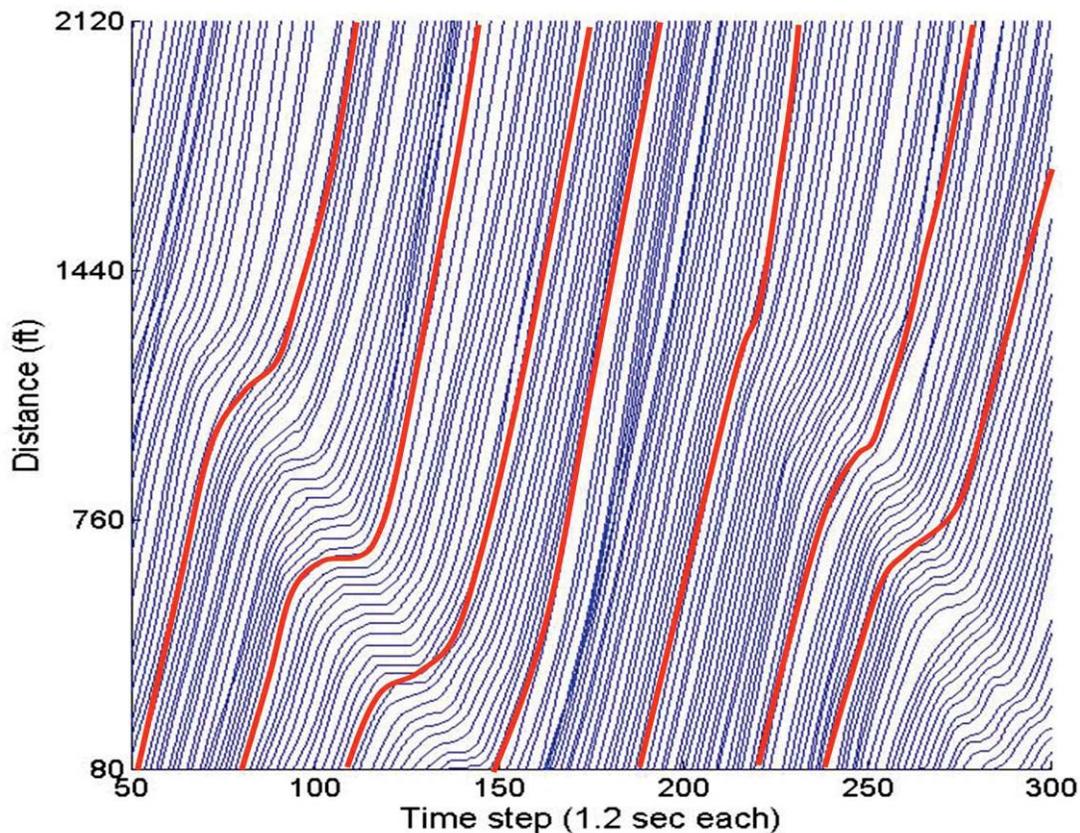


Figure 3.1: Vehicle trajectories from NGSIM. Shown in red, a flow fraction, θ , of trajectories are randomly designated as probes.

Problem formulation. The problem is posed in the following way. Assume that a fraction θ of vehicles on the expressway are equipped with GPS-enabled phones. In Figure 3.1, these equipped vehicles, shown in red, flow along with general expressway traffic, shown in blue. An example trajectory from one probe vehicle is displayed in Figure 3.2. It is assumed that the equipped vehicles can calculate average velocity over a time period τ . In addition, the probes report their velocity and position once every T seconds. The penetration rate θ and the sampling rate T determine the quantity of data available to the reconstruction algorithms. For

each of the four test scenarios defined in Table 3.1, the number of observation samples arising from each parameter set is listed.

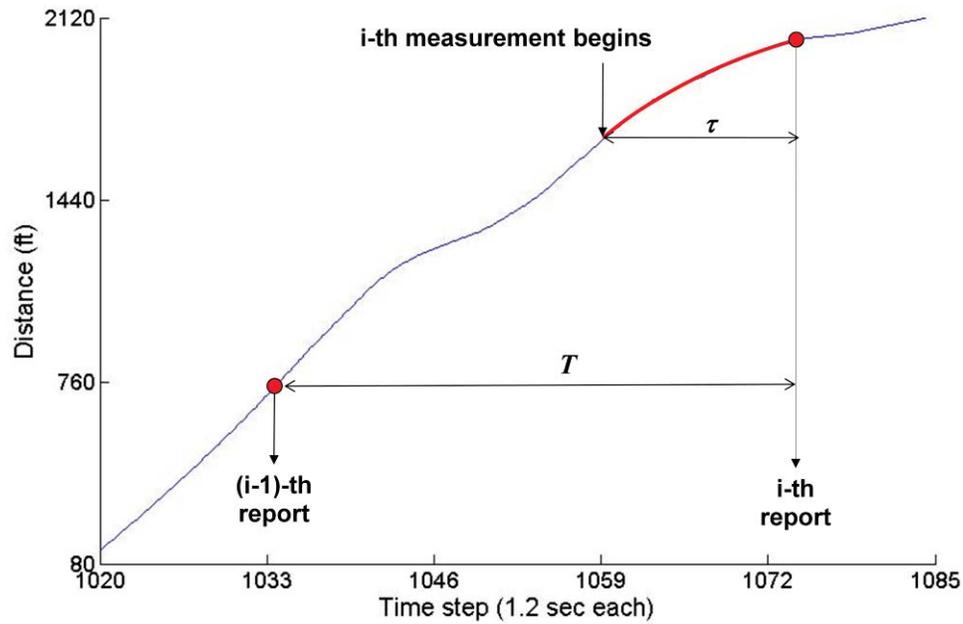


Figure 3.2: One vehicle trajectory. Parameters are shown for real-time probe data reports.

Table 3.1: Test cases for traffic reconstruction.

Case	θ (%)	T (sec)	t (sec)	# of Lag. observations
1	5	600	30	117
2	5	10	10	773
3	20	600	30	417
4	20	10	10	2577

Reconstruct traffic using measurements distributed in space and time. For the sake of this discussion, it is assumed that a communication network has been implemented to collect data from multiple cell phones and to deliver it to some central server where the algorithms will be run. The observations are dispersed in space and time as shown in Figure 3.3. The immediate goal is to reconstruct traffic flow based on these data. For this purpose, the expressway is discretized and the cell transmission model is employed to solve³ the conservation partial differential equation (PDE), the underlying model for this entire study.⁴

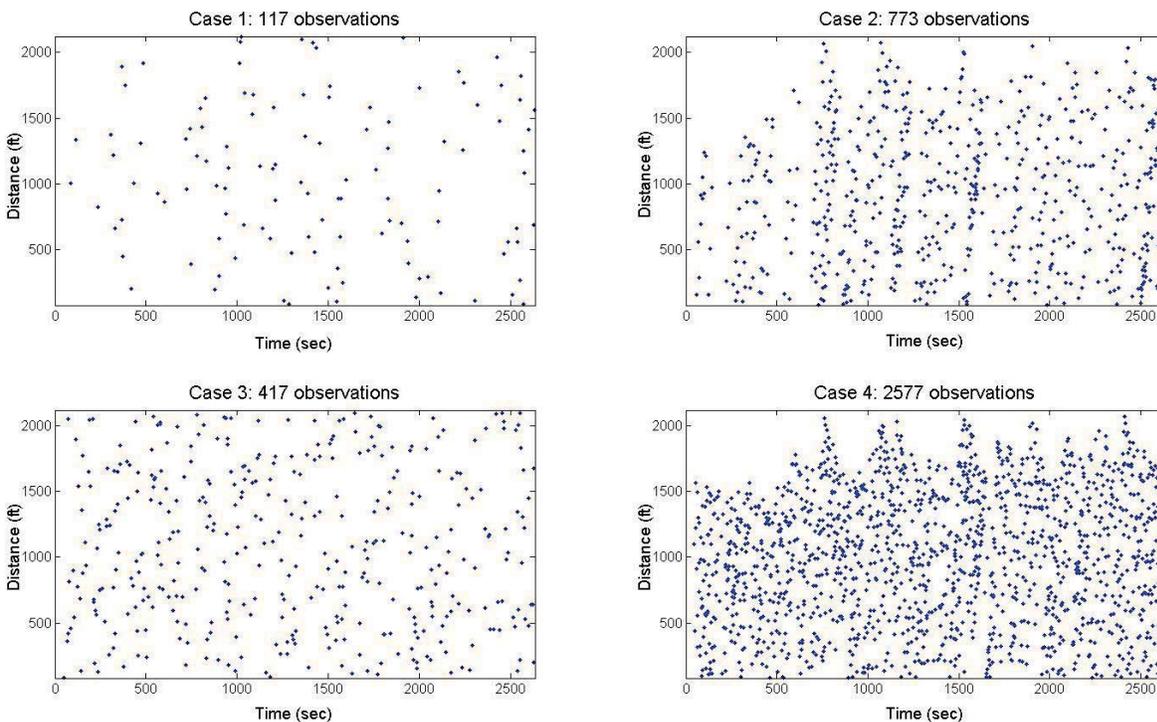


Figure 3.3: Spatio-temporal dispersion of probe measurements for different combinations of penetration and sampling rates.

Solving the PDE. Assuming only initial and boundary conditions, the PDE can be solved. This solution is referred to as the EDO solution, for Eulerian data only. Typically, the EDO solution has poor accuracy at spatio-temporal regions far from the initial and boundary conditions. As the PDE solver steps through time, probe data becomes available. This data needs to be

³ Methods are described in detail in Chapter 5.

⁴ Based on traffic flow physics, [79] and [97] independently proposed a first order partial differential equation (referred to as the LWR PDE) to describe traffic evolution over time and space.

incorporated into the solution so as to improve accuracy. Physically, this corresponds to adding or deleting vehicles in the cells so that the solution agrees with the additional data.

Incorporating Lagrangian data in the solution. Two algorithms were used to incorporate the Lagrangian data into the traffic reconstruction. The results were used to calculate measures comparable to the ground truth. Figure 3.4 shows how estimates of vehicular accumulations are improved as Lagrangian data is incorporated. Figure 3.5 shows that as more Lagrangian data is incorporated, the performance of the reconstruction algorithm improves. Figure 3.6 displays estimates of vehicular density at the middle of the modeled expressway section. Figure 3.7 displays estimates of travel time. Details of the mathematical tools employed to generate the estimates shown in these figures are presented in Chapter 5.

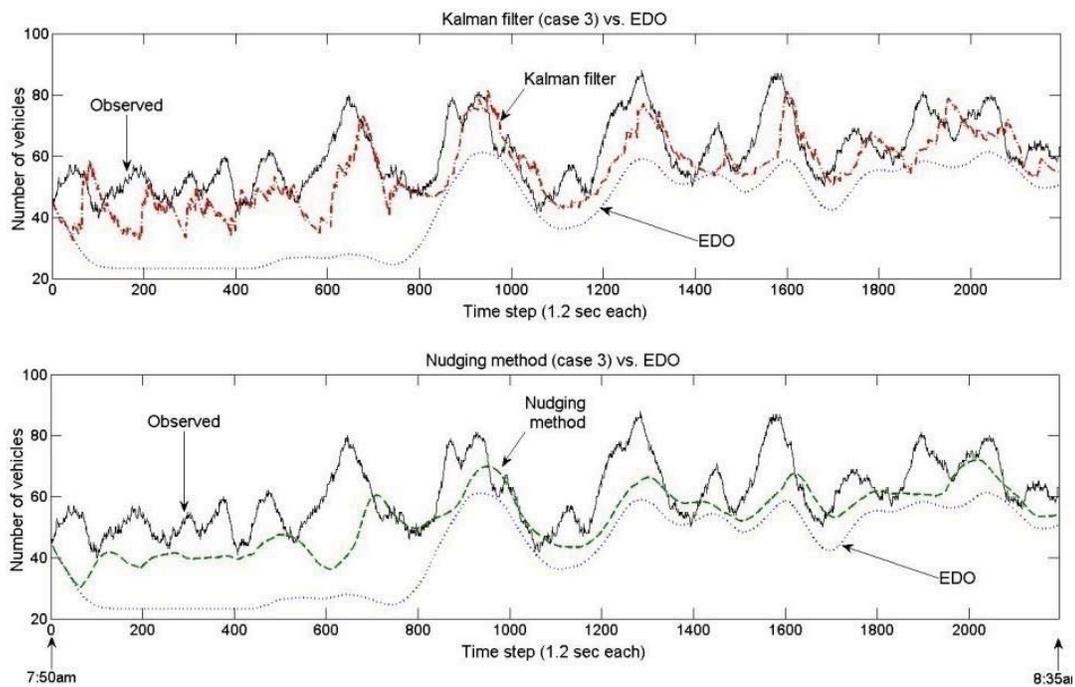


Figure 3.4: Vehicle accumulations. Comparison of estimates using the Kalman filter method (top), and Nudging method (bottom). In both cases, the incorporation of Lagrangian data results in improved estimates over those using Eulerian data only.

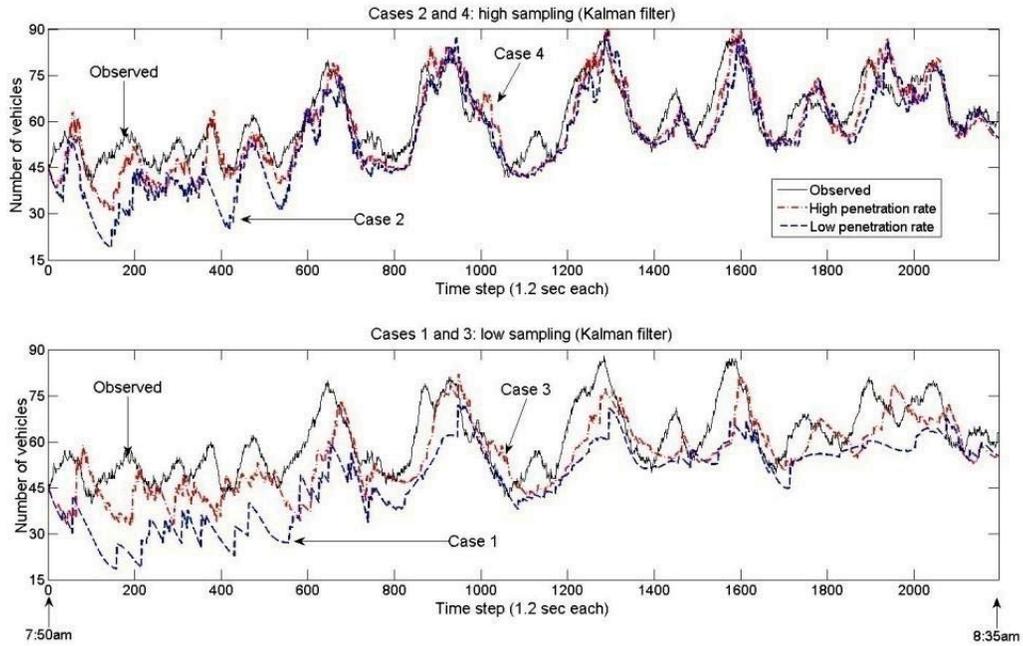


Figure 3.5: Vehicle accumulations. The higher the sampling rate, the higher the fidelity of the accumulation estimate.

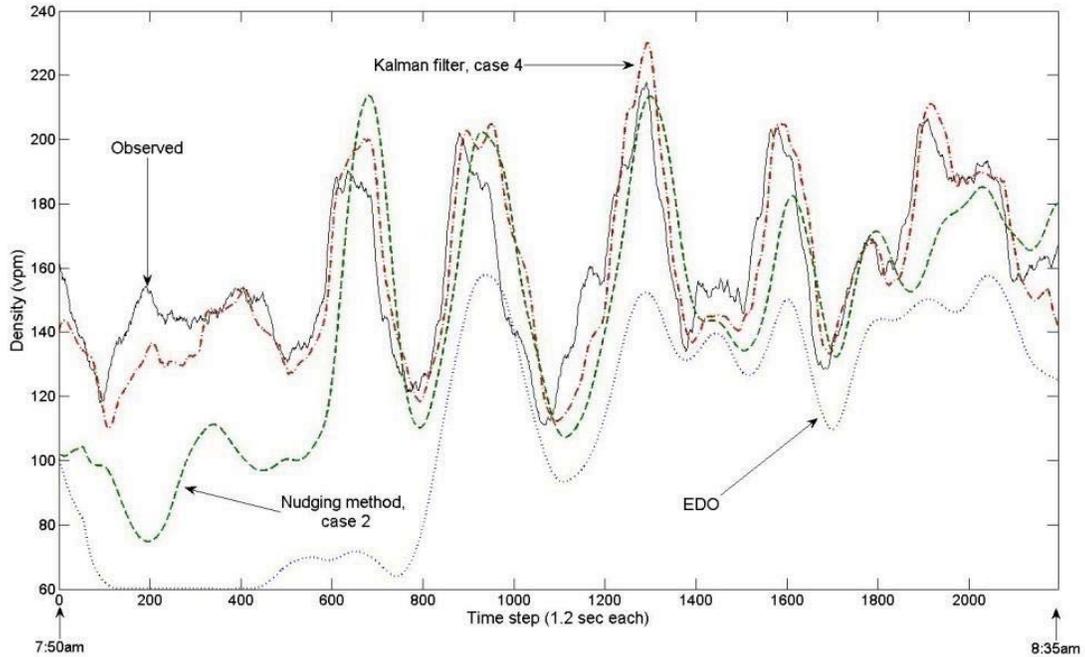


Figure 3.6: Vehicle density estimated at the middle of the modeled expressway section.

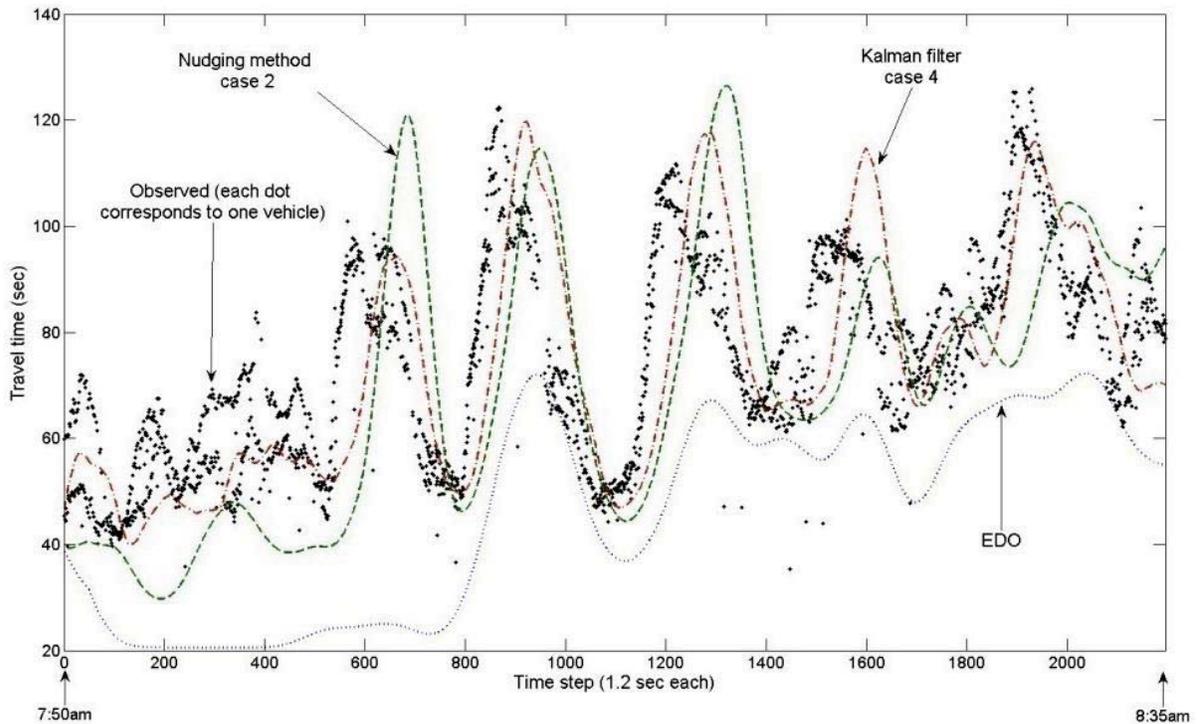


Figure 3.7: Travel time estimation for the modeled expressway section.

Discussion of results. The algorithms used here, and described later in the report, incorporate Lagrangian data in the modeling, and the accuracy of the estimates improve as more Lagrangian data are made available. Based on preliminary findings, Kalman Filtering slightly outperforms the nudging method. For this reason, Kalman Filtering was eventually chosen for the actual probe deployment. As indicated above, more details on these methods are furnished in Chapter 5, and Chapter 10.

Sampling rate issues. There is a clear design trade-off between penetration rate and sampling rate. The lower the penetration rate, the greater the necessary sampling rate for a given necessary quantity of data. Of course, a greater sampling rate can be more privacy intrusive. This issue leads to yet another trade-off between quality of data for traffic monitoring and privacy preservation; these issues are discussed in detail in Chapter 4.

3.2 Scale of Experiment

Scale of experiment. An experiment with ten vehicles would be too small, and an experiment with 1000 vehicles would not be feasible. To a first order, the scale of the experiment was a function of available funding. How many vehicles can be rented, and how many drivers can be employed within the funding constraints? For the clarity of exposition, we assume an experiment involving 100 vehicles, and investigate the consequences of this choice in terms of logistics and experiment design.

Formulation of constraints. Assuming that probe vehicles circulate in loops along an expressway, the cycle time, C , is given by:

$$C = L \left(\frac{1}{v_1} + \frac{1}{v_2} \right) + T_1 + T_2 \quad (3-1)$$

Where L is the length of the section (one way), v_i is the average speed for an equipped vehicle on direction i , and T_i is the lost time incurred when exiting the expressway and entering again in the opposite direction. The specification of a minimum penetration rate, θ , places a constraint on the number of equipped vehicles, N , the maximum flow in either direction, F , and the cycle time, C , as

$$\theta \leq \frac{N}{CF} \quad (3-2)$$

Intuitively, the constraint is met when N stays high, and F stays low. Constraints may also be formulated in terms of occupancy. For example,

$$L = \frac{NG}{2r\theta} \quad (3-3)$$

Where G is the average effective vehicle length, and r is the occupancy. Assuming $G = 20$ ft, $\theta = 0.2$ and $N = 100$, the maximum feasible expressway length can be calculated. In free-flow conditions near capacity, one might expect $r = 0.1$, and $L = 9.5$ miles. With congestion in both directions, one might expect $r = 0.3$, and $L = 3.2$ miles.

Consequences of 100 vehicles. A moderately scaled experiment with 100 participating probe vehicles has consequences. Chief among these is that a base station becomes necessary for staging purposes. The base camp requires significant space; i.e., a large parking lot that one might find at a shopping center. In addition, an experiment of this size requires non-trivial logistical support. The participation of probe vehicles, together with the penetration requirement, determines the feasible area that can be monitored.

Capture free flow and congestion conditions by design. To be a feasible paradigm for traffic monitoring, it is crucial to show that the methods and algorithms will work in both congestion and in free flow. For this purpose, a location must be chosen that can serve as a dependable source of recurrent congestion.

Variable spatial zone vs. variable number of probes. As shown above, the monitoring capability with a fixed number of vehicles depends greatly on the traffic conditions to be monitored. If the spatial zone of the experiment were constant, than one would need a variable number of probe vehicles as traffic conditions vary. Instead, it was proposed to define

the spatial zone of the experiment to correspond appropriately with the expected traffic conditions during the day.

Compare status quo monitoring with capabilities of new paradigm. Another requirement is to make a fair comparison with start-of-the-art traffic monitoring systems such as 511.org that make use of ILDs. For this purpose, an expressway segment with good detector coverage is necessary. In addition, both systems must be benchmarked with ground-truth travel time data. Therefore, it was necessary to instrument the endpoints of the expressway section with video cameras capable of re-identifying vehicles.

Effective measurement, without disrupting traffic, by design. Previous work with Caltrans on a PARAMICS model for I-880 was leveraged to refine the experimental protocol. With this model, the incremental effect of adding 100 extra cars onto I-880 was investigated. The question was whether the additional 100 vehicles would cause an undue influence on the data to be gathered. From simulation, it became clear that with 100 extra vehicles traveling in loops, that it was possible for on-ramps or off-ramps to become oversaturated. One outcome of the simulation was to break up the 100 vehicles into three separate groups with three separate routes. In this way, the probe vehicles would sense traffic, but not unduly affect the traffic being measured.

Privacy preservation. Having three routes in which vehicles may pass each other improves the data set for the purposes of investigating the privacy threats inherent in the data gathering mechanisms. This issue was of particular concern for one of the corporate partners, Nokia. With only one circuit, and with few opportunities for vehicles to pass each other, re-identification becomes trivial. Three circuits provide for overtaking and mixing, as would be the case with less contrived driving situations.

3.3 Practical Considerations

Legal requirements affecting experimental protocol. As the experiment began to take shape, the formal role of the drivers to be hired needed further clarification. In particular, it was necessary to understand how federal regulations and policies for the protection of human subjects might impact the experimental protocol.

Office for the Protection of Human Subjects. It is essentially for this purpose that the Office for the Protection of Human Subjects (OPHS) exists at the University of California, Berkeley. This office coordinates with the Committee for Protection of Human Subjects (CPHS) of which there are two groups: CPHS-1; and CPHS-2, who serve as Institutional Review Boards (IRBs). These IRBs ensure the protection of the rights and welfare of all human participants in research conducted by university faculty, staff and students.

Initial protocol to comply with federal and university requirements. In particular, a CPHS Narrative Form detailing the proposed protocol was submitted for approval. This document is included as Appendix 1 in its entirety. Although the final protocol for the February 2008

experiment differs from what was proposed in March of 2007, Appendix 1 is useful as a historical document that captures a snapshot of the evolving experiment that came to be known as Mobile Century.

Practical considerations. The driver schedule was defined based on the most restrictive truck driver regulations worldwide, stipulating at least 45 minutes of rest for every 4 hours of driving. In addition, to limit the total number of hours worked, it was deemed infeasible to begin the experiment early enough in the morning to capture both the morning rush and the evening rush. Instead, only the evening rush was intended to be captured.

Institutional collaboration. The collaborative model for the present work was that industry provides implementation expertise. The NRC team was tasked to build the back-end infrastructure, the cell phone client, the real-time sampling scheme, and the servers to get data off the phones, and software to visualize the results. The role for academia was to provide new processing techniques, scientific expertise, and algorithm development. The UCB team was to perform the traffic experiment design, propose a suitable site, and build the algorithms to estimate traffic conditions. The UCB team consisted of about twenty to thirty people during the evolution of the experiment design. Meetings were regularly held between members of both teams to coordinate the details of data-handling. In particular, data from the cell phones needed to be supplied in a useable form for the UCB team's algorithms and then passed back to the NRC team's software for visualization.

Choosing a date. When choosing a date for the experiment, the key concern was to find a part of the semester when graduate students would be available to participate. Weather was also a concern, but avoiding finals week and vacation weekends was crucial. Taking these factors into account, the second week of February was chosen as the target date.

3.4 *Alternate Site*

A site in Danville, pictured in Figure 3.8, was initially envisioned as a possible venue for the 100-car deployment due to its proximity to Berkeley and ease of access. However, detailed inspection of the traffic properties along this site revealed that the presence of recurrent congestion was not reliable. In addition, this site suffered from a lack of reliable PeMS coverage available at that time.

The I-880 site eventually chosen for the 100-car deployment had excellent PeMS coverage, which proved to be crucial for analysis purposes.

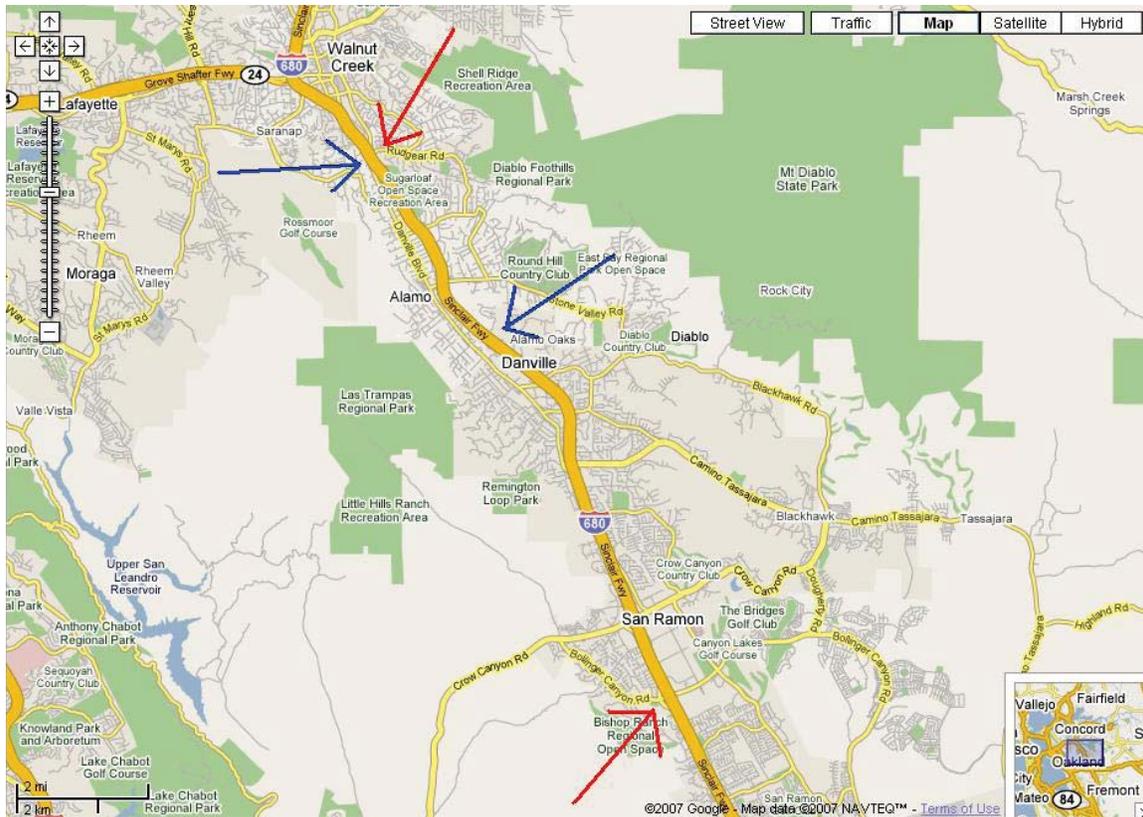


Figure 3.8: Alternate I-680 site.

4 VTL Traffic Monitoring

This chapter begins with an overview of privacy risks in Section 4.1. To address these risks, the VTL concept is introduced in Section 4.2 as a spatial sampling scheme. In Section 4.3, a traffic monitoring architecture is designed and implemented according to the VTL scheme.

Next, two prototyping efforts are described. The data collected in these small-scale deployments are examined briefly in Section 4.1. The privacy risk and quality of travel time estimation is a function of the number and placement of VTLs. This trade-off is explored in Section 4.2. Privacy is addressed by defining exclusion zones and a minimum spacing. Travel time quality is addressed by optimal placement of VTLs. The VTL scheme is compared against a temporal sampling strategy. In Section 4.3, the tradeoff between privacy and data quality is further explored. Key results are summarized in Section 4.4.

4.1 Privacy Risks and Threat Model

Privacy concerns for participating drivers. Traffic monitoring through GPS-equipped cell phones raises significant privacy concerns for participating users. Social acceptance of such monitoring is less likely if location traces are detailed enough to infer medical conditions, political affiliations, speeding, or involvement in traffic accidents.

Threat Model and Assumptions. The present work assumes that adversaries can compromise any single infrastructure component to extract information and can eavesdrop on network communications. We assume that different infrastructure parties do not collude and that a driver's own handset is trustworthy. We believe this model is useful in light of the many data breaches that occur due to dishonest insiders, hacked servers, stolen computers, or lost storage media (see [4] for an extensive list, including a dishonest insider case that released 4500 records from California's FasTrak automated road toll collection system). These cases usually involve the compromise of log files or databases in a single system component and motivate our approach of ensuring that no single infrastructure component can accumulate sensitive information.

Naive anonymization is insufficient to protect privacy. We consider sensitive information to be any information from which the precise location of an individual at a given time can be inferred. Traffic monitoring requires at least aggregated statistics from a large number of probe vehicles, but does not require individual node identities. Therefore, one obvious privacy measure would be to anonymize the location data by removing identifiers such as network addresses. This approach is insufficient, however, because drivers can often be re-identified by correlating anonymous location traces with identified data from other sources. For example, home locations can be identified from anonymous GPS traces [54, 73], which may be correlated with address databases to infer the likely driver. Similarly, records on work locations or automatic toll booth records could help identify drivers. Even if anonymous point location samples from several drivers are mixed, it is possible to reconstruct individual traces because successive

samples from the same vehicle inherently share a high spatio-temporal correlation. If overall vehicle density is low, samples that are close in time and space likely originate from the same vehicle. This approach is formalized in target tracking models [96].

Example formulation of a tracking model. As an example of tracking anonymous samples, consider the following problem: given a time series of anonymous location and speed samples mixed from multiple users, extract a subset of samples generated by the same vehicle. Toward this end, an adversary can predict the location of the next sample $x_{t+\Delta t} = v_t \cdot \Delta t + x_t$ based on the reported speed of the previous sample, where x_t and $x_{t+\Delta t}$ are locations at time t and $t + \Delta t$, respectively, and v_t is the reported speed at t . The adversary then associates the prior location sample with the next sample closest to the prediction, or more formally with the most likely sample, where likelihood can be described through a conditional probability $P(x_{t+\Delta t} | x_t)$ that primarily depends on spatial and temporal proximity to the prediction. The probability can be modeled through a probability density function (pdf) of distance (or time) differences between the predicted sample and an actual sample (under the assumption that the distance difference is independent of the given location sample).

Speed patterns correlate with route choice, and provide clues to an adversary. Knowing speed patterns further helps tracking anonymous location samples if it is combined with map information. For example, consider the traffic scenarios depicted in Figure 4.1. On straight sections (a) vehicles on high-occupancy vehicle (HOV) or overtaking lanes often experience lower variance in speed. Vehicles entering at an on-ramp (b) or exiting after an off-ramp (c) usually drive slower than main road traffic. These general observations can be formally introduced into the tracking model by assigning an a priori probability derived from the speed deviations. For example, to identify the next location sample after an on-ramp for a vehicle that generated x_t on the main route before the ramp, an adversary could assign a lower probability to location samples with low speed. These low speed samples are likely generated by vehicles that just entered after the ramp.

Privacy Metrics. As observed in [55], the degree of privacy risk depends on how long an adversary successfully tracks a vehicle. Longer tracking increases the likelihood that an adversary can identify a vehicle and observe it visiting sensitive places. We thus adopt the time-to-confusion [55] metric and its variant distance-to-confusion, which measures the time or distance over which tracking may be possible. Distance-to-confusion is defined as the travel distance until tracking uncertainty rises above a defined threshold. Tracking uncertainty is calculated separately for each location sample in a trace as the entropy $H = -\sum p_i \log p_i$, where the p_i are the normalized probabilities derived from the likelihood values described later. These likelihood values are calculated for every location sample generated within a temporal and spatial window after the location sample under consideration.

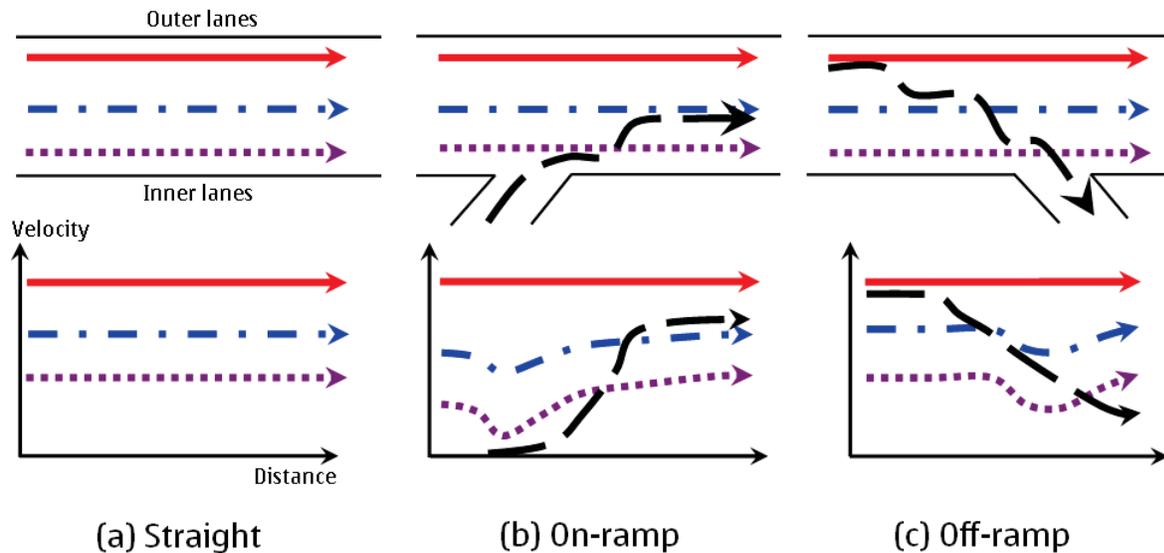


Figure 4.1: Driving Patterns and Speed Variations in Highway Traffic.

4.2 Preserving Privacy with Virtual Trip Lines

We introduce the concept of virtual trip lines (VTLs) for privacy-preserving monitoring and describe an architecture that embodies it.

4.2.1 Design Goals

Tradeoff between quality information and privacy protection. The big-picture challenge is to balance the tradeoff between two conflicting requirements. On one hand, quality traffic information needs to be acquired from each smartphone client. On the other hand, all gathered information must be limited and structured in such a way that it is unwieldy, or difficult, to exploit for unintended purposes. In particular, we address issues of privacy invasion that have the potential to hinder widespread social acceptance of such a system.

Privacy. We aim to achieve privacy protection by design so that the compromise of a single entity, even by an insider at the service provider, does not allow the identifying or tracking of users.

Data Integrity. The system should not allow adversaries to insert spoofed data, which would reduce the data quality of traffic information. This is especially challenging because it conflicts with the desire for anonymity.

Smartphone Client. The client software must cope with the resource constraints of current smartphone platforms. For energy consumption, we mainly focus on designing a light-weight component that filters noisy GPS samples and computes trip-line measurements.

4.2.2 Virtual Trip Line Concept

Definition of VTL. The proposed traffic monitoring system builds on the concept of virtual trip lines and the notion of separating the communication and traffic monitoring responsibilities (as introduced in [54]). A VTL is a line in geographic space that, when crossed, triggers a client to send a VTL sample to the traffic monitoring server. More specifically, it is defined by:

$$[volid, x_1, y_1, x_2, y_2, d] \quad (4-1)$$

where $volid$ is the virtual trip line ID, $x_1, y_1, x_2,$ and y_2 are the (x, y) coordinates of two line endpoints, and d is a default direction vector (e.g., N-S or E-W). When a vehicle traverses the trip line its VTL sample comprises time, trip line ID, speed, and the direction of crossing. The trip lines are pre-generated and downloaded and stored in clients.

Spatial sampling preferred over temporal sampling. Virtual trip lines control disclosure of location data by sampling in space rather than sampling in time, since clients generate VTL samples at predefined geographic locations, compared to sending samples at periodic time intervals. The rationale for this approach is that in certain locations traffic information is more valuable and certain locations are more privacy-sensitive than others. Through careful placement of trip lines the system can thus better manage data quality and privacy than through a uniform temporal sampling interval. In addition, the ability to store trip lines on the clients can reduce the dependency on trustworthy infrastructure for coordination. These concepts are revisited in Section 4.2.

4.2.3 Architecture for Probabilistic Privacy

Strict separation of identity information (for communication) and location information (for traffic monitoring). To achieve the anonymization of VTL samples from clients while authenticating the sender of VTL samples, we split the actions of authentication and data processing onto two different entities, an ID proxy server and a traffic monitoring server. By separately encrypting the identification information and the sensing measurements (i.e., trip line ID, speed, and direction) with different keys, we prevent each entity from observing both the identification and the sensing measurements.

Overview of system architecture. Figure 4.2 shows the resulting system architecture eventually implemented for the field experiment. It comprises four key entities: probe vehicles with the cell phone handsets, an ID proxy server, a traffic monitoring service provider, and a VTL generator. Each probe vehicle carries a GPS-enabled mobile handset that executes the client application. This application is responsible for the following functions: downloading and caching

trip lines from the VTL server, detecting trip line traversal, and sending measurements to the service provider. To determine trip line traversals, probe vehicles check if the line between the current GPS position and the previous GPS position intersects with any of the trip lines in its cache. Upon traversal, handsets create a VTL sample comprising trip line ID, speed readings, timestamps, and the direction of traversal and encrypt it with the VTL server's public key. Handsets then transmit this sample to the ID proxy server over an encrypted and authenticated communication link set up for each handset separately. Each handset and the ID proxy share an authentication key in advance.

Figure 4.2: Virtual Trip Line: Privacy-Preserving Traffic monitoring System Architecture. This system was implemented and ran for the entire duration of the 100-vehicle deployment of the Mobile Century experiment

ID proxy server handles identity information. The ID proxy's responsibility is to first authenticate each client to prevent unauthorized VTL samples and then forward anonymized samples to the VTL server. Since the VTL sample is encrypted with the VTL server's key, the ID proxy server cannot access the VTL sample content. It has knowledge of which phone transmitted a VTL sample, but no knowledge of the phones position. The ID proxy server strips off the identifying information and forwards the anonymous VTL sample to the VTL server over another secure communication link.

VTL server handles location information. The VTL server aggregates samples from a large number of probe vehicles and uses them for estimating the real-time traffic status. The VTL generator determines the position of trip lines, stores them in a database, and distributes trip lines to probe vehicles when any download request from probe vehicles is received. Similar to the ID proxy, each handset and the VTL generator should share an authentication key in advance. The VTL generator first authenticates each download requester to prevent unauthorized requests and can encrypts trip lines with a key agreed upon between the requester and the VTL generator. Both the download request message and the response message are integrity protected by a message authentication code.

Advantages of this architecture. The above architecture improves location privacy of probe vehicle drivers through several mechanisms. First, the VTL server must follow specific restrictions on trip line placements that we will describe in Section 4.2. This means that a handset will only generate samples in areas that are deemed less sensitive and not send any information in other areas. By splitting identity-related and location-related processing, a breach at any single entity would not reveal the precise position of an identified individual. A breach at the ID proxy would only reveal which phones are generating samples (or are moving) but not their precise positions. Similarly, a breach at the VTL server would provide precise position samples but not the individual's identities. Separating the VTL server from the VTL generator prevents active attacks that modify trip line placement to obtain more sensitive data. This is, however, only a probabilistic guarantee because tracking and eventual identification of outlier trips may still be possible. For example, tracking would be straightforward for a single probe vehicle driving along on empty roadway at night [55]. The outlier problem in sparse traffic situations can be alleviated by changing trip lines based on traffic density heuristics. Trip lines could be locally deactivated by the client based on time of day or the clients speed. They could also be deactivated by the VTL generator based on traffic observations from other sources such as loop detectors.

4.3 Implementation

The architecture described above was implemented using Nokia N95 smartphone handsets, which include a full Global Positioning System receiver that can be accessed by application software.

4.3.1 Map Tiles and Trip Lines

Quadrant representation. In our system, we recursively divide the geographic region of interest into four smaller rectangles (or quadrants), and the minimum quadrant size is 1m by 1m. We convert the GPS location of a user into a Mercator projection using the WSG84 world model. Mercator projects the world into a square planar surface. A zoom of 25 is assumed to be the maximum precision that location can be specified in. By default every GPS location is converted into 25 bit x and y values with zoom set to 25. By using the quadrant representation the mobile device can efficiently control the granularity by simply changing the zoom level. In this

encoding, the world is treated as a square grid of four quadrants with zoom level 2, where x and y are the offsets from the top left corner of the world.

VTLs contained in map tiles. This representation makes it easy to specify the specific map tile. We define a map tile as a container that groups all trip lines within it. When a client wants to download all virtual trip lines within the San Francisco Bay Area, it sends the VTL server the triplet, $(zoom, x, y)$ for the corresponding region. In our implementation, we choose 12 as the default zoom level, which corresponds to an 8 km by 8 km square.

Memory requirements. This representation also helps in reducing storage size and bandwidth consumption. Since the general area is identified by the quadrant, we only store the 13 least significant bits of the trip line end point coordinates instead of the full 25 bits used for typical UTM coordinates. This decreases storage consumption to 68bits (15 bit id, 1 bit direction, $4 \cdot 13$ bits coordinates) per trip line. As an example of required storage and bandwidth consumption, consider the San Francisco Bay Area, the total road network of which contains about 20,000 road segments, according to the Digital Line Graph 1:24K scale maps of the San Francisco Bay Area Regional Database (BARD [1], managed by USGS). Assuming that the system on average places one trip line per segment this results in 166KB of storage.

4.3.2 Client Device and Software

Client hardware and software. We implemented the client software using J2ME (Java Platform, Micro Edition) on an Nokia N95 handset. This Symbian OS handset uses an ARM11-based Texas Instruments OMAP2420 processor running at 330MHz, and it contains 64MB RAM and 160MB internal memory. Its storage can be expanded up to 8GB with flash memory. We use the JSR 179 library (Location API for J2ME) [2] for communicating with the internal TI GPS5300 Navilink 4.0 single-chip GPS/A-GPS module to set the sampling period and retrieve the position readings. This setup did not provide speed information. Instead, we calculate the mean speed using two successive location readings (in our implementation, every 3 seconds). The client software registers the task for checking the traversal of trip lines as an event handler for GPS module location samples, which is automatically invoked whenever a new position reading becomes available.

Communication protocol. The communication between the handset and the ID proxy server, to send updated lists of VTLs or to request VTL downloads, is implemented via HTTPS GET/POST messages. The client software encrypts the message content but not the handset identification information using the public key of the VTL server so that only the VTL server with the corresponding private key can decrypt the message. To save network bandwidth and to reduce delay, we cache the downloaded trip lines for the nine map tiles closest to the current position in local memory. When a vehicle crosses a tile boundary, it initiates VTL download background threads for the missing tiles.

4.3.3 Servers and Databases

VTL database server. At the bottom of the hierarchy of our server implementation is a backend database server. The database server contains two databases. First is a VTL database which holds GPS coordinates of all trip lines. In future we plan to enhance our trip line database to hold meta data associated with that trip line. For instance, the meta data for a trip line can contain the posted speed limit at that trip line which can be used by the client application to decide if it is going over the speed limit in which case the client application can disable the transmission of VTL samples. Write access to this database is restricted only to traffic administrators who can add, delete or update a VTL.



Figure 4.3: Road networks extracted from Bay Area DLG files (Left) and Trip Lines per road segment in Palo Alto CA (Right).

Traffic database server. The second database is the VTL sample measurement database. This database stores the VTL samples sent by the mobile device whenever the mobile device chooses to send a sample after crossing a VTL. The sample database simply appends every VTL sample along with a time stamp on when the sample was received. To sanitize bogus VTL samples from the clients, the VTL sample database also keeps both the encrypted and decrypted versions of the VTL sample for further investigation in collaboration with the ID proxy server. When bogus VTL samples are detected in the VTL sample database, their encrypted versions are compared to the encrypted version stored in the ID proxy server to blacklist the originator of bogus VTL samples.

Database implementation. We use Microsoft SQL to implement the databases, and we develop the VTL server using J2EE (Java Platform, Enterprise Edition) and JDBC (Java Database Connectivity) to control the SQL databases that are connected to the VTL server. While we have used only a single DB server in this prototype, the two databases should ideally be implemented by different entities to prevent active trip line modification attacks by a compromised traffic monitoring entity.

ID Proxy Server. On top of the database server is the ID Proxy server. The identification proxy server is envisioned to be operated by an entity that is independent of the traffic service provider. We implement the ID proxy server as a servlet-based web server that takes in HTTPS GET/POST messages from clients and forwards messages to the VTL server. The HTTP message received by the proxy server from the client has two components. The first component contains the mobile device identification information, namely phone number of the message origin. This component of the message is required for all cell phone communications as operator needs to appropriately charge for data communication costs. The second component of the message contains information that is intended for the database server. The proxy server strips all the identification information from the message, namely the first component of the message, and passes on the second component of the message to the application server. We implemented the secure channel between ID proxy server and the VTL server using WSDL (Web Service Definition Language)-RPC (Remote Procedure Call) over J2EE Server.

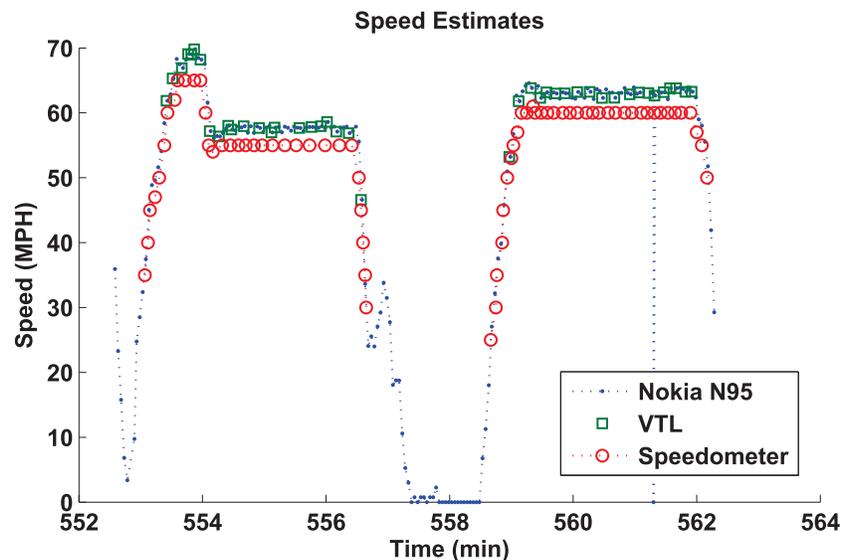


Figure 4.4: Comparison of the speed measurements recorded from the N95 (dots), the VTLs (boxes) and the vehicle speedometer (circles) as a function of time.

4.1 Experimental Deployment

The implementation described above was used for several experimental deployments. The correct operation of the traffic monitoring system was first demonstrated with an initial test along I-80. A second test involving twenty cars was performed to measure data quality and to inform the design of the 100 vehicle deployment.



Figure 4.5: Satellite image of the first experiment site I-80 near Berkeley, CA. The red lines represent the locations of the VTLs, the blue squares show the speed recorded by the VTL, and the green squares represent the position and speed stored in the phone log. The brown circles represent the readings from the vehicle speedometer.

4.1.1 Velocity Measurement Accuracy

GPS speed and position accuracy. A first experiment was performed to estimate the position and speed accuracy of a single cell phone carried onboard a vehicle. The experiment route consisted of a single 7-mile loop on I-80 near Berkeley, CA. VTLs were placed evenly on the highway every 0.2 miles. Speed and position measurements were stored locally on the phone every 3 seconds, and speed measurements were sent over the wireless access provider's data network every time a VTL was crossed. The speed measurements were computed using two consecutive position measurements. In order to substantiate the correctness of the data,

vehicle speed was also recorded directly from the speedometer on a laptop with a clock synchronized with the N95. In Figure 4.4, the speed measured directly from the vehicle speedometer is compared to the speeds measured by the VTLs and the speed stored in the phone log. Timestamp of each record denotes the elapsed time since midnight of the experiment day. On average, the vehicle odometer reported a speed 3 mph slower than the GPS. The position data was accurate enough to correctly place the vehicle on either the correct or neighboring lane of travel.

Disambiguation of closely-spaced parallel roads. To further evaluate the position accuracy of GPS enabled cell phones, the vehicle was driven on a frontage road along the highway. The purpose was to determine whether erroneous VTL samples would be generated. Frontage roads typically have slow moving traffic with speed limits of 25 or 35 mph and run alongside the freeway. Without high precision position accuracy, this traffic can be incorrectly identified as freeway traffic. VTLs were only placed on the freeway, not the frontage road. Although the frontage road is separated by as little as 30 ft from the freeway, no erroneous VTL samples were generated. This test was conducted using only a single phone, but it presented promise that the technology can be used for advanced traffic monitoring applications.

4.1.2 Twenty-vehicle Experiment

Experimental Setup.⁵ For two hours, twenty vehicles were driven back and forth on a 4-mile section of I-880 south of Oakland CA as shown in Figure 4.6. The length of the test road segment was chosen to have 1% to 2% penetration rate given twenty participants and approximate round travel time. In order to observe a more natural mixing phenomena (in which vehicles pass each other) half of the drivers were instructed to drive a slightly shorter, 3 mile section of the highway (red circle) after the completion of the first lap. The location of this experiment was specifically selected because it featured both free flowing traffic at greater than 50 mph, and congested, stop and go traffic. An accident just north of the experiment site further added to the complexity of the northbound traffic flow. As observed by the drivers of the experiment, this accident created “shear” in the traffic flow, where vehicles in adjacent lanes of traffic were traveling at significantly different speeds.

⁵ For a detailed description of the experimental setup for this prototyping effort, refer to the collection of materials in Appendix 2. The overall strategy and objectives of the twenty-vehicle deployment are described in appendix 2.1. The actual protocol is provided in appendix 2.2. Participants were divided into two groups (even, and odd). Detailed instructions for each group are given in appendices 2.3 and 2.4.

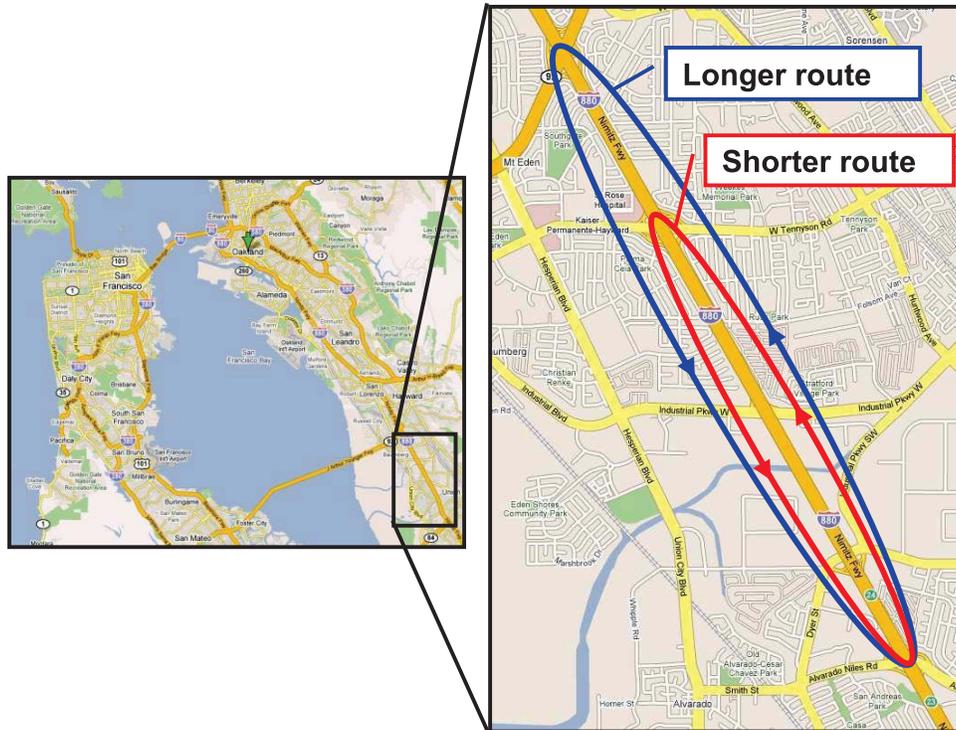


Figure 4.6: I880 Highway Segment for Twenty Car Experiment.

Cellphone usage protocol. During the twenty-vehicle experiments, cell phones were placed on the dashboard as shown on the right side of Figure 4.7. In future deployments, however, actual users may place their phones into a dashboard car holder (depicted in the left side of Figure 4.7) to be able to view navigation and travel time information while driving.



Figure 4.7: Experimental Setup in a Car for Twenty Car Experiment.

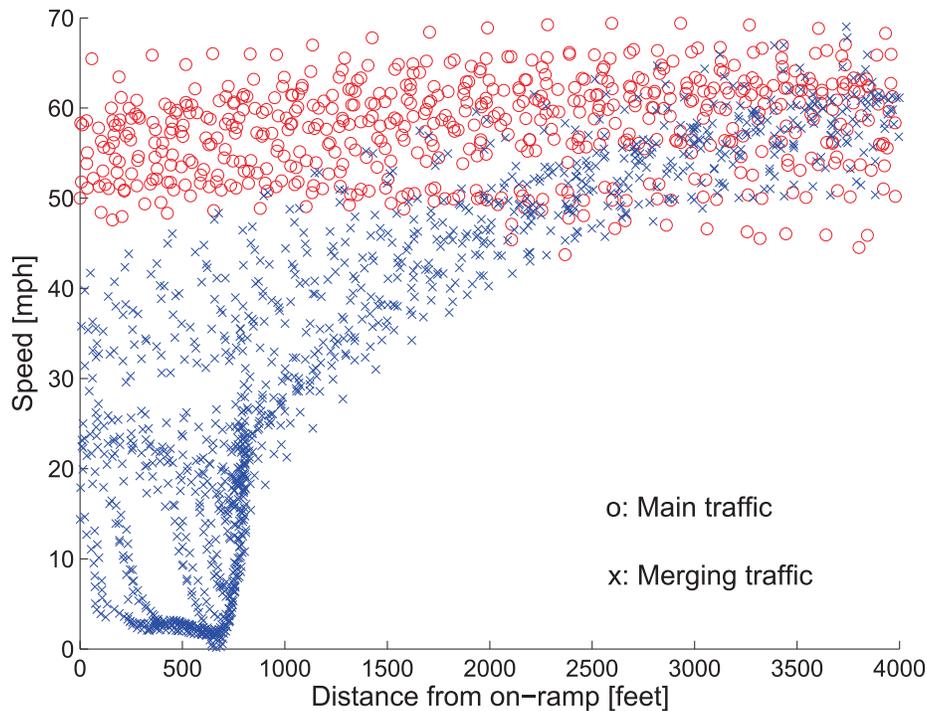


Figure 4.8: Speed Measurements over Distance.

Empirical speed patterns at an on-ramp. Figure 4.8 shows the difference in speed between the merging traffic (red circles) and the main traffic (blue crosses) along the on-ramp. As expected, main-line and merging traffic can be easily distinguished by a simple speed threshold near the on-ramp. The graph also shows that the distribution of speeds from the main line and onramp overlap at about 3000 ft from the on-ramp. In the next section, we use this empirical result to define an exclusion zone.

Empirical speed patterns at a merge location over time. Figure 4.9 depicts a weekday time interval between 2:15 PM to 4:14 PM, during which the mean speed of the test road segment decreases due to the increasing congestion, as reported by the PeMS highway measurement database [91]. As evident, the speed variation between main route traffic and merging traffic increases with congestion. This may require longer exclusion during congestion.

4.2 Trip Line Placement

This section describes several VTL placement procedures and two main types of constraints that must be determined. The first constraint is a minimum spacing between VTLs. This constraint reduces tracking ability along straight roadways. The second constraint is the extent of exclusion areas to be placed around intersections and merges.

The basic VTL placement procedure is spatial periodic sampling, in which VTLs are evenly-spaced. The second procedure is to define exclusion zones, and then to distribute equidistant trip lines orthogonally along the remaining road segments.

A third procedure is called optimal placement, in which VTLs are placed on the road network where the velocity field tends to change (e.g., at recurrent bottlenecks).

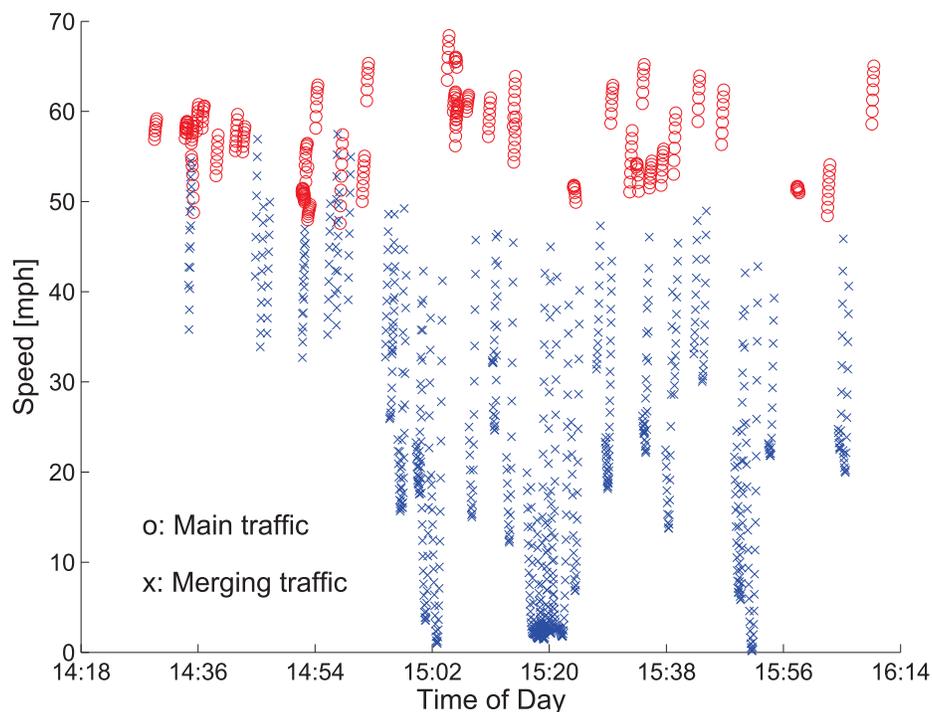


Figure 4.9: Speed Measurements over Time.

4.2.1 Minimum Spacing Constraint

Determining Minimum Spacing. The minimum spacing constraint is particularly important on highways, where more regular traffic flows increase the tracking risks. Thus, we focus our derivations on straight highway scenarios. Minimum spacing for longer road segments is determined based on a tracking uncertainty threshold. Recall that to prevent linking compromises, an adversary should not be able to determine with high confidence that two anonymous VTL samples were generated by the same handset.

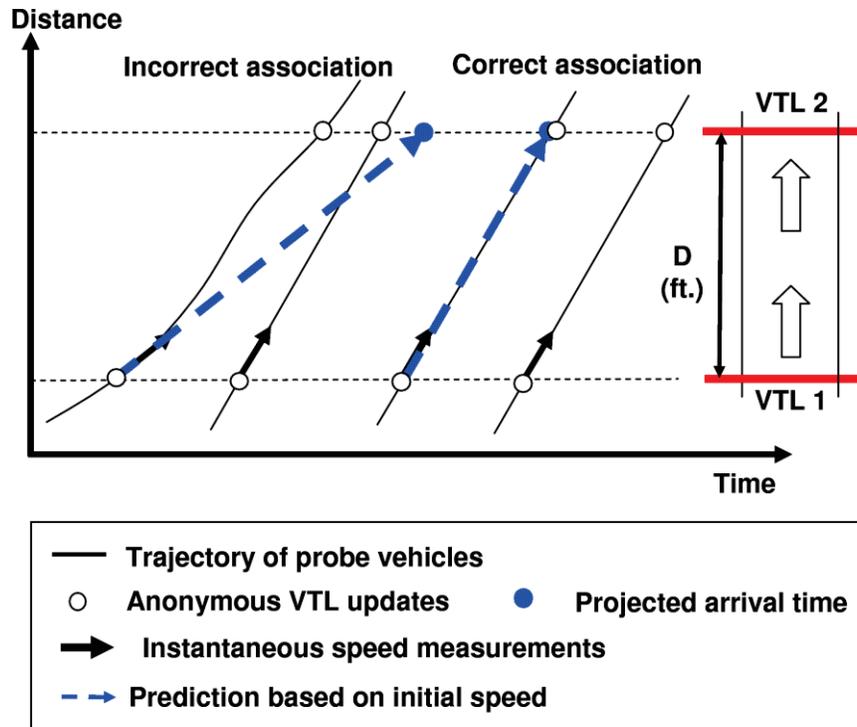


Figure 4.10: Linking prediction on a straight highway section.

Define tracking uncertainty and pdf for travel time variability. Tracking uncertainty defines the level of confusion that an adversary encounters when associating two successive anonymous VTL samples to each other. We define tracking uncertainty as the entropy $H = -\sum_i p_i \log p_i$,

where p_i denotes the probability (from the adversary's perspective) that anonymous VTL sample i at the next trip line was generated by the same phone as a given anonymous VTL sample at a previous trip line. The probability p_i is calculated based on an empirically derived pdf model that takes into account the time difference between the predicted arrival time at the next trip line and the actual timestamp of VTL sample i . We fit an empirical pdf of time deviation with an exponential function, $p_i = \frac{1}{\alpha} e^{-\frac{t}{\beta}}$, where we obtain the values of α and β by using unconstrained nonlinear minimization.

Example linking based on prediction. Consider the example scenario in Figure 4.10. In the scenario on the left, the adversary projects the arrival time at VTL 2 based on the phone's speed report at VTL 1. The projected arrival time is the endpoint of the dashed line (the solid lines indicate phones' actual paths). There are two actual flow samples at VTL 2 (indicated with circles). The adversary would calculate the time difference between the projected arrival times, assign probabilities p_1 and p_2 using the pdf, and determine entropy. Compared to the second

example to the right, entropy is high indicating that an adversary cannot determine the correct VTL sample with high confidence. In fact, the closest sample would be incorrect in this scenario. Tracking uncertainty calculated through the entropy is maximized when there is more than one anonymous VTL sample with similar time-differences. In other words, lower values of H indicate more certainty or lower privacy. To reduce computational complexity, we only consider VTL samples within a time window w of the projected arrival time during the entropy calculation.

Physical intuition for VTL placement and resulting uncertainty. In general, tracking uncertainty is dependent on the spacing between VTLs, the penetration rate, and speed variations of vehicles. If speed remains constant, as in the second example of Figure 4.10, the projected arrival times match well and tracking uncertainty is low. Higher penetration rates lead to more VTL samples around the projected arrival time, which decreases certainty. As spacing increases, the likelihood that speeds and the order of vehicles remain unchanged decreases, leading to more uncertainty. Speed variations on highways are frequently caused by congestion—thus road segments with lower average speed tend to increase tracking uncertainty.

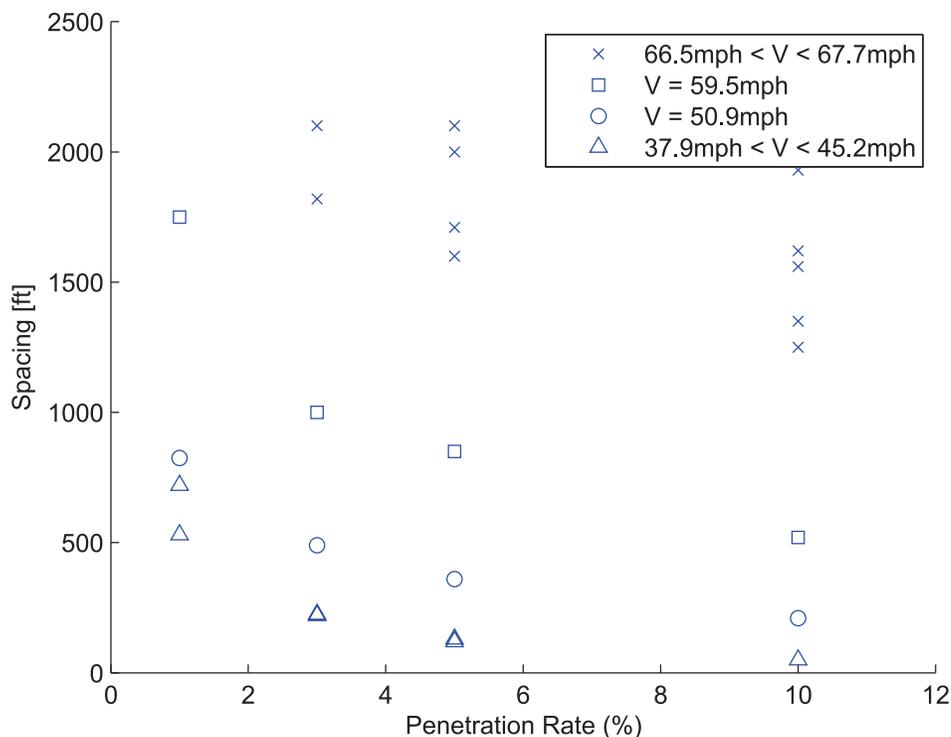


Figure 4.11: Minimum Spacing Constraints for Straight Highway Section.

Simulation results for minimum spacing as a function of penetration rate and prevailing speeds in congestion. We empirically substantiate these observations through simulations using the PARAMICS vehicle traffic simulator [3]. Figure 4.11 depicts the minimum spacing required to

achieve a minimum mean tracking entropy of $H = 0.2$ for different penetration rates and different levels of congestion (or mean speed of traffic). We choose a reasonably low uncertainty threshold, which ensures to an adversary a longer tracking that could have privacy events such as two different places (e.g., origin and destination).⁶ The uncertainty value of 0.2 corresponds to an obvious tracking case in which the most likely hypothesis has a likelihood of 0.97. The penetration rates used were 1%, 3%, 5% and 10%. To evaluate different levels of congestion, we used traces from seven 15-minute time periods distributed over one day. We also used three different highway sections (between the junction of CA92 and the junction of Tennyson Rd., between the junction of Tennyson Rd. and the junction of Industrial Rd., and between the junction of Industrial Rd. and the junction of Alvarado-Niles Rd.) to reduce location-dependent effects. The simulations show that the needed minimum spacing decreases with slower average speed and higher penetration rate. The clear dependency of the tracking uncertainty on the penetration rate and the average speed allows creating a model that provides the required minimum spacing for a given penetration rate and the average speed of the target road segment.

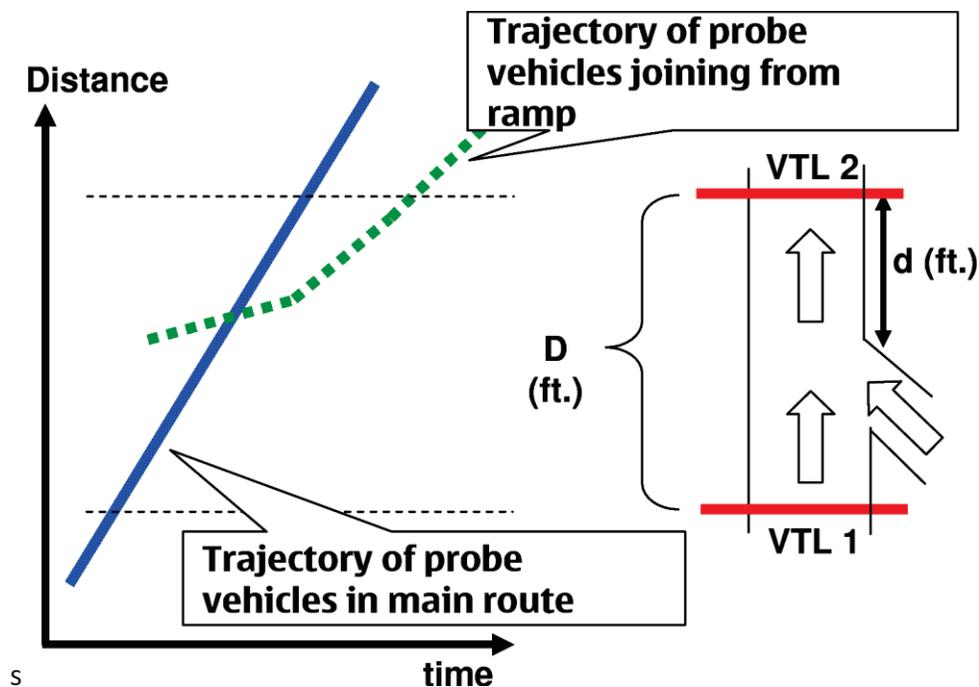


Figure 4.12: Linking attack near an on-ramp.

⁶Two recent studies, [54] and [73], observe about 15 minutes as a median trip time.

4.2.2 Determining Exclusion Areas

Determining Exclusion Areas. Additional tracking risks are present at ramps and many intersections because of large speed variations. Vehicles leaving the main direction of travel slow down, vehicles joining the main direction of travel from ramps accelerate from a very low speed, and vehicles staying on the main direction maintain their (relatively high) speed. Figure 4.12 illustrates an example. If trip lines are placed immediately before or after intersections, an adversary may be able to follow vehicle paths based on speed patterns.

Exclusion zone near a ramp is about 3000 ft. Empirical measurements of these speed patterns were presented in Figure 4.8. From this, we suggest an exclusion zone of around 3000 ft from an on-ramp.

Candidate areas for exclusion. In addition to merges and intersections, where detailed information would be especially important for an adversary to track which path a vehicle takes, exclusion zones can also be allocated around other sensitive places. These may be places that could allow sensitive inferences, such as a medical clinics, or locations where the owner of a handset may be identified, such as suburban home locations [73] or automatic toll booth plazas.

4.2.3 Optimal Placement for Traffic Flow Estimation Accuracy

Optimal Placement of VTLs. If the number of VTLs is fixed, then it is best to allocate them in space where the information content is greatest. Specifically, one would like to place VTLs where changes in the velocity field tend to originate such as recurrent bottlenecks.

Set-up for Optimal Placement Algorithm. An optimal placement algorithm is described in [10] and summarized here. To determine the optimal VTL locations for a given number K of VTLs on a given route, we first evenly divide the route into N segments, called sections. In practice, each section is assumed to be 50 to 100 ft long. We can reasonably assume that speed is approximately constant along each short section. Each VTL is associated with a link, which consists of one or multiple consecutive sections. We assume that the VTL is always in the middle of its link. This way, the studied route is divided into K links, where K is less than N . The travel time of the route is equal to the summation of all link travel times. Under the above assumptions, the optimal VTL location problem becomes how to determine the optimal starting and ending sections of each of the links.

Optimal Placement. Posed as DP Deploying VTLs can be solved optimally in a staged process, with one VTL deployed in each stage. This is a key feature in the formulation of the problem as a DP model. In particular, if we are to deploy K VTLs, the model will contain K stages. The state variable of each stage is the starting location of the link associated with the VTL of the stage, and the decision variable is the ending location of the link. The objective is to minimize

the summation of mean squared errors of travel time estimates of all links for a given number of “representative” vehicles.

Optimal Placement Solved by Finding Shortest Path. It is shown in [10] that such a DP model can be represented as an acyclic graph, and solving the optimal VTL location problem is equivalent to finding the shortest path in the graph. The complexity of the algorithm is $O(KN^2)$, which is polynomial and can be used for solving large scale problems. Exclusion zones and minimum spacing constraints can be incorporated into the optimal placement algorithm by modifying the graph representation of the DP model, eliminating proper arcs. The complexity of solving the model with these constraints remains $O(KN^2)$.

4.3 Estimating traffic

Estimation of travel times. To estimate the travel time, the instantaneous travel time was computed, which assumes traffic conditions remain unchanged on every link⁷ from the time the vehicle enters the link until it leaves the link. Therefore, the travel time of the section can be computed by simply summing those of the constituent links at the time a vehicle enters the route. The travel time of each link is computed with the length of a link and the mean speed that is obtained by averaging out speed readings from probe vehicles during an aggregation interval.

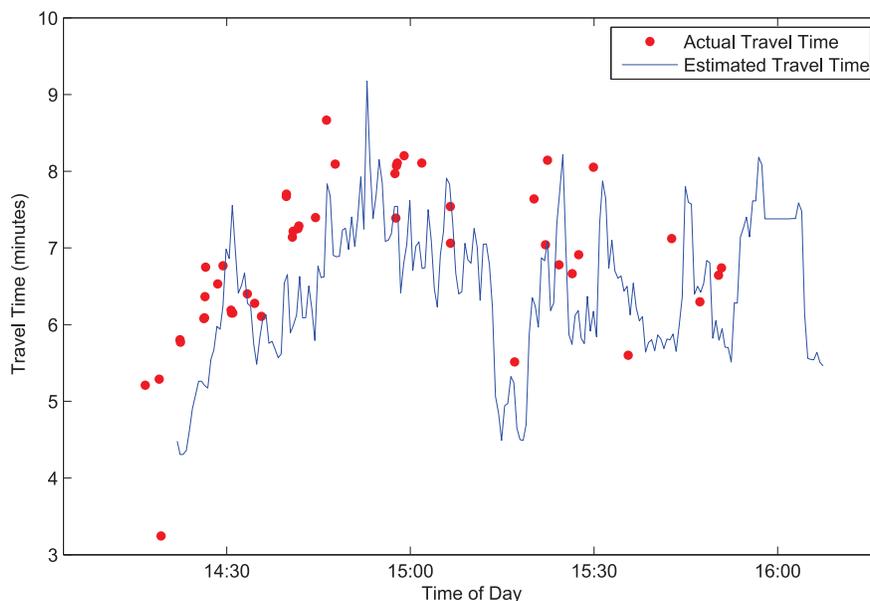


Figure 4.13: Actual travel times compared with an estimate given by the instantaneous method (30 second aggregation interval).

⁷ Each VTL is placed in the middle of its respective link and the conditions on the entire link are given by the VTL reading.

RMS Used for Estimation Error Metric Optimal VTL locations are calculated according to the DP algorithm described above. Using the instantaneous travel time method, we plot actual travel times versus predicted travel times in

Figure 4.13. Ground-truth travel times are obtained by checking logged data of each probe vehicle. Since the errors between actual travel times and predicted travel times are positive and negative, we calculate Root Mean Squared (RMS) error between two sets to see the expected magnitude of a travel time estimation error. For a given 30 second aggregation interval, we achieved a RMS error of about 80 seconds.

Spatial sampling vs. temporal sampling. We compare the accuracy of travel time estimation among the optimal placement approach and the temporal periodic sampling techniques. To conduct a fair comparison, we choose parameters such as a sampling period for temporal periodic sampling techniques and the minimum spacing for the optimal placement to let both two approaches have the same number of total anonymous samples. Figure 4.15 shows RMS error of travel time estimation for both approaches. We choose 15 seconds to 60 seconds for temporal sampling techniques and 10 to 15 trip lines for the optimal placement approach. The positions of trip lines are determined by the optimal placement algorithm. Against the same number of anonymous VTL samples, the optimal placement outperforms the temporal sampling techniques.

Intuitively, the temporal approach may provide data in uninteresting spatial locations that provide less information content. The spatial sampling approach ensures that data is collected in the most useful parts of the network.

4.3.1 Privacy-Accuracy Tradeoffs

This section analyzes the travel time estimation accuracy and privacy preservation of our probabilistic approach, the VTL-based placement algorithm.

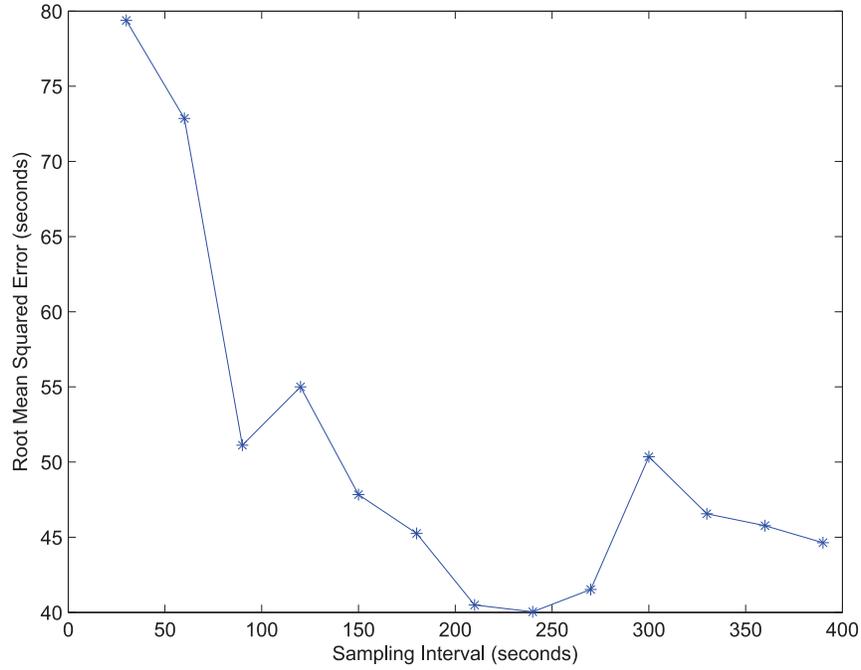


Figure 4.14: Travel time estimate errors by different sampling intervals using 15 VTLs.

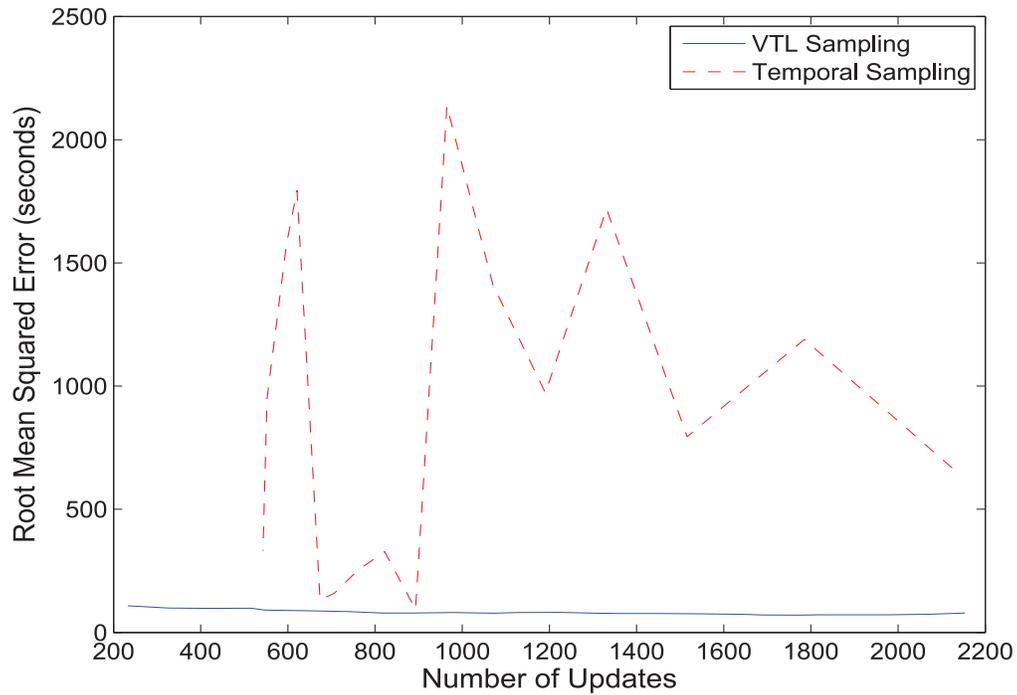


Figure 4.15: Comparison between VTL-based spatial sampling and temporal periodic sampling against the same number of total anonymous samples.

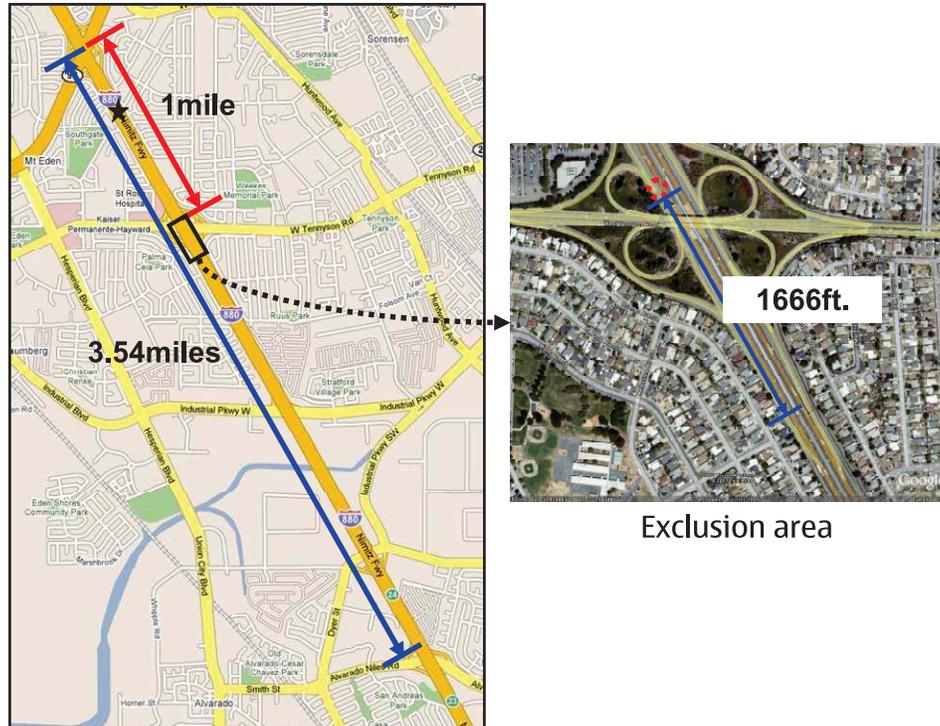


Figure 4.16: Exclusion Area on Test Road Segment. Tracking starts from the point marked by star.

Experimental setup and measurement conditions. To analyze privacy, we measure the distance-to-confusion with two different sets of anonymous VTL samples from both the evenly spaced VTLs (with exclusion area) and the evenly spaced VTLs (without exclusion area). We call the latter spatial periodic sampling. We use the repeated southbound trips of the twenty probe vehicles, which contain the effect of merging traffic from the shorter loop (see Figure 4.16). The southbound direction also has lower traffic volume than the northbound direction, providing a more challenging environment to protect privacy. On the experiment day, we verified from the PeMS [91] highway measurement database that our test road segment (southbound from Route 92 to Alvarado) experienced about 5000 vehs/h as a traffic volume and an average speed of 55 mph. Because we have 88 traces from twenty probe vehicles during our 100 min test period, the penetration rate is about 1% to 2%. Based on the reported average speed and the penetration rate, we obtain the approximate value of the required minimum spacing (800 ft.) from the empirical result graph as shown in Figure 4.11. At the on-ramp, we define a 1670 ft (500 meters) exclusion area. Given the fixed exclusion area, we generated different sets of equidistant trip lines with minimum spacing varying from 333 ft (100 meters) to 1670 ft (500 meters).

Importance of exclusion zone. When we measure the distance-to-confusion, we use an uncertainty threshold of $H = 0.2$, meaning that tracking stops when it incorrectly links samples from different handsets, or when the uncertainty at any step rises above this threshold. We choose the probe vehicles of the main route as the test vectors. Among the set of anonymous samples that are reported at the same VTL, we measure the time deviation of each of them from the projected arrival time of the target probe vehicle, then we calculate the entropy using the empirically obtained probability distribution function of the time deviation between the projected arrival time and each timestamps of anonymous samples at the corresponding VTL. This empirical pdf was measured from the PARAMICS traffic dataset that have similar average speed and traffic volume. In linking anonymous VTL samples that the spatial periodic sampling techniques generate, the adversary removes from the candidate set several anonymous VTL samples that have speed measurements less than 40 mph within the 500 meter distance from the on-ramp, because an adversary has a knowledge on general trend around on-ramps as shown in Figure 4.8. This leads to better tracking performance by reducing the number of likely hypotheses.

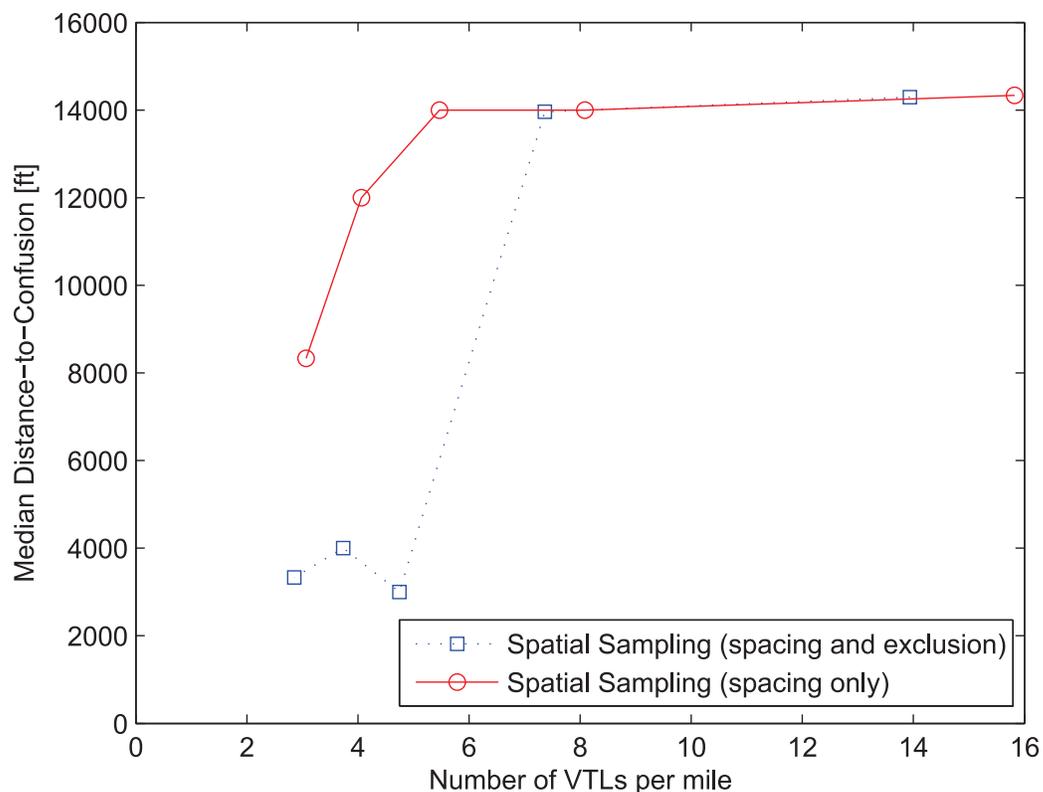


Figure 4.17: Spatial sampling and the benefit of an exclusion zone.

Explanation of figure. The results are shown in Figure 4.17 which plots the median distance-to-confusion against spatial sampling rate. The dotted curve shows the VTL-based placement cases, 1666, 1333, 1000, 666, and 333 feet from the left to the right. The solid curve shows the spatial periodic sampling techniques for the same spacings. We observe that the dotted line drops at spacing of 1000 feet. As we expect from Figure 4.11, two successive anonymous samples that are sampled longer than 800 feet apart experience high tracking uncertainty. Another major reason for the drop in the curve is the existence of the exclusion area. The anonymous samples from the merging traffic can cause high uncertainty outside the exclusion area since the speed measurements look similar to those from the main route traffic.

To study the traffic flow estimation accuracy tradeoff incurred by larger VTL spacings, Figure 4.18 shows the root-mean-square error over the same range of VTL spacings. The travel time estimation generally improves with an increasing number of VTLs, both for evenly spaced and for optimally placed VTLs. In particular, one can see that the error from optimally placed VTLs from the DP algorithm is lower than the naive approach of evenly spaced VTLs.

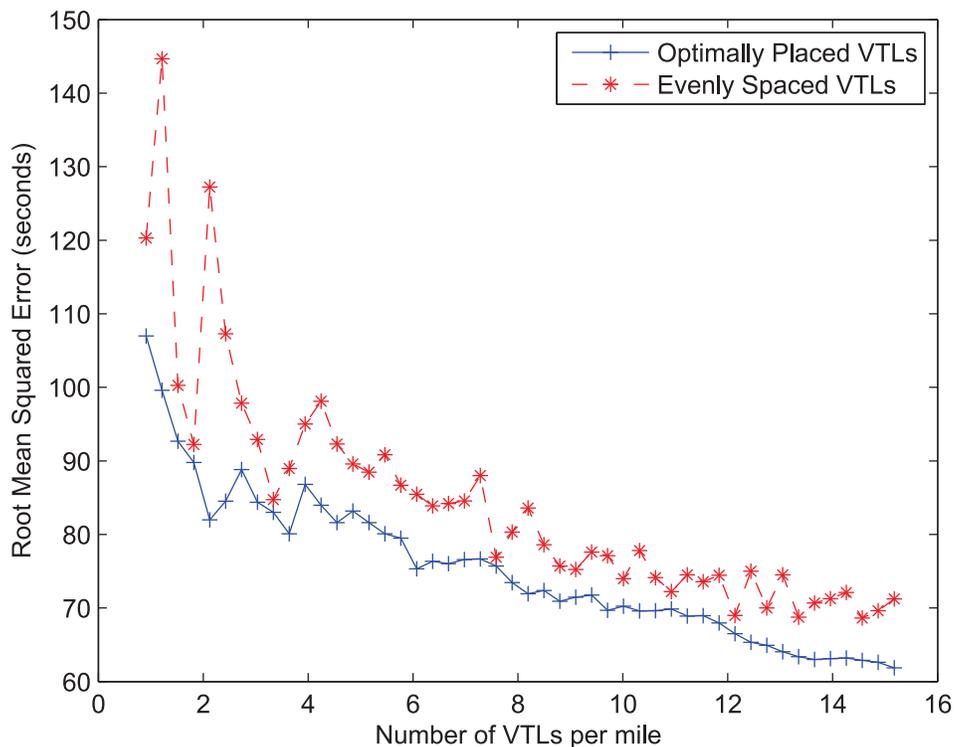


Figure 4.18: Travel time accuracy plotted vs. VTL spacing.

4.4 Discussion

Advantages of VTL based monitoring. In addition to the privacy benefits, a key advantage of virtual trip lines over physical traffic sensors is the flexibility with which they can be deployed. For example, when roadwork is performed, VTLs can be deployed throughout the construction region, providing accurate travel time estimates in an area which often creates significant congestion. Because there is almost no additional cost to deploy the VTLs, and they do not interfere with the construction work or the highway traffic, they can be placed to adjust to the temporarily changed traffic patterns. One could even envision a VTL placement strategy which changes on a much shorter time period, with optimal placement strategies [10] for the morning and evening rush hours, or holiday traffic patterns.

Advantages of the architecture in the present work. The proposed architecture significantly improves privacy protection over earlier proposals, by distributing the traffic monitoring functions among multiple entities, none of which have access to both location and identity records.

Long-term passive attacks. The system protects privacy against passive attacks under the assumption that only a single infrastructure component is compromised. One passive attack that remains an open problem for further study is timing analysis by network eavesdroppers or by the ID Proxy. Given knowledge of the exact trip line locations, which every handset could learn over time, and public travel time information on the road network an adversary could estimate the time needed to travel between any two trip lines. The adversary could then attempt to match a sequence of observed VTL sample message inter-arrival times to these trip line locations. One may expect that the natural variability of driving times provides some protection against this approach. Protection could be further strengthened against network eavesdroppers by inserting random message delays.

The need for periodic VTL re-allocation. If trip line identifiers are used for extended durations, an adversary may match them to actual VTL positions based on the sequence in which probe vehicles pass them. This threat can be alleviated through frequent VTL ID changes. Quantifying these threats and choosing exact tile size and sample frequency parameters to balance privacy and network overhead concerns remain open research problems.

The same methods also offer protection against spoofing attacks that seek to reduce the accuracy of traffic monitoring data. The system does not offer full protection against any active attack on traffic monitoring accuracy, however. For example, a compromised ID proxy could drop messages to reduce accuracy. These challenges remain an open problem for further work.

5 Assimilating Lagrangian Data

This chapter addresses challenges of assimilating Lagrangian data (e.g., velocity measurements from GPS-enabled smartphones) in the density domain, and furnishes a refined presentation of the two methods foreshadowed in Chapter 3: Newtonian Relaxation (NR), and Kalman Filtering (KF). It should be clarified upfront that both methods assume knowledge of the fundamental diagram and the conditions at both boundaries of the section of interest. In addition, the fundamental diagram is used to convert velocity measurements into density measurements compatible with the density based model.

The organization of this chapter is as follows: Section 5.1 provides a literature review in the larger context, contrasting purely statistical methods with those based on physical models. Section 5.2 provides an important discussion on boundary conditions and discretization procedures. The methods are fully described, and tradeoffs between them are presented in Section 5.3. Finally, Section 5.4 assesses the performance of the methods with the NGSIM data set. This chapter closes with Section 5.5, in which a brief numerical example of the nudging factor (peculiar to the the NR technique) is explained.

5.1 Background

Traffic state estimation requires data, which may have different characteristics depending on the type of sensors used to collect it. Numerous freeways around the world, in particular in the US, are equipped with loop detector stations embedded in the pavement to collect traffic data. For each lane, these detectors aggregate information at a given sample time (usually 20 or 30 seconds). Vehicle counts, occupancy and speed are among the information that a detector can collect. These measurements provided by static sensors are traditionally referred to as Eulerian measurements, which means that the detector measures flow through a fixed control volume. On the other hand, Lagrangian sensors collect measurements of the system along a particle trajectory (motion of a car). RFID transponders, smart phones and GPS devices onboard vehicles providing position and/or velocity are examples of mobile (or Lagrangian) sensors.

Data can be used with different modeling approaches for traffic reconstruction. One class of models does not make use of traffic flow physics, and the estimation is based on statistics (current and historical data). Real-time traffic reports are sometimes based on this class of models. Use of these has been a common practice in studies that use smartphones as traffic sensors, in which one of the main goals has been speed or travel time estimation on a stretch of road [13, 72, 100, 115, 117]. Note that four of the aforementioned studies use the cell tower signal to obtain cell phone position, which is less accurate than GPS positioning. In [72] the authors have investigated the use of machine learning techniques to reconstruct travel times on graphs based on sparse measurements collected from GPS devices embedded in smartphones and automobiles.

[113] proposes a method that filters Eulerian data collected from loop detectors (density or velocity) to reconstruct traffic flow. In practice, density can be extracted from occupancy; therefore, density based approaches are possible. The filter is such that in free flow, traffic information propagates downstream, while in congestion it travels upstream. Even though this method takes into account the way in which information propagates in traffic streams, it does not make use of any flow model.

Another class of models is based on traffic flow physics. [79] and [97] independently proposed a first order partial differential equation (referred to as the LWR PDE) to describe traffic evolution over time and space. The LWR PDE is a scalar hyperbolic conservation law, which relates changes in density over time to changes in flow over space. This law is based on the conservation of vehicles principle. Extensions of this model include a second equation accounting for the fact that vehicles do not accelerate/decelerate instantaneously; these are known as second order models [9, 122]. Numerical schemes, such as the Godunov scheme [42], can be used to discretize these continuous models (both first and second order models). While the discussions with respect to first order and second order models led to specific conclusions on both sides, the method in the present work applies to models of any order. In this case, we apply it to a first order model (the LWR model); extensions to second order models are straightforward.

The cell transmission model (CTM), [25, 28] is a discretization of the LWR PDE⁸. The CTM divides the highway into “cells” of length Δx and computes the state of the system (vehicle accumulation or density) every Δt units of time in each cell according to the conservation of vehicles principle. The CTM transforms the nonlinear flux function into a nonlinear discrete operator. Modifications to the CTM can be found in the literature. [86] uses a hybrid system framework to develop the switching-mode model (SMM), which combines discrete event dynamics estimation (mode identification) with nonlinear continuous dynamic state estimation (density estimation). [44] modifies the original merge rule of the CTM and proposes the asymmetric CTM (ACTM).

For traffic state estimation, dynamic flow models of the system can be combined with data collected by sensors, a process known as data assimilation. However, there exist various techniques to perform data assimilation. [40, 107, 109] have used Eulerian measurements to perform data assimilation using Kalman Filtering techniques. While both [40] and [109] use the conservation of vehicles as a model, [107] uses the SMM. [22] and [87] have also performed data assimilation using Kalman Filtering techniques, but with simulated Lagrangian data. [22] uses the conservation of vehicles equation and Kalman Filtering techniques to combine point detection data and probe vehicle data into the travel time estimation, while a second order model is used in [87] to perform data assimilation.

⁸ The CTM is a special case of the Godunov scheme when the fundamental relation between flow and density is assumed to be triangular.

Data assimilation using Lagrangian data to estimate the state of a system is common in other fields such as meteorology and oceanography. In these fields, data assimilation methods range from simple, sub-optimal techniques such as direct insertion, statistical correction, and statistical interpolation to more sophisticated, optimal algorithms such as inverse modeling, variational techniques and a family of methods based on Kalman Filtering [89]. The extrapolation of these techniques to transportation engineering problems appears to be very promising.

A known and simple method used in oceanography is the Newtonian relaxation (or nudging) method [8], used in the present work. The Newtonian relaxation method relaxes the dynamic model of the system toward the observations. To this end, a source term proportional to the difference between the predicted and observed state is included in the constitutive equation of the model (in the present case the LWR PDE).

Most data assimilation methods use boundary conditions at the boundary of the physical domain of interest. These are assumed to be known at some specific locations (usually at the boundaries of the computational domain). Thus, these methods make use of Eulerian data as well.

This chapter presents two methods for traffic state estimation that can handle both Eulerian and Lagrangian measurements, using a first order flow model. The following are addressed:

- The development of a first order flow model that integrates a nudging term (to perform data assimilation) and its corresponding discretization using the Godunov scheme,
- An extension of the Kalman Filtering based method used to estimate traffic state, following the work of [106], which can incorporate Lagrangian observations,
- A Lagrangian data selection procedure to incorporate the measurement data into the discrete model, and,
- An implementation of both algorithms on an extensive dataset (the NGSIM data) for which ground truth is known.

5.2 Description of Preliminary Concepts

The problem of interest is the incorporation of Lagrangian data into traffic flow models for freeways. This section describes the new approach proposed to address this problem and an extension of the Kalman Filtering approach specific to our problem. A description the flow model is followed by a description of the two methods.

5.2.1 Continuous Flow Model

The continuous flow model proposed in the 1950's by Lighthill and Whitman (1955) and Richards (1956) describes the evolution of traffic on an infinite road [79], [97]. The LWR PDE relates density on the road and its flow:

$$\frac{\partial k(x,t)}{\partial t} + \frac{\partial}{\partial x}[q(k(x,t))] = 0 \quad (5-1)$$

The function $k(x,t)$ represents the vehicle density in vehicles per unit length and $q(k(x,t))$ is the flux function in vehicles per unit time (or the fundamental diagram, which is assumed to be triangular and time-space invariant in this work) at location x and time t .

The initial condition corresponds to the density along the road at the beginning of the period of analysis:

$$k(x,0) = k_0(x) \text{ for all } x \quad (5-2)$$

Given the dynamics of the traffic, the lack of accurate knowledge of initial conditions can be counter balanced by the “flush out” effect, i.e. for sufficiently large periods of time, the influence of the initial conditions becomes negligible.

For a finite road, the boundary conditions correspond to the density at the upstream and downstream ends of the section ($x = a$ and $x = b$, respectively), and they could be obtained from sensors at these locations, such as loop detectors. Let $k_a(t)$ and $k_b(t)$ be the desired density or the density measurements collected at time t by these sensors at boundaries a and b . Let $k(a,t)$ and $k(b,t)$ be the modeled density or the density computed with equation (5-1) at time t at the boundaries a and b . If boundary conditions are to be applied in the strong sense, the desired density and the modeled density should be equal all the time. That is, $k(a,t) = k_a(t)$ and $k(b,t) = k_b(t)$ for all t .

However, for a proper characterization of the solution to equation (5-1), weak boundary conditions are required, which implies that the previous equalities between the desired and modeled densities do not hold all the time. Boundary conditions only apply (that is, the modeled density is made equal to the desired density) on the boundary of the section where characteristics are entering the computational domain. Otherwise, the modeled density will not necessarily be equal to the desired density.

The way in which characteristics move depends on the traffic conditions. In the free flow regime, characteristics travel downstream (i.e. in the same direction as the flow) while in the congested regime they travel upstream. Characteristics enter the domain from the upstream (resp. downstream) boundary at location a (resp. b) if free-flow (resp. congested) conditions prevail at that location. Therefore, the upstream boundary condition is relevant (and thus can be applied in the strong sense) only when a free flow condition is observed at that point. Otherwise, the boundary condition is irrelevant and conditions are dictated by downstream traffic. The opposite is true for the downstream boundary, in which the boundary condition is relevant only if congestion is observed at that point.

Mathematically, the weak boundary conditions for the specific case of the LWR PDE can be found in [12] for general hyperbolic conservation laws. In the present context, they can be written in simpler form as follows:

$$\left. \begin{array}{l} \text{(i) } k(a, t) = k_a(t) \quad \text{and} \quad q'(k_a(t)) \geq 0 \quad \text{or} \\ \text{(ii) } q'(k(a, t)) \leq 0 \quad \text{and} \quad q'(k_a(t)) \leq 0 \quad \text{or} \\ \text{(iii) } q'(k(a, t)) \leq 0 \quad \text{and} \quad q'(k_a(t)) \geq 0 \quad \text{and} \quad q(k(a, t)) \leq q(k_a(t)) \end{array} \right\} \text{for all } t \quad (5-3)$$

and

$$\left. \begin{array}{l} \text{(i) } k(b, t) = k_b(t) \quad \text{and} \quad q'(k_b(t)) \leq 0 \quad \text{or} \\ \text{(ii) } q'(k(b, t)) \geq 0 \quad \text{and} \quad q'(k_b(t)) \geq 0 \quad \text{or} \\ \text{(iii) } q'(k(b, t)) \geq 0 \quad \text{and} \quad q'(k_b(t)) \leq 0 \quad \text{and} \quad q(k(b, t)) \leq q(k_b(t)) \end{array} \right\} \text{for all } t \quad (5-4)$$

In equations (5-3) and (5-4), $q'(k)$ is the slope of the flux function $q(k)$, defined as $q'(k) = \frac{dq}{dk}$.

These conditions were derived by [28] from physical principles. Note that only (i) implies that the modeled density is equal to the desired density. In the other two cases, the modeled density is driven by the model and is not necessarily equal to the desired density.

The interpretation of the conditions in (5-3) is as follows (interpretation for condition (5-4) is analogous). One of three alternatives is possible, based on the state of the system directly downstream $x = a$ (i.e. at $x = a + \varepsilon$, where ε is small):

Either (i). This means that it is possible to impose the desired density $k_a(t)$ upstream from the section of interest, therefore $k(a, t) = k_a(t)$. For example, a low flow situation in which the upstream inflow is dictated by the demand falls into this category.

Or (ii). Both the desired inflow and modeled inflow are congested flows, given by the negative slope of $q(\cdot)$. In this case, the solution (the boundary condition) is driven by the state, i.e. the inflow allowed by the current state of congestion.

Or (iii). The modeled flow is congested, but the desired inflow (demand) is not. However, the modeled flow value is less than the desired flow value on the highway. In this case, the flow acceptable by the highway is smaller than the flow which the boundary condition seeks to push. Therefore, the boundary condition cannot drive the inflow.

While the physical reasoning between these three cases could probably be organized with a different Boolean logic (see in particular [28] for a physical analysis) equations (5-3) and (5-4)

above are mathematically required for the problem to be well-posed. In particular, they ensure existence and uniqueness of an entropy solution to this equation [12, 88] for a bounded domain, from which follows convergence of the Godunov discretization scheme, which is key to ensure the proper numerical solution to the problem. The first known instantiation of this Boolean type condition is due to [75], and more recently adapted for highway specific flux functions by [105].

In practice, these boundary conditions are implemented using ghost cells. These cells correspond to the input and output cells as proposed by [25], and are presented in the subsequent sections.

5.2.2 Discretization Method

For implementation purposes, the LWR PDE needs to be discretized. To this end, the freeway section is divided into I cells (each one of length Δx distance units, and indexed by i starting from upstream). Time is divided into H time steps (each one of length Δt time units, and indexed by h). In order to meet the Courant-Friedrichs-Lewy (CFL) stability condition [76], which states that a vehicle traveling at the free flow speed v_f cannot traverse more than one cell in one time step, the condition $\Delta t \cdot v_f \leq \Delta x$ should be met. At every time step, the model estimates the density in each cell according to the following expression:

$$k_i^{h+1} = k_i^h - r(q_{i+1}^h - q_i^h) \quad i = 1, 2, \dots, I \quad \text{and} \quad h = 0, 1, \dots, H - 1 \quad (5-5)$$

The parameter r is the inverse of the speed needed to travel one cell in exactly one time step (i.e. $r = \Delta t / \Delta x$), while k_i^h is the discrete approximation of the density in cell i at time step h . The variable q_i^h is the flow into cell i between time h and $h + 1$, and depends nonlinearly on the density of cells $i - 1$ and i . It can be computed using the Godunov scheme as follows:

$$q_i^h = \begin{cases} q(k_i^h) & \text{if } k_c \leq k_i^h \leq k_{i-1}^h \\ q(k_c) & \text{if } k_i^h \leq k_c \leq k_{i-1}^h \\ q(k_{i-1}^h) & \text{if } k_i^h \leq k_{i-1}^h \leq k_c \\ \min\{q(k_{i-1}^h), q(k_i^h)\} & \text{if } k_{i-1}^h \leq k_i^h \end{cases} \quad (5-6)$$

In equation (5-6), k_c is the critical density, which corresponds to the point at which the flux reaches its only maximum (the flux function is concave).

As was explained earlier, weak boundary conditions are required for a proper characterization of the solution of the LWR PDE (5-1). In the implementation, one ghost cell is inserted at each boundary of the section. Ghost cells are outside of the physical domain, and are virtual cells on

which the Godunov update equations (5-5) and (5-6) automatically apply weak boundary conditions (5-3) and (5-4), by construction of the Godunov scheme [76, 105]. The ghost cells contain the boundary conditions and allow the first and last cells of the computational domain to be updated depending on the existing traffic conditions (free flow, congested, or a combination of the two). This approach was successfully implemented and tested in earlier work [105] with traffic data collected from loop detectors. The equations for the ghost cells ($i = 0$ and $i = I + 1$) and for the initial conditions ($h = 0$) are given by equations (5-7) and (5-8), respectively:

$$k_0^h = \frac{1}{\Delta t} \int_{(h-1)\Delta t}^{h\Delta t} k_a(t) dt \quad \text{and} \quad k_{I+1}^h = \frac{1}{\Delta t} \int_{(h-1)\Delta t}^{h\Delta t} k_b(t) dt \quad h = 0, 1, \dots, H-1 \quad (5-7)$$

$$k_i^0 = \frac{1}{\Delta x} \int_{x_{i-1}}^{x_i} k_0(x) dx \quad i = 1, 2, \dots, I \quad (5-8)$$

5.2.3 Definitions: State, Estimated, and Observed Variables

The state variable characterizes the state of a dynamical system. In the present case, it corresponds to vehicle density, and it is denoted by $k(x, t)$. By definition, it satisfies the physical model, which is the LWR PDE (5-1) in the present case.

The state estimate (or estimated variable) is the result of the state estimation process, and it is distinguished from the model variable by the use of a hat. We denote $\hat{k}(x, t)$ as the state estimate.

To perform the estimation, it is assumed that some variable is measured. In the present case, the available data consists of position and velocity measurements from GPS-equipped vehicles at different times. Therefore, a relationship between the velocity and the density is needed in order to relate the measured or observed variable and the estimated variable. This issue is addressed in the following paragraphs.

The observed position and velocity of an equipped vehicle labeled j at time t are denoted by $s_j^o(t)$ and $v_j^o(t)$, respectively (superscript o denotes *observation*). It is assumed that the observed velocity at the corresponding location and time is provided by the individual probe

velocity measurement. Using $v(x, t)$ to denote the average velocity field on the freeway, we define the observed velocity at time t and location $s_j^o(t)$ by:

$$v^o(s_j^o(t), t) = v_j^o(t) \quad (5-9)$$

This assumption may not be appropriate when dealing with multi-lane freeways, where different lanes might have different speeds. In these cases, this problem requires a specific treatment, which should use more sophisticated models, and are out of the scope of this work⁹. As a first approach, we treat the freeway as a single traffic stream.

The fundamental diagram relates the flow, the density and the velocity of the flow on a section of road. In particular, this relationship can be used to infer the density from the velocity. That is, using the fundamental diagram, $v^o(s_j^o(t), t)$ can be converted into the observed density at the point of measurement, denoted by $k^o(s_j^o(t), t)$. In reality, $k^o(s_j^o(t), t)$ is not the observed density but an estimate based on the speed measurement and the fundamental diagram. Since the fundamental diagram is only a model, and in reality flow-density points (or traffic states) do not necessarily lay on a line, this conversion is expected to introduce error in the observed density.

If a triangular fundamental diagram is used, a problem arises when the speed $v_j^o(t)$ is high. In fact, if the measured speed $v_j^o(t) \geq v_f$, where v_f is the free flow speed, different combinations of flow and density have the same free flow speed. Indeed, under free flow conditions it is not possible to observe the local density through speed measurements in the way described in the previous paragraph. For these cases, free flow conditions will be assumed and a “free flow density” value (denoted k^{FF}) will be used. The value for k^{FF} can change in time and space, and can be obtained from historical data. This seems reasonable considering that our main interest is to obtain accurate density estimates specifically when congestion arises.

Therefore, the observed density is given by:

⁹ In multi-lane freeways, more than one vehicle could send a measurement from the same location $s_j^o(t)$ at the same time t . In this case, the observed velocity is the average of all the measurements from location $s_j^o(t)$ at time t .

$$k^o(s_j^o(t), t) = \begin{cases} \frac{w \cdot k_j}{v^o(s_j^o(t), t) + w} & \text{if } v^o(s_j^o(t), t) < v_f \\ k^{FF} & \text{otherwise} \end{cases} \quad (5-10)$$

Parameters k_j and w are the jam density and the slope of the right branch of the triangular fundamental diagram, respectively. The observed velocity $v^o(s_j^o(t), t)$ comes from equation (5-9).

5.3 Explanation of Proposed Methods

In order to perform traffic state estimation (vehicle accumulation or density, in this case) along the section of interest using Lagrangian data, specific data assimilation methods need to be developed. If accurate traffic data from all locations was available at all times, traffic state estimation would not be needed since the state could be directly inferred from the data. In practice, however, data is not always available from all locations and is, in general, very sparse.

As explained below, different approaches can be used to obtain estimates for locations without measurements. The approaches adopted here use traffic flow models, based on vehicle accumulation (or density).

Two methods are proposed and discussed in the following sections, which rely on the following assumptions:

- GPS-enabled mobile phones traveling on the section of interest are sampled in time. They report their position and velocity at specific time intervals T .
- Boundary conditions are known. That is, the density at both ends of the section of interest is available. This data can be provided by loop detectors. However, as we shall see later, the availability of loop detector data at the boundaries is not a critical requirement.
- The fundamental diagram is assumed to be triangular and known.
- Information from intermediate ramps is not required. In fact, this constitutes one of the main features of the proposed methods, since information from ramps is rarely available.

5.3.1 Newtonian Relaxation Method

The Newtonian relaxation (or nudging) method is a simple heuristic method that has been used for data assimilation in the field of environmental fluid mechanics [8]. In oceanography, GPS-

equipped drifters are used to estimate the velocity field of rivers, using shallow water models. The extension of this technique to transportation engineering problems appears to be very promising, since GPS-equipped vehicles are similar to the drifters, and we aim to estimate the state of the freeway in terms of the vehicle density.

The Newtonian relaxation method relaxes the dynamic model of the system toward the observations. To this end a source term, called the nudging term, proportional to the difference between the estimated and observed state, is included in the constitutive equation of the model, which in the present case is the LWR PDE (5-1):

$$\frac{\partial \hat{k}}{\partial t} + \frac{\partial q(\hat{k})}{\partial x} = - \sum_{j=1}^J \sum_{t_j^p \in \Omega_j^t} \lambda(x - s_j^o(t_j^p), t - t_j^p) \cdot [\hat{k}(s_j^o(t_j^p), t_j^p) - k^o(s_j^o(t_j^p), t_j^p)] \quad (5-11)$$

The summation over the index j in the RHS of equation (5-11) accounts for the J different vehicles equipped with Lagrangian sensors, while the second summation includes all the observations sent by each vehicle j before the current time t . The expression in equation (5-11) assumes that vehicle j sends observations at times $t_j^p \in \Omega_j^t$, where Ω_j^t represents the set of times until t at which measurements from vehicle j are performed and used for data assimilation (note that necessarily $t_j^p < t$ since only observations from the past are available at the current time t).

The nudging factor $\lambda(\delta_x, \delta_t)$ represents the weight of each observation to be applied to the solution. This weight is expected to become negligible (and eventually to become zero) away from the measurement location and after the measurement time. For this reason, we have adopted here an expression for the nudging factor $\lambda(\delta_x, \delta_t)$ that takes this into account, and can be found in [60]:

$$\lambda(\delta_x, \delta_t) = \begin{cases} \frac{1}{T_a} \exp\left(-\left(\frac{\delta_x}{X_{\text{nudge}}}\right)^2\right) \exp\left(-\frac{\delta_t}{T_d}\right) & \text{if } |\delta_x| \leq X_{\text{nudge}} \text{ and } 0 < \delta_t \leq T_d \\ 0 & \text{otherwise} \end{cases} \quad (5-12)$$

The factor dies out on a space and time scale of X_{nudge} and T_d , respectively. Therefore, close to where and when the observation is made, $\lambda(\delta_x, \delta_t)$ nudges the solution towards the observations. The parameter T_a has units of time and determines the strength of the nudging

factor. A numerical example to show the effect of the nudging factor is provided at the end of this chapter.

In the present context, the nudging term “adds” or “removes” vehicles from the state of the flow model depending on whether the model underestimates or overestimates the number of vehicles on the freeway. In Section 5.3.2 we will discuss the implication of adding or removing vehicles on the conservation of vehicles principle.

5.3.1.1 Numerical Implementation: Discrete Model

The discretization of the nudging term in the RHS of equation (5-11) needs to be added to equation (5-5). The final expression for the discretized model is as follows:

$$\hat{k}_i^{h+1} = \hat{k}_i^h - r(\hat{q}_{i+1}^h - \hat{q}_i^h) - \Delta t \sum_{j=1}^J \sum_{t_j^p \in \Omega_j^{h,\Delta t}} \lambda(x_i - s_j^o(t_j^p), h\Delta t - t_j^p) \cdot \left[\hat{k}_{c_{jp}}^{m_{jp}} - k_{c_{jp}}^{o,m_{jp}} \right] \quad (5-13)$$

$i = 1, 2, \dots, I$ and $h = 0, 1, \dots, H - 1$

The notation introduced in equation (5-13) is explained below:

- \hat{q}_i^h : estimated flow into cell i in time step h . This flow is obtained by application of (5-6) to the set of \hat{k}_i^h ,
- x_i : location of the beginning of cell i , $x_i = x_0 + (i - 1) \Delta x$, for $i = 1, 2, \dots, I$, where x_0 is the beginning of the section of interest,
- c_{jp} : cell index corresponding to location $s_j^o(t_j^p)$, which is the location where vehicle j is at time t_{jp} when its p -th observation is sent, $c_{jp} = \left\lceil \frac{s_j^o(t_j^p) - x_0}{\Delta x} \right\rceil$, where $s_j^o(t_j^p) > x_0$,
- m_{jp} : time step corresponding to location t_j^p , which is the time (in time units) when the p -th observation from vehicle j occurs, $m_{jp} = \left\lceil \frac{t_j^p}{\Delta t} \right\rceil$.

The last term in the RHS of equation (5-13) is the discretization of the nudging factor times the difference between estimated and measured density (terms inside the square brackets). The nudging factor is given by the expression in (5-12).

5.3.2 Kalman Filtering Based Method

Kalman Filtering is a recursive method used to estimate the state of a discrete process governed by a linear stochastic dynamical system [14] in the presence of noisy measurements. The method assumes that the way in which the state of the system (density in this case) evolves is linear and known, and is referred to as the dynamics, or state equation (which includes a process noise). Noisy measurements of the output of the system (i.e. observed density) are available, and the measurement or observation equation relates the output and the state of the system. Knowing the covariance of both the process and the measurement error, the method obtains the best estimate of the state of the system in the sense of least squares.

Kalman Filtering techniques have been proposed to perform traffic state estimation for loop detector data [106, 107]. Given the nonlinearity of equation (5-6), the model in (5-5) is nonlinear as well. Therefore, conventional Kalman Filtering can only be applied to linear subsets of these dynamics. This technique can be extended to cases in which data is provided by mobile sensors, described next.

5.3.2.1 State space representation

The state space representation of the system consists of two equations: the dynamics (or state) equation and the measurement (or observation) equation.

The dynamics equation describes how the state of the process (density) evolves over time and space. Since the constitutive equation in (5-5) is nonlinear, it needs to be linearized first. The hybrid system framework used in [86] is adopted for this purpose. Discrete event dynamic estimation is performed to identify the traffic condition or mode of the section of interest. That is, the mode of each cell (i.e., free flow or congested) needs to be determined at the beginning of each time interval.

Previous studies [86, 107] have proposed different ways to identify the mode on a short section of highway. For longer sections, the number of possible modes increases, adding complexity to mode identification.

In the present study, state estimates (and indirectly Lagrangian measurements) are used to identify the mode. At the end of time interval h , density estimates for every cell are available. These estimates consider all the observations collected until time step h (inclusive) and are referred to as the *a-posteriori* estimates at time step h . The *a-posteriori* estimate at cell i at time step h is denoted by $\hat{k}_i^{+,h}$. The mode chosen for each cell will depend on the value of $\hat{k}_i^{+,h}$. In other words, if $\hat{k}_i^{+,h} > k_c$, cell i is congested; otherwise, cell i is in free flow (k_c is the critical density). Note that the mode is being identified by using the state estimates and not by measuring the actual state of the system.

Once the mode for each cell has been identified, the flow into cell i between time step h and $h + 1$, q_i^h , becomes linear in the densities in cells $i - 1$ and i . Therefore, equation (5-5) can be written as follows:

$$\mathbf{k}_{h+1} = \mathbf{A}_h \cdot \mathbf{k}_h + \mathbf{B}_h \cdot \mathbf{u}_h + \mathbf{B}_h^J \cdot k_J + \mathbf{B}_h^Q \cdot q_{\max} + \mathbf{w}_h \quad (5-14)$$

Bold letters represent matrix or vector notation. The scalars k_J and q_{\max} are the jam density and the maximum flow, respectively. The vector \mathbf{u}_h is the input vector at time h , which includes the density at the boundaries of the domain, and \mathbf{w}_h is the process error (caused for instance by the fact that not all the entry/exit counts are available). Equation (5-14) is linear in the state \mathbf{k}_h , which is defined as $\mathbf{k}_h = [k_1^h \ k_2^h \ \dots \ k_I^h]^T$. The matrices \mathbf{A}_h , \mathbf{B}_h , \mathbf{B}_h^J , and \mathbf{B}_h^Q , depend on the traffic conditions or mode of the section, which shows the importance of the mode identification step. The algebraic expression of these matrices for specific cases can be found in [86].

The measurement, or observation equation, projects the state vector into the measurements provided by Lagrangian sensors, \mathbf{y}_h , with the one predicted by the model:

$$\mathbf{y}_h = \mathbf{C}_h \cdot \mathbf{k}_h + \mathbf{v}_h \quad (5-15)$$

The vector \mathbf{v}_h represents the measurement noise. The matrix \mathbf{C}_h is time dependent, and its size and elements depend on the location of the measurements. It only contains zeros and ones, whose position in the matrix \mathbf{C}_h depend on the locations at which the measurements are taken. The time dependency of \mathbf{C}_h is a major challenge and is directly linked to the Lagrangian aspect of the measurements.

In summary, the dynamics or state equation and the measurement or observation equation of the system are given by equations (5-14) and (5-15), respectively.

5.3.2.2 Kalman Filtering

The following notation is used:

- $\hat{\mathbf{k}}_h$: *a-priori* state estimate of \mathbf{k}_h , where $\hat{\mathbf{k}}_h = [\hat{k}_1^h \ \hat{k}_2^h \ \dots \ \hat{k}_I^h]^T$,
- $\hat{\mathbf{k}}_h^+$: *a-posteriori* state estimate of \mathbf{k}_h , where $\hat{\mathbf{k}}_h^+ = [\hat{k}_1^{+,h} \ \hat{k}_2^{+,h} \ \dots \ \hat{k}_I^{+,h}]^T$,
- \mathbf{P}_h : *a-priori* estimate error covariance, where $\mathbf{e}_h = \mathbf{k}_h - \hat{\mathbf{k}}_h$ is the *a-priori* estimate error,
- \mathbf{P}_h^+ : *a-posteriori* estimate error covariance, where $\mathbf{e}_h^+ = \mathbf{k}_h - \hat{\mathbf{k}}_h^+$ is the *a-posteriori*

estimate error.

The difference between the *a-priori* and *a-posteriori* estimates at time step h is the fact that the *a-priori* estimates do not take into account the observations collected at time step h , while the *a-posteriori* estimates do. That is, the *a-posteriori* estimate is an updated version of the *a-priori* estimate.

Kalman Filtering provides a set of recursive equations to estimate the vector state. The equations are as follows [14]:

$$\hat{\mathbf{k}}_{h+1} = \mathbf{A}_h \cdot \hat{\mathbf{k}}_h^+ + \mathbf{B}_h \cdot \mathbf{u}_h + \mathbf{B}_h^J \cdot k_J + \mathbf{B}_h^Q \cdot q_{\max} \quad (5-16)$$

$$\mathbf{P}_{h+1} = \mathbf{A}_h \cdot \mathbf{P}_h^+ \cdot \mathbf{A}_h^T + \mathbf{Q} \quad (5-17)$$

$$\mathbf{F}_{h+1} = \mathbf{P}_{h+1} \cdot \mathbf{C}_{h+1}^T [\mathbf{C}_{h+1} \cdot \mathbf{P}_{h+1} \cdot \mathbf{C}_{h+1}^T + \mathbf{R}]^{-1} \quad (5-18)$$

$$\hat{\mathbf{k}}_{h+1}^+ = \hat{\mathbf{k}}_{h+1} + \mathbf{F}_{h+1} \cdot (\mathbf{y}_{h+1} - \mathbf{C}_{h+1} \hat{\mathbf{k}}_{h+1}) \quad (5-19)$$

$$\mathbf{P}_{h+1}^+ = (\mathbf{I} - \mathbf{F}_{h+1} \cdot \mathbf{C}_{h+1}) \cdot \mathbf{P}_{h+1} \quad (5-20)$$

Initial conditions $\hat{\mathbf{k}}_0$ and \mathbf{P}_0 are assumed to be known. \mathbf{Q} and \mathbf{R} are the covariance matrices of the process and measurement error, respectively. \mathbf{F}_h is known as the Kalman gain at time step h . Note that there is a striking similarity between the two methods: the second term in equation (5-19) represents a non physical source (correction) term, which modifies the *a priori* value of the state of the system, similar to the nudging term.

5.3.2.3 Implementation

The state vector contains the density in each cell. At the beginning of time step $h+1$, *a-posteriori* estimates at time step h are available. The traffic conditions on the network at the end of time step h need to be identified in order to determine which set of matrices \mathbf{A}_h , \mathbf{B}_h , \mathbf{B}_h^J , and \mathbf{B}_h^Q to use. The mode will be identified with the process outlined previously, which indirectly uses the Lagrangian observations collected. Note, however, that the mode identification is the most challenging task in the implementation of this method.

Once the mode has been identified at the beginning of time step $h+1$, equation (5-16) and (5-17) are used to obtain the *a-priori* density estimate and its covariance, respectively. At this point, Lagrangian data becomes available to the model, i.e. the observed local density at time $h+1$ will be known for some cells (the quantity and position of the Lagrangian sensors at $h+1$ will determine how many and for which cells the density is observed). With this information, the observed vector \mathbf{y}_{h+1} and matrix \mathbf{C}_{h+1} can be constructed. Then, the Kalman gain is computed using equation (5-18). Finally, the *a-posteriori* density estimate and its covariance are obtained using equation (5-19) and (5-20), respectively. In the event that no

Lagrangian observation is available at time $h + 1$, the matrix \mathbf{C}_{h+1} is set equal to zero, which implies that the Kalman gain is also zero. In this case, the *a-priori* and *a-posteriori* density estimates are the same.

5.3.3 Comments on the Proposed Methods

Note first that in both methods, the observed density is computed using equation(5-10), which includes at least three error sources. The first source of error has to do with equation (5-9), when the velocity of an individual vehicle is assumed to correspond to the velocity at that given location. The second source of error is related to the fact that the fundamental diagram is not exact. Thus, the velocity-to-density conversion, as expressed in equation (5-10), introduces error. This error may not be negligible for free flow because of the approximation made for these cases.¹⁰ The third source of error corresponds to the measurement error in the velocity $v_j^o(t)$, which is expected to be small given the accuracy of GPS.

Both methods are conceptually similar. They add or remove vehicles depending on the difference between an estimated density and the observed density computed using GPS data. For this, they used a so-called “observer equation”, which is derived from the flow model, but includes modifications which integrate the measurements. The methods differ, however, in the way this difference between measurement and estimate is used. The Kalman Filtering based method assumes that an observation obtained from cell i at time step h only corrects or updates the density on the corresponding cell and time step. The flow model propagates the effect in time and space. On the other hand, the Newtonian relaxation method uses the observation to directly affect the density of neighboring cells and for future times. Therefore, the effect of each observation is propagated in time and space directly through the nudging term, but also through the flow model. This may be useful when the number of observations is low, because if cell i is congested at time step h (according to the observation), neighboring locations are expected to also be congested for a certain period of time (congestion takes time to vanish).

Some tradeoffs between the two methods include the following:

Modeling complexity. The Newtonian relaxation method is trivial to develop on almost any flow model, since it only consists in adding a source term weighed properly by the nudging factor. For Kalman Filtering, the difficulty consists in finding the proper linearization, which in the present case cannot be obtained by linearizing the dynamics directly, but by identifying the proper modes (hence the hybrid system approach), which is a research topic in itself.

¹⁰ Note that if density estimates are used to obtain travel times (by computing the velocity), the value used for k^{FF} would not affect the travel time estimates significantly. Therefore, travel time estimates might be more accurate than density estimates under free flow conditions.

Implementation difficulty. The benefit of the Newtonian relaxation method is such that any solver used for forward simulations (of the model) can be directly used and modified to include the nudging term, a feature which has been extensively used in other fields, in particular in oceanography. For Kalman Filtering, additional update equations (outlined earlier) need to be added.

Parameter tuning. The nudging term requires tuning three parameters, which are usually chosen based on physical considerations, while the Kalman Filtering approach requires assumptions on the covariance matrices.

Computational cost. Both methods have roughly the same cost (i.e. the cost of a forward computation), with additional computations required for the Kalman filter because of the observer equation.

Convergence of the estimator. While the convergence of the Kalman filter can be assessed theoretically [14], Newtonian relaxation is a heuristic technique for which no convergence results are known to the authors.

Convergence of the model. By definition of the Godunov scheme used, the discrete model used in Newtonian relaxation without the nudging term can be shown to converge to the solution of the LWR PDE. An extension of this property can be used to develop a convergence result for a PDE which includes a continuous nudging term. No such counterpart is known to the authors for the Kalman filter.

As was briefly explained before, mass (vehicle) conservation is not satisfied by the methods proposed in this article. This is a common feature of estimation techniques, so the next paragraphs explain the considerations which have to be taken into account when doing estimation, and how our methods fit in them.

Traditionally, estimation makes the assumption that the physics of a phenomenon are governed by a constitutive model [14]. The equations for such models can be of various nature, in particular partial differential equations like equation (5-1) or difference equations like equation (5-5). Because of modeling inaccuracy (models are never perfect), and because of measurement noise (measurements are never noise free), empirically measured data almost never satisfies a model perfectly, i.e. plugging data in the model will violate the model. In the present case, the model is represented by equation (5-1) with corresponding continuous state variable $k(x, t)$, and the corresponding discrete model by equation (5-5), with discrete state variable k_I^h . With empirically measured data, $k(x, t)$ would never satisfy equation (5-1) exactly, and k_I^h would never satisfy equation (5-5) exactly.

The field of estimation has produced a large number of techniques which are based on so-called observer equations, which are traditionally denoted by hats (therefore $\hat{k}(x, t)$ and \hat{k}_I^h), which integrate measurements into constitutive equations, but usually do not satisfy modeling

assumptions anymore (i.e., in the present case, $\hat{k}(x, t)$ and \hat{k}_i^h will not satisfy mass – or vehicle – conservation). The field of estimation and data assimilation is based on constructing evolution equations for these estimator states ($\hat{k}(x, t)$ and \hat{k}_i^h in the present case), which are shown to minimize discrepancy error between measurements and model estimates. For example, Kalman Filtering equations add a corrective term to the model equations (therefore violate the model), and provide a least square estimator of the state of the system.

Both the Newtonian relaxation method and the Kalman Filtering based method add or remove vehicles from a cell depending on the relative values of the observed and estimated density. If the counts were known exactly at every entry and exit point of the network, this addition/removal of vehicles would violate the conservation of vehicles. In practice however, on- and off-ramps are rarely equipped with loop detectors (in addition to the fact that specifically to the case of the highway, the observed quantities are never perfect because of measurement errors). In particular, loss or gain of vehicles in between loop detectors is a well known problem in traffic engineering.¹¹ The present methods can thus be used to incorporate Lagrangian data in place of this missing loop detector data. The methods thus conveniently bypass the modeling of networks: for mainlines with on- and off-ramps only (no major intersection between highways), they replace the merge-diverge junctions by Lagrangian data incorporation.

When no loop detector data is available at ramps, numerical simulations are simply underdetermined, because of the lack of inflow and outflow information. When data from ramps is available, because of measurement errors, the data might be inconsistent with the model, in particular vehicle conservation. The methods presented steers the state of the model locally (in x and t) towards the Lagrangian measurements, which is a way to reestablish a value of the state closer to the actual state of the system wherever such measurements are available. Therefore, it compensates for the lack of inflow and outflow information or the corresponding inaccuracy.

Because of the considerations discussed above, the conservation of vehicles that incorporates the on- and off-ramps is replaced by the nudging term (in the Newtonian relaxation method) or the correction term (in the Kalman Filtering based method), which perform the required local adjustments based on measurements.

5.4 Assessment of the Methods

The Newtonian relaxation method (NR) and the Kalman Filtering method (KF) are applied. In the NGSIM data set, we have full knowledge of all positions and speeds of all vehicles during

¹¹ There are algorithms to correct for the measurement errors. For instance, miscounts arising during congestion can be corrected once free flow conditions are restored. See the appendix of [23] for more details.

the entire experiment, which enables an extensive validation of the method against “ground truth”.

Traffic data from the Next Generation Simulation (NGSIM) project¹² is used to evaluate the proposed approach. NGSIM data has been extracted from video, which provides ground truth trajectories for all vehicles. The data consist of the trajectories of all vehicles entering a stretch of US Highway 101S in Los Angeles, CA, during 45 minutes (from 7:50am to 8:35am on June 15, 2005). The site is approximately 0.4 miles in length, with 5 mainline lanes (see Figure 5.1). An auxiliary lane exists between the on-ramp and the off-ramp. Transition from free flow to congestion happens during the first 10-12 minutes of the data set (i.e., part of the section was in the free flow mode while the rest was in the congested mode).

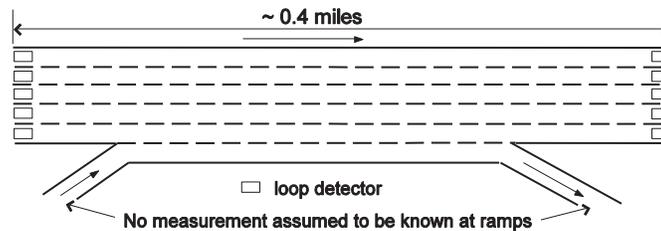


Figure 5.1 Highway US101 S used for the NGSIM dataset.

For implementation purposes, only the five mainline lanes were considered. The full section (1920 ft) has been divided into $I = 16$ cells of $\Delta x = 120$ ft each, and the total simulation time (2610 sec) has been divided into $H = 2175$ time steps of $\Delta t = 1.2$ seconds each. Parameters of the fundamental diagram – maximum flow (q_{\max}), jam density (k_j), critical density (k_c), free flow speed (v_f) and wave speed (w) – were extracted from the PeMS system¹³.

The data is processed in order to emulate loop detector data at both boundaries of the section. Since all vehicle trajectories are known, this can easily be done without usual measurement error associated with loop detectors. These emulated detectors will provide the boundary conditions. Ramps counts could be inferred in the same manner, but they are assumed to be unknown, a realistic assumption given the lack of available infrastructure for most freeways in North America.

¹² <http://ngsim.camsys.com/>

¹³ Assuming a triangular shape of the fundamental diagram: $q_{\max} = 2040$ vphpl, $k_j = 205$ vpmpl, $k_c = 30$ vpmpl, $v_f = 68$ mph, and $w = 11.7$ mph.

5.4.1 Scenarios investigated

Penetration rate and sampling strategies. The implementation of the algorithms selects a subset of the NGSIM data and treats it as Lagrangian data used for the assimilation. Twelve different scenarios were investigated to account for different penetration rates and sampling strategies. The penetration rate P is the proportion of trajectories that are chosen (randomly) as equipped vehicles. These equipped vehicles report their position and speed every time interval T . The reported speed is the average speed over the last τ seconds (Figure 5.2 sketches this sampling strategy). The values chosen for P , T , and τ determine the total number of Lagrangian measurements created for each case investigated. Five different values of P and two values of T were investigated (τ was assumed 6 seconds in all scenarios). Table 5.1 shows the scenarios investigated and the average number of Lagrangian measurements per mile-lane per minute for each one of them.

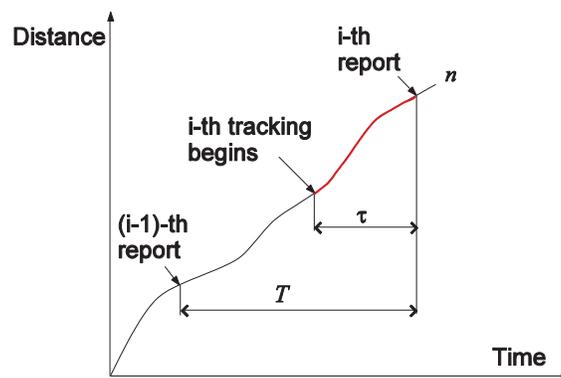


Figure 5.2 Schematic of the sampling strategy on equipped vehicle n .

The travel time for the section of interest is around 2 minutes under congested conditions. Thus, scenarios with $T = 150\text{s}$ assume that each equipped vehicle sends only one report while it is traveling the section. Almost continuous tracking for equipped vehicles is assumed for scenarios 7 to 12 ($T = 10\text{s}$ and $\tau = 6\text{s}$). Because of privacy issues, it is not clear if such a sampling strategy would be socially acceptable for a commercial product, a problem addressed in [53]. This issue is still open and generates ongoing debates. These scenarios were investigated to evaluate the traffic reconstruction potential of the proposed method.

Parameters selection. The nudging factor $(\delta_x \delta_t)$ depends on three parameters. Nudging parameters are typically set heuristically based on physical considerations. Parameters X_{nudge} and T_d determine how far in space and time, respectively, an observation influences the solution. For instance, if congestion is observed at location x at time t , it is expected that congestion also exists at locations $[x - X_{\text{nudge}}, x + X_{\text{nudge}}]$ and between times $[t, t + T_d]$. For

the present case, the values for X_{nudge} and T_d are 180 ft and 15 seconds, respectively. The parameter T_a can be seen as a gain that determines the strength of the nudging term. Therefore, it has to be tuned as a control parameter of the model. For the present study, values of 10, 20 and 30 seconds have provided good results for the scenarios investigated.

5.4.2 Results

The twelve scenarios presented in Table 5.1 were implemented following the methods described in in Section 5.3. For comparison purposes, we use a scenario which only incorporates information from boundary detectors, which we will refer to as Eulerian data only (EDO). This scenario was implemented according to the numerical scheme presented earlier, and does not make use of the ramp counts for the estimation.

Table 5.1: Twelve scenarios investigated in the NGSIM data set.

Case	P (%)	T (sec)	# of Lagrangian Measurements per mile-lane per minute
1	2	150	1
2	5	150	3
3	10	150	6
4	15	150	10
5	20	150	13
6	25	150	16
7	2	10	9
8	5	10	22
9	10	10	42
10	15	10	62
11	20	10	80
12	25	10	97

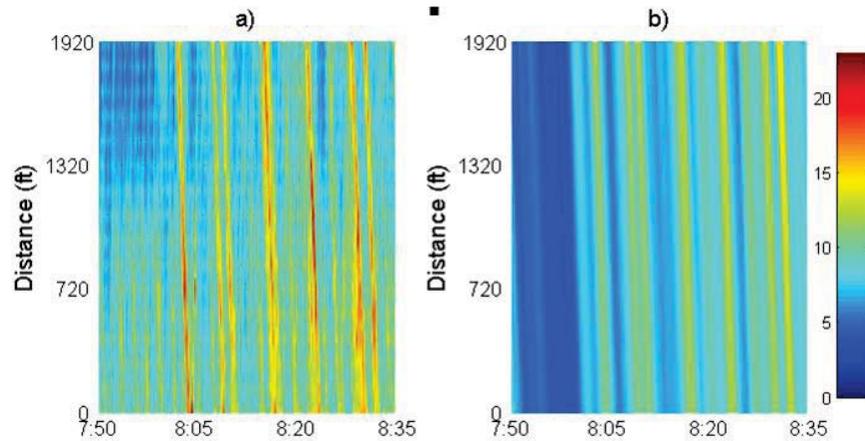


Figure 5.3. Vehicle accumulation per cell for a) ground truth, and b) EDO case.

The *Root Mean Square Error*¹⁴(RMSE) of the accumulation of vehicles (i.e., number of vehicles per cell per time step) is used as a metric for accuracy of the method. The true accumulation of vehicles can easily be computed since all vehicle trajectories are known. The RMSE for the EDO case is 2.6.

Figure 5.3 shows the ground truth and estimated evolution over time (horizontal axis) and space (vertical axis) of the accumulation of vehicles for the EDO case. Colors close to red correspond to high accumulation of vehicles (i.e., congestion). Figure 5.4 shows the same information for scenarios 1, 3, and 9 using the NR (left) and the KF (right). It can be seen that all the estimates are able to capture the main shockwaves traversing the section of interest. The intensity of the shockwave, however, is not equally captured by different scenarios.

¹⁴ The RMSE is defined as: $RMSE = \sqrt{\frac{\sum_h (\hat{z}_h - z_h)^2}{H}}$, where \hat{z}_h and z_h are the estimator and its actual value at time step h , respectively, and H is the total number of observations.

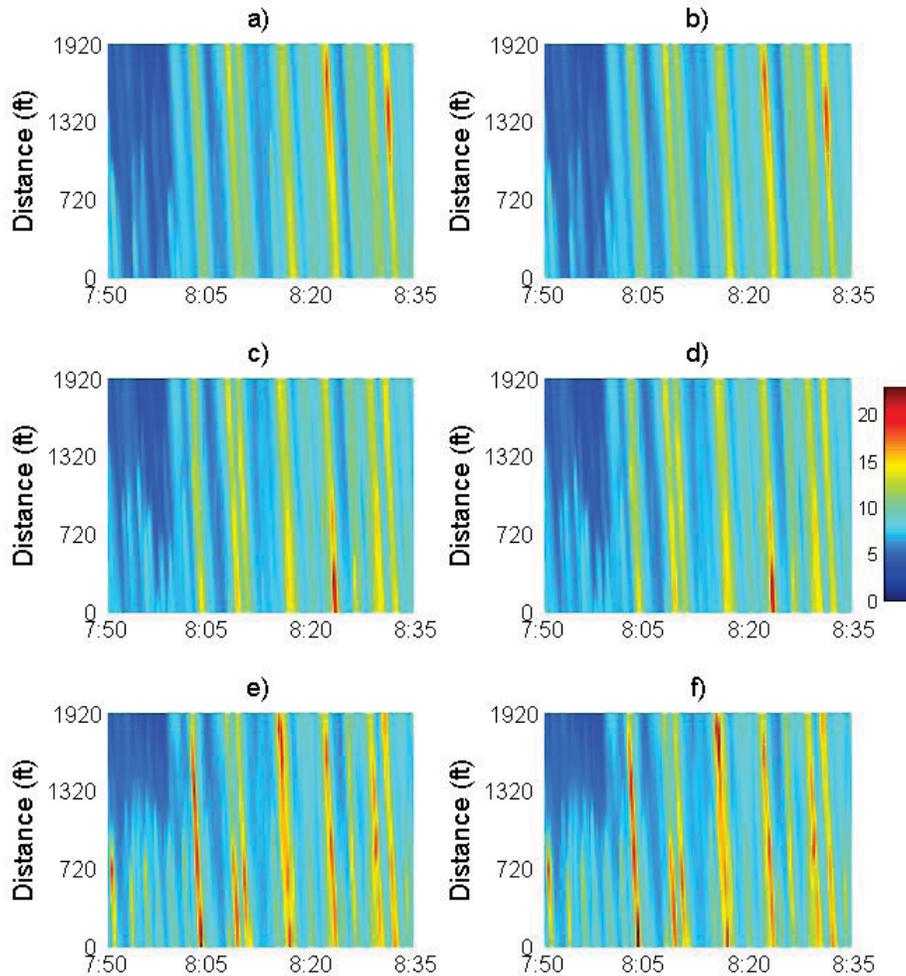


Figure 5.4 Vehicle accumulation (vehicles per cell) estimated using Newtonian relaxation method (left) and Kalman Filtering techniques (right) for scenarios 1 (top), 3 (middle), and 9 (bottom).

Figure 5.3 and Figure 5.4 suggest that the main difference among the estimations occurs during the first 10-12 minutes (before 8:00am), when some waves emanate from intermediate locations. In fact, the fundamental difference between the EDO scenario and scenarios with Lagrangian measurements – in terms of the RMSE – happens during this period of time. The added value of Lagrangian measurements is thus clear. They enable the methods to capture phenomena otherwise not detectable with loop detectors only. For the remaining period of time (until 8:30am), the scenarios with Lagrangian measurements do not show a significant improvement when compared with the EDO. Figure 5.5 shows the evolution of the actual total accumulation of vehicles on the entire section over time and its corresponding estimate with EDO and scenario 5 for NR and KF. This graph confirms that the main difference occurs during the first 10-12 minutes.

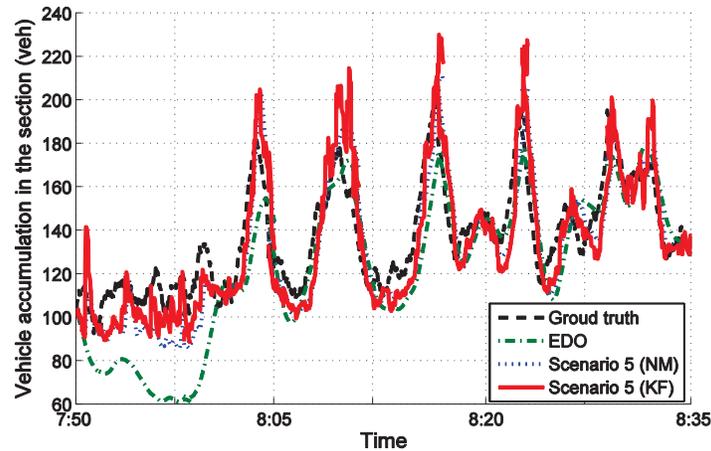


Figure 5.5 Total vehicle accumulation on the entire section.

It is expected that for longer sections between detectors, with more intermediate ramps and probably more waves emanating from intermediate locations, the difference between EDO and scenarios with Lagrangian measurements would be larger.

Table 5.2 shows the RMSE for each scenario investigated and its corresponding Percentage of Improvement (PoI) when compared with the EDO case. The PoI for scenario i is computed as

$$PoI_i = \frac{RMSE_{EDO} - RMSE_i}{RMSE_{EDO}} \cdot 100.$$

Note that the RMSE for each scenario corresponds to the average over 20 different realizations¹⁵. Figure 5.6a shows the same information as a function of the total number of Lagrangian measurements used.

¹⁵ For the same penetration rate P and sampling strategy (P, τ) , different realizations consider different vehicles as equipped vehicles and different times when measurements are sent.

Table 5.2: RMSE for each scenario and its improvement with respect to the EDO case.

Scenario	Newtonian relaxation		Kalman Filtering	
	RMSE	Improvement (%)	RMSE	Improvement (%)
1	2.44	6.1	2.33	10.1
2	2.35	9.8	2.18	16.2
3	2.21	14.9	2.04	21.6
4	2.16	16.7	1.98	23.7
5	2.14	17.7	1.94	25.1
6	2.11	18.9	1.90	27.0
7	2.27	12.8	2.00	23.2
8	2.18	16.2	1.89	27.4
9	2.06	20.6	1.81	30.3
10	2.02	22.3	1.80	30.7
11	2.03	22.0	1.78	31.3
12	1.98	23.8	1.78	31.6

Figure 5.6a suggests an increasing performance of the algorithms as the number of observations increases. The performance reaches a saturation level for large number of observations (more than 40 observations per mile-lane per minute in this case) at about 22% for the NR and 31% for the KF. The solid line shows the performance for scenarios 1 to 6 (one observation per equipped vehicle), while the dotted line corresponds to scenarios 7 to 12 (almost continuous tracking). For the NR, the lines in Figure 5.6a also suggest that for similar numbers of observations (scenarios 4 and 7), having more vehicles sending observations less frequently seems to produce more accurate results than having few vehicles being continuously tracked. This gap is not observed in the KF implementation.

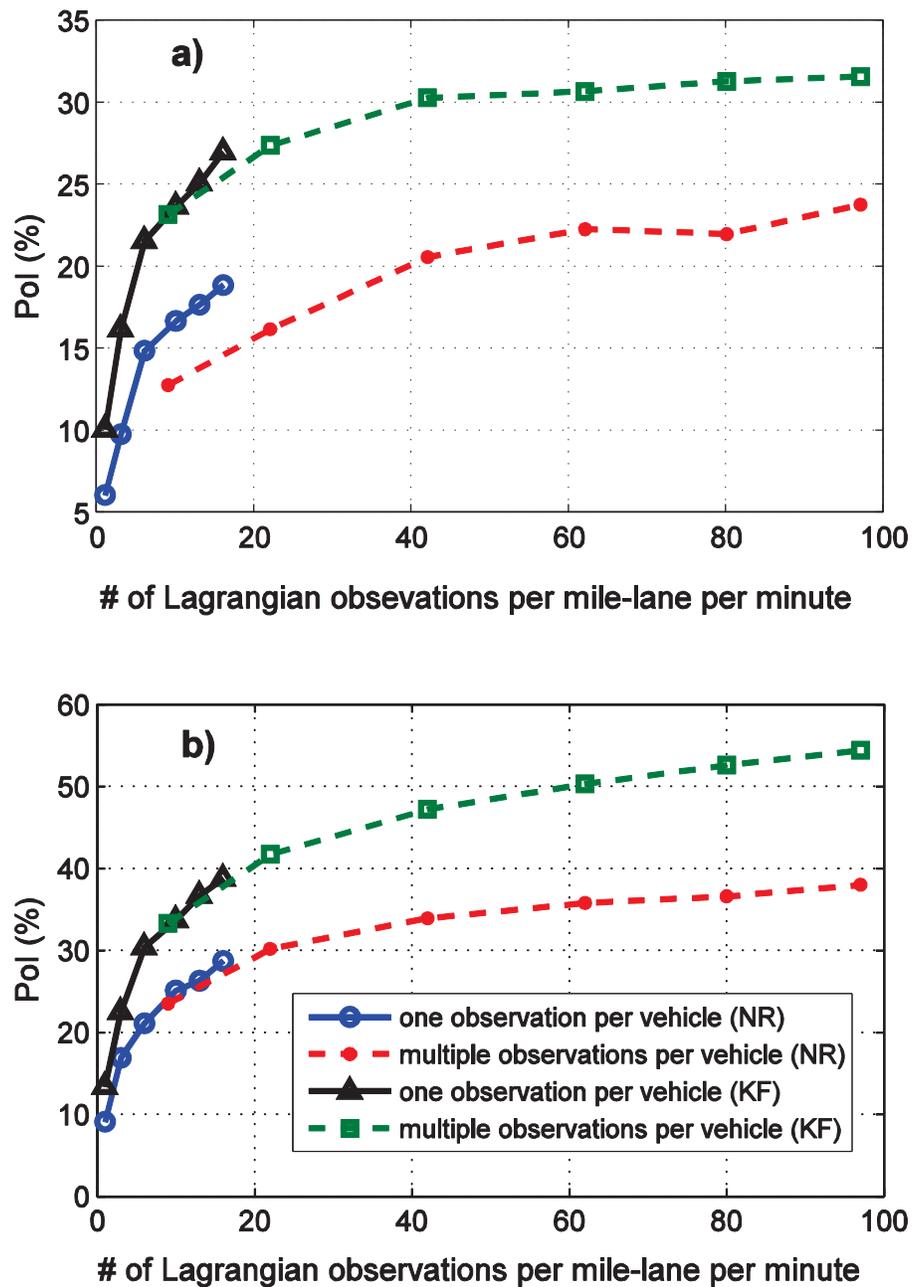


Figure 5.6 Percentage of Improvement (Pol) in the RMSE as the number of Lagrangian measurements increases: (a) computing observed density using the fundamental diagram (Section 5.2.3), (b) using the actual density computed from vehicle trajectories as the observed density. Note that the two graphs are at different scales.

Ideally, the observed density computed using the observed velocity and the fundamental diagram (Section 5.2.3) should perfectly match the ground truth density. As expected, it is not the case in practice. Figure 5.7 shows the true density (computed from vehicle trajectories) vs. the observed density (computed according to Section 5.2.3) for scenarios 3 and 9 (the trend is very similar for the other scenarios). This figure reveals the differences between the true and the observed density, suggesting that (i) the speed reported (and used to compute the observed density) is not representative of the actual state on the cell, and/or (ii) in congestion, a given density can be achieved at different velocities (or same velocity yields different densities). We believe that the main source of error is related with (ii). Indeed, it is well known that the congested branch of the fundamental diagram is not a line of points but a cloud of points for non-stationary traffic [20].

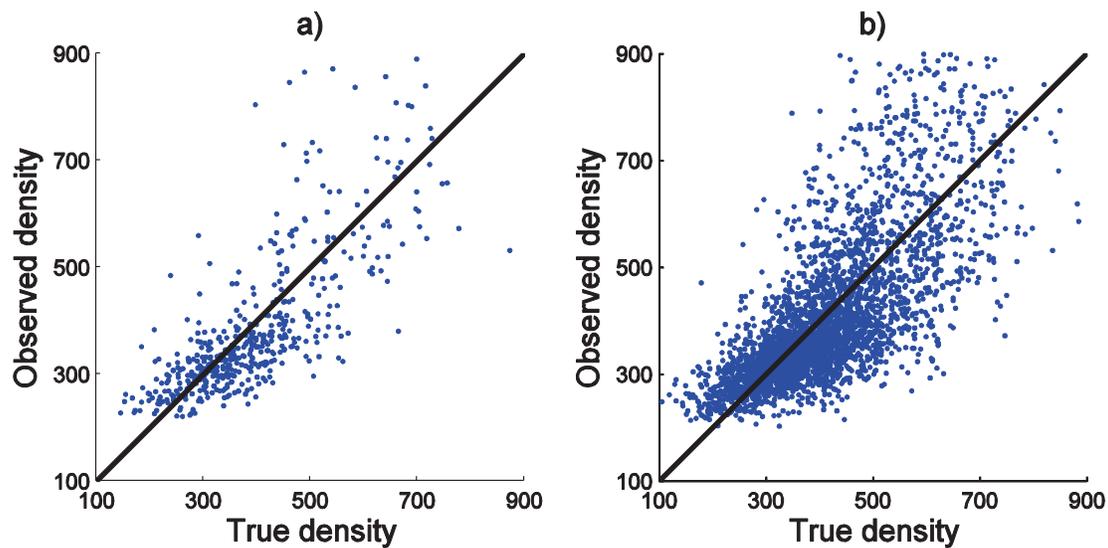


Figure 5.7 *True density* (computed from vehicle trajectories) versus *Observed density* (computed using the fundamental diagram as described in Section 5.2.3) for scenario 3 (left) and 9 (right).

To determine how the error in the observed density affects the performance of the methods, all scenarios were investigated assuming that the density computed from the Lagrangian observations was the actual density. That is, the observations are perfect (error-free). The improvement of the RMSE with respect to the EDO case is shown in Figure 5.6b. As expected, the accuracy of the estimates improves, achieving up to 40% (NR) and 50% (KF) of total improvement compared with the EDO case. The gap between the solid and the dotted line observed in Figure 5.6a for the NR method is not observed in part b of the figure, suggesting

that when the Lagrangian observations report the correct density, the sampling strategy does not really matter.

Statistical approaches such as the ones mentioned in Section 5.1 can be used to compute the velocity field. Using the fundamental diagram, the velocity field can be converted into the density field. This process is useful to compare the results obtained with the NR method and the results obtained with a simple statistical approach – in this case, the method based on the approach proposed by [100]. In terms of the RMSE, only scenarios 11 and 12 show an improvement compared to the EDO case (3.4% and 5.7%, respectively). The rest of the scenarios yield a RMSE higher than the one obtained for the EDO case. That is, for this set of data, a method based on a traffic flow model produces more accurate estimates of the vehicle accumulation than a simple statistical method. This is not surprising since the statistical method is meant to estimate velocity and not vehicle accumulation.

The main difference between the estimates from the EDO scenario and from scenarios using Lagrangian measurements was observed during the periods in which shockwaves emanate from intermediate points. This result is insightful regarding the benefits of using Lagrangian measurements. By collecting data from individual vehicles at different times and locations, we were able to capture the shockwave generated in the middle of the section (between detectors). Therefore, in the presence of Lagrangian data, inter-detector spacing could be increased. Unfortunately, the NGSIM section is too short to test this statement.

5.5 Numerical example of the NR nudging factor

This section provides a numerical example on the effect of the nudging factor. For illustration purposes, assume the freeway section has been divided into cells of length $\Delta x = 0.02$ miles and time into time steps of duration $\Delta t = 2$ seconds. Let us assume that the estimated density using the model at cell i at time step h is $\hat{k}_i^h = 70$. Assume also, observation is available at the same location and at the same time, $k_i^{o,h} = 100$. Since $\hat{k}_i^h < k_i^{o,h}$, vehicles will be added, and the nudging factor will determine how many.

Let us also assume that $X_{\text{nudge}} = 0.04$ miles and $T_d = 4$ seconds. That is, observation from cell i at time step h affects two cells around cell i during the next two time steps (i.e. all the cells between $i - 2$ and $i + 2$, both inclusive, for time steps $h + 1$ and $h + 2$). Table 5.3 shows the values taken by the nudging factor λ for the affected cells during the next two time steps (assuming $T_a = 10$ seconds). It can be seen how the effect of the observation from cell i at time step h reaches its maximum at $\lambda(0, \Delta t)$, and decreases with distance and time.

The total number of vehicles added because of the observation from cell i at time step h is $0.398 \cdot (k_i^{o,h} - \hat{k}_i^h) = 11.9$ vehicles. According to equation (5-12) different values for X_{nudge} , T_d ,

and T_a would affect neighboring cells in a different way, adding a different number of vehicles in the end. Remember that other observations are also adding (removing) vehicles to (from) the affected cells during the next two time steps.

Table 5.3: Values taken by the nudging factor for a specific case.

		Time steps affected		Σ
		$h + 1$	$h + 2$	
Cells affected	$i - 1$	0.039	0.024	0.063
	$i - 2$	0.054	0.033	0.087
	i	0.061	0.037	0.098
	$i + 1$	0.054	0.033	0.087
	$i + 2$	0.039	0.024	0.063
	Σ	0.247	0.151	0.398

6 Velocity-based Modeling

This chapter introduces an alternative to the data assimilation methods (i.e., the reconstruction of the state of a system using data [36]) presented in Chapter 5. The crucial difference is that in this alternative method, Lagrangian data (velocity measurements from GPS-enabled smartphones) is directly fed to a velocity-based flow model (instead of a more traditional density-based model). This approach (implemented for the real time system deployed during the 100-vehicle Mobile Century experiment) obviates the difficulties associated with converting velocity measurements into density measurements as discussed in Section 5.4.2.

To reiterate briefly the main difficulty, consider that although smartphone-based sensing provides accurate position and speed information (as shown in Chapter 4) from a sample population of equipped vehicles, density cannot be measured and cannot be extrapolated easily (see Section 5.4.2). As a result, traditional highway traffic theory such as the Lighthill-Whitham-Richards (LWR) partial differential equation (PDE) [79, 98] and related density based algorithms such as [48] cannot be used as such for data assimilation. The principle objective of this chapter is to describe a new inverse modeling technique specifically designed to use velocity measurements as inputs. These velocity measurements are incorporated into a flow model for velocity, and used to produce an optimal estimation of travel time.

This chapter is organized as follows. Section 6.1 furnishes a brief review of related work in the context of the problem at hand. In Section 6.2, we propose a new velocity flow model for highways based on the LWR PDE and discretize it using a Godunov scheme. We detail the EnKF algorithm which enables us to maintain the nonlinearities of the velocity CTM- v ¹⁶ in section 6.3. Finally, in section 6.4, this new data assimilation algorithm for highway velocity field reconstruction is implemented in microsimulation.

6.1 Related Work

Kalman Filtering (KF) has been widely used for traffic state estimation in earlier studies in its various forms. In [106, 107], Mixture Kalman Filtering (MKF) was applied to the Cell Transmission Model (CTM) to estimate traffic densities for ramp metering. The nonlinear CTM was transformed into a switching state space model, which enabled the use of a set of linear equations to describe the state evolution for the distinct flow regimes on the highway (e.g. highway is in free-flow or congestion). In [51], a Kalman Filter was used to incorporate Lagrangian velocity trajectories into a density based CTM for highway traffic. A real time algorithm for traffic estimation based on the Extended Kalman Filter (EKF) using second order PDE as a flow model was used in [114]. Other treatments of traffic monitoring include adjoint based data assimilation in [61, 62], and particle filtering in [101].

¹⁶ The nomenclature “CTM- v ” emphasizes the fact that the method described here is a translation of the classic CTM scheme into the velocity domain.

A common factor for the CTM based methods [51, 106, 107] described above is that the evolution of traffic state (typically density, not velocity) relies on a set of linearized equations which are needed in order to use the KF or EKF techniques. However, the approach proposed in this work employs Ensemble Kalman Filtering (EnKF) which enables the use of fully nonlinear equations such as the discretization of the new flow model proposed in this article. Using Monte Carlo integrations, it can maintain the nonlinear features of error statistics that can be lost if using linear approximations. Furthermore, by employing a fully nonlinear model, no highway-mode-selection-algorithms or simplifications to the equations are needed in this work.

6.2 Highway Traffic Flow Model

6.2.1 Speed Evolution Equation

To address the problem of reconstruction of the velocity field on the highway, we introduce a new first order hyperbolic PDE similar to the LWR PDE, which models the evolution of speed on the highway. This PDE can be shown to be equivalent to the kinematic wave theory for the Greenshields flux function [18]. This PDE is referred to as LWR-v PDE (“v” for velocity). Proper weak boundary conditions are defined to formulate the well posedness of an initial boundary value problem.

The seminal LWR equation proposed in [78] and [98] to model traffic on highways reads:

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0 \quad (6-1)$$

where $q(x,t)$ and $\rho(x,t)$ respectively denote the flow of vehicles and their density at location x and time t . Additionally we denote $v(x,t)$ the velocity field on the highway. Equation (6-1) is derived from hydrodynamics theory and expresses the conservation of mass for a fluid of density ρ and of flux q and is considered relevant to model traffic on a highway (see [27, 26], for more background).

In order to express the flow q as a function of the density ρ traffic theory uses an empirical relation called the fundamental diagram:

$$q(x,t) = Q(\rho(x,t)) \quad (6-2)$$

where Q is the flux function, which is assumed to be independent from time and space.

One of the seminal flux functions used is the Greenshields flux function [45] which expresses a linear relation between ρ and v as:

$$v = v_{\max} \left(1 - \frac{\rho}{\rho_{\max}} \right) \quad (6-3)$$

where v_{\max} and ρ_{\max} denote respectively the maximal velocity and the maximal density allowed by the model.

When the flux function is a Greenshields flux function, it is possible to invert the speed-density function and express ρ as a function of v , namely $\rho = \rho_{\max} \left(1 - \frac{v}{v_{\max}} \right)$. Thus inserting this expression of ρ in (6-1), one can rewrite the LWR PDE on ρ as a LWR-v PDE:

$$\frac{\partial v}{\partial t} + \frac{\partial}{\partial x} (R(v)) = 0 \quad (6-4)$$

where $R(v) = (v)^2 - v_{\max} v$, using the notation $(v)^2 := vv$ to avoid confusion with discretization indices introduced later.

The simple variable change $v = v - \frac{v_{\max}}{2}$ transforms equation (6-4) into:

$$\frac{\partial v}{\partial t} + \frac{\partial}{\partial x} (v)^2 = 0 \quad (6-5)$$

on the domain $(x, t) \in]a, b[\times]0, T[$. We recognize a Burgers equation with a factor $\frac{1}{2}$ omitted (see [33]). The initial condition and boundary conditions in weak form read:

$$\forall x \in]a, b[\quad v(x, 0) = v_0(x) \quad (6-6)$$

and

$$\begin{cases} v(a, t) = v_a(t) \text{ or} \\ v(a, t) \leq 0 \text{ and } (v(a, t))^2 \geq (v_a(t))^2 \end{cases} \quad (6-7)$$

and

$$\begin{cases} v(b, t) = v_b(t) \text{ or} \\ v(b, t) \geq 0 \text{ and } (v(b, t))^2 \geq (v_b(t))^2 \end{cases} \quad (6-8)$$

where $v_a(t)$ and $v_b(t)$ denote the boundary conditions that are applied but which are not always active, as described by the above equations. As demonstrated in [12], the PDE (6-5) with the initial condition (6-6) and the weak boundary conditions (6-7) (6-8) admits a unique entropy solution in the space $BV([a, b] \times]0, T[)$.

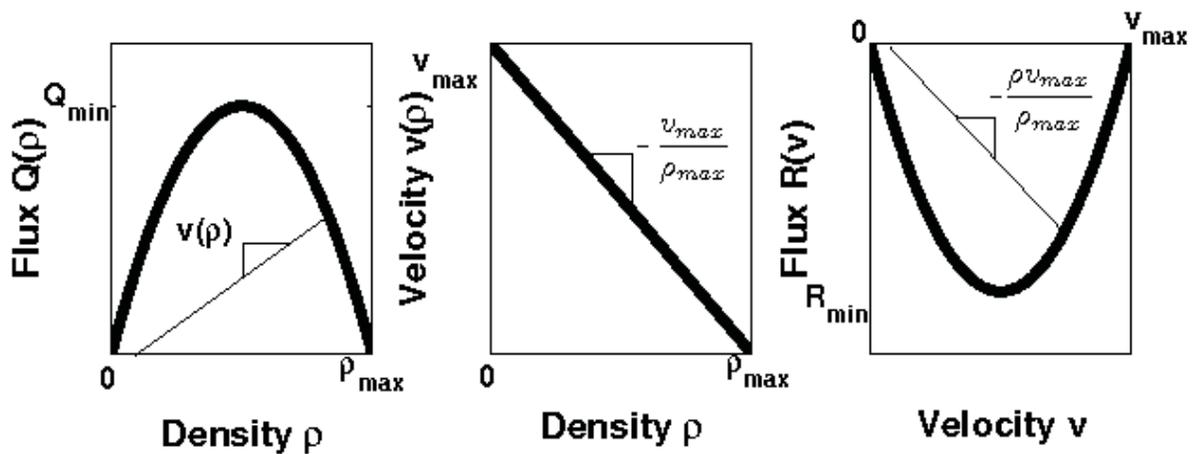


Figure 6.1: Greenshields model. Left: Classical fundamental diagram (parabolic). Center: Linear relation between speed and density. Right: Flux function for the LWR-v PDE (6-4). The flux is parabolic with negative values.

The fundamental properties of the LWR PDE (6-1) are conserved in the LWR-v PDE (6-4). First, the speed of a characteristic, given for a state (ρ_0, v_0) by the derivative of the flux function in (ρ_0, v_0) is the same for both PDEs. Indeed, for the LWR PDE it is:

$$Q'(\rho_0) = v_{\max} - 2\rho_0 \frac{v_{\max}}{\rho_{\max}} \quad (6-9)$$

whereas for the LWR-v PDE it is:

$$R'(v_0) = 2v_0 - v_{\max} \quad (6-10)$$

and these two expressions are equivalent given relation (6-3). Second, the Rankine-Hugoniot relation giving the speed of shocks for conservation laws is also preserved. Indeed, the speed of a shock between a state (ρ_1, v_1) and a state (ρ_2, v_2) is given for the LWR PDE by:

$$\frac{Q(\rho_1) - Q(\rho_2)}{\rho_1 - \rho_2} = v_{\max} - \frac{v_{\max}}{\rho_{\max}}(\rho_1 + \rho_2) \quad (6-11)$$

whereas for the LWR-v PDE it is given by:

$$\frac{R(v_1) - R(v_2)}{v_1 - v_2} = (v_1 + v_2) - v_{\max} \quad (6-12)$$

Given the speed-density relation (6-3) expressed by the Greenshields model, these two expressions can also be checked to give the same speed for a shock.

6.2.2 Numerical Discretization

For practical implementation, the LWR-v PDE is discretized using a Godunov numerical scheme to obtain a velocity cell transmission model (CTM-v) [27, 26, 42]. In this section we detail the use of the Godunov scheme (see [42, 77]) for equation (6-5). This scheme is known to be convergent for convex and concave flux functions [42, 77, 105] such as the Greenshields flux function.

$N, M \in \mathbb{Z}$ the set of integers, we discretize time and space in N time steps J_n ($0 \leq n \leq N$) of length $\Delta t = \frac{T}{N}$ and M space cells I_i ($0 \leq i \leq M$) of length $\Delta x = \frac{b-a}{M}$. We call v_i^n the discrete value of v on $I_i \times J_n$. According to the Godunov scheme, at each time step v_i^{n+1} is computed from the previous time step by the following formula:

$$v_i^{n+1} = v_i^n - \frac{\Delta t}{\Delta x} \left(g(v_i^n, v_{i+1}^n) - g(v_{i-1}^n, v_i^n) \right) \quad (6-13)$$

where the numerical flow g is defined, as follows:

$$g(v_1, v_2) = \begin{cases} (v_2)^2 & \text{if } v_1 \leq v_2 \leq v_c \\ (v_c)^2 & \text{if } v_1 \leq v_c \leq v_2 \\ (v_1)^2 & \text{if } v_c \leq v_1 \leq v_2 \\ \max((v_1)^2, (v_2)^2) & \text{if } v_1 \geq v_2. \end{cases} \quad (6-14)$$

with v_c defined to be the minimum of the convex flux function from equation (6-5) (For example: $v_c = 0$ for $R(v) = v^2$, and $v_c = \frac{v_{\max}}{2}$ for $R(v) = (v)^2 - v_{\max}v$). For stability of the numerical discretization, the spatial and temporal step sizes must obey the Courant-Friedrichs-Lewy (CFL) condition:

$$\left| \frac{\Delta t}{\Delta x} v_{\max} \right| \leq 1 \quad (6-15)$$

In order to implement the weak boundary conditions defined in the previous section, we use ghost cells placed at each side of the domain defined by the strong boundary conditions we would like to be satisfied, namely:

$$v_{-1}^n = \frac{1}{\Delta t} \int_{J_n} v_a(t) dt \text{ and } v_{M+1}^n = \frac{1}{\Delta t} \int_{J_n} v_b(t) dt \quad (6-16)$$

with $J_n = \left[n \frac{T}{N}, (n+1) \frac{T}{N} \right]$. The choice of the Godunov scheme to solve a first order scalar hyperbolic conservation law such as (6-5) is standard in literature. Note that equations (6-13), (6-14) and (6-16) could also be viewed as a counterpart of the cell transmission model for speed.

This model is thus a nonlinear dynamical system, in which the state of the system $v^n = [v_0^n, v_2^n, \dots, v_M^n]$ is the vector of velocities in all cells at time step n . Letting M represent the nonlinear discrete time dynamical system model equations (6-13) (6-14) for the full state vector v^n , and for the state noise η^n , the state dynamics of the system can be written as:

$$v^{n+1} = M[v^n] + \eta^n. \quad (6-17)$$

Here the state noise η^n represents the modeling error introduced by discretization and uncertain boundary conditions. For more information on modeling errors and state noise covariance estimation, see for example [66, 103].

The observation equation can be written as follows:

$$y^n = \mathbf{H}^n[v^n] + \varepsilon^n = \mathbf{H}^n v^n + \varepsilon^n \quad (6-18)$$

where $\mathbf{H}^n \in \{0,1\}^{p^n \times M}$ encodes the p^n discrete cells on the highway for which the velocity is observed during discrete time step n , and ε^n is the Gaussian observation noise with covariance \mathbf{R}^n associated with the observation. In the event no equipped vehicles are in the spatial domain during a particular timestep, \mathbf{H}^n reduces to the zero matrix.

6.3 Speed Estimation

6.3.1 Ensemble Kalman Filter

A common approach to solving the inverse modeling problem for linear time invariant (LTI) systems is to implement a KF algorithm [68], which is particularly well suited for real-time algorithms because of its recursive nature. Only the previous state of the system is needed to optimally integrate new measurements in the Minimum Mean Square Error (MMSE) sense. Due to the nonlinearity of the CTM-v, the standard KF cannot be used. Furthermore, due to the nonlinearity in the flux function of CTM-v the EKF has limited applicability. Hence, we extend this framework by implementing an Ensemble Kalman Filter (EnKF) algorithm [35], first introduced in [34].

EnKF is a sequential data assimilation method, which uses Monte Carlo or ensemble integrations. By integrating an ensemble of model states forward in time, it is possible to compute the mean and error covariances needed at analysis times (measurement update) [19, 50]. The analysis scheme in the EnKF uses traditional update equations of the KF, except that the Kalman gain is computed using the error covariances provided by the ensemble of model states.

The nonlinearities introduced by the CTM-v equation are captured well by the EnKF because of the sample based computation of covariance matrices in contrast to tangent linear models (Jacobian matrix) used in EKF.

The EnKF algorithm can be broken into three phases [50]:

First, we generate an ensemble of model states by drawing K samples $\xi_k^0 \in \mathbb{R}^{M+1}$, $k = 1, \dots, K$, from a Gaussian prior distribution to initialize the algorithm. These samples represent our prior knowledge of the initial velocity field v^0 on the highway. We assume that the initial velocity profile on the highway is smooth, without shocks. The construction of a smooth prior is made using a framework proposed in [67]. With this approach we can generate initial model states which maintain the same correlation properties regardless of the discretization level (number of cells) of the CTM-v model.

Next, a prediction $\hat{\xi}_k^n$ of ξ_k^n is made from the CTM-v model:

$$\hat{\xi}_k^n = M[\xi_k^{n-1}] + \eta_k^{n-1} \quad (6-19)$$

We then compute the mean of the ensembles:

$$v^n = \frac{1}{K} \sum_{k=1}^K \hat{\xi}_k^n \quad (6-20)$$

from which the covariance \mathbf{P}^n of the predicted state can be computed as:

$$\mathbf{P}^n = \frac{1}{K-1} \mathbf{E}^n (\mathbf{E}^n)^T \quad (6-21)$$

where matrix \mathbf{E}^n is defined as:

$$\mathbf{E}^n = [\hat{\xi}_1^n - v^n, \dots, \hat{\xi}_K^n - v^n]. \quad (6-22)$$

Next we compute the Kalman gain \mathbf{G}^n :

$$\mathbf{G}^n = \mathbf{P}^n (\mathbf{H}^n)^T [\mathbf{H}^n \mathbf{P}^n (\mathbf{H}^n)^T + \mathbf{R}^n]^{-1}. \quad (6-23)$$

Finally, the ensemble ξ_k^n is updated with new measurements y^n as follows:

$$\xi_k^n = \hat{\xi}_k^n + \mathbf{G}^n [y^n - \mathbf{H}^n \hat{\xi}_k^n + \varepsilon_k^n]. \quad (6-24)$$

The presence of ε^n in (24) is important both from a physical interpretation, as well as for the stability of the convergence of the EnKF routine. When the state of the highway is observed from GPS measurements, ε^n accounts for the GPS position and speed error. During field testing with the Nokia N95 mobile device, a mean velocity error of 3 mph has been observed giving approximate lane level position accuracy. From an algorithmic viewpoint, the random error is shown in [19] to be necessary to maintain sufficient variance in the ensemble and to prevent filter divergence.

6.4 Implementation and Validation

6.4.1 Paramics Microsimulation

We analyze the performance of the EnKF CTM-v algorithm using the Paramics Microscopic Traffic Simulation software [95], calibrated for highway I-880 south of Oakland, CA. The I-880 calibrated model produces individual trajectories on the highway for each vehicle in the simulation, and it has previously been used for bottleneck identification in [11]. For this experiment, a subset of the vehicles are randomly selected as vehicles which are equipped with GPS phones. The percentage of equipped vehicles relative to the total traffic flow is known as the penetration rate. Ten VTLs are placed evenly between Industrial Parkway (milepost 24.917) and Tennyson Rd. (milepost 25.767) which cause equipped northbound vehicles to report speeds and positions after crossing the VTL.

The highway is discretized into ten spatial cells, and we take a timestep of two seconds in order to maintain stability in the Godunov numerical scheme (6-15). A maximum speed of 70 mph is assumed for the Greenshields equivalent velocity flux function. The simulation is run for two hours, over which time congestion increases and the speed of flow decreases. Just after 3:30 pm in the simulation, the four-lane averaged speeds decrease from the 65-70 mph free flow speed to speeds ranging between 20-40 mph. The congestion and corresponding slowdown is captured in the ground truth velocity contour shown at the top of Figure 6.2.

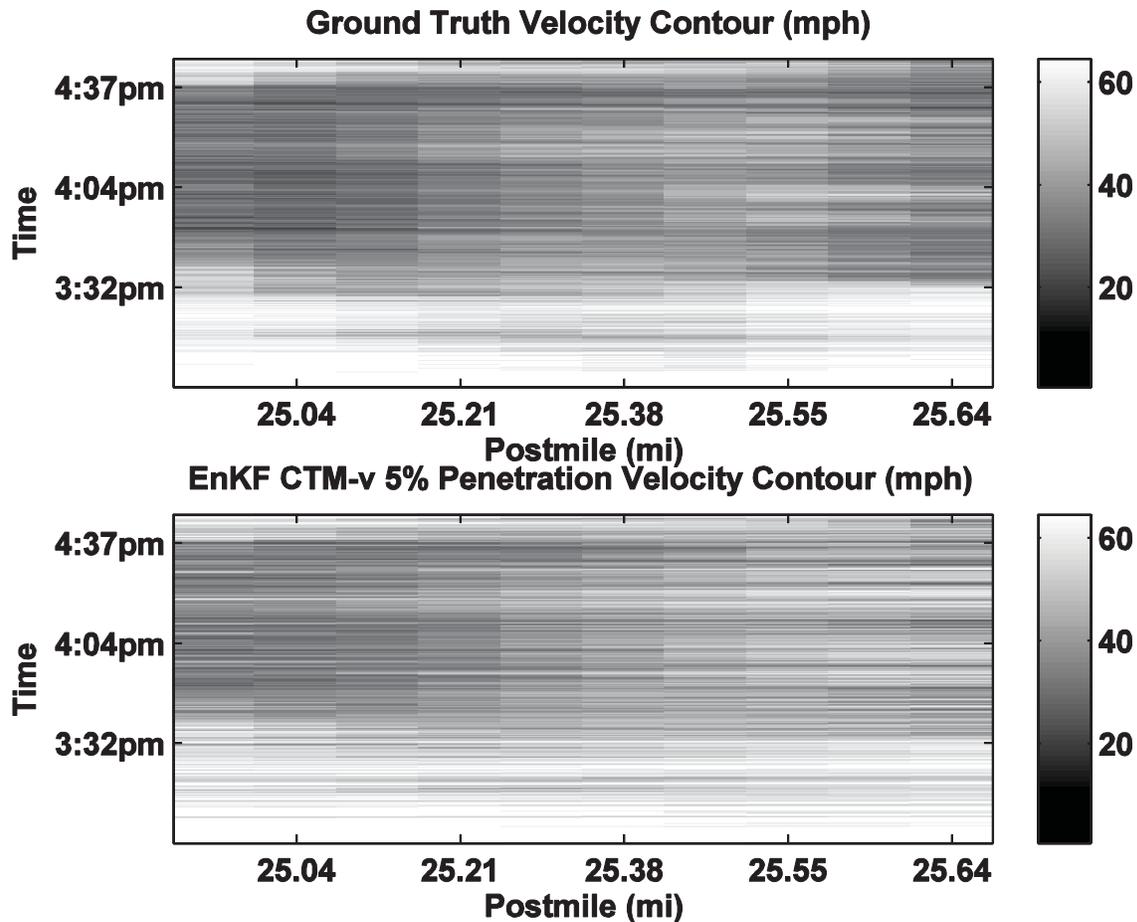


Figure 6.2: Paramics velocity contours. Top: Ground truth velocity contour average across all vehicles. Bottom: Estimated velocity contour from the EnKF CTM-v algorithm (6-19) through (6-24) at 5% penetration rate. X-axis: position along highway in milepost; Y-axis: time of day.

In order to evaluate the performance of the EnKF CTM-v algorithm, a simple averaging-based estimation scheme is introduced. For this scheme, the velocity v_i^n in the discrete cell $I_i \times J_n$ is computed from the average of all measurements observed $(v_{\text{obs}})_i^n$ in each discrete cell as:

$$v_i^n = \begin{cases} v_i^{n-1} & \text{if } v_{obs} = \emptyset \\ (v_{obs})_i^n & \text{otherwise} \end{cases} \quad (6-25)$$

The mean point-wise \mathcal{L}_1 relative error v_{re} between the estimated velocity v_{est} and the ground truth velocity v_{gt} is computed by:

$$v_{re} = \frac{1}{MN} \sum_{i=0}^M \sum_{n=0}^N \frac{\left| (v_{est})_i^n - (v_{gt})_i^n \right|}{(v_{gt})_i^n} \quad (6-26)$$

and the mean point-wise \mathcal{L}_1 absolute error v_{ae} of the discrete density field is computed as:

$$v_{ae} = \frac{1}{MN} \sum_{i=0}^M \sum_{n=0}^N \left| (v_{est})_i^n - (v_{gt})_i^n \right|. \quad (6-27)$$

In Figure 6.2 the plot of the estimate from the EnKF relative to the ground truth shows that the main features of the shock wave are captured, even with the relatively low penetration rate of 5%. It is worth noting that the ground-truth highway exhibits lane shearing, where vehicles in each lane have different mean speeds. The result is that vehicles sampled from the same discrete space and time cell, but from different lanes may have significant variance relative to the lane-averaged mean speed. As the penetration rate increases (see Figure 6.3), the sampled vehicles become more representative of the flows on each lane, and thus more accurately predict the lane-averaged mean speed.

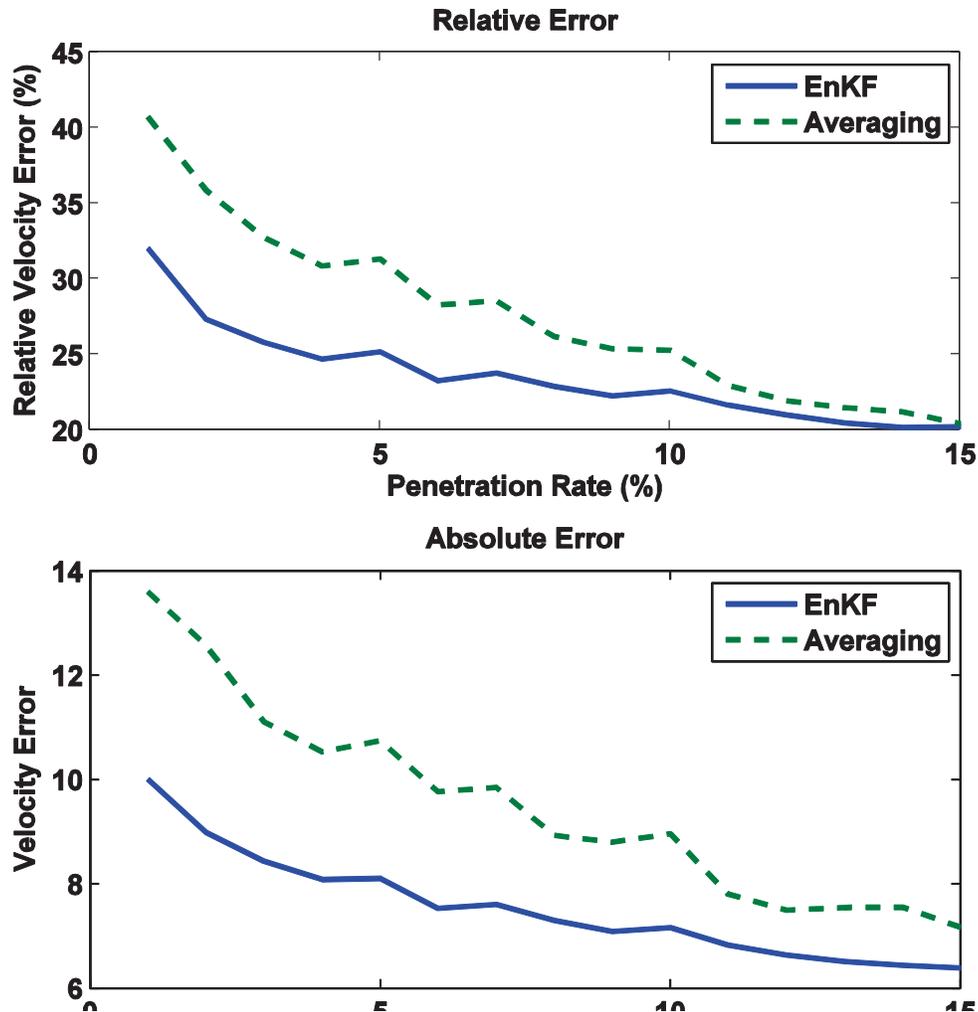


Figure 6.3: Error comparison of the EnKF CTM-v scheme, equations (6-19) through (6-24), (solid) and the averaging scheme (6-25), (dashed) using Paramics. Top: Relative error computed from (6-26) as a function of penetration rate. Bottom: Absolute error computed from (6-27) as a function of penetration rate.

The added value of the EnKF CTM-v algorithm is shown in Figure 6.3, relative to a simple averaging estimate based on tracking. For this comparison, the complete trajectories of the equipped vehicles are observed and used in the simple averaging scheme (6-25), representing a privacy intrusive method in which the complete vehicle path is known. Alternatively, the EnKF CTM-v algorithm only integrates the velocities observed as a result of the equipped vehicle crossing the VTL. Even by assimilating fewer data, the EnKF CTM-v estimate has less error than the averaging scheme. At low penetration rates, the EnKF CTM-v algorithm reduces the relative error by 8%, or three miles per hour, and as the penetration rate increases, the simple averaging estimate converges towards the EnKF CTM-v estimate. Although the relative error remains large at low penetration rates, the error occurs in the congested regime where the ground truth speed is slower and absolute errors are magnified.

6.5 Conclusion and Future Work

This work presents a method for assimilating GPS speed and position data into a new velocity model derived from the LWR PDE. By working directly with the velocity PDE, conversions to density for data assimilation and back to velocity for travel time computations are eliminated. At low penetration rates, the method implemented using the VTL framework outperforms trajectory averaging based on tracking, despite using fewer data. Given that all data must be transmitted across a cellular network, optimal data assimilation methods will be increasingly important to efficiently use limited data streams. Furthermore, the recursive structure of the EnKF method is well suited for real-time applications

In addition to improving estimates at low penetration rates, the EnKF CTM-v algorithm has additional features which we intend to highlight in future work. The framework can be run forward in time to produce forecasts of the traffic state, in addition the current state estimates presented here. This will be important to compute dynamic travel times that account for changes in the traffic state as the vehicle travels on the highway.

Finally, as historic data is collected, the model accuracy can be improved, by computing more accurate state noise and observation noise covariance matrices used to model the system dynamics. In a way similar to the accumulation of historical data for the PeMS system, the availability of training data will become greater as does penetration rates of GPS-equipped cellular phones on the highway.

7 Video validation

This chapter details the video validation effort that occurred in parallel with the planning, execution, and evaluation of the 100-probe vehicle deployment described in the following chapters. During the initial planning stages, it was clear that ground truth trip times would be necessary to evaluate the quality of data received from the cell phones, and also to evaluate the quality of the algorithms used to estimate traffic state. Video cameras were ultimately deployed with the intent of re-identifying license plates (vehicles) during the experiment. The successful deployment of cameras resulted in hours of video data that were manually processed to re-identify about 20% of the total vehicle flow, thus providing the required ground-truth travel times.

This chapter is organized as follows. In Section 7.1, resolution requirements are determined and a selection of standard off-the-shelf cameras and camcorders are evaluated. Rationale for the final choice of HD camcorders is provided in Section 7.2. Practical considerations such as data capacity and battery life are taken into account in Section 7.3. Trial tests during a range of weather and lighting conditions were run to inform the protocol for the actual experiment, as described in Section 7.4. The protocol for camera deployment is furnished in Section 7.5. A brief narrative of the deployment is told in Section 7.6. Finally, in Section 7.7, the post-processing of the video data is explained.

7.1 Resolution Requirements and Camera Capabilities

Requirements. First, it was necessary to evaluate the requirements for cameras capable of capturing images in sufficient detail to recover license plate numbers. Several factors were considered, including the resolution limitations from the number of pixels, blurring due to longitudinal and lateral motion, and luminosity effects.



Figure 7.1: Source image for resolution analysis.

Resolution based on number of pixels: A typical lane is about 5 m wide, and a license plate is approximately 15 x 30 cm. If a camera is set up to view only one lane, then the 5-m width of

the lane will map to the horizontal width of the camera field of view. The observational capabilities for standard off-the-shelf components are simulated in Figure 7.2 for the one-lane configuration. For the case where a camera is configured to view all five lanes, then the resolution is reduced by a factor of five. For this latter case, the simulated result is displayed in Figure 7.3. The results summarized in Table 7.1 represent best-case scenarios for unmoving vehicles, and do not take into account any other factors. From the figures, it is clear that sub-centimeter resolution is required to ensure the legibility of license plates.



Figure 7.2: Simulated resolution for three camera types, assuming one camera per lane.

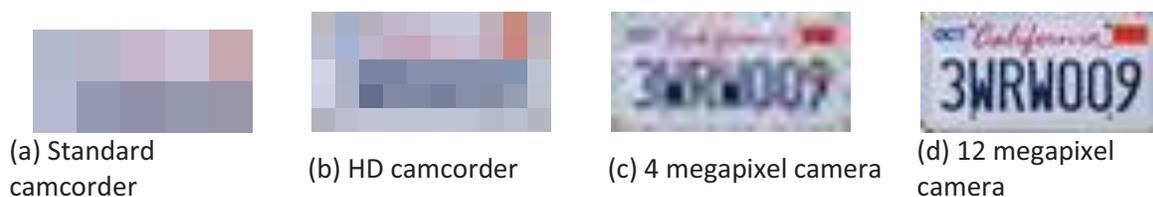


Figure 7.3: Simulated resolution for four camera types, assuming one camera for five lanes.

Table 7.1: Resolution capability for a menu of cameras and configurations

Type	Pixel width	One-Lane Configuration Resolution	Five-Lane Configuration Resolution
Standard camcorder	600	1.0 cm	5.0 cm
HD camcorder	1080	0.5 cm	2.5 cm
4 megapixel camera	2500	0.2 cm	1.0 cm
12 megapixel camera	4000	0.1 cm	0.5 cm

Longitudinal Motion and Depth of Field. The depth-of-field does not pose a problem for digital camcorders, because the resolution is low and the frame-rate is high. Each vehicle will appear in focus in several consecutive frames. However still-frame, high-resolution cameras must be used in burst mode. The resolution is very high and the frame-rate is very low. Therefore, there is a danger that a vehicle may pass through the plane of focus in between frames. To

avoid this, the depth of field must be greater than 10 m (assuming the maximum vehicle speed is about 30 m/s, and the frame-rate is 3 images per second). This requirement for depth of field will determine a maximum aperture size, or equivalently, a minimal value for the aperture number (f-stop).

Blurring Due to Lateral Motion. The shutter speed (exposure time) must be set to minimize blur from lateral motion. Assuming the maximum lateral speed of a car is on the order of 5 m/s (7m/s for a brutal lane change with 0.5 g lateral acceleration), the exposure time for a standard camcorder must be no more than 1/500 s simply to maintain the 1.0-cm resolution for the one-lane configuration in Table 7.1. However, the shutter speed for a typical camcorder ranges from 1/60 s to 1/250 s; and, therefore blurring from motion would reduce resolution to about 2.0 cm as shown in Figure 7.4. This issue can be problematic for high resolution camcorders as well. The only option is to have a high-speed camcorder with selectable shutter speeds greater than (exposure times less than) 1/1000 s. There should be no problems of lateral speed blur for a still-frame digital camera, because exposure times can be set to 1/2500 s or less. For our application, the resolution of a 12-mega pixel camera capturing five lanes would be limited by the number of pixels, not the exposure time.



Figure 7.4: Image of license plate with blurring to simulate 2.0 cm resolution

Luminosity analysis in Low-light Conditions. Fog or rain scatters light; the reduction in resolution depends on the intensity of the rain, and must be checked empirically. However, back-of-the envelope calculations allow a discussion of several rules of thumb. During low-light conditions such as rain, the exposure time must increase (in order to keep the noise levels down), but this in turn risks forward and lateral speed blur. A rough estimate of the required shutter speed for a full aperture during cloudy conditions is 1/500 s. This imposes a constraint in which a lens with a longer focal length and narrower field of view are required (a telephoto lens). For this application however, filming was to take place from a low angle. At high zoom, there is a danger that tall vehicles will occlude the license plates of following vehicles. Perhaps the only way to address this problem is to choose a camcorder with a high S/N ratio.

Luminosity analysis in Bright-light Conditions. On a clear, sunny day, the danger is that the contrast of the license plate can be too high (for example if there are sun reflections on the license plate). This limitation will in turn define a feasible time of day in which the experiment can take place. It is also advisable to have a cover for preventing excessive sunlight to enter the camera if the sun is low in the sky or if the camera is pointed in a direction such that glare from reflections is intense.

7.2 Final Selection of Video Cameras

Standard camcorder. If one standard camcorder is devoted to each lane, and each camcorder is focused on a 2-m wide area corresponding to the average width of a car, then a standard camcorder is capable of providing the required image quality. However, vehicles changing lanes, or somehow failing to drive near the center of the lane would not be identified. A standard camcorder configured to capture the entire lane would yield 1-cm resolution in best-case conditions, which is still not enough to read a license plate number. For this reason, standard camcorders were disqualified for this application.

Still-frame camera. The problem of reading license plate numbers on multiple lanes is much more challenging, even with a high-quality, still-frame camera in burst mode. One problem with still-frame cameras is that in burst mode, the sampling rate cannot be controlled precisely. There is another practical issue with regards to the data rate (and amount of data); this will be discussed later. For now, consider a 10 Megapixel camera with low focal length, and in burst mode (for instance in B/W and high JPEG compression to minimize the data transfer rate). On a typical SLR, this data rate is around 50 Mbits per second. Over a multi-hour experiment, the amount of data would be enormous. In addition, the data rate itself far exceeds USB data-transfer capabilities.

HD camcorder. Ultimately, HD camcorders affixed to tripods were chosen as the recording device. One HD camcorder was used to record 2-3 lanes of traffic. Specifically, the Canon VIXIA HV20 was chosen for being the most cost effective solution at the desired resolution.



Simulated resolution using a standard camcorder



Simulated view using a HD camcorder:

Figure 7.5: Comparison of standard vs. HD camcorder image quality

7.3 Practical Considerations for Camera Deployment

Practical Considerations. Once having chosen HD camcorders, there were several practical problems to overcome: (1) limited data storage capacity of the cassette, (2) battery life of only two hours, and (3) synchronization of cameras for accurate travel time estimates.

Capacity. The capacity problem was solved by saving the video data directly to the hard drive of a laptop computer. The freely downloadable HDVSplit program¹⁷ was employed for this purpose. This program was chosen because it was free, lightweight, and enabled the streaming of images from camera to computer via Firewire with hardly any frame-drops. In this configuration, there was about 10 GB per hour of filming.

Battery life. Power inverters were used to supply power to both the PCs and the HD camcorders. Each surveillance team was supplied with two Xantrex inverters and one spare.

Synchronization. Another problem was to synchronize all cameras. Several steps were taken to accomplish this. First, the system time of each computer was synchronized. From the Windows operating system, the time of day when the file was started is stored (of course this time of day gets overwritten when the file is copied). For safety, the starting time of each video was recorded with pencil and paper.

7.4 Trial Tests

Informal Trial Testing. Several trial tests were performed to assess the quality of the equipment and to inform the protocol. One trial test was performed with low speed cars on an arterial on a sunny day. Another test was performed to record vehicles on a highway. There was one test that occurred during rain. This test resulted in the end of life for one computer. There was one night test, and one bad glare test. The latter test revealed that glare is a significant problem. For this reason, east-west orientations can be problematic.

Dry Run. One final test took place one week before the MC experiment with a sample setup the same as in the actual MC. Cameras were set up facing south, and with the sun high in the sky, there were no problems with glare. However, a new problem was encountered. In our application, cameras were positioned on an overpass in which the expressway could be viewed from the other side of a chain-link fence. In this configuration, the position of the fence mesh causes blurring in fixed regions of the camera field of view. It was crucial to check the position of the cameras and to make certain that the mesh was not positioned directly over the center of the lanes to be recorded.

Encroachment Permit. For both the dry run and for the field experiment, an encroachment permit was acquired to comply with CalTrans procedures.

¹⁷ <http://strony.aster.pl/paviko/hdvsplit.htm>



Figure 7.6: Video frame with fence mesh between lanes.

7.5 Protocol for Camera Deployment

This section details the deployment recommendations in the protocol for the camera teams. The figures in the section display excerpts from the Canon VIXIA HV20 instruction manual.

Recording mode. The HDV mode (with resolution of 1920x1080) was chosen for the experiment. This mode samples at 30 frames per second. An alternative for future deployments might be HDV PF24, with a sampling rate of 24 frames per second. The ultimate choice would depend on the consequences of “cinema mode” on image quality.

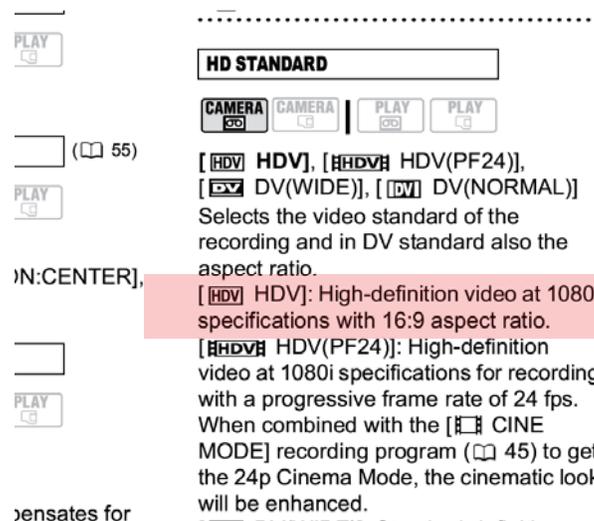


Figure 7.7: HD Camcorder recording mode.

Shutter priority. For this application it was crucial to control the shutter speed manually. Therefore the automatic shutter setting was turned off as shown in Figure 7.8. In this configuration, the camera is free to choose the aperture. The possible configurations are shown in Figure 7.9. Unfortunately, it was not possible to manually control both the shutter speed and aperture, independently of each other for the cameras used in the present work.

Procedure to set appropriate shutter speed. Camcorder operators were instructed to test video quality using a range of shutter speeds: 1/250 s, 1/500 s, 1/1000 s and 1/2000 s. The expectation was that if the day was very sunny, image detail would be best at about 1/2000 s.

Shutter Speed Considerations for Cloudy Conditions. As noted above, there is a trade off with the S/N ratio when the sky is cloudy, and also because the camcorder will try to maximize its aperture, which can cause chromatic aberration or other problems. Nonetheless, camcorder operators were instructed to choose a shutter speed of at least 1/250 s.



Figure 7.8: Shutter priority switch.

P : Flexible Recording Programs (46)

[P] PROGRAM AE

The camcorder sets the aperture and shutter speed.

[TV] SHUTTER-PRIO. AE

You select the shutter speed.

[AV] APERTURE-PRIO. AE

You select the aperture value.

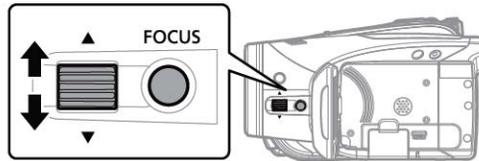
Figure 7.9: Shutter control vs aperture control.

Focusing. Focus is a critical issue. In order to see the license plate and not to be thwarted by occlusion, low zoom is the only option. Unfortunately, focus cannot be perfect on multiple lanes at the same time. The more lanes in the field of view, the worse the focusing will be. Since the shutter speed will be high, the camera will compensate with a lower aperture. The danger of autofocus is that the camera will constantly be changing focus as vehicles move along. Camcorder operators were instructed to use manual focus with the controls illustrated in Figure 7.10, and to set focus at infinity.



Adjust the zoom before you start the procedure.

Mode switch: P



- 1 Press **FOCUS**.
"MF" appears.
- 2 **Adjust the focus as necessary with the FOCUS dial until the image appears focused.**
Pressing **FOCUS** again will return the camcorder to autofocus.

Figure 7.10: Focus control.

Sharpening and contrast. As shown in Figure 7.11 and Figure 7.12, settings for contrast and sharpening can be adjusted. This is accomplished by selecting the custom mode, and changing the sharpness and contrast parameters. There is a trade-off, however: the more the sharpness or contrast is increased, the more noise and jpeg artifacts will be seen. Camcorder operators were instructed to adjust these settings during setup as appropriate for that day. Slightly increased contrast was expected to be helpful for this application, whereas all the other parameters we expected to be appropriate by default.

Image Effects

You can use the image effects to change the color saturation and contrast of your recordings.



Mode switch: P (except for the Special Scene recording programs)

Figure 7.11: Custom image effects.

[CUSTOM]	
Allows you to adjust the brightness, contrast, sharpness, and color depth of the image.	
[COL.DEPTH]:	(-) Shallower colors, (+) Richer color tonality
[BRIGHTN.]:	(-) Darker picture, (+) Brighter picture
[CONTRAST]:	(-) Flatter picture, (+) More defined light and shade
[SHARPNESS]:	(-) Vaguer outlines, (+) Sharper outlines

Figure 7.12: Controls for contrast and sharpness.

7.6 Narrative of Camera Deployment

Procedures during deployment. On the day of the field experiment, a total of nine people in three teams were tasked with recording northbound traffic. The three team leaders were Christian Claudel, Saurabh Amin, and Tim Racine. The groups departed from CCIT at 8:30 am and drove to one of three vantage sites shown in Figure 7.13. Setup began at 9:30 am and was completed by 10 am. Note in the figure that the cameras were pointing southward, and there were no significant glare problems despite the sunny weather. Filming began at 10:45am, and ran straight through the experiment with no interruptions. During the actual experiment, all equipment at any one site could be monitored by only one person, so there were no logistical difficulties with breaks or meals. Ultimately, the pencil and paper documentation for film start times for each camcorder was used for the actual synchronization. Backup copies of all video data were made the next day.

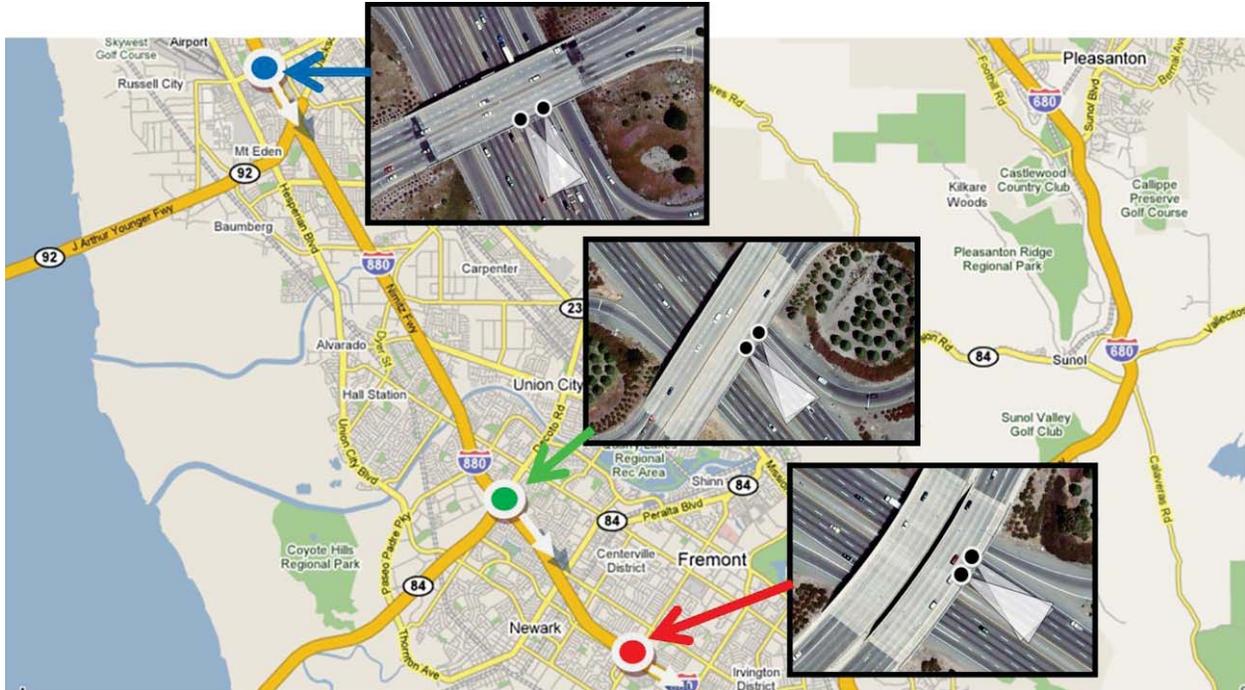


Figure 7.13: Camera deployment during MC experiment. Stevenson and I880 (South bridge), Decoto and I880 (Central bridge), Winton and I880 (North bridge).

7.7 Post-processing of Video Data

Collected data. At the conclusion of the field experiment, the video teams ended up with six large video file sets: one video file set from each camcorder from three bridges with two cameras at each bridge. For safety, backup copies of all video data were made. There were approximately six to seven hours of video per camera. Videos were viewable with the VLC media player.¹⁸ However, the video data was cumbersome to view.

Re-identification of vehicles. The problem to be solved is to extract travel times by re-identifying license plates. The first step toward accomplishing this was to record the license plate number of each vehicle observed in the videos.

Re-identification procedures. The FFMPEG program¹⁹ was used to subsample the video at four frames per second. A huge text file was generated with one time-stamped line per image frame. Undergraduate students were tasked with entering the license plate number of each

¹⁸ <http://www.gofree.com/vlcmediaplayer.php?gclid=Clvtq7a3waQCFRkjiQodZkWqEA>

¹⁹ <http://ffmpeg.org/>

vehicle as it approached the bottom of the video frame in the appropriate line in the text file. This work was labor intensive, and video data could only be human-processed at a rate 10-times below that of the real-time viewing rate of the video. Approximately 12 undergraduate students were hired to perform this task over about one month.

Re-identification result. The process resulted in a set of text files in which time-stamped appearances of license plate numbers were recorded at each bridge location. A script was written to automatically parse the text files and to match license plate numbers with time-stamps. The re-identification rate was about 20% of the total flow. The processed video data is freely available for download from the Mobile Millennium website.²⁰

²⁰ <http://traffic.berkeley.edu/downloads/>

8 Probe Vehicle Deployment

Nicknamed the Mobile Century experiment, on February 8, 2008, the privacy-preserving data collection system (presented in Chapter 4) was built and used to estimate traffic conditions for a day on I-880 near San Francisco, CA. With the help of over 200 UC Berkeley students, 100 vehicles carrying Nokia N95 phones were driven in repeated loops of six to ten miles in length continuously for nine hours. Drivers were instructed to drive as they would normally, on one of three routes. These vehicles represented approximately 3% of the total flow of traffic on the highway during the experiment.

The purpose of this section is to describe as faithfully as possible the deployment of probe vehicles during the field experiment. In particular, the resources employed and the procedures used are explained.

8.1 Resources

Drivers. The pool of drivers consisted of 165 UC Berkeley students over the age of 21. Participants were selected and hired according to the criteria and procedures in Appendix 1. For example, participants were required to have valid driver's licenses, to have more than 100 miles of driving experience in California, to be fluent in English, and to be in good physical condition. Participants were instructed to gather at CCIT at on the morning of the experiment. Participants were delivered to the Base Camp via buses. At the conclusion of the experiment, participants were returned to CCIT by bus.

Support Staff. Support staff, including about 40 student field officers, was hired to care for the drivers.

<i>Position</i>	<i>Staff Member</i>
<i>Experiment Director (ED)</i>	Alexandre Bayen, Assistant Professor UC Berkeley
<i>Experiment Supervisor (ES)</i>	Dan Work, PhD Student, CEE UC Berkeley
<i>Experiment Host (EH)</i>	JD Margulici, Associate Director CCIT
<i>Human Logistics Officer (HLO)</i>	Steve Andrews, CCIT
<i>Site Logistics Officer (SLO)</i>	Ryan Herring, GSR, CCIT
<i>Nokia Tech. Manager (NTM)</i>	Quinn Jacobson
<i>Site Managers (SM)</i>	Dr. Ali Mortazavi, Senior Dev. Engineer, CCIT Marika Benko, Staff Research Associate
<i>Safety Officers (SO)</i>	Olli-Pekka Tossavainen, Post Doctoral Researcher Osama Elhamshary, CCIT
<i>Assistant Human Logistics Officer (AHLO)</i>	Tia Dodson, CCIT
<i>Human Logistics Assistants (HLA)</i>	Chris Flens-Batina, UC Berkeley
<i>Site Logistics Assistant (SLA)</i>	Xavier Litrico, IGRF Research Fellow, Fr.

Position	Staff Member (continued)
<i>Site Assistants (SA)</i>	Alan Wang Alexis Clinet Anurag Sridharan Carl Misra Charlotte Wong Christina Sedighi Elizabeth Kincaid Irene Kwan Jennifer Chang Jessica Ariani Julie Percelay Karl David Cruz Kristen Ray Madeline Ziser Matthew Vaggione Nick Semon Sandy Do Swe Shin Maung Timothy Racine Trucy Phan Tyler Moser
<i>Head Personnel Manager (HPM)</i>	Juan Carlos Herrera, PhD Student, CEE UCB
<i>Phone Operators (PO)</i>	Nokia Staff
<i>Team Leaders (TL)</i>	Jean Parks (UCB, Red Team) Kristen Parrish (UL, Yellow Team) Josh Pilachowski (UL, Red Team) Justin Pope (UCB, Orange Team) Matt Vaggione (UCB, Yellow Team) Megan Smirti (UL, Orange Team)
<i>Team Personnel Assistants (TPA)</i>	Tarek Rabbani Qingfang Wu Arthur Wiedmer Manju Kumar (Backup)
<i>Team Traffic Assistants (TTA)</i>	Alexander Alshanetsky (UCB, Yellow Team) Negin Aryaee (UCB, Orange Team) Timmy Siuaw (UL, Yellow Team) Emma Strong (UCB, Red Team) Jason Wexlar (UL, Orange Team) Anthony Patire (UL, Red Team)
<i>Nokia Monitors (NM)</i>	Nokia Staff
<i>Raffle Team</i>	Coralie Claudel Andrew Tinka

Rental cars. Rental cars were provided by Enterprise and delivered to the Union Landing parking lot on the evening before the experiment. At the conclusion of the experiment, rental cars were returned to the parking lot and retrieved by the rental car company. For ease of identification, all rental cars were marked with prominent decals. The decals consisted of door logos, and a color-coded hood adhesive to designate the vehicle number and to match teams and routes. A total of 99 vehicles were used as probes. Additional vehicles were available as spares or to enable management functions.

GPS-enabled phones. Nokia N95 phones, Bluetooth headsets, and pre-paid voice and data plans were provided by Nokia. The Bluetooth headsets were used to provide a communication link between support staff and drivers to ensure smooth operation of the experiment. Phones were pre-loaded with client software to collect detailed trajectory data on the phone, and to communicate a subset of that data in real time.

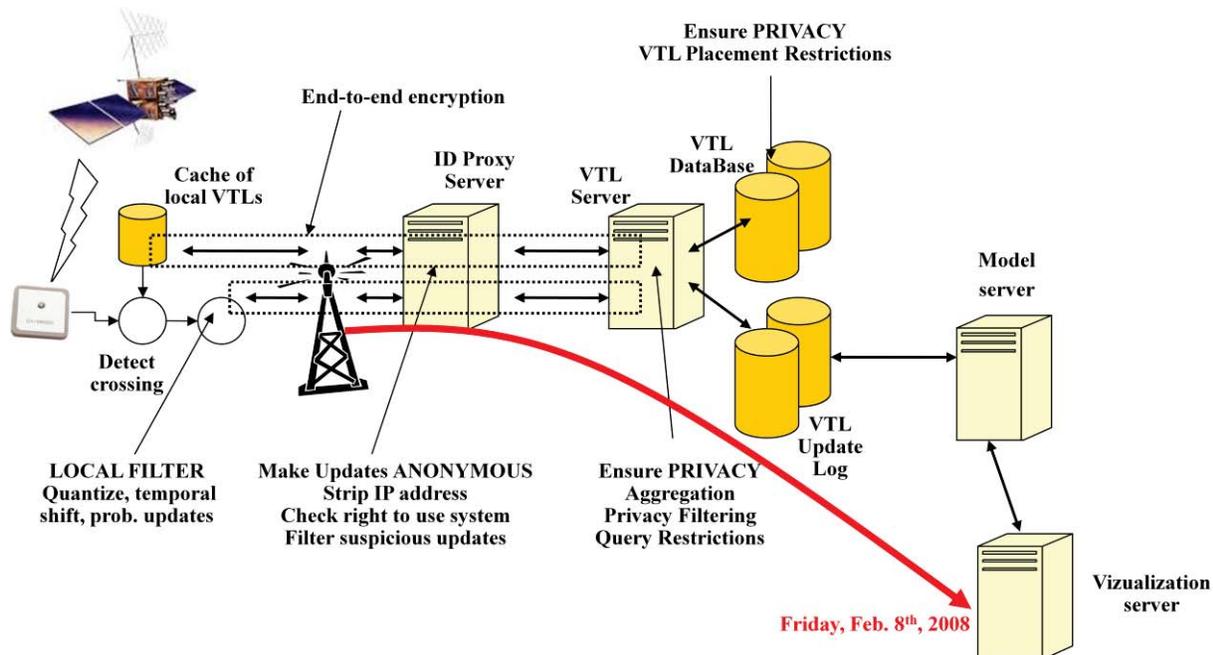


Figure 8.1: Traffic monitoring infrastructure built for field experiment.

Traffic monitoring infrastructure. The privacy-preserving data collection system was built according to the principles described in Chapter 4. For the experiment, however, no data was sent back to the phones. Cell phones were strictly used as sensors, and the flow of information was one-way. VTL positions were pre-loaded in the phone. Backup servers were located at Nokia labs in Palo Alto.

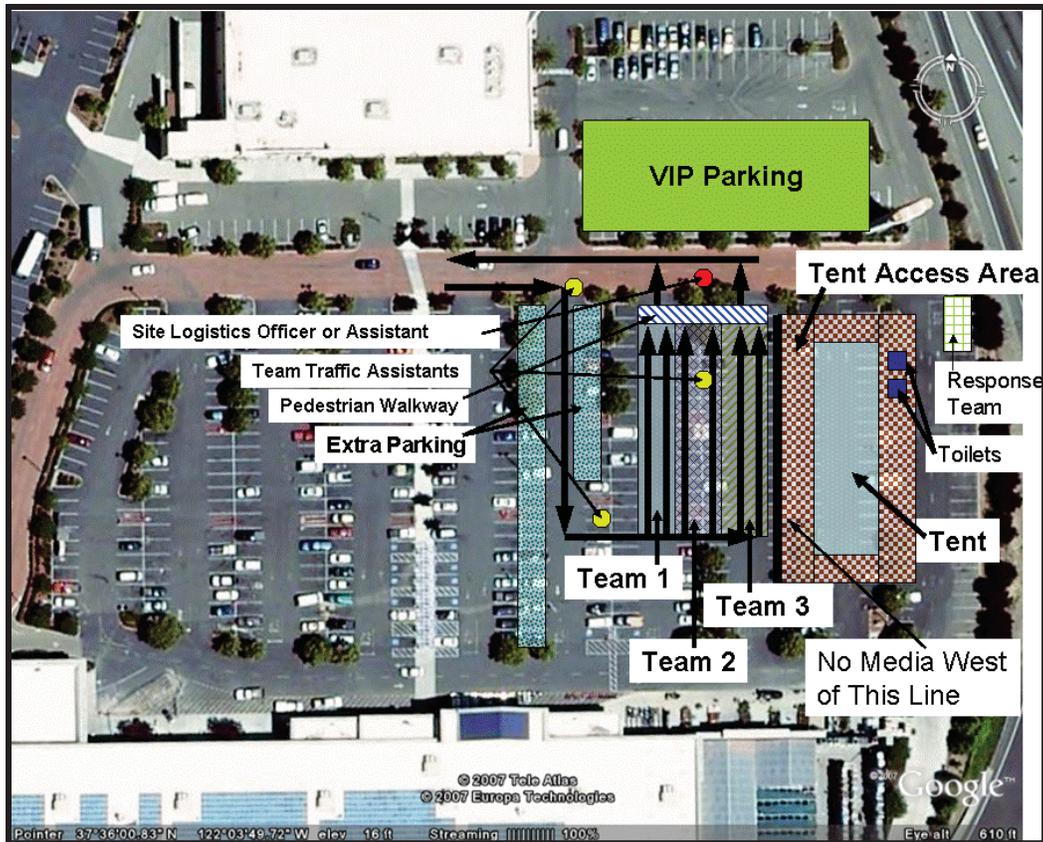


Figure 8.2: Layout of Base Camp on the Union Landing Parking Lot.

Base Camp. The parking lot at the Union Landing shopping center was utilized as the base camp for this experiment. The base camp shown in Figure 8.2 provided a variety of functions for participant care, driver briefing, safety monitoring, press briefing, and experiment staging. Restrooms and power generation equipment were installed for the day.



Date: 02/08/08
 Project: CCIT
 Owner: Andrew Sutton
 Designer: Andrew Sutton

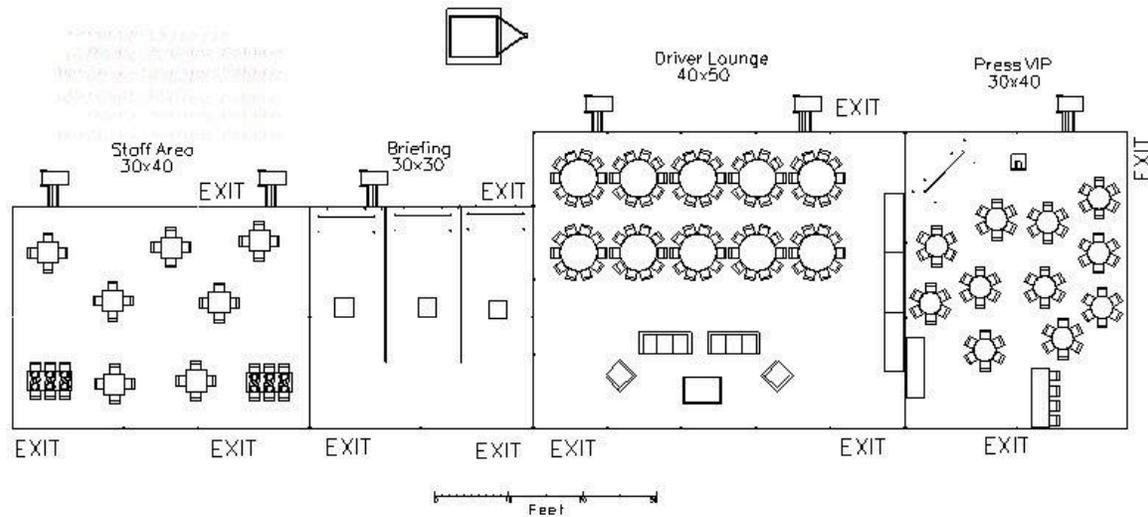


Figure 8.3: Tent Layout.

Tent. The tent housed a lounge for drivers to rest and to eat meals between driving shifts. In addition, space was available for driver briefing, press briefing, and a command and control center, labeled “Staff Area” in Figure 8.3. The CCC, equipped with computers and internet connectivity, was used to monitor the experiment and maintain safe operations. Restroom facilities and power generators were installed just outside the tent.

Response Team. In addition to a CCC, two paramedics and one tow truck were kept on site to assist if needed.

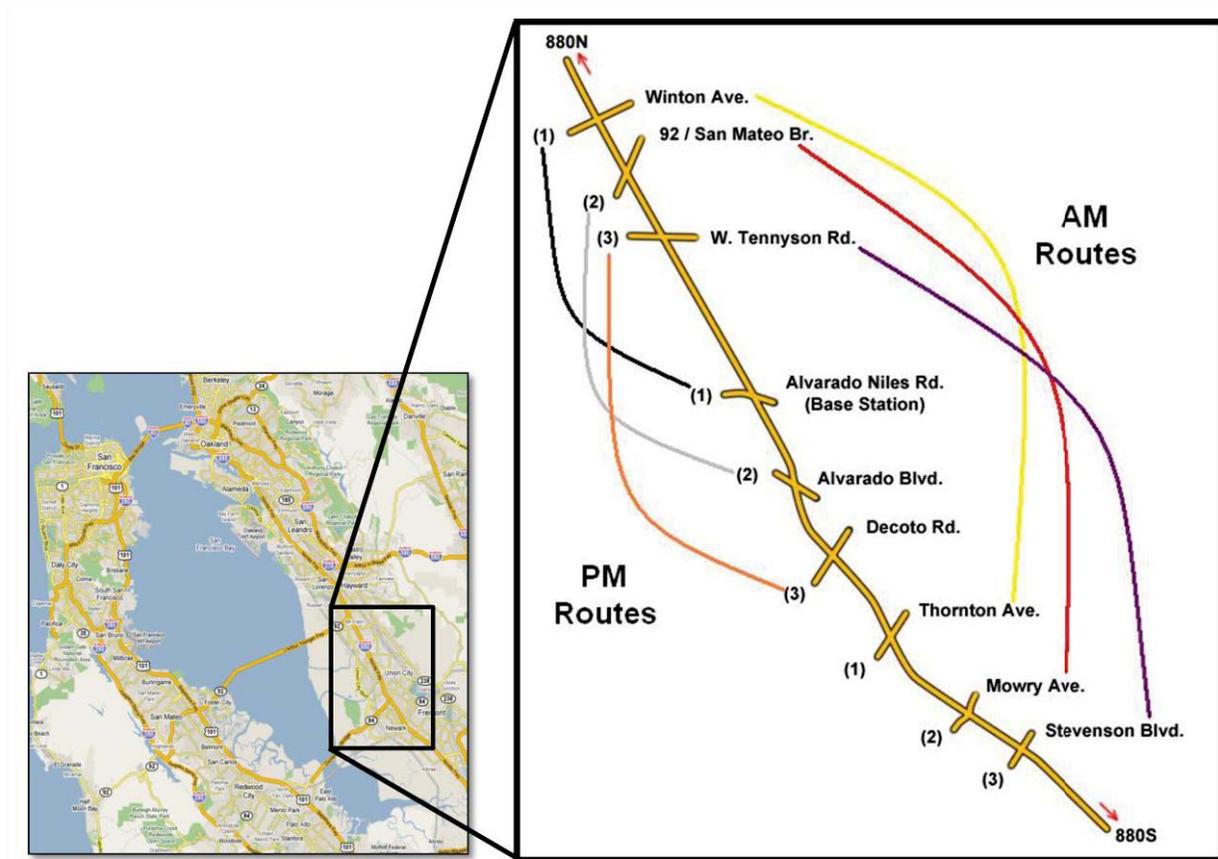


Figure 8.4: Study site and driver routes.

Study Site. The experiment was conducted on Highway I-880 between Winton Ave. to the North and Stevenson Blvd. to the South on February 8, 2008. This 10-mile long section was selected for its traffic properties, an existing knowledge-base for this particular highway from traffic simulations, and for its proximity to UC Berkeley. For logistic purposes, convenient facilities were available in the vicinity of this highway, including access to parking, gasoline, and food.

This section of highway has four (and sometimes five) lanes, the leftmost one being a HOV lane. Alternating periods of free-flow and congestion occur throughout the day. In particular, the northbound (NB) direction presents a recurrent and severe bottleneck between Tennyson Rd. and CA92 during the afternoon. The section is also well covered with dual loop detector stations (25 ILDs in the NB direction).

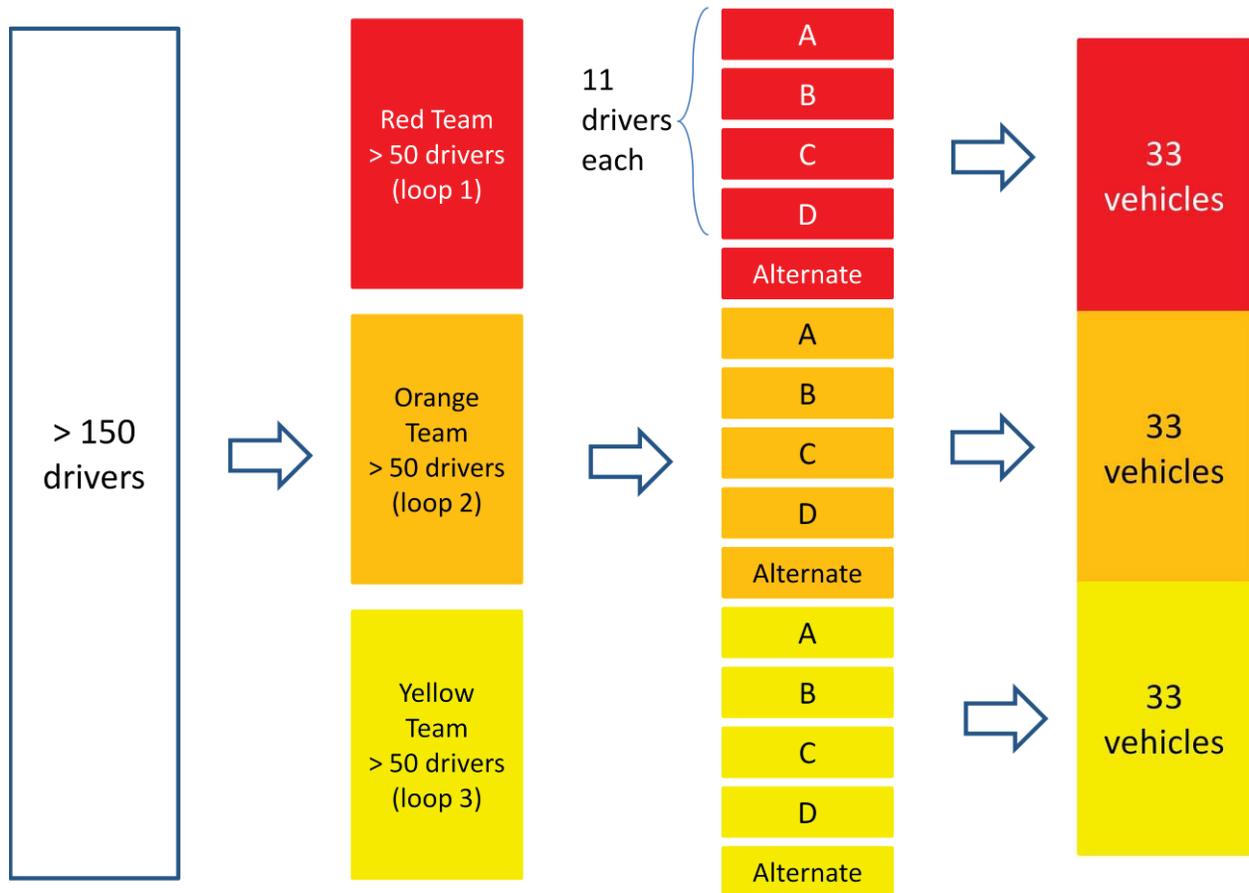


Figure 8.5: Assigning drivers to routes and cars.

8.2 Procedures

Mapping drivers to routes and cars. Drivers were divided into 3 teams: Yellow, Orange and Red. Each driver on the same team was assigned to drive on the same loop. The 99 rental cars were also divided into 3 groups of 33, and assigned a color according to the same loops and teams above. For example, every member of the red team was always assigned to a red car and to drive on a red loop.

Mapping drivers to schedules. Drivers were also divided according to a schedule so that within any team, there would always be one group resting and three groups driving. Five groups were defined: A, B, C, D and Alternates. The first 4 groups had 11 drivers each, and the last one had any remaining drivers. Each member of group A (cutting across all teams; yellow, orange, and red) had the same time schedule for driving time and breaks. Alternate drivers were available to fill in for scheduled drivers as necessary. As shown in Table 2, driving shifts were scheduled with a maximum duration of three hours.

Table 8.1: Driving Schedules Sorted by Car Color

	10:00	10:30	11:00	11:30	12:00	12:30	1:00	1:30	2:00	2:30	3:00	3:30	4:00	4:30	5:00	5:30	6:00	6:30
<i>Car Color:</i>																		
Red, A		BREAK		AM	AM	AM	AM	PM	PM	LUNCH		PM	PM	PM	PM	PM	PM	
Red, B		AM	AM	LUNCH		AM	AM	PM	PM	PM	PM	LUNCH		PM	PM	PM	PM	
Red, C		AM	AM	AM	AM	LUNCH		PM	PM	PM	PM	PM	PM	LUNCH		PM	PM	
Red, D		AM	AM	AM	AM	AM	AM	LUNCH		PM	PM	PM	PM	PM	PM	LUNCH		
Orange, A		BREAK		AM	AM	AM	AM	PM	PM	LUNCH		PM	PM	PM	PM	PM	PM	
Orange, B		AM	AM	LUNCH		AM	AM	PM	PM	PM	PM	LUNCH		PM	PM	PM	PM	
Orange, C		AM	AM	AM	AM	LUNCH		PM	PM	PM	PM	PM	PM	LUNCH		PM	PM	
Orange, D		AM	AM	AM	AM	AM	AM	LUNCH		PM	PM	PM	PM	PM	PM	LUNCH		
Yellow, A		BREAK		AM	AM	AM	AM	PM	PM	LUNCH		PM	PM	PM	PM	PM	PM	
Yellow, B		AM	AM	LUNCH		AM	AM	PM	PM	PM	PM	LUNCH		PM	PM	PM	PM	
Yellow, C		AM	AM	AM	AM	LUNCH		PM	PM	PM	PM	PM	PM	LUNCH		PM	PM	
Yellow, D		AM	AM	AM	AM	AM	AM	LUNCH		PM	PM	PM	PM	PM	PM	LUNCH		

AM Route = Long Route (before 1:30), PM Route = Short Route (after 1:30)

Probe vehicle routes. Loops were chosen to maintain a desirable penetration rate near 3% of the total volume of traffic on the highway during the experiment. Given that the total flow expected on the section of interest is approximately 6000 vehicles per hour (obtained from PeMS) and the number of equipped vehicles is 100, the necessary cycle time to achieve the desired rate is about 20 minutes. Knowing the expected speed throughout the day and cycle time is sufficient to determine appropriate lengths for the loops during the day. In the NB direction, free flow conditions are historically expected during the morning until 2-3pm, when the recurrent bottleneck between Tennyson Rd. and CA92 activates. Free flow is expected during most of the day for the southbound direction. For this reason, long loops (or AM) loops were chosen during the morning and short (or PM) loops were used during the afternoon. The change was scheduled to start at 1:30pm. Table 8.2 presents the main features of the loops used during the experiment, also shown in Figure 8.4. Three different loops of almost the same length were used during the experiment so as not to over-saturate any of the ramps being used.

Monitoring and Safety Procedures. A line of command for emergency situations was defined according to Appendices 3.1, 3.2, and 3.3.

Table 8.2: Features of the loops (routes) used in the experiment.

Loop type	North end	South end	One-way distance (miles)
	CA92 (San Mateo Br.)	Mowry Ave.	8.6
AM (long)	Winton Ave	Thorton Rd.	9.4
	Tennyson Rd.	Stevenson Blvd.	9.3
PM (short)	Winton Ave.	Alvarado-Niles	4.5
	CA92 (San Mateo Br.)	Alvarado Blvd.	5.2
	Tennyson Rd.	Decoto Rd.	5.4

8.3 Collection of Probe Data

The probe data were collected in several ways during the experiment.

Phone logs. First, each Nokia N95 cell phone stored its position and velocity log once every 3 seconds, allowing for the computation of every equipped vehicle's trajectory. These data were gathered locally on the phones for analysis purposes, and were not part of the data gathering process of the system presented in the previous section. They became available only once the experiment was finished, and were useful to test the accuracy of the sampling strategy a posteriori.

VTL crossings (real-time). The privacy-aware architecture described in Chapter 4 collected data from the 30 VTLs deployed between Stevenson Blvd. and Winton Ave. (each VTL covers both travel directions). These data were used to produce real-time travel time and speed estimates, and helped to assess the feasibility of a system based solely on Lagrangian data.

8.4 Narrative of Experiment

As-built vs as-planned

The purpose of this section is to explain as faithfully as possible the apparent discrepancies between the experiment design as explained in the previous section with the experiment design as planned in the historical documents included in the Appendices. In particular, we address the protection of human subjects in Appendix 1, and the planned protocol in Appendix 6.

Protection of Human Subjects

A preliminary version of the study procedures and planned protocol appear in the CPHS protocol narrative form included in Appendix 1. This document captures the spirit of the operation with regards to the care of the participating drivers. In particular, this document

describes the selection and hiring criteria for the participating drivers. It also describes considerations taken into account for the safety and comfort of all participants.

The preparation and conclusion phases of the study procedures evolved substantially from the initial conception documented in Appendix 1.

Planned Protocol

The actual execution of the Mobile Century experiment differed somewhat from the final version of the Protocol. For historical purposes, the January 30, 2008 version of the Protocol document is included as Appendix 6. This document encapsulates that overall spirit of the operation, the roles and responsibilities of the main actors, and detailed schedules for the coordination of resources. The purpose of this section is to identify the main deviations between the plan and the experiment as it was actually executed.

The largest difference is in the logistics of the rental cars and drivers. The protocol document describes a plan wherein rental cars were to be delivered to a parking lot in downtown Berkeley. Drivers were scheduled to meet and to drive the cars to the base camp before the beginning of the experiment. To simplify operations, Enterprise was tasked to deliver and pick up the rental cars directly from the base camp. This finesse eliminated the risk of drivers getting lost or involved in accidents before the experiment could start. In addition, it reduced risk by restricting the experimenters' responsibility for the vehicles between the base camp and the experimental site. All tasks associated with the checking-in and checking-out of drivers at the beginning and end of the day were accomplished at the base camp.

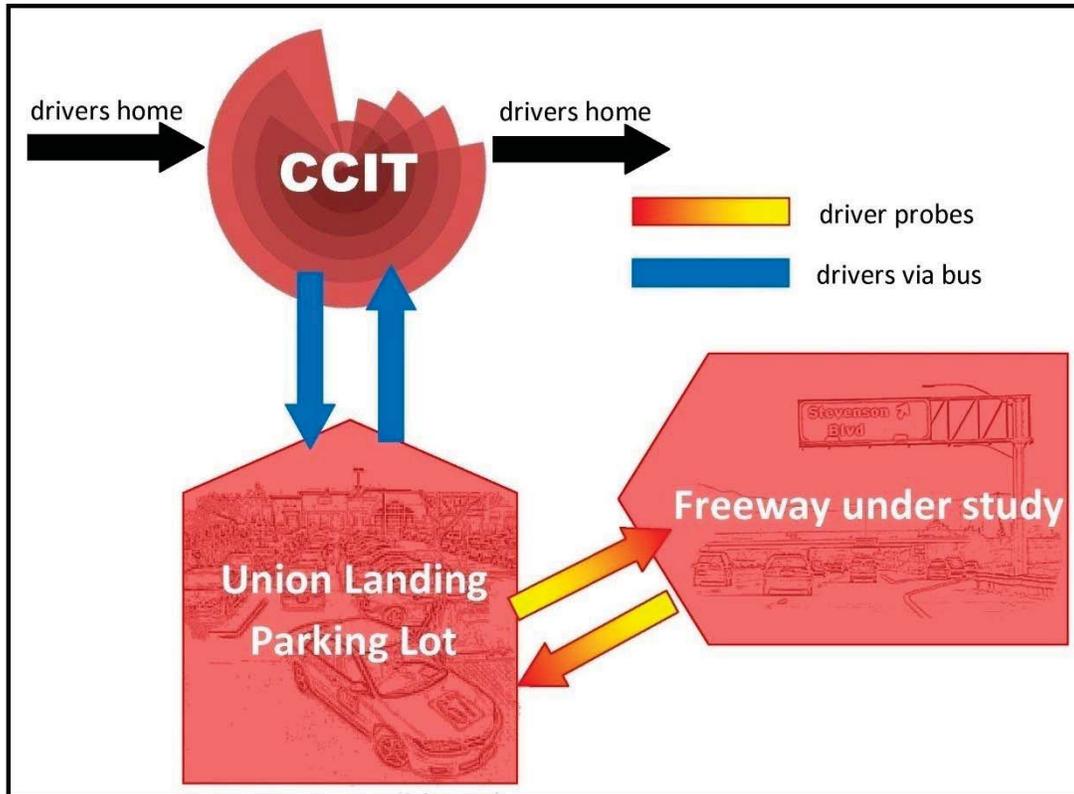


Figure 8.6: Drivers logistics

Guest Program

Guests from the press, industry, and government were invited to observe the experiment. One bus was reserved to provide transportation between CCIT and the base camp. The experimental kick-off at the base camp was followed by a press conference and guided tours of the CCC. Afterward, lunch was served and presentations were held at CCIT. Finally, a scientific seminar was held in conjunction with the ITS Friday seminar. Details of the program are provided in Appendices 7.1 and 7.2.

Unplanned Occurrences

At around 10:30 am, a multiple-car accident created significant unanticipated congestion on northbound traffic south of CA-92. The EnKF algorithm, running in real-time during the experiment, detected the accident's resulting bottleneck and corresponding shock wave.

Use of Emergency Procedures

During the actual experiment, there was one flat tire, and one non-working rental car. The tow truck was promptly deployed to retrieve the vehicle with the flat tire. The effect on traffic was minimized.

Lessons Learned

There was one significant problem that was undetected until after the experiment had concluded. About 20 phones never recorded or sent any data. This was undetected because when all the vehicles were at the lot, they appeared directly on top of each other on the map. There were about 4 monitors assigned to keep track of some subset of vehicles, but there was no single place where the status of all phones could be checked at once.

This could have been solved had there been one human readable keeping track of all of the phones, and providing a time when the last packet of data was received from each phone. Then, it would have been easy to notice the phones on the list that hadn't provided any updates within a time window (the last ten minutes).

8.5 Data Pre-processing

At the conclusion of the experimental 100-vehicle deployment, a significant amount of data was collected. Pre-processing of the data to make it usable for research was a considerable effort. This chapter describes the pre-processing steps for the collected data.

8.5.1 Phone Log Data

Each of the smartphones employed during the experiment had recorded trajectories in the form of GPS coordinates stored at a rate of one sample every three seconds. These data required significant processing to be used for research purposes. The pre-processing steps are summarized:

- Data download. The data was downloaded from the phones into a data server and stored in individual phone files.
- Anonymization of the data. Information linking position with vehicle ID for the day of the experiment was disposed.
- Analysis of the client-side filtering of GPS data. As described in Chapter 4, the smartphone client needed to detect VTL crossings and transmit correct velocity estimates to the VTL server. This data was processed during the experiment by the client on the phone. The filtering techniques employed needed to be analyzed to enable further refinement.

- **Filtering.** Errors not detected and caught by the client-side filtering onboard the smartphone were removed (outliers, erroneous time sampling leading to incorrect speed estimates, etc.) Speed computations (performed on the phone as well as during the experiment) were filtered in the same way.
- **Trajectory sequencing.** Because of the cyclical nature of the probe vehicle routes, trajectories were cut into tracks corresponding to the northbound and southbound portions of the trajectories of the cars. The remaining parts of the trajectories were disposed.
- **Database creation.** A database of tracks was created to access the northbound and southbound portions of the trajectories separately.
- **Map-matching.** The tracks were matched to mileposts to serve as inputs to the models. While this was done during the experiment, a database of milepost tracks was created.

8.5.2 VTL Server Data

During the experiment, data in the form of VTL crossings were transmitted from each phone to the VTL server. These VTL crossings served as the only data source for the real-time reconstruction of travel times and speeds. The processing of this data followed the steps outlined below:

- **Data download.** The data was downloaded from the VTL server into another data server and stored in VTL files.
- **Integrity check.** The data from the phone logs were compared with the data received by the VTL server. This comparison was used to quantify any discrepancies caused by transmission.
- **Filtering.** The data was filtered in exactly the same way as the data from the phone logs (identical steps, except that this data is a subset of that from the phone logs).

8.5.3 Video Data

The video data collected during the experiment was processed frame by frame in order to identify vehicle license plates. License plates were re-identified between camera locations to compute travel time. This effort provided a ground-truth database consisting of all vehicles that traveled the entirety of the expressway section studied on the day of the experiment. The data processing consisted of the following steps:

- **Database creation.** Because of the high resolution of video used during the experiment, the sizes of the files were enormous. These files were therefore divided into smaller files to be archived and processed separately.

- File analysis. For each vehicle entering the field of view a camera, a license plate number and an arrival time were entered in a database, camera by camera, lane by lane.
- Vehicle re-identification. Once the database was created, vehicles were re-identified. For this purpose, a script file was run to match license plates up to three times in the database (three bridges) and compute the corresponding travel time.
- Computation of counts. Additionally, counts were computed to obtain the full vehicle flow for the day of the experiment at all three bridge locations.

8.5.4 Database Tools

In addition to the processing described above, a set of database tools was necessary to use the processed data in a meaningful way. In order to analyze the data and study it thoroughly, an integrated set of tools was created.

- Interface. An interface was created to enable access to the database for all fields of interest (trajectories, logs, VTL readings, travel times, etc.) In particular, vehicles needed to be selectable by feature (AM, PM, lane based, time based, north bound, south bound, loop number etc.)
- Visualization. Vehicles needed to be locatable on the map (map-matching) to study trajectories. In addition, movies and animations were created in order to visualize traffic (at least for the probe vehicles).
- MATLAB interface. The database was made to be accessible through MATLAB, one of the key tools used by engineering students at UC Berkeley. A MATLAB interface was created to enable this access.
- Statistics tools. Basic statistics tools were created to process the data, and to run tests on the data before using it for practical purposes.

9 Experimental results

This chapter presents the main results of the experiment. Unless otherwise noted, this section focuses on the highway segment covered by the routes in the northbound direction. This section consists of the portion of highway between Stevenson Blvd. to the south (milepost 17) and Winton Ave. to the north (milepost 27.5).

Section 9.1 presents real time data that was obtained (and augmented with the assimilation technique of Chapter 6) and displayed live during the experiment. Sections 9.2 and 9.3 present an analysis of trajectories, and velocity fields, that were constructed from the raw data without augmentation by data assimilation techniques. Section 9.4 compares collected data with ground truth travel times recovered from the video validation efforts of Chapter 7. Section 9.5 explores disparities between ILD and VTL data. The achieved penetration rate of probes is evaluated in Section 9.6. Finally, Section 9.7 closes this chapter with a discussion of parameters inferred after the conclusion of the experiment from the shockwave speed captured by trajectories of probe vehicles.

9.1 Real-Time Traffic Monitoring Using Only Cell Phone Data

Real-time broadcast of traffic conditions. During the experiment, thirty VTLs were allocated along the expressway section. Every time a smartphone probe crossed a geographical VTL, it sent a VTL report through the cell phone network. This VTL report was collected by the traffic monitoring infrastructure described in Chapter 4 and processed in real time. The EnKF of Chapter 6 was employed to incorporate the velocity measurement from each VTL report into the evolution of the expressway traffic model. The resulting travel time and velocity estimates from the model were broadcast for eight hours. Figure 9.1 illustrates the interface used to broadcast travel time and speed during the day. The figure shows traffic soon after an accident occurred in the northbound direction between Tennyson Rd. and CA-92. The inset of Figure 9.1 shows the 511.org traffic display at the same time.



Figure 9.1: Snapshot of the live traffic feed provided by the system in the present work (and from 511.org in the inset) at 10:52am on February 8, 2008. Traffic conditions after an incident on the northbound direction of I-880 are displayed. Numbers in circles correspond to speed in mph.

Cell phone data vs. 511 status quo monitoring. The spatial extent of congestion reported by the system in the present work was comparable to that reported by 511.org on the day of the experiment. Recall that the present work makes use of spatially sampled velocity data from smartphone probes only. On the other hand, 511.org makes use of a combination of data sources including loop detectors, FasTrak-equipped vehicles, and radars. However, 511.org only provided speed estimates in discrete increments (e.g., the black color represents “stop and go” and red means “heavy traffic”), while the present work generated speed estimates with a finer scale. This capability is important because it allows a more accurate identification of the boundary between zones with different traffic states (i.e., the location of the shockwave). Comparisons with the 511.org speed map at other times during the experiment showed compatible results. This result confirms that the system in the present work can produce reasonable speed estimates for an expressway with only spatially sampled data from smartphone probes.

9.2 Trajectory data

Features of trajectory data. After the experiment, trajectories from the phone logs were processed in order to conduct a more detailed analysis of the quality of the data collected by the GPS-enabled smartphones. Figure 9.2 shows 50% of the gathered trajectories between Stevenson Blvd. (milepost 17) and Winton Ave. (milepost 27.5) in the northbound direction. The transition from the AM loops to the PM loops that occurred at 1:30 pm can be clearly seen in the figure, as well as the fact that different vehicles were using different ramps to get on and off the highway (as shown in Figure 9.2). The propagation of the shockwave generated by the incident is clearly identified from this plot as well.

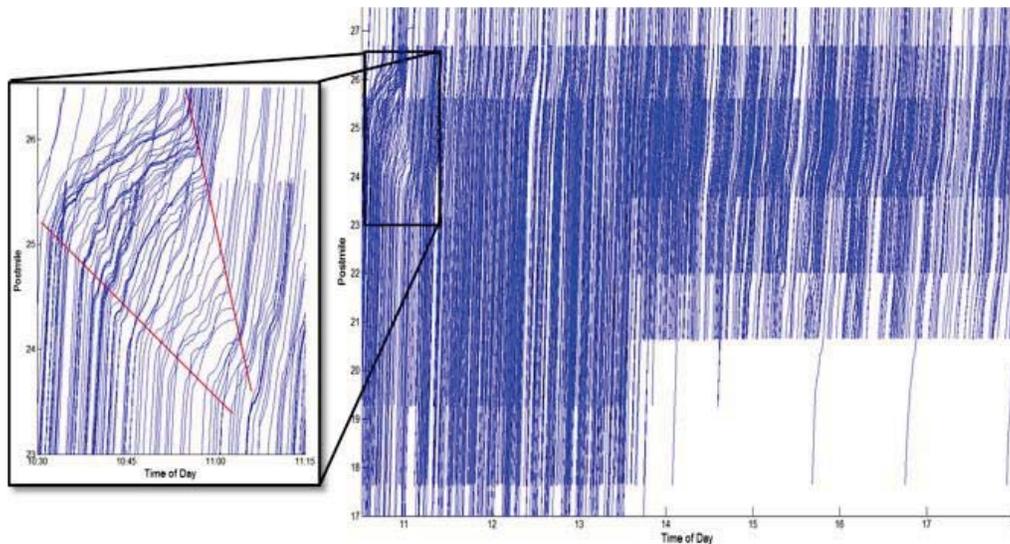


Figure 9.2: Vehicle trajectories in the northbound direction extracted from the data stored by 50% of the cell phones. The propagation of the shockwave from the accident can clearly be identified from this plot. The red lines in the close-up were drawn by hand by fitting a line through the points where trajectories change slope.

Shockwaves from traffic accident. The red lines, drawn by hand in Figure 9.2, represent the approximate propagation of the shockwaves generated by the incident. Information about the propagation of shockwaves can be used to infer parameters of the fundamental diagram for this stretch of expressway (assuming a triangular relationship between flow and density). As explained at the end of this Chapter, the speeds of the shockwaves relate driving parameters of the equipped vehicles with those of all vehicles experiencing the shock. This information may be used to infer flows and densities in the absence of loop detectors, despite the fact that mobile sensors are unable to directly measure flows and densities.

9.3 Velocity Field

Velocity field calculated from PeMS data. Loop detectors collect flow and occupancy data for each lane every 30 seconds. Every 5 minutes, a flow-weighted average velocity is computed from these measurements. The velocity field from PeMS data is shown in Figure 9.4(a). The method used in PeMS associates a spatial area with each detector station. The assumption is that measurements from the detector station are valid for the entire spatial area. The size of the spatial area depends on the proximity of neighboring detector stations. The closer the neighboring detectors, the more reasonable the assumption, and the better the estimates that can be obtained using this method.

Velocity field calculated from trajectory data. Using only the trajectories of equipped vehicles, the velocity field can be computed using Edie's generalized definition [32]. The corresponding result is shown in Figure 9.4(b). The qualitative agreement between subfigures (a) and (b) is evident in terms of the bottleneck locations, and their spatial and temporal extent. Note that

trajectory data from about 3% of the total flow is adequate to generate a velocity field comparable to the one generated from loop detector data on this relatively well instrumented section of expressway.

Feature of time-sampling with mobile sensors. When sampled in time (every 3 seconds in this case), mobile sensors can provide highly detailed information – such as the backward propagation of congestion – that would only otherwise be available with a high density of loop detector stations. Note that with a temporal sampling strategy, more observations – reporting low velocities – are expected to be available during congestion, because vehicles spend more time in it (and there are more vehicles per unit length).

Redefine VTLs a posteriori. The VTL data collection strategy used in the present work can be evaluated in more detail by considering the VTL reports that would have been generated had the VTL positions been different. This is possible because complete trajectories sampled at 0.33 Hz were stored on each of the cell phones during the experiment. Using the trajectory data, and desired VTL positions, a new set of VTL reports can be generated a posteriori.

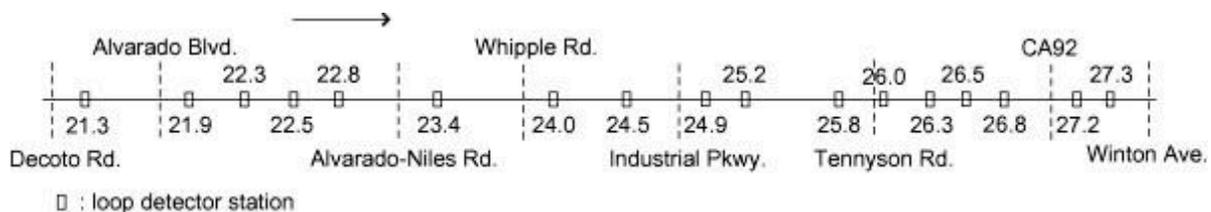


Figure 9.3: Loop detector locations along the northbound direction. Numbers indicating mileposts increase in the direction of traffic flow from left to right.

Set VTL location equal to ILD location. VTL reports were synthesized at the 17 loop detector locations shown in Figure 9.3. These reports were aggregated in 5-minute periods and the mean temporal speed was computed so as to be consistent in method with loop detector measurements from PeMS. This exercise yielded VTL measurements at the same 17 locations and sampled at the same rate as the loop detector measurements.

Compare velocity fields calculated from VTLs and ILDs at the same positions. Using data from the 17 VTLs above, the velocity field is reconstructed with same method employed for the loop detectors in Figure 9.4(c). The velocity map exhibits the same main features captured by the loop detector velocity field. Even though both sensors provide qualitatively similar information, there are some discrepancies in the velocity values they report (suggested by the difference in colors observed at certain times and locations).

Velocity field calculated from thirty VTLs. The velocity field in Figure 9.4(d) was constructed using data from the thirty VTLs originally allocated during the experiment, and is shown here for comparison. The greater detail in this plot is a consequence of the higher spatial sampling rate.

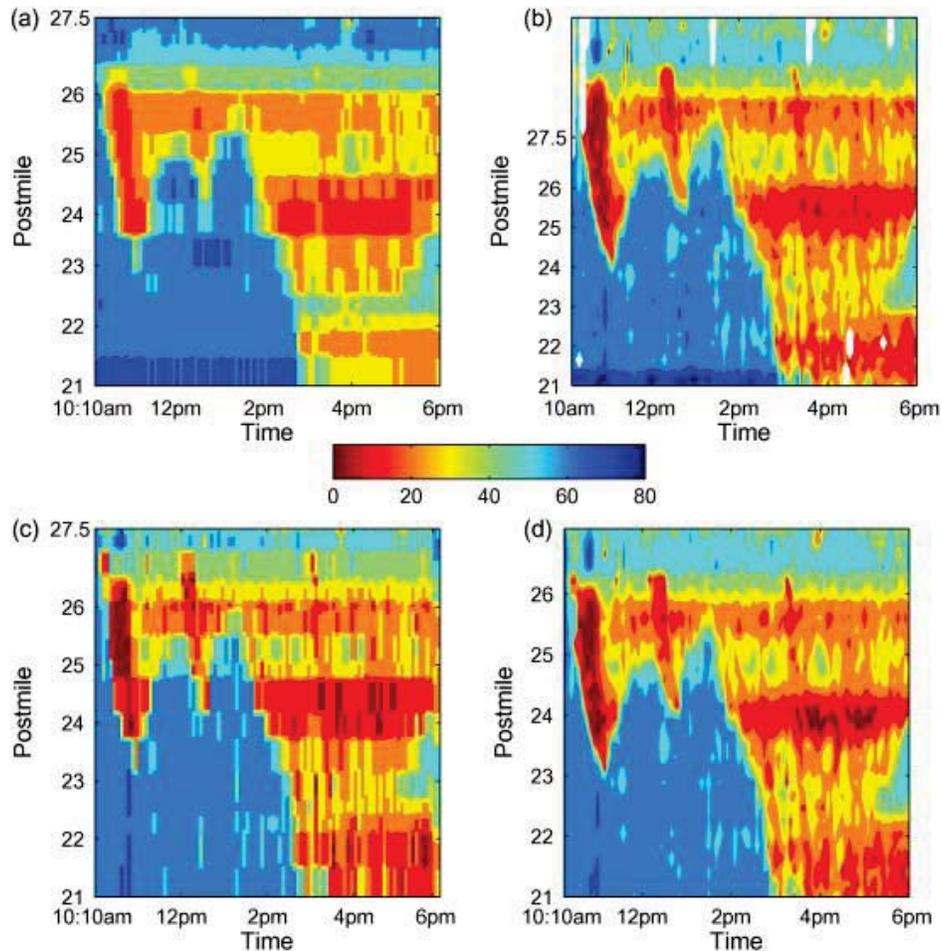


Figure 9.4: Velocity fields in mph using: (a) 17 loop detector stations; (b) vehicle trajectories and Edie's generalized definition; (c) 17 VTLs at the loop detector locations; and (d) 30 equally spaced VTLs.

9.4 Ground-truth Travel Times

Given the difference in the nature of the velocity measurements provided by VTLs and loop detectors, one fundamental question is to determine which measurements are more accurate. Since ground truth velocity is not known for the present experiment, the accuracy of VTL velocity measurements cannot be directly assessed. Note that loop detector measurements are usually considered as ground truth. However velocities on this stretch of expressway are estimated from single loop detector measurements, and it is known that these estimates may include substantial errors, depending on the algorithm used [63],[64].

Travel time measured with cameras. Individual travel times for vehicles traveling from Decoto Rd. to Winton Ave. between 10:45 am and 5 pm were extracted from high-definition video cameras by re-identifying license plates as described in Chapter 7. A total of 4789 vehicles were matched between 10:40 am and 5 pm. However 521 matches yielded unusually high travel times, suggesting that these trips were interrupted by at least one excursion from the expressway. These outliers were excluded from further analysis. The remaining 4268 travel times, plotted as grey dots in Figure 9.5, represent at least 10% of all vehicles traveling the entire section between 10:40 am and 5 pm.

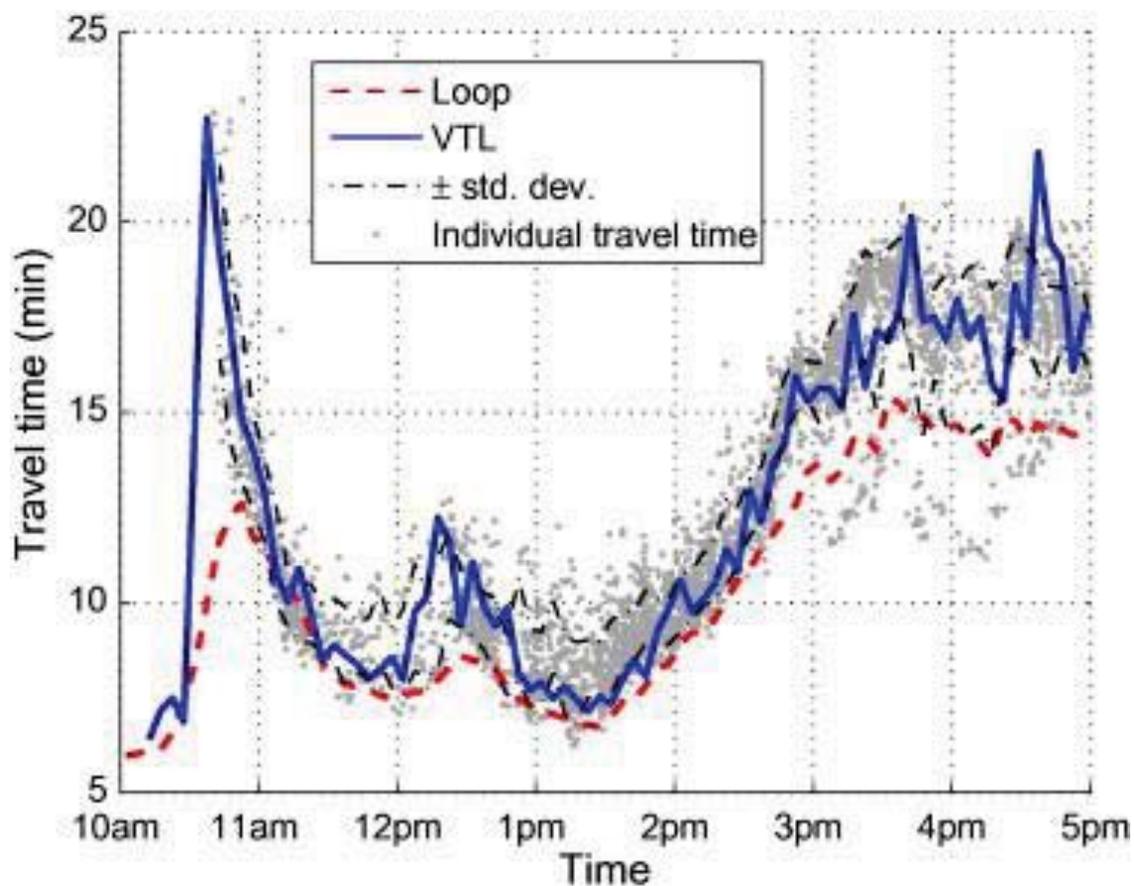


Figure 9.5: Travel time (in minutes) between Decoto Rd. and Winton Ave. The x-axis indicates arrival time at Decoto Rd. Dots correspond to individual vehicle travel times (4268 in total), collected manually using video. Black dash-dotted lines correspond to the standard deviation of travel times obtained from video cameras in 5-minute windows.

Travel time calculated from velocity fields. Velocity fields can be integrated to compute travel time.²¹ This can be used to assess which velocity measurements are more likely to be closer to ground truth. Figure 9.5 displays travel times computed by integrating the VTL and loop detector velocity fields in blue and red, respectively. All travel times reported above are plotted against vehicle arrival time at Decoto Rd. (i.e., when a vehicle enters the section and begins the northward trip). Note that after 3 pm, there are a number of grey dots beneath the curves that indicate unusually low trip times. These observations reflect the fact that at 3 pm, the left-most lane becomes a HOV lane.

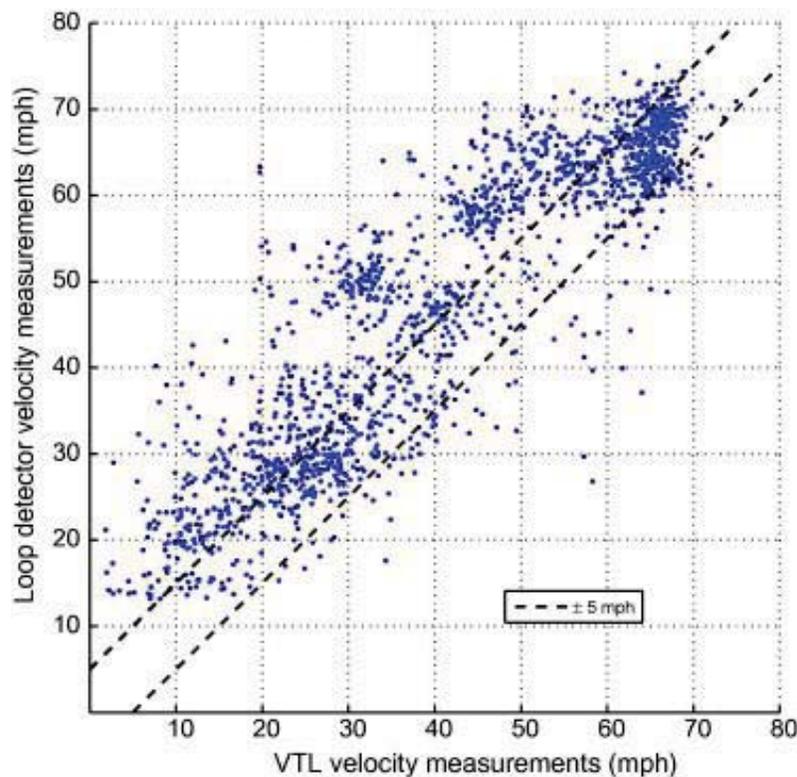


Figure 9.6: Loop detector vs. VTL velocity measurements (all locations). Dotted lines are the ± 5 mph thresholds.

²¹ This a-posteriori travel time estimation method is also known as dynamic travel time or the walk the speed matrix method.

Loops underestimate travel time in congestion. Both estimates reflect the main trend observed in travel times during the day. The VTL estimates, however, reproduce the observed travel times more accurately. Loop detector estimates tend to underestimate travel times, implying that they tend to overestimate velocities. In fact, the VTL estimates are almost always within one standard deviation of the average observed travel times obtained from the video cameras in 5-minute windows (represented by the two black dash-dotted lines in the figure). In stark contrast, the travel times estimated from loop detectors are almost always more than one standard deviation below this average.

VTL travel times are closer to ground truth. Travel times computed with the VTL velocity field are in greater agreement with the observed travel times than those calculated from loop detector data. This suggests that the VTL velocity field is more likely to be closer to the actual velocity experienced by the vehicles, and is therefore more accurate than the loop detector velocity field. That is, accuracy afforded by GPS technology is such that a sample of speed measurements from a small proportion of vehicles can provide more accurate estimates of velocity for the entire population than can loop detectors (even though loop detectors sample the velocity of all vehicles, eventually). This has to be kept in mind when loop detector data are considered as ground truth, especially for an assessment of alternative data sources.

Overall comparison of loop vs. VTL velocity data. Figure 9.6 plots the 5-minute aggregated velocity measurements from VTL data versus loop detector data for all the observations collected at the 17 locations. For high velocities (over 55 mph) about 70% of measurements have an absolute difference of less than 5 mph. However, for low velocities (below 40 mph) only 31% of measurements have an absolute difference of less than 5 mph. In most cases, loop detector velocity measurements tend to be higher than VTL measurements, and the discrepancy is greater for lower velocities. This difference explains the smaller travel times computed with the loop detector velocity field, shown above in Figure 9.5.

Compare ILD vs. VTL data over space and penetration rate. These matters are further explored in Figure 9.7. Figure 9.7 shows a time-series of loop detector and VTL velocity measurements for four different locations with time-varying penetration rates during the day, shown in Figure 9.7(d). At some locations, a bias can be observed near the off-ramps used during the experiment, where VTL velocity measurements are always lower than the loop detector velocity measurements, for example milepost 27.3 as shown in Figure 9.7(c). However, this bias is not observed at some other locations, for example milepost 24.0 shown in Figure 9.7(d). The level of discrepancy varies with time, location, penetration rate, and traffic conditions (i.e., velocity).

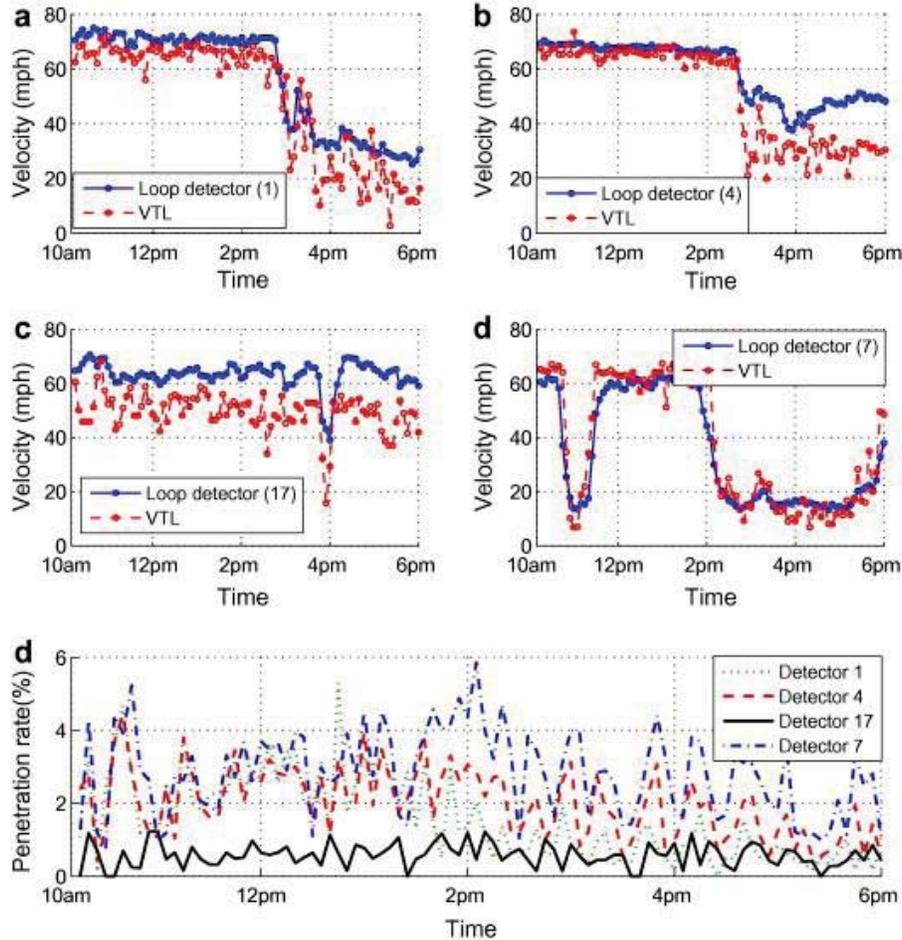


Figure 9.7: Loop detector and VTL velocity data collected at: (a) milepost 21.3, downstream of Decoto Rd.; (b) milepost 22.5, half-way between Alvarado Blvd. and Alvarado Niles Rd., (c) milepost 27.3, the most downstream detector near the Winton Ave. exit; and (d) milepost 24.0, downstream of Whipple Rd. Subfigure (e) shows the penetration rate at each of these four locations during the day.

9.5 Reasons For Disparity Between Loop and VTL Data

There are several reasons to discuss: (1) selection bias due to sample participants; (2) selection bias due to low penetration; (3) bias in loop detectors; (4) bias in test driver dynamics before exiting or after entering the mainline of the highway.

Selection bias because participants are not representative of the greater population. The difference between both types of measurements raises the question about the presence of selectivity bias in the sample chosen for this experiment. Drivers hired for the experiment are not necessarily a proper statistical sample of the population. The 165 drivers were UC Berkeley students over 21, which may constitute a biased sample of the driving population. In addition to this, the driving behavior may be biased with respect to the rest of the traffic for other reasons, including fatigue or knowledge gained of the location and driving conditions (which may be similar to the expertise gained by regular commuters).

Bias due to low penetration. Figure 9.7(a) and Figure 9.7(c) have the lowest penetration rates; i.e., the VTL measurements survey a smaller proportion of the total flow. VTLs collect velocity from only a proportion of all vehicles crossing that location, while loop detectors collect data from (eventually) all the vehicles. If this proportion is too small, it might not be statistically representative of the entire population.

Bia due to driver dynamics. In addition, the data from Figure 9.7(a) is close to the Decoto Rd. freeway entrance, and the data from Figure 9.7(c) is near the Winton Ave. exit. Lower speeds in the VTL measurements may reflect a bias in test driver dynamics before exiting or after entering the main-line of the highway.

Bia due to loop detector bias. Figure 9.7(b) and Figure 9.7(d) have mid- to high-penetration rates, and are not located near any on- or off-ramps. Nonetheless, significant discrepancies exist between loop and VTL data for Figure 9.7(b). This suggests that some loop detectors are either biased or not computing the velocity properly. Loop detectors and VTLs compute velocity in different ways, and they have different measurement errors. While the loop detectors on this site currently estimate lane speeds from 5-minute flow and occupancy measurements and then compute the flow average of all lanes to obtain a single value, VTLs obtain the average of individual GPS-computed velocities.

9.6 Achieved Penetration Rate During Experiment

Penetration rate definition. Penetration rate of equipped vehicles refers to the proportion of equipped vehicles in the total flow. This proportion can be computed by placing VTLs on each of the 17 existing loop detector locations and dividing the VTL count by the loop detector count every 5 minutes.

Continuous flow of probe vehicles achieved. During the morning, less than 3% of the 5-minute periods have no VTL observations, and in the afternoon that number decreases to less than 1%. In addition, 50% of the 5-minute periods in the morning have a penetration rate of at least 2%, while in the afternoon only 35% of the periods meet this condition. This suggests that a steady flow of equipped vehicles was achieved in which at least one probe vehicle crossed an ILD in each 5-minute period at each location.

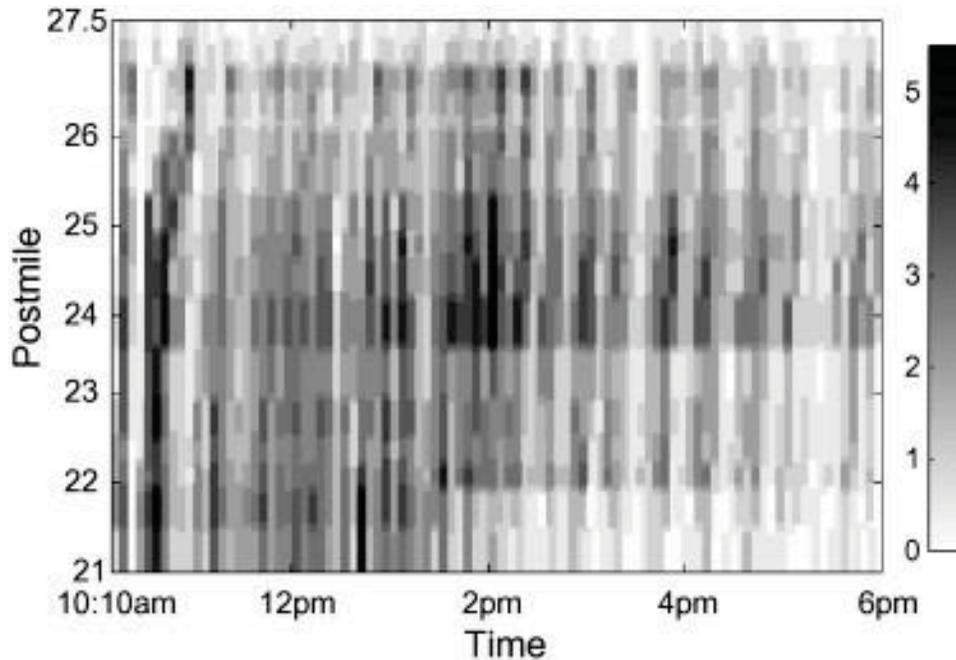


Figure 9.8: Penetration rate map

Penetration rate map. During the experiment, the penetration rate changed over time and space, as shown in the penetration rate map of Figure 9.8. In the morning, all three probe vehicle routes passed between mileposts 21 and 26 (Decoto Rd., and Tennyson Rd.) In the afternoon, all probe vehicle routes passed between mileposts 23.3 and 26 (Alvarado-Niles Rd. and Tennyson Rd.) As a result these sections experienced the highest penetration rate (i.e., proportion of equipped vehicles) during the day. Locations at the edges of the study section, such as downstream of milepost 27 (between CA-92 and Winton Ave.), were visited by only one third of the equipped vehicles during the whole day and thus experience the lowest penetration rates.

Penetration rate histogram. The penetration rates for all locations between Decoto Rd. and Winton Ave. are displayed in Figure 9.9. Circles in (a) and (c) of the figure represent the average penetration rate along the study section during the morning and the afternoon, respectively. The range corresponds to one standard deviation around the mean. The histograms in (b) and (d) show the penetration rates measured at all of the 17 locations during the morning periods, and afternoon periods, respectively.

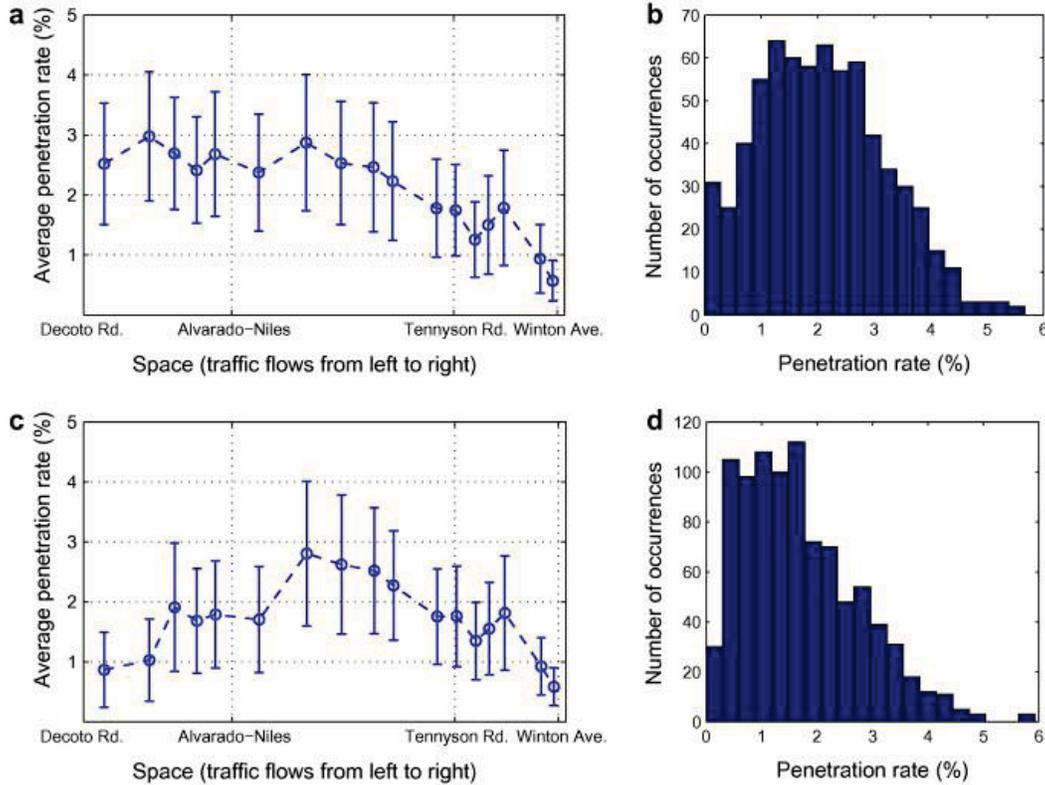


Figure 9.9: (a) and (c): Average penetration rate over time at existing detector station locations during the morning and the afternoon. The range is one standard deviation below and over the mean. Traffic flows from left to right. (b) and (d): Histogram of the penetration rate including all the 17 locations during the morning routes and the afternoon routes, respectively.

9.7 Inferring parameters from shockwave speed

This section shows how the information about the propagation of shockwaves (illustrated in Figure 9.2) can be used to infer parameters of the fundamental diagram (assuming a triangular relationship), as well as flows and densities that mobile sensors are not able to capture directly.

We start by assuming that a vehicle spans s_J feet when stopped at a traffic jam. Therefore, $k_J = \frac{5280}{s_J}$ the jam density is vpmpl (5280 is a unit conversion factor). For instance, $s_J = 26$ ft (8 meters) yields a jam density around $k_J = 200$ vpmpl. This can be seen as a standard value for jam density. The other two parameters needed to fully characterize the triangular fundamental diagram correspond to the free flow speed v_f and the shockwave speed w , which are obtained from the data.

The free flow speed corresponds to the speed of the vehicles before or after the incident (v_f is 65 mph in Figure 9.2). The shockwave speed is the speed of the second wave traveling

upstream in Figure 9.2 (the steepest red line in the figure, which has been manually drawn by connecting the points where vehicles approximately change velocity), which is $w = -15.6$ mph. With this information and some basic geometry, we can conclude that the critical density and the maximum flow are around $k_C = 40$ vpmpl ($k_C = \frac{wk_J}{w-v_f}$) and $q_{\max} = 2570$ vphpl ($q_{\max} = k_C \cdot v_f$), respectively.

From the data, velocity in the queue can also be obtained. Most of the speeds range from 3 mph to 7 mph, although few vehicles with speed in the order of 12 mph can be found. The difference in the speed among vehicles can be explained by the lane used by each vehicle. An average value of $v_{\text{queue}} = 6$ mph can be used for the speed in the queue. Using the triangular fundamental diagram obtained before, the speed in the queue is sufficient to characterize the traffic state in the queue, which in this case correspond to $q_{\text{queue}} = 867$ vphpl and $k_{\text{queue}} = 144$ vpmpl. The flow is very close to the flow reported by PeMS using loop detectors, which is 850 vphpl.

This information can be used to infer the flow before the accident occurred using the Rankine-Hugoniot condition, which relates the speed of the shockwave u_s (inferred from the data) with the flows and densities at both sides of the shockwave. Since the shockwave is traveling backwards, the state in front of the shockwave corresponds to the state before the accident happens, and the state behind the shockwave is the queued state obtained previously.

Therefore, we have:

$$u_s = \frac{q_{\text{queue}} - q_{\text{front}}}{k_{\text{queue}} - k_{\text{front}}} \quad (9-1)$$

Both the flow and density before the accident (that is, in front of the shockwave) q_{front} and k_{front} , respectively, can be obtained using equation (9-1) and knowing that $q_{\text{front}} = k_{\text{front}} \cdot v_f$. In this case, and speed of the first shockwave (drawn in the same way as the previous one) is $u_s = -3.6$ mph, which yields a traffic state with a flow $q_{\text{front}} = 1300$ vphpl and $k_{\text{front}} = 20$ vpmpl. The value for flow obtained in this way is similar to the flow of 1100 vphpl collected with loop detectors before the accident.

We started by assuming a specific jam density based on the space used by vehicles when fully stopped ($s_J = 26$ feet). If different values of jam density are tried, the results will change but still be valuable. For instance, for $k_J = 175$ vpmpl ($s_J = 30$ feet), $q_{\text{queue}} = 760$ vphpl and $q_{\text{before}} = 1140$ vphpl; for $k_J = 225$ vpmpl ($s_J = 23$ feet), $q_{\text{queue}} = 975$ vphpl and $q_{\text{before}} = 1465$ vphpl. However, the flows are still reasonably close to the flows measured with loop detectors.

Considering that the flows are obtained using data from an unknown proportion of the total flow, the information is valuable.

10 Revisiting the Density Based Methods

This chapter revisits the performance assessment of Newtonian Relaxation and Kalman Filtering introduced in Chapter 5. An implementation of both algorithms is investigated using the data set from the Mobile Century experiment (presented in Chapter 9), and a from PeMS data²². It is worthwhile to contrast the focus of this chapter with that of the previous one. Whereas Chapter 9 examined the raw data from the Mobile Century experiment, this chapter investigates assimilation techniques that can be used to enhance that data.

For operational reasons, lanes 1 to 4 (lane 1 being the leftmost lane) on the section between Decoto Rd. and Winton Ave. in the NB direction were considered. The period of time analyzed starts at 10am and ends at 6pm. The 6.5 miles were divided into $I = 44$ cells of $\Delta x = 780$ ft each, and the 8-hour period of meaningful data was divided into $H = 3600$ time steps of $\Delta t = 8$ seconds each. As in the NGSIM case (in Chapter 5), parameters of the fundamental diagram were extracted from PeMS²³.

10.1 Different Scenarios

The N95 Nokia phones used during the Mobile Century experiment stored position and velocity logs every 3 seconds. The penetration rate P is given by the number of smart phones on the road. During the day, this rate was sustained near 3% of the total flow. Using this data, we changed the values of T to construct different scenarios following the method outlined in the previous section ($\tau = 2$ for all three scenarios, where τ corresponds to the number of logs over which the average speed is computed). Table 10.1 shows the scenarios investigated in this case and the total number of Lagrangian measurements per lane-mile per hour that each one produces.

The parameter choice for X_{nudge} and T_d is the same as for the NGSIM case (180 ft and 15s, respectively). The values of T_a chosen for this experiment are 10s for the first two scenarios and 30s for the third one.

²² California Performance Measurement System: <http://pems.eecs.berkeley.edu/>

²³ Assuming a triangular shape (and using the same notation as before): $q_{\max} = 2275$ vphpl, $k_j = 152$ vpmp, $k_c = 35$ vpmp, $v_f = 65$ mph, and $w = -19.4$ mph.

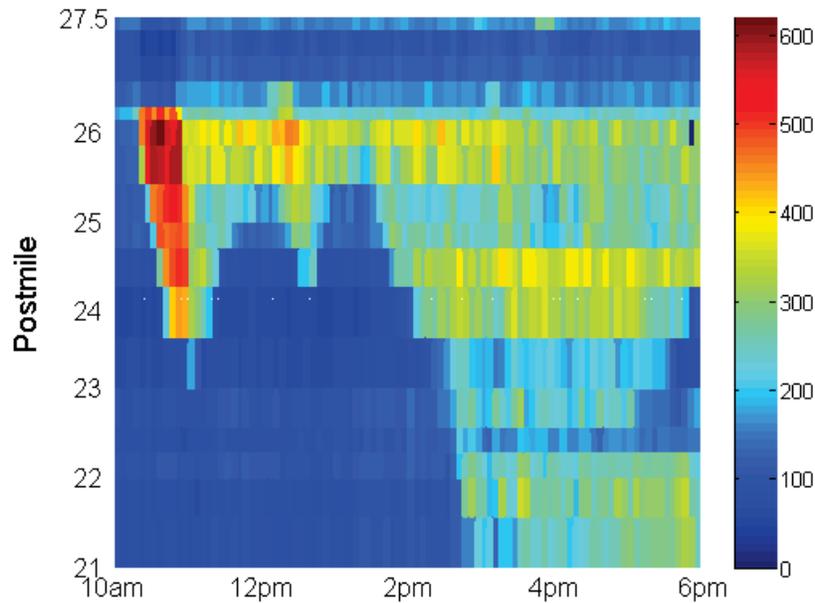


Figure 10.1; Density field (in vpm) using 17 loop detector stations deployed in the section of interest (obtained through PeMS).

10.1 Results

Unlike for the NGSIM data, in this case we only have access to trajectories of the equipped subset of vehicles. Therefore, all trajectories and vehicle accumulation ground truth are not known. To evaluate the results we use information collected by 17 loop detector stations installed along the section of interest using the PeMS system.

Figure 10.1 shows the density field produced by PeMS. The density field collected by PeMS captures the morning accident at milepost 26 (between Tennyson Rd and CA92). It also shows the spatial and temporal extent of the queue that formed upstream. The situation at this location does not fully recover, and the bottleneck remains active until the evening. A short wave propagates upstream for about 2 miles at 12:30pm. Congestion starts at 2pm, and by 3pm the entire section between milepost 21 (Decoto Rd) and 26 exhibit serious levels of congestion. Note that the severity of congestion around milepost 24 (between Whipple Rd. and Industrial Blvd.) is higher than for some other locations, suggesting the presence of a second active bottleneck in series at this milepost.

Table 10.1: Scenarios investigated using data from field experiment.

Case	T (sec)	# of Lagrangian measurements per mile-lane per minute
1	300	0.2
2	60	1.1
3	10	5.1

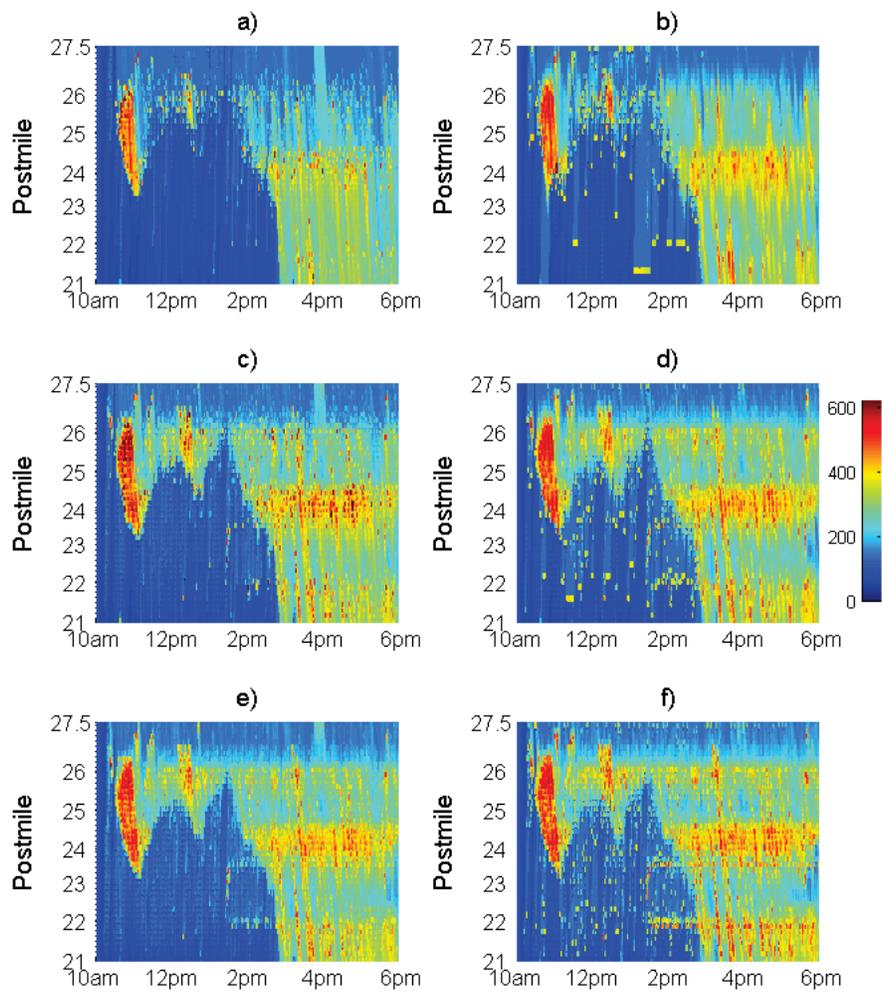


Figure 10.2: Density field (in vpm) using the Newtonian relaxation method (left) and the Kalman Filtering techniques (right) for scenario 1 (top), 2 (middle), and 3 (bottom). For each scenario, the boundary data is provided by loop detectors.

Figure 10.2 shows estimated density for each scenario in Table 10.1 using NR (left) and KF (right). Recall that each scenario only makes use of only two detector stations (at both boundaries of the section) in addition to the Lagrangian data provided by the cell phones. The main congested pattern observed from detectors is well captured in the three scenarios investigated and for the two methods. All the scenarios captured the accident at milepost 26 at 10:30am, even though not all the equipped vehicles were on the highway by then (vehicles were released between 10-10:30am). The location, intensity and duration of the incident are well replicated in each scenario. The short wave at 12:30pm is also observed on the three scenarios, but with different duration and intensity. The main congestion that starts at 2pm can be seen in the three scenarios as well. The more severe congestion observed during the afternoon from detectors around milepost 24 is also observed from Lagrangian data.

Figure 10.2 also shows that KF results appear noisier than the NR results, especially during the free flow conditions. This can be seen by the yellow points observed in the graphs upstream of milepost 23 and before 3 pm, which correspond to observations with high density. In these cases the effect is visually magnified by the fact that traffic is freely flowing in the surroundings. This difference between NR and KF can be explained by the way in which each method uses each observation. While NR distributes the effect of the observation among several cells, KF directly affects only the observed cell. This means that the addition of vehicles in the KF is concentrated mainly in the observed cell.

In summary, there is a good agreement between the density field provided by PeMS and the reconstruction using Lagrangian data. As expected, the agreement with PeMS data improves with the number of Lagrangian observations.

Recall, the value chosen for k^{FF} in equation (5-10) was arbitrarily set to 25 vpmpl. Different values of k^{FF} change the estimates, but not significantly. The main difference observed for different values of k^{FF} occurs during the transition between free flow and congested periods. Cases with higher values of k^{FF} tend to exhibit congestion more easily than lower values.

Note that the estimates shown in Figure 10.2 make use of only two detector stations located at both boundaries of the section. However, in the absence of detector data at the boundaries of the section, the results are still very accurate. For instance, we can assume that the Eulerian boundary data is also provided by equipped vehicles. To this end, the concept of *virtual trip lines* (VTL) introduced in [53] is useful. A VTL can be thought of a virtual loop detector that records speed of equipped vehicles as they cross it. In our case, every 5 minutes a VTL at each boundary of the section of interest computes the average speed from GPS measurements from different vehicles. Obviously, the more equipped vehicles per interval, the more representative the estimate of the speed should be. Knowing the velocity, the fundamental diagram can be used to compute the density at the boundaries, which is the input to the model. When VTLs are used to provide boundary data the results do not change significantly. There is still a good agreement between the density field from PeMS and the estimates for each scenario. In fact,

when boundary data is assumed to be totally unknown and some reasonable but constant value is assumed, the qualitative results still agree well with those obtained from PeMS.

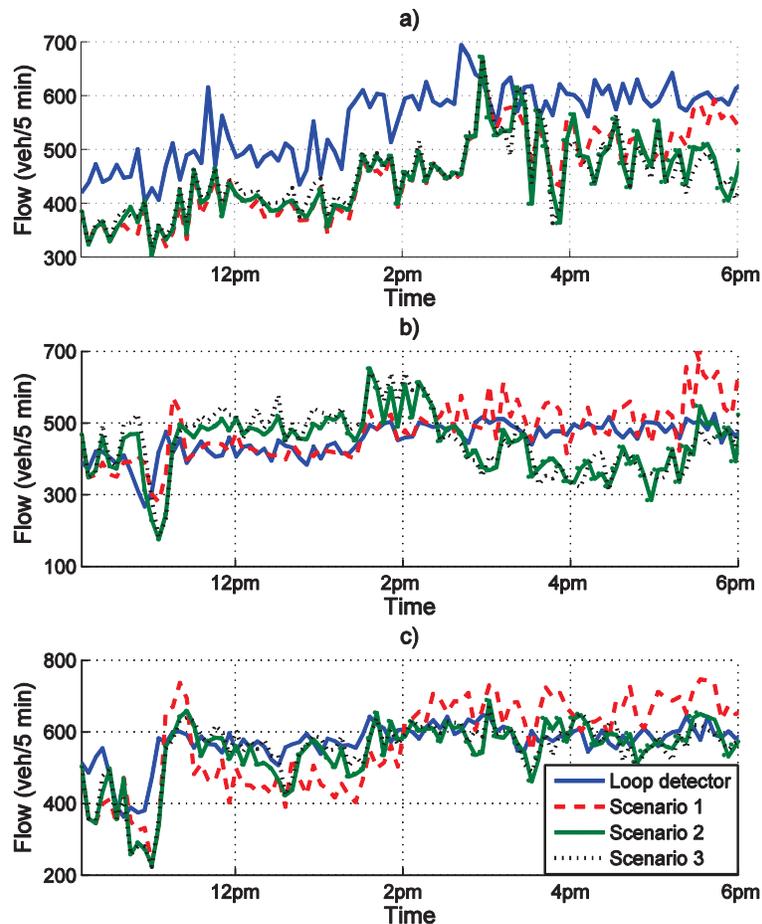


Figure 10.3: Flow comparison at mileposts (a) 21.3 (detector 1), (b) 24 (detector 7), and (c) 25.2 (detector 10).

Finally, the flows between consecutive cells can also be used for comparison. Figure 10.3 shows a flow comparison at three existing loop detector locations. Each graph shows the flow measured by the corresponding loop detector and the three flows yielded by each scenario using NR. Given that the correction terms in both methods are in part correcting the error in the flows between cells and the measurement error in loop detectors, a perfect agreement between the estimated flows and the flows measured by loop detectors cannot be expected. Despite the previous reason, and the fact that only two detectors located 6.5 miles from each other are being used in addition to the Lagrangian data, the estimated flows are in reasonable agreement with the flows measured by the detectors.

For this data, ground truth density is not known. The density field obtained using the 17 loop detector stations shown in Figure 10.1 is available. Measurements from these stations, however, have errors. For this reason, a quantitative analysis cannot use loop detector measurements as ground truth. Moreover, loop detectors measure the occupancy, which is converted into density using some approximations.

The procedure adopted to assess quantitatively the results consists of using 20% of the trajectories obtained in the Mobile Century experiment to provide ground truth velocity. The ground truth velocity is computed as the average of all the velocity logs in the corresponding cell at the corresponding time step. Only 20% of the trajectories, however, are not enough to provide a velocity measurement for every cell at every time step. Therefore, velocity ground truth is only known for some cells at some time steps, and the comparison can only be performed at these cells and time steps (around 10% of all the cell and time step points in the grid).

Ground truth is in terms of velocity, but the output of the methods is an estimate of the density field. The estimated velocity field is obtained using the fundamental diagram and by reversing the process described in Section 5.2.3. Note that for traffic state estimation using Lagrangian data, the other 80% of the trajectories are used in the estimation, reducing by 20% the number of observations for all the scenarios.

The metric used to assess the performance of the methods corresponds to the proportion of the estimates with less than μ mph of absolute difference with respect to the corresponding ground truth measurement, denoted by M_μ . The value for μ ranges between 5 and 20 mph.

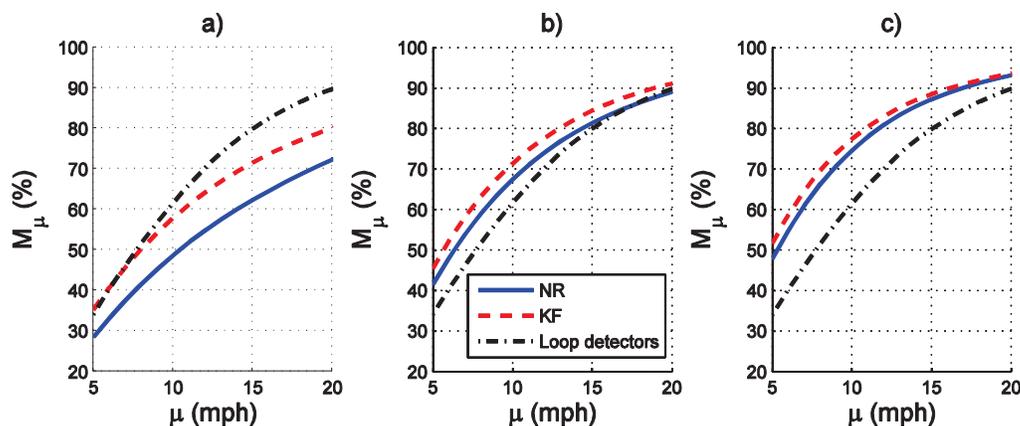


Figure 10.4: Results of the quantitative analysis for the Newtonian relaxation method and the Kalman Filtering technique for scenario 1 (left), 2 (center), and 3 (right). Results obtained using loop detector data are also included for comparison.

Figure 10.4 shows the results. Note that the values presented are the average obtained from 20 realizations, where each realization considers a different set of vehicles in the 80% used for the estimation (and thus in the 20% used for the validation as well). For comparison purposes, results obtained using loop detector velocity measurements have also been included. For scenarios 2 and 3, velocity fields obtained with the methods proposed have a better agreement with the velocity field obtained from the validation set than the velocity field obtained using loop detector measurements from 17 stations. For high values of μ , all three methods tend to converge in these cases. When the number of observations is very low (like in scenario 1), the estimated velocity field using measurements from 17 loop detector stations is preferred. The results also suggest that Kalman Filtering techniques outperform the Newtonian relaxation method in all the scenarios investigated. Interestingly, both methods in the presence of Lagrangian data achieve similar (or even better) performance levels than the level achieved by a good coverage of loop detector stations.

10.2 Conclusions

Two methods (introduced in Sections 5.3.1 and 5.3.2) to integrate Lagrangian measurements into a traffic flow model to perform traffic state estimation have been assessed. Lagrangian measurements used in the estimation consist in position and velocity readings from GPS-enabled smartphones traveling on the road. The methods were tested on two data sets: the NGSIM data set for initial prototyping and validation purposes (as explained in Chapter 5), and the Mobile Century data set to explore more completely the results of the proof-of-concept system.

The main conclusions of this work are:

- The Newtonian relaxation method (NR) and the Kalman Filtering techniques (KF) appropriately incorporate the Lagrangian measurements into flow models for traffic state estimation purposes. This is true even for a low number of observations per lane-mile per minute. This statement validates the value of both methods for traffic state estimation in the presence of Lagrangian data.
- Lagrangian measurements from vehicles (i.e. speed) are converted into density using the fundamental diagram, which introduces some error. However despite this error, the methods proposed produces accurate estimates. When the actual density is obtained from Lagrangian measurements, the accuracy of the results improves.
- In both methods, the accuracy of the final estimate grows with the number of Lagrangian measurements. The improvement in the estimates reaches a saturation level for a high number of Lagrangian measurements. Marginal benefits are obtained from new observations when the number of measurements becomes large.
- In both implementations, KF yields better results than the NR. In the implementation using the NGSIM dataset (in Chapter 5), the measure of performance was the percentage of improvement of the root mean square error of the density field. This index for the KF is approximately 1.5 times the index of the NR. In the implementation

with the Mobile Century data, the measure of performance was the proportion of the estimates with less than certain level of absolute difference with respect to the corresponding ground truth measurement. The index for the KF is greater than the index for the NR, although the difference decreases when more Lagrangian observations are available. The differences observed between both methods can be explained by the fact that KF provides an optimal estimate in the sense of least squares, while the estimates obtained with the NR are not necessarily optimal. Note again that in the KF algorithm, a major assumption is made when the mode is identified directly from the hybrid system model. This assumption might explain why for these datasets the method performs better (in fact, in the case of the NGSIM dataset, no mode identification is required since the traffic is congested all the time).

- Even though having more observations is better, results from the NGSIM analysis, and using the Newtonian relaxation method, suggest that it is better to have less frequent sampling on more vehicles, than frequent sampling on few vehicles. This is explained by a better spatial and temporal coverage of the first strategy.
- The proposed method makes use of density measured at upstream and downstream boundaries. This information is assumed to be provided by loop detectors. However, when no detector data is available, GPS-enabled smartphones can provide this information without affecting significantly the performance of the method. Therefore, in the presence of Lagrangian data, loop detector stations are not essential for traffic state estimation.

The present work thus demonstrates the added value of Lagrangian information in the system. It also highlights the potential of this type of data by itself (self sufficient for traffic reconstruction at high enough penetration rates). It finally underlines the tremendous potential of this type of data assimilation: even without fixed detectors, this data has the potential of supplying transportation agencies with loop detector quality (and potentially above) data, at no additional public infrastructure cost.

10.2.1 Future Work

The nudging factor determines the desired effect of each Lagrangian measurement over time and space on the estimates of density. Potentially, a nudging factor that takes into account the way in which traffic information flows in different conditions could yield better results. For instance, an observation reporting congestion at location x_c at time t_c should influence locations upstream of x_c at times $t > t_c$. Based on this, a different functional form for the nudging factor can be investigated to determine its influence on the results.

The use of Kalman Filtering techniques to incorporate Lagrangian measurements in traffic state estimation as described here follow from the work of [107]. This approach requires the identification of the mode at each time step. Ideally, the mode should be directly observed from the observations collected. Since observations are sparse in time and space, however, the mode identification is a challenging task (in addition, the knowledge of the error covariances is

assumed). A heuristic can be developed to circumvent these issues, in which the gain to be used is not necessarily optimal but depends on the number and accuracy of the observations. This idea is being currently analyzed.

As mentioned before, the methods do not require the knowledge of ramp counts. It would be interesting to analyze if (and how) these counts can be derived from the density estimates obtained.

11 Discussion

Mobile Century proof-of-concept. The 100-vehicle deployment introduced in Chapter 3 and presented in detail in Chapter 8 of this report was conceived as a proof of concept for a traffic monitoring system based on GPS-enabled mobile phones. The experiment demonstrated the feasibility of the proposed system for real-time traffic monitoring, in which GPS-enabled mobile phones can be used as traffic sensors, providing their velocity at different points on the freeway. By design, an uninterrupted and dispersed flow of probes was sustained along the study site. This prototype system exploited the extensive coverage thus provided and the high accuracy in position and velocity measurements provided by commercially available GPS units.

New paradigm realistic in near future. An average penetration rate near 3% was achieved during the experiment, which is viewed as realistic in the near future, considering the increasing penetration of GPS-enabled cellular devices. It is expected that GPS-enabled cell phones will penetrate the market rapidly in the near future, and the quality of measurements will increase with the evolution of GPS technology itself, thus opening new opportunities for smartphone-based monitoring systems. A near-future scenario can be imagined (see Chapter 2) in which probe vehicles become the primary source for traffic data.

Reduced costs and increased coverage. The benefits of the new paradigm are manifold. First, it promises to reduce significantly the cost of collecting traffic information as compared to the status quo. Carefully designed, the collection of mobile probe data over existing cellular networks will only incur minimal incremental costs to the operations of those networks. Second, in addition to lowering costs, this method will augment coverage to thousands of miles of roads that are currently not instrumented with any kinds of sensors.

Benefits for traveling public and roadway operators. The opportunity for government agencies is significant: the availability of data will empower the traveling public with real-time access to current traffic conditions, while transportation operators will gain access to an unprecedented wealth of information to help them better manage road networks. The massive availability of traffic data at a lower cost will have important consequences for both the traveling public and roadway operators. The dissemination of traffic information will enable a form of system 'self-management', in which individual commuters can make informed travel decisions. Not only will each user benefit personally, but the entire driving community will enjoy more balanced loads across the road network. Roadway operators will also benefit tremendously from collecting rich performance data that can be turned into a wide range of strategies, from demand management to dynamic traffic control.

11.1 Conclusions

Added value of probe vehicles. As established in Chapter 9, raw data from probe vehicles alone was sufficient to infer traffic features, i.e., to construct an accurate velocity map over time and space. In addition, such data enable faster detection of incidents as was demonstrated in real-time during the experiment, when an incident in the morning was detected on the northbound segment of I-880. The incident appeared immediately on the monitors at the command center before its presence was published on the 511.org website.

Validation against video data. Much of the scientific merit of the present work, the assessment of GPS data quality, and the confidence in the accuracy of these results derive from the certainty afforded by video data. As described in Chapter 7, northbound vehicles were captured on videotape and re-identified by their license plate numbers. As a result, ground-truth travel times were recovered for a substantial fraction of traffic flow during the course of the experiment. In this context, notions of velocity and travel times must be viewed as distributions due to the heterogeneity of driver behavior on the highway. However, when the velocity fields produced from VTLs and loop detector data are integrated to estimate travel times, the travel times produced from VTLs are more likely to fall within one standard deviation of the mean travel time recovered from the videos.

Data quality assessment. From this result, we conclude that the GPS capabilities of present-day smartphones were adequate as a raw data source for the expressway section used in the present work (GPS accuracy is highly dependent on the number of satellites in view, and accuracy could be adversely affected in urban or natural canyons). However, there remains much work to be done in data filtering, and in defining requirements for a full-scale deployment.

VTL monitoring predicts travel times accurately. As presented in Chapter 9, a 3% penetration rate of probes provided better data for travel time prediction than that of the PeMS loop detectors spaced at an average distance of 0.35 mi. The detailed comparisons in Chapter 9 suggest the presence of some bias in the velocity estimation for some loop detectors, showing sometimes substantial differences with the VTL measurements.

VTL monitoring is low cost and maintains privacy. The sampling strategy presented in Chapter 4 is based on the use of VTLs, and provides enough data for traffic monitoring purposes while managing the privacy of participants. By its very nature, the proposed architecture separates identity information from traffic information. However, future effort is necessary to quantify more completely the risks and feasible privacy guarantees. In addition to the higher accuracy achieved with this technology, the proposed traffic monitoring system has other advantages over current systems based on loop detectors. From the standpoint of transportation agencies, for example, the system comes at almost no installation and maintenance cost.

VTL monitoring is flexible. In addition to the privacy benefits, a key advantage of virtual trip lines over physical traffic sensors is the flexibility with which they can be deployed. For example, when roadwork is performed, VTLs can be deployed throughout the construction region, providing accurate travel time estimates in an area which often creates significant congestion. Because there is almost no additional cost to deploy the VTLs, and they do not interfere with the construction work or the highway traffic, they can be placed to adjust to the temporarily changed traffic patterns. One could even envision a VTL placement strategy which changes on a much shorter time period, with optimal placement strategies [11] for the morning and evening rush hours, or holiday traffic patterns.

11.2 Challenges

Participation of traveling public is crucial for success. In order to create and maintain the desired service quality, a large number of participants must be recruited and sustained. Thus, when launching the service, either a large number of users must be recruited simultaneously, or the service must be bootstrapped with additional data. In the case of traffic monitoring, bootstrapping can be achieved by leveraging data from existing sensor networks in areas where they currently exist. Spatial coverage can increase as does the volume of data from the pool of users.

Incentives for participation are needed. The proliferation of GPS-equipped smartphones will not be an issue in future years. However, a full-scale, long-term deployment cannot be successful unless users are provided with continuing incentives to participate. The creation of self-sustaining business models is outside the scope of this work. However, these issues do present substantial challenges.

Premature deployment is counterproductive. One can imagine a worst case scenario in which a deployment fails for lack of public interest and participation. One prior version of a proposal document expresses interest in the potential data obtainable (through a probe vehicle deployment) in Caltrans District 3. While this district might be a suitable candidate for the final system, incremental deployments should occur only in highly instrumented regions where redundant sources of data are available for scientific evaluation of results. While the present work is extremely promising, scaling up something that began an academic experiment to be a highly reliable, 24/7 system will be a significant effort.

11.3 Future Goals

Mobile Millennium. Nokia, Navteq and UC Berkeley have proceeded with a field operational test that extends the Mobile Century system to an urban network. The Mobile Millennium field operational test consisted, in part, of free distribution of traffic software (such as the smartphone client presented in Chapter 4) to regular commuters, and the collection of traffic data (travel times) for several months, principally covering Northern California.

Fusion of data. A system that fuses both static (loop detectors) and mobile sensors (GPS-enabled mobile phones) would be expected to provide more accurate estimation of traffic (density and flow, as well as velocity) than would a system using only a single sensor type, as suggested by [115]. Besides real-time traffic monitoring, the data collected could also be used for traffic state estimation and/or planning purposes. The role of previously deployed sensing infrastructure should be examined holistically in the context of a future system. Augmenting PeMS, for example, is possible but requires future work.

Inverse modeling. Any future traffic information system would need to include inverse modeling and data assimilation algorithms (see Chapters 5 and 6) aimed at: (1) circumventing the potential deficiencies of available data, and (2) augmenting the value of the data by careful use of appropriate models. Any potential errors, inaccuracies, and/or biases discovered in the data would need to be characterized. Improved flow models of highway traffic and inverse modeling techniques would be applied to calculate such quantities as travel times on links or routes, robust ranges of arrival times, and variances in travel times along links or routes. The data shown in the present work are rich enough to extract these features, and methods for these calculations are the subject of ongoing work.

Arterial traffic modeling. In comparison to highway traffic, arterials come with additional challenges. The underlying flow physics that governs them is more complex (traffic lights, often with unknown cycles, intersections, stop signals, parallel queues). While the present work explicitly derives techniques to reconstruct traffic from VTL type data, such a reconstruction becomes harder for arterials. Also, while macroscopic flow models exist and can be used for the arterial network, their parameters are in general unknown or inaccessible and only documented for few cities, making their use impossible without going to the field and measuring them. In addition, even if they were known, the complexity of the underlying flows makes it challenging to perform estimation of the full macroscopic state of the system at low penetration rates. In light of these challenges, statistical approaches for characterizing a subset of the macroscopic state (for example travel times and aggregated speeds) may be a viable alternative to traffic flow modeling based on traffic reconstruction. This is an area of open research, and continued effort is necessary.

11.4 Steps Toward Deployment

Deployment path. Stated simply, any future deployment will require substantial research and development. We recommend an evolutionary progression of field operational tests, so that lessons learned during any particular iteration may be incorporated into subsequent efforts. In addition, we recommend a deployment path that involves a gradual and continuous hand-off from the research team to an industrial product development group. The Mobile Millennium FOT will be a precursor to a future commercial service, and its success will depend on both the technical parameters and the business model. However, development of an industrial version of the mobile probe data collection platform will by no means signal the end of the research program. Rather, both development tracks will coexist and complement one another. The research track will continue to spearhead systems development by designing advanced data

processing techniques and adding features. The commercial track will trail behind by several months, but will also bring invaluable feedback to instruct and steer further studies. For instance, the correct interpretation of probe data collected on signalized arterials is likely to require substantial research efforts. Initially, it may be acceptable to launch a crude version among a selected group of participants (beta testers) while further investigations are conducted. Eventually, improved algorithms can be rolled into commercial services.

Traffic server development. A substantial information technology infrastructure must be implemented to support the computational modeling requirements of a scaled-up system. In addition to interstate expressways, we intend to monitor urban arterials and secondary highways as well. A considerable data storage infrastructure alongside a set of analysis and prototyping tools will form the foundations of this system. These tools will be used on an ongoing basis to evaluate the relevance of the data being collected, the performance of the real-time algorithms, and the development and calibration of novel traffic data processing techniques, including the fusion of mobile probe and fixed sensor data. A large part of this task will consist in designing, implementing, and operating live feeds and a database for location and speed data, adding automated processing capabilities as well as analysis and visualization functions.

Leverage trends in mobile computing. Any future deployment effort must proceed in a way that complements existing trends in mobile computing. Vehicular applications combining cellular and other wireless technologies with GPS are enabled by the growing ubiquity of wireless communications that enable connections to vehicles in either broadcast (e.g., FM, digital, and satellite radio) or point-to-point modes (e.g., cellular, e-tolling, WiFi). The development of multi-media platforms has created a new software-hardware environment in which a single device can encompass dozens of functionalities. Ultimately, increasingly integrated car accessories and personal handsets will spur convergence of multiple wireless technologies, opening even more applications. Key mobility goals can be readily served by embracing prevalent wireless services, and through policies that foster public-private partnerships in this milieu.

Industry collaboration. We recommend that future work have a strong industry component. Our partnership with Nokia ensures that future field operational tests will benefit from this collaboration. Nokia's deep commitment toward ongoing development in this field speaks to the short-term commercial potential of using GPS-equipped mobile phones as traffic probes. Ultimately, it is the role of industry partners to turn our joint research and development effort into a robust service sustained by a viable business model. Yet each entity (academic, industrial, and governmental) contributes its own strength and carries a unique responsibility toward that goal.

Data sharing modalities. Future work as a part of this program has direct application toward the strategic goals of Caltrans in the area of operational data collection; therefore, long-term synergies should be developed. In particular, it will be critical to explore the data sharing modalities that can be agreed upon between industrial companies, Caltrans, and local public

agencies. In addition, a mobile probe system could be expanded to other modes of travel, including public transportation and even pedestrian traffic.

Applicability to future practice and information dissemination. Traffic data collected as part of the next iteration of this program will be processed to generate traveler information and served back to the mobile phone users who originally generated the data. In future iterations, that same information could be disseminated much more broadly to feed regular traveler information channels, including broadcast media, internet websites, Personal Navigation Devices (PND) and other mobile handsets, as well as roadside Changeable Message Signs (CMS). This probe data may also be fused with existing traffic data from fixed sensors. Ultimately, this application of technology has the potential to alleviate traffic congestion, improve traffic data and traveler information, and reduce unnecessary fuel consumption.

References

- [1] <http://bard.wr.usgs.gov/>.
- [2] <http://jcp.org/en/jsr/detail?id=179/>.
- [3] <http://www.paramics-online.com>.
- [4] <http://www.privacyrights.org/ar/ChronDataBreaches.htm>.
- [5] CROWDSOURCING. <http://googleblog.blogspot.com/2009/08/bright-side-of-sitting-in-traffic.html>, 2009.
- [6] D. AGRAWAL and C. C. AGGARWAL. On the design and quantification of privacy preserving data mining algorithms. In *Symposium on Principles of Database Systems*, 2001.
- [7] R. AGRAWAL and R. SRIKANT. Privacy-preserving data mining. In *Proc. of the ACM SIGMOD Conference on Management of Data*, pages 439–450. ACM Press, May 2000.
- [8] R. A. ANTHES. Data assimilation and initialization of hurricane prediction models. *J. Atmos. Sci.*, 31:702–719, 1974.
- [9] A. AW and M. RASCLE. Resurrection of “second order” models of traffic flow. *SIAM Journal on Applied Mathematics*, 60(3):916–938, 2000.
- [10] X. BAN, R. HERRING, J. MARGULICI, and A. BAYEN. Optimal sensor placement for freeway travel time estimation. In *18th International Symposium on Transportation and Traffic Theory (ISTTT)*, pages 697–721, 2009.
- [11] X.J. BAN, L. CHU, and H. BENOUAR. Bottleneck identification and calibration for corridor management planning. *Transportation Research Record*, 1999:40–53, 2007.
- [12] C. BARDOS, A. Y. LEROUX, and J. C. NEDELEC. First order quasilinear equations with boundary conditions. *Communications in partial differential equations*, 4(9):1017–1034, 1979.
- [13] H. BAR-GERA. Evaluation of a cellular phone-based system for measurements of traffic speeds and travel times: A case study from Israel. *Transportation Research C*, 15(6):380–391, 2007.
- [14] Y. BAR-SHALOM and X.R. LI. Estimation and tracking- Principles, techniques, and software. Norwood, MA: Artech House, Inc., 1993.
- [15] A. BERESFORD and F. STAJANO. Mix zones: User privacy in location-aware services. In *IEEE PerSec*, 2004.

- [16] R. BERTINI, M. LASKY, and C. MONSERE. Validating predicted rural corridor travel times from an automated license plate recognition system: Oregon's frontier project. In *Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems*, pages 706–711, Vienna, Austria, 2005.
- [17] R. BERTINI and S. TANTIYANUGULCHAI. Transit buses as traffic probes: empirical evaluation using geo-location data. *Transportation Research Record: Journal of the Transportation Research Board*, 1870:35–45, 2004.
- [18] S. BLANDIN and A. BAYEN. A first order hyperbolic partial differential equation formulation of speed evolution on highways. in preparation for Transportation Research B., 2009.
- [19] G. BURGERS, P. VAN LEEUWEN, and G. EVENSEN. Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126(6):1719–1724, 1998.
- [20] M. CASSIDY. Bivariate relations in nearly stationary highway traffic. *Transportation Research B*, 32(1):49–59, 1998.
- [21] R. CAYFORD and T. JOHNSON. Operational parameters affecting the use of anonymous cell phone tracking for generating traffic information. *Transportation Research Board 82nd Annual Meeting*, 1(3):03–3865, 2003.
- [22] L. CHU, S. OH, and W. RECKER. Adaptive Kalman filter based freeway travel time estimation. In *84th TRB Annual Meeting*, Washington D.C., January 9-13 2005. Transportation Research Board.
- [23] K. CHUNG, J. RUDJANAKANOKNAD, and M. CASSIDY. Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research B*, 41(1):82–95, 2007.
- [24] B. COIFMAN. Improved velocity estimation using single loop detectors. *Transportation Research Part A*, 35(10):863–880, 2001.
- [25] C. DAGANZO. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research B*, 28(4):269–287, 1994.
- [26] C. DAGANZO. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B*, 28(4):269–287, 1994.
- [27] C. DAGANZO. The cell transmission model, part II: network traffic. *Transportation Research Part B*, 29(2):79–93, 1995.
- [28] C. DAGANZO. The cell transmission model, Part II: Network traffic. *Transportation Research B*, 29(2):79–93, 1995.

- [29] X. DAI, M. FERMAN, and R. ROESSER. A simulation evaluation of a real-time traffic information system using probe vehicles. In *Proceedings of the IEEE Intelligent Transportation Systems*, pages 475–480, 2003.
- [30] A. DEMERS, G. F. LIST, W. A. WALLACE, E. E. LEE, and J. M. WOJTOWICZ. Probes as path seekers: a new paradigm. *Transportation Research Record: Journal of the Transportation Research Board*, 1944:107–114, 2006.
- [31] R. DINGLEDINE, N. MATHEWSON, and P. F. SYVERSON. Tor: The second-generation onion router. In *USENIX Security Symposium*, pages 303–320, 2004.
- [32] L. EDIE. Discussion on traffic stream measurements and definitions. In *Proc. of the Second Int. Symp. on the Theory of Traffic Flow*, pages 139–154, Paris, France, 1965.
- [33] L. EVANS. *Partial differential equations*. American Mathematical Society, Providence, USA, 1998.
- [34] G. EVENSEN. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5):10143–10162, 1994.
- [35] G. EVENSEN. The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, 2003.
- [36] G. EVENSEN. *Data Assimilation: The Ensemble Kalman Filter*. Springer-Verlag, Secaucus, NJ, 2006.
- [37] M. FERMAN, D. BLUMENFELD, and X. DAI. A simple analytical model of a probe-based traffic information system. In *Proceedings of the IEEE Intelligent Transportation Systems*, pages 263–268, 2003.
- [38] M. FONTAINE and B. SMITH. Investigation of the performance of wireless location technology-based traffic monitoring systems. *Journal of Transportation Engineering*, 133(3):157–165, 2007.
- [39] A. GAL and V. ATLURI. An authorization model for temporal data. In *Proceedings of the 7th ACM CCS*, pages 144–153, New York, NY, USA, 2000. ACM Press.
- [40] D. GAZIS and C. KNAPP. On-line estimation of traffic densities from time-series of flow and speed data. *Transportation Science*, 5(3):282–301, 1971.
- [41] B. GEDIK and L. LIU. Location privacy in mobile systems: A personalized anonymization model. In *Proceedings of the 25th IEEE ICDCS'05*, pages 620–629, Washington, DC, USA, 2005. IEEE Computer Society.

- [42] S. GODUNOV. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306, 1959.
- [43] D. GOLDSCHLAG, M. REED, and P. SYVERSON. Onion routing for anonymous and private internet connections. *Communications of the ACM (USA)*, 42(2):39–41, 1999.
- [44] G. GOMES and R. HOROWITZ. Optimal freeway ramp metering using the asymmetric cell transmission model. *Transportation Research C*, 14(4):244–262, 2006.
- [45] B.D. GREENSHIELDS. A study of traffic capacity. *Highway Research Board*, 14:448–477, 1935.
- [46] M. GRUTESER and D. GRUNWALD. Anonymous usage of location-based services through spatial and temporal cloaking. In *Proceedings of the ACM MobiSys*, 2003.
- [47] M. GRUTESER and B. HOH. On the anonymity of periodic location samples. In *Proceedings of the Second International Conference on Security in Pervasive Computing*, 2005.
- [48] H. HAJ-SALEM. and J.P. LEBACQUE. Reconstruction of false and missing data with first-order traffic flow model. *Transportation Research Record*, 1802:155–165, 2002.
- [49] F. HALL and B. PERSAUD. Evaluation of speed estimates made with single-detector data from freeway traffic management systems. *Transportation Research Board*, Transportation Research Record 1232:9–16, 1989.
- [50] A.W. HEEMINK, M. VERLAAN, and A.J. SEGERS. Variance reduced ensemble Kalman filtering. *Monthly Weather Review*, 129:1718–1728, 2001.
- [51] J.C. HERRERA and A.M.BAYEN. Traffic flow reconstruction using mobile sensors and loop detector data. In *87th TRB Annual Meeting*, Washington D.C., January 12-16 2008. Transportation Research Board.
- [52] HICOMP. Highway congestion monitoring program. <http://www.dot.ca.gov/hq/traffops/sysmgtp/HICOMP/index.htm>, 2007.
- [53] B. HOH, M. GRUTESER, R. HERRING, J. BAN, D. WORK, J.-C. HERRERA, A. BAYEN, and Q. JACOBSON. Virtual trip lines for distributed privacy-preserving traffic monitoring. In *6th International Conference on Mobile Systems, Applications, and Services*, pages 15–28, Breckenridge, CO, June 17-18 2008.
- [54] B. HOH, M. GRUTESER, H. XIONG, and A. ALRABADY. Enhancing security and privacy in traffic-monitoring systems. *IEEE Pervasive Computing*, 5(4):38–46, 2006.

- [55] B. HOH, M. GRUTESER, H. XIONG, and A. ALRABADY. Preserving privacy in gps traces via uncertainty-aware path cloaking. In *Proceedings of ACM CCS 2007*, October 2007.
- [56] <http://www.waze.com/>, 2010. Waze.
- [57] B. HULL, V. BYCHKOVSKY, Y. ZHANG, K. CHEN, M. GORACZKO, A. K. MIU, E. SHIH, H. BALAKRISHNAN, and S. MADDEN. CarTel: A Distributed Mobile Sensor Computing System. In *4th ACM SenSys*, Boulder, CO, November 2006.
- [58] INRIX. <http://www.inrix.com/>, 2006. Inrix.
- [59] INTELLIONE. Intellione. <http://www.intellione.com/>, 2006.
- [60] Y. ISHIKAWA, T. AWAJI, and K. AKIMOTO. Successive correction of the mean sea surface height by the simultaneous assimilation of drifting buoy and altimetric data. *Journal of Physical Oceanography*, 26:2381–2397, 1996.
- [61] D. JACQUET, C. CANUDAS DE WIT, and D. KOENIG. Traffic control and monitoring with a macroscopic model in the presence of strong congestion waves. In *Proc. of the 44th IEEE Conference on Decision and Control, and European Control Conference*, pages 2164–2169, Sevilla, Spain, 2005.
- [62] D. JACQUET, M. KRSTIC, and C. CANUDAS DE WIT. Optimal control of scalar one-dimensional conservation laws. In *25th American Control Conference, 2006 Minneapolis, MN, USA*, pages 5213–5218, 2006.
- [63] M. JAIN and B. COIFMAN. Improved speed estimates from freeway traffic detectors. *ASCE Journal of Transportation Engineering*, 483(7), 2005.
- [64] Z. JIA, C. CHEN, B. COIFMAN, and P. VARAIYA. The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors. In *4th IEEE Conference on Intelligent Transportation Systems*, Oakland, CA, August 2001.
- [65] T. JIANG, H. WANG, and Y. HU. Preserving location privacy in wireless lans. In *Proceedings of the 5th ACM MobiSys*, New York, NY, USA, 2007. ACM Press.
- [66] J. KAIPIO and E. SOMERSALO. *Statistical and Computational Inverse Problems*. Springer, New York, NY, 2005.
- [67] J. KAIPIO and E. SOMERSALO. Statistical inverse problems: Discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics*, 198(2):493–504, 2007.
- [68] R.E. KALMAN. A new approach to linear filtering and prediction problems. *Transactions of the ASME Journal of Basic Engineering*, 82:35–45, 1960.

- [69] P. KALNIS, G. GHINITA, K. MOURATIDIS, and D. PAPADIAS. Preventing location-based identity inference in anonymous spatial queries. *IEEE Transactions on Knowledge and Data Engineering*, 19(12):1719–1733, 2007.
- [70] H. KARGUPTA, S. DATTA, Q. WANG, and K. SIVAKUMAR. Random data perturbation techniques and privacy preserving data mining. In *IEEE ICDM*. IEEE Press, 2003.
- [71] T. KOHNO, A. BROIDO, and K. C. CLAFFY. Remote physical device fingerprinting. In *SP '05: Proceedings of the 2005 IEEE Symposium on Security and Privacy*, pages 211–225, Washington, DC, USA, 2005. IEEE Computer Society.
- [72] A. KRAUSE, E. HORVITZ, A. KANSAL, and F. ZHAO. Toward community sensing. In *ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, St.Louis, MO, 2008.
- [73] J. KRUMM. Inference attacks on location tracks. In *Fifth International Conference on Pervasive Computing (Pervasive2007)*, Toronto, Ontario, Canada, 2007.
- [74] J. KWON, B. MCCULLOUGH, K. PETTY, and P. VARAIYA. Evaluation of PeMS to improve the congestion monitoring program. California PATH Research Report UCB-ITS-PRR-2007-6, Institute of Transportation Studies, University of California, Berkeley, CA, 2007.
- [75] P. LEFLOCH. Explicit formula for scalar non-linear conservation laws with boundary conditions. *Mathematical methods in the applied sciences*, 10(3):265–287, 1988.
- [76] R. LEVEQUE. *Finite volume methods for hyperbolic problems*. Cambridge University Press, Cambridge, UK, 2002.
- [77] R.J. LEVEQUE. *Numerical Methods for Conservation Laws*. Birkhäuser Verlag, Basel, Switzerland, 1992.
- [78] M. LIGHTHILL. and G. WHITHAM. On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 229(1178):317–345, 1955.
- [79] M. LIGHTHILL and G. WHITHAM. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 229(1178):317–345, May 1955.
- [80] H. LIU, A. DANCZYK, R. BREWER, and R. STARR. Evaluation of cell phone traffic data in minnesota. *Transportation Research Record*, 2086:1–7, December 2008.
- [81] D. LOVELL. Accuracy of speed measurements from cellular phone vehicle location systems. *Journal of Intelligent Transportation Systems*, 6(4):303–325, 2001.

- [82] A. MACHANAVAJHALA, J. GEHRKE, D. KIFER, and M. VENKITASUBRAMANIAM. I-diversity: Privacy beyond k-anonymity. In *Proceedings of the 22nd International Conference on Data Engineering (ICDE'06)*, page 24, 2006.
- [83] B. MIKHALKIN, H. PAYNE, and L. ISAKSEN. Estimation of speed from presence detectors. *Highway Research Record*, 388:73–83, 1972.
- [84] M. F. MOKBEL, C. CHOW, and W. G. AREF. The new casper: query processing for location services without compromising privacy. In *Proceedings of the 32nd VLDB'2006*, pages 763–774. VLDB Endowment, 2006.
- [85] J. MOORE, S. CHO, and A. BASU. Use of Los Angeles Freeway Service Patrol vehicles as probe vehicles. Technical report, Berkeley, CA, 2001.
- [86] L. MUÑOZ, X. SUN, R. HOROWITZ, and L. ALVAREZ. Traffic density estimation with the cell transmission model. In *American Control Conference, 2003. Proceedings of the 2003*, volume 5, 2003.
- [87] C. NANTHAWICHIT, T. NAKATSUJI, and H. SUZUKI. Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway. *Transportation Research Record*, 1855:49–59, 2003.
- [88] O. OLEINIK. Discontinuous solutions of non-linear differential equations. *Uspekhi Matematicheskikh Nauk*, 12(3):3–73, 1957.
- [89] C. PANICONI, M. MARROCU, M. PUTTI, and M. VERBUNT. Newtonian nudging for a Richards equation-based distributed hydrological model. *Advances in Water Resources*, 26(2):161–178, 2003.
- [90] H. PAYNE and S. THOMPSON. Malfunction detection and data repair for induction-loop sensors using i-880 data base. *Transportation Research Board, Transportation Research Record* 1570(-1):191–201, 1997.
- [91] PEMS. Freeway Performance Measurement System. <http://pems.eecs.berkeley.edu/Public/>, 2010.
- [92] PNT. Information about the end of gps selective availability. <http://pnt.gov/public/sa/>, May 2010.
- [93] A. PUSHKAR, F. HALL, and J. ACHA-DAZA. Estimation of speeds from single-loop freeway flow and occupancy detectors using cusp catastrophe theory model. *Transportation Research Board, Transportation Research Record* 1457:149–157, 1994.

- [94] Z. QIU, P. CHEN, J. JING, and B. RAN. Cellular probe technology applied in advanced traveler information. *International Journal of Technology Management*, 2007. Accepted for publication.
- [95] QUADSTONE LIMITED. PARAMICS Modeller Version 5.1 Reference Manual. 2004.
- [96] D. REID. An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, 24(6):843–854, Dec 1979.
- [97] P. RICHARDS. Shock waves on the highway. *Operations Research*, 4(1):42–51, February 1956.
- [98] P.I. RICHARDS. Shock waves on the highway. *Operations Research*, 4(1):42–51, 1956.
- [99] P. SAMARATI and L. SWEENEY. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. In *Proceedings of IEEE Symposium on Research in Security and Privacy*, 1998.
- [100] K.K. SANWAL and J. WALRAND. Vehicles as probes. California PATH Working Paper UCB-ITS-PWP-95-11, Institute of Transportation Studies, University of California, Berkeley, CA, 1995.
- [101] J. SAU, N.E. EL FAOUZI, A. BEN ASSA, and O. DE MOUZON. Particle filter-based real-time estimation and prediction of traffic conditions. *Applied Stochastic Models and Data Analysis*, 12, 2007.
- [102] A. SCHWARZENEGGER, D. E. BONNER, W. KEMPTON, and R. COPP. State Highway Congestion Monitoring Program (HICOMP), Annual Data Compilation. Technical report, Caltrans, Sacramento, CA, 2008.
- [103] A. SEPPÄNEN, M. VAUHKONEN, E. SOMERSALO, and J. KAIPIO. State space models in process tomography – approximation of state noise covariance. *Inverse Problems in Engineering*, 9:561–585, 2001.
- [104] STL. Wireless Location Technology-Based Traffic Monitoring Demonstration and Evaluation Project, Evaluation Final Report. Technical report, Samrt Travel Laboratory, Center for Transportation Studies, University of Virginia, Blacksburg, VA, 2006.
- [105] I. STRUB and A. BAYEN. Weak formulation of boundary conditions for scalar conservation laws: an application to highway traffic modelling. *International Journal of Robust Nonlinear Control*, 16(16):733–748, 2006.
- [106] X. SUN, L. MUÑOZ, and R. HOROWITZ. Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control. In *Proc. of the 42nd IEEE Conference on Decision and Control*, volume 6, Maui, HI, 2003.

- [107] X. SUN, L. MUÑOZ, and R. HOROWITZ. Mixture Kalman filter based highway congestion mode and vehicle density estimator and its application. In *Proc. of the American Control Conference*, volume 3, pages 2098 – 2103, Boston, MA, 2004.
- [108] L. SWEENEY. Achieving k -Anonymity Privacy Protection Using Generalization and Suppression. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10(5):571–588, 2002.
- [109] M. SZETO and D. GAZIS. Application of Kalman filtering to the surveillance and control of traffic systems. *Transportation Science*, 6(4):419–439, 1972.
- [110] TELENAV. TeleNav. <http://www.telenav.net/>, 2004.
- [111] A. THIAGARAJAN, L. R. SIVALINGAM, K. LACURTS, S. TOLEDO, J. ERIKSSON, S. MADDEN, and H. BALAKRISHNAN. VTrack: Accurate, Energy-Aware Traffic Delay Estimation Using Mobile Phones. In *7th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Berkeley, CA, November 2009.
- [112] TRAFFICMASTER. <http://www.trafficmaster.co.uk/>, 2010.
- [113] M. TREIBER and D. HELBING. Reconstructing the spatio-temporal traffic dynamics from stationary detector data. *Cooperative Transportation Dynamics*, 1:3.1–3.24, 2002.
- [114] Y. WANG and M. PAPAGEORGIOU. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. *Transportation Research Part B*, 39(2):141–167, 2005.
- [115] M. WESTERMAN, R. LITJENS, and J-P. LINNARTZ. Integration of probe vehicle and induction loop data – estimation of travel times and automatic incident detection. PATH Research Report UCB-ITS-PRR-96-13, Institute of Transportation Studies, University of California, Berkeley, CA, 1996.
- [116] J. WRIGHT and J. DAHLGREN. Using vehicles equipped with toll tags as probes for providing travel times. California PATH Working Paper UCB-ITS-PWP-2001-13, Institute of Transportation Studies, University of California, Berkeley, CA, 2001.
- [117] J-L. YGNACE, C. DRANE, Y.B. YIM, and R. DE LACVIVIER. Travel time estimation on the San Francisco Bay Area network using cellular phones as probes. California PATH Working Paper UCB-ITS-PWP-2000-18, Institute of Transportation Studies, University of California, Berkeley, CA, 2000.
- [118] Y.B. YIM. The state of cellular probes. California PATH Working Paper UCB-ITS-PRR-2003-25, Institute of Transportation Studies, University of California, Berkeley, CA, 2003.

- [119] Y.B. YIM and R. CAYFORD. Investigation of vehicles as probes using global positioning system and cellular phone tracking: field operational test. California PATH Working Paper UCB-ITS-PWP-2001-9, Institute of Transportation Studies, University of California, Berkeley, 2001.
- [120] J. YOON, B. NOBLE, and M. LIU. Surface street traffic estimation. In *MobiSys '07: Proceedings of the 5th international conference on Mobile systems, applications and services*, pages 220–232, New York, NY, USA, 2007. ACM.
- [121] M. YOUSSEF, V. ATLURI, and N. R. ADAM. Preserving mobile customer privacy: an access control system for moving objects and customer profiles. In *Proceedings of the 6th MDM '05*, pages 67–76, New York, NY, USA, 2005. ACM Press.
- [122] H. ZHANG. A theory of nonequilibrium traffic flow. *Transportation Research Part B*, 32(7):485–498, 1998.
- [123] R. ZITO, G. D'ESTE, and M. TAYLOR. Global positioning systems in the time domain: how useful a tool for intelligent vehicle-highway systems? *Transportation Research C*, 3(4):193–209, 1995.

Appendices

1 CPHS Protocol Narrative Form

Committee for Protection of Human Subjects (CPHS) Office for the Protection of Human Subjects (OPHS) University of California Berkeley 2160 Shattuck Ave., Suite 515 Berkeley, CA 94704-6840		Phone: 510/842-7491 Fax: 510/842-6275 E-mail: ophs@berkeley.edu http://ohps.berkeley.edu	
CPHS PROTOCOL NARRATIVE FORM			
Instructions: Complete all applicable sections of this form. If requesting Eminent Status, see instructions on Eminent Request form. Please type, using a different font than the one in this form. Handwritten or incomplete forms will be returned. Use language that is clear, concise, and non-technical whenever possible, and define all acronyms. For revisions or amendments, highlight all changes from the previously approved version on one copy. A grant proposal or thesis will not be accepted in place of a protocol written according to this format.			
Lead Investigator:	Prof. Alexandre Bayen	CPHS #:	
Protocol Title:	Testing Cellular Phones as Traffic Probes	CPHS #:	
Related CPHS Title:		CPHS #:	
Project(s)? Title:	Deployment of Value-Added Mobile Traffic Probe	CPHS #:	
SECTION 1: PURPOSE AND BACKGROUND OF STUDY			
<ul style="list-style-type: none"> Purpose: Provide a brief explanation of the proposed research, including specific study hypothesis, objectives, and rationale. 			
<p>This experiment intends to provide a proof of concept for traffic-flow reconstruction from probe vehicle measurements. To this end, 100 drivers will be asked to drive on a stretch of highway. Their vehicles will be equipped with GPS devices that send information such as speed and position to a server gathering all the data. These probe vehicles, driving in general traffic conditions, will provide local traffic information. The scientific challenge is to reconstruct the profile of the whole flow from these mobile measurements. This has been achieved in simulation so far, and advancing theoretical developments lead us to believe that this can be done in experimental conditions.</p>			
Study Hypothesis			
<p>The study hypothesis is that the section of highway and the time-range of the experiment are subjected to representative traffic patterns. It is also assumed that traffic theory will apply to the data obtained during the experiment.</p>			
Objectives			
<p>The objective of the study is to obtain a set of GPS measurements from the highway for free-flow and congestion conditions and test different algorithms to prove that it is possible to reconstruct the whole flow-profile from discrete GPS measurements from probe cars.</p>			
Rationale			
<p>A successful experiment would be a milestone in the process of providing the traffic community with algorithms for real-time estimation of traffic phenomena. It will give traffic management operators scientific and technical support for using GPS measurements in the context of flow reconstruction. This would provide a more accurate description of the state of traffic in real-time, both to drivers and transportation agencies, which would give a deeper understanding of congestion problems. Since GPS</p>			
Protocol Narrative Form	Page 1 of 39	March 2007	

This Appendix contains the protocol as it was originally submitted to the Office for the Protection of Human Subjects in September of 2007. It is included here as a historical document to record the initial logistical requirements that were considered as the plan for the 100-vehicle deployment was taking shape. Details of the final protocol differ somewhat from the procedures outlined in this Appendix. However, the recruiting materials, informed consent form, and overall hiring process are accurate as described.

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

CPHS PROTOCOL NARRATIVE FORM

Instructions: Complete all applicable sections of this form. (If requesting Exempt Status, see instructions on Exempt Request form). Please type, using a different font than the one in this form. **Handwritten or incomplete forms will be returned.** Use language that is clear, concise, and non-technical wherever possible, and define all acronyms. **For renewals or amendments, highlight all changes from the previously approved version on one copy.** A grant proposal or thesis will not be accepted in place of a protocol written according to this format.

Lead Investigator: Prof. Alexandre Bayen

Protocol Title: Testing Cellular Phones as Traffic Probes

CPHS #: _____

Related CPHS Title: _____

CPHS #: _____

Project(s)? Title: Deployment of Value-Added Mobile Traffic Probe

CPHS #: _____

SECTION 1: PURPOSE AND BACKGROUND OF STUDY

- **Purpose:** Provide a brief explanation of the proposed research, including specific *study hypothesis, objectives, and rationale.*

This experiment intends to provide a proof of concept for traffic-flow reconstruction from probe vehicle measurements. To this end, 100 drivers will be asked to drive on a stretch of highway. Their vehicles will be equipped with GPS devices that send information such as speed and position to a server gathering all the data. These probe vehicles, flowing in general traffic conditions, will provide local traffic information. The scientific challenge is to reconstruct the profile of the whole flow from these mobile measurements. This has been achieved in simulation so far, and advancing theoretical developments lead us to believe that this can be done in experimental conditions

Study Hypothesis

The study hypothesis is that the section of highway and the time-range of the experiment are subjected to representative traffic patterns. It is also assumed that traffic theory will apply to the data obtained during the experiment.

Objectives

The objective of the study is to obtain a set of GPS measurements from the highway for free-flow and congestion conditions and test different algorithms to prove that it is possible to reconstruct the whole flow-profile from discrete GPS measurements from probe cars.

Rationale

A successful experiment would be a milestone in the process of providing the traffic community with algorithms for real-time estimation of traffic phenomena. It will give traffic management operators scientific and technical support for using GPS measurements in the context of flow reconstruction. This would provide a more accurate description of the state of traffic in real-time, both to drivers and transportation agencies, which would give a deeper understanding of congestion problems. Since GPS

devices will be standard equipment in cars, mobile phones, and other electronic devices in a few years, this experiment would start a new era in traffic engineering.

- **Background:** Give relevant background (e.g., summarize previous/current related studies) on condition, procedure, product, etc. under investigation, including citations (with attached bibliography) if applicable.

Using GPS measurements in order to reconstruct global phenomena is a growing research area. The accuracy of the information provided and the simplicity of applications of this technology is far beyond what has been done in the past. In traffic engineering, the potential improvement from the use of GPS measurements is even higher.

Currently, traffic knowledge is gathered by fixed detectors placed on the highway where they provide a count of vehicles passing by. This method provides sparse, inaccurate information and requires high maintenance costs from public agencies such as California's Department of Transportation (CalTrans). But the penetration of GPS technologies and the incorporation of GPS chips in cell phones can also provide accurate information about traffic flow. Numerous studies worldwide have incorporated additional knowledge provided by probe vehicles to accurately estimate travel-timeⁱ and provide a reliable method for incident detection.ⁱⁱⁱ

The California Center for Innovative Transportation (CCIT) is currently conducting a study for the California Department of Transportation intended to improve congestion level assessment on California highways through the use of probe vehicles. Several research groups (including Ph.D. students) at UC Berkeley are also working on the improvement of algorithmic methods to extract relevant information from probe measurements, and similar experiments are conducted in rivers in the Bay area for reconstructing the velocity profile of shallow-water flows.

- *International research:* If research will be done outside the U.S., see [CPHS Guidelines on Conducting Research Abroad—Demonstrating Knowledge of “Local Research Context.”](#)

N/A

- *Collaborative research:* If any non-UCB institutions or individuals are collaborating in the research, discuss here and complete CPHS Cover Sheet, Part IV, *attaching any relevant IRB approvals.*

N/A

SECTION 2: QUALIFICATIONS OF STUDY PERSONNEL

- **Expertise:** Explain expertise of Lead Investigator, Faculty Advisor (if applicable), any co-investigators or other key personnel listed in the application, and how it relates to their specific roles on the study team.

Faculty

Professor Alexandre Bayen: Alex is working in the area of Lagrangian sensor networks, i.e. sensor networks which are mobile and transmit their measurements through wireless communication channels. His area of expertise is optimization based control and estimation for algorithm design. His role in this project is to architect the experiment around algorithms which will produce estimation algorithms applicable to the highway problem of interest. He is supervising and helping the team to construct the experiment in a way which will provide enough data compatible with algorithms available for the highway traffic flow reconstruction.

Xavier Litrico: Xavier is a visiting faculty member in the Department of Civil and Environmental Engineering at the University of California, Berkeley. He is an IGREF research Fellow at CEMAGREF, France, and is an expert in the field of real time control and estimation algorithms for distributed parameter systems.

CCIT

J.D. Margulici: J.D. is the Associate Director of the California Center for Innovative Transportation (CCIT), University of California, Berkeley. He holds 6 years of experience with transportation systems and software product development. He worked for the French air traffic control system before becoming the director of product development at Edatis, a Paris-based software company. Prior to joining CCIT, he developed simulation models at Leigh Fisher Associates, a leading airport consulting company located in the San Francisco Bay Area. He joined CCIT in 2004, and has managed nearly a dozen ITS projects. One of his most notable accomplishments at CCIT is the deployment of a system to automatically display travel times on changeable message signs in the Bay Area. He holds a Bachelor of Science degree in Mathematics and Physics from the Ecole Polytechnique in Paris, France and a Master of Science degree in Transportation Engineering from the University of California, Berkeley.

Ali Mortazavi: Ali is a Senior Development Engineer at the California Center for Innovative Transportation (CCIT), University of California, Berkeley. Ali holds a Doctorate degree in Transportation Engineering from George Washington University. His research focused on vehicle active safety and driver drowsiness detection systems. At CCIT, he is responsible for traveler information systems and the assessment and deployment of new traffic data collection technology. One of his projects includes the improvement and further implementation of travel times on Changeable Message Signs. He spent the last four years as a graduate student research associate at the Center for Intelligent System Research at the George Washington University. He also spent several years as an engineer in the private sector.

Jeff Ban: Jeff is an Assistant Research Engineer at CCIT at the University of California, Berkeley. He received an M.S. degree in Computer Sciences and a Ph.D. degree in Transportation Engineering, both from the University of Wisconsin at Madison. His research interests are in the areas of traffic network modeling, traffic simulation, mathematical programming and optimization, and traffic data analysis and mining. During his Ph.D. study, he developed optimization models for computing static and dynamic traffic equilibria. Prior to joining CCIT, he worked as a research associate at the Utah State University for the California SR-41 simulation study project sponsored by Caltrans and CCIT. He is currently involved with projects for Corridor Management Demo for I-5 and I-880, Developing Expert Systems for ITS Decision, Optimal Sensor Deployment, and Optimal Use of Changeable Message Signs (CMS) for Displaying Travel Times.

Post-Doctoral Researchers

Olli-Pekka Tossavainen: Olli-Pekka is a visiting post-doctoral researcher in department of Civil and Environmental Engineering at the University of California, Berkeley. He earned his PhD in physics from the University of Kuopio, Finland, and is an expert in the field of inverse modeling for Lagrangian data in distributed parameter systems.

Annalisa Scacchioli: Annalisa is a post-doctoral researcher in the Department of Civil and Environmental Engineering at the University of California, Berkeley. She received her PhD in Electrical Engineering and Computer Science from the University of L'Aquila, Italy. Her field of expertise is the prediction of error propagation in uncertain systems, and has previous research

experience in automotive control systems.

Graduate and Visiting Students

Daniel Work: Daniel is a PhD student in the department of Civil and Environmental Engineering at the University of California, Berkeley. His research is focused on using cellular phones for traffic estimation. He has previously worked for the City of Westerville, OH Engineering Department, where he was responsible for collecting traffic data at city intersections and examining the safety and efficiency of the city's signalized intersections.

Juan Carlos Herrera: Juan Carlos is a PhD student in the Civil and Environmental Engineering Department at the University of California, Berkeley. His background is in transportation engineering, and in particular in traffic flow theory. He is currently working on the use of Lagrangian sensors in transportation networks as part of his PhD dissertation.

Ryan Herring: Ryan is a Ph.D. student in the department of Industrial Engineering and Operations Research at the University of California, Berkeley. He is currently studying optimization, and he has expertise in the optimal placement of sensors for estimating travel times.

Sebastien Blandin: Sebastien is a visiting mathematics and transportation graduate student from the Ecole Polytechnique in France. His research interests are traffic engineering and data assimilation. His roles in this project are experiment design and algorithm development to reconstruct traffic flow from collected measurements.

Arthur Wiedmer: Arthur is a graduate student in the Civil and Environmental Engineering M.S. program at UC Berkeley, graduated from the Ecole Polytechnique in France. He has research experience in general and sensor-related data analysis.

Christian Claudel: Christian is a PhD student in the Electrical Engineering and Computer Science department at the University of California, Berkeley. His background is in numerical models for traffic congestion and real time control.

Tarek Rabbani: Tarek is a PhD student in the Mechanical Engineering department at the University of California, Berkeley. His research focus is in real time control for onboard sensing of ground vehicles.

Qingfang Wu: Qingfang is a PhD student in the department of Civil and Environmental Engineering at the University of California, Berkeley. Her background is in inverse modeling for GPS trajectories evolving in moving fields.

- **Training:** For *graduate or undergraduate students* who are Lead Investigator or key personnel of the study, confirm training to conduct research with human subjects (required for all student researchers—see CPHS Cover Sheet, Part VI). *Attach copy of completion report for each individual, unless submitted previously.*

Copy of completion report for each individual is attached.

SECTION 3: SUBJECTS (Persons/Records/Specimens)

- **Eligibility:** Describe proposed subject population, including criteria for study inclusion and exclusion (e.g., age, health status, language). If any inclusion/exclusion criteria are based on gender, race, or ethnicity, explain rationale for the restrictions. Indicate how, when, and by whom prospective subjects will be identified and eligibility determined (provide fuller discussion of recruitment, screening, and consent process in Sections 4-6). Describe randomization or other assignment method for intervention and control groups.

The proposed subject population will be a set of 135 students from UC Berkeley. The criteria for inclusion are: to be a student from UC Berkeley, to have a valid driver's license, have a fluent understanding of the English language, having already driven a minimum of 100 miles on California highways, be in good physical health so that they can handle driving for a duration of 3 consecutive hours, 21 years or older. Pregnancy constitutes exclusion criteria due to the stress of having to drive for 8 hours. No inclusion/exclusion criteria will be based on gender, race, or ethnicity. Once selected, the subjects will have to be trained about the detailed procedure of the experiment. Failing to complete this training before the day of the experiment will be a criteria of exclusion.

We request subjects be at least 21 years of age because they will likely have more experience driving than subjects below the age of 21. We ask that subjects be able to communicate in English, because this will allow information to be more easily communicated to and from the researchers and subjects. We select students from UC Berkeley at the suggestion of the Office of Risk Management, since this population is already covered under the University insurance policy. This issue of liability was discussed with Ms. Barbara VanCleave Smith from the Office of Risk Management. The use of rental cars was suggested. The vehicles and the drivers are automatically covered by the University if they are driving rental cars on University business.

- **Number:** State total number of subjects planned for the study and how many must be recruited to obtain this sample size. Explain how number of subjects needed to answer the research question was determined.

Subjects will be hired into one of two groups. The first 135 subjects which are verified by the Institute for Transportation Studies (ITS) human resources personnel to meet the age, license, and student status qualifications will be hired as *drivers*, while subjects 136 through 150 will be hired as alternates. The number of drivers has been chosen in order to allow drivers sufficient breaks while still maintaining nearly 100 vehicles (and drivers) on the freeway at any given time. 100 of the drivers will be driving at any given time and 35 will be resting. An indicative rest time has been defined in accordance with legal times for truck drivers worldwide. The total number of vehicles (100) included in the study was selected to allow a rich data set from which we can reconstruct the traffic conditions. This number is based on our own research, and depends on the length of the section used for the experiment and of the flow of vehicles usually on the highway at the time of experiment. The time duration of the experiment is fixed by the congestion phenomena we want to capture and the number of subjects to hire was calculated accordingly.

- **Vulnerable Subject Groups:** Indicate whether any proposed subjects are children/minors, prisoners, pregnant women, those with physical or cognitive impairments, or others who are considered vulnerable to coercion or undue influence.

The subjects are not children/minors, prisoners, pregnant women, those with physical or cognitive impairments, or others who are considered vulnerable to coercion or undue influence.

SECTION 4: RECRUITMENT

- **Summary:** Explain how, where, when, and by whom prospective subjects will be identified/selected and approached for study participation. NOTE: If researcher is subject's instructor, physician, or job supervisor,

or if vulnerable subject groups will be recruited, explain what precautions will be taken to minimize potential coercion or undue influence to participate.

Recruitment of prospective subjects will begin starting from the date of approval for the experiment from CPHS through the completion of recruitment for the required number of subjects. Subjects will be recruited via multiple forms of advertisement (fliers, email, and website). Subjects will be asked to submit an email indicating their *Interest to Participate* in the experiment to NextGenCell@ce.berkeley.edu (Only key personnel will have access to this email account). Key personnel will respond to the email, asking for confirmation that they meet the qualifications listed on the website, flier and *Statement of Informed Consent to Participate in Research*, and to select from a list of four dates all days that they will be able to attend a hiring session, to be conducted by the ITS human resources personnel. If the subject answers that they meet these qualifications, another email will be sent indicating the date and location of the hiring session, further indicating that only the first 135 qualified drivers will be hired as drivers, with an additional 15 subjects hired as alternates. At the completion of the hiring of 135 drivers and 15 alternates, all 150 subjects will be notified by email to which group (driver or alternate) they belong, and given further instructions to complete an online training course no later than 10 days before the date of the experiment.

The online training course will provide the participants with an overview of the section of the highway that will be driven. Maps will be used to indicate entrances and exits to and from the highway as well as how to navigate the turn-around points. A driving schedule, as well as a schedule for the day will also be provided. The training course will also include instructions on proper driving behavior during the experiment (follow all laws, etc.) All 150 participants will need to complete a quiz at the end of the training course to ensure they understand the important details of the experiment for which they will be responsible. Any participant who fails to successfully complete the training will be notified at the end of the test, and will not be allowed to participate on the day of the experiment and will not be paid.

135 subjects will be recruited based on the criteria above. Additionally, 15 alternates will be recruited. These alternates will be required to go through the same steps as described above, but will be informed ahead of time that they will only participate in the driving phase of the experiment if there are not enough people from the original pool of 135 (for example, if some of the original 135 do not show up on the day of the experiment, or a driver requests an unscheduled break). Alternates will receive the same instruction, care, and compensation as drivers, regardless of whether they are asked to drive at some point during the experiment.

- **Recruitment Materials:** Describe and *attach samples of any recruitment materials* (e.g., letters, flyers, advertisements [note type of media/where posted], scripts for verbal recruitment, etc.).

There will be three primary recruitment materials: fliers, email, and a web page. The flier and recruitment email are attached to the present document, and the website is located at <http://NextGenCell.googlepages.com>. The flier will be posted only on the UC Berkeley campus, and only on bulletin boards designated for fliers. The email recruitment letter will be used on departmental list serves, with the approval of the departments. The website will contain an overview of the experiment, procedures for participating in the experiment, and the Statement of Intent to Participate in Research

- **Permissions:** If applicable, describe and *attach IRB approval or letter of permission/ cooperation* from institutions, agencies or organizations where off-site subject recruitment will take place (e.g., another UC campus, clinic, school district).

N/A

SECTION 5: SCREENING PROCEDURES

- **Summary:** If prospective subjects will be screened via tests, interviews, etc., prior to entry into the “main” study, explain how, where, when, and by whom screening will be done. NOTE: *Consent must be obtained for screening procedures as well as “main” study procedures. As appropriate, either: 1) create a separate “Screening Consent Form;” or 2) include screening information within the consent form for the main study (see Section 6).*

As outlined above, we will recruit subjects who 1) are UC Berkeley students, 2) are at least 21 years old, 3) have a valid driver’s license, 4) communicate fluently in English, 5) have driven at least 100 miles on California highways 6) are in good physical health 7) are not pregnant and 8) willing to participate in the experiment.

Applicants will be screened for qualifications 1-8 by asking them (by an email from NextGenCell@ce.berkeley.edu) if they meet these requirements, after receiving an email indicating the subject’s Interest to Participate. If subjects respond that they do not meet these requirements (or do not respond), they will be notified that they are not eligible to participate by an email from key personnel. If the subjects respond that they meet these requirements, requirements 1-3 will be verified by the ITS Human Resources Office during the hiring of the subject.

Key qualifications are also included on the website and flier recruitment materials.

- **Identifiable Personal Information:** Indicate if identifiable personal information will be obtained as part of the screening process. (Confidentiality issues should be addressed in Section 11).

Identifiable personal information will be obtained as a part of the screening process. The name and email address of interested subjects will be obtained by key personnel, and potentially an email indicating if the subject meets all qualifications. Additional personal information will be obtained by the ITS human resources offices for the purposes of hiring the subjects. The related confidentiality issues are addressed in Section 11.

SECTION 6: INFORMED CONSENT

NOTE: See [CPHS Informed Consent Guidelines](#) before completing this section.

- **Summary:** Explain how, where, when, and by whom informed consent and/or assent will be obtained. NOTE: If any vulnerable subject groups/other special circumstances are involved (e.g., use of surrogate consent), address considerations appropriately.

Informed consent will be obtained at the time of hiring by the ITS human resources personnel. Applicants will be asked to read the Statement of Informed Consent to Participate in Research, and sign it only if they agree to its terms.

- **Consent Materials:** Describe any consent/assent form(s) to be used, and *attach copies*.

If *screening procedures* will be done for the study, see above. Whichever method is used (separate consent or part of the main consent), the form should include a statement regarding what will happen to screening information collected for individuals who do not enter the study.

If any *vulnerable subject groups* will be involved, address appropriately (e.g., if study includes minors, both an assent form for the child and a consent/permission form for the parent(s) may be required).

For *international research*, provide for and describe *local contacts* in the area.

The Statement of Informed Consent to Participate in Research is attached to this document. It describes the purpose and procedures of the experiment, benefits, risks and discomforts, hiring procedure, confidentiality of the data, compensation, costs, treatment and compensation for injury, the voluntary nature of the experiment, necessary qualifications, and contact information for additional questions. This form will also be posted on the recruitment website of the experiment. Subjects will be asked to read and sign it by the ITS human resources personnel during the hiring procedure.

- **Request for Waiver of Consent:** If you are requesting waiver of any of the required elements of informed consent, or waiver of documented consent, or waiver of parental consent or child's assent, provide justification and describe plans for any additional safeguards. (See [CPHS Informed Consent Guidelines](#)).

N/A

SECTION 7: STUDY PROCEDURES

- **Summary:** Describe how the research will be conducted, providing information about all study procedures (e.g., interventions/interactions with subjects, randomization, photographing, audio- and/or videotaping, data collection), including follow-up procedures. (Screening procedures should be discussed in Section 5). Be sure to make clear what the sequence of study procedures is (i.e., describe in chronological order).

Overview

The goal of the experiment is to test a system designed to collect velocity and position data from vehicles carrying Nokia N95 cellular phones on-board. This data will be used to develop tools to estimate traffic conditions on urban freeways. In order to collect this data, 100 vehicles will make loops on a limited stretch of I-880 near Union City, CA, and the cellular phones will record the speed and position of each vehicle at all times.

In order to obtain relevant results, the experiment must be conducted during usual daily traffic conditions, from light traffic to heavy congestion. Because of the varying traffic conditions, the experiment will consist of two phases of data collection. *Phase 1* will occur during un-congested or free-flow traffic conditions, while *phase 2* will occur during congestion. Phase 1 of the experiment will be conducted on a 10 mile long section of I-880, whereas phase 2 will be conducted on a shorter, 3 mile long sub-section. In addition to the two phases of data collection on the day of the experiment, the subjects will assist with an experiment *preparation phase* and an experiment *conclusion phase*. The preparation phase includes shuttling the subjects to a rental car company, where they will pick up their vehicles and drive them to the experiment site, as well as a debriefing on the experiment procedures. The conclusion phase involves the subjects returning the vehicles to the rental agency before being shuttled back to the UC Berkeley campus. The participation of the subject will last from 6:45am through 10:00pm on the day of the experiment. No further involvement will be requested after the day of the experiment, and all contact with the subject will stop at the end of this day.

Location

The experiment will be conducted on Highway I-880 between Highway 92 to the North and Mowry Avenue to the South, and is scheduled to take place on February 8, 2008. This 10-mile long section has been selected for its traffic properties, an existing knowledgebase for this particular highway from

traffic simulations, and for its proximity to UC Berkeley. There are also a number of facilities available in the vicinity of the highway, including access to parking, gasoline, and food, have also been taken into account.

Phase 1 of the experiment will take place on the entire 10-mile section, from Mowry Avenue to Highway 92, whereas phase 2 will take place between Whipple road and Road 92. A base station located at the Union Landing retail entertainment center parking lot on Whipple Road will serve as the main coordination site during the experiment, as well as the rest site for the drivers. An overview of the experiment location is attached to this document.

Time

Prior to the experiment, the subjects hired will be given details of the procedure of the experiment online. This step will be crucial for the safety of the subjects and for ensuring that every feature of the operational process is understood. This will be achieved through a secured online website dedicated to this task. The training will be taken by every subject at least 10 days prior to the experiment and is a requirement for participation in the experiment.

On the day of the experiment, February 8, 2008, the drivers will be required to report at 6:45 am at the West Gate of UC Berkeley, returning to the same location at 10 pm after the culmination of the experiment. The preparation phase will last from 6:45 am to 9 am, during which time the subjects will be driven by shuttle to the rental agency to pick up the vehicles, before proceeding to drive the rental cars to the experiment site. Because there are 135 drivers and an additional 15 alternates and only 100 vehicles, the busses will drive the 50 subjects without vehicles to the experiment site. Drivers will be briefed on the experiment before beginning the main data collection phases.

Phase 1 and phase 2 of the experiment will be conducted between 10 am and 8 pm on the stretch of highway described previously. Phase 1 will be conducted between 10 am and 1 pm to capture free flow traffic conditions. Phase 2, which is aimed at capturing the congestion phenomena, will last from 1 pm to 8 pm.

The conclusion phase will last from 8 pm to 10 pm, and will include the subjects returning the vehicles to the rental agency before receiving a bus shuttle back to the West Gate of UC Berkeley.

Preparation phase

On the day of the experiment, the first step will be to bring the drivers, the cars, the key personnel, and the devices to the location of the experiment.

From UC Berkeley to the vehicles

All students hired for the experiment will be gathered at the West Gate of UC Berkeley campus in the morning (6:45 am) and will be driven by buses to the rental agency leaving at 7:00 am. The cars will be picked up from a rental car company at the Oakland International Airport, which is 16 miles south of UC Berkeley.

Prior to the students' arrivals the cell phones will be distributed to the cars and put in a small sealed box and placed on the passenger seat of the vehicle. In order to prevent the box from moving, the passenger seatbelt will be fastened, and a special clip on the box will attach to the seatbelt. The phone must be placed near the front console so that the phone car charger may be plugged in, but the box will

completely conceal the phone and prevent tampering, eliminating any potential distraction to the driver. This particular point will be emphasized before the drivers enter the vehicles.

The drivers and alternates will be divided into four groups (33 or 34 drivers, plus 3 or 4 alternates), each with a group leader (a member of the key personnel). The group leader will be in charge of accompanying the students during the whole day, and will address any questions or concerns from the students.

From the vehicles to the base station

After picking up the vehicles, 100 subjects will drive the vehicles under police escort to the experiment base station on Whipple Road, while the remaining 35 subjects will arrive by bus.

The road from the airport to the base station is a 12-mile section of I-880.

Briefing at Union Landing retail entertainment center

The convoy will arrive at the Union Landing retail entertainment center parking lot at 9:00 am. The site manager (a member of the key personnel) will have already arrived there and put a sign from the main road to the parking location, which is less than two hundred meters from the main road.

At the arrival of the drivers, each group leader will address the drivers in order to clarify every step of the loop driving procedure, and the schedule for the day. The drivers within each group will have the same driving schedule (Figure 1), and the schedule will be discussed at this time.

Driving and Resting Schedule										
	10 AM to 11 AM	11 AM to 12 AM	12 AM to 1 PM	1 PM to 2 PM	2 PM to 3 PM	3 PM to 4 PM	4 PM to 5 PM	5 PM to 6 PM	6 PM to 7 PM	7 PM to 8 PM
Group 1	Rest	Driving	Driving	Driving	Rest	Driving	Driving	Driving	Rest	Driving
Group 2	Driving	Rest	Driving	Driving	Driving	Rest	Driving	Driving	Driving	Rest
Group 3	Driving	Driving	Rest	Driving	Driving	Driving	Rest	Driving	Driving	Driving
Group 4	Driving	Driving	Driving	Rest	Driving	Driving	Driving	Rest	Driving	Driving

Figure 1. Driving and Resting Schedule for Each Group

At this time, each driver and alternate will be provided a Bluetooth head set in order to communicate with the group leader throughout the duration of the experiment, if such communication is needed.

Preparation to release vehicles onto the highway

Prior to sending the car on the highway, the proper function of the cell phone, GPS component, and Bluetooth head set will be checked to ensure functionality.

Signage will be posted at each end of the highway segment, in consultation with the California Highway Patrol, in order to ensure a safe initial deployment and that drivers adhere to the loop pattern. All cars will have been marked with a number so that it is easy for safety personnel to identify them on the highway. This mark will be placed on the body of the vehicle, where it will not reduce the visibility of the driver.

Phase 1

At 10:00 am, phase 1 of the experiment will start from the base station with 100 drivers released onto the highway. The cars will be released from the base station with a time interval of 12 seconds between two consecutive cars. This time delay allows for a relatively uniform distribution of the vehicles on the 10-mile section.

The path from the base station to the highway will be indicated by signs, and one of the extra group leaders will verify that drivers take the correct entrance to the highway.

The releasing step would last until approximately 10:20 am, at which point there will be 100 drivers on the highway, each tracked in near real-time by the cellular phone. At this point 35 of the drivers and 15 alternates will be resting at the base station, while the risk management personnel (members of the key personnel) and group leaders monitor (to be described subsequently) the groups of drivers and respond to any unforeseen logistical issues that might arise. The base station is located in the Union Landing retail entertainment center parking lot, and was selected taking into consideration the fact that 35 drivers would rest during the whole day at that location. The drivers will be able to use the center's facilities and restrooms while they are on break.

Description of the loops

Each driver will drive on the indicated section of highway and will be asked to drive as they would normally, while being sure to follow all traffic laws. Each driver is allowed to pass other drivers, and perform any legal actions that would normally occur while driving on the freeway. This includes taking breaks at their own discretion if they no longer wish to drive the vehicle, by returning it to the base station and notifying their group leader. The driver will be replaced by an alternate, who will be asked to complete the rest of the driving shift.

At the end of the highway section each driver will take the specified off-ramp and reenter the highway in the opposite direction. At the northern endpoint of the section, this can be achieved without passing through any intersection. At the southern endpoint of the section (during phase 1), the driver will need to wait at a stop light, then to turn left, and take the onramp in the opposite direction. At the southern point of the section (during phase 2) the drivers will be able to change direction by taking the off-ramp, turning right, and then taking the on-ramp.

After returning to the opposite boundary of the segment, the driver will change direction again, starting the second loop. Each loop will take approximately 20 minutes to complete, and the driver will complete these loops during three hours. During this time, the vehicle will need refueled. Drivers are given specific instructions to exit towards the base station and refuel at the gas station immediately adjacent to the base station as soon as the vehicle has less than ¼ tank of gas remaining. The amount of fuel remaining will be checked each time a vehicle is brought in to change drivers. The gas at the gas station will be purchased by the researchers in advance.

Phase 2

Phase 1 of the experiment will be conducted during free-flow conditions from 10 am to 1 pm, during which time drivers will complete long loops from Mowry Avenue to Road 92 (10-mile section). The shorter, 3-mile sub-section from Whipple Road to Road 92 will be driven during phase 2 of the experiment from 1 pm to 8 pm.

It is important to describe the way this change from phase 1 to phase 2 will be achieved. Drivers will have been briefed several times (during the online training and during the preparation phase) on the

time the change will occur, and they will receive a reminder through the Bluetooth headset at 12:30. All drivers will be asked to complete the loop they are currently driving. At this point they will take the Whipple Rd. off-ramp (southbound), change directions, and take the on-ramp to begin heading north and start the shorter phase 2 loops. Since the short loop travel time should be around 20 minutes at this point during the day (based on increased congestion), the pattern change will be completed by approximately 1 pm. This shift from the longer (in length), phase 1 loop to the shorter phase 2 loop should enable a relatively uniform distribution of vehicles on the short loop after completion of the transition. A group leader who is not directly managing one of the four groups will be sent to the exit ramp to observe the correct transition of vehicles. In the event a vehicle does not correctly transition, the vehicle will be contacted via the Bluetooth headset, and will be asked to exit at the next available exit and change directions.

Conclusion phase

Drivers will be given precise indications about the time they should stop their loop during the online training and during the morning preparation phase. At this time (7:40 pm) they will be asked to complete their loop until they reach the off-ramp at Whipple Road, and come to park at the parking lot. The group leaders will wait for their groups and check that all drivers are safe. A group leader who is not directly managing one of the four groups will be sent to the exit ramp to observe that all vehicles return to the base station.

When all drivers have gathered in the parking lot a formation will be constituted and escorted by police officers to the rental agency at the Oakland International Airport. The remaining 35 drivers not returning a vehicle will be transported to the rental agency by bus. At the Oakland International Airport, drivers will drop off the vehicles. The mobile phones and secure boxes will be collected from the vehicles by the group leaders and will be brought to CCIT by the lead investigator. Once all group leaders have verified that their drivers have returned their cars, all students will be bussed back to UC Berkeley.

Upon arriving at the West Gate of UC Berkeley, drivers will exit the buses and thanked for their participation. Each driver will receive a Bluetooth headset to keep, and the winners of the N95 cell phone raffle will be announced. The winners will receive the phone at this time, and no further communication with the subjects will occur after this time.

Monitoring the drivers

During the experiment, the drivers will be monitored through several channels in order to avoid any problems. Being able to know where the drivers are, how they feel, and to provide them with any assistance if needed will be the principle concern of all key personnel throughout the experiment.

First, the location of every vehicle will be known in near real-time (maximum few minutes delay) from GPS measurements sent from the phone to the secure server, which will be accessed on site one of the two risk management personnel. The risk management personnel will be able to check drivers' positions in near real-time, and software will automatically detect if a vehicle exhibits inappropriate behavior (fails to exit at off-ramp, or driving at a dangerously high speed (more than 15 mph over the posted speed limit, since the GPS may have small errors)). In the event inappropriate behavior is detected, risk management personnel will contact the group leaders, who will identify the driver and ask them (via Bluetooth) to correct the behavior. If the driver is unable or unwilling to comply, the police will be notified to respond accordingly.

Second, the group leaders will have the capability to contact their drivers in the event it becomes necessary. This will be achieved with no disturbance to the driver, since message will be received on a hands-free Bluetooth headset. This communication will provide a channel for important safety information to be conveyed to the driver. Additionally, the driver may contact the group leader through the headset to report any problems with the vehicle, or if they need to take a break. These communications are designed to help ensure the safety of the driver.

Third, a member of key personnel will be posted at the highway entrances and exits during the start of phase 1, the transition between phase 1 and 2, and at the end of phase 2, which will provide immediate relevant information at this location. This will allow for real-time feedback during these transitions.

Fourth, hourly patrol on the highway will be done by a member of key personnel who will make a loop to check that the experiment is proceeding accordingly and that nothing has happened on the highway that was not detected by the GPS and could reduce the safety of the drivers. In the event a vehicle has broken down, the police will be notified and the on-site tow truck will be called to respond to the vehicle, and return the driver to the base station. If the driver requests or is believed to need medical attention, an ambulance waiting at the base station will be prepared to respond.

Fifth, each hour 33 or 34 drivers will be returning from the road for a break (almost continuously), and will be asked to mention anything perceived as potentially dangerous by drivers to the group leader. Group leaders will immediately report any problems to one of the risk management personnel, who will determine an appropriate course of action in consultation with the on-site police officer.

- **Study Personnel, Location, Time:** Explain who will conduct the procedures, where and when they will take place. Indicate frequency and duration of visits/sessions, as well as total time commitment for the study.

The study personnel will consist of a UC Berkeley professor, a visiting professor, three CCIT researchers, two UC Berkeley post-doctoral researchers, and seven graduate or visiting students at UC Berkeley. Professor Alexandre Bayen will be the experiment supervisor, while visiting professor Xavier Litrico will be in charge of oversight of the experiment. One post-doctoral researcher and one CCIT researcher will be risk management personnel, with responsibilities to monitor the safety of the experiment and will be the primary contact with the on-site police, ambulance and tow truck. One post-doctoral researcher and one CCIT researcher will be site managers, ensuring logistical operations at the site run smoothly. One student will be an experiment manager, and will be available to assist other key personnel throughout the experiment. Six additional students will be the group leaders of the four groups. The extra two group leaders will provide each group leader with breaks, and will be in charge of monitoring the vehicles at the entrance/exit ramps during transition phases.

Resting Schedule

It is important to define a rotating shift to enable drivers to take rest while ensuring that the vehicles can collect as much data as possible. Phase 1 and phase 2 of the experiment will last 10 hours, and each driver will be asked drive between seven and eight hours. A schedule allowing each driver one hour of rest after every three hours of driving will be distributed to each subject during the online test, and during the preparation phase. The drivers will be free to rest in the buses parked on the base station parking lot during the whole day, or to enjoy the facilities offered at the Union Landing retail entertainment center. Every sanitary accommodation will be available in the shopping center. Both lunch and dinner is catered and will be available to drivers during their breaks at no cost to them.

This driving schedule has been defined based on the most restrictive truck driver regulations worldwide, which stipulates rest at least 45 minutes every 4 hours of driving. One may consider that truck drivers drive long distances on monotonic trip with few changes in directions or contact with other people. The students will complete loops in approximately twenty minutes, which means that they will have to change directions every ten minutes. These frequent changes will help avoid the monotony of the driving.

In the event a driver wishes to take an unscheduled break, they may return the vehicle to the base station. At this time, an alternate, if available, will be asked to begin driving until the end of the shift. In the event more drivers request breaks than alternates are available, the vehicle will remain at the base station until the driver is willing to resume the experiment.

- **Experimental vs. Standard Procedures:** Identify any procedures that are experimental/ investigational and explain how they differ from standard procedures (medical, psychological, educational). If applicable, distinguish between procedures that the subject would undergo regardless of enrollment in the study and procedures done specifically for study purposes.

N/A

- **Deception:** This includes both “*active deception*” (deliberately giving false information about study purpose and/or procedures to subjects) and “*lack of full disclosure*” (withholding complete information about the study from subjects.) If any type of deception will be used, explain what it will entail, why it is justified, and what the plans are to debrief subjects. Also, attach debriefing forms(s)/materials. (NOTE: If study involves significant deception at time of subject enrollment/consent, the CPHS may require a post-study re-consent as part of debriefing process).

N/A

- **Drugs/Devices:** If study involves an experimental drug or device, complete IND/IDE information on CPHS Cover Sheet. Describe any study drug here, including generic and/or chemical name, how it is supplied (e.g., powder, capsule, liquid), administration method and schedule, etc.

N/A

- **Placebo:** If placebo will be used, provide rationale and explain why active control is not appropriate.

N/A

- **Data Collection Instruments:** If interviews, questionnaires, surveys, or focus groups will be conducted for the study, provide citations for standard instruments and *attach 1 copy of any non-standard instruments to be used.*

The data collection instrument will consist of a cell phone equipped with a GPS connected to a storage device. The GPS will record position and speed during the whole experiment. The data collected will be stored on a storage device located in the sealed box with the cell phone. At the same time, the measurement from the cell phones will be sent to a server located at UC Berkeley over the cellular phone network. The server will be protected from any non-authorized access, both in the real and virtual world. The storage device will be kept in the sealed box during the whole duration of the experiment, and will be available for processing only after a randomization step. This will be achieved by randomly changing the IDs of the vehicles. This way, even key personnel on site during the experiment who would know both a vehicle’s number and the name of his driver will not be able to

match the data collected by the device to the driver after the experiment is complete, since the number of the vehicle in the database would have been randomized.

- **Identifiable Personal Information:** Indicate if identifiable personal information will be obtained from/about subjects. (Confidentiality issues should be addressed in Section 11).

The data collected during the experiment will be limited to the cell phone/vehicle ID, the time of measurement, the location of measurement, and the speed measured. For safety purposes, the lead investigator will have at his disposal a list of the names of each driver, and a member of the ITS human resources department will be available by phone to provide emergency contact information to the lead investigator in the event of an accident. Each group leader will maintain an evolving list of the first and last name of the drivers, and the cell phone/vehicle ID they are currently operating. This list will be updated throughout the day as the drivers rotate shifts, and will be stored in a secure file (locked, only the lead investigator will have access) at CCIT after the experiment is complete.

Personal information will be collected by the ITS human resources personnel in order hire the drivers. Subjects will need to fill out the Employee Data form, and provide emergency contact information. If the subject is not a UC Berkeley employee, they will need to provide a valid passport or California driver's license and a social security card, and fill out the university hiring paperwork. This data will be collected and maintained by the ITS human resources personnel, following standard University procedures.

SECTION 8: RISKS/DISCOMFORTS

- **Summary:** Describe all known risks, discomforts, and/or side effects of study procedures, whether physical, psychological, or social (e.g., pain, stress, invasion of privacy), noting probability and magnitude of potential harm. Include risks of randomization and placebo if applicable.

The experiment poses no additional risks to the subjects other than those normally encountered in operating a vehicle on an urban freeway. These risks include roughly a 0.08% chance to be involved in an automobile accident, and roughly a 0.0001% chance of being involved in a fatal accident, based on national highway statistics and the mileage they will be asked to drive. For comparison, the risk of being involved in an accident while riding a bicycle is roughly 0.031%, with roughly a 0.00005% chance it will be fatal. These figures have been calculated assuming the subject rides a bicycle for 8 hours instead of driving a car for 8 hours. The annual number of bicycle-hours, accidents and fatalities in the US are obtained from the Consumer Product Safety Commission^{iv}, and from the number of car accidents, fatalities and vehicle-miles traveled are obtained from the Bureau of Transportation Statistics^v.

- **Measures to Minimize Risks/Discomforts:** Discuss measures that will be taken to minimize risks or discomforts to subjects.

The main measure to minimize risks to the subjects is the monitoring of the vehicles throughout the day (this is described in detail in section 7 "monitoring the drivers"). Additionally, the driving breaks are designed to keep the drivers safe (see section 7 "resting schedule"). Finally the screening procedure (including age requirement and driving experience) is designed to select drivers who are willing and able to operate a vehicle safely on a California highway.

Two members of the key personnel will be dedicated as risk management personnel, with the sole purpose of maintaining the safety of the drivers throughout the experiment. As an additional safety measure, a police officer and an ambulance will be at the base station to respond to any unforeseen events that may occur.

- **Currently Unknown Risks:** If applicable, indicate if a particular study treatment or procedure may involve risks to the subject (or to the embryo or fetus, if the subject is or may become pregnant) that are currently unforeseeable.

N/A

SECTION 9: BENEFITS

- **Summary:** Describe any potential benefits to the individual subject, group of subjects, and/or society. If subjects will not benefit directly from study procedures, this should be stated. NOTE: Do not include compensation/ payment of subjects in this section, as remuneration is not considered a "benefit" of participation in research (compensation/ payment should be addressed in Section 12).

There is no direct benefit to your participation in this experiment. It is hoped that the participation will enable researchers to develop the next generation of highway traffic sensing technologies, from which the subjects will benefit indirectly.

SECTION 10: ALTERNATIVES TO PARTICIPATION

- **Summary:** Describe appropriate alternative resources, procedures, courses of treatment, if any, that are available to prospective subjects. If there are no appropriate alternatives to study participation, this should be stated. If the study does not involve treatment/intervention, put "N/A" here.

N/A

SECTION 11: CONFIDENTIALITY

NOTE: See [CPHS Data Security Policy](#) before completing this section.

- **Summary:** Explain how subject privacy will be protected and how confidentiality of subject information will be maintained.

People who volunteer to participate as subjects in research do so with the understanding that the researchers will protect their identity and the information that is obtained from them from inadvertent or inappropriate disclosure. The researchers will make sure to take appropriate actions to protect participants' personal privacies. A secure filing cabinet will be dedicated to the experiment documents in CCIT. The researchers will follow all the recommendations in 'Security of Research Subjects' Personally Identifiable Data Held by Researchers' policy provided by CPHS.

In the hiring process, personal information will be collected. All hiring forms which include confidential information will be filed and stored at ITS.

During the experiment each group leader will maintain a list of the names of the drivers and the vehicles in their group. This list will link the driver to the cell phone and vehicle which they are currently driving. It is the only data collected during the experiment which can link the driver to the mobile phone/vehicle they are driving. At the end of experiment, these lists will be collected by the lead investor and stored in the secure filing cabinet at CCIT.

The raw data collected during the experiment will be stored in two locations. The first is on the storage device attached to each cellular phone in the experiment. The second is on a secure server at UC Berkeley. At the end of the experiment, all raw electronic research data will be encrypted using secure data encryption algorithms, and stored permanently on a hard drive which will be placed in the secure

filing cabinet at CCIT. All raw data on the storage devices attached to the phone will then be erased. At this time, the raw research data will be processed to remove the vehicle and phone ID numbers, and new random vehicle ID numbers will be assigned. The researchers will analyze the de-identified data for the study. This data will also be stored on the secure UC Berkeley server.

- **Access to/Security of Study Records:** Discuss who will have access to study records/specimens and how the records will be secured. Address all applicable points below:
 - Will subjects be asked to give permission for release of identifiable data (e.g., information, videotapes), now or in future? If so, explain here and include appropriate statements in consent materials.
 - Will data be collected anonymously (i.e., no identifying information from subjects will be collected/recorded that can be linked to the study data)? (NOTE: Data is not collected anonymously if there is a code linking it to personally identifiable information).
 - If using existing data/biological specimens, will the researchers have access to a code linking the data to personally identifiable information?
 - If identifying information will be collected and linked to data/specimens, explain at what stage identifiers will be removed from the data/specimens.
 - If identifiers will be retained, explain why this is necessary and how confidentiality will be protected.
 - If the data is coded, explain where the key to identifiers will be stored, how it will be protected, and who will have access to it.
 - Indicate whether research data/specimens will be destroyed at the end of the study. If data will not be destroyed, explain why, where, in what format, and for how long it will be retained.
 - Explain how data collection instruments, audiotapes, videotapes, photographs, etc. will be stored and who will have access to them. Indicate at what point they will be transcribed and/or destroyed (if ever).

The principle investigator will have the access to the records. The records will be stored in a secure filing cabinet protected by a lock in CCIT. Only the lead investigator will hold the key.

Portions of the experiment may be recorded (film/video), and subjects will be asked to provide photographic consent. The standard University photographic consent form will be used in this regard. This form is attached to this document. All photographs and video will be stored in the secure filing cabinet at CCIT.

The data will not be collected anonymously, and this is stated in the consent form. During the experiment the GPS data will be recorded as well as the cell phone/vehicle IDs. This data will be accessed from the secure server during the experiment in near real-time by the risk management personnel. Only the group leaders will have lists of the current driver in each vehicle during the experiment, but this information may be shared with the risk management personnel, lead investigator, and police in the event of an adverse event. Therefore, the collected data set will be identifiable. Proper actions will be taken to de-identify the data before the data analysis.

During the experiment, the vehicle data from the GPS will be collected and stored in a secure server. The drivers will be in interaction with other researchers and staff. Therefore, to avoid any possible

correlation of the data to any driver, at the end of the experiment a key list will be generated and the vehicle ids in the data set will be encrypted. The old data set will be saved into a hard drive and will be returned into the Lead investigator. The hard drive and the key lists will be stored in the experiment cabinet.

There is no need to retain the identifier in future. In case, the lead investigator will be the only person who hold the cabinet key and can access the documents.

NOTE: The CPHS does not require that researchers destroy their human subjects data at the completion of their research. *Whenever appropriate, researchers may retain study data for future use/ other research purposes as long as they make provision in the protocol and consent documents for such use.* Researchers must spell out in the protocol how confidentiality will be maintained vis-à-vis long-term storage of data and/or granting of access to other researchers, and the consent forms must clearly ask subjects for permissions in this regard.

- **HIPAA:** If any of the study data sources are covered entities under HIPAA (Health Insurance Portability and Accountability Act), explain what arrangements have been made to comply with the Privacy Rule regarding subjects' "protected health information." (See [CPHS website for HIPAA guidance](#)).

N/A

- **Reportable information:** If it is reasonably foreseeable that the study will collect information which state or federal law requires to be reported to other officials (e.g., child or elder abuse) and/or ethically requires action (e.g., suicidal ideation), discuss here and reference reporting requirements in consent documents.

N/A

- **Certificate of Confidentiality:** In certain circumstances, researchers may plan to protect research records from subpoena by seeking a Certificate of Confidentiality (<http://grants.nih.gov/grants/policy/coc/index.htm>). If a Certificate of Confidentiality will be sought for this study, indicate here and reference in consent documents.

N/A

SECTION 12: FINANCIAL CONSIDERATIONS

- **Compensation/payment:** Describe plan for compensation of subjects by addressing points below. If no compensation will be provided, this should be stated.
 - If subjects will be compensated for their participation, explain in detail about the amount and methods/terms of payment.
 - Include any provisions for partial payment if subject withdraws before study is complete.
 - When subjects are required to provide Social Security number in order to be paid, this data must be collected separately from consent documentation. If applicable, describe security measures that will be used to protect subject confidentiality.

As a participant in this experiment, subjects will receive payment of \$250, which they will receive on their next scheduled paycheck from the university (if already an employee). If the subject is not already an employee, a check will issued through ITS. Subjects will be paid in full upon completion of the experiment. In the event a subject withdraws before the study is complete, they will not receive the \$250 compensation.

As previously described, subjects will be required to provide personal information to ITS human resources, including a social security number, in order to be paid. This data will be collected and

maintained by the ITS human resources personnel. Data linking the subjects' names to the cellular phone and vehicle ID numbers will be collected for logistical purposes during the experiment. This information will be permanently stored in a secure file (paper form only) at CCIT. Only the principle investigator will have access to this file.

- If non-monetary compensation (e.g., course credit, services) will be offered, explain how it will be provided.

All 135 drivers and 15 alternates will receive a hands-free cellular phone kit as additional compensation for participating. Additionally, all participants and alternates who show up at 6:45am on the day of the experiment will be eligible to win one of four Nokia N95 cell phones. A drawing will take place after the end of the experiment.

- Discuss reasoning behind amount/method/terms of compensation, including appropriateness of compensation for the study population and avoiding undue influence to participate.

The compensation for the drivers is based on roughly a \$15 an hour wage for each hour from of the experiment, from 6:45-10:00pm. Payment may not be received immediately to allow for processing of the checks. Drivers will not receive any compensation if they withdraw before completion of the experiment, since the primary job for which they are hired is to drive the vehicle. More specifically, the terms of the compensation package are designed to prevent subjects who are unwilling to drive from enrolling in the experiment with the sole intent of earning a portion of the compensation.

There are measures in place to ensure drivers who find themselves unwilling or unable to complete their driving responsibilities during the day of the experiment. In addition to mandatory breaks after at most three hours of driving, drivers may request breaks or request exemption from the remaining driving portion of the experiment at any time. These requests will be automatically granted to ensure driver safety, and will not jeopardize the subject's compensation, provided they do not completely withdraw from the experiment.

The hands free cellular phone kit is included as a gift for participating in the experiment. This is an appropriate gift, since the devices must be used while talking on mobile phones while driving in the state of California. Furthermore, since the devices will be used during the experiment, they will have little residual value if collected after the experiment.

The raffle to win one of 4 Nokia N95 cellular phones (valued at \$650) is designed encourage the drivers and alternates to arrive on time on the day of the experiment. If a subject must withdraw from the experiment, they will be eligible to win the cellular phone, but will not receive the \$250 compensation or Bluetooth headset.

- **Costs to Subjects:** If applicable, describe any costs/charges which subjects or their insurance carriers will be expected to pay. (If there are no costs to subjects or their insurers, this should be stated.)

N/A

- **Treatment and Compensation for Injury:** *If the study involves more than minimal risk*, indicate that the researchers are familiar with and will follow University of California policy in this regard, and will use recommended wording on any consent forms (see [CPHS Informed Consent Guidelines](#)).

The experiment poses no additional risks to subjects other than those normally encountered in operating a vehicle on an urban freeway. Key personnel believe this risk to be minimal, but have put

additional measures in place to insure the safety of the drivers, which is necessary for the successful operation of the experiment.

Key personnel are familiar with and will follow University of California policy in this regard. As an added precaution, the following language will be used on the Informed Consent to Participate in Research:

If you are injured as a result of taking part in this study, medical care and treatment will be available. The costs of this care may or may not be covered by the University of California, depending on a number of factors. If you have any questions regarding this assurance, you may contact the Committee for the Protection of Human Subjects: University of California, Berkeley, 2150 Shattuck Ave., Suite 313, University of California, Berkeley, CA 94720-5940 (510) 642-7461, subjects@berkeley.edu.

SECTION 13: ADVERSE EVENT MANAGEMENT/REPORTING

- Explain how unanticipated negative outcomes/experiences or serious adverse events will be managed. (NOTE: This may apply in social-behavioral as well as biomedical research (e.g., undue stress or anxiety of subject, breach of confidentiality via loss of laptop computer with study data.) Provisions should be made and described here *if applicable*.)
- Describe plans for provision of treatment for study-related injuries, and how costs of injury treatment will be covered (see "Treatment and Compensation for Injury" above).
- Discuss plans for reporting unanticipated or serious adverse events to CPHS (see [CPHS Adverse Events](#)). (This applies to all types of research.)

In order to be responsive to adverse events that may occur, an ambulance with emergency medical personnel and a squad car with a police officer will be at the base station throughout the experiment. Risk management personnel will be under instruction to immediately inform the on-site police officer of any adverse event, who will provide counsel on the appropriate response. In case of emergency on the road, the ambulance and police officers will be able to respond immediately. They will also be able to respond to any medical or safety issues that arise at the base station.

If appropriate (as determined in consultation with the on-site police officer) drivers on the road will be notified of the adverse event through the Bluetooth headset.

To provide any needed roadside assistance, a tow truck will be available at the base station during the whole experiment. Although drivers will be provided scheduled breaks, and may take additional breaks if needed, it is possible a driver may decide to no longer participate in the experiment. The driver will be provided two options: 1) Stay at the experiment site and take the bus back to UC Berkeley at the end of the experiment, or 2) withdraw from the experiment completely. A member of the key personnel will be responsible for driving any subjects who withdraw back to UC Berkeley. Subjects selecting option 1 will still receive full compensation (as they have not withdrawn from the experiment), while subjects selecting option 2 will not receive compensation (they must indicate their intent to withdraw to any member of key personnel)

Provision of treatment for study-related injuries and costs of injury treatment will follow the University of California policy in this regard.

The report of any unanticipated or serious adverse events will be completed as soon as known by the principal investigator through the “Report of Serious Adverse Event or Unanticipated Problem” form in accordance with the Committee for Protection of Human Subject procedure.

SECTION 14: ATTACHMENTS

- Please list all attachments (e.g., consent forms, survey instruments, recruitment materials, appendices) included with your submission.

The following items are attached to this document in Appendix A:

- 1) Statement of Informed Consent
- 2) Photographic Consent Form
- 3) Experiment Location
- 4) CITI Training certificates
- 5) Recruitment Flier
- 6) Recruitment Email

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

Appendix A

1) Statement of Informed Consent

Please see next page.

UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO



SAN FRANCISCO • SANTA BARBARA • SANTA CRUZ

Informed Consent to Participate in Research
Cellular Phones for Traffic Estimation

Purpose and Background

A group of researchers (herein referred to as “we”) working under the supervision of Professor Alexandre Bayen in the Civil and Environmental Engineering Department at the University of California, Berkeley (UC Berkeley) and in collaboration with the California Center for Innovative Transportation (CCIT) are conducting a research experiment (herein referred to as “the experiment”) to collect speed data from 100 vehicles on a highway using cellular phones equipped with GPS units.

Procedures: What will happen to me if I take part?

If you volunteer to participate in this study, you will be asked to drive a rental car (provided to you) on a ten mile segment of I-880, near Union City, CA. On the day of the experiment, February 8, 2008, you will be asked to drive the segment (or a portion of the segment) continuously for at most 8 hours. One hour rest breaks will be provided after three hours of consecutive driving, although you may request additional breaks if necessary. With your permission, the vehicle will have a cellular phone on-board, which will be used to calculate and record the vehicle’s position and speed at all times while it is in your possession. This phone will be placed in a secure box to prevent you from accessing it. You will not be able to talk on or otherwise use this phone. Additionally, some or all of the position and speed data and your vehicle and cell phone ID number will be sent over the cell phone network to a secure server at UC Berkeley. This server will be accessed in near real-time by the researchers, and will be used to monitor your vehicle on the day of the experiment. You will be asked to wear a hands-free Bluetooth headset while you are driving to allow us to communicate with you while you are driving.

If you volunteer to participate in this research, you will be asked to meet at the West Gate of UC Berkeley at 6:45am on the day of the experiment. After a debriefing of the procedure for the day, we will provide you with a bus shuttle to a rental car agency at the Oakland International Airport. At this point, you will likely be asked to pick up a vehicle and drive it to the experiment site approximately 12 miles away near Union City, CA. All drivers will be asked to travel together under a police escort between the rental agency and the experiment site. If you are not asked to drive a vehicle at this stage, you will be shuttled to the experiment base station at the Union Landing retail entertainment center parking lot by bus. After arriving at the experiment site, you will be asked to drive the vehicle in scheduled shifts of no more than three hours. You will be asked to refill the gas tank at a specified gas station when the tank is approximately ¼ full, at no cost to you. You will be provided an hour break in between driving shifts, and you will not be asked to drive the segment for more than a total of 8 hours. At the conclusion of this phase of the experiment, all drivers will return to the base station before being asked to return their vehicles to the rental car agency, again under police escort. A shuttle will provide transportation from the rental agency back to UC Berkeley, and should arrive no later than 10:00pm.

Hiring Procedure

In order to participate in this experiment, you will be hired as an employee of UC Berkeley for the day of the experiment. You will be asked to complete hiring forms by the UC Berkeley Institute for Transportation Studies (ITS) human resources personnel. The first 135 subjects who meet the qualifications outlined below and complete the hiring paperwork will be considered "drivers" for the experiment. An additional 15 "alternates" will be hired for the purposes of filling in for drivers in the event they request an unscheduled break or withdraw from the experiment, and will be asked to be at the base station throughout experiment. Subjects will be notified of their initial status as a driver or alternate one week after hiring has been completed. All 135 drivers and 15 alternates are expected to complete an online training course 10 days prior to the experiment and arrive at the West Gate of UC Berkeley at 6:45am on the day of the experiment.

Benefits

There is no direct benefit to your participation in this experiment. It is hoped that your participation will enable researchers to develop the next generation of highway traffic sensing technologies, from which you will benefit indirectly.

Risks/Discomforts

The experiment poses no additional risks to you other than those normally encountered in operating a vehicle on an urban freeway. These risks include roughly a 0.08% chance to be involved in an automobile accident, and roughly a 0.0001% chance of being involved in a fatal accident, based on national highway statistics and the mileage they will be asked to drive. For comparison, the risk of being involved in an accident while riding a bicycle is roughly 0.031%, with roughly a 0.00005% chance it will be fatal. These figures have been calculated assuming you choose to ride a bicycle for 8 hours instead of driving a car for 8 hours. The annual number of bicycle-hours, accidents and fatalities in the US were obtained from the Consumer Product Safety Commission, and from the number of car accidents, fatalities and vehicle-miles traveled are obtained from the Bureau of Transportation Statistics. Although we cannot prevent all accidents, we will make every effort to help you minimize your risk.

As part of the experiment, your vehicle will be monitored in near real-time through the cellular phone in your vehicle. You are asked to drive as you would normally drive, in accordance with all traffic laws. If it is determined that you are driving at speeds in excess of 15 miles per hour over the posted speed limit as reported by the cellular phone in your vehicle, you will be asked to reduce your speed. If you choose not to comply, an on-site police officer will be notified, who may then choose to verify the speed of your vehicle using traditional tools. Because we recognize there are errors in the speeds reported by the phone, we will not use the data we collect in near-real time for the purposes of issuing traffic citations. Furthermore, this is not the objective of this research experiment.

In accordance with University of California, Berkeley standard hiring procedures, you will be asked to provide personal information, such as your name, phone number, email address, mailing address, social security number, and UC Berkeley student/employee ID number. All hiring information will be collected and stored by the ITS human resources office, and will be kept confidential.

As a driver, you will be assigned an ID number linking your name to the cellular phone recording your speed and position data during the experiment. The correspondence between your name and ID number will be kept confidential.

As with all research experiments, there is a small chance that the confidentiality of the data collected could be compromised, although we will take care to prevent this from happening.

Confidentiality: Who has access to the data?

Your voluntary participation in this experiment may involve a loss of privacy, but your personal data will be handled as confidentially as possible. The primary purpose of the experiment is to collect speed and location data of each vehicle, and this data will be used by researchers in presentations, reports, and publications. We will assign new random ID numbers to the data we collect as a further measure to protect your privacy. We may choose to sublicense portions of this de-identified speed and position data to other parties. However, we will not release any of your personally identifiable information, or the correspondence between your name and ID number.

All information related to your hiring will be collected and stored by the UC Berkeley ITS human resources personnel.

Data linking your name to the cellular phone and vehicle ID numbers will be collected for logistical purposes during the experiment. This information will be permanently stored in a secure file (paper form only) at the CCIT. Only the lead investigator will have access to this file.

If you decide not to participate in this research, all of your communication sent to NextGenCell@ce.berkeley.edu will be deleted no later than two weeks after the completion of the experiment.

We will implement all Committee for the Protection of Human Subjects recommendations to secure any personally identifiable data collected during this experiment.

Compensation

As a driver in this experiment you will receive a payment of \$250. You will receive your payment on your next scheduled paycheck from the university, if you are already a UC Berkeley employee. If you are not already a UC Berkeley employee, a check will be issued to you by ITS. You will be paid in full upon your completion of the experiment. In the event you withdraw before the experiment is complete, you will not receive the \$250 compensation.

All drivers will receive a hands-free Bluetooth headset, used during the experiment, as additional compensation for participating.

All drivers and alternates who show up at 6:45am on the day of the experiment will be eligible to win one of four Nokia N95 cell phones, valued at \$650. A drawing will take place after all drivers return to UC Berkeley on the day of the experiment.

If you must withdraw from the experiment on the day of the experiment, you will not be eligible for the \$250 compensation or the Bluetooth headset, but you will remain eligible for the Nokia N95 drawing.

As a driver or alternate, you will receive a catered lunch and dinner at the experiment base station during your rest breaks.

Costs to you

There are no expected costs to you for participating in the experiment.

Treatment and Compensation for Injury:

If you are injured as a result of taking part in this study, medical care and treatment will be available. The costs of this care may or may not be covered by the University of California, depending on a number of factors. If you have any questions regarding this assurance, you may contact the office of the Committee for Protection of Human Subjects: University of California, Berkeley, 2150 Shattuck Avenue, Suite 313, Berkeley, CA 94704-5940, (510) 642-7461, subjects@berkeley.edu.

Taking Part in this Study

Your participation in this research is voluntary, and you are free to refuse to participate or quit the experiment at any time. If you are unable to continue driving, you will be provided transportation back to UC Berkeley by bus at the end of the experiment. If you choose to withdraw from the experiment, a staff member will be available to drive you back to UC Berkeley. As outlined above, if you withdraw from the experiment, you will not receive the \$250 compensation or Bluetooth headset, but will still be eligible for the N95 phone raffle. Whether or not you chose to participate will have no bearing in relation to your standing in any department of the University of California, Berkeley.

Qualifications

To be eligible to participate, you must 1) be a UC Berkeley student, 2) be at least 21 years old, 3) have a valid driver's license, 4) be able to communicate fluently in English, 5) have driven at least 100 miles on California highways, and 6) be in good physical health. Please do not participate if you are pregnant, or think that you may otherwise be unable or unwilling to perform the duties briefly outlined above. We will ask that you be well rested on the day of the experiment. You will be asked to drive the vehicle as you would normally drive your own vehicle on the highway, in accordance with all traffic laws. We ask that you understand and comply with these laws. No inclusion/exclusion criteria will be based on gender, race, or ethnicity.

Photography

We may wish to photograph or film portions of this experiment. If you are willing to allow us to photograph or film you while you participate in this research, please indicate your consent by reading and signing the Photographic Consent Form. You may choose to participate regardless of your decision to authorize photographic consent.

Questions

If you have questions about the research, you may contact us at (510)-642-6060, or by email at nextGenCell@ce.berkeley.edu. Additionally, you may visit <http://nextgencell.googlepages.com> for additional information about this experiment. You may keep a copy of this form for reference.

Consent

I have read this consent form, considered all the pertinent information presented, meet all qualifications, and I wish to accept these terms and participate in this research.

Signature

Date

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

2) Photographic Consent Form

Photographic Consent Form

The undersigned does hereby authorize

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

and/or its associates, assistants, or subcontractors to photograph/film

Name (please print)

The undersigned authorizes The Regents of the University of California to permit the use and display of said photographs in any publication, multimedia production, display, advertisement or World-Wide Web Publication for **Civil and Environmental Engineering** or its constituent departments.

The undersigned agrees that the Regents of The University of California may use name, likeness, or biographical information supplied by the undersigned.

The undersigned releases and forever discharges The Regents of the University of California, its agents, officers and employees from any and all claims and demands arising out of or in connection with the use of said photographs / images, including but not limited to, any claims for invasion of privacy or defamation.

Accepted and Agreed:

Signature of Subject

Signature of Witness

Date

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

3) Experiment Location

Experiment Location



Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

4) CITI Training Certificate

CITI Collaborative Institutional Training Initiative

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.
Professor, University of Miami
Director Office of Research Education
CITI Course Coordinator

Human Research Curriculum Completion Report

Printed on Sunday, September 16, 2007

Learner: Christian Claudel (username: predictor_m9)

Institution: University of California, Berkeley

Contact Information: Phone: 1-510-725-8828

Email: christian.claudel@gmail.com

Group 2: Social and Behavioral Research Investigators and Key Personnel

Stage 1. Basic Course Passed on 09/16/07 (Ref# 1280946)

Required Modules

Date

completed Score

Introduction 09/16/07 no quiz

History and Ethical Principles - SBR 09/16/07 4/5 (80%)

Defining Research with Human Subjects - SBR 09/16/07 5/5 (100%)

The Regulations and The Social and Behavioral Sciences - SBR 09/16/07 4/6 (67%)

Assessing Risk in Social and Behavioral Sciences - SBR 09/16/07 5/5 (100%)

Informed Consent - SBR 09/16/07 4/5 (80%)

Privacy and Confidentiality - SBR 09/16/07 3/5 (60%)

Research with Prisoners - SBR 09/16/07 5/5 (100%)

Research with Children - SBR 09/16/07 5/5 (100%)

Research in Public Elementary and Secondary Schools - SBR 09/16/07 5/5 (100%)

International Research - SBR 09/16/07 3/5 (60%)

Internet Research - SBR 09/16/07 4/4 (100%)

Group Harms: Research With Culturally or Medically Vulnerable
Groups

09/16/07 3/3 (100%)

Workers as Research Subjects-A Vulnerable Population 09/16/07 4/4 (100%)

Hot Topics 09/16/07 no quiz

Conflicts of Interest in Research Involving Human Subjects 09/16/07 1/2 (50%)

University of California, Berkeley 09/16/07 no quiz

[Return](#)

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

CITI Collaborative Institutional Training Initiative

Human Research Curriculum Completion Report

Printed on Sunday, September 16, 2007

Learner: Juan Herrera (username: jchm)

Institution: University of California, Berkeley

Contact Information: Phone: 510-6427390

Email: jcherrera@berkeley.edu

Group 2: Social and Behavioral Research Investigators and Key Personnel

Stage 1. Basic Course Passed on 09/15/07 (Ref # 1273924)

Required Modules Date completed Score

Introduction 09/14/07 no quiz

History and Ethical Principles - SBR 09/14/07 4/5 (80%)

Defining Research with Human Subjects - SBR 09/15/07 5/5 (100%)

The Regulations and The Social and Behavioral

Sciences - SBR

09/15/07 6/6 (100%)

Assessing Risk in Social and Behavioral Sciences -

SBR

09/14/07 4/5 (80%)

Informed Consent - SBR 09/14/07 3/5 (60%)

Privacy and Confidentiality - SBR 09/14/07 3/5 (60%)

Research with Prisoners - SBR 09/14/07 3/5 (60%)

Research with Children - SBR 09/14/07 3/5 (60%)

Research in Public Elementary and Secondary

Schools - SBR

09/14/07 4/5 (80%)

International Research - SBR 09/14/07 4/5 (80%)

Internet Research - SBR 09/14/07 3/4 (75%)

Group Harms: Research With Culturally or

Medically Vulnerable Groups

09/14/07 3/3 (100%)

Workers as Research Subjects-A Vulnerable

Population

09/14/07 4/4 (100%)

Hot Topics 09/14/07 no quiz

Conflicts of Interest in Research Involving Human

Subjects

09/15/07 2/2 (100%)

University of California, Berkeley 09/15/07 no quiz

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

CITI Collaborative Institutional Training Initiative

Human Research Curriculum Completion Report

Printed on Sunday, September 16, 2007

Learner: Sebastien Blandin (username: sebastien)

Institution: University of California, Berkeley

Contact Information: Phone: 510-642-9569

Email: sebastie@gmail.com

Group 2: Social and Behavioral Research Investigators and Key Personnel

Stage 1. Basic Course Passed on 09/16/07 (Ref # 1280920)

Required Modules Date

completed Score

Introduction 09/16/07 no quiz

History and Ethical Principles - SBR 09/16/07 4/5 (80%)

Defining Research with Human Subjects - SBR 09/16/07 5/5 (100%)

The Regulations and The Social and Behavioral Sciences -

SBR 09/16/07 6/6 (100%)

Assessing Risk in Social and Behavioral Sciences - SBR 09/16/07 5/5 (100%)

Informed Consent - SBR 09/16/07 4/5 (80%)

Privacy and Confidentiality - SBR 09/16/07 4/5 (80%)

Research with Prisoners - SBR 09/16/07 5/5 (100%)

Research with Children - SBR 09/16/07 4/5 (80%)

Research in Public Elementary and Secondary Schools -

SBR 09/16/07 5/5 (100%)

International Research - SBR 09/16/07 5/5 (100%)

Internet Research - SBR 09/16/07 4/4 (100%)

Group Harms: Research With Culturally or Medically

Vulnerable Groups 09/16/07 3/3 (100%)

Workers as Research Subjects-A Vulnerable Population 09/16/07 4/4 (100%)

Hot Topics 09/16/07 no quiz

Conflicts of Interest in Research Involving Human Subjects 09/16/07 2/2 (100%)

University of California, Berkeley 09/16/07 no quiz

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.

Professor, University of Miami

Director Office of Research Education

CITI Course Coordinator

[Return](#)

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

CITI Collaborative Institutional Training Initiative

Human Research Curriculum Completion Report

Printed on Sunday, September 16, 2007

Learner: Arthur Wiedmer (username: arthur.wiedmer)

Institution: University of California, Berkeley

Contact Information: Department: CEE

Phone: 5106424522

Email: arthur.wiedmer@berkeley.edu

Group 2: Social and Behavioral Research Investigators and Key Personnel

Stage 1. Basic Course Passed on 09/16/07 (Ref# 1272823)

Required Modules

Date

completed Score

Introduction 09/12/07 no quiz

History and Ethical Principles - SBR 09/12/07 5/5 (100%)

Defining Research with Human Subjects - SBR 09/12/07 5/5 (100%)

The Regulations and The Social and Behavioral Sciences

- SBR

09/16/07 5/6 (83%)

Assessing Risk in Social and Behavioral Sciences - SBR 09/16/07 5/5 (100%)

Informed Consent - SBR 09/16/07 5/5 (100%)

Privacy and Confidentiality - SBR 09/16/07 4/5 (80%)

Research with Prisoners - SBR 09/16/07 5/5 (100%)

Research with Children - SBR 09/16/07 5/5 (100%)

Research in Public Elementary and Secondary Schools -

SBR

09/16/07 4/5 (80%)

International Research - SBR 09/16/07 5/5 (100%)

Internet Research - SBR 09/16/07 4/4 (100%)

Group Harms: Research With Culturally or Medically

Vulnerable Groups

09/16/07 3/3 (100%)

Workers as Research Subjects-A Vulnerable Population 09/16/07 4/4 (100%)

Hot Topics 09/16/07 no quiz

Conflicts of Interest in Research Involving Human

Subjects

09/16/07 1/2 (50%)

University of California, Berkeley 09/16/07 no quiz

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.

Professor, University of Miami

Director Office of Research Education

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

CITI Collaborative Institutional Training Initiative

Human Research Curriculum Completion Report

Printed on Friday, September 14, 2007

Learner: Dan Work (username: dbwork)

Institution: University of California, Berkeley

Contact Information: Department: Civil and Environmental Engineering

Phone: 510) 642-9278

Email: dbwork@berkeley.edu

Group 2: Social and Behavioral Research Investigators and Key Personnel

Stage 1. Basic Course Passed on 09/14/07 (Ref # 1272584)

Required Modules

Date

completed Score

Introduction 09/12/07 no quiz

History and Ethical Principles - SBR 09/12/07 4/5 (80%)

Defining Research with Human Subjects - SBR 09/12/07 3/5 (60%)

The Regulations and The Social and Behavioral Sciences -

SBR

09/12/07 6/6 (100%)

Assessing Risk in Social and Behavioral Sciences - SBR 09/12/07 4/5 (80%)

Informed Consent - SBR 09/12/07 4/5 (80%)

Privacy and Confidentiality - SBR 09/12/07 5/5 (100%)

Research with Prisoners - SBR 09/12/07 4/5 (80%)

Research with Children - SBR 09/14/07 4/5 (80%)

Research in Public Elementary and Secondary Schools -

SBR

09/14/07 4/5 (80%)

International Research - SBR 09/14/07 5/5 (100%)

Internet Research - SBR 09/14/07 3/4 (75%)

Group Harms: Research With Culturally or Medically

Vulnerable Groups

09/14/07 2/3 (67%)

Workers as Research Subjects-A Vulnerable Population 09/14/07 3/4 (75%)

Hot Topics 09/14/07 no quiz

Conflicts of Interest in Research Involving Human Subjects 09/14/07 1/2 (50%)

University of California, Berkeley 09/12/07 no quiz

<https://www.citiprogram.org/members/learners/...=94A3C66A-486D-4924-8C24-03F3D921D071-1327656> (1 of 2) [9/13/2007 10:26:19 PM]

Completion Report

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.

Professor, University of Miami

Director Office of Research Education

CITI Course Coordinator

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

CITI Collaborative Institutional Training Initiative

Human Research Curriculum Completion Report

Printed on Sunday, September 16, 2007

Learner: Qingfang Wu (username: qingfang)

Institution: University of California, Berkeley

Contact Information: Phone: 510-642-9569

Email: qingfangwu@berkeley.edu

Group 2: Social and Behavioral Research Investigators and Key Personnel

Stage 1. Basic Course Passed on 09/16/07 (Ref# 1281127)

Required Modules

Date

completed Score

Introduction 09/16/07 no quiz

History and Ethical Principles - SBR 09/16/07 5/5 (100%)

Defining Research with Human Subjects - SBR 09/16/07 5/5 (100%)

The Regulations and The Social and Behavioral Sciences - SBR 09/16/07 6/6 (100%)

Assessing Risk in Social and Behavioral Sciences - SBR 09/16/07 5/5 (100%)

Informed Consent - SBR 09/16/07 5/5 (100%)

Privacy and Confidentiality - SBR 09/16/07 5/5 (100%)

Research with Prisoners - SBR 09/16/07 5/5 (100%)

Research with Children - SBR 09/16/07 5/5 (100%)

Research in Public Elementary and Secondary Schools - SBR 09/16/07 5/5 (100%)

International Research - SBR 09/16/07 5/5 (100%)

Internet Research - SBR 09/16/07 4/4 (100%)

Group Harms: Research With Culturally or Medically Vulnerable

Groups

09/16/07 3/3 (100%)

Workers as Research Subjects-A Vulnerable Population 09/16/07 4/4 (100%)

Hot Topics 09/16/07 no quiz

Conflicts of Interest in Research Involving Human Subjects 09/16/07 2/2 (100%)

University of California, Berkeley 09/16/07 no quiz

For this Completion Report to be valid, the learner listed above must be affiliated

with a CITI participating institution. Falsified information and unauthorized use of

file:///C:/Users/Dan Work/AppData/Local/Temp/Temp1_qingfang.zip/crbystage.asp.htm (1 of 2)

[9/17/2007 4:40:25 AM]

Completion Report

the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.

Professor, University of Miami

Director Office of Research Education

CITI Course Coordinator

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

CITI Collaborative Institutional Training Initiative

Human Research Curriculum Completion Report

Printed on Sunday, September 16, 2007

Learner: Tarek Rabbani (username: trabbani)

Institution: University of California, Berkeley

Contact

Information:

1709 shattuck ave.

Berkeley, CA 94709 USA

Phone: 4084605810

Email: trabbani@berkeley.edu

Group 2: Social and Behavioral Research Investigators and Key Personnel

Stage 1. Basic Course Passed on 09/16/07 (Ref# 1281186)

Required Modules

Date

completed Score

Introduction 09/16/07 no quiz

History and Ethical Principles - SBR 09/16/07 5/5 (100%)

Defining Research with Human Subjects - SBR 09/16/07 3/5 (60%)

The Regulations and The Social and Behavioral Sciences -

SBR

09/16/07 6/6 (100%)

Assessing Risk in Social and Behavioral Sciences - SBR 09/16/07 5/5 (100%)

Informed Consent - SBR 09/16/07 5/5 (100%)

Privacy and Confidentiality - SBR 09/16/07 4/5 (80%)

Research with Prisoners - SBR 09/16/07 5/5 (100%)

Research with Children - SBR 09/16/07 5/5 (100%)

Research in Public Elementary and Secondary Schools -

SBR

09/16/07 5/5 (100%)

International Research - SBR 09/16/07 5/5 (100%)

Internet Research - SBR 09/16/07 4/4 (100%)

Group Harms: Research With Culturally or Medically

Vulnerable Groups

09/16/07 3/3 (100%)

Workers as Research Subjects-A Vulnerable Population 09/16/07 4/4 (100%)

Hot Topics 09/16/07 no quiz

Conflicts of Interest in Research Involving Human Subjects 09/16/07 2/2 (100%)

University of California, Berkeley 09/16/07 no quiz

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.

Professor, University of Miami

Director Office of Research Education

CITI Course Coordinator

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

5) *Recruitment Flier*

GPS Cell Phone for traffic estimation
Research Experiment

Earn \$250 for a day of driving
A chance to win 1 of 4 Nokia N95 phones (with GPS) !!!

Help create the next generation of cell phone "location-based services"

Contribute to a better understanding of highway congestion in California
Participate in the first large scale experiment using cell phones to measure real-time highway congestion

- 135 Drivers Needed for Research Experiment
- Must be at least 21, and a UC Berkeley student
- Must have valid driver's license and good driving record

• UC Berkeley is developing tools to estimate traffic conditions using cellular phones in vehicles. You will drive a rental car for 8 hours in 3 shifts to help us collect data and test our system.

• For more information:
email nextGenCell@ce.berkeley.edu
or visit nextGenCell.googlepages.com

Committee for Protection of Human Subjects (CPHS)
Office for the Protection of Human Subjects (OPHS)
University of California, Berkeley
2150 Shattuck Ave., Suite 313, Berkeley, CA 94704-5940

Phone: 510/642-7461
Fax: 510/643-6272
E-mail: ophs@berkeley.edu
<http://cphs.berkeley.edu>

6) Recruitment Email

Recruitment Email

Dear colleagues,

My research group will be conducting a cell-phoned based highway monitoring experiment, for which we will recruit 135 drivers. There is a compensation of \$250 for each driver, a free Bluetooth headset, and a chance to win a Nokia N95 cell phone (value \$650). I would be grateful if you could circulate this to your students, along with the attached flyer and URL which contain the details of the experiments and procedures.

<http://nextGenCell.googlepages.com/>

Thank you!
Sincerely,

Alex Bayen

--

Alexandre M. Bayen
Assistant Professor
Department of Civil and Environmental Engineering
University of California at Berkeley
711 Davis Hall
Berkeley, CA 94720-1710

Tel: (510) 642-2468
Fax: (510) 643-5264
Email: bayen@berkeley.edu
URL: <http://www.ce.berkeley.edu/~bayen/>

[Attachment: The recruitment flier will be attached to this email]

Reference

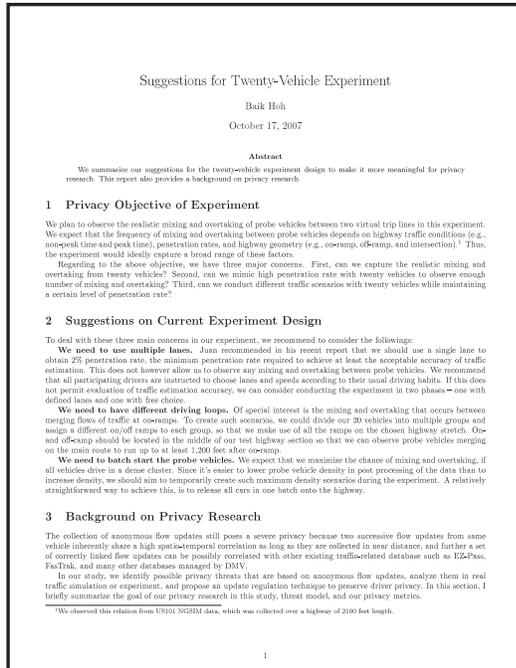
-
- i T. Nakatsuji C. Nanthawichit and H. Suzuki. Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway. *Transportation Research Record*, 1855:49-59, 2003
 - ii Y.B. Yim J-L. Ygnace, C. Drane and R. de Lacvivier. Travel time estimation on the San Francisco Bay Area network using cellular phones as probes. California PATH Working Paper UCB-ITS-PWP-2000-18, Institute of Transportation Studies, University of California, Berkeley, 2000
 - iii R. Litjens M. Westerman and J-P. Linnartz. Integration of probe vehicle and induction loop data estimation of travel times and automatic incident detection. PATH Research Report UCB-ITS-PRR-96-13, Institute of Transportation Studies, University of California, Berkeley, June 1996
 - iv G. B. Rodgers. An Overview of the Bicycle Study. Bicycle Safety Publications, US Consumer Product Safety Commission. Available online:
<http://www.cpsc.gov/CPSCPUB/PUBS/bikepubs.html>
 - v. Table 2-17: Motor Vehicle Safety Data. National Transportation Statistics. Bureau of Transportation Statistics, 2002. Available online:
http://www.bts.gov/publications/national_transportation_statistics/2002/html/table_02_17.htm

2 Deployment Prototyping

Collected in Appendix 2 are descriptions of prototyping efforts that took place during the initial planning stages of Mobile Century. Documentation for the twenty-vehicle deployment of November 2, 2007 is presented in four parts:

The overall strategy and objectives of the twenty-vehicle deployment are described in Appendix 2.1. The actual protocol is provided in Appendix 2.2. Participants were divided into two groups (even, and odd). Detailed instructions for each group are given in Appendices 2.3 and 2.4.

2.1 Strategy and Objectives



Suggestions for Twenty-Vehicle Experiment

Baik Hoh

October 17, 2007

Abstract

We summarize our suggestions for the twenty-vehicle experiment design to make it more meaningful for privacy research. This report also provides a background on privacy research.

1 Privacy Objective of Experiment

We plan to observe the realistic mixing and overtaking of probe vehicles between two virtual trip lines in this experiment. We expect that the frequency of mixing and overtaking between probe vehicles depends on highway traffic conditions (e.g., non-peak time and peak time), penetration rates, and highway geometry (e.g., on-ramp, off-ramp, and intersection).¹ Thus, the experiment would ideally capture a broad range of these factors.

Regarding to the above objective, we have three major concerns. First, can we capture the realistic mixing and overtaking from twenty vehicles? Second, can we mimic high penetration rate with twenty vehicles to observe enough number of mixing and overtaking? Third, can we conduct different traffic scenarios with twenty vehicles while maintaining a certain level of penetration rate?

2 Suggestions on Current Experiment Design

To deal with these three main concerns in our experiment, we recommend to consider the followings:

We need to use multiple lanes. Juan recommended in his recent report that we should use a single lane to obtain 2% penetration rate, the minimum penetration rate required to achieve at least the acceptable accuracy of traffic estimation. This does not however allow us to observe any mixing and overtaking between probe vehicles. We recommend that all participating drivers are instructed to choose lanes and speeds according to their usual driving habits. If this does not permit evaluation of traffic estimation accuracy, we can consider conducting the experiment in two phases – one with defined lanes and one with free choice.

We need to have different driving loops. Of special interest is the mixing and overtaking that occurs between merging flows of traffic at on-ramps. To create such scenarios, we could divide our 20 vehicles into multiple groups and assign a different on/off ramps to each group, so that we make use of all the ramps on the chosen highway stretch. On- and off-ramp should be located in the middle of our test highway section so that we can observe probe vehicles merging on the main route to run up to at least 1,200 feet after on-ramp.

We need to batch start the probe vehicles. We expect that we maximize the chance of mixing and overtaking, if all vehicles drive in a dense cluster. Since it's easier to lower probe vehicle density in post processing of the data than to increase density, we should aim to temporarily create such maximum density scenarios during the experiment. A relatively straightforward way to achieve this, is to release all cars in one batch onto the highway.

3 Background on Privacy Research

The collection of anonymous flow updates still poses a severe privacy because two successive flow updates from same vehicle inherently share a high spatio-temporal correlation as long as they are collected in near distance, and further a set of correctly linked flow updates can be possibly correlated with other existing traffic-related database such as EZ-Pass, FasTrak, and many other databases managed by DMV.

In our study, we identify possible privacy threats that are based on anonymous flow updates, analyze them in real traffic simulation or experiment, and propose an update regulation technique to preserve driver privacy. In this section, I briefly summarize the goal of our privacy research in this study, threat model, and our privacy metrics.

¹We observed this relation from US101 NGSIM data, which was collected over a highway of 2100 feet length.

3.1 Design Goals

We aim to achieve privacy protection by design so that no one, not even service provider, can identify or track a user from anonymous flow updates. We design a privacy-preserving placement algorithm that determines the positions of virtual trip lines in real time while observing the exclusion area and the minimum spacing between two successive virtual trip lines to preserve driver privacy. On top of the placement algorithm, we derive additional mechanism that controls the flow updates (e.g., shifting timestamp, quantizing speed, or suppressing the update) on a client side depending on both the traffic disturbance and the density of probe vehicles, where traffic disturbance indicates how many mixing and overtaking take place. Key idea underlying the update control mechanism is that clients need to accurately estimate both traffic disturbance and the density of probe vehicles by themselves instead of having a trustworthy location proxy server² that distributes those information to users.

3.2 Threat Model

We assume in our study two different attacks, reconstructing the speed profile and reconstructing the complete trajectory of a probe vehicle from anonymous flow updates. An adversary (or highway patrol) can reconstruct the speed history of a probe vehicle up to a few miles back by linking several consecutive anonymous flow updates. If two successive virtual trip lines are placed close to each other, two successive flow updates of the same probe vehicles have high correlation as long as not a severe external disturbance (e.g., high traffic from ramps) or turbulence (e.g., car accidents) is applied to traffic stream. We expect that the correct association of two anonymous flow updates from the same probe vehicles depends on penetration rate, traffic conditions (e.g., traffic volume, density, and average speed), and geometry of highways. We refer to the probability of a correct association as *tracking probability*. The figure 1 shows a simple example of two probe vehicles in which an adversary guesses the correct match of four anonymous flow updates collected from two successive virtual trip lines.

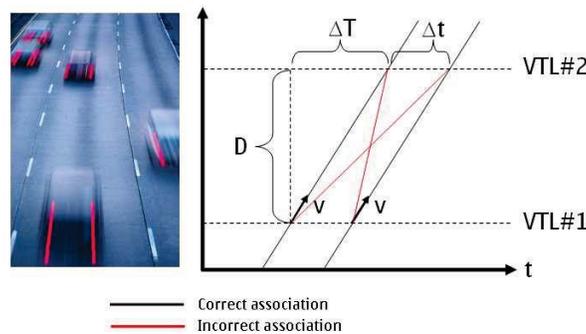


Figure 1: Linking of Two Anonymous VTL Flow Updates

An adversary can also identify which flow updates are from probe vehicles merging from on-ramp or diverging to off-ramp by looking at reported speeds. It is known that there exists an average speed difference between merging traffic and main traffic around ramps, and this speed difference decreases as distance from ramps increases. The figure 2 shows the typical trajectories of three different traffic flows at an on-ramp: the red straight line is from main traffic in inner lanes, another red curved line is from main traffic in outer lanes, the other blue curved line is from the ramp traffic. If a virtual trip line is placed near an on-ramp, an adversary can tell three different traffics apart from their recorded speed measurements. An adversary finally concludes that the flow update with the lowest speed measurement is likely to originate from the nearest entrance (or ramp), which means that the origin or the destination of highway trip can be identified at least in terms of Exit numbers. Even though it is out of scope in our study, the Exit numbers of origin and destination can be further correlated with EZ-Pass (in New Jersey) or FasTrak database (in California).

3.3 Example Metrics

To evaluate privacy, we plan to measure from this experiment how typical mixing of highway traffic and the distance between virtual trip lines affect an adversary's ability to reconstruct the trace. For example, give the scenario in figure 1,

²We proposed the centralized approach with the help of a trustworthy location proxy server in ACM CCS2007 in automotive traffic monitoring applications. In our current study, we aim to achieve similar level of privacy protection in a decentralized way since the existence of a trustworthy entity is not preferable to participants of traffic monitoring applications, who are afraid of insider attack.

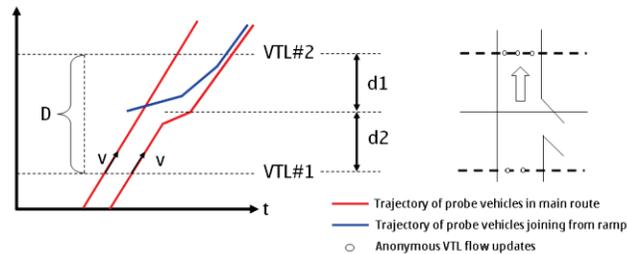


Figure 2: Identifying VTL Flow Updates from Ramps

we can vary the distance, D between the trip lines and measure how it affects the ability to reconstruct paths. An example metric we may use is the **distance to confusion**. It is defined as the distance between both ends of correctly linked (anonymous) flow updates by an adversary with the confidence level that is higher than a pre-defined threshold. We also measure the distribution of tracking distance to identify the median and the percentage of outliers.

In contrast to figure 1, we are more interested in minimum spacings ($d1$ and $d2$) from ramps than D when an adversary figures out the flow updates from ramp traffic as shown in figure 2. In this case, we measure the ability of the correct identification by the **ratio of correct identification**. It is defined as the ratio ($\frac{k}{n}$) of the number of flow updates, k that are correctly identified as flow updates from merging traffic or diverging traffic by an adversary with enough high confidence to a given n anonymous flow updates recorded on each virtual trip line. We need to investigate the effect of $D(= d1 + d2)$ on the number of mixings of probe vehicles. Whether increasing D around ramps helps confuse an adversary or not is of another interest during our experiment.

In both metrics, we measure the privacy risks against the pre-defined confidence level. Confidence level indicates how certain an adversary is about his or her association in hypothesis testing. This metric captures how much information an adversary achieves compared to the case of uniform probabilities over all possible associations by using the uncertainty concept in information theory.

2.2 Protocol for 20-Vehicle Deployment

Protocol for the 20 cars experiment

Objectives of the experiment

- Observe the mixing and overtaking behavior of probe vehicles between two virtual trip lines.
- Find any problem that might arise in data collection/storage process.
- Collect traffic data to evaluate the algorithms to reconstruct density/velocity field.

Section to be used

The experiment will take place in a section of I880, between SR-92 to the north and Alvarado-Niles to the south (Figure 1). The length of this section is 4 miles.



1

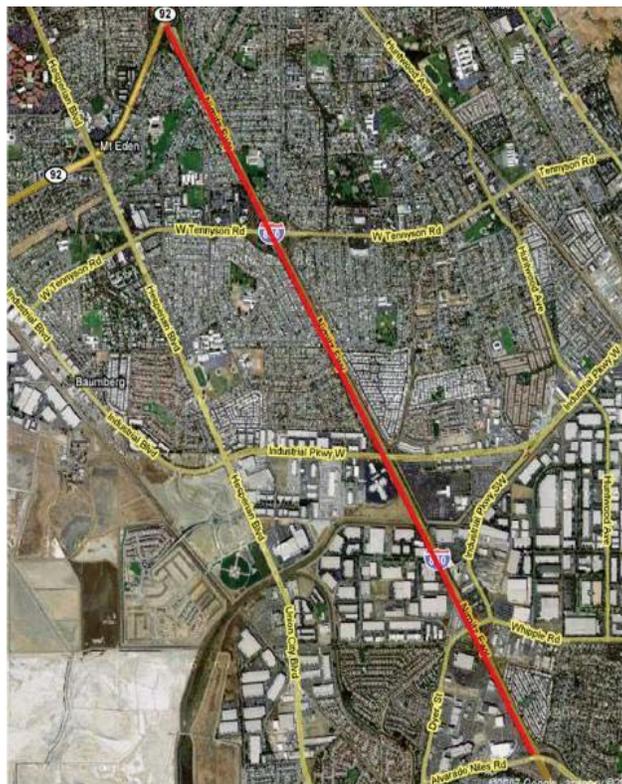
Protocol for the 20 cars experiment

Objectives of the experiment

- Observe the mixing and overtaking behavior of probe vehicles between two virtual trip lines.
- Find any problem that might arise in data collection/storage process.
- Collect traffic data to evaluate the algorithms to reconstruct density/velocity field.

Section to be used

The experiment will take place in a section of I880, between SR-92 to the north and Alvarado-Niles to the south (Figure 1). The length of this section is 4 miles.



A2.2 Figure 1: The 4 mile section to be used in the 20 cars experiment.

Day/time of experiment

The experiment will be done on November 2nd. It will begin at 1:45pm and it will last 1.5 hours (until 3:15pm). At the time of the experiment, the NB direction is expected to present congestion, while the traffic in the SB direction should be smooth.

Loops description

There will be two loops. The *long loop* (LL) goes between Alvarado-Niles and SR92, and the *short loop* (SL) goes between Alvarado-Niles and Tennyson Rd. Table 1 shows the length and expected travel time of each loop.

Table 2: Characteristics of each loop.

Loop	Length (miles)	Cycle time (min)
LL	8	20
SL	6	15

Figures 2 to 4 show in detail the end of both loops.



A2.2 Figure 2: Detail of the Alvarado-Niles interchange and the mall.

The day of the experiment

All drivers will meet at 1:35pm at the parking lot of the mall located next to I880S in Union City, between Alvarado-Niles and Whipple Rd. (the star in Figure 5 shows the exact meeting point). At this point, the N95 cell-phones will be distributed among the drivers. Cell-phones will be



A2.2 Figure 3: Detail of the Tennyson Rd. interchange.

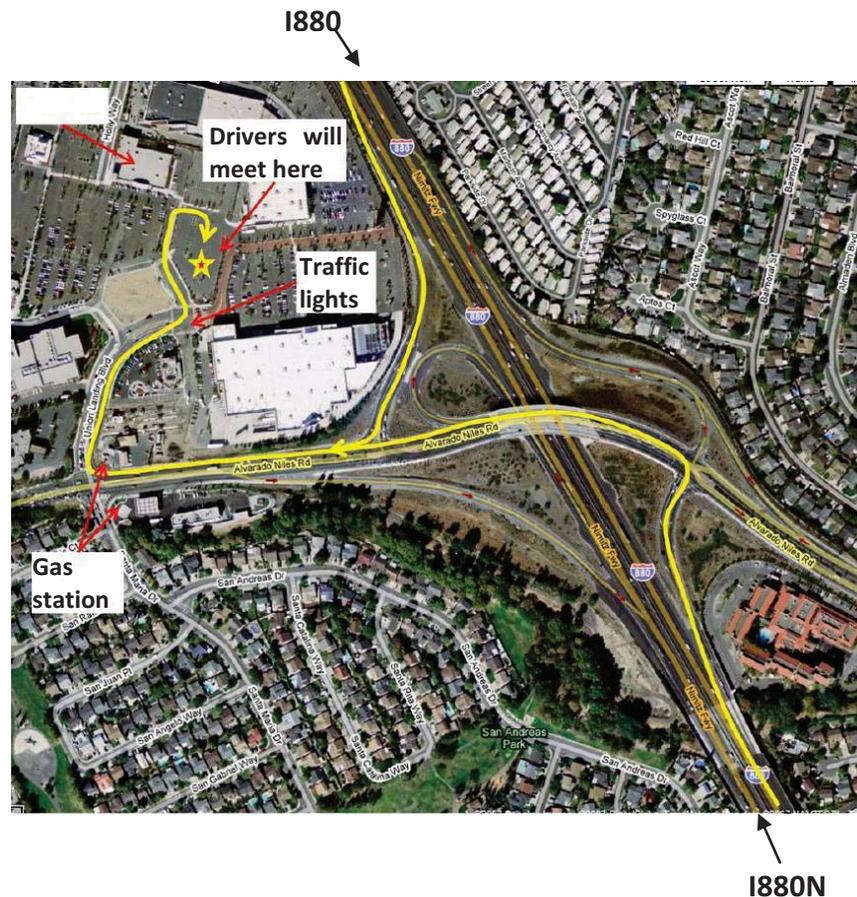


A2.2 Figure 4: Detail of the SR92 interchange.

ready-to-go, so drivers only need to tape the cell-phone to the dashboard (material will be provided).

Each cell-phone will contain an emergency number in its memory. The phone number is (650) 521 7940. In case of any problem during the experiment (problem with the phone, with car, with the instructions, etc.), drivers can call this number.

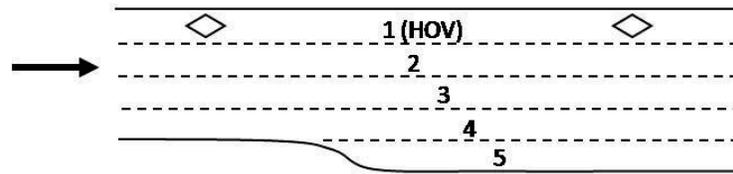
Let us identify the 10 drivers from Nokia as **Nokia01-10** and the 10 drivers from Berkeley as UCB01-10.



A2.2 Figure 5: From I880S and I880N to the parking lot..

Release of vehicles and driving. At the very beginning of the experiment, vehicles will be entering the section from the Alvarado-Niles on-ramp (see Figure 2). First, all drivers from Nokia will be released from the parking lot at a rate of 1 vehicles per minute (i.e. **Nokia01, Nokia02...**

Nokia10). Then, all UCB driver will start entering the section at the same rate (i.e. UCB01, UCB02... UCB10). **All the vehicles will drive on the LL during the first lap using only lane 4²⁴.**



A2.2 Figure 6: Lane numbering.

Once the first lap is completed, drivers will be divided into two sets. *Even* drivers²⁵ will start driving on the SL for the rest of the experiment. The rest of the drivers, i.e. the *odd* drivers²⁶, will keep driving on the LL. Only for the second lap, all the drivers will be still using lane 4 only. **Once the second lap is completed, drivers will have the freedom to choose lane (among lane 2, and/or 4), and they can change lanes while traveling the section if they want to.**

Each even driver (either from Nokia or UCB) is expected to complete 5 SLs after the completion of the first LL. Odd drivers from Nokia and UCB are expected to complete 5 LLs in total. Table 2 summarizes the experiment.

Table 3: Description of the laps during the experiment.

		Lap						Expected Driving time (min)
		1	2	3	4	5	6	
<i>Even drivers</i>	Loop	LL	SL	SL	SL	SL	SL	95
	Lane	4	4	2-3-4	2-3-4	2-3-4	2-3-4	
<i>Odd drivers</i>	Loop	LL	LL	LL	LL	LL	-	100
	Lane	4	4	2-3-4	2-3-4	2-3-4	-	

During the first two laps, we expect to achieve a penetration rate of 3% of the total flow on lane 4 (NB direction). When multiple lanes are used (laps 3 to 6), the penetration rate will go down to 1-2%.The total flow of probe vehicles between Alvarado-Niles and Tennyson Rd. (per direction) is expected to be around 75 vph, and the expected number of probe vehicles per direction at a given time is close to 9.

²⁴ Lane 1 corresponds to the leftmost lane, which is a HOV lane. See Figure 6

²⁵ Nokia and UCB02-04-06-08-10.

²⁶ Nokia and UCB01-03-05-07-09.

Because of traffic conditions (and the use of two loops), we expect that mixing and overtaking will happen. We also expect some vehicles to group naturally as the experiment progresses. Vehicles in the SL will be exiting and entering the section at a middle point, which is also a desirable feature for privacy preservation test.

At the end of the experiment, all drivers will meet again at the parking lot. Cell-phones will be returned at this point. The cell phones will be taken back to NRCPA.

If the driver has not completed all the required loops after 2 hours of driving, he/she will take the most convenient route in order to go back to the parking lot. The countdown starts with the first driver leaving the parking lot.

2.3 Instructions for Odd Drivers

Instructions and directions: Odd drivers

Even drivers are the following: Nokia01, Nokia 03, Nokia05, Nokia 07, Nokia09, UC001, UC003, UC005, UC007, and UC009 (see Appendix 1).

Drivers will meet at the Union Landing retail entertainment center parking lot at 1:35pm on Friday November 2nd (star in Figure 1). The entertainment center is in Union City, next to I8805 between Whipple Rd and Alvarado-Niles Rd.

Directions to get to the parking lot (see Figure 1):

- From 8805:
 - Take exit 23 for Alvarado-Niles Rd.
 - Turn right at Alvarado-Niles Rd.
 - Turn right at Union Landing Blvd.
- From 8809:
 - Take exit 23 for Alvarado-Niles Rd.
 - Turn left at Alvarado-Niles Rd.
 - Turn right at Union Landing Blvd.

Once at Union Landing Blvd., turn left at the traffic light (see Figure 1).

At the parking lot, drivers will receive the N95 cell phone, which has to be attached to the dashboard of the vehicle (material to attach the cell-phone to the dashboard will be provided).

Instructions on how to call the emergency number using the N95 will be also provided at this point (one click call). The phone number is (650) 521 7940.

Instructions and directions: *Odd* drivers

Odd drivers are the following: Nokia01, Nokia 03, Nokia05, Nokia 07, Nokia09, UCB01, UCB03, UCB05, UCB07, and UCB09 (see Appendix 1).

Drivers will meet at the Union Landing retail entertainment center parking lot at 1:35pm on Friday November 2nd (star in Figure 1). The entertainment center is in Union City, next to I880S between Whipple Rd and Alvarado-Niles Rd.

Directions to get to the parking lot (see Figure 1):

- *From I880S:*
 - Take exit 23 for Alvarado-Niles Rd.
 - Turn right at Alvarado-Niles Rd.
 - Turn right at Union Landing Blvd.
- *From I880N:*
 - Take exit 23 for Alvarado-Niles Rd.
 - Turn left at Alvarado-Niles Rd.
 - Turn right at Union Landing Blvd.

Once at Union Landing Blvd., turn left at the traffic light (see Figure 1).

At the parking lot, drivers will receive the N95 cell phone, which has to be attached to the dashboard of the vehicle (material to attach the cell-phone to the dashboard will be provided).

Instructions on how to call the emergency number using the N95 will be also provided at this point (one click call). **The phone number is (650) 521 7940.**



A2.3 Figure 1: From 1880S and I880M to the parking lot.

Drivers will be told when to leave the parking lot, but they should be ready to leave the parking lot at 1:45pm. That is, the cell-phone has to be installed inside the vehicle, and the vehicle should have enough gas to drive 40 miles²⁷. The very first vehicle will be released at 1:45pm.

From the parking lot to the freeway (Figure 2):

- Leave the mall using Union Landing Blvd.
- Turn left at Alvarado-Niles Rd.
- Turn left to merge onto I880N toward Oakland.

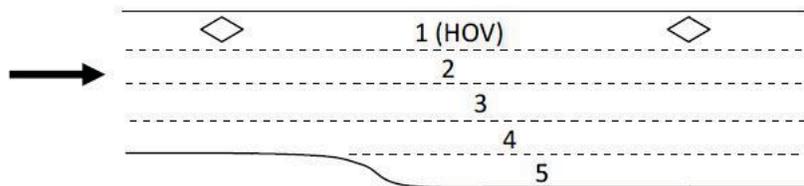
²⁷ There are two gas stations at the corner of Alvarado-Niles Rd. and Union Landing Blvd.



Drivers will meet here

A2.3 Figure 2: From the parking lot onto the freeway.

First, all drivers will go on the long loop using lane 4 (lane 1 is the leftmost lane, HOV lane in our case. See Figure 3).



A2.3 Figure 3: Lane numbering.

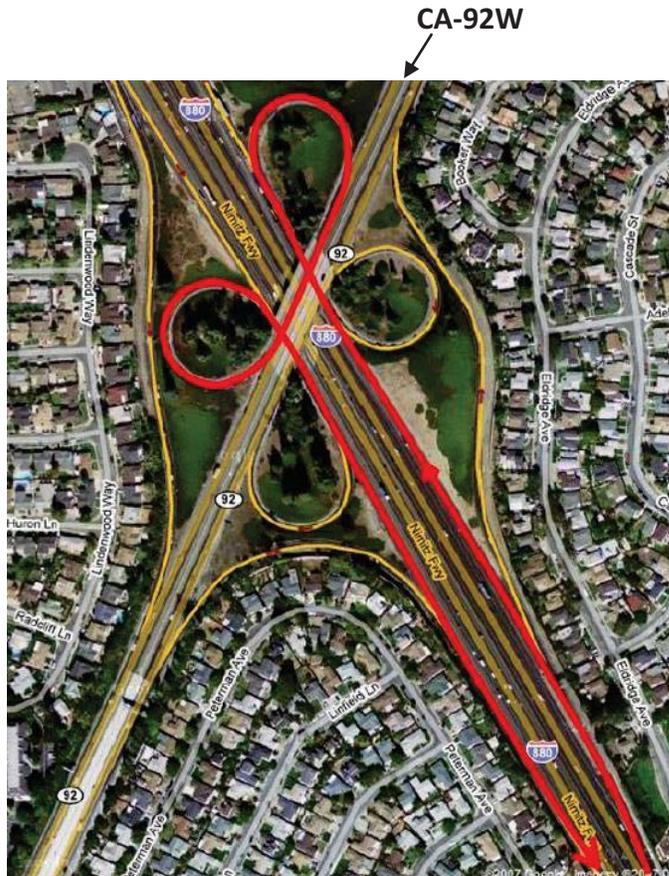
Directions for the Long loop:

- Once on I880NN, drive for 3.5 miles approx.
- Take exit onto CA-92W toward San Mateo Bridge
- Take the I880SS exit toward San Jose. Drive for 3.5 miles approx.

} See Figure 4

- Take exit 23 for Alvarado-Niles Rd.
- Turn left at Alvarado-Niles Rd.
- Turn left to merge onto I880N toward Oakland.

} See Figure 6



A2.3 Figure 4: Turn back at CA-92 (north end of the long loop).

Odd drivers have to complete 5 long loops in total. For the other 3 long loops (i.e., from lap 3 to 5), *odd* drivers can use either lane 2,3 and/or 4. (see table 1).

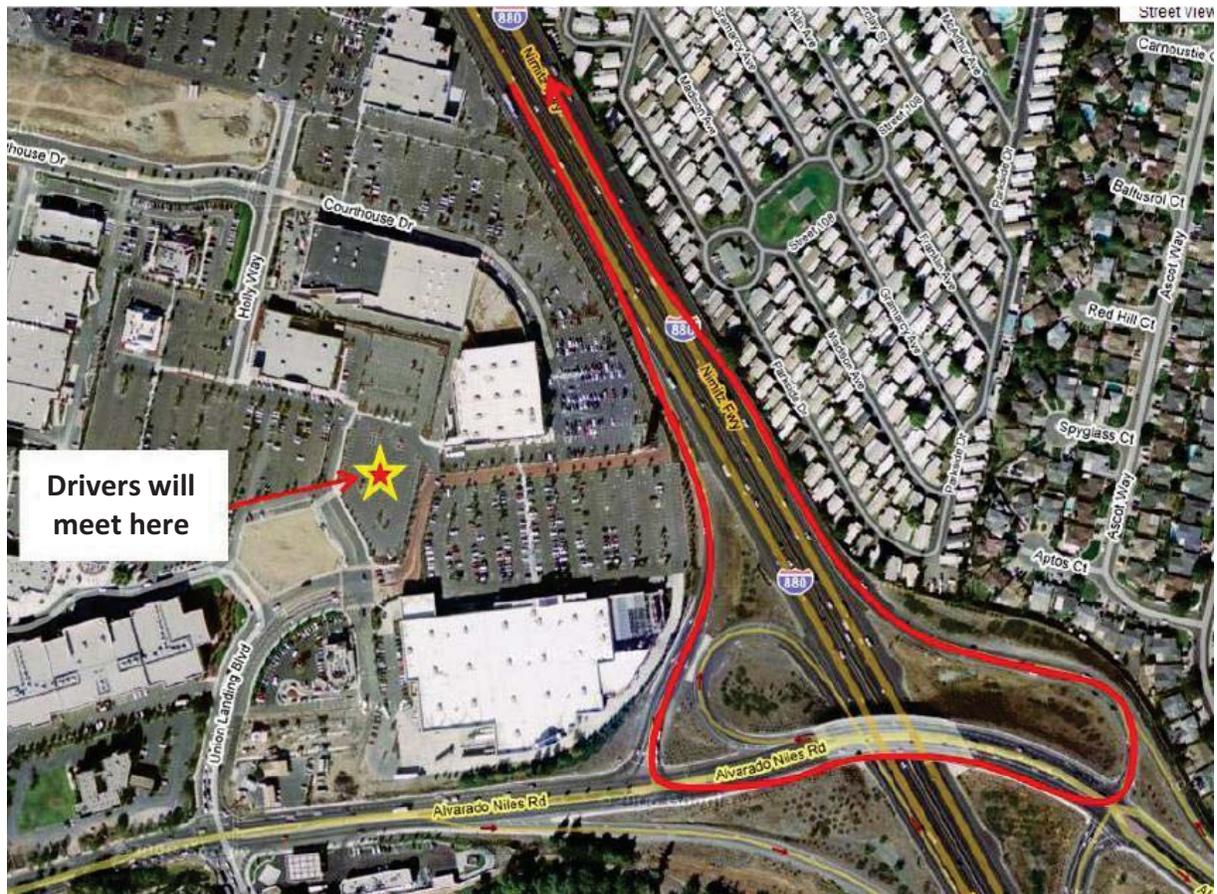
Table 4: Driving Instructions for odd drivers.

Lap	Loop	Lane
1	Long	4
2	Long	4
3	Long	2-3-4
4	Long	2-3-4
5	Long	2-3-4

At the end of the last lap, each vehicle will go back to the parking lot. They will be driving on I880S, so they have to follow the directions given before to go from I880S to the parking lot (Figure 1).

At the parking lot, drivers will return the cell-phone, which concludes their participation in the experiment. After this, drivers are free to go.

If the driver has not completed all the 5 laps after 2 hours of driving, he/she will take the most convenient route in order to go back to the parking lot. The countdown starts with the first driver leaving the parking lot.



A2.3 Figure 5: Turn back at Alvarado-Niles Rd. (sound end of both the long and short loops).

What if the driver misses the off-ramp?

- If the driver misses exit 23 for Alvarado-Niles at the south end of loops, follow the directions outlined in Figure 6:

-

- Take the Alvarado Blvd./Fremont Blvd. exit (around 1.5 miles from the Alvarado-Niles off-ramp).
- Turn left at Alvarado Blvd./Fremont Blvd.
- Turn left to merge onto I880N toward Oakland.



A2.3 Figure 6: Turn back at Alvarado Blvd./Fremont Blvd. in case driver misses exit 23.

- If the driver misses off-ramp to Tennyson Rd. at the north end of short loop, follow the directions outlined in Figure 3:
 - Take the Winton Ave. exit (around 0.7 miles from the CA-92W off-ramp).
 - Keep left at fork. Follow signs for Heald college/Winton Ave. W and merge onto W Winton Rd.
 - Take the I880S ramp to San Jose.



A2.3 Figure 7: Turn back at Winton Ave. in case driver misses off-ramp to CA-92W.

What if the driver gets lost, or has any kind of problem with the vehicle or the cell-phone?

In case of any problem during the experiment, drivers can use the N95 phone to call an emergency number that will be in the phone memory. **The phone number is (650) 521 7940.** This number will connect with a person at Nokia that will provide some help. Instruction on how to use the N95 phone for this purpose will be provided before the experiment begins (at the parking lot).

Appendix 1: UCB *odd* drivers.

UCB01	Andrew Tinka
UCB03	Annalisa Scacchioli
UCB05	Tarek Rabbani
UCB07	Olli-Pekka Tossavainen
UCB09	Alex Bayen

2.4 Instructions for Even Drivers

Instructions and directions: *EVEN* drivers

Even drivers are the following: Nokia02, Nokia 04, Nokia06, Nokia 08, Nokia10, UCB02, UCB04, UCB06, UCB08, and UCB10 (see Appendix 1).

Drivers will meet at the Union Landing retail entertainment center parking lot at 1:35pm on Friday November 2nd (star in Figure 1). The entertainment center is in Union City, next to I805 between Whipple Rd and Alvarado-Niles Rd.

Directions to get to the parking lot (see Figure 1):

- From I805:
 - Take exit 23 for Alvarado-Niles Rd.
 - Turn right at Alvarado-Niles Rd.
 - Turn right at Union Landing Blvd.
- From I880V:
 - Take exit 23 for Alvarado-Niles Rd.
 - Turn left at Alvarado-Niles Rd.
 - Turn right at Union Landing Blvd.

Once at Union Landing Blvd., turn left at the traffic light (see Figure 1).

At the parking lot, drivers will receive the N95 cell phone, which has to be attached to the dashboard of the vehicle (material to attach the cell-phone to the dashboard will be provided).

Instructions on how to call the emergency number using the N95 will be also provided at this point (one click call). The phone number is (650) 521 7940.

Instructions and directions: *EVEN* drivers

Even drivers are the following: Nokia02, Nokia 04, Nokia06, Nokia 08, Nokia10, UCB02, UCB04, UCB06, UCB08, and UCB10 (see Appendix 1).

Drivers will meet at the Union Landing retail entertainment center parking lot at 1:35pm on Friday November 2nd (star in Figure 1). The entertainment center is in Union City, next to I880S between Whipple Rd and Alvarado-Niles Rd.

Directions to get to the parking lot (see Figure 1):

- *From I880S:*
 - Take exit 23 for Alvarado-Niles Rd.
 - Turn right at Alvarado-Niles Rd.
 - Turn right at Union Landing Blvd.
- *From I880N:*
 - Take exit 23 for Alvarado-Niles Rd.
 - Turn left at Alvarado-Niles Rd.
 - Turn right at Union Landing Blvd.

Once at Union Landing Blvd., turn left at the traffic light (see Figure 1).

At the parking lot, drivers will receive the N95 cell phone, which has to be attached to the dashboard of the vehicle (material to attach the cell-phone to the dashboard will be provided).

Instructions on how to call the emergency number using the N95 will be also provided at this point (one click call). **The phone number is (650) 521 7940.**



A2.4 Figure 1: From I-880S and I-880M to the parking lot.

Drivers will be told when to leave the parking lot, but they should be ready to leave the parking lot at 1:45pm. That is, the cell-phone has to be installed inside the vehicle, and the vehicle should have enough gas to drive 40 miles²⁸. The very first vehicle will be released at 1:45pm.

From the parking lot to the freeway (Figure 2):

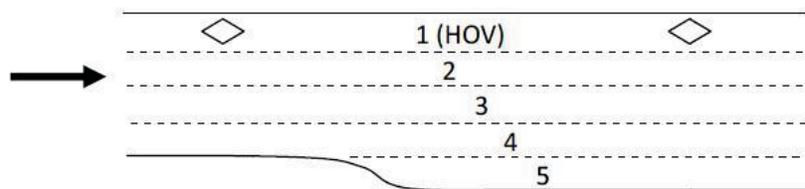
- Leave the mall using Union Landing Blvd.
- Turn left at Alvarado-Niles Rd.
- Turn left to merge onto I880N toward Oakland.

²⁸ There are two gas stations at the corner of Alvarado-Niles Rd. and Union Landing Blvd.



A2.4 Figure 2: From the parking lot onto the freeway.

First, all drivers will go on the long loop using lane 4 (lane 1 is the leftmost lane, HOV lane in our case. See Figure 3).



A2.4 Figure 3: Lane numbering.

Directions for the Long loop:

- Once on I880NN, drive for 3.5 miles approx.
 - Take exit onto CA-92W toward San Mateo Bridge
 - Take the I880SS exit toward San Jose. Drive for 3.5 miles approx.
- } See Figure 4

- Take exit 23 for Alvarado-Niles Rd.
- Turn left at Alvarado-Niles Rd.
- Turn left to merge onto I880N toward Oakland.

} See Figure 6



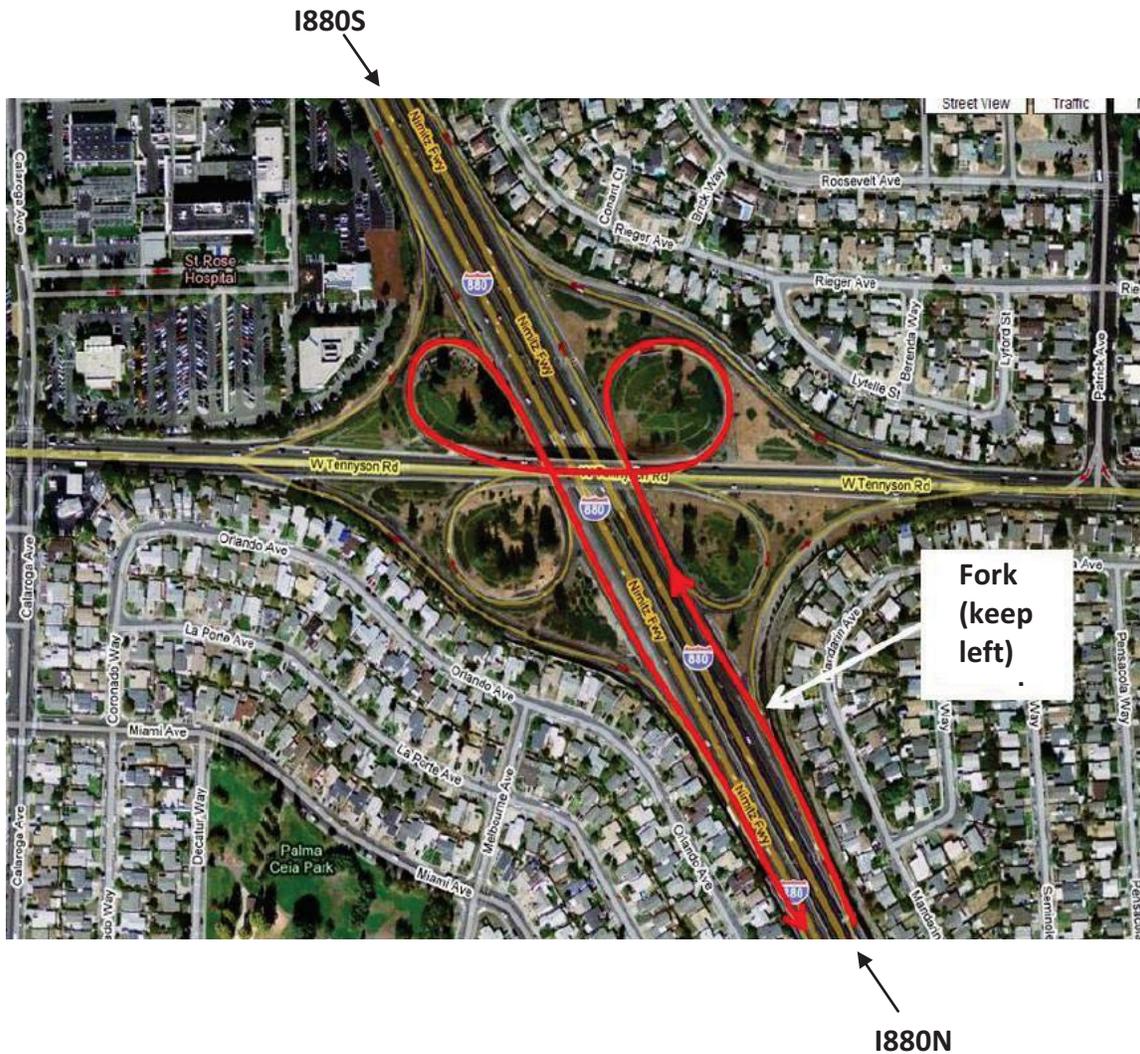
A2.4 Figure 4: Turn back at CA-92 (north end of the long loop).

For the second lap, *even* drivers will keep driving on lane 4, but they will start driving on the short loop.

Directions for the Short loop:

- Once on I880NN, drive for 2.5 miles approx.
- Take Tennyson Rd. exit.
- Keep left at fork. Follow sign to Tennyson Rd. W and merge onto W Tennyson Rd.
- Take the I880SS ramp to San Jose. Drive for 2.5 miles approx.
- Take exit 23 for Alvarado-Niles Rd.
- Turn left at Alvarado-Niles Rd.
- Turn left to merge onto I880N toward Oakland.

} See Figure 5
 } See Figure 6



A2.4 Figure 5: Turn back at W Tennyson Rd. (north end of the short loop).

After the completion of the second lap, even drivers will do 4 more short loops, but they can use either lane 2, 3 and/or 4 (see Table 1).

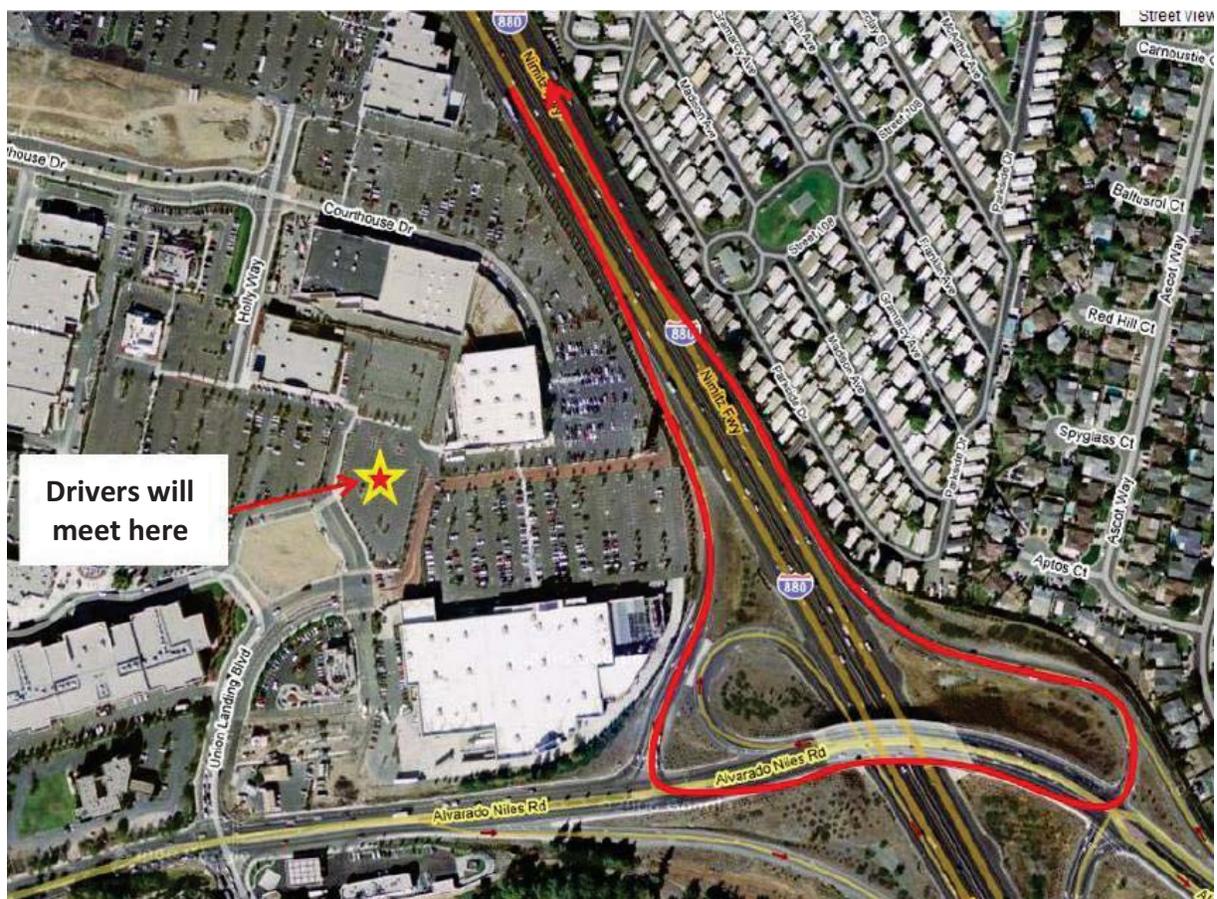
Table 5: Driving instructions for *even* drivers.

Lap	Loop	Lane
1	Long	4
2	Short	4
3	Short	2-3-4
4	Short	2-3-4
5	Short	2-3-4
6	Short	2-3-4

At the end of the last lap, each vehicle will go back to the parking lot. They will be driving on I880S, so they have to follow the directions given before to go from I880S to the parking lot (Figure 1).

At the parking lot, drivers will return the cell-phone, which concludes their participation in the experiment. After this, drivers are free to go.

If the driver has not completed all the 6 laps after 2 hours of driving, he/she will take the most convenient route in order to go back to the parking lot. The countdown starts with the first driver leaving the parking lot.



A2.4 Figure 6: Turn back at Alvarado-Niles Rd. (sound end of both the long and short loops).

What if the driver misses the off-ramp?

- If the driver misses exit 23 for Alvarado-Niles at the south end of loops, follow the directions outlined in Figure 7:
 - Take the Alvarado Blvd./Fremont Blvd. exit (around 1.5 miles from the Alvarado-Niles off-ramp).
 - Turn left at Alvarado Blvd./Fremont Blvd.
 - Turn left to merge onto I880N toward Oakland.



A2.4 Figure 7: Turn back at Alvarado Blvd./Fremont Blvd. in case driver misses exit 23.

- If the driver misses off-ramp to Tennyson Rd. at the north end of short loop, follow the directions outlined in Figure 3:
 - Take the next exit as described in the long loop (because it corresponds to the off-ramp to CA-92). See Figure 3.
- If the driver misses off-ramp to CA-92W at the north end of long loop, then follow the directions outlined in Figure 8:
 - Take the Winton Ave. exit (around 0.7 miles from the CA-92W off-ramp).
 - Keep left at fork. Follow signs for Heald college/Winton Ave. W and merge onto W Winton Rd.
 - Take the I880S ramp to San Jose.



A2.4 Figure 8: Turn back at Winton Ave. in case driver misses off-ramp to CA-92W.

What if the driver gets lost, or has any kind of problem with the vehicle or the cell-phone?

In case of any problem during the experiment, drivers can use the N95 phone to call an emergency number that will be in the phone memory. **The phone number is (650) 521 7940.** This number will connect with a person at Nokia that will provide some help. Instruction on how to use the N95 phone for this purpose will be provided before the experiment begins (at the parking lot).

Appendix 2: UCB even drivers.

UCB02	Christian Claudel
UCB04	Ryan Herring
UCB06	Marika Benko
UCB08	Dan Work
UCB10	Juan Carlos Herrera

3 Contingency Plans

Appendix 3 documents the contingency plans that made explicit the pre-meditated responses to foreseeable risks such as accidents, injuries, fatalities, or other situations such as those arising from extreme weather. Responses were categorized as “go,” “abort,” and “recall.” The management team and decision structure for each risk are specified in Appendix 3.1. The power-point briefing for the safety and emergency response team is furnished in Appendix 3.2. Finally, Appendix 3.3 provides step-by-step instructions for the phone operators to handle contingencies and to support the participating drivers.

3.1 Risk Management

1. Scope

The purpose of this document is to identify possible risks/threats to conducting the "mobile century" experiment and provide corresponding responses and strategies against the risks/threats. The experiment is slated to take place on February 8, 2008 on a 10-mile stretch of I-880. The outlined strategies only accept the consequences of a particular risk. In other words, the risk management strategies focus only on the following experiment alternatives:

- Go
- Recall
- Abort

The alternatives will be explained in section 3.

Prior to, or on the day of experiment, the operation management will use this document as a guideline to make appropriate decisions in case of hazardous events that may threaten the safety and results of the experiment, i.e. accidents, adverse weather, etc.

Efficient responses to the risks must be approved by the appropriate level of management. In this regard, the management team (listed in section 2) reviewed the threats and approved appropriate response for each identified risk.

The document follows two major steps in risk management:

1. *Identification and Assessment.* At this stage, major threats are identified and listed. A preliminary assessment of probabilities/likelihood of some threats is also analyzed.
2. *Risk Management Plan Response.* Based on management team decision, an appropriate response to each risk is proposed.

Scope

The purpose of this document is to identify possible risks/threats to conducting the “mobile century” experiment and provide corresponding responses and strategies against the risks/threats. The experiment is slated to take place on February 8, 2008 on a 10-mile stretch of I-880. The outlined strategies only accept the consequences of a particular risk. In other words, the risk management strategies focus only on the following experiment alternatives:

- Go
- Recall
- Abort

The alternatives will be explained in section 3.

Prior to, or on the day of experiment, the operation management will use this document as a guideline to make appropriate decisions in case of hazardous events that may threaten the safety and results of the experiment, i.e. accidents, adverse weather, etc.

Efficient responses to the risks must be approved by the appropriate level of management. In this regard, the management team (listed in section 0) reviewed the threats and approved appropriate response for each identified risk.

The document follows two major steps in risk management:

1. *Identification and Assessment.* At this stage, major threats are identified and listed. A preliminary assessment of probabilities/likelihood of some threats is also analyzed.
2. *Risk Management Plan/Response.* Based on management team decision, an appropriate response to each risk is proposed.

Management Team

The list of management team is as follows:

- Alex Bayen (Experiment Director)
- Ali Mortazavi (Experiment Supervisor)
- Dan Work (Experiment Supervisor Assistant)
- Tom West

- JD Margulici

Threats/Risks

In this section, we list possible risks that could threaten the execution of the experiment. Each risk can be classified based on the time of occurrence and its level of threat. For this experiment, it is important to consider the element of time in our calculations. For example, the response to a forecasted heavy rain three days before the experiment will not be similar to the response to an unexpected heavy rain on the day of experiment. To avoid possible losses, it might be rational to abort the experiment's execution if the risk is known in advance. However, it may be wise to continue the experiment's execution if the threat occurs during the experiment, after all logistical costs have been incurred.

In this section, we try to assess the probabilities of occurrence for specific risks, present different scenarios, and address appropriate responses to the scenarios. Considering the time element, the responses can be classified for two major scenarios:

1. Occurring before the day of experiment
2. Occurring on the day of experiment

Three possible responses are presented:

1. *GO*. There is a low possibility of risk, and it is ok to conduct the experiment.
2. *ABORT*. The risk is a significant threat to safety or it may render the experiment result useless. The management team will decide on aborting the experiment.
3. *RECALL*. This action is the response to a threat during the experiment. Experiment Supervisor (Ali) authorizes the recall after consulting with Experiment Director (Alex). If the threat requires a quick response, Experiment Supervisor (Ali) can make the decision without consulting Experiment Director (Alex). The drivers are asked to come back to the base station temporary and be prepared / alert to continue the experiment when the threat diminishes. The management team will decide on continuing the experiment.

Please see section 0 for RECALL instructions. ABORT instruction will be similar to RECALL instructions.

Adverse Weather

We consider three types of risks associated with weather:

1. Storm
2. Rain
3. Fog

For each category, the historic weather data for 7 days — Feb 5th to Feb 11th — over past 10 years were analyzed and the probability of occurrence (potential severity) for each type is assessed.

Table 6 shows average minimum and maximum temperature of the historic data.

Table 6: Average min and max temperature

Average Min Temperature	43°F
Average Max Temperature	61°F

Storm

A strict meteorological definition of a terrestrial storm is a wind speed of 24.5 m/s (89 km/h, 55 mph) or more; however, popular usage is not so restrictive. Storms can last anywhere from 12 to 200 hours, depending on season and geography.

Assessment

Based on historic data, no storm was recorded for 10 years. Therefore, storm threat falls into low risk category.

Response

Scenario A. Before the day of experiment

- Tuesday news forecasts a heavy storm (sustained wind speed higher than 45 mph) on Friday (2/08/08) — Consult with management team / possibly **ABORT**
- Heavy storm on Wednesday (2/06/08) or Thursday (2/07/08), trees are down and block the road — Consult with management team / possibly **ABORT**
- Otherwise — **GO**

Scenario B. On the day of experiment

Sustained wind speed higher than 45 mph — **RECALL**

Rain**Classification**

Based on the amount of precipitation, rain can be classified into 6 categories:

- *Very light rain* — when the precipitation rate is < 0.25 mm/hour (0.01 inch/hour)
- *Light rain* — when the precipitation rate is between 0.25 mm/hour (0.01 inch/hour) - 1.0 mm/hour (0.04 inch/hour)
- *Moderate rain* — when the precipitation rate is between 1.0 mm/hour (0.04 inch/hour) - 4.0 mm/hour (0.16 inch/hour)
- *Heavy rain* — when the precipitation rate is between 4.0 mm/hour (0.16 inch/hour)- 16.0 mm/hour (0.63 inch/hour)
- *Very heavy rain* — when the precipitation rate is between 16.0 mm/hour (0.63 inch/hour)- 50 mm/hour (1.97 inch/hour)
- *Extreme rain* — when the precipitation rate is > 50.0 mm/hour (1.97 inch/hour)

Assessment

Table 7 shows the likelihood of each category based on historic data. The data shows the possibility of having heavy rain is extremely low.

Table 7: Likelihood of having rain based on historic data

	Very light	Light	Moderate	Heavy	Very heavy	Extreme
Likelihood	0.17	0.17	0.01	0	0	0

Response**Scenario A. Before the day of experiment**

Responses to the following scenarios forecasted on Tuesday news will be as follow:

- *Very light rain* — **GO**
- *Light rain* — **GO**

- *Moderate rain* — **GO**
- *Heavy rain less than 3 hours* — **GO**

The management team needs to rearrange the experiment timeline.

- *Sustained heavy rain over 3 hours* — **GO**
- *Very heavy rain* — Consult with management team / possibly **ABORT**
- *Extreme rain* — Consult with management team / possibly **ABORT**

Scenario B. On the day of experiment

For sustained very heavy or extremely heavy rain — **RECALL**

Fog

Classification

The international definition of fog is a visibility of less than 1 kilometer (3,300 ft); mist is a visibility of from 1 kilometer (0.62 mi) to 2 kilometers (1.2 mi) and haze from 2 kilometers (1.2 mi) to 5 kilometers (3.1 mi). Visibility of less than 100 meters (330 ft) is usually reported as zero. In these conditions, roads may be closed. This risk can be classified into:

1. Visibility greater than 0.6 mile
2. Fog (with visibility less than 0.6 mile)

Assessment

The historic data shows the likelihood/possibility of having fog at the day of experiment is 0.13, which put this threat as low risk category.

Response

Scenario B. On the day of experiment

- Visibility greater than 0.6 mile — **GO**
- Visibility less than 0.6 mile — **RECALL**

Incidents

This type of risk includes any incident/accident that may block the flow of the traffic and affect the execution of the experiment. It may also cause adverse psychological effect on the experiment participants which can jeopardize the safety of the drivers.

Road Accident

This type of risk consists of any traffic accident that can affect the traffic flow.

Classification

We classify the Road Accident risk into five major categories:

1. *Minor accident.* It is a typical fender bender. It may slow down the traffic, but can be cleared quickly and the effect is not long.
2. *Short moderate accident.* It blocks at least 1-3 lanes of the highway and significantly slows down the traffic flow which may last no longer than 3 hours. The vehicles are still able to move on the highway.
3. *Long moderate accident.* It blocks at least 1-3 lanes of the highway and significantly slows down the traffic flow, which may last longer than 3 hours. The vehicles are still able to move on the highway.
4. *Short major accident.* It causes complete road blockage for less than 3 hours. The traffic is completely stopped.
5. *Long major accident.* It causes complete road blockage for more than 3 hours. The traffic is completely stopped

Assessment

- *Minor accident* — **GO**
- *Short moderate accident* — **GO**
- *Long moderate accident* — **RECALL**
- *Short major accident* — **RECALL**
- *Long major accident* — **RECALL**

Response*Injuries*

This risk only includes the injuries related to the experiment participants.

Classification

Injuries can be classified, based on the level of intensity, into three levels:

1. *Level 1.* The injured driver does not need to be transferred to a hospital.
2. *Level 2.* The injured driver has to be transferred to a hospital or clinic, but the injury is not life threatening.
3. *Level 3.* The injured driver has to be transferred to a hospital and the injury is life threatening.

Assessment

The incident data on I-880 over 5 years (2003-2007) was collected from the PeMS database. There is a 71% chance of an accident on I-880 (NB and SB). It should be noted that the result could not be localized for the stretch of I-880 in which the experiment will be conducted. (It is for whole I-880 highway.) Therefore, an accurate conclusion cannot be made.

Response

- *Level 1* — **GO**
- *Level 2* — **GO**
- *Level 3* — **RECALL**

Fatality

This risk only considers the experiment drivers.

Response

Any fatality — **RECALL**. Experiment Supervisor (Ali) can authorize the RECALL response immediately without consulting the Experiment Director. No management decision needed.

Other Incidents

These risks include incidents other than road accidents, injuries, and fatalities, i.e. natural disaster. Due to a low probability and a broad spectrum of possible scenarios, the response will be made by management team on the fly.

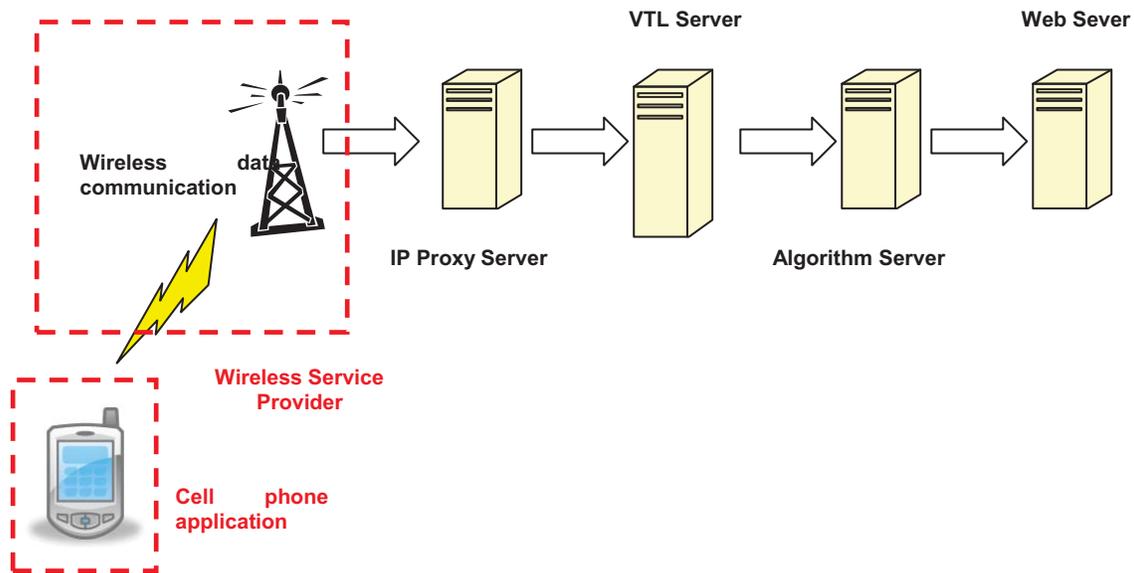
Technical problems

There are different technical risks which can jeopardize the project. Major concerns are focused on the subsystems that:

- Collect the data;
- Calculate the travel times.

Figure 1 displays the architecture of mobile probe data gathering system. Similar architecture will be used to collect traffic data during the experiment. According the architecture, possible threats can be identified in different layers of the system:

1. Cell phone application failure
2. Wireless data connection failure
3. Encryption server failure
4. VTL server failure
5. Algorithm server failure
6. Web server failure



A3 Figure 1: Architecture of mobile probe data collection / travel time calculation system

Cell phone application failure

Response

A list of possible risks / threats is as follows:

- *Large scale loss of GPS tracks on monitoring screens for extended period of time* — **RECALL** after being advised by Jeff (CCIT technical staff) at Nokia, or any of the other Nokia staff (in case Ali cannot reach Alex or Dan).
- *Loss of server capabilities, VTL transmission capabilities or broadcast capabilities* — **GO** (capabilities will be restored).
- *Server failure leading to loss of GPS tracks. In this case* — Experiment Supervisor will be notified by Jeff to recall **progressively** the cars. Client restart procedure will be performed on each of the cars individually as they come in and out of the lot. No recall is necessary. However, the staff onsite will check all the phones one by one before they are sent to the highway again.
- *Catastrophic event (for example loss of tower)* — In this case, the phone operators probably do not have contact with drivers anymore. The only way the phone operators

can communicate with the drivers is their own phones (if they are on another network).
The Experiment Supervisor (Ali) has to inform Experiment Director (Alex).

RECALL Instructions

For Emergency staff:

1. Based on the conditions, Experiment Supervisor (Ali) identifies the threat.
2. Experiment Supervisor (Ali) calls/meets Experiment Director (Alex) to finalize the decision on sending 'recall' message to the drivers. Experiment Director is available through his cell phone.
3. After Experiment Director's (Alex's) confirmation, the experiment supervisor (Ali) authorizes the recall procedure by informing the Emergency/Safety team supervisor (Osama) and/or designated phone operator for recall message.
4. Using the teleconference feature, the phone operator informs all the drivers simultaneously. On each call he/she repeats the recall message five times. The message is: "Attention all the drivers, please proceed with the recall procedure." The operators repeat the call every one minute for 5 minutes to make sure all the drivers get the message.
5. The monitoring team identifies the remaining cars and informs the phone operators to contact the remaining drivers on the road.

For the drivers:

1. React immediately to the message.
2. Remain calm.
3. Using the off-ramps (exits) designated to your team, drive back to the Union Landing Mall.
4. Stay alert for other upcoming calls from the operators. Be aware the operators may contact you and your team repeatedly and announce the recall message. Please answer all the receiving calls.

5. Upon arrival to the base station, follow the Team Leader Assistants' instructions and park the vehicles in designated spots. Be extra careful to avoid an accident.
6. Once parked, report to the Team Personnel Assistants.
7. Go to the drivers lounge area and wait for further instructions.

3.2 *Emergency Response*

***Powerpoint Presentation: Safety and Emergency Response
Team: Process and Procedures***

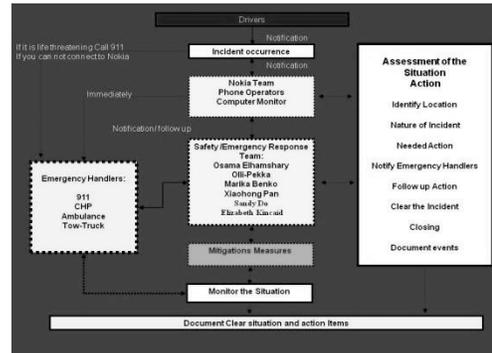


Safety and Emergency Response Team: Process and Procedures

1

Safety and Emergency Response Team Process and Procedures

2



3

Emergency response plan Scenario 1. Whenever second party involved

In any event of critical emergency (accident, injury or fatal/life threatening),
The driver Should:

- Step 1: Call (911) immediately by using all necessary means (personal cell if available or the Nokia phone);
- Step 2: Notify the Nokia phone operators by pushing the button on the Bluetooth headset; If it fails to connect, call the safety team contact number listed on you sign checklist sheet.

Nokia team Will:

- Identify and record the vehicle number and location;
- Call the emergency handler (911) or ambulance immediately, if needed;
- Notify the emergency response team for follow up.

Emergency Response Team will:

- Assess the situation according to the emergency response plan protocol (see attached flowchart);
- Follow up, monitor, and document the situation;
- Make sure the incident is cleared;
- If needed, the emergency response team will contact the Enterprise rental company and coordinate with the operational team for necessary actions.

4

Emergency response plan Scenario 2.0. General safety instructions

Scenario 2.0. General safety instructions

To Logistic site team, Nokia Team and Emergency Response Team:

- All drivers will be instructed about possible emergency scenarios and how to handle the communication.
- Phone ID, vehicle ID and vehicle license plate should be recorded before driving onto the highway.
- Copy of all those records should be available to the emergency response team.
- Emergency response team should have a comprehensive contact list of all outside emergency handlers.

5

Emergency response plan Scenario 2. Routine emergency response plan (continue)

To All Drivers:

- In any event of critical emergency (accident, injury or fatal/life threatening), call 911 immediately by using all necessary means (including your personal cell if available or Nokia call). And then contact Nokia phone operators to report your status.
- If the incident involves second party, driver should stay in the scene until it's cleared by the proper authority (Local Police, CHP, etc).
- All non-life threatening emergency will be reported to and will be handle by Nokia phone operators and the emergency response team.
- If your car breakdown, all drivers should stay away from the paved lane; if possible stay behind the road guardrail and away from the traffic.
- If you have to walk, always walk facing the traffic, be alert and minimize the chance to interrupt the traffic.
- Always follow the emergency handlers' instruction and follow the proper safety driving rules and regulations.

6

Emergency response plan Scenario 2.1. On the transit route between Berkeley and Union Landing

Emergency Response Team need:
Designate a staff on the bus to handle any event of emergency;
Call 911 immediately;
Report back to emergency response team member.

7

Emergency response plan Scenario 2.1. During the process of the experiment

Type of Incident	Actions		
	Drivers	Nokia Team	Emergency Response Team
1. Car crash (accident, injury or fatal/life threatening):	<ul style="list-style-type: none"> Call 911 immediately by using necessary means (your personal phone if available or Nokia phone); Notify Nokia phone operators by pushing the button on the Bluetooth headset or using necessary means; If possible or needed, pull aside your car from the paved road to the road shoulder and stay away from the traffic. 	<ul style="list-style-type: none"> Identify and record the vehicle number and location; Call the emergency handler (911) or ambulance immediately, if needed; Notify the safety team for follow up. 	<ul style="list-style-type: none"> Assess the situation according to the emergency response plan; Follow up, monitor, and document the situation; Make sure the incident is cleared; If needed, the emergency response team will contact the Enterprise agency and coordinate with the operational team for necessary actions.

8

Emergency response plan Scenario 2.1. During the process of the experiment

Type of Incident	Actions		
	Drivers	Nokia Team	Emergency Response Team
2. Car breakdown or Flat tire or Out of gas	<ul style="list-style-type: none"> Pull aside your car from the paved road to the road shoulder and stay away from the traffic; Use your Bluetooth to report the situation to Nokia phone operators; Wait for instructions or help; If the car is completely unusable, try to move the car away from the traffic, lane an abandoned you car in a safe way, stay away from the traffic; call 911 first and then notify Nokia phone operators. 	<ul style="list-style-type: none"> Identify and record the vehicle number and location; Notify the safety team for follow up. 	<ul style="list-style-type: none"> Assess the situation according to the emergency response plan; May need to send the tow-truck and possible another car to pick the driver; If needed, the emergency response team will contact the Enterprise rental car company and coordinate with the operational team for necessary actions; Follow up, monitor, and document the situation, and make sure the incident is cleared.

9

Emergency response plan

Scenario 2.1. During the process of the experiment

Type of Incident	Actions		
	Drivers	Nokia Team	Emergency Response Team
3. Violation of traffic rules	<ul style="list-style-type: none"> - Pull aside your car from the paved road to the road shoulder. - Listen to the authorities and inform them about the nature of the experiment, and provide all needed documents. - If the authority allowed you to continue your driver, make sure to follow the traffic regulations. - Notify the Nokia phone operator. 	<ul style="list-style-type: none"> - Identify and record the vehicle number and location. - Notify the safety team for follow up. 	<ul style="list-style-type: none"> - Assess the situation according to the emergency response plan. - We may need to replace the driver by sending another car to pick the driver for a short rest for the whole duration. - If needed, the emergency response team will contact the Enterprise agency and coordinate with the operational team for necessary actions. - Follow up, monitor, and document the situation, and make sure the incident is cleared.

10

Emergency response plan

Scenario 2.1. During the process of the experiment

Type of Incident	Actions		
	Drivers	Nokia Team	Emergency Response Team
4. Driver got lost	<ul style="list-style-type: none"> - Use the nearest exit and continue driving to the nearest tram, and find a safe spot to park your car away from the traffic. - Notify Nokia phone operators and wait for instruction. 	<ul style="list-style-type: none"> - Identify and record the vehicle number and location. - Notify the safety team for follow up. - Nokia team will always track the drivers and immediately notify the safety team regarding any possible route deviations. 	<ul style="list-style-type: none"> - Assess the situation according to the emergency response plan. - Use the Nokia monitor to locate the driver. - Figure out the best way to retrieve the lost driver. Use the safe map. - May need to coordinate with the operational team for necessary actions. - Follow up, monitor, and document the situation, and make sure the incident is cleared.

11

Emergency response plan

Scenario 2.1. During the process of the experiment

Type of Incident	Actions		
	Drivers	Nokia Team	Emergency Response Team
5. Severe road or weather conditions	<ul style="list-style-type: none"> - Notify Nokia phone operators and wait for instruction. - If notified by Nokia phone operators or public authorities, follow the instructions and use all of your safety tips. 	<ul style="list-style-type: none"> - Identify the situation and notify all the drivers. - Notify the safety team for follow up. 	<ul style="list-style-type: none"> - Assess the situation according to the emergency response plan. - collect all related information by all available means, weather forecast or local news. - Coordinate with other teams (Ryan team) to make sure the safety of all experimenters. - Follow up, monitor, and document the situation, and make sure the incident is cleared.

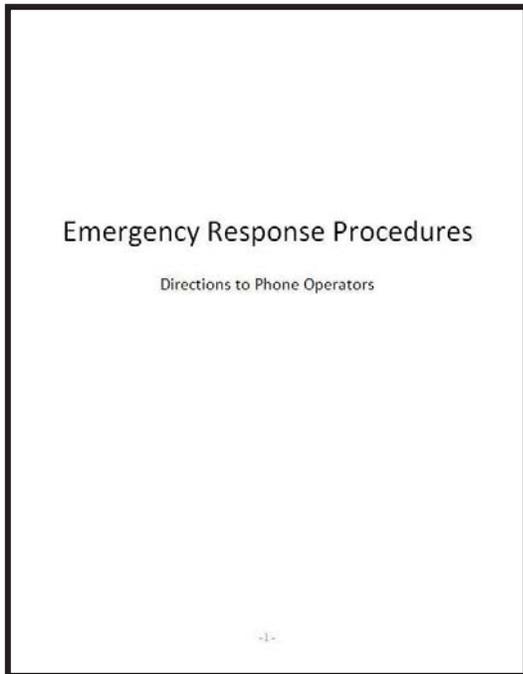
12

Emergency response plan

Scenario 2.1. During the process of the experiment

Type of Incident	Actions		
	Drivers	Nokia Team	Emergency Response Team
6. Driver illness	<ul style="list-style-type: none"> - If possible, use the nearest exit and continue driving to the nearest landmark and find a safe spot to park your car away from the traffic. - If not, pull aside your car from the paved road to the road shoulder. - Notify Nokia phone operators and wait for instruction. 	<ul style="list-style-type: none"> - Identify and record the vehicle number and location. - Call the emergency handler (911) or ambulance immediately, if needed. - Notify the safety team for follow up and instructions for the driver. 	<ul style="list-style-type: none"> - Assess the situation according to the emergency response plan. - Send out the ambulance if necessary. - If necessary, we may need to replace the driver by sending another car to pick the driver for a rest at Union landing. - Follow up, monitor, and document the situation, and make sure the incident is cleared.

3.3 *Directions for Phone Operators*



Emergency Response Procedures

Directions to Phone Operators

Emergency response team (ERT)

Name	Title	Nokia phone number
Osama Elhamshary	Safety/emergency response supervisor	650-644-9645
Olli-Pekka Tossavainen	Safety officer	650-644-9840
Xiaohong Pan	Safety officer	650-644-9328
Andrew Tinka	Team member	650-644-9761
Sandy Do	Team member	
Elizabeth Kincaid	Team member	
Nick Semon	Team member	
Trucy Phan	Team member	
Christopher Flens-Batina	Team member	

Note: Some team members will act as phone operators if needed.

Notes:

1. If there is not a guardrail nearby or a safe path to it, drivers stuck on the road should NOT exit the car. It is much safer inside a car with a seatbelt on than outside on the road. Remember: people smash. Cars just crunch.
2. If Nokia phone is used instead of double-clicking the Bluetooth button, inform the driver that s/he need to come back to union landing ASAP to have the GPS system reconnected.

Car crash -- whenever second party is involved

- Begin to fill out incident reporting sheet
- Get driver information
 - Vehicle number
 - Location (if known)
 - Note type of incident
- Call 911 or ambulance if incident is life-threatening and driver cannot personally dial 911
- Notify emergency handlers if necessary
 - 911
 - California Highway Patrol (CHP)
 - Ambulance
 - Tow-truck
- Fill out incident recoding sheet
- Give driver instructions (as needed)
 - Pull off road
 - Take nearest exit, park at safe location
 - Pull onto shoulder
 - Turn on four-way flashers
 - Wait for help
 - Stay in car with seatbelt on (preferable)
 - Get out of car and move as far away from traffic as possible
- Notify Emergency Response Team
- Follow up
 - Call driver if needed
 - Make sure incident is cleared
 - Call Enterprise rental company if necessary

Car breakdown, flat tire, or out of gas

- Begin to fill out incident reporting sheet
- Get driver information
 - Vehicle number
 - Location (if known)
 - Note type of incident
- Give driver instructions
 - Breakdown
 - Pull off road if possible. Turn on four-way flashers.
 - If fully inoperable and in roadway,
 - Stay in car with seatbelt on if busy traffic.
 - If not busy traffic and guardrail exists with, consider exiting car and standing on opposite side of guardrail. Exercise caution.
 - Flat tire
 - Pull off road carefully and turn on four-way flashers.
 - If very near exit, consider taking exit and parking at a safe location nearby, driving very slowly.
 - Wait in car with seatbelt on for help to arrive if on roadside.
 - Out of gas
 - If nearly out of gas:
 - Take nearby exit, if possible, and park in a safe location.
 - Pull to side of road and turn on four-way flashers if they have no confidence about reaching an exit.
 - If stuck in traffic jam with no chance of getting to side of road, run until actually out of gas while trying to get to the side.
 - If fully out of gas
 - If on side of road or in road, put on four-way flashers and wait for help.
 - If in safe location, relax and wait for help to arrive.
- Notify Emergency Response Team
- Follow up
 - Call driver if needed
 - Make sure incident is cleared
 - Call Enterprise rental company if necessary

Violation of traffic rules

- Begin to fill out incident reporting sheet
- Get driver information
 - Vehicle number
 - Location (if known)
 - Note type of incident -- presumably pulled over by police
- Give driver instructions (where are the police?)
 - Police are in their car, running plates
 - Reassure driver
 - Prep them to fully disclose experiment details
 - Ask them to call back when police business finished
 - Police are at the car, want to talk to someone in charge of the experiment
 - Talk to police, provide details
 - Hand over to Osama or Olli-Pekka if needed
 - Police have left, driver is calling back as instructed
 - Instruct driver to come back to base, take 1 hr break now
 - ERT: Talk to driver when they return; evaluate fitness to continue
 - ERT: Notify Driver Team Leader to adjust driver scheduling
 - Driver is at car, cannot continue driving (police instructions or personal choice)
 - Instruct driver to stay at car
 - ERT: Assemble pickup team, retrieve driver
 - ERT: Notify Driver Team Leader to adjust driver scheduling
 - Driver is getting arrested by the police
 - Get as many details as possible: which police station, where will car go, bail/pickup details
 - ERT: If possible, send team to pick up/bail out driver
 - ERT: Notify Driver Team Leader to adjust driver scheduling
 - ERT: If possible, send team to pick up car
- Follow up
 - Call driver if needed
 - Make sure incident is cleared

Driver got lost

- Begin to fill out incident reporting sheet
- Get driver information
 - Car info – vehicle number
 - Location (if known, but if calling about being lost, probably won't know)
 - Note type of incident
- Give driver instructions:
 - Might be lost
 - Pull off at the next safe exit and look at your map to try and find your location. If you are still lost, call back.
 - What is last exit or the next one?
 - Missed the exit for Union Landing
 - Take the next exit. Get on the freeway going in the opposite direction. Keep driving until you reach the Union Landing exit and resume following the printed directions.
 - Missed turn
 - Take the next exit and continue your loop
 - No idea where I am
 - Ask if still driving. If yes, hang up and get off at the nearest safe exit, park in a visible and secure area and call back so we can find you on the map
 - What exit did you get off at?
 - *If no idea:* Okay, describe your surroundings to us (like a unique building or company) and the street names around you
 - *If person is panicking:* then say it is okay, we can also find you using the GPS monitor. Just relax and wait for further instructions.
 - Ramp is closed
 - Go to the next exit and continue your loop
 - If person does not want to continue driving, contact relevant help to send to driver
- Notify Emergency Response Team
- Follow up
 - Call driver if needed
 - Make sure incident is cleared

Driver illness

- Begin to fill out incident reporting sheet
- Get driver information.
 - Vehicle number
 - Location (if known)
 - Note type of incident
- Give driver instructions:
 - Severe illness
 - Ensure safety in the traffic
 - Dial 911 if possible (if not, phone operator dial 911 and notify ERT immediately)
 - Medium illness
 - If possible, take the nearest exit after hanging up
 - Continue driving to the nearest landmark and find a safe spot to park your car away from the traffic; and then call phone operator again for instructions and help.
 - If not, pull off road. Turn on four-way flashers. Wait for help.
 - Wait in car with seatbelt on for help to arrive or consider exiting car and standing on opposite side of guardrail. Exercise caution.
 - Light illness
 - If possible, take the nearest exit after hanging up
 - Continue driving to the nearest landmark and find a safe spot to park your car away from the traffic; and then call phone operator again for instructions and help.
 - If not, pull off road. Turn on four-way flashers. Wait for help.
 - Wait in car with seatbelt on for help to arrive or consider exiting car and standing on opposite side of guardrail Exercise caution.
- Notify Emergency Response Team
- Follow up
 - Call driver if needed
 - Make sure incident is cleared

Severe road or weather conditions (recall procedure)

- If receive such type of reporting from driver
 - Reassure driver
 - Notify ERT immediately!
 - Wait for project manager's decision on if trigger the recall procedure
- If project manager decide to trigger the record procedure
 - Help to contact all drivers about the procedure by any necessary means
- Follow up
 - Make sure incident is cleared

Emergency/Logistical Contacts

1. 911 (CHP) for all emergencies/Serious accidents

2. Nearest Hospital with an Emergency Room

*911 will handle all ambulance/emergency calls.

Name: _____ Eden _____ Medical _____ Center
 Address: 20103 Lake Chabot Road
 Castro Valley, CA 94546-5367
 Phone: 510-537-1234
 E-mail: edeninfo@sutterhealth.org

3. Towing Service to be Onsite All Day: 8:00 AM-8:00 PM.

* Rick will be onsite for minor issues (someone ran out of gas, locked themselves out of their car, etc.).

Name: Rick
 Phone: 510-776-5661.

4. Marika Benko is the point of contact for Enterprise

Enterprise

contact

*Call for replacement a vehicle, if the vehicle is no longer fit to drive

Name: Sarah Peterson
 Phone: 415-720-7437 cell

Second Enterprise Contact

Name: Doug Watt
 Phone: 415-720-8987 cell

5. Marika Benko is the point of contact for any tent/equipment issues

Tent/Furniture/Set Up Issues:

Andrew and his team is setting up and providing the furniture, tent, electricity, portable toilets, generator, etc.

Name: Andrew Sutton
 Phone: cell: 650-533-3922

6. Ambulance service is scheduled to be onsite All Day: 8 AM-8 PM

Ambulance and EMT will be onsite if someone panics or gets injured in the parking lot area. 911/CHP will be notified for an ambulance, should there be an accident on the freeway or a serious injury in the parking lot.

Royal Ambulance
 Name: John Eric Henry
 Phone: 510-568-6161 office

7. Union City Police

Non-emergency phone number: (510) 471-1365
 Address: 34009 Alvarado-Niles Road
 Union City, California 94587

Union City Officer: Marika is the point-person for this contact.

Lt. Kelly Musgrove
 Field Operations Command
 Phone: (510) 675.5262

8. Non-Emergency Highway Patrol

Phone: (510) 489-1500

Location: 2434 Whipple Rd., Hayward

*Also see letter from Officer Oscar Johnson approving the experiment. The letter is in the emergency file cabinet, to be stationed at the emergency team table.

4 Driver Briefings

Appendix 4 contains the PowerPoint presentations that were used during the driver briefing sessions. Topics covered include the schedule for the day, emergency procedures, how to use the Nokia phone, what to do in case of an emergency, and instructions in the event of a recall. The bulk of the presentation covered the route assigned to the team in the briefing. There were a total of three teams.

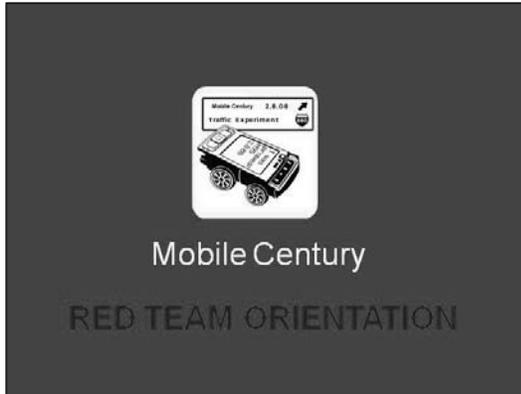
The Red Team was assigned to routes that extended as far north as Winton Ave., and as far south as Thornton Ave. and Alvarado Niles Rd. for the morning and afternoon routes, respectively. This briefing is contained in Appendix 4.1.

The Yellow Team was assigned to routes that extended as far north as SR-92 (San Mateo Bridge), and as far south as Mowry Ave. and Alvarado Blvd. for the morning and afternoon routes, respectively. This briefing is contained in Appendix 4.2.

The Orange Team was assigned to routes that extended as far north as Tennyson Rd., and as far south as Stevenson Blvd. and Decoto Rd. for the morning and afternoon routes, respectively. This briefing is contained in Appendix 4.3.

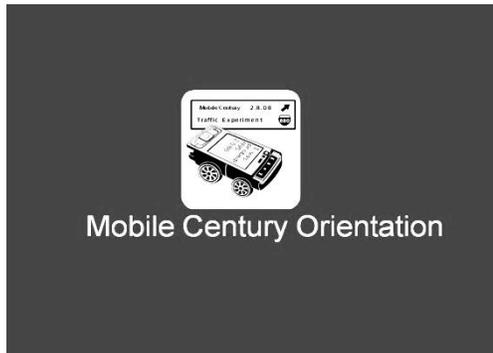
4.1 Red Team

Power Point Presentation: Red Team Orientation

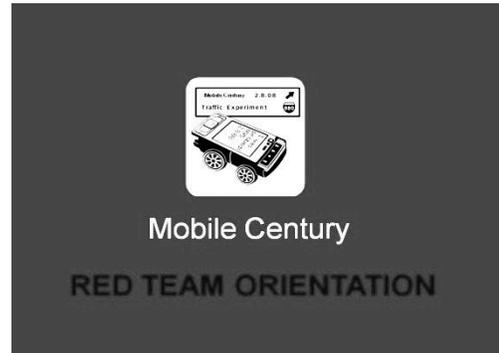


Mobile Century: Red Team Orientation

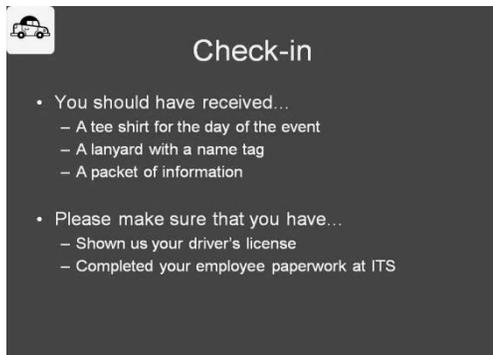
1



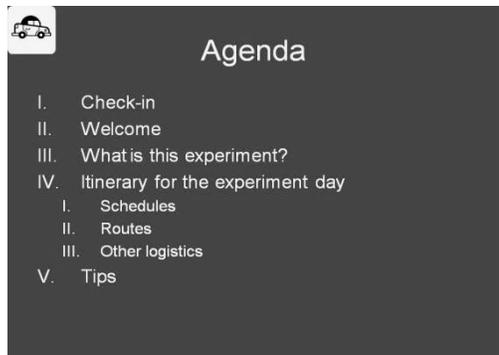
2



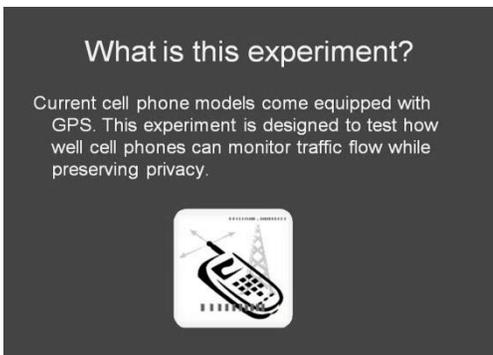
3



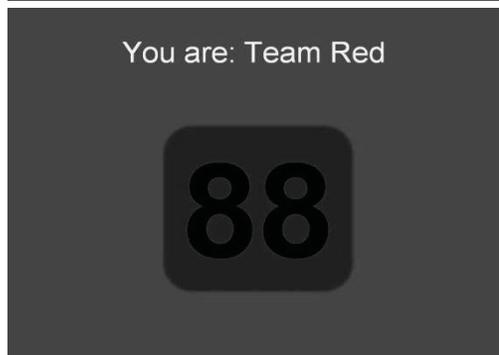
4



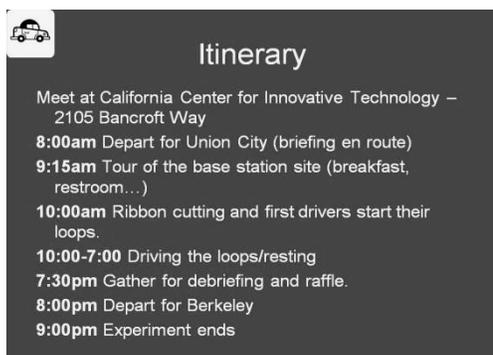
5



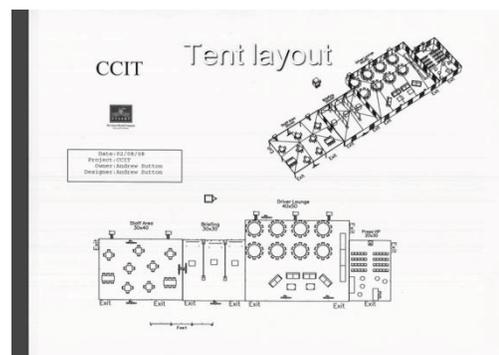
6



7



8



9

Start of a shift

Before your driving shift the team leader will brief drivers of upcoming routes
 These briefings will include information about:
 – Upcoming loop and directions
 After the briefing you will check in with the Team Personnel Assistant.
 – Assign cars, distribute tickets and timers
 Keys remain in the cars! Do not lock them in!
 There will be a Bluetooth headset in the car for communications while driving.
 – Keep the earpiece with you all day.

10



Gas station

- The team traffic assistant will tell you whether to refuel
- The gas station is located across the street from Union Landing lot
- There will be an assistant with a credit card at the gas station to help you.

11

Map of gas station



12

End of a shift

- Enter the parking lot
- Park
- Check out with the personnel assistant
 - Return timer and ticket
- Rest and Relax
- Another briefing will occur before your next shift

13

Schedules

There are four different driving schedules. These will be assigned to you when you check in on the morning of the experiment.



Schedules are based on four letters.

14

Rest time and meals

Driver lounge
 -with heat, TV, couches, coffee and snacks

Meals
 -breakfast and lunch are provided
 -lunch time depends on your driving schedule

Restrooms
 -restroom breaks will be difficult during driving (go before you drive)
 -there will be restroom facilities near the tent.

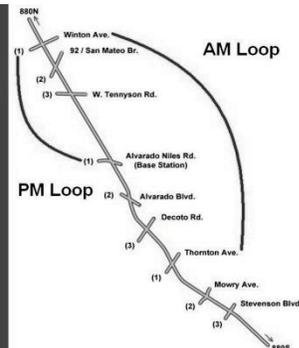
15

Routes

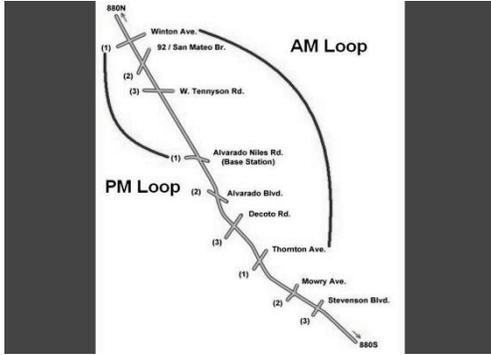


- You will be driving a longer loop in the morning and a shorter loop in the afternoon.
- The northern exit will be the same but the southern one will change.

16



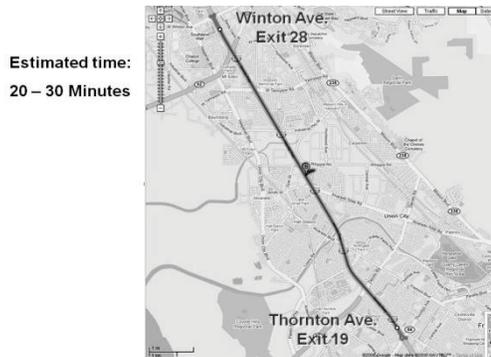
17



18



19



20



21



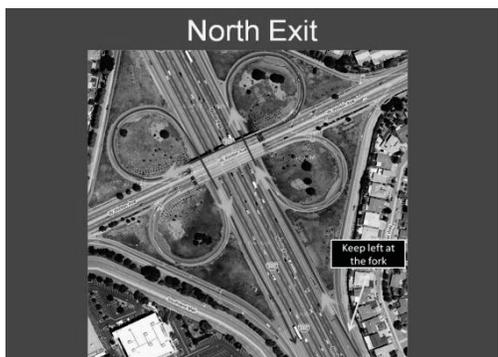
22



23



24



25



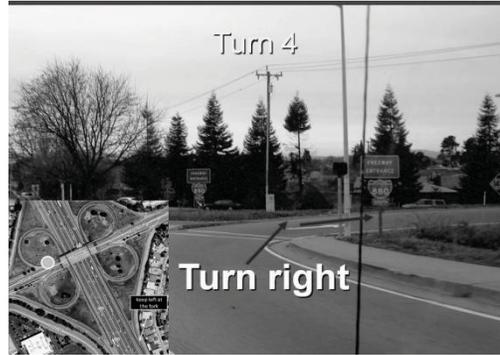
26



27



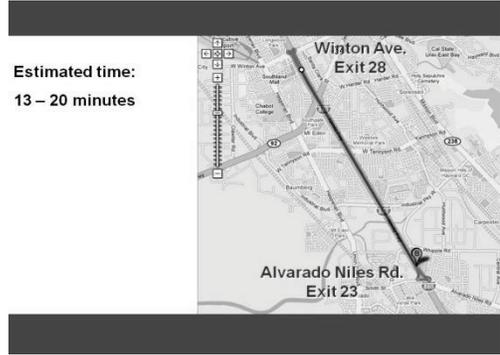
28



29



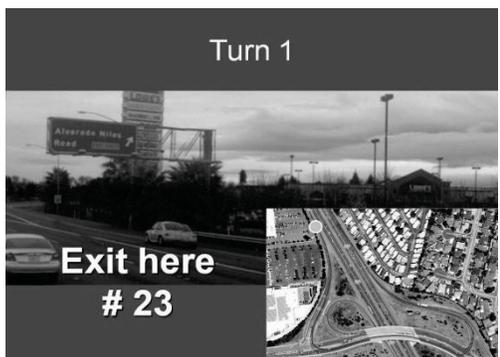
30



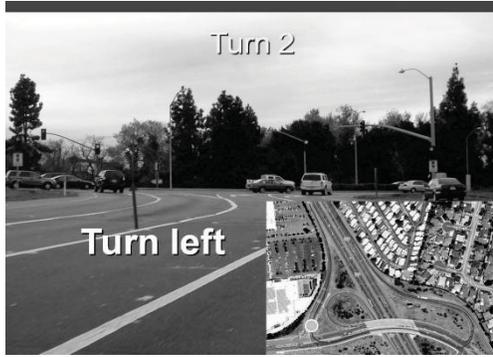
31



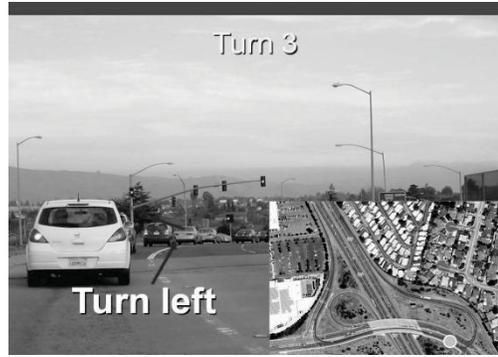
32



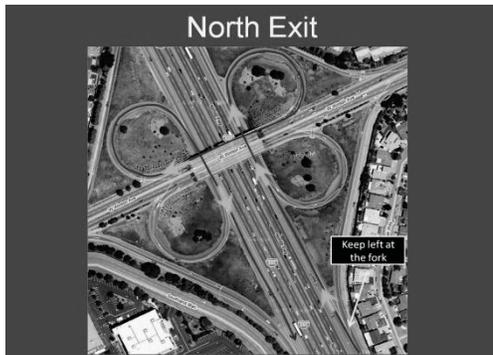
33



34



35



36



37



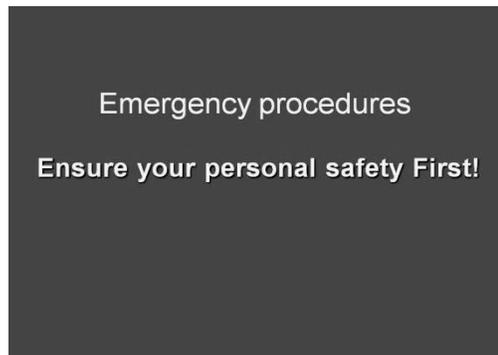
38



39



40



41

Call 911 if your safety is in risk!

- Options available to call 911:
 - Nokia phone
 - Your Bluetooth to communicate with phone operator
 - Your personal phone if available
 - Calling box in the roadside

42

Other Emergency

- The car runs out of gas
- The car breaks down
- Flat tire
- Violation of traffic rules
- I get lost
- I am not feeling well
-



43

Other Emergency

- Contact experiment staff using:
 - Your Bluetooth to communicate with phone operator
 - If the headset fails, you may use your personal cell phone or the Nokia phone.
- We have an ambulance, and tow-truck on hand to assist if necessary

44

How to Use Nokia Phone

To use the Nokia phone:

- Press the red button before dialing
- Dial the number
- And then push the green button (send)

For example:



45

Re-establish contact with phone operators

- After a 911 situation is cleared and released by an authority
 - Connect to phone operator using the number in the dash board
- Then drive back to Union Landing, if possible.



(see your dashboard)



46

How to Use Bluetooth

To use the Bluetooth:

- Double click your Bluetooth button
- Report your situation to the phone operator
- Follow the instructions

47

General safety instructions

- In any event of critical emergency (accident, injury, fatal, or life threatening), call 911 immediately by using all necessary means (Including your personal cell phone if available or Nokia cell phone)!
- If the incident involves second party, driver should stay in the scene until it's cleared by the proper authority (Local Police, CHP...etc).
- All non-life threatening emergency will be reported to and will be handled by Nokia phone operators and the emergency response team.

48

Recall Instructions

- React immediately to the message.
- Remain calm.
- Using the off-ramps (exits) designated to your team, drive back to the Union Landing Mall.
- Stay alert for other upcoming calls from the operators. Be aware the operators may contact you and your team repeatedly and announce the recall message. Please answer all the receiving calls.

49

Recall Instructions

- React immediately to the message.
- Remain calm.
- Using the off-ramps (exits) designated to your team, drive back to the Union Landing Mall.
- Stay alert for other upcoming calls from the operators. Be aware the operators may contact you and your team repeatedly and announce the recall message. Please answer all the receiving calls.

50

Recall Instructions (cont.)

- Upon their arrival to the base station, follow the Team Leader Assistants instructions and park the vehicles in the designated spots. Be extra careful to avoid accident.
- Once parked, report to the Team personnel assistants.
- Go to the drivers lounge area and wait for further instructions.

51



Safety tips



- Get a good night's sleep
- Don't party the night before
- Wear comfortable clothes
- Drive normally and obey traffic laws.
- Wear your seat belt

52

Don't forget



- Driver's license
- Sunglasses/corrective lenses
- Name tag
- The packet materials
- Tee Shirt
- Music (some cars will have CD and MP3 jacks)

53

Additional safety considerations



You will be advised about all pertinent weather and traffic information on the day of the experiment.

54

THANK YOU!

We are so grateful that you are participating!

4.2 Yellow Team

Power Point Presentation: Yellow Team Orientation



Mobile Century: Yellow Team Orientation

1



Mobile Century
YELLOW TEAM ORIENTATION

2



Check-in

- You should have received...
 - A tee shirt for the day of the event
 - A lanyard with a name tag
 - A packet of information
- Please make sure that you have...
 - Shown us your driver's license
 - Completed your employee paperwork at ITS

3



Agenda

- Check-in
- Welcome
- What is this experiment?
- Itinerary for the experiment day
 - Schedules
 - Routes
 - Other logistics
- Tips

4

What is this experiment?

Current cell phone models come equipped with GPS. This experiment is designed to test how well cell phones can monitor traffic flow while preserving privacy.



5

You are: Team Yellow



6



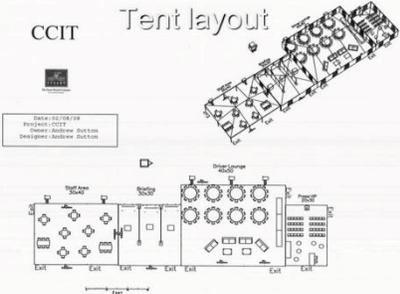
Itinerary

Meet at California Center for Innovative Technology – 2105 Bancroft Way

- 8:00am** Depart for Union City (briefing en route)
- 9:15am** Tour of the base station site (breakfast, restroom...)
- 10:00am** Ribbon cutting and first drivers start their loops.
- 10:00-7:00** Driving the loops/resting
- 7:30pm** Gather for debriefing and raffle.
- 8:00pm** Depart for Berkeley
- 9:00pm** Experiment ends

7

CCIT Tent layout



CCIT
Date: 10/24/08
Project: CCIT
Client: Andrew Mathson
Designer: Andrew Mathson

8

Start of a shift

Before your driving shift the team leader will brief drivers of upcoming routes

These briefings will include information about:

- Upcoming loop and directions

After the briefing you will check in with the Team Personnel Assistant.

- Assign cars, distribute tickets and timers

Keys remain in the cars! Do not lock them in!

There will be a Bluetooth headset in the car for communications while driving.

- Keep the earpiece with you all day.

9



Gas station

- The team traffic assistant will tell you whether to refuel
- The gas station is located across the street from Union Landing lot
- There will be an assistant with a credit card at the gas station to help you.

10

Map of gas station



11

End of a shift

- Enter the parking lot
- Park
- Check out with the personnel assistant
 - Return timer and ticket
- Rest and Relax
- Another briefing will occur before your next shift

12

Schedules

There are four different driving schedules. These will be assigned to you when you check in on the morning of the experiment.



Schedules are based on four letters.

13

Rest time and meals

- Driver lounge
-with heat, TV, couches, coffee and snacks
- Meals
-breakfast and lunch are provided
-lunch time depends on your driving schedule
- Restrooms
-restroom breaks will be difficult during driving (go before you drive)
-there will be restroom facilities near the tent.

14

During rest time...

- | | |
|---|--|
| <p>DO</p> <ul style="list-style-type: none"> • Relax • Watch TV • Eat • Read or study • Visit the restroom | <p>DON'T</p> <ul style="list-style-type: none"> • Leave the tent area without notifying your team personnel assistant |
|---|--|



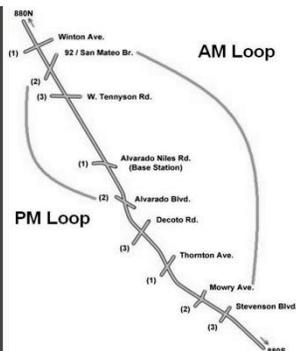
15

Routes



- You will be driving a longer loop in the morning and a shorter loop in the afternoon.
- The northern exit will be the same but the southern one will change.

16

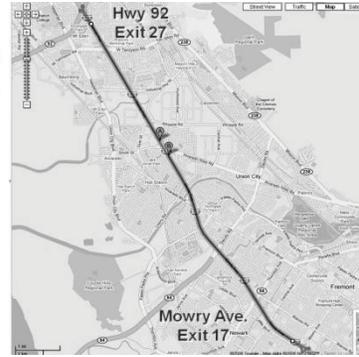


17



18

Estimated time:
20 – 30 Minutes



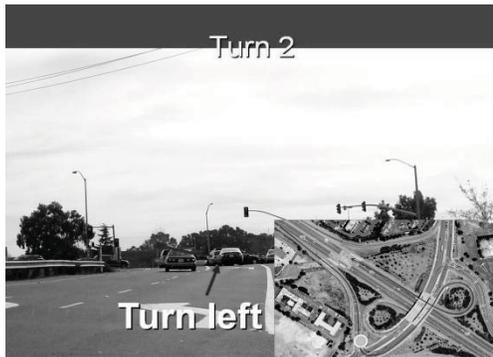
19



20



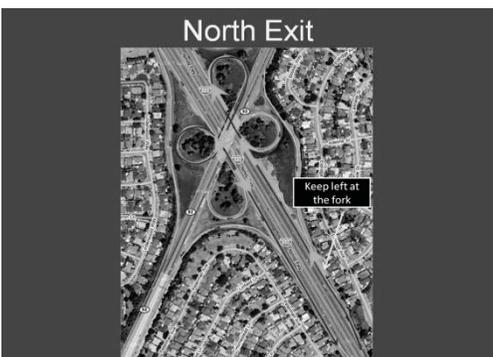
21



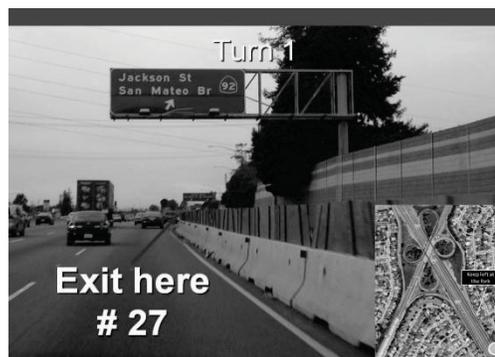
22



23



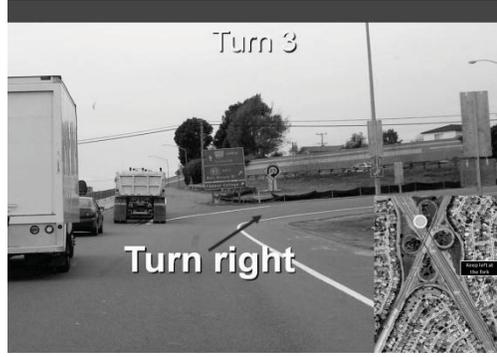
24



25



26



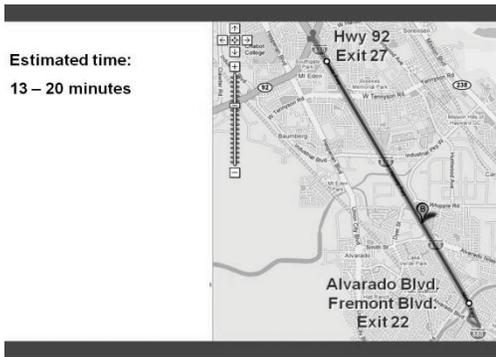
27



28



29



30



31



32



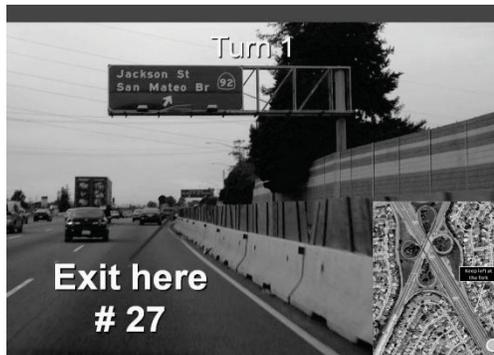
33



34



35



36



37



38



39

Emergency procedures
Ensure your personal safety First!

40

Call 911 if your safety is in risk!

- Options available to call 911:
 - Nokia phone
 - Your Bluetooth to communicate with phone operator
 - Your personal phone if available
 - Calling box in the roadside

41

Other Emergency

- The car runs out of gas
- The car breaks down
- Flat tire
- Violation of traffic rules
- I get lost
- I am not feeling well
-



42

Other Emergency

- Contact experiment staff using:
 - Your Bluetooth to communicate with phone operator
 - If the headset fails, you may use your personal cell phone or the Nokia phone.
- We have an ambulance, and tow-truck on hand to assist if necessary

43

How to Use Nokia Phone

To use the Nokia phone:

- Press the red button before dialing
- Dial the number
- And then push the green button (send)

For example:

RED 9 1 1 GREEN



44

Re-establish contact with phone operators

- After a 911 situation is cleared and released by an authority
 - Connect to phone operator using the number in the dash board
- RED Operator Phone Number GREEN
- (see your dashboard)*
- Then drive back to Union Landing, if possible.



45

How to Use Bluetooth

To use the Bluetooth:

- Double click your Bluetooth button
- Report your situation to the phone operator
- Follow the instructions

46

General safety instructions

- In any event of critical emergency (accident, injury, fatal, or life threatening), call 911 immediately by using all necessary means (including your personal cell phone if available or Nokia cell phone!)
- If the incident involves second party, driver should stay in the scene until it's cleared by the proper authority (Local Police, CHP... etc).
- All non-life threatening emergency will be reported to and will be handled by Nokia phone operators and the emergency response team.

47

General safety instructions (cont.)

- If encountering car breakdown, flat tire, or out of gas, please pull aside your car from the paved road to the road shoulder; if possible stay behind the road guardrail and stay away from the traffic.
- If you have to walk, always walk facing the traffic, be alert and try not to interrupt traffic.
- Always follow the emergency handlers' instruction and follow the proper safety driving rules and regulations.

48

Recall Instructions

- React immediately to the message.
- Remain calm.
- Using the off-ramps (exits) designated to your team, drive back to the Union Landing Mall.
- Stay alert for other upcoming calls from the operators. Be aware the operators may contact you and your team repeatedly and announce the recall message. Please answer all the receiving calls.

49

Recall Instructions (cont.)

- Upon their arrival to the base station, follow the Team Leader Assistants instructions and park the vehicles in the designated spots. Be extra careful to avoid accident.
- Once parked, report to the Team personnel assistants.
- Go to the drivers lounge area and wait for further instructions.

50



Safety tips



- Get a good night's sleep
- Don't party the night before
- Wear comfortable clothes
- Drive normally and obey traffic laws.
- Wear your seat belt

51

Don't forget



- Driver's license
- Sunglasses/corrective lenses
- Name tag
- The packet materials
- Tee Shirt
- Music (some cars will have CD and MP3 jacks)

52

Additional safety considerations



You will be advised about all pertinent weather and traffic information on the day of the experiment.

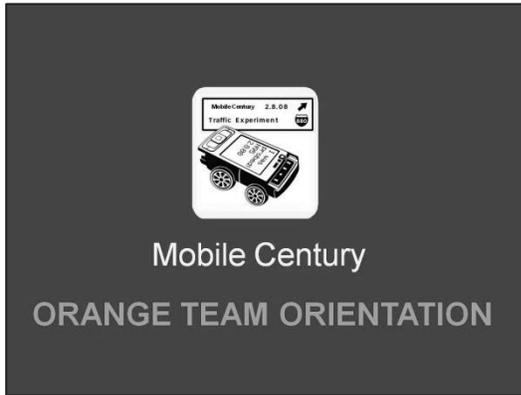
53

THANK YOU!

We are so grateful that you are participating!

4.3 Orange Team

PowerPoint Presentation: Orange Team Orientation



Mobile Century: Orange Team Orientation

1



Mobile Century
ORANGE TEAM ORIENTATION

2



Check-in

- You should have received...
 - A tee shirt for the day of the event
 - A lanyard with a name tag
 - A packet of information
- Please make sure that you have...
 - Shown us your driver's license
 - Completed your employee paperwork at ITS

3



Agenda

- Check-in
- Welcome
- What is this experiment?
- Itinerary for the experiment day
 - Schedules
 - Routes
 - Other logistics
- Tips

4

What is this experiment?

Current cell phone models come equipped with GPS. This experiment is designed to test how well cell phones can monitor traffic flow while preserving privacy.



5

You are: Team Orange



6



Itinerary

Meet at California Center for Innovative Technology – 2105 Bancroft Way

8:00am Depart for Union City (briefing en route)

9:15am Tour of the base station site (breakfast, restroom...)

10:00am Ribbon cutting and first drivers start their loops.

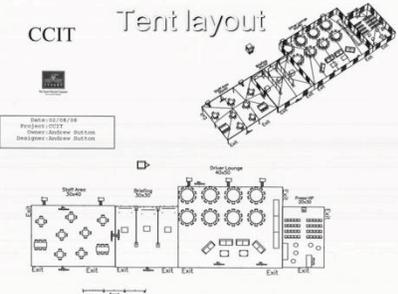
10:00-7:00 Driving the loops/resting

7:30pm Gather for debriefing and raffle.

8:00pm Depart for Berkeley

9:00pm Experiment ends

7



CCIT Tent layout

DATE: 11/27/08
PROJECT: 0127
OWNER: Caltrans - 2105 Bancroft
DRAWN BY: Andrew Smith

8

Start of a shift

Before your driving shift the team leader will brief drivers of upcoming routes

These briefings will include information about:

- Upcoming loop and directions

After the briefing you will check in with the Team Personnel Assistant.

- Assign cars, distribute tickets and timers

Keys remain in the cars! Do not lock them in!

There will be a Bluetooth headset in the car for communications while driving.

- Keep the earpiece with you all day.

9



Gas station

- The team traffic assistant will tell you whether to refuel
- The gas station is located across the street from Union Landing lot
- There will be an assistant with a credit card at the gas station to help you.

10

Map of gas station



11

End of a shift

- Enter the parking lot
- Park
- Check out with the personnel assistant
 - Return timer and ticket
- Rest and Relax
- Another briefing will occur before your next shift

12

Schedules

There are four different driving schedules. These will be assigned to you when you check in on the morning of the experiment.



Schedules are based on four letters.

13

Rest time and meals

- Driver lounge
- with heat, TV, couches, coffee and snacks
- Meals
- breakfast and lunch are provided
 - lunch time depends on your driving schedule
- Restrooms
- restroom breaks will be difficult during driving (go before you drive)
 - there will be restroom facilities near the tent.

14

During rest time...

- | | |
|---|--|
| <p>DO</p> <ul style="list-style-type: none"> • Relax • Watch TV • Eat • Read or study • Visit the restroom | <p>DON'T</p> <ul style="list-style-type: none"> • Leave the tent area without notifying your team personnel assistant |
|---|--|



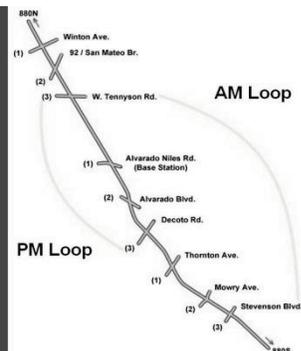
15

Routes

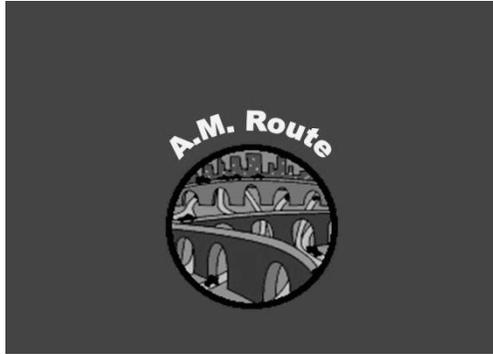


- You will be driving a longer loop in the morning and a shorter loop in the afternoon.
- The northern exit will be the same but the southern one will change.

16



17



18

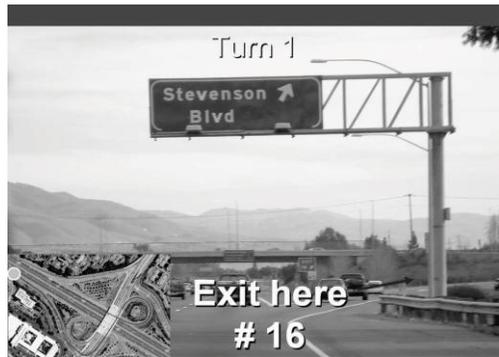
Estimated time:
20 – 30 Minutes



19



20



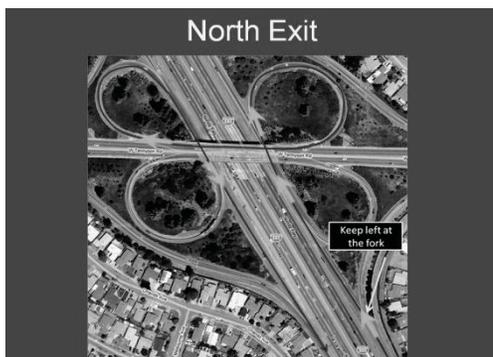
21



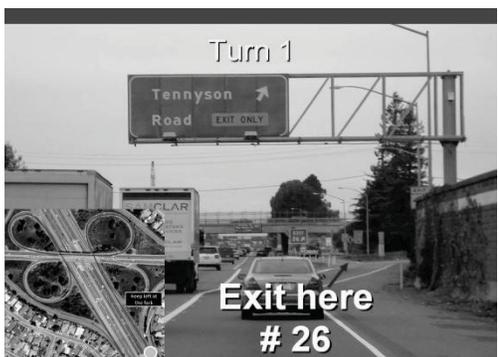
22



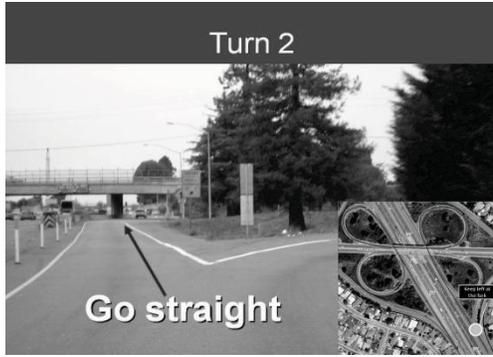
23



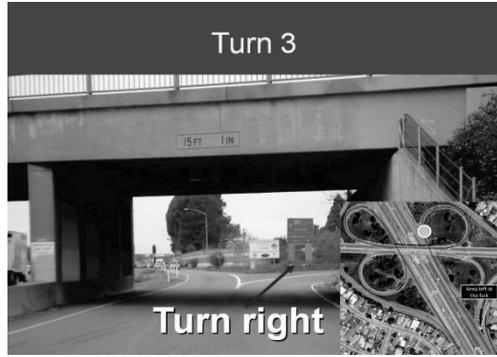
24



25



26



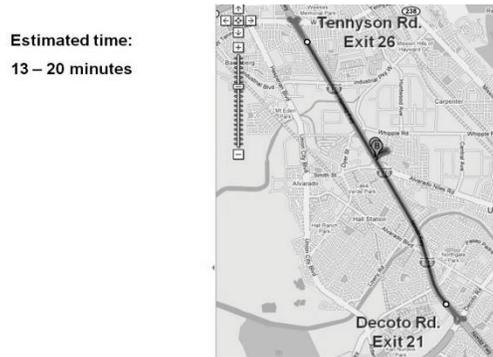
27



28



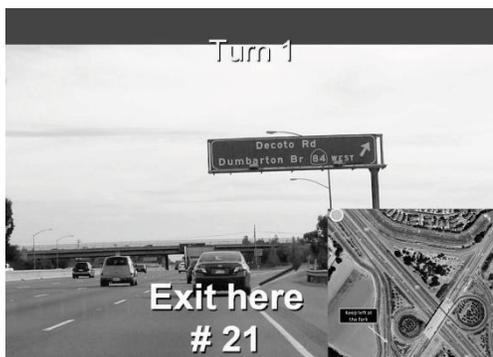
29



30



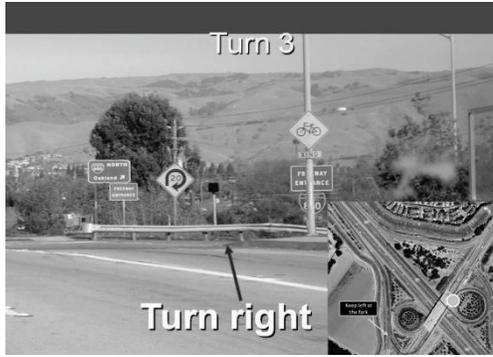
31



32



33



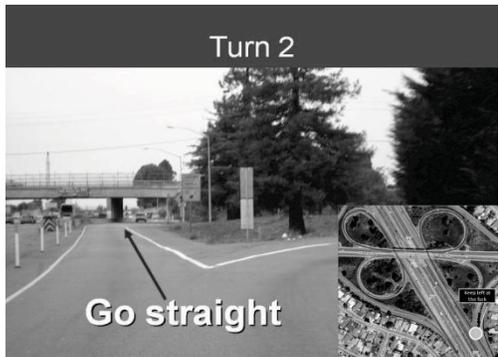
34



35



36



37



38



39

Emergency procedures
Ensure your personal safety First!

40

- Call 911 if your safety is in risk!**
- Options available to call 911:
 - Nokia phone
 - Your Bluetooth to communicate with phone operator
 - Your personal phone if available
 - Calling box in the roadside

41

Other Emergency



The car runs out of gas
 The car breaks down
 Flat tire
 Violation of traffic rules
 I get lost
 I am not feeling well



42

Other Emergency

- Contact experiment staff using:
 - Your Bluetooth to communicate with phone operator
 - If the headset fails, you may use your personal cell phone or the Nokia phone.
- We have an ambulance, and tow-truck on hand to assist if necessary

43

How to Use Nokia Phone

To use the Nokia phone:

- Press the red button before dialing
- Dial the number
- And then push the green button (send)

For example:

RED 9 1 1 GREEN



44

Re-establish contact with phone operators

- After a 911 situation is cleared and released by an authority
 - Connect to phone operator using the number in the dash board

RED Operator Phone Number GREEN

(see your dashboard)

- Then drive back to Union Landing, if possible.



45

How to Use Bluetooth

To use the Bluetooth:

- Double click your Bluetooth button
- Report your situation to the phone operator
- Follow the instructions

46

General safety instructions

- In any event of critical emergency (accident, injury, fatal, or life threatening), call 911 immediately by using all necessary means (including your personal cell phone if available or Nokia cell phone!)
- If the incident involves second party, driver should stay in the scene until it's cleared by the proper authority (Local Police, CHP... etc).
- All non-life threatening emergency will be reported to and will be handled by Nokia phone operators and the emergency response team.

47

General safety instructions (cont.)

- If encountering car breakdown, flat tire, or out of gas, please pull aside your car from the paved road to the road shoulder; if possible stay behind the road guardrail and stay away from the traffic.
- If you have to walk, always walk facing the traffic, be alert and try not to interrupt traffic.
- Always follow the emergency handlers' instruction and follow the proper safety driving rules and regulations.

48

Recall Instructions

- React immediately to the message.
- Remain calm.
- Using the off-ramps (exits) designated to your team, drive back to the Union Landing Mall.
- Stay alert for other upcoming calls from the operators. Be aware the operators may contact you and your team repeatedly and announce the recall message. Please answer all the receiving calls.

49

Recall Instructions (cont.)

- Upon their arrival to the base station, follow the Team Leader Assistants instructions and park the vehicles in the designated spots. Be extra careful to avoid accident.
- Once parked, report to the Team personnel assistants.
- Go to the drivers lounge area and wait for further instructions.

50



Safety tips



- Get a good night's sleep
- Don't party the night before
- Wear comfortable clothes
- Drive normally and obey traffic laws.
- Wear your seat belt

51

Don't forget



- Driver's license
- Sunglasses/corrective lenses
- Name tag
- The packet materials
- Tee Shirt
- Music (some cars will have CD and MP3 jacks)

52

Additional safety considerations



You will be advised about all pertinent weather and traffic information on the day of the experiment.

53

THANK YOU!

We are so grateful that you are participating!

5 Driver Instructions

Appendix 5 contains handouts that were provided to each participating driver according to his team color. These handouts provide a Google map inset in addition to a street-view photo of decision points at the northern and southern edges of each route. There are two versions for each team, corresponding to the morning and afternoon routes.

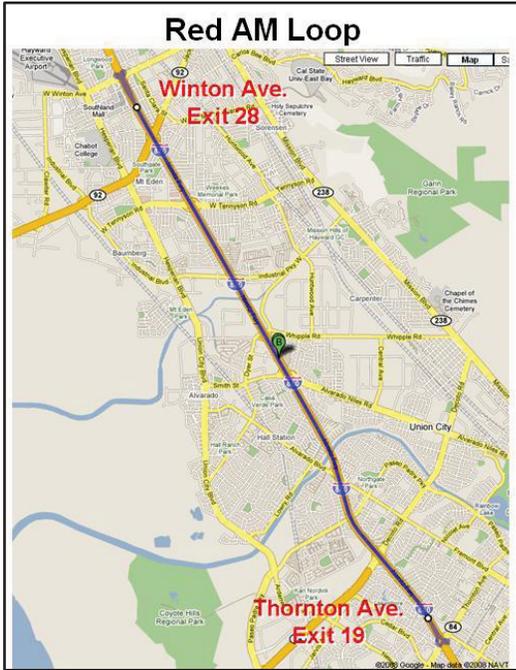
The Red Team was assigned to routes that extended as far north as Winton Ave., and as far south as Thornton Ave. and Alvarado Niles Rd. for the morning and afternoon routes as illustrated in Appendices 5.1 and 5.2, respectively.

The Yellow Team was assigned to routes that extended as far north as SR-92 (San Mateo Bridge), and as far south as Mowry Ave. and Alvarado Blvd. for the morning and afternoon routes, as illustrated in Appendices 5.3 and 5.4, respectively.

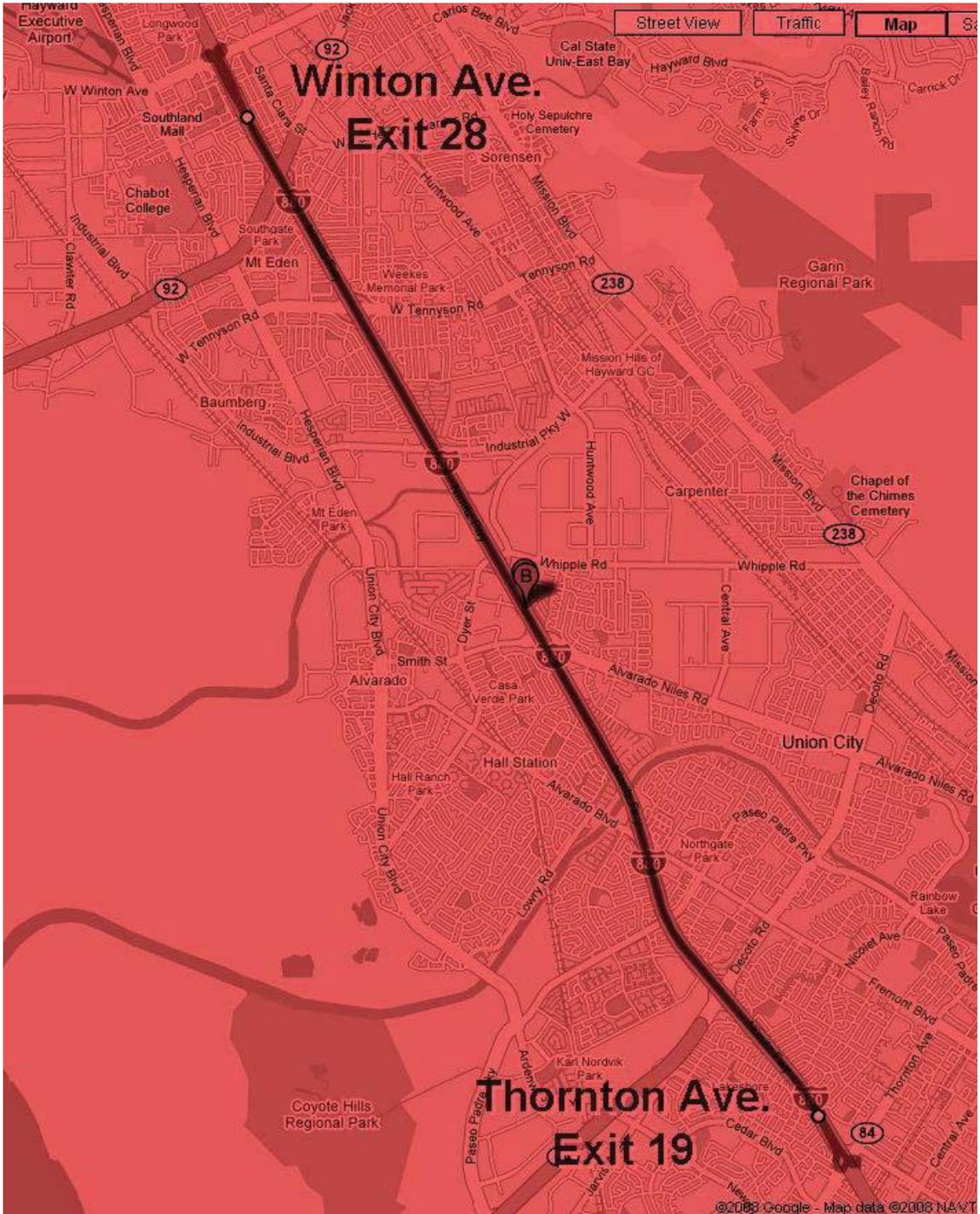
The Orange Team was assigned to routes that extended as far north as Tennyson Rd., and as far south as Stevenson Blvd. and Decoto Rd. for the morning and afternoon routes, as illustrated in Appendices 5.5 and 5.6, respectively.

5.1 Red AM route

Driver Handout for Red AM Loop



Driver Handout: Red AM Loop



Winton Ave. Exit 28

Turn
1



Exit at Winton Ave.

Turn
2



Keep left and go
under the overpass.

Turn
3



Turn right at the off-
ramp.

Turn
4



Immediately turn right at the on-ramp.

Thornton Ave. Exit 19

Turn
1



Exit at Thornton Ave.

Turn
2



Turn left and go over the overpass.

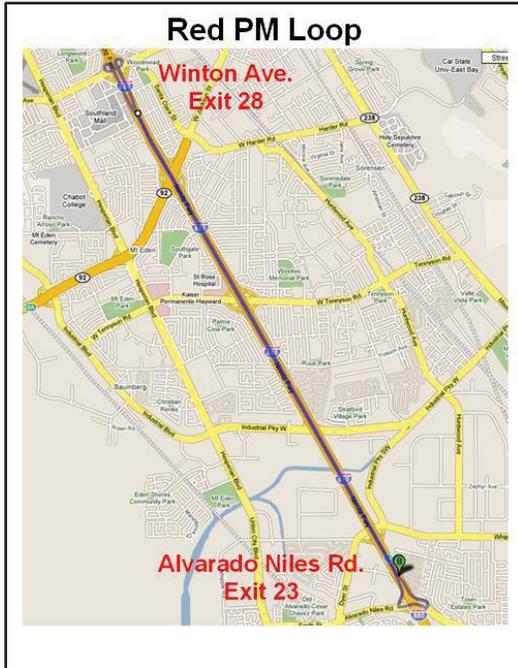
Turn
3



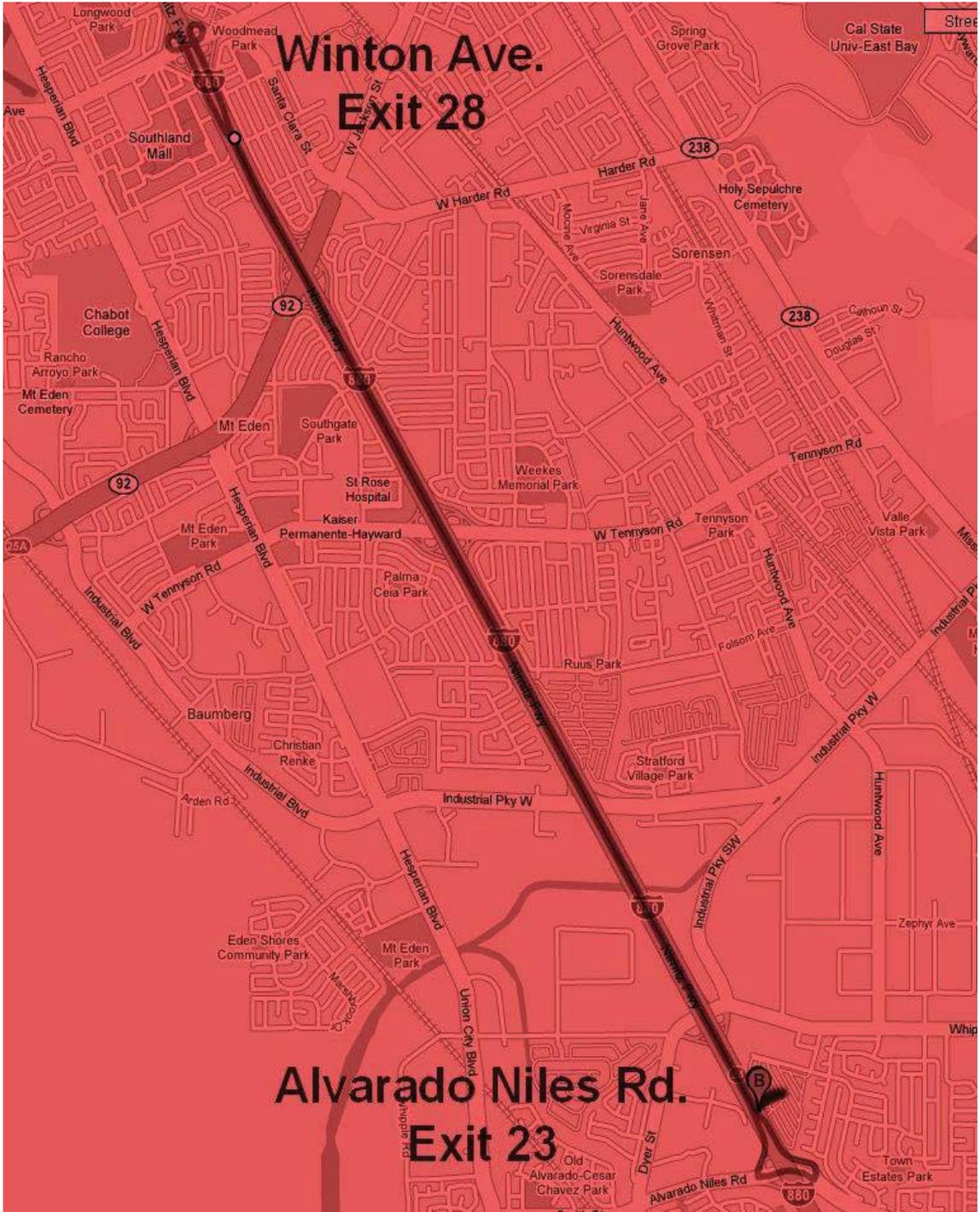
Turn right at the on-ramp.

5.2 Red PM route

Driver Handout for Red PM Loop



Driver Handout: Red PM Loop



Winston Ave. Exit 28

Turn
1



Exit at Winton Ave.

Turn
2



Keep left and go
under the overpass.

Turn
3



Turn right at the off-
ramp.

Turn
4



Immediately turn right at the on-ramp.

Alvarado Niles Road Exit 23

Turn
1



Exit at Alvarado Niles Road

Turn
2



Turn left and go over the overpass.

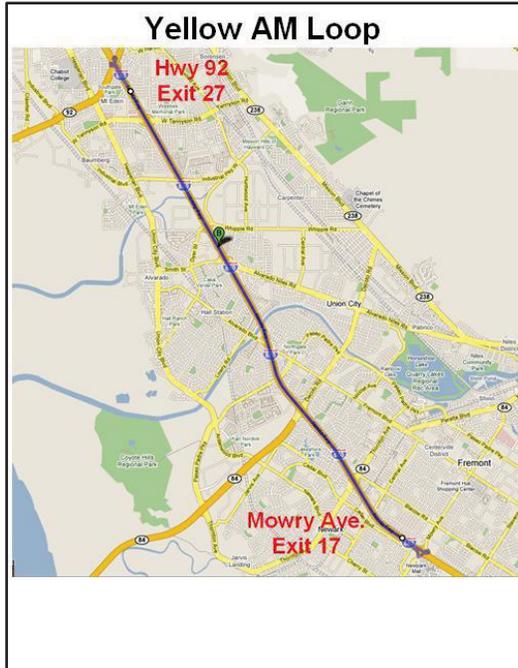
Turn
3



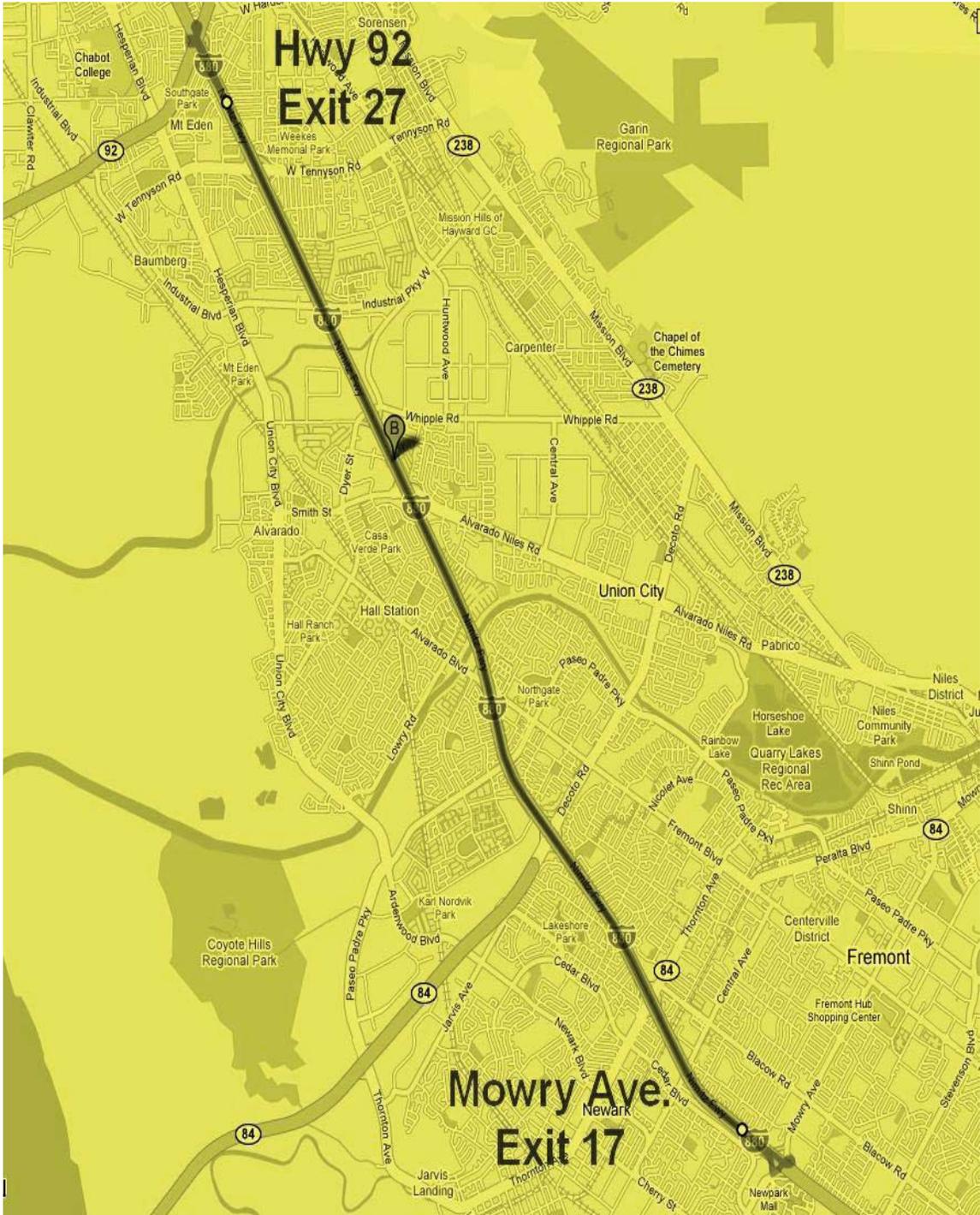
Turn left at the on-ramp.

5.3 Yellow AM route

Driver Handout for Yellow AM Loop



Driver Handout: Yellow AM Loop



Highway 92, Jackson St. Exit 27

Turn
1



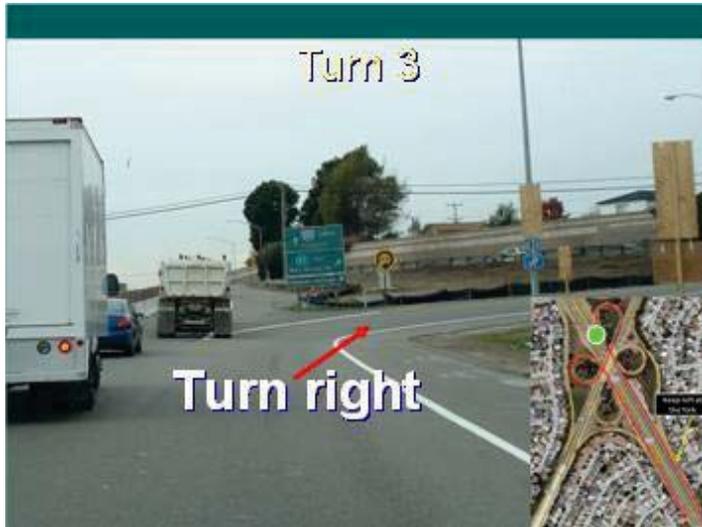
Exit at Highway 92,
Jackson St.

Turn
2



Keep left and go
under the overpass.

Turn
3



Turn right at the off-ramp.

Turn
4



Immediately turn right at the on-ramp.

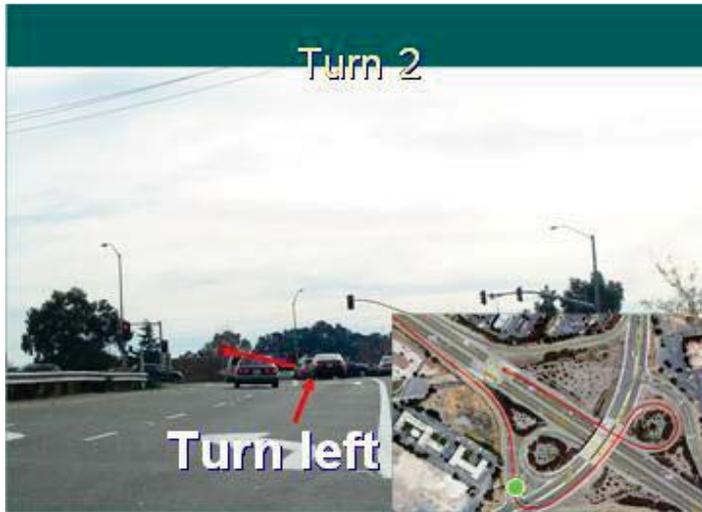
Mowry Ave. Exit 17

Turn
1



Exit at Mowry Ave.

Turn
2



Turn left and go over the overpass.

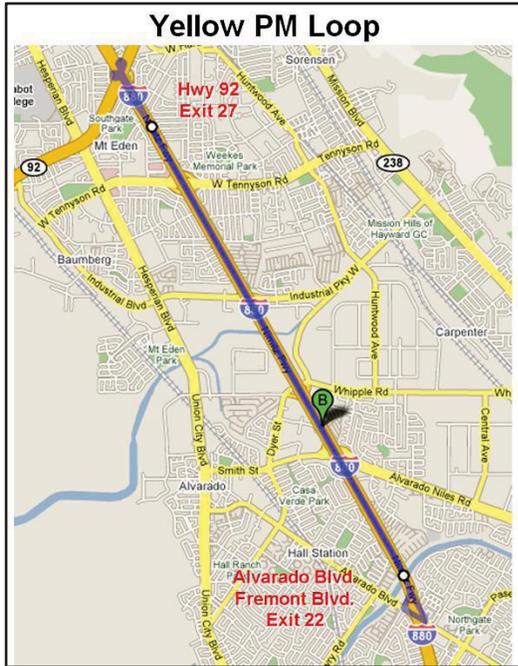
Turn
3



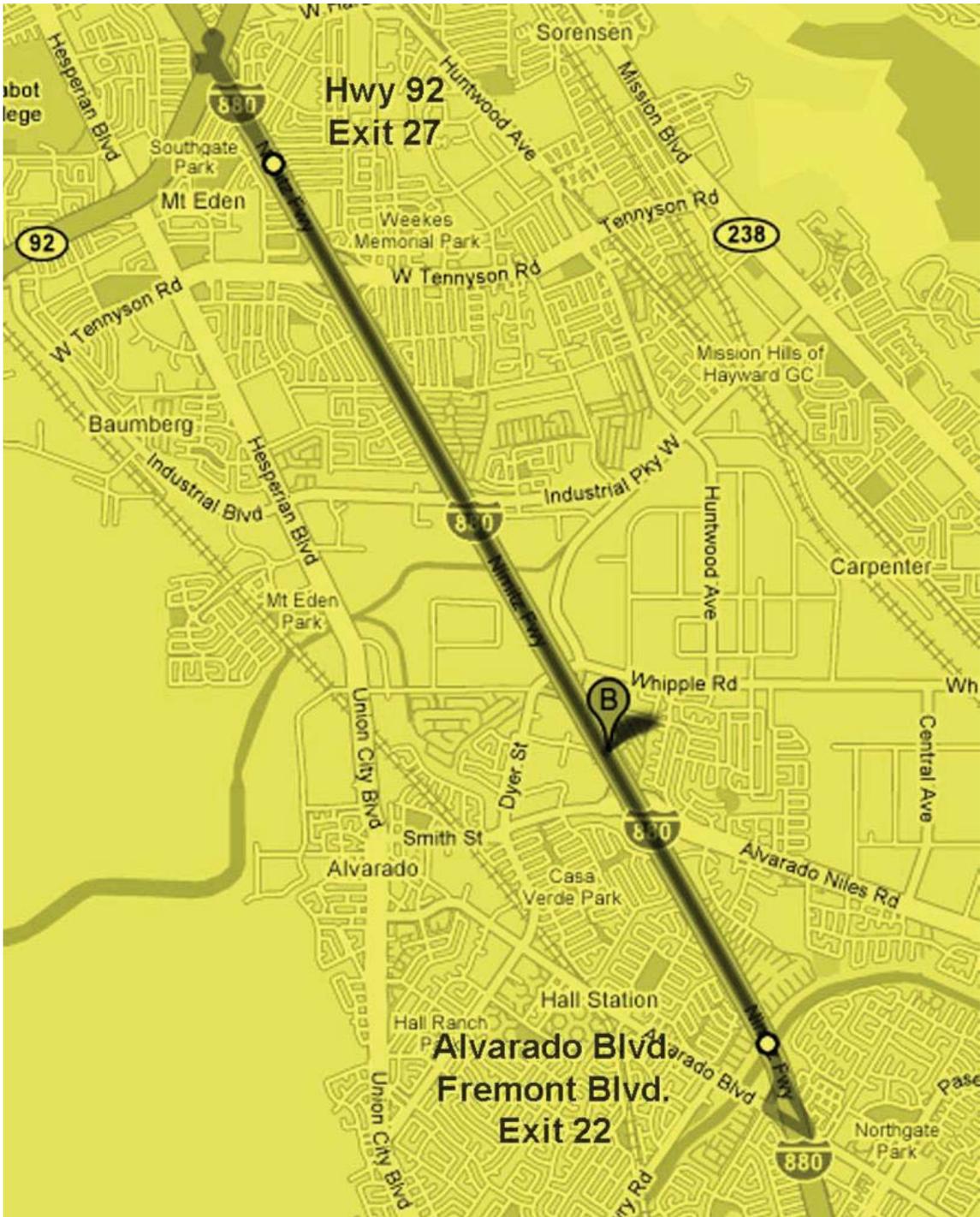
Turn left at the on-ramp.

5.4 Yellow PM route

Driver Handout for Yellow PM Loop



Driver Handout: Yellow PM Loop



Highway 92, Jackson St. Exit 27

Turn
1



Exit at Highway 92,
Jackson St.

Turn
2



Keep left and go
under the overpass.

Turn
3



Turn right at the off-
ramp.

Turn
4



Immediately turn right at the on-ramp.

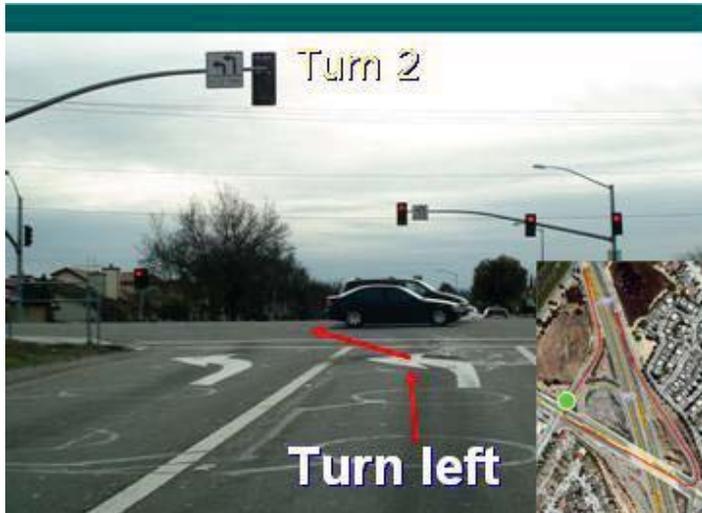
Alvarado Blvd. Fremont Blvd. Exit 22

Turn
1



Exit at Fremont Blvd, Alvarado Blvd.

Turn
2



Turn left and go over the overpass.

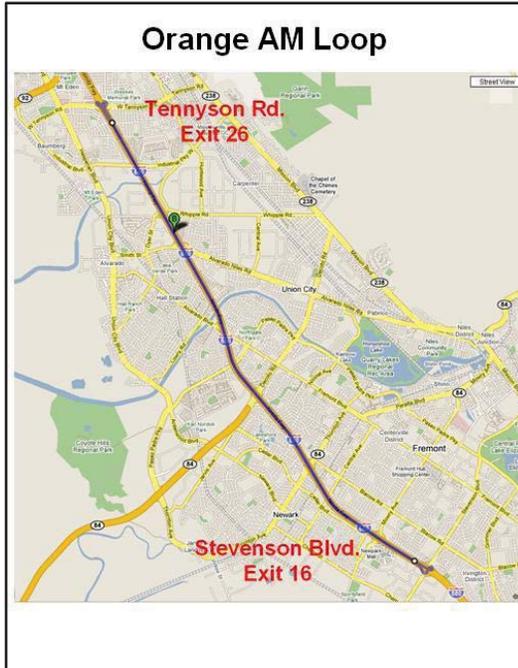
Turn
3



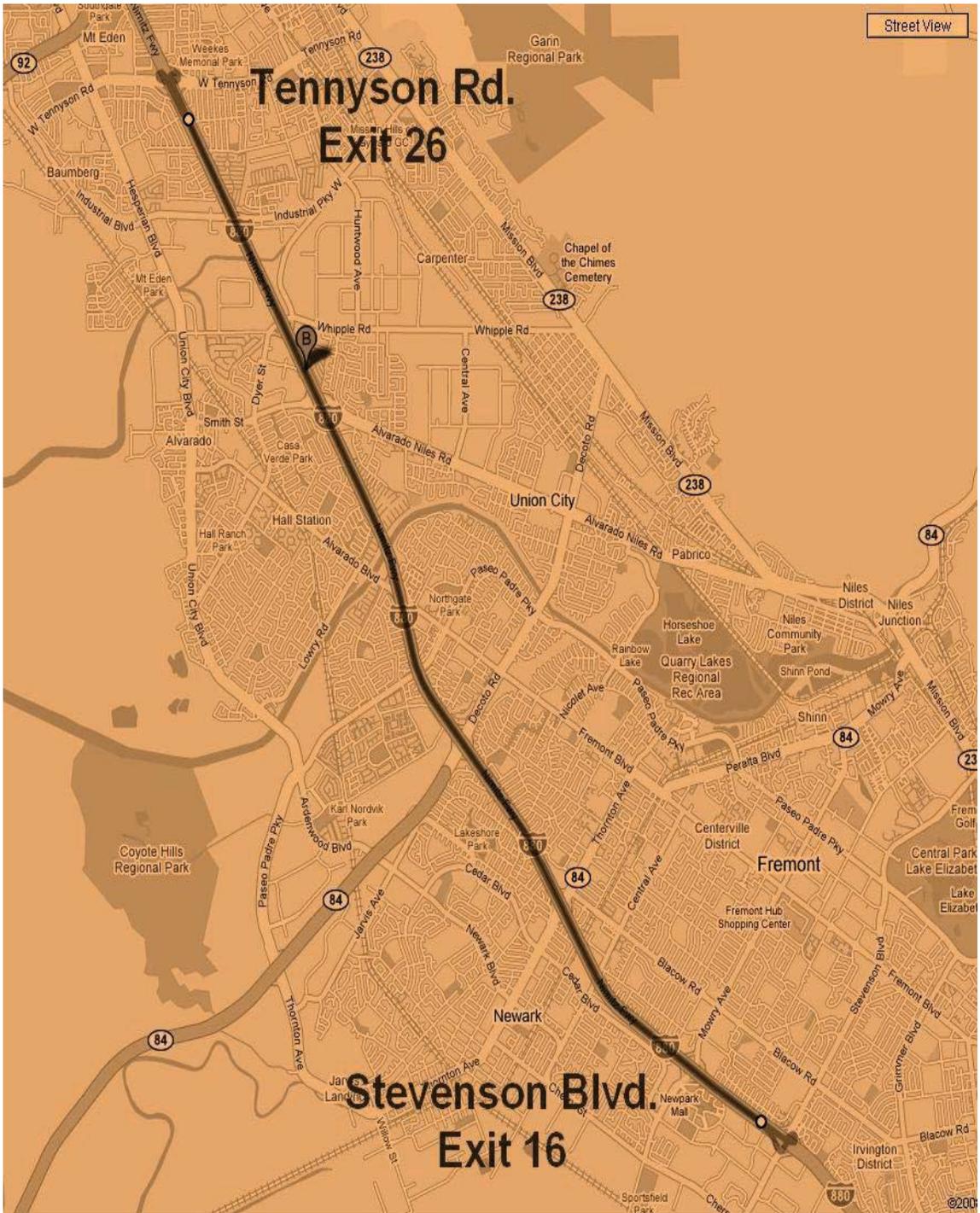
Turn left at the on-ramp.

5.5 Orange AM route

Driver Handout for Orange AM Loop



Driver Handout: Orange AM Loop



Tennyson Road Exit 26

Turn
1



Exit at Tennyson Road.

Turn
2



Keep left and go under the overpass.

Turn
3



Turn right at the off-ramp.

Turn
4



Immediately turn right at the on-ramp.

Stevenson Blvd. Exit 16

Turn
1



Exit at Stevenson Blvd.

Turn
2



Turn left and go over the overpass.

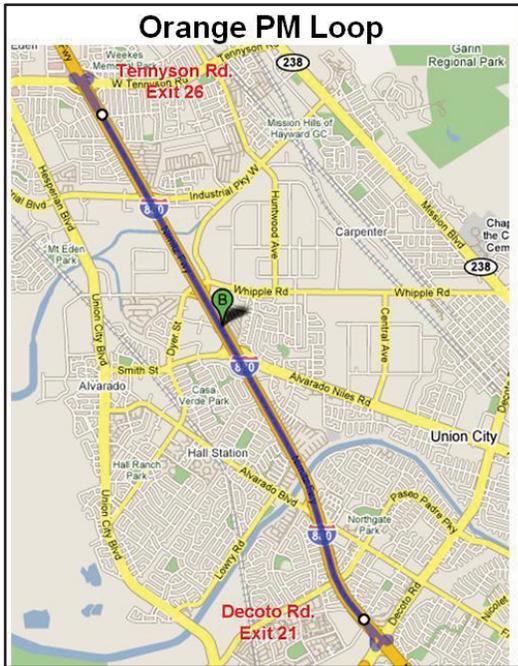
Turn
3



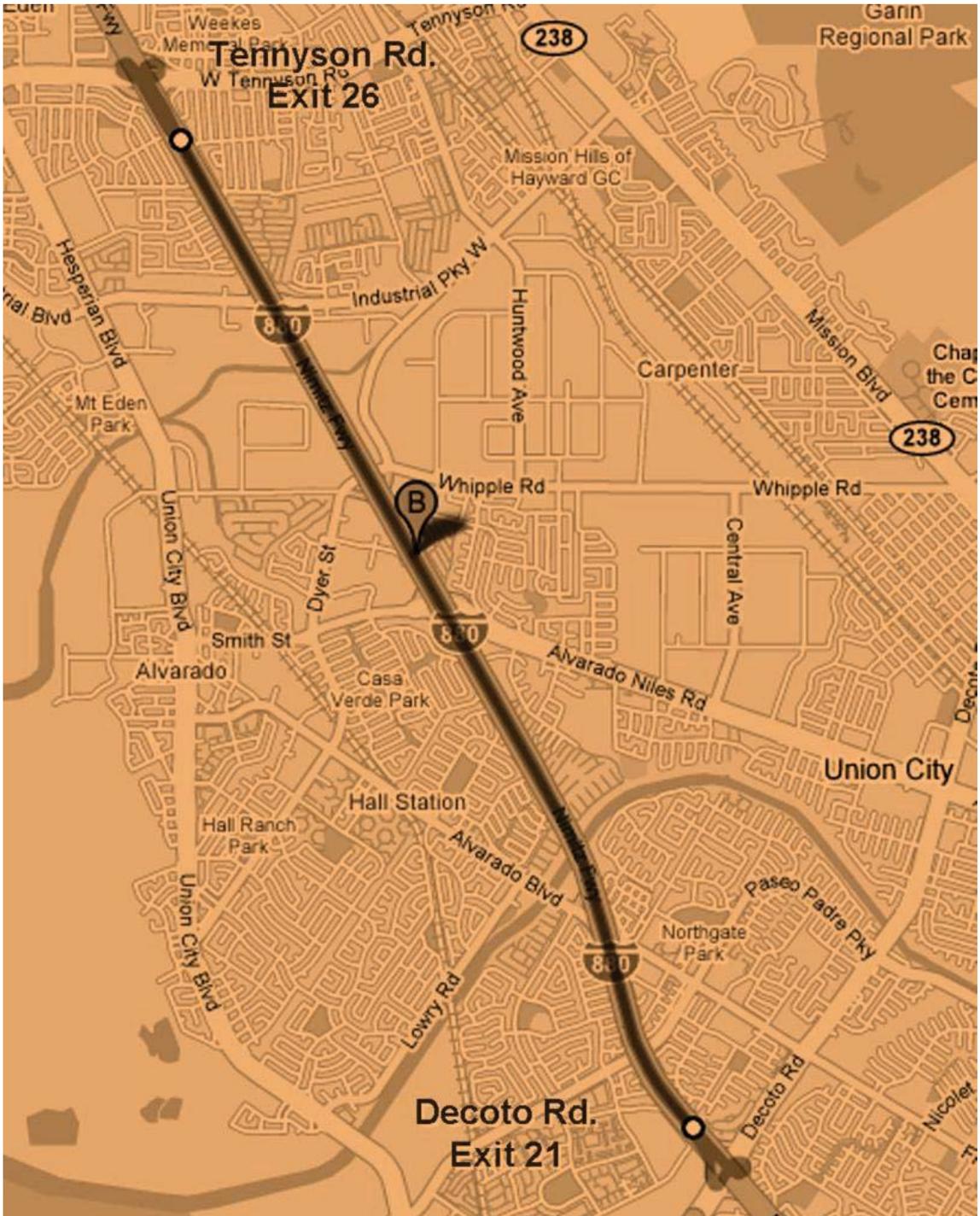
Turn right at the on-ramp.

5.6 Orange PM route

Driver Handout for Orange PM Loop



Driver Handout: Orange PM Loop



Tennyson Road Exit 26

Turn
1



Exit at Tennyson Road.

Turn
2



Keep left and go under the overpass.

Turn
3



Turn right at the off-ramp.

Turn
4



Immediately turn right at the on-ramp.

Decoto Road, Exit 21

Turn
1



Exit at Decoto Road, Dumbarton Bridge

Turn
2



Turn left and go over the overpass.

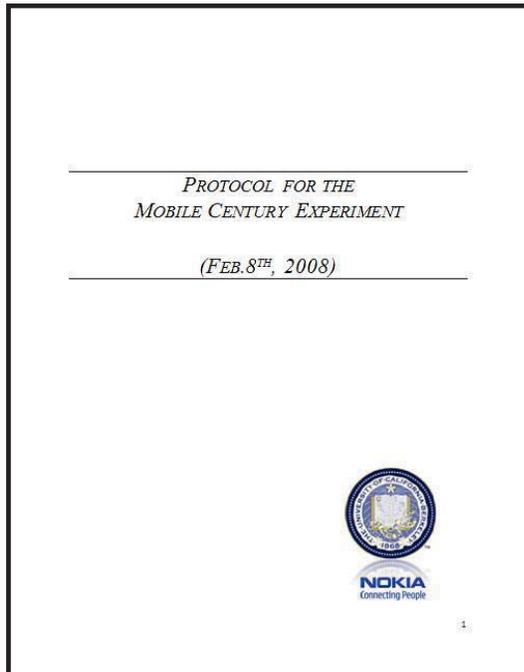
Turn
3



Turn right at the on-ramp.

6 Field Experiment Protocol

Protocol for the Mobile Century Experiment



This Appendix contains the last formal version of the Mobile Century experimental protocol as it was written on January 30, 2008. It is included here for historical purposes, and is presented in the form in which it was recovered. Portions highlighted in yellow or formatted with red text were preserved here. This document encapsulates the overall spirit of the operation, the roles and responsibilities of the main actors, and detailed schedules for the coordination of resources.

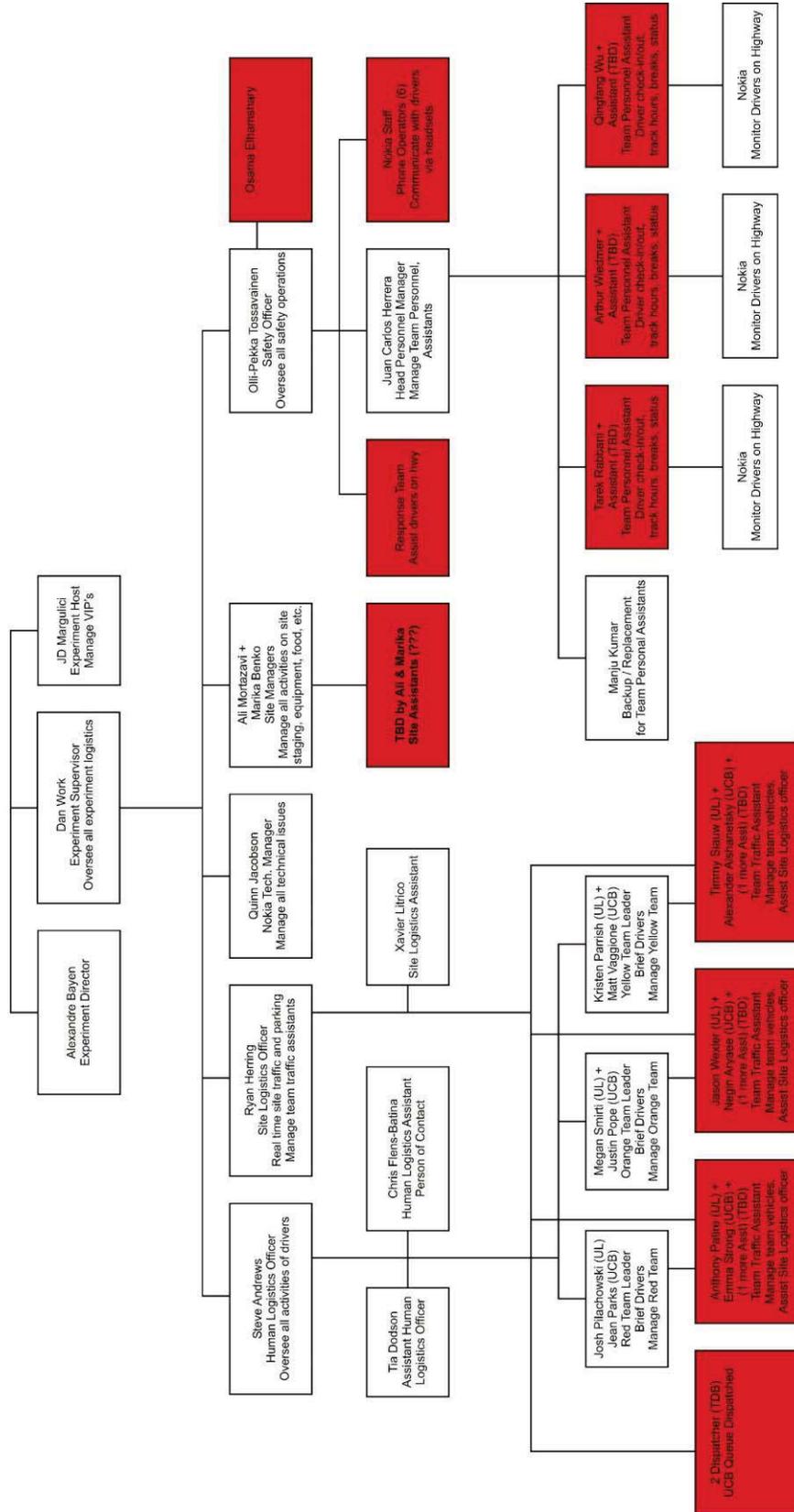
The actual execution of the Mobile Century experiment differed somewhat from the final version of the protocol as provided here. The main deviations between the plan and the experiment as it was actually executed are explained in Chapter 8.

*PROTOCOL FOR THE
MOBILE CENTURY EXPERIMENT*

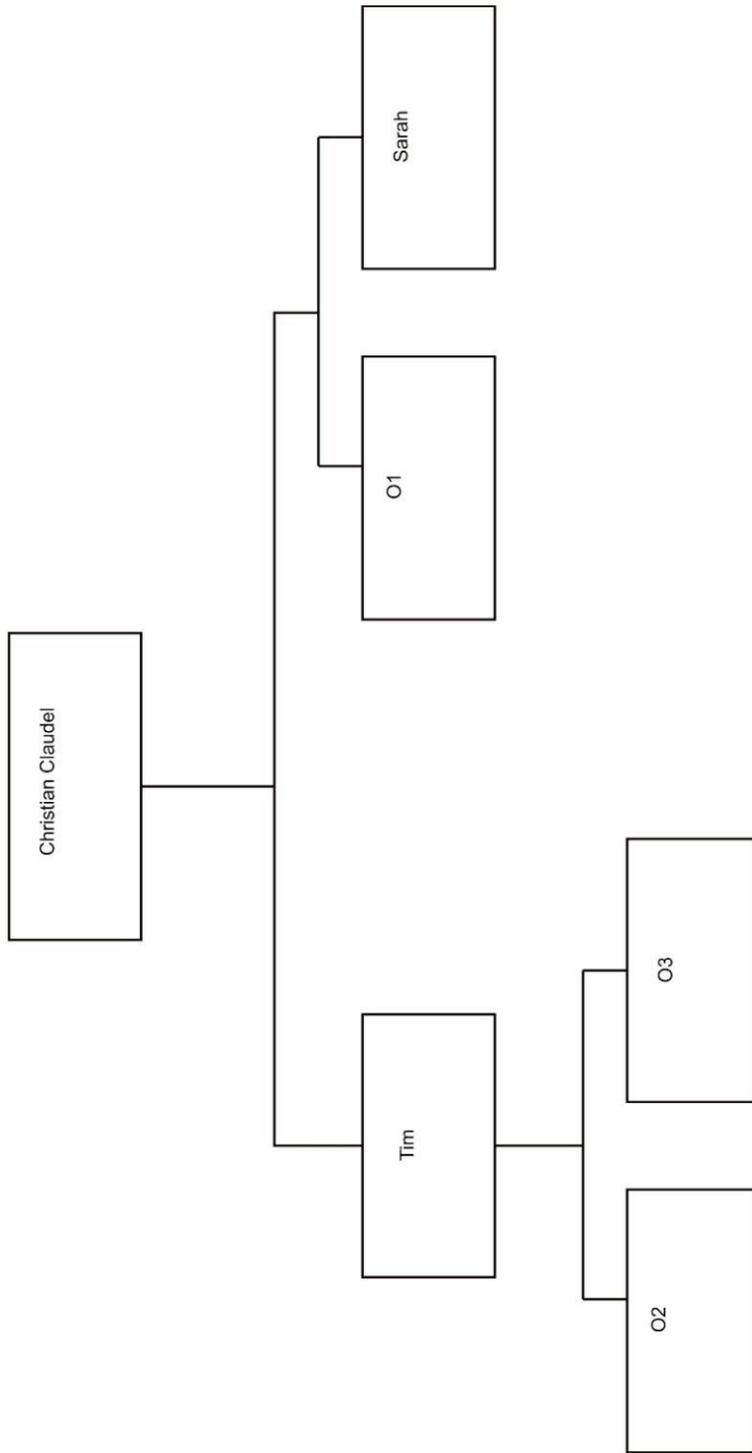
(FEB. 8TH, 2008)



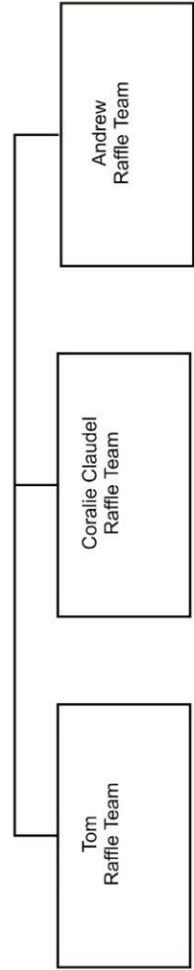
Mobile Century Experiment Organization Chart



Mobile Century Experiment Organization Chart (cont.)



Raffle Team



Positions	Staff Members
Experiment Director (ED)	Alexandre Bayen, Assistant Professor UC Berkeley.
Experiment Supervisor (ES)	Dan Work, PhD Student, CEE UC Berkeley.
Experiment Host (EH)	JD Margulici, Associate Director CCIT.
Site Managers (SM)	Dr. Ali Mortazavi, Senior Dev. Engineer, CCIT.
Human Logistics Officer (HLO)	Marika Benko, Staff Research Associate.
Site Logistics Officer (SLO)	Steve Andrews, CCIT
Nokia Tech. Manager (NTM)	Ryan Herring, GSR, CCIT.
Site Managers (SM)	Quinn Jacobson
	Dr. Ali Mortazavi, Senior Dev. Engineer, CCIT.
Safety Officer (SO)	Marika Benko, Staff Research Associate.
TITLE?	Olli-Pekka Tossavainen, Post Doctoral Student
Assistant Human Logistics Officer (AHLO)	Osama Elhamshary (TITLE?)
Human Logistics Assistants (HLA)	Tia Dodson, CCIT
Sire Logistics Assistant (SLA)	Chris Flens-Batina, UC Berkeley
Site Assistants (SA)	Xavier Litrico, IGREF Research Fellow, Fr.
Head Personnel Manager (HPM)	ASCE Officers.
Phone Operators (PO)	Juan Carlos Herrera, PhD Student, CEE UC Berkeley.
Team Leaders (TL)	Nokia Staff.
	Jean Parks (UCB, Red Team)
	Kristen Parrish (UL, Yellow Team)
	Josh Pilachowski (UL, Red Team)
	Justin Pope (UCB, Orange Team)
	Matt Vaggione (UCB, Yellow Team)
	Megan Smirti (UL, Orange Team)
Team Personnel Assistants (TPA)	Tarek Rabbani (TITLE?)
	Qingfang Wu (TITLE?)
	Arthur Wiedmer (TITLE?)
	Manju Kumar (Backup) (TITLE?)
	(+3 more staff TBD)

Team Traffic Assistants (TTA) Alexander Alshanetsky (UCB, Yellow Team)

Negin Aryaee (UCB, Orange Team)

Timmy Siuaw (UL, Yellow Team)

Emma Strong (UCB, Red Team)

Jason Wexlar (UL, Orange Team)

Anthony Patire (UL, Red Team)

+ 3 more Officers (TBD?)

UCB Queue Dispatchers

(+2 Officers TBD)

Nokia Monitors (NM)

Nokia Staff

TITLE?

Christian Claudel (TITLE?)

TITLES?

Tim (LAST NAME, TITLE?)

Sarah (LAST NAME, TITLE?)

(+ 3 Officers TBD)

Raffle Team

Coralie Claudel (TITLE?)

Tom (LAST NAME, TITLE?)

Andrew (LAST NAME, TITLE?)

OVERVIEW

THURSDAY FEBRUARY 7, 2008.

Cars will begin arriving on the parking lot on Thursday at 9am. By Thursday evening all the cars should be on the lot. Cell-phones will be installed in the cars by Thursday night.

The 99 cars/cell-phones will be numbered and divided into three groups (33 vehicles each): V1 (1-33), V2 (34-66), and V3 (67-99). Each car key will have the number of the vehicle. Each one of the groups will be monitor by one Nokia Monitor. That is, there will be three Nokia Monitors, each of which will be in charge of the same group of 33 vehicles during the entire experiment. Each car will have its number written on the hood of the vehicle.

There are **3 Teams** each corresponding to the **Car Colors (Red, Orange, Yellow)**. Additionally, note that there are **4 Schedules Blocks (A, B, C, D)**. Please note that AM designates a long route, and a PM designates a short route. **All drivers on the road before 1:30 PM will be driving an AM (long) route, and all drivers on the road at 1:30 and after will be driving a PM (short) route.** At the beginning of each shift, all drivers should begin their route by going to the SOUTH end of the loop first.

Driving Schedule Sorted by Schedule Blocks

In Table 1, note that that at any given time, there are 3 Schedule Blocks on the road (e.g. at 3:00, cars with Schedules B, C, and D are on the road). In this way, there is always one Schedule Color on break / at lunch, and there will be 99 cars on the road (33 for each schedule color).

Driving Schedule Sorted by Car Color

In Table 2, we have expanded table 1 to show each individual Schedule Block (A,B,C,D) for each of the 3 Teams / Car Colors. Each Car Color and Schedule contains a group of 11 drivers (e.g. there are 11 drivers with an Orange Car Color AND Schedule C). Therefore, at any given time period, there are 33 drivers from each CAR COLOR, giving a total of 99 cars on the road at any time (i.e. ALL cars are on the road at any given time).

Furthermore, at any given time, there are 33 drives from each SCHEDULE COLOR on the road at any time (e.g. at 10:30, there are 11 Red cars, The table above is useful for figuring out which teams or schedules must be on the road at any given time, and which route (AM or PM) they should be driving. This can be seen by looking at any column (i.e. time).

Table 1 - Driving Schedule sorted by Schedule Color

	10:00	10:30	11:00	11:30	12:00	12:30	1:00	1:30	2:00	2:30	3:00	3:30	4:00	4:30	5:00	5:30	6:00	6:30
<i>Schedule Color:</i>																		
A		BREAK	AM	AM	AM	AM	AM	PM	PM	LUNCH	PM	PM	PM	PM	PM	PM	PM	
B		AM	AM	LUNCH	AM	AM	AM	PM	PM	PM	PM	BREAK	PM	PM	PM	PM	PM	
C		AM	AM	AM	AM	LUNCH	AM	PM	PM	PM	PM	PM	PM	BREAK	BREAK	PM	PM	
D		AM	AM	AM	AM	AM	AM	LUNCH	LUNCH	PM	PM	PM	PM	PM	PM	PM	BREAK	

AM Route = Long Route (before 1:30), PM Route = Short Route (after 1:30)

Table 2 - Driving Schedule sorted by Car Color

	10:00	10:30	11:00	11:30	12:00	12:30	1:00	1:30	2:00	2:30	3:00	3:30	4:00	4:30	5:00	5:30	6:00	6:30
<i>Car Color:</i>																		
Red, A		BREAK	AM	AM	AM	AM	AM	PM	PM	LUNCH	PM	PM	PM	PM	PM	PM	PM	
Red, B		AM	AM	LUNCH	AM	AM	AM	PM	PM	PM	BREAK	PM	PM	PM	PM	PM	PM	
Red, C		AM	AM	AM	AM	LUNCH	AM	PM	PM	PM	PM	PM	PM	BREAK	BREAK	PM	PM	
Red, D		AM	AM	AM	AM	AM	AM	LUNCH	LUNCH	PM	PM	PM	PM	PM	PM	PM	BREAK	
Orange, A		BREAK	AM	AM	AM	AM	AM	PM	PM	LUNCH	PM	PM	PM	PM	PM	PM	PM	
Orange, B		AM	AM	LUNCH	AM	AM	AM	PM	PM	PM	BREAK	PM	PM	PM	PM	PM	PM	
Orange, C		AM	AM	AM	AM	LUNCH	AM	PM	PM	PM	PM	PM	PM	BREAK	BREAK	PM	PM	
Orange, D		AM	AM	AM	AM	AM	AM	LUNCH	LUNCH	PM	PM	PM	PM	PM	PM	PM	BREAK	
Yellow, A		BREAK	AM	AM	AM	AM	AM	PM	PM	LUNCH	PM	PM	PM	PM	PM	PM	PM	
Yellow, B		AM	AM	LUNCH	AM	AM	AM	PM	PM	PM	PM	BREAK	PM	PM	PM	PM	PM	
Yellow, C		AM	AM	AM	AM	LUNCH	AM	PM	PM	PM	PM	PM	PM	BREAK	BREAK	PM	PM	
Yellow, D		AM	AM	AM	AM	AM	AM	LUNCH	LUNCH	PM	PM	PM	PM	PM	PM	PM	BREAK	

AM Route = Long Route (before 1:30), PM Route = Short Route (after 1:30)

Preparation Phase

FRIDAY FEBRUARY 8, 2008.

Location: UC Berkeley (UCB) Lot

7:00am On site staff: Experiment Director (Alex), Experiment Supervisor (Dan), Site Managers (Ali and Marika, and assistants), Site Logistic Officer (Ryan), Human Logistic Officer (Steve), Head Personnel Manager (Juan Carlos), Team Leader (Jean, Justin, Matt), Team Traffic Assistant (Emma, Negin, Alexander), Team Personnel Assistant (Tarek, Arthur, Qingfang).

8:00am TPAs will welcome drivers upon their arrival to the UC Berkeley lot. Drivers will be divided into three teams (Red, Orange, or Yellow) during their training sessions prior to the experiment date. During the check-in on the experiment date, each driver will receive the following:

- A card with the number of the team they belong to. The same team will drive the same loop during the experiment. That is, T1 will drive on loop 1 (in both Phase I and II) and so on. Each team should have about 44 drivers, which accounts for a total of 132 drivers.
- Car key: The matching between drivers and cars will be as follows:
 - o Drivers from T1 (Red) will receive a vehicle from V1.
 - o Drivers from T2 (Orange) will receive a vehicle from V2.
 - o Drivers from T3 (Yellow) will receive a vehicle from V3.

There will be one Bluetooth rubber headset per phone, so drivers will have to share the headset between shifts. At the end of the check-in, drivers will be sent with the corresponding Team Leader.

8:30am Alex welcomes drivers, introduces the team and gives a quick overview of the day.

(Alex to produce script for introduction / welcome)

8:45am The Team Leaders for UCB will give a detailed overview of the experiment to the drivers, which include:

- Introduction of the Team Traffic Assistants (Jean, Justin, and Matt) and the Team Personnel Assistants (Tarek, Arthur, Qingfang).
- Instructions on how to get from (to) UCB lot to (from) the Base Station (BS) at Union

- Landing lot (B1 and B8)
- Safety briefing (B9)

8:50am Ryan, Juan Carlos and the TPAs (Tarek, Qingfang, Arthur) will start driving to the Base Station (BS) at Union Landing lot. Before leaving, Ryan will give specific instructions to the UCB TTAs (Emma, Negin, Alexander) on how to release vehicles from the UCB lot to the BS. Late-show ups will go in the last vehicles departing the UCB lot. The release schedule will be ad hoc, so that 15 cars will be released simultaneously.

9:00am Nokia monitors and phone operators should be ready to go at Union Landing by this time.

9:30am Drivers will start leaving the parking lot at a rate of 3 vehicles per minute if possible. Each minute, a vehicle from T1 will leave first, followed by one from T2, and finally one from T3. The first 33 drivers leaving the parking lot will ride one of the last 33 drivers to register to the BS (because they do not have a car). See the release schedule on Page 10 for details on exact method of releasing cars and extra drivers from UCB parking lot.

If after several briefings some drivers still do not understand how to get to the BS, they will leave the UCB lot at the end and will be escorted to the BS by staff members.

UCB to BS: DEFINE HOW THIS IS GOING TO BE DONE (CHP? FOLLOW THE VEHICLE IN FRONT? DIRECTIONS?).

10:15am Expected arrival time to the BS at the Landing Union lot of the first driver.

Location: Base Station at the Union Landing Lot

4:00am Tents and equipment to be on site, ready for assembly

7:00am On site staff: ??? Team Leader (Josh, Megan, Kristen), Team Traffic Assistant (Anthony, Jason, Timmy)

The tents have to be mounted, generator has to be operational, headquarters and

rest area have to be assembled, and the food place ready. NMs and POs should be operative before 9:30am, when the first vehicle starts moving at the UCB lot.

9:30am Ryan (SLO), the Juan Carlos, and the TPAs (Tarek, Qingfang, Arthur) should have arrived to the BS. Ryan will give specific instructions to the Base Station TTAs (Anthony, Jason, Timmy) on where to park the cars once they have arrived. The TPAs will get in contact with the corresponding NM to see if some drivers get lost. In case any driver gets lost, they have to contact the HPM.

See section release schedule below for protocol on how to release cars from the UCB Parking lot to the Base Station, and how to assign schedules to all cars and drivers.

Each of the Base Station TLs (Josh, Megan, Kristen) will be continuously briefing drivers as they arrive. Once a driver has been briefed, he / she will contact the corresponding TPA (Tarek, Qingfang, or Arthur), which will assign a vehicle to a driver. The TPA, with the help of the NM, will track driving times and the status of the gas tank of the corresponding 33 vehicles during the experiment. A form like the one in Table 3 may be helpful.

Table 3: Form-prototype to be used by the TPAs to keep track of the driving times.

Veh. #	Driver name	Team	Schedule	Start time	End time	Gas	Action	Status
15	John	Red	A	11:02	13:20	Full	No	On the road
10	Paul	Yellow	D	13:08		¼	Refill	On the road
15	Lisa	Orange	B	13:25		¾	No	On the road
...

Release Schedule: Assigning Drivers to Car Colors and Schedules

Each Team Color will consist of 44 drivers (11 drivers for each Schedule A, B, C, and D). 3 Team Colors times 44 drivers per team means **at least 132 drivers total** needed for the experiment. Because Schedule B, C, and D all start driving at 10:30, the first cars to arrive at the Base Station should be assigned to schedules B, C, and D first (i.e. the last 11 drivers for a Team Color to arrive should be assigned to schedule A).

When the experiment cars begin arrive at the Base Station, **11 cars from each Team Color will hold 2 passengers; the remaining 22 cars from a Team Color will have single passengers.** We propose the following method of assigning schedules:

Prior to the Experiment Date (During Training Sessions):

- 44 drivers will be assigned to the Red Team, 44 to the Orange Team, and 44 to the Yellow Team

Leaving the UCB Parking Lot:

In order to maximize the number of drivers that arrive at the Base Station, the cars carrying two passengers should leave the UCB Parking lot first. That is, the following should occur:

- 11 cars from the Red Team holding 2 passengers leave first
- 11 cars from the Orange Team holding 2 passengers leave next
- 11 cars from the Yellow Team holding 2 passengers leave next
- The remaining 22 cars from the Red Team leave next
- The remaining 22 cars from the Orange Team leave next
- The remaining 22 cars from the Yellow Team leave last

Upon arriving at the Base Station at Union Landing:

- All 11 passengers from the Red Team who were driven as a passenger to the Base Station will be assigned a D schedule
- The 11 drivers from the Red Team who drove passengers to the BS should be assigned a B Schedule
- The next 11 drivers from the Red team arriving at the BS should be assigned a C schedule
- The last 11 drivers from the Red Team arriving at the BS should be assigned an A schedule.

Repeat this process for the arriving members of the Orange Team and the Yellow Team.

The reason for this is that if any drivers arrive late, they should be assigned to schedule A since drivers on Schedule A start an hour later than Schedules B, C, or D. The remaining drivers that arrive will be on reserve.

Extra / Backup Drivers

After the initial 132 drivers have been equally split amongst the 3 Teams Colors, the remaining extra drivers should also be split evenly amongst the 3 Team Colors. **These extra drivers will not be assigned a schedule block, and must be ready to replace any driver on their Team Color at any schedule block.** The extra drivers should be subject to the following rules:

- Extra drivers must not drive more than 3 hours without an hour break following
- Extra drivers should not drive to Union Landing from UCB; i.e. they must be passengers.
- Extra drivers should attend all briefings for their Team Color so that they are ready to drive if they must substitute another driver during any schedule block, including in the middle of a shift.
- Extra drivers are allowed to replace drivers who are unable to begin their next shift, assuming that a Team Leader is notified and approved the substitution before the shift begins

The purpose of extra drivers is to relieve drivers who arrive late and must take a break before their next shift. This will help reduce delays in the schedule and prevent over-exhaustion from other drivers. They will receive the same compensation as other drivers for their participation on the experiment date.

Phase I

Location: Base Station at the Union Landing Lot

10:30am Expected time when first drivers are sent to the freeway.

11:00am All the 99 vehicles should be on the road. Drivers have to follow instructions given by the Base Union landing TTAs (Anthony, Jason, Timmy). The rest of the team (Alex, Dan, Steve, Jean, Kristen, Josh, Justin, Matt, and Megan) should arrive to the BS no later than 11:00am.

Start/end of a shift:

Each Team Leader (either UCB or UL) will brief the sub team (11 drivers) that is resting just before they start a new driving shift. The briefing should follow B3.

After the end of a shift, if a driver has missed the off-ramp that leads to the BS (NM), he/she will be contacted by the corresponding PO.

At their arrival to the BS, TTAs (Anthony, Emma, Jason, Negin, Timmy, or Alexander) will instruct drivers where to park the car. Once the car is parked, TTAs will check gas tank status and will determine whether or not a refill is needed before entering the freeway. The TPAs (Arthur, Qingfang, or Tarek) will keep record of the new driver start time and will show the driver how to open the gas tank.

Juan Carlos will coordinate with the Team Leaders to make sure that the resting drivers are ready to go once the driver comes back to the BS.

Phase II

Location: Base Station at the Union Landing Lot (?)

1:30pm Phase II begins. Drivers will be instructed during their last briefing to change loops at 1:30pm. NM will check that all the drivers start doing the short loops corresponding to the Phase II. Drivers who do not start driving the short loop will be contacted by the PO.

1:50pm All the drivers on the road should be driving their corresponding short loop by this time.

The start/end of shifts will work in the same way as in Phase I.

Conclusion Phase

Location: Base Station at the Union Landing Lot (?)

- 5:30pm** All **A, B, C** schedule drivers should be finishing their last shift. All teams will rest at the BS until 6:30pm, where they will be debriefed and released back to the UCB lot.
- 6:30pm** **A schedule** drivers will pair with **D** drivers up from the base station to return to UCB lot, and **D** drivers are expected to drive on the way back. **B** and **C** drivers will also be told to return to the BS. At the BS they will check with the TPAs (Tarek, Qingfang, Arthur) their departure time to the UCB lot.
- 7:30pm** Expected arrival time of drivers to the UCB lot. The Human Logistics Team (Chris, James, Steve) will remain at the Union Landing lot until the last vehicles return to UCB before they return to UCB themselves.

Ryan and the UCB TTAs (Emma, Negin, Alexander) have to be back at the UCB lot by 6:00pm, in order to coordinate the arrival of the drivers.

The rest of the team (Alex, Dan, Josh, Megan, Kristen, Anthony, Jason, Timmy) will drive back to the UCB lot with the last drivers leaving the BS.

Site Managers (Ali, Marika, + assistant), Olli-Pekka, JC, TPAs, NMs and POs will stay at the BS after the last driver leaves. The tent will start to be dismantled. However, note that the NMs will track their corresponding vehicles until they are parked at the UCB lot. Once the last vehicle is parked at the UCB lot, Olli-Pekka, JC, TPAs, NMs and POs can go back to UCB lot. The Site Managers will stay until everything has been dismantled, ideally by 9 PM.

Location: UC Berkeley (UCB) Lot

- 6:00pm** On-site staff: Site Managers (Ali, Marika, and assistant), Ryan, UCB Team Traffic Assistants (Emma, Negin, Alexander), Nokia Staff.
- 7:00pm** Earliest possible show up time for first vehicle arriving to UCB lot. Emma, Negin, Alexander will instruct drivers on where to park the car, following Ryan's directions.

Site assistants will be collecting car keys as vehicles arrive. Once the driver has returned the car key, he/she will receive a certificate of completion of the experiment (necessary for processing paychecks) and the free Bluetooth headset.

- 7:45pm** Expected arrival time of part of the rest of the team (Alex, Dan, BS Team Leaders (Josh, Megan, Kristen), BS Team Traffic Assistants (Anthony, Jason, Timmy)) the UCB lot.

- 8:00pm** The lottery drawing for the free N95 phones (4) will be done at 8pm (or earlier if all drivers

and team members have already arrived to the UCB lot). After this, participants are free to go!

Nokia staff will be in charge of removing the N95 phones from the vehicles and taking them back to Palo Alto.

SATURDAY FEBRUARY 9, 2008.

Vehicles will be picked up by the rent-a-car company staff.

POSITIONS

Experiment Supervisor (Dan Work)

Responsibilities Oversee all aspects of experiment

- Equipment**
- Walkie-talkie
 - Private desk space
 - Phone with pre-programmed numbers of all management staff.
-

- Support Staff**
- Support Staff (Patrick Saint-Pierre)
 - Site Logistics Officer
 - Site Managers
 - Human Logistics Officer
 - Nokia
 - Safety Officer
-

Experiment Host (JD Margulici)

Responsibilities Host all Caltrans, Nokia, Berkeley, CCIT, etc. VIPs.

Equipment To be determined by

Support Staff No dedicated staff

Support Staff Experiment Supervisor (P. St-Pierre)

Responsibilities Assist Experiment Supervisor with tasks as needed for emergency response. Will be on site with no pre-assigned function ready for act if needed.

Equipment	Regular Phones
Support Staff	None

Site Logistics Officer (Ryan Herring, Xavier Litrico)

- Responsibilities** Manage team traffic assistants
- Training before experiment
 - Briefing at Center Street Garage at 7:00
 - Briefing at Union Landing at 9:30 or 11:00 depending on group
- Oversee all vehicular traffic at Union Landing site
- Instruct drivers where to park and what route to take in the parking lot
 - Ensure vehicles park in the correct zone and that team traffic assistants are checking gas levels in vehicles before leaving

-
- Equipment** Numbers/markings for each vehicle
Walkie-talkie for communication with Team Traffic Assistants only
Signs for Union Landing site:
1. From the freeway to the parking lot
 - Right onto Alvarado Niles
 - Right onto Union Landing
 - Straight through intersection (at Marcia Drive)
 - Left into the parking lot
 2. Signs for zones 1, 2 and 3 at their correct locations on the parking lot
 3. From parking lot to freeway
 - Right onto Union Landing
 - Straight through intersection (at Marcia Drive)
 - Left onto Alvarado Niles
 - Left onto freeway
 4. Signs for Center Street garage site
 - Towards the exit of the lot
 - Right onto Allston Way
 - Left at MLK

-
- Support Staff** Site Logistics Assistant (1) (Xavier Litrico)
- This person will have the same detailed outline and timeline as the Site Logistics Officer, will be prepared to handle the same duties, and will generally assist the Site Logistics Officer as needed throughout the day)

Team traffic assistants (6) (see detailed outline for this position)

-
- Preparation Work** Prepare training material for Team Traffic Assistants for briefing one week prior to experiment
- Traffic Assistants will be divided into 2 groups (Y and Z). Create a detailed timeline for each group.

-
- The Y group consists of 4 team traffic assistants
 - Arrive to Union Landing early with Ryan to handle initial arrival of drivers
 - Stay at Union Landing until all drivers have completed the experiment and are on their way back to Berkeley
 - The Z group consists of 2 team traffic assistants and the Site Logistics Assistant (Xavier Litrico)
 - Responsible for initial release of drivers from Center Street garage
 - Drive to Union Landing after all drivers have left Berkeley
 - Arrive back in Berkeley early with Ryan to handle returning drivers
 - Instructions for Center Street garage logistics (morning)
 - Details on the timing of the release of the vehicles (1 vehicle from team 1, then 1 vehicle from team 2, then 1 vehicle from team 3, etc.)
 - Directions to Union Landing including group Y leaves at 9:00am and group Z leaves after all drivers have left
 - Instructions for Union Landing logistics
 - Make sure they understand the site logistics map (see section 2a. organization of the site), including the routes to and from the parking zones they are responsible for
 - Details on timing
 - Drivers start to arrive at 10:00am, be prepared for high volume initially that will require them to be quick and decisive in getting everyone to an appropriate parking space
 - Drivers will begin leaving the parking lot while others are still arriving. Make sure they understand how to manage the traffic flow in and out of their parking zone, paying particular attention to the timing of when to release vehicles from their zone based on the Site Logistics Officer's (Ryan) orders
 - Explain the timing of the end of the day. Group Z will leave the site at 5:30pm. Group Y will leave at 7:30pm ensuring that all drivers have already left to return to Berkeley.
 - Instruction for Center Street garage (evening)
 - Go over map of the parking area and give instructions for where drivers will park.
 - Instructions on timing, group Z will be present when drivers begin returning to the lot.

Timelines
7:00am *(Arrive at Center Street Garage)*

- Brief Group Z Team Traffic Assistants

(Supervisor)

- Instruct Group Y Team Traffic Assistants to leave for Union Landing by 8:00am
- 7:30am** Drive to Union Landing with Group Y Team Traffic Assistants
- 8:30am** Arrive to Union Landing, make final preparations for arrival of drivers
- 9:00am** Brief Group Y Team Traffic Assistants
- Go over site logistics map in detail and explain the routes in and out of their parking zone.
 - Explain the details on when drivers will be arriving and that there will be an initial surge of drivers at the beginning of the experiment.
 - Give instructions for when it is appropriate to take breaks and to ask Ryan if they need to leave their parking zone.
- 10:00am** *(Drivers start arriving)*
- Oversee parking operations by Team Traffic Assistants
- 11:00am** *(Group Z Team Traffic Assistants arrive at Union Landing)*
- Brief group Z team traffic assistants
- Go over site logistics map in detail and explain the routes in and out of their parking zone.
 - Give instructions for when it is appropriate to take breaks and to ask Ryan if they need to leave their parking zone.
- 11:00am to 5:30pm** Oversee traffic operations
- 5:30pm to 6:30pm** Drive back to Center Street garage
- 6:30pm** Direct vehicles back into parking spaces
- Brief group Z team traffic assistants on what to do when drivers return
 - Instructions for how to direct drivers to the proper parking

location

- Oversee parking of returning drivers and ensure they return keys/phones

Timeline (Assistant)	7:00am	<i>(Arrive at Center Street Garage)</i>
	7:30am	In charge of Group Z Team Traffic Assistants
	9:30am	Oversee exiting vehicles from Center Street garage:
		<ul style="list-style-type: none"> • Posted at exit of garage giving visual signals to drivers for when it is okay to exit
	10:30am	Drive to Union Landing
	11:00am	<i>(Arrive at Union Landing)</i>
	to 5:30pm	Assist Site Logistics Officer with traffic flow logistics
	5:30pm	Take over site logistics for Union Landing
	to 7:30pm	<ul style="list-style-type: none"> • Make sure drivers come back to Union Landing one last time before heading back to Berkeley
	7:30pm	Drive back to Berkeley
	to 9:00pm	<ul style="list-style-type: none"> • Assist with final check out procedures

Site Managers (Ali Mortazavi & Marika Benko)

Responsibilities (Ali)

- Supervising the site staff
- Union City site manager
- Coordinating Tent set-up
- Staff seating layout
- Participants seating layout

Parking markup

Coordinate tent dismantling

Manage site clean-up

Responsibilities (Marika) Assisting Ali in managing the sites.

Berkeley site manager

Preparing the Berkeley site

Providing breakfast

Facilitating required equipment in the parking lot

Supervising two assistant in the site

Responsibilities (Assistants) Food

Setting up the site

Site clean-up

Parking mark-up

Set-up tables and chairs

Numbering cars

- Equipment**
- Color copier and printer (1) 1/08
 - Extension cords (12-Costco) 1/08
 - Power strips (20-Costco) 1/08
 - Tool kits (1-Ace Hardware) 1/08
 - Duct tape (3 rolls) 1/08
 - Generator (supplied by tent company) *Finalized 12/07
 - Pens, pencils(10 boxes-CCIT will order 1/08)
 - Paper(3 packages-CCIT will order 1/08)
 - Masking tape(3-CCIT will order 1/08)
 - File folders (20-CCIT will order 1/08)
 - Clip board(30--CCIT will order 1/08)
 - Scissors (5-CCIT will order 1/08)
 - Recorders for vehicle tracking (10-CCIT will order 1/08)
 - 2 8 gallon trash cans; One roll of industrial trash bags

- 2 waste paper baskets
- Breakfast: Coffee-Starbucks in shopping center *Order in 1/08
- Lunch: ToGos sandwiches in shopping center *Order in 1/08
- Dinner: Snacks from Costco—Pretzels, fruit, bread, peanut butter and jelly, cookies *Buy in 1/08
- Utensils: 200 plates and forks; cups for coffee *Order in 1/08
- Coffee: All day from Starbucks or we Make? *Order in 1/08
- Popcorn machine from tent rental supply company *Finalized 12/12
- Water, Soda (200-from Costco) *Order in 1/08
- Umbrellas 5-foldup *Order in 1/08
- Heater (1 from tent rental supply company) *Finalized 12/07
- Tent (1 from tent rental supply company-see site plan) *Finalized 12/07
- Tables (27 – 9 for staff, 2 for food, 1 check in, 15 for participants-from tent rental supply company-see site plan) *Finalized 12/07
We are working on fitting the tables into the participants area. The tent company will help us on that.
- Rope: (50 feet Ace Hardware)
- Traffic cones: (23 \$230—Ali has website)
- Chairs (90 from tent rental supply company-see Ali’s site plan) *Finalized 12/07
- Bull horns (2-CCIT will order 1/08)
- Safety vests (20-CCIT will order 1/08)
- Name tags (200-CCIT will order 1/08)
- One 6 foot table 2 chairs for Berkeley parking lot

Support Staff 2 assistants for Union City site, 2 assistants for Berkeley site

Preparation Work

- Finalize the mall contract
- Obtaining formal estimate for rental car
- Rental car paper work process Mapping out the rest rooms on 12/13 during a site visit and photo taking opportunity for training materials
- Select or hire the supporting staff
- Briefing the staff on the tasks
- Finalize the table sizes and locations in the participants area with the tent company
- Contract with the tent company
- Purchase the required items for the day of experiment: water, traffic cones...

Timeline (Before D-Day)

- 02/06/2008 11: 00pm** 100 spaces will be set aside (via caution tape/roped off area) for Berkeley
- 02/07/2008 9:00am**
- Enterprise will begin delivering cars
 - Marika and other Berkeley staff will coordinate with Enterprise to number cars / keys starting at 9 AM, continuing throughout day
 - **Numbering system:**

- Each key, phone, and car will have the same number, respectively
- Nokia will have numbered the phones before 2/07 at 4:00 PM, when they come to Berkeley

02/07/2008 4:00pm Nokia will have pre-numbered phones ready in Berkeley

02/07/2008 5:00pm Berkeley staff and Nokia staff will begin to install phones in cars, each car # is the same as its phone #

02/07/2008 6:00pm Nokia staff will check software

**Timeline
(D-Day)**

UC Berkeley Site

6:00am Marika and two helpers will set up coffee and pastries; one folding table; three folding chairs. One helper will have picked up table and chairs from CCIT and brought it over. Marika will have picked up coffee and pastries from Starbucks.

7:30am *(Participants arrive to the location)*

Marika and 4 helpers will check participants in; check drivers licenses again, make sure we have a copy of consent form, answer questions, hand out maps, etc.

8:30am Nokia staff will arrive to check software again.

Drivers will be briefed, directed out of lot, etc.

10:00am Take down will begin (table, food leftovers, and chairs)

11:00am Parking lot will be cleared

5:00pm Marika and two helpers will arrive at Berkeley parking lot to make sure rented area is cordoned off.

Nokia will arrive to collect phones.

6:00pm Participants may start to arrive; Marika and Nokia will collect phones and keys

Union Landing Site

- 6:00 am** Staff on site (Ali and 2 site staff)
Coordination with the tent crew (Ali)
- 6:30am** Mark up the parking area including the Response Team Parking. (Ali and 2 staff). Helping Nokia to set up their equipment (Ali and 2 staff)
- 7:00am** Checking the tent staging (Ali)
- 7:30am** Inside tent set-up includes setting up the tables, roping up the briefing and training area, setting up the staff area, placing the office supplies on the tables. (Ali and 2 staff)
- 8:00am** Wiring (Ali, 1 site staff and tent crew)
- 8:15am** Water , coffee and snacks set up(2 staff)
- 9:30am** Arrangements with the catering (Marika)
- 9:45am** Finalizing the staff staging (Ali, Marika, 2 site staff)
- 11:00am** Marika on site
Marika handles the coffee
- Food is served to drivers based on their driving schedule (Ali, Marika, 2 staff)
- 2:00pm** Last food is served (Ali, Marika, 2 staff)
- 4:00pm** Marika leaves the site to Berkeley parking lot
- 5:00pm** Cleaning starts (2 site staff, we need two more people)
- 6:00pm** Clean the office supplies and wires (Ali and staff)

7:00pm Cleaning up the cones (2 site staff)

8:30pm Site will be fully cleaned / cleared (2 staff + 2 extra traffic assistants)

8:45pm Dispose the trash bags into dumpsters

Timeline	02/09/2008	Marika will meet with Enterprise staff and give them the keys.
(Day after)	9:00am	Enterprise will collect the cars throughout the day on Saturday, ending at 5:00 PM.

Briefings Briefing the assistants a week before D-Day

Human Logistics Officer (Steve Andrews)

Responsibilities Manage all team leaders (6)
 Handle all training related questions that cannot be addressed by team leaders
 Manage any driver questions not dealt with by team leaders
 Handle any “problem drivers”
 Add more here...

Equipment Walkie-talkies to communicate with team leaders etc.
 Personal Vehicle to get from UC Berkeley to Union Landing and back
 Will be out working in the field – Waterproof gear.
 Folder with all training material, list of contacts
 Add more here...

Support Staff Assistant (1)
 Team Leaders (6)

Preparation Work

- Make sure the Team leads are fully staffed by Jan 15
- Hold a team lead meeting during first week of class
 - Opportunity for team leads to meet each other
 - Explain Team lead responsibilities throughout the day of the experiment
 - Have scripts prepared for Hearst briefing
 - Safety
 - Overview for the day
 - Directions to Union Landing
- First briefing at Union Landing

-
- Explain looping behavior again
 - Explain the site layout and locations of facilities
 - Explain protocol for requesting unexpected breaks
 - Explain protocol for checking in and out with the team personnel manager
 - Gas protocol
 - Subsequent briefings
 - Explain the looping again (repeat ramps)
 - All Team Lead Training material must be completed before Jan 15
 - Insist the team leads practice scripts and know information.
 - Have communication devices ready to explain how to use them
 - Explain the chain of command and how to delegate tasks
 - Problems with car -> team traffic
 - Need more break time -> team personnel
 - Want to quit -> team personnel
 - Problem with phone -> Nokia Tech
 - Etc
 - Contact team leads via email week of experiment to answer any additional questions
 - Send reminder email the evening before

Timeline

7:00am Show up at UC Berkeley lot before 7:00am

7:30am Hold staff meeting with team leads to go over procedure again

7:30am to 8:45am Float time, assist where needed

8:45am to 9:00am Team leads train teams. Monitor progress, address problems
 Delegate additional UC Berkeley site responsibilities to assistant
(Must Leave at 9:00 am for Union Landing)

10:00am *Arrive no later than 10:00 (arrival of first cars)*
 3 team leads will already be on site (one will be in charge of all drivers until human logistics officer arrives)

Prepare to respond to drivers as they arrive (run 15 min briefings until last cars arrive)

- 10:00am** Supervise all trainings until all drivers arrive
- to 11:30am** Work in team leads from UC Berkeley as they arrive

- 12:00pm** rest 3 team leads for 1 hour for lunch

- 1:00pm** rest the other 3 team leads

- 1:00pm** address Team lead problems as they arise, spot check drivers to make sure
- to 4:30pm** there are no complaints

- 4:30pm** send 3 team leads back to UC Berkeley to await return of drivers

- 5:00pm** stay until last drivers return, or all drivers are accounted for.
- to 7:00pm**

- 7:00pm** return to UC Berkeley

Briefings: B9b Safety

Human Logistics Assistant (Chris Flens-Batina, James Lew)

Responsibilities Act as a centralized point of information for all drivers
 Keep the drivers happy
 Assist the Human Logistics Officer in managing the drivers and the team leaders

Equipment Desk in the tent with laptop
 Copies of all training literature
 Copies of site layout, etc.

Support Staff None

- Preparation Work**
- Scenario planning, predict common driver questions and figure out how to address them
 - Meet with human logistics officer for all training sessions
 - Need to develop a protocol for drivers who want to go home or quit early before January staff training session

Timeline	7:00am	Show up at UC Berkeley lot before 7:00am
	7:30am	Attend staff meeting with team leads to go over procedure again
	7:30am to 8:45am	Float time, assist where needed
	8:45am	<i>Must Leave at 9:00 am for Union Landing</i>
	to 10:30*	Delegate additional UC Berkeley site responsibilities to assistant <i>*10:30 or time of last driver leaving</i>
	11:30am	Arrive at Union Landing
	11:30am to 6:30pm	Run an information center for the drivers, respond to complaints
	5:00pm	Stay until last drivers return, or all drivers are accounted for
	to 7:00pm	
	7:00pm	Return to UC Berkeley

Briefings: No briefings given but have all of the briefings on hand in case a driver has a question

Safety Officer (Olli-Pekka Tssavainen)

Responsibilities Oversee all safety operations
 Develop protocols for appropriate responses to “likely” events
 Coordinate response efforts with police/ambulance/tow-truck
 Manage Response Team, Phone Operators, and Head Personnel Manager
 More

Equipment

- Walkie-talkies
- Empty 2-gallon gas canister (filled only if necessary for safety)
- Working space near phones and monitors
- Laptop + Charger

- Mobile phone + Charger ; Pre-programmed contacts for staff, police, ambulance, tow-truck
- Folder with all paper documents; safety protocols, contact information
- Transportation to Union Landing (own vehicle)

Support Staff Response Team
 Phone Operators (3-6)
 Head Personnel Manager

- Preparation Work**
- Make sure Response Team and Phone Operators are fully staffed by Jan 15
 - Ask Marika where the broken rental cars should be towed
 - In which hospital possibly injured/ill is taken
 - Prepare / Distribute / Explain Emergency Responses document to all staff
 - “In case of emergency” document to drivers
 - Arrange a meeting / email / phone call with Phone Operators, Head Personnel Manager, and Team Personnel Assistants
 - Chain of command
 - Explain Emergency Responses
 - Additional questions (Check by email right before the experiment)

Timeline	5:30am to 6:00am	Show up at Union Landing (5:30 am) 6:00 am and setup all communications/devices
	7:00am	hold a staff meeting to see that everyone has arrived and answer additional questions, extra training, breakfast
	8:30am	ready to monitor first vehicles leaving UC Berkeley (9:30 am)
	9:30am	start safety protocols / monitoring of vehicles coming from UC Berkeley and be prepared to response accidents/lost vehicles
	6:00pm	Maintain /supervise safety communications when Response Team etc. is moving out from tent
	6:00pm to 10:00pm	monitor until the last vehicle is reported as arrived to UC Berkeley Release Response Team, Phone Operators and HPM

Nokia Tech Manager (Quinn)

Responsibilities	Manage Nokia Staff and Technology: Staff in charge of installing (removing) the cell-phones in (from) the cars Respond to any cell-phone-related problem (cell-phone is not working, vehicles cannot be tracked, etc)
Equipment	Mobile phone + Charger ; Pre-programmed contacts for staff Laptop
Support Staff	Nokia staff
Preparation Work	Make sure the system is working properly (server tests, etc). 20-car experiment Meet with site managers to define the needs of the Nokia staff at both the UCB lot and the Base Station. Get all the cell-phones installed in cars by Thursday night and make sure they are working properly.
Timeline	<p>7:00am Show up with the Nokia staff at the UCB lot in case any problem arises with the cell-phones</p> <p>9:00am Drive to the BS at Union Landing with part of the Nokia Staff. The rest of the Nokia staff will stay at the UCB lot in case last-minute problem arises. They will drive to the Base Station once the last vehicle is released.</p> <p>6:00pm Drive back to the UCB lot with part of the Nokia Staff.</p> <p>7:00pm start removal of cell-phones as vehicles arrive to the UCB lot.</p> <p>8:00pm Stay for the four N95 lottery.</p> <p>8:30pm take all the cell-phones back to Palo Alto.</p>

Team Leaders (Jean Parks, Kristen Parrish, Josh Pilachowski, Justin Pope, Matt Vaggione, Megan Smirti)

Responsibilities Manage a team of drivers (T1, T2, or T3)
 On-site training of drivers:

- Brief drivers on details of experiment at UCB Lot
- Brief drivers before beginning loops at Base Station
- Brief drivers throughout day regarding next steps in experiment

Main point of contact for drivers

Equipment Mobile phone + Charger: Pre-programmed contacts for staff (list of contacts)
 Will be out working in the field – Waterproof gear.
 Folder with all training material (briefings) and paper documents

Support Staff Team traffic assistants, team personnel assistant, Nokia monitor

Preparation Work Hold a meeting with Dan and the other Team Leaders during first week of class

- Opportunity for Team Leaders to know each other and about their responsibilities throughout the day of the experiment

Have read all the briefings they are supposed to give to drivers by the last week of January.
 Contact Dan if questions arise or something is not clear.
 The week of the experiment: Hold a meeting with the Dan (+ Assistant), Team Traffic Assistant and Personal Traffic Assistant of the corresponding team:

- Opportunity to know each other
- Make sure everyone knows their own responsibilities during the experiment.
- Explain the chain of command and how to delegate tasks
 - Problems with car → team traffic
 - Need more break time → team personnel
 - Want to quit → team personnel
 - Problem with phone → Nokia Tech

Timeline

Team Leaders Starting at UCB Lot (TL_UCB)

7:00am Show up at UC Berkeley lot before 7:00am

7:30am Hold staff meeting with Dan to go over procedure again (have breakfast)

8:45am Train teams.

to 9:15am

9:30am Help with the release of vehicles
 Leave the UCB lot once all drivers have left to the Base Station.

11:00am Arrive to Base Station. Rest (and lunch) until 1pm.

1:00pm to 4:00pm Brief drivers before they start driving, address any problem that may arise, etc.

4:30pm Go back to UCB lot to await return of drivers. Make sure drivers return car key, and that they get the certificate of completion of the experiment and the free Bluetooth headset.

Team Leaders Starting at Base Station (TL_BS):

7:00am Show up at the Base Station at 7:00am, have breakfast.

8:00am Hold staff meeting with Human Logistic Assistant to go over procedure again

8:30am meet the Team Traffic Assistant and the Team Personnel Assistant to make sure everything is set up.

10:00am Brief drivers before they start driving, address any problem that may arise, etc.

1:00pm to 2:00pm Lunch time.

2:00pm to 4:00pm Rest time, but ready to help if needed.

4:00pm to 6:30pm Brief drivers before they start driving, address any problem that may arise, etc. Stay at the Base Station until last drivers return, or all drivers are accounted for.

7:00pm Return to UCB lot.

Briefings: Briefing 1

Users: Team Leaders

To: Drivers

Time: 8:45 am, UCB lot

Description: Detailed explanation with pictures:
Shattuck → MLK → CA24W → I880S → Exit 23, Alvarado-Niles Rd.

Support Mat'l: One-pager with Google map information (UCB to BS) for drivers to use it on-board.

Briefing 2

Users: Team Leaders

To: Drivers

Time: 8:45 am, UCB lot, before a new shift starts at the BS.

Description: Description of the loops. Each team will do the same loop during each phase. Therefore, the Team Leaders only need to explain one loop

during each briefing.

Each team leader will receive 2 scripts, one for each phase. Each script will have pictures showing how to turn around at both ends of each loop.

Support Mat'l: One-pager for the drivers -with the name of the ramps they have to use- to look at it on-board.

Briefing 3

Users: Team Leaders

To: Drivers

Time: Before a new shift starts at the BS.

Description: Information provided to drivers every time a new shift begins. Each team leader will brief the 11 drivers that are resting just before they start a new driving shift. The briefing includes:

- How to get from (to) the BS to (from) the freeway, in case it is the first briefing.
- Description of the loop drivers will be doing and the schedule (start/end time).
- Instructions on how to refill the gas tank if needed (where to refill, how to operate the gas pump, and payment method).
- How to get from the BS to the UCB lot, in case it is the last briefing. For those drivers that have to come back to pick other drivers up, this has to be said during their last briefing.

JC will coordinate with the team leaders to make sure that resting drivers are ready to go once the driver comes back to the BS.

Support Mat'l: Directions to go from (to) the BS to (from) the freeway
Gas refill instructions
Protocol for requesting unexpected breaks

Briefing 4

Users: Team Leaders

To: Drivers

Time: At the beginning of a break, BS

Description: Information on what drivers are allowed to do during their breaks. Each break lasts 1 hour. At the beginning of the break, the team leader will tell the drivers the start time of the next shift. During the break drivers are expected to stay at the base station, without interfering researchers/staff work. Drivers can bring book/papers/magazines to read or portable games to play, use their own cell-phone to talk, listen

music (with headphones). Drivers can bring their laptop at their own risk, but they are not allowed to use any of the outlets available in the tent (there won't be internet access at the BS). If a driver wants to go for a walk, he/she needs to let his/her team leader know.

Briefing 5

Users: Team Leaders
To: Drivers
Time: At the last briefing, BS
Description: Script document explaining how to go from BS to UC Berkeley.
Support Mat'l: One-pager with Google map information (UCB to BS) for drivers to use it on-board.

Briefing B9a (Safety)

Users: Team Leaders
To: Drivers
Time: 8:45 am, UCB lot, before a new shift starts at the BS.
Description: What a driver has to do in case of emergency.
 In case any problem arises on the road, drivers will contact the corresponding Phone Operator using the N95 phone. This briefing will contain instructions on how to use the phone to this end. The Phone Operator will communicate directly with the Safety Officer, which will decide the action to take.

Team Personnel Assistants (Tarek Rabbani, Arthur Wiedmer, Qingfang Wu, 3 additional staff TBD)

Responsibilities Maintain spreadsheet showing status of each driver:

- Driving, resting, lost, etc.
- Which car the driver is using
- When car was last refueled (info from team traffic assistant)
- Total hours driven by each driver

Check in and checkout of each driver to the experiment site
 Register replacement driver if/when needed
 Ensure communication with the Emergency Response Team.

Track late drivers

Equipment

- Table space at the entrance of each briefing room.
- Laptop (x4)

-
- Walkie-talkie (x3 we only need them at the union site.)
 - Paper and pencils just in case...
 - Color stickers to the color of each team (-> easier for driver/staff to remember.)
 - Access to a printer might be a plus. (To be able to write the names and times and display them in an easy way for the drivers.)
-

Support Staff None

**Special
Procedures**

Keeping track of drivers and vehicles

The generation of spreadsheet can be done from a single text file with one name per line. Arthur assumes the cars will be attributed for a specific team only and that there will not be exchange of cars between teams. Each team corresponds to 1 loop.

The following "fields" will appear on the driver worksheet.

- Number ID
- Name
- Shift Team (A, B, C, D)
- Current Car ID
- Status
- Beginning time of the current shift (driving or rest)
- Expected time of return
- Total driving time (in minutes)

Another worksheet should have a list containing the schedule for refueling:

- Car ID
 - Expected time for refueling
 - Refueling done?
-

Check in Procedure at UCB

According to the current outline, drivers should arrive between 7:30 and 8:00 to check-in. We should prepare to begin check-in earlier to do this as painless as possible.

- 1) They should be directed to their team personnel assistant (TPA) (Berkeley Team). Each team personnel assistant should have an excel file with all the names to be able to guide someone mislead at first.
 - 2) Once in front of the right TPA, the TPA checks they have their US driving licence. They also have to check no one is under influence. In case of a problem refer to team manager.
 - 3) They are assigned an ID number, a car/cell-phone ID number (for those who will drive at first). They are given the keys to the cars and they should go
-

listen to the PI speech.

- 4) Remaining drivers that haven't checked-in by 8:30 should be quick checked-in (step 1) and 2)) asap when they come in and finish check-in after the last PI speech. That is between 9:30 and 10:00.

Drivers way off schedule (arriving after the briefings) need to be either turned away or we still can have them as alternate drivers for later. Anyway last check-in should not happen after 10:00.

The excel-file containing should be sent by email to the Union team so they can prepare as soon as possible.

-
- | | |
|-------------------------|---|
| Preparation Work | <ul style="list-style-type: none"> • Excel spreadsheet formation before the d-day. Need only a 40 minutes meeting with all the team to explain quickly how it will work. NEED TO BE PLANIFIED in the officers training (Arthur) • Quick reference guide of the problems the team personnel assistants could be encountering and who should be contacted in each case. I'm working on it (Arthur/Dan) • There should be a team meeting so that every officer in each team (Red, Yellow, Orange) know each other and can see physically who they should be talking to for each problem. Officers training... • Collect final version of drivers list with their cell phone. And finalize spreadsheet. (Arthur/Marika/Chris) • Get kitchen timers. • Print car numbers on a token/cardboard/paper. |
|-------------------------|---|

Timeline

The TPAs will split into two groups: TPA1 and TPA2

TPA1 consists of Qingfang, Tarek, and 1 additional staff

TPA2 consists of Arthur and 2 additional staff

Team Personnel Assistant (Berkeley Lot)

- 6:50am** Arrive at Center Street parking lot
Installation on the parking lot. Setting up equipment.
- 7:00am** Ready.
- 7:30am** Check-in procedure.

to 8:30am

8:30am Quick check-in procedure for late drivers.

to 8:50am

8:50am All but 1 TPA 2 leave the parking lot for Union Landing; 1 TPA2 remains for late checking.

9:30am Finalizing check-in procedure for late-comers

to 9:45 am

9:45am Last check-ins possible.

to 9:55am

9:55am Send the file

9:55am Packing

to 10:15am

10:30am Leave parking lot with clean-up teams.

11:00am Arrival on the Union Landing site.

11:30am Follow Turns Table of TPA "Tracking Duty" below

to 8:30pm

Team Personnel Assistant (Union Landing)

9:30am Expected arrival time at BS of TPA1, Arthur, +1 TPA2

9:30am Communication established with Nokia Monitors.

to 9:40am

9:45am Start tracking the drivers

Conclusion Phase

~5:00pm 1 TPA will leave with Ryan and the TTAs to be ready for check-out the first drivers coming back to the parking lot and getting the car keys back.

A second one will follow with the first drivers to go.

Arthur + 3 will keep tracking the drivers/ insuring the link with the NM. They will help the Site Managers ASAP.

Table of TPA "Tracking Duty" Turns																		
	10:00	10:30	11:00	11:30	12:00	12:30	1:00	1:30	2:00	2:30	3:00	3:30	4:00	4:30	5:00	5:30	6:00	6:30
<i>Schedule Color:</i>																		
A		BREAK	AM	AM	AM	AM	AM	PM	PM	LUNCH	PM	PM	PM	PM	PM	PM	PM	
B		AM	AM	LUNCH	AM	AM	PM	PM	PM	PM	BREAK	PM	PM	PM	PM	PM	PM	
C		AM	AM	AM	AM	LUNCH	PM	PM	PM	PM	PM	PM	BREAK	PM	PM	PM	PM	
D		AM	AM	AM	AM	AM	AM	LUNCH	PM	PM	PM	PM	PM	PM	PM	BREAK	PM	PM
					LUNCH													
								LUNCH										
AM Route = Long Route (before 1:30), PM Route = Short Route (after 1:30)																		

The team of TPA which is not on tracking duty has one person dedicated to the communication with the emergency response team and the NM, one person tracking the late drivers and one person in break/having lunch.

Team Traffic Assistants (Alexander Alshanetsky, Negin Aryaee, Anthony Patire, Timmy Siau, Emma Strong, Jason Wexlar)

Responsibilities Report to Site Logistics Manager (Ryan) and follow instructions

Manage the driving behavior for one team of drivers

Guide drivers to exit Center Street garage

Guide drivers into appropriate parking spots at Union Landing

Assist drivers returning to Center Street Garage

Guide drivers into appropriate parking spots

Equipment Cell phone with pre-programmed numbers

Walkie-talkies (6) for communication with Site Logistics Officer and other Team Traffic Assistants

Support Staff None

Preparation Work Training one week before experiment

Will be briefed by Site Logistics Officer (Ryan) - details of this training under prep work of Site Logistics Officer section

Timeline Individual Timeline (TTAs will be divided into groups A and B, each consisting of 3

members who will be responsible for one driving team each)

- 7:00am** *Report to Center Street Garage*
- Briefing from Site Logistics Officer (Ryan)
- Overview of the whole day
- Specific instructions on how to direct drivers out of parking lot (Group Z)
- Instructions on how to get to Union Landing
- Prepare for the arrival of drivers
- 8:00am** Be prepared to help drivers find their cars (Group Z)
Drive to Union Landing (Group Y)
- 9:00am** *Arrive to Union Landing (Group Y)*
Briefing from Site Logistics Manager (Ryan)
- Instructions on how the flow in and out of the parking lot will work
- Instructions on what zone they are to cover and what their role is
- 9:30am** Guide drivers toward the exit of the Center Street Garage (Group B)
- 10:00am** Drivers start arriving to Union Landing (Group Y)
- Guide drivers to proper parking areas
- Guide drivers out of the lot at the appropriate time
- 10:30am** Drive to Union Landing (Group Z)
- 11:00am** *Arrive to Union Landing (Group Z)*
- Briefing from Site Logistics Manager (Ryan)
- Instructions on how the flow in and out of the parking lot will work
- Instructions on what zone they are to cover and what their role is

- 11:00am to 7:00pm** Manage incoming and outgoing vehicles
- Guide drivers to proper parking areas
- Guide drivers out of the lot at the appropriate time
- Check gas levels before drivers leave and instruct to fill up if necessary
- Alternate breaks so that only one traffic assistant is gone at any given time
- Check with Site Logistics Manager (Ryan) when leaving and returning to work
- 5:30pm** Drive back to Center Street garage in Berkeley (Group Z)
- Prepare for drivers return trip
- 7:00pm** Drivers start arriving back at Center Street garage (Group Z)
- Guide drivers to parking spots
- 7:30pm** Drive back to Center Street garage in Berkeley (Group Y)
- Help with final items

Nokia Monitor (4 Nokia Staff)

Responsibilities 2 teams of 2 people for Thursday, February 7th

4 teams of 2 people on Friday, February 8

Monitor in real time the same set of vehicles during the entire experiment. This includes:

Trip from UCB lot to BS at the Union Landing lot.

All loops done at the I880 site.

Trip from BS at the Union Landing lot back to UCB lot.

Install and test all chargers, bluetooth headsets, and phone equipment in all cars

Identify lost or problem vehicles, and update JC (or Olli-Pekka) with "lost" vehicles.

Equipment	Table space near Olli-Pekka, JC, TPAs and the Phone Operators Laptop Bluetooth headsets Phone chargers Experiment phones
Support Staff	None
Preparation Work	Make sure cell-phones can be tracked after they are taped in vehicles (Thursday, 02/08) Make sure all chargers and Bluetooth headsets are working cars by the day of experiment
Timeline	<p><i>Thursday, February 7, 2008</i></p> <p>4:00pm Begin installing and testing phones, chargers, and Bluetooth headsets in all cars (~ couple minutes per car). Finish shortly after 6:00pm.</p> <p><i>Friday, February 8, 2008</i></p> <p>7:30am Show up time at the Base Station in Union Landing.</p> <p>8:00am Start application on phones in all cars. Make sure equipment (headsets, chargers, phones) are ready to start tracking vehicles by 9:00am.</p> <p>9:30am Last vehicle is parked at the UCB lot Track vehicles.</p> <p>Once the last vehicle is parked at the UCB lot, Nokia Monitor is free</p>

to go.

Note that there are only 3 monitors considered, which gives no rest time for them. We may need to consider having at least one more monitor so they can have a one hour break (ask Quinn).

Head Personnel Manager (Juan Carlos Herrera)

Responsibilities	<p>Monitor and manage Team Personnel Assistants</p> <p>Make sure that driver's status is up-to-date</p> <p>Maximum driving times are not exceeded</p> <p>Determine appropriate responses to lost vehicles</p> <p>Supervise that lost driver gets back on I-880 with assistance of Phone Operators</p> <p>Monitor real travel times</p> <p>Supervise that drivers exceeding time limits are called back</p> <p>More</p>
Equipment	<p>Table space near Team Personnel Assistants, Nokia Monitors, and Phone Operators</p> <p>Mobile phone + Charger ; Pre-programmed contacts for staff, police, ambulance, tow-truck</p> <p>Folder with all paper documents; safety protocols, contact information</p> <p>Transportation to Union Landing (own vehicle)</p>
Support Staff	<p>Team Personnel Assistants (Arthur, Qingfang, Tarek)</p>
Preparation Work	<p>Divide hired drivers into 3 teams.</p> <p>Brief Team Personnel Assistants on their main tasks.</p>

Prepare “Responses to lost vehicles”-protocol for support staff

Guidance using visual monitoring of vehicle movements (turn left etc)

Meet with Safety Officer and Team Personnel Assistants

Create spreadsheet to keep track driving times, gas status, etc.

Contact TPAs and Nokia Monitors by email and ask if any questions about “Responses to lost vehicles” –protocol.

Timeline	7:00am	Show up at UCB lot Setup all communications/devices with staff.
	8:50am	Start driving –with the TPAs- to the BS.
	9:30am	Arrive at BS Supervise TPAs during the first driver/car matching.
	10:45am	Supervise shift schedules, travel times, driving times and response to lost vehicles. Report any accidents/incidents to Safety Officer. Leave Union Landing with Safety Officer and TPAs when last car arrives to UC Berkeley

Phone Operator (6 Nokia staff)

Responsibilities Contact lost or problem team drivers

Implement strategy determined by JC to get vehicles back to base station/ UCB lot via phone

Handle inbound calls from drivers. In case of emergency, contact Olli-Pekka immediately.

Equipment Table space near Olli-Pekka, JC, TPAs and the Nokia Monitors

Dedicated phones with all the 100 cell-phone numbers

	Laptop
	Maps
Support Staff	None
Preparation Work	Meet with Olli-Pekka, JC and the TPA to check protocol: How to help lost drivers (direct them back to Base Station from the freeway, etc).
Timeline	<p>8:00am Show up time at the Base Station. Make sure they are ready to start calling drivers by 9:00am.</p> <p>9:30am Last vehicle is parked at the UCB lot: contact drivers by phone upon request.</p> <p>Once the last vehicle is parked at the UCB lot, Phone Operator is free to go.</p> <p>Note that there are only 3 operators considered, which gives no rest time for them. We may need to consider having at least one more operator so they can have a one hour break (ask Quinn).</p>

Response Team

Composition	Tow-truck team (1 unit)
	Continuous presence not necessarily possible
	Information of nearest possible services available for Safety Officer
	Tow-truck takes away UC Berkeley rental cars broken/involved in an accident
	Ambulance team (1 unit)
	Continuous presence on the Union Landing (if possible)
	Helps drivers in case of emergency
	CHP
	Is called to assist in an accident/emergency by Safety Officer

Constant presence impossible in practice?

Responsibilities	<p>Assist drivers in the field</p> <p>Respond to accidents, car trouble, etc.</p> <p>Help get vehicles back to base station/ home base, etc.</p> <p>Add more here...</p>
Equipment	<p>1 police car</p> <p>Add more here...</p>
Support Staff	<p>None</p>
Preparation Work	<p>None</p>

SITES

UC Berkeley Lot

Berkeley Site Preparation

WEDNESDAY NIGHT, 02/07/2008

11:00pm to finish 11:00pm 100 spaces will be set aside (via caution tape/roped off area) for Berkeley, starting 11:00pm

THURSDAY MORNING, 02/07/2008

9:00am Marika and other Berkeley staff will coordinate with Enterprise to number cars and keys starting at 9 AM and continuing throughout the day.

9:30am Last vehicle is parked at the UCB lot: contact drivers by phone upon request.

THURSDAY AFTERNOON, 02/07/2008

Before 4:00pm **Numbering system:**

- Nokia will have numbered the phones before 4:00 PM, when they come to Berkeley.
- Each key, phone, and car will have the same number, respectively.

4:00pm Nokia will have pre-numbered phones ready in Berkeley

5:00pm Berkeley staff and Nokia staff will begin to install phones in cars, each car # is the same as its phone #

6:00pm Nokia staff will check software

FRIDAY MORNING, 02/08/2008

6:30am Marika and two helpers will set up coffee and pastries; one folding table; three folding chairs. One helper will have picked up table and chairs from CCIT and brought it over. Marika will have picked up coffee and pastries from Starbucks.

7:30am Participants will come to the location
Marika and 4 helpers will check participants in; check drivers licenses again, make sure we have a copy of consent form, answer questions, hand out maps, etc

8:30am Nokia staff will arrive to check software again.
Drivers will be briefed, directed out of lot, etc.

10:00am Take down will begin (table, food leftovers, and chairs)

11:00am Parking lot will be cleared

FRIDAY AFTERNOON, 02/08/2008

- 5:00pm** Marika and two helpers will arrive at Berkeley parking lot to make sure rented area is cordoned off
- 5:00pm** Nokia will arrive to collect phones.
- 6:00pm** Participants may start to arrive; Marika and Nokia will collect phones and keys.

SATURDAY MORNING, 02/09/2008

- 9:00am** Marika will meet with Enterprise staff and then the keys. Enterprise will collect the cars throughout the day on Saturday, ending at 5:00 PM.

Preparation Work for Berkeley Site

Make sure paperwork is ready to bring to site: consent forms, check in sheets with driver's info, checklist to make sure they have a current license on them at the site, etc.

Center Street Parking Lot Logistics

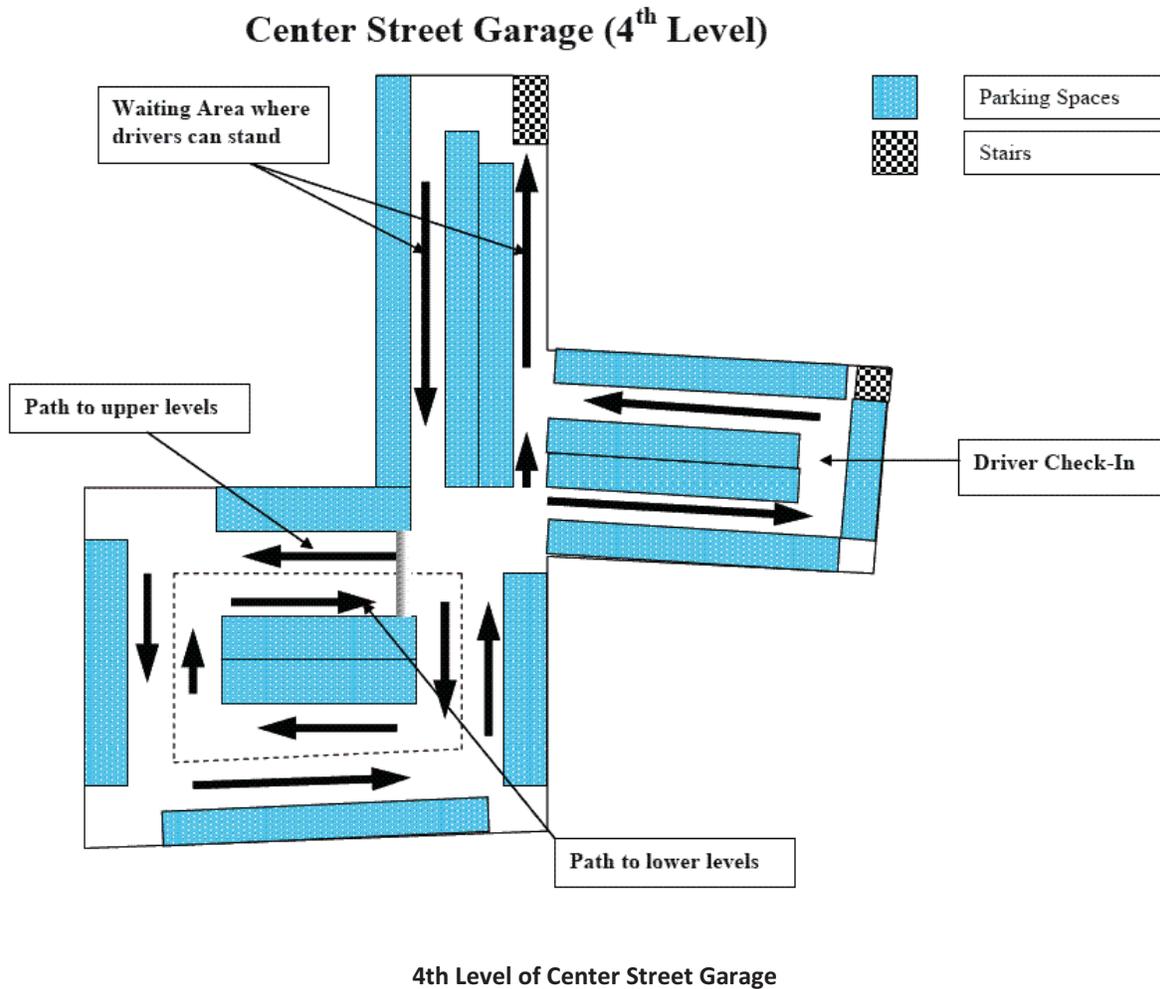
Site Layout (see map of 4rd level of garage)

15 of the reserved parking spaces will be used to setup a table for checking in drivers.

Cars will be parked backwards so that cars can drive forward out of the parking spaces.

Drivers will be able to wait for their various briefings in the marked areas of the map. This will not interfere with other (non-experiment) vehicles in the garage because we will have this entire section reserved for us (and roped-off from outside vehicles).

Vehicles will be numbered in a logical order to make it easy to direct drivers to their cars and to facilitate placing the phones in the cars.



Release of Vehicles from the parking lot (see overview map of the Center Street garage)

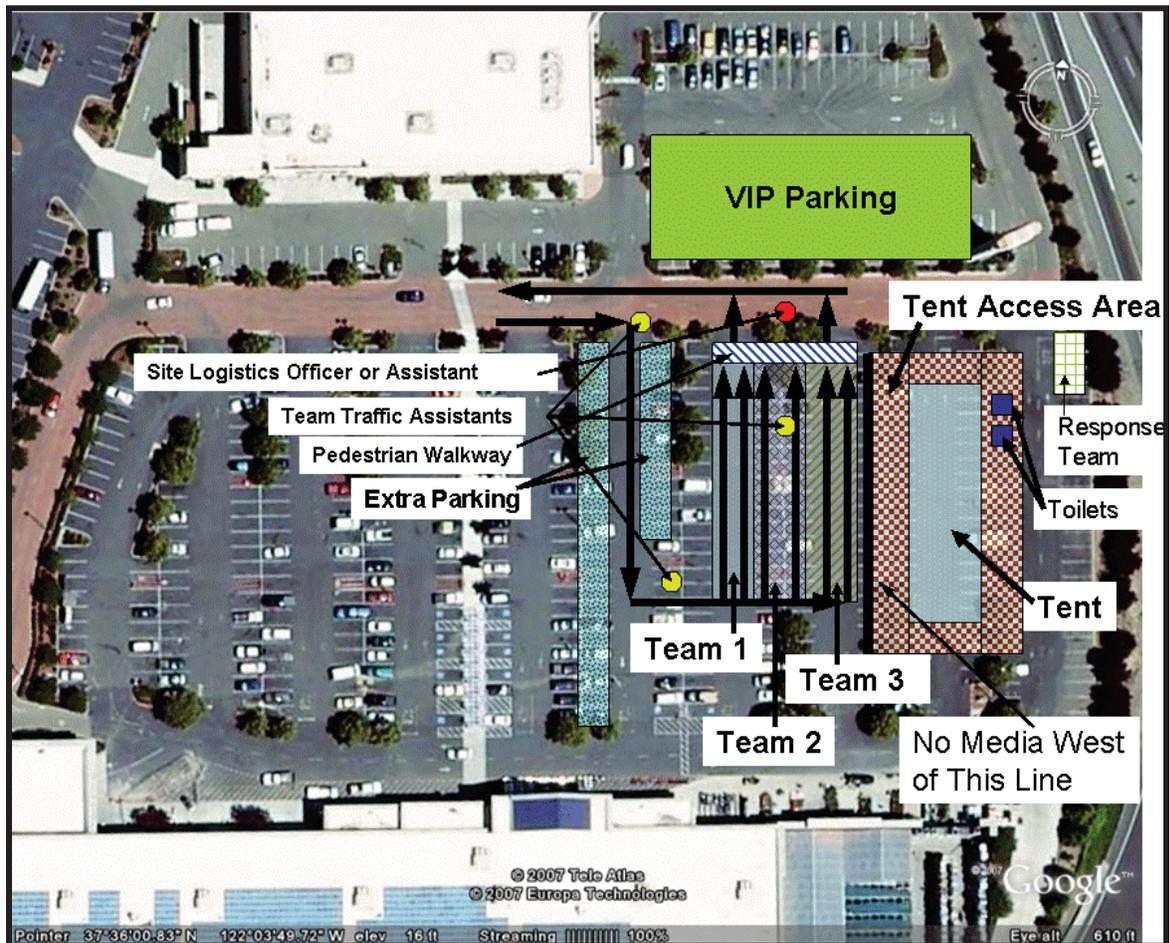
Vehicles will be released at a rate of approximately 2 per minute as allowed by traffic conditions on Allston Way.

Vehicles will exit onto Allston Way by turning right out of the parking garage. Then turn left onto MLK and go south until road turns into Highway 24, bear right onto I-980 and continue onto I-880.

Two team traffic assistants and site logistics assistant will guide the drivers towards the exit of the parking lot. They will release vehicles by team, letting one member from team 1 leave, then one member from team 2, then one member from team 3 and so on.

Union Landing Lot

Organization of the Site



Overview of Union Landing Parking Lot

Tent General Set-Up

Area I (Staff Area):

This place is designated for staff (office style). It holds 32 people, and includes 7 x 4-person and 2 x 6-persons tables. The area has to provide power for 16 cell phone chargers and 25 laptops at the same time. Minimum 1 exit door to the outside is required (see the general layout). We envision having a wireless router in the staff area to create a local area network.

Area II (Briefing Area):

This area is designated for briefing the drivers and consists of 3 compartments. Each compartment hosts 15 drivers. The desired layout is displayed in the general layout. In this area, the drivers are divided into three groups of 15 people and briefed for 10 minutes in each compartment. We prefer to have walled/separated compartments to avoid any mixing. The drivers enter from the driver lounge (see the general layout) and exit from the back of the tent, located in Area II (marked in the layout). This area requires power for three laptops and three projectors. Total three tables are required. The drivers will stand in the compartments. We are ok with a compact area that can hold total 3x15 people standing. An aisle is needed to facilitate people entering and exiting the area (see the layout).

Area III (Driver Lounge):

This area is designated for drivers resting during the experiment. During the normal operation, this area should be capable of hosting 50 drivers. There is a possible scenario in which we need to host about 100 drivers in the lounge during the morning briefing. We don't want to provide chairs for all the 100 drivers. They can stand. This area will also require power for 2 TVs and 10 laptops.

*Equipment***Technical and Mechanical Tools**

Laptops (2-for logistics). Others for Nokia. No ordering required

Server? (1) 1/08

Color printer (1) 1/08

Extension cords (12-Costco) 1/08

Power strips (20-Costco) 1/08

Tool kits (1-Ace Hardware) 1/08

Duct tape (3 rolls) 1/08

Generator (supplied by tent company) *Finalized 12/07

Office Supplies

Notebook (20-CCIT will order 1/08)

Pens, pencils(10 boxes-CCIT will order 1/08)

Paper(3 packages-CCIT will order 1/08)

Masking tape(3-CCIT will order 1/08)

File folders (20-CCIT will order 1/08)

Clip board(30--CCIT will order 1/08)

Scissors (5-CCIT will order 1/08)

Recorders for vehicle tracking (10-CCIT will order 1/08)

2 8 gallon trash cans; One roll of industrial trash bags

2 waste paper baskets

Entertainment, food, etc.

Breakfast: Coffee-Starbucks in shopping center *Order in 1/08

Lunch: Togos sandwiches in shopping center *Order in 1/08

Dinner: Snacks from Costco—Pretzels, fruit, bread, peanut butter and jelly, cookies *Buy in 1/08

Utensils: 200 plates and forks; cups for coffee *Order in 1/08

Coffee: All day from Starbucks or we Make? *Order in 1/08

Popcorn machine from tent rental supply company *Finalized 12/12

Water, Soda (200-from Costco) *Order in 1/08

Site set-up

Umbrellas 5-foldup *Order in 1/08

Heater (1 from tent rental supply company) *Finalized 12/07

Tent (1 from tent rental supply company-see site plan) *Finalized 12/07

Tables (27 – 9 for staff, 2 for food, 1 check in, 15 for participants-from tent rental supply company-see site plan) *Finalized 12/07

Mobile Century 390



We are working on fitting the tables into the participants area. The tent company will help us on that.

Rope: (50 feet Ace Hardware)

Traffic cones: (23 \$230—Ali has website)

Chairs (90 from tent rental supply company-see Ali's site plan) *Finalized 12/07

Miscellaneous

Bull horns (2-CCIT will order 1/08)

'orange vests' (50 from Tim)

Name tags (200-CCIT will order 1/08)

One 6 foot table 2 chairs for Berkeley parking lot.

TimeLine of the Site

Before D-Day (Thursday, 02/07/2008)

3:00pm Marika will call the mall and double check the parking status
3:30pm Marika contact the tent company

D-Day (Friday, 02/08/2008)

12:00am Mall authorities clear the parking lot in Union City.
6:00am Staff on site (Ali and 2 site staff)
6:00am Tent people to set up the tent.
6:30am Ali and the staff will mark up the parking area (including the Response Team Parking)
7:00am Tent staging is done.
7:30am Tent set-up is finished. (includes table set-up, rope-up the briefing and training area, set-up the staff area) The staff will place the required equipment/office supplies.
7:30am Generator is installed. (tent staff)
8:00am Wiring/ heating. (Ali, 2 site staff and tend crew)

8:30am	Nokia staff arrives. Site logistics officer (Ryan), Human Logistics Officer (Steve) and personnel manager arrive
8:45am	Nokia staff set up their laptops.
9:00am	Coffee/snack and water. (1 site staff)
9:00am	Response team arrives.
9:30am	First ASCE team arrives
9:30am	First personnel assistant arrives
9:30am	First Team personnel arrives
10:00am	All the staff are seated
10:00am	first driver shows up
10:00am	The traffic team starts guiding the vehicles
10:05am	The reception starts
10:10am	The Team Leads start briefing the drivers
11:00am	Marika on site
11:00am	Marika handles the coffee
11:30am	Last driver shows up
11:30am	Food is served to drivers based on their driving schedule (Ali, Marika, 2 staff)
2:00pm	Last food is served (Ali, Marika, 2 staff)
4:00pm	Marika leaves the site to Berkeley parking lot
5:00pm	Cleaning starts (2 site staff, we need two more people)
6:00pm	Clean the office supplies and wires (Ali and staff)
6:30pm	Drivers start returning to Berkeley
7:00pm	Cleaning up the cones (2 site staff)
7:30pm	Last driver leaves the site
8:30pm	Site will be fully cleaned and cleared (2 staff plus 2 extra traffic assistants)
8:45pm	Dispose the trash bags into dumpsters

Union Landing Site Preparation (02/08/2008):

Explained in the last section

Union Landing Site Arrangements (02/09/2008):

No Arrangement for the next day

Preparation Work for Union Landing Site

Make sure paperwork (if required) is ready to bring to site: paper work related to tenting, check-in sheets with driver's info, etc.

The site managers has to have a list of drivers and their driving schedule to coordinate lunch

Union Landing Site Logistics

Entering and exiting vehicles (Refer to site layout map):

Vehicles will enter the lot using a single row as depicted on the map.

When drivers arrive at the entrance to the parking lot, one of the Team Traffic Assistants will identify the team to which that driver belongs by looking at the number marked on the hood of the car. This assistant will direct drivers south along the western most parking corridor to one of the other traffic assistants and also inform the other TTAs of the incoming driver's team (by walkie-talkie).

3 other TTAs (stationed as depicted on the map) will guide the driver into the appropriate queue for that driver's team.

We will have 6 parallel queues heading towards the exit of the parking lot, 2 for each team. The 6 queues will funnel through 2 exit points from the parking lot.

Each queue will be able to hold about 8 vehicles and still allow cars to drive across the parking lot. This means that ~48 vehicles will be able to be in queue at any one time. If more vehicles than 48 are present, we will extend the queues south until we reach the passageway in front of Lowe's which will require entering vehicles to drive around the end of a row to enter their queue. If we exceed this limit, we will use the space designated "Extra Parking" on the map.

Drivers, once parked, will be directed towards the tent and told to be cautious as there will be other entering and exiting vehicles nearby.

The TTAs will also be responsible for moving cars forward as necessary to ensure the queues do not get backed up.

The Site Logistics Officer (Ryan) will manage the departing vehicles to make sure an equal number of drivers from each team is on the road at any given time. The SLO will direct drivers to the exit at the appropriate time.

The SLO will also check the gas level of each vehicle before it exits to ensure that the vehicle has enough gas for 3 hours of driving. If not, the driver will be instructed to go to the gas station before starting on the appropriate highway loop.

Other

Parking Lot:

Center Street Garage-Finalized. Contract will be drawn up by 12/12. The parking lot will be reserved and cleared on the evening of 2/07 by mall personnel.

Car Rental:

Obtaining formal estimate for procurement.

Restrooms:

Six in Barnes and Nobel: open to the public and does not require keys.

One in property management office: open to the public and does not require a key.

One in Starbucks: open to the public and does not require a key.

**Marika will map out other rest rooms on 12/13 during a site visit and photo taking opportunity for training material.

EMERGENCY PROCEDURES

Response Procedure for Lost Vehicles During Transfer Periods and Actual Experiment

“If a driver gets lost and cannot get immediately connected to Phone Operator, driver must pull over TO A SAFE LOCATION and wait until Phone Operator is available”

If there is an emergency, drivers should not call the staff to alert them of the accident in order to avoid phone congestion. Rather, the staff will handle calling 911 or emergency services, assuming those services do not respond first already.

During transfer period towards Union Landing or UC Berkeley

Phone Operators prepare to calculate a route from driver’s current location to Union Landing using navigation software (Nokia navigation software, Google maps, most suitable). Since the driver is not allowed to operate N95 phone, Phone Operators must provide voice guidance until driver is back on I-880 going into right direction.

Nokia Monitors try to spot unexpected routes

- Report to Head Personnel Manager
- Phone Operators co contact driver and calculate route back to I-880
- When close to destination (Union Landing or UC Berkeley) but lost guidance to this destination or until driver is comfortable with route

Drivers can call Phone Operators in case they get lost

During experiment loops

Nokia Monitors try to spot unexpected routes

- Report to Head Personnel Manager
- Phone Operators contact driver and calculate route back to I-880

Drivers can call Phone Operators in case they get lost

- Loop information is available in driver's package
- Right loop also obtained from Team Personnel Assistants

Responses to different scenarios

Driver gets lost

Nokia Monitors spot a lost driver

- Report to Head Personnel Manager
- Phone operators call driver and assist back to I-880 using "Responses to lost vehicles"

Driver contacts Phone Operator after getting lost

- Report to Head Personnel Manager
- Phone operators assist driver back to I-880 using "Responses to lost vehicles"

Car runs out of gas

Driver calls Phone Operator

Report to Head Personnel Manager

Car gets broken

Driver calls Phone Operator

Report to Head Personnel Manager

Tow-truck is sent to the scene

Tow-truck takes car away and then gets back to Union Landing

CHP or tow-truck picks the driver off the highway and brings back to Union Landing

Driver gets suddenly ill

Driver contacts Phone Operator

Report to Head Personnel Manager

Medical care on the scene and possibly CHP assisting

If driver can return to Union Landing without assistance he/she will be treated at the parking lot

Nokia Monitor spots a stopped vehicle

Report to Head Personnel Manager

Try to contact driver

If no response and if necessary ambulance and CH

Driver involved in a car accident

In case of a serious car accident in which the car is no longer drivable, please

Call Sarah Peterson at Enterprise on her cell phone: 415-720-7437; request a replacement vehicle to be delivered ASAP to the Union Landing mall lot.

Call the towing company LA Tech: 510-234-5044; Ask for Jeanne

Send staff out to driver to bring him or her back to Union Landing lot, if driver is still at accident site.

Driver calls Phone Operator

Report to Head Personnel Manager

CHP and ambulance sent to scene

Driver/CHP presents Insurance Certificate of rental car

Nokia Monitor spots a stopped vehicle

Report to Head Personnel Manager

Try to contact driver

If no response and if necessary ambulance and CH

Driver/CHP presents Insurance Certificate of rental car

Driver's license gets suspended

Driver is brought to Union Landing by CHP

General instructions to Safety Team

Nokia Monitors report any suspicious behavior of cars to Head Personnel Manager

Suspicious behavior is for example

Driver is lost

Driver doesn't move in 20 minutes

Car doesn't move and gets overtaken continuously

Head Personnel Manager reports all possible incidents to Safety Officer who will make the final decisions of responses, i.e., sending CHP, ambulance

PREPARATION WORK

TO DO LIST

Ali	
Marika	
Quinn	
Arthur	
Dan	
JC	
Ryan	
Olli-Pekka	
Jeff	
Other	
Christian+JC	

ANNEXES

Loop Scheduling

THE SAME SET OF VEHICLES IS ASSIGNED TO THE SAME TEAM DURING THE ENTIRE EXPERIMENT
 EACH NOKIA MONITOR MONITORS ONLY ONE LOOP ALL THE TIME (EASIER TO SEE WHEN A DRIVER GETS LOST)

Lunch:

A: 14:30-15:30

D: 11:30-12:30

B: 12:30-13:30

C: 13:30-14:30

		AM PHASE (3 hours)											
		10:30	10:45	11:00	11:15	11:30	11:45	12:00	12:15	12:30	12:45	13:00	13:15
A	Red	BREAK				AM							
B	Red	AM	AM	AM	AM	LUNCH				AM	AM	AM	AM
C	Red	AM	AM	AM	AM	AM	AM	AM	AM	LUNCH			
D	Red	AM	AM	AM	AM	AM	AM	AM	AM	AM	AM	AM	AM
A	Orange	BREAK				AM							
B	Orange	AM	AM	AM	AM	BREAK				AM	AM	AM	AM
C	Orange	AM	AM	AM	AM	AM	AM	AM	AM	BREAK			
D	Orange	AM	AM	AM	AM	AM	AM	AM	AM	AM	AM	AM	AM
A	Yellow	BREAK				AM							
B	Yellow	AM	AM	AM	AM	BREAK				AM	AM	AM	AM
C	Yellow	AM	AM	AM	AM	AM	AM	AM	AM	BREAK			
D	Yellow	AM	AM	AM	AM	AM	AM	AM	AM	AM	AM	AM	AM

		PM PHASE (5 hours)																					
		13:30	13:45	14:00	14:15	14:30	14:45	15:00	15:15	15:30	15:45	16:00	16:15	16:30	16:45	17:00	17:15	17:30	17:45	18:00	18:15		
A	Red	PM	PM	PM	PM	BREAK				PM													
B	Red	PM	PM	PM	PM	PM	PM	PM	PM	BREAK				PM	PM								
C	Red	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	BREAK				PM	PM	PM	PM		
D	Red	LUNCH				PM	BREAK																
A	Orange	PM	PM	PM	PM	BREAK				PM	PM												
B	Orange	PM	PM	PM	PM	PM	PM	PM	PM	BREAK				PM	PM								
C	Orange	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	BREAK				PM	PM	PM	PM		
D	Orange	LUNCH				PM	BREAK																
A	Yellow	PM	PM	PM	PM	BREAK				PM	PM												
B	Yellow	PM	PM	PM	PM	PM	PM	PM	PM	BREAK				PM	PM								
C	Yellow	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	BREAK				PM	PM	PM	PM		
D	Yellow	LUNCH				PM	BREAK																

		CONCLUSION PHASE (1 hour)				Total hours of driving	
		18:30	18:45	19:00	19:15		
A	Red	return to BS, driven to UC Berkeley lot by Red-D	0	0	0	0	6
B	Red	return directly to UC Berkeley lot	1	1	1	1	7
C	Red	return directly to UC Berkeley lot	1	1	1	1	7
D	Red	drive Red-A to UC Berkeley lot	1	1	1	1	7
A	Orange	return to BS, driven to UC Berkeley lot by Orange-D	0	0	0	0	6
B	Orange	return directly to UC Berkeley lot	2	2	2	2	7
C	Orange	return directly to UC Berkeley lot	2	2	2	2	7
D	Orange	drive Orange-A to UC Berkeley lot	2	2	2	2	7
A	Yellow	return to BS, driven to UC Berkeley lot by Y-D	0	0	0	0	6
B	Yellow	return directly to UC Berkeley lot	3	3	3	3	7
C	Yellow	return directly to UC Berkeley lot	3	3	3	3	7
D	Yellow	drive Yellow-A to UC Berkeley lot	3	3	3	3	7

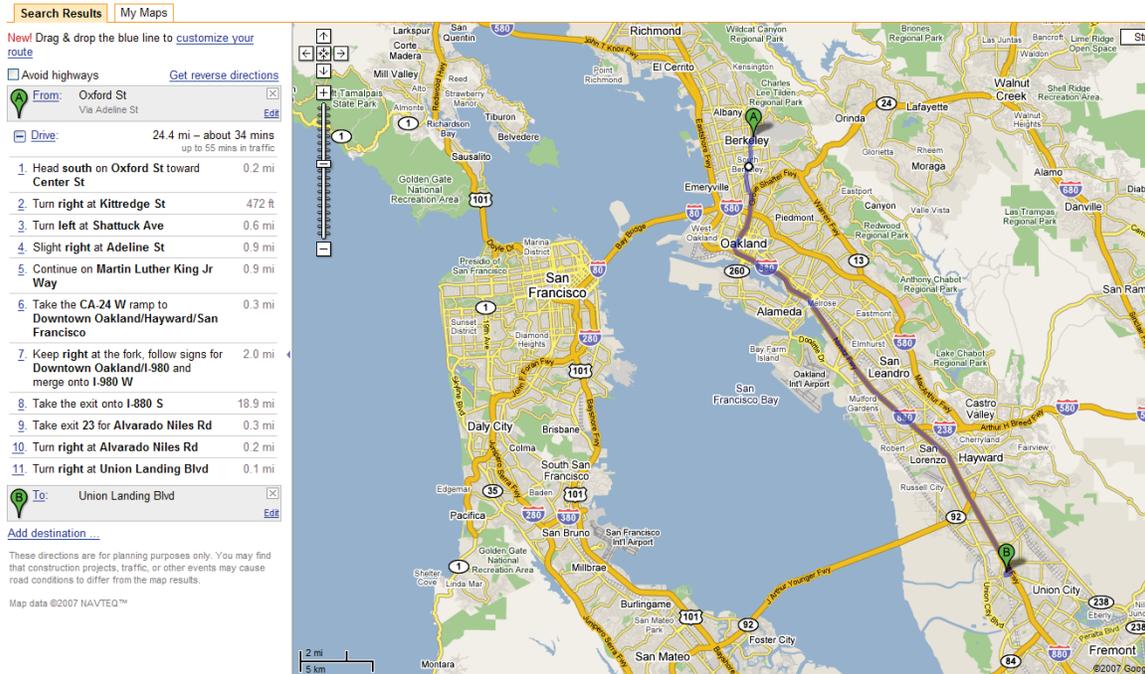
APPENDIX

Briefing 1

Users: Team Leaders
To: Drivers
Time: 8:45 am, UCB lot

Description: Detailed explanation with pictures:
 Shattuck → MLK → CA24W → I880S → Exit 23, Alvarado-Niles Rd.

Support Mat'l: One-pager with Google map information (UCB to BS) for drivers to use it on-board.



Briefing 2

Users: Team Leaders
To: Drivers
Time: 8:45 am, UCB lot, before a new shift starts at the BS.

Description: Description of the loops. Each team will do the same loop during each phase. Therefore, the Team Leaders only need to explain one loop during each briefing.

Each team leader will receive 2 scripts, one for each phase. Each script will have pictures showing how to turn around at both ends of each loop.

Support Mat'l: One-pager for the drivers -with the name of the ramps they have to use- to look at it on-board.

Phase I (10:30am – 1:30pm)

Loop	From (to the north)	To (to the south)	Distance (miles)
1	Winton Ave.	Thorton Ave.	9.4
2	CA92	Mowry Ave.	8.6
3	Tennyson Rd.	Stevenson Blvd.	9.3

Phase II (1:30pm – 6:30pm)

Loop	From (to the north)	To (to the south)	Distance (miles)
------	---------------------	-------------------	------------------

1	Winton Ave.	Alvarado-Niles Rd.	4.5
2	CA92	Alvarado Blvd.	5.2
3	Tennyson Rd.	Decoto Rd.	5.4

Briefing 3

Users: Team Leaders
To: Drivers
Time: Before a new shift starts at the BS.

Description: Information provided to drivers every time a new shift begins. Each team leader will brief the 11 drivers that are resting just before they start a new driving shift. The briefing includes:

- How to get from (to) the BS to (from) the freeway, in case it is the first briefing.
- Description of the loop drivers will be doing and the schedule (start/end time).
- Instructions on how to refill the gas tank if needed (where to refill, how to operate the gas pump, and payment method).
- How to get from the BS to the UCB lot, in case it is the last briefing. For those drivers that have to come back to pick other drivers up, this has to be said during their last briefing.

JC will coordinate with the team leaders to make sure that resting drivers are ready to go once the driver comes back to the BS.

Support Mat'l: Directions to go from (to) the BS to (from) the freeway:

- Detailed explanation with pictures: Landing Union Blvd, turn left at Alvarado-Niles Rd., turn right to take the on-ramp to I880S to San Jose (or turn left to take the on-ramp to I880N to Oakland).
- Detailed explanation with pictures: take exit 23 to Alvarado-Niles Rd., turn right at Alvarado-Niles Rd. if coming from I880S (or turn left if coming from I880N), turn right at Landing Union Blvd.

Gas refill instructions:

- The Team Personnel Assistant will tell the driver if a gas refill is needed before entering the freeway. If the gas tank is less than a half, a gas refill will be requested. The team leaders are responsible for briefing the drivers regarding the location of the gas station, payment method, and how to operate the gas pump. TPA will show the driver how to open the gas tank before driver leaves.
- Location: The gas station is located at the South-East corner of Alvarado-Niles Rd./Union Landing Blvd.
- Payment method: prepaid card?

Protocol for requesting unexpected breaks:

- In case a driver needs to stop driving before the shift has ended, he/she will contact the phone operator (emergency call). The phone operator will tell the driver how to proceed. The phone operator will contact JC, who will give specific instructions on how to come back to the BS → JC contact Team leader to brief an alternate driver.

Briefing 4

Users: Team Leaders
To: Drivers
Time: At the beginning of a break, BS

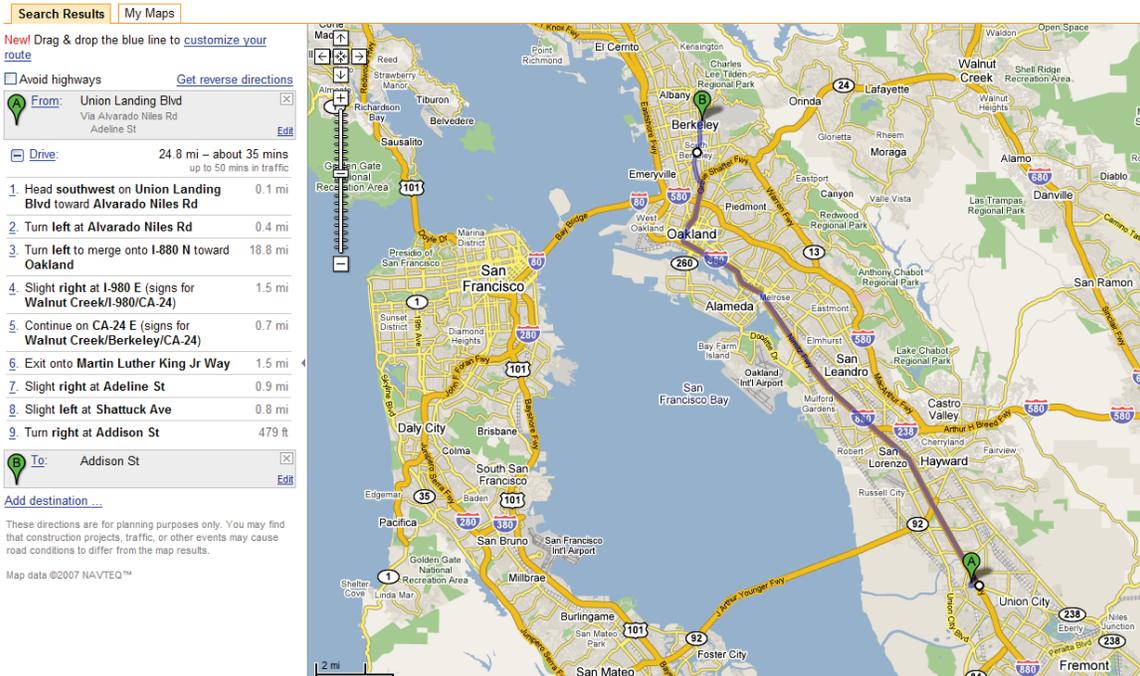
Description: Information on what drivers are allowed to do during their breaks. Each break lasts 1 hour. At the beginning of the break, the team leader will tell the drivers the start time of the next shift. During the break drivers are expected to stay at the base station, without interfering researchers/staff work. Drivers can bring book/papers/magazines to read or portable games to play, use their own cell-phone to talk, listen music (with headphones). Drivers can bring their laptop at their own risk, but they are not allowed to use any of the outlets available in the tent (there won't be internet access at the BS). If a driver wants to go for a walk, he/she needs to let his/her team leader know.

Briefing 5

Users: Team Leaders
To: Drivers
Time: At the last briefing, BS

Description: Script document explaining how to go from BS to UC Berkeley
 Detailed explanation with pictures:
 I880N → CA24E → exit at 51 St. → MLK (left) → Shattuck

Support Mat'l: One-pager with Google map information (UCB to BS) for drivers to use it on-board.



Briefing B9a (Safety)

Users: Team Leaders

To: Drivers

Time: 8:45 am, UCB lot, before a new shift starts at the BS.

Description: What a driver has to do in case of emergency.

In case any problem arises on the road, drivers will contact the corresponding Phone Operator using the N95 phone. This briefing will contain instructions on how to use the phone to this end. The Phone Operator will communicate directly with the Safety Officer, which will decide the action to take.

7 Press Release Materials

Appendix 7 contains press release materials provided to visitors from local governments and from the press. The fact sheet containing a summary of the deployment experiment is furnished in Appendix 7.1. The guest program is provided in Appendix 7.2.

7.1 Fact Sheet

Mobile Century Fact Sheet



MOBILE CENTURY – FEBRUARY 8, 2008

CORRESPONDANCE
mobilecentury@calccit.org

Contact information

J.D. Margulies
Associate Director
California Center for
Innovative Transportation
(510) 642-5929
jd@calccit.org

Alexandre M. Bayen
Assistant Professor
Department of Civil and
Environmental Engineering
University of California,
Berkeley
(510) 642-2468
bayen@berkeley.edu

Quinn Jacobson
Research Leader
Nokia Research Center –
Palo Alto
(650) 521-3243
quinn.jacobson@nokia.com

Using GPS Mobile Phones as Traffic Sensors: A Field Experiment





Final Program

10:00 AM - 11:30 AM *Mobile Century Kick-Off and Press Conference*
Union Landing Mall, Union City, California

PRESS CONFERENCE
Alexandre M. Bayen, Assistant Professor of Civil and Environmental Engineering, University of California, Berkeley
Bob Imrecci, Chief Technical Officer and Head of Nokia Research Center, Nokia
Renele Havelle, Chief Deputy Director, California Department of Transportation

GUIDED TOURS OF THE COMMAND CENTER AND MINGLING TIME

12:30 PM - 2:30 PM *Launch and Presentations*
290 Hearst Memorial Mining Bldg., UC Berkeley

WELCOME ADDRESS
Lisa Alvarez-Cohen, Fred and Claire Sayer Professor and Faculty Chair, Civil and Environmental Engineering, University of California, Berkeley

PRESENTATIONS
J.D. Margulies, Associate Director, CCIT
The California Center for Innovative Transportation Model
John Paul Chen, Director, Nokia Research Center, Palo Alto
The Nokia Research Center, Palo Alto: Mobility, Community, and Serendipity
Paul Wright, Acting Director of CITRIS, & Martin Berlin Professor of Mechanical Engineering, University of California, Berkeley
The Center for Information Technology Research in the Interest of Society
Jeremy Walz, Government & Industry Relations, NAVTEQ
NAVTEQ and NAVTEQ Traffic
Quinn Jacobson, Research Leader, Nokia Research Center, Palo Alto
The technology behind the Mobile Century experiment

3:00 PM - 4:00 PM - OPTIONAL TOUR OF THE BERKELEY CAMPUS

4:00 PM - 5:00 PM *Scientific Seminar*
290 Hearst Memorial Mining Bldg., UC Berkeley

INTRODUCTION
Paul Wright, Acting Director of CITRIS

TALK
Alexandre M. Bayen, Assistant Professor of Civil and Environmental Engineering, University of California, Berkeley
Using GPS Mobile Phones as Traffic Sensors: A Field Experiment

CALIFORNIA CENTER FOR INNOVATIVE TRANSPORTATION
UNIVERSITY OF CALIFORNIA, BERKELEY - 2305 BANCROFT WAY, SUITE 300 - BERKELEY, CA 94720-3830
PHONE: (510) 642-4522 - FAX: (510) 642-0910 - HTTP://WWW.CALCCIT.ORG



MOBILE CENTURY – FEBRUARY 8, 2008

Using GPS Mobile Phones as Traffic Sensors: A Field Experiment

Objectives at a glance:

A field experiment to collect traffic data from GPS-enabled cell-phones

Provides a proof of concept for traffic-flow reconstruction from probe vehicle measurements



Collecting Traffic Data from Mobile Probes

The potential of cell phones to operate as traffic data collection devices has been considered by the Intelligent Transportation Systems (ITS) community for several years. Government agencies currently deploy networks of traffic sensors that are expensive to install and maintain. Leveraging commercial cellular networks could drastically cut the ongoing costs of traffic monitoring and expand coverage to

thousands of miles of highways and urban arterials for which sensors are not even considered an option. Available methods to collect data from cell phones rely on approximate positioning provided by the cellular networks and have shown limited accuracy to date. However, GPS chips are now built into more and more handsets and they will soon become as ubiquitous as cell phone cameras. For instance, Nokia will stop producing

cellular phones without GPS in less than 18 months. The prospect of large numbers of GPS-equipped cell phones reporting position and speed with 10 meter / 3 mph accuracy at regular intervals represents a huge leap forward. Yet its implementation requires addressing key questions regarding individual privacy, data ownership, network load, and proper traffic flow estimation techniques.



The experiment will test a novel traffic monitoring system designed to collect velocity and position data from GPS-enabled Nokia N95 cellular phones.

The Experiment

Under the umbrella of the California Center for Innovative Transportation (CCIT), Caltrans, Nokia, and UC Berkeley's Department of Civil and Environmental Engineering are collaborating to conduct an unprecedented experiment in the area of traffic monitoring. For an entire day, 100 vehicles carrying the GPS-equipped Nokia N95 will drive along a 10-mile stretch of I-880 between Hayward and Fremont,

California. Given their number, those vehicles will constitute up to 5% of the traffic traveling along this section, a penetration rate that adequately represents the potential of the market for GPS-equipped cell phones in the near future. The data obtained in the experiment will be processed by a team of UC researchers led by Professor Alex Bayen to determine the trade-offs between data volumes, information quality and

privacy concerns. The experiment will thus underline the value-added available from cellular phones, which could rapidly complement existing traffic sensors. This work will ultimately guide the design and sharing modalities of future traffic information collection systems that can be operated by the private sector and offer substantial benefits to government agencies and the traveling public.

CALIFORNIA CENTER FOR INNOVATIVE TRANSPORTATION
UNIVERSITY OF CALIFORNIA, BERKELEY · 2105 BANCROFT WAY, SUITE 300 · BERKELEY, CA 94720-3830
PHONE: (510) 642-4522 · FAX: (510) 642-0910 · HTTP://WWW.CALCCIT.ORG

2 USING GPS MOBILE PHONES AS TRAFFIC SENSORS: A FIELD EXPERIMENT

Logistics



100 vehicles carrying the GPS-equipped Nokia N95 will drive along a 10-mile stretch of I-880 between Hayward and Fremont, California.

The experiment, nicknamed *Mobile Century*, will take place on Friday, February 8, 2008, from 9:30am until 6:30pm. 150 hired UC students will drive 100 vehicles in a loop along I-880 between Winton Ave. to the North and Stevenson Blvd. to the South. This 10-mile-long section was selected for its traffic properties, the availability of an existing knowledge base for this particular highway from

traffic simulations, and for its proximity to UC Berkeley. There are a number of conveniences available along the section, including access to parking, gasoline, and food. Drivers will be taking 1-hour breaks throughout the day. The driving pattern will ensure a penetration of 4-5% of the total flow. Each vehicle will carry a Nokia N95 phone, which will store speed and position information every 3 seconds.

The measurements from the cell phones will be sent wirelessly to a server for real-time processing. Cameras located on bridges at both ends of the loop will be used to record the actual travel times of all vehicles, including those not participating in the experiment, which will provide a 'ground truth' reference that can later be compared with the estimates produced from the GPS data.

Capturing the Event

A public relations event will be organized in the morning at the experiment command center for representatives from government agencies, industry, and academia. This will be followed by a reception and lunch at CCIT headquarters in Berkeley starting at 12:30 pm. Come share the excitement of a Berkeley-scale event with societal-scale transformation potential!



A Google Earth rendering of collected GPS data. Virtual trip lines measuring speed upon crossing of the vehicles are shown in red.

VIP and Media Event:

February 8, 2008
10 am – 2:30 pm
Union Landing, California
and UC Berkeley

Contact Information

Alexandre M. Bayen
Assistant Professor
Department of Civil and
Environmental Engineering
University of California,
Berkeley
(510) 642-2468
bayen@berkeley.edu

Quinn Jacobson
Research Leader
Nokia Research Center – Palo
Alto
(650) 521 3243
quinn.jacobson@nokia.com

J.D. Margulic
Associate Director
California Center for
Innovative Transportation
(510) 642 5929
jd@calccit.org

7.2 Guest Program

Mobile Century Final Guest Program



MOBILE CENTURY – FEBRUARY 8, 2008

Using GPS Mobile Phones as Traffic Sensors: A Field Experiment

Final Program

10:00 AM *Mobile Century Kick-Off and Press Conference*
11:30 AM *Union Landing Mall, Union City, California*

12:30 PM *Lunch and Presentations*
2:30 PM *290 Hearst Memorial Mining Bldg., UC Berkeley*

4:00 PM *Scientific Seminar*
5:00 PM *290 Hearst Memorial Mining Bldg., UC Berkeley*

CONTACT INFORMATION:
 I.D. Margulicci
 Associate Director
 California Center for
 Innovative Transportation
 (510) 642-5929
 idm@calccit.org

SPONSORS:




PRESENTATIONS:
 I.D. Margulicci, Associate Director, CCIT
The California Center for Innovative Transportation Model
 John Paul Chen, Director, Nokia Research Center, Palo Alto
The Nokia Research Center, Palo Alto: Mobility, Community, and Serendipity
 Paul Wright, Acting Director of CITRIS, A Martin Berlin Professor of Mechanical
 Engineering, University of California, Berkeley
The Center for Information Technology Research in the Interest of Society
 Jeremy Wolstan, Government & Industry Relations, NAVTEQ
NAVTEQ and NAVTEQ Traffic
 Quinn Jacobson, Research Leader, Nokia Research Center, Palo Alto
The technology behind the Mobile Century experiment
 3:00 PM – 4:00 PM – OPTIONAL TOUR OF THE BERKELEY CAMPUS

WELCOME ADDRESS:
 Lisa Alvarez-Cohen, Fred and Claire Saver Professor and Faculty Chair, Civil and
 Environmental Engineering, University of California, Berkeley

GUIDED TOURS OF THE COMMAND CENTER AND MINGLING TIME

INTRODUCTION:
 Paul Wright, Acting Director of CITRIS

TALK:
 Alexandre M. Bayen, Assistant Professor of Civil and Environmental Engineering,
 University of California, Berkeley
Using GPS Mobile Phones as Traffic Sensors: A Field Experiment

CALIFORNIA CENTER FOR INNOVATIVE TRANSPORTATION
 UNIVERSITY OF CALIFORNIA, BERKELEY 3105 BANCROFT WAY, SUITE 500 BERKELEY, CA 94720-3830
 PHONE: (510) 642-4522 FAX: (510) 642-0910 HTTP://WWW.CALCCIT.ORG



MOBILE CENTURY – FEBRUARY 8, 2008

CORRESPONDANCE
mobilecentury@calccit.org

Contact Information

J.D. Margulici
 Associate Director
 California Center for
 Innovative Transportation
 (510) 642 5929
jd@calccit.org

Alexandre M. Bayen
 Assistant Professor
 Department of Civil and
 Environmental Engineering
 University of California,
 Berkeley
 (510) 642-2468
bayen@berkeley.edu

Quinn Jacobson
 Research Leader
 Nokia Research Center –
 Palo Alto
 (650) 521 3243
quinn.jacobson@nokia.com

Using GPS Mobile Phones as Traffic Sensors: A Field Experiment



Final Program

10:00 AM – 11:30 AM *Mobile Century Kick-Off and Press Conference* Union Landing Mall, Union City, California

PRESS CONFERENCE

Alexandre M. Bayen, Assistant Professor of Civil and Environmental Engineering, University of California, Berkeley

Bob Iannucci, Chief Technical Officer and Head of Nokia Research Center, Nokia
 Randell Iwasaki, Chief Deputy Director, California Department of Transportation

GUIDED TOURS OF THE COMMAND CENTER AND MINGLING TIME

12:30 PM – 2:30 PM *Lunch and Presentations* 290 Hearst Memorial Mining Bldg., UC Berkeley

WELCOME ADDRESS

Lisa Alvarez-Cohen, Fred and Claire Sauer Professor and Faculty Chair, Civil and Environmental Engineering, University of California, Berkeley

PRESENTATIONS

J.D. Margulici, Associate Director, CCIT
The California Center for Innovative Transportation Model

John Paul Chen, Director, Nokia Research Center, Palo Alto
The Nokia Research Center, Palo Alto: Mobility, Community, and Serendipity

Paul Wright, Acting Director of CITRIS, A. Martin Berlin Professor of Mechanical Engineering, University of California, Berkeley
The Center for Information Technology Research in the Interest of Society

Jeremy Wolstan, Government & Industry Relations, NAVTEQ
NAVTEQ and NAVTEQ Traffic

Quinn Jacobson, Research Leader, Nokia Research Center, Palo Alto
The technology behind the Mobile Century experiment

3:00 PM – 4:00 PM – OPTIONAL TOUR OF THE BERKELEY CAMPUS

4:00 PM – 5:00 PM *Scientific Seminar* 290 Hearst Memorial Mining Bldg., UC Berkeley

INTRODUCTION

Paul Wright, Acting Director of CITRIS

TALK

Alexandre M. Bayen, Assistant Professor of Civil and Environmental Engineering, University of California, Berkeley

Using GPS Mobile Phones as Traffic Sensors: A Field Experiment

CALIFORNIA CENTER FOR INNOVATIVE TRANSPORTATION
 UNIVERSITY OF CALIFORNIA, BERKELEY · 2105 BANCROFT WAY, SUITE 300 · BERKELEY, CA 94720-3830
 PHONE: (510) 642-4522 · FAX: (510) 642-0910 · [HTTP://WWW.CALCCIT.ORG](http://www.calccit.org)

Short Bios

Alexandre M. Bayen

Alexandre Bayen has been an Assistant Professor in the Department of Civil and Environmental Engineering at UC Berkeley since January 2005. Until December 2004, he worked as the Research Director of the Autonomous Navigation Laboratory at the Laboratoire de Recherches Balistiques et Aerodynamiques in Vernon, France. Pr. Bayen was a Visiting Researcher at NASA Ames Research Center from 2000 to 2003. He received the Engineering Degree in applied mathematics from the Ecole Polytechnique, France, in July 1998, the M.S. degree in aeronautics and astronautics from Stanford University in June 1999, and the Ph.D. in aeronautics and astronautics from Stanford University in December 2003.

Lisa Alvarez-Cohen

Professor Alvarez-Cohen is the Fred and Claire Sauer Professor of Environmental Engineering and chair of the department of Civil and Environmental Engineering at UC Berkeley. Her research interests are on the microbial degradation of environmental contaminants in natural and engineered systems with application to innovative hazardous waste treatment technologies and groundwater remediation. Professor Alvarez-Cohen recently co-authored a textbook entitled Environmental Engineering Science, available from John Wiley and Sons. She holds a B.A. in Engineering and Applied Science from Harvard University and an M.S. and Ph.D. in Environmental Engineering and Science from Stanford University.

Bob Iannucci

Dr. Bob Iannucci is Chief Technology Officer of Nokia and Head of Nokia Research Center (NRC). As Head of NRC he has initiated many open research partnerships with universities and beta communities, and has been instrumental in assembling open, flexible and geographically focused research teams to take NRC's tradition of breakthrough innovation into exciting new areas. Throughout his career Dr. Iannucci has consistently shown an ability to balance scientific and technological understanding with business insight. In his role as CTO, he is intent on bringing new dimensions of renewal to Nokia by continuing to drive its research strategy and innovation processes in new ways and new directions. Dr. Iannucci holds a Ph.D. in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology, and is the author and co-author of two books and several academic papers, as well as five patents.

Randell H. Iwasaki

Randell "Randy" Iwasaki is the Chief Deputy Director of the California Department of Transportation (Caltrans). Iwasaki manages the day-to-day operations of the Department, including an operating budget of nearly \$10 billion and more than 21,000 employees. A licensed civil engineer, he has been with Caltrans for more than 20 years, serving in a number of high profile engineering and management positions. Iwasaki is a member of a public advisory committee for the ITS Caucus for the U. S. Congress. He is also a member of the National Academy of Sciences panel that is looking at impacts to transportation from Global Climate Changes. Iwasaki earned his bachelor's degree in Engineering from California Polytechnic State University, San Luis Obispo, and a Master's in Engineering from California State University, Fresno.

Quinn Jacobson

Quinn Jacobson leads the Mobile Internet Services Systems team at the Nokia Research Center - Palo Alto. His expertise is in architecting and prototyping mobile services, including privacy and security related challenges. Over the last year Quinn has been leading an effort to look at how to construct collaborative services, where individuals can share information in real-time, while preserving privacy. He joined NRC Palo Alto in October 2006. Before joining Nokia, Quinn worked in Intel's Microarchitecture Research Lab, and before that he was the chief architect for Sun Microsystems' UltraSPARC IV family of processors. Quinn holds a Ph.D. in Electrical and Computer Engineering from the University of Wisconsin - Madison.

3 MOBILE CENTURY: GUEST PROGRAM AND LOGISTICS

Short Bios (cont.)

J.D. Margulici

J.D. Margulici is the Associate Director of the California Center for Innovative Transportation (CCIT), at the University of California, Berkeley. He has nearly 10 years of experience in the areas of transportation, software design and marketing, as well as systems engineering and deployment within large organizations. J.D. worked for the French air traffic control agency before becoming the Director of product marketing at Edatis, a Paris-based software company. Prior to joining CCIT, he was a consultant in airport operations with Jacobs Consulting. J.D. holds a B.S., Mathematics and Physics, from the Ecole Polytechnique in Paris, France, and a M.S. in Transportation Engineering from the University of California, Berkeley. He is currently a MBA candidate in the Evening and Weekend program of the Haas School of Business at Berkeley.

John Paul Shen

John Paul Shen is the Head of Nokia Research Center Palo Alto. NRC Palo Alto lab was established in the fall of 2006 and currently houses over 50 researchers pursuing a wide range of research projects on large-scale mobile systems, applications, and services. Prior to joining Nokia in 2006, John was the Director of the Microarchitecture Research Lab at Intel. Before joining Intel in 2000, John was a Professor at Carnegie Mellon University, where he supervised a total of 17 PhD students and received multiple teaching awards. He is currently an Adjunct Professor at Carnegie Mellon University West Campus. John is the author or co-author of numerous published articles and two books, including *Modern Processor Design: Fundamentals of Superscalar Processors*, McGraw-Hill 2005. John received his BS degree from the University of Michigan and his MS and PhD degrees from the University of Southern California, all in Electrical Engineering.

Jeremy Wolstan

Jeremy Wolstan manages the Government & Industry Relations program at NAVTEQ. He has been at NAVTEQ for 5 years, working for 2 years in the Operations group, before moving into Government & Industry in early 2005. At NAVTEQ, Jeremy focuses on VII and other programs related to the wireless distribution of map data and related content, such as traffic and parking location data. He is also responsible for managing NAVTEQ's University Relationship programs. Jeremy's involvement with the Connected Vehicle Trade Association and Connected Vehicle Proving Center had him named a CVTA fellow in 2007. Jeremy holds a B.S. in Business Administration from the University of Southern California.

Paul K. Wright

Paul Wright is the A. Martin Berlin Professor of Mechanical Engineering at the University of California, Berkeley and the Acting Director of the Center for Information Technology Research in the Interest of Society (CITRIS). In addition to his position at CITRIS, Paul Wright serves as the Associate Dean of the College of Engineering, Co-Chair of the Management of Technology Program (a joint program with the Haas School of Business) and Co-Director of the Berkeley Manufacturing Institute. He is a member of the National Academy of Engineering. Director Wright received his B.Sc. and Ph.D. degrees from University of Birmingham, England, in industrial metallurgy. His publications span the areas of metal processing, especially machining, robotics and its applications in flexible manufacturing systems, the development of expert systems for manufacturing, and rapid prototyping. His current research is in the fields of energy scavenging and storage for micro-electronics.

8 Supplemental Tasks

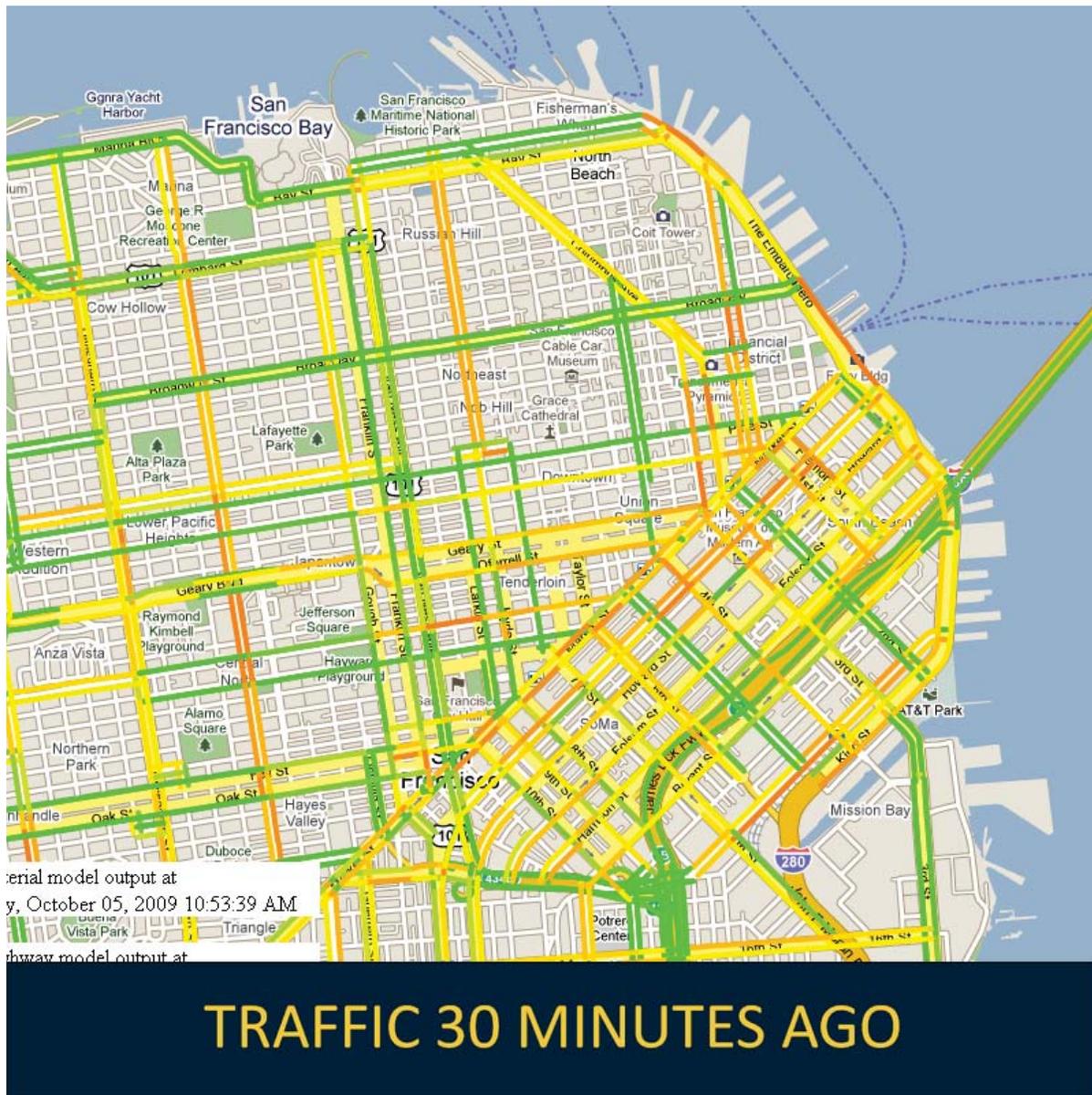
Appendix 8 addresses supplemental tasks included in amendments to the task orders. As noted in Section 1.3, these tasks fall into one of three categories:

- (1) AASHTO presentation
- (2) Mobile Millennium Planning, Design, and Server Development
- (3) Mobile Millennium Arterial Modeling

Herein, work performed in accordance with the AASHTO presentation of Mobile Millennium software is summarized. In addition, the initial stages of work toward tasks (2) and (3) as funded by monies allocated from TO 1021 and TO 1029 are outlined. The ultimate success of tasks (2) and (3) lie outside the scope of Mobile Century, and will be addressed in detail in the forthcoming Final Report for Mobile Millennium.

8.1 AASHTO presentation

The Mobile Millennium system was demonstrated at the October 2009 AASHTO Annual Meeting and Trade Fair in Palm Desert, California. This demonstration required substantial resources, including internet connectivity, computers, and displays to visualize traffic state at the CCIT booth. An example of the visualization is shown below.



A special invitation was extended to encourage participants of the conference to test drive the Mobile Millennium software on their cell phones. Those with Blackberry or Nokia smartphones and serviced on the AT&T network (or select Blackberry models on the Verizon network) with unlimited data plans were able to download the Mobile Millennium application directly to their phones. In addition, Nokia cell phones were available on loan during the conference. Sufficient quantities of phones were available to loan one phone per state as well as an additional 50 phones on a first come, first serve basis.

8.2 Mobile Millennium Planning, Design, and Server Development – Initial Stages

As part of the initial scale-up for Mobile Millennium, a server at CCIT was brought online. This work included the data feeds, servers, the preliminary data base structure with archival capabilities, and a core java code version of the traffic algorithm.

The capabilities of this server were tested during August and September of 2008 in both downtown Berkeley and in San Francisco. Data was sent from the client cellphones and successfully received at the NRC server in Palo Alto and at the CCIT server in Berkeley.

Traffic collection features of the server were completed in August. Data collection was run and performed online. Traffic reconstruction algorithms were linked to data feeds inside the server. The data assimilation algorithm was ready to operate starting from August 10, 2008.

Data assimilation with the first version of the arterial algorithm was ready to operate from August 26, 2008. An end-to-end test was performed to assess the validity of the feeds. This included a San Francisco test drive with 20 cars took on September 3, 2008.

On September 21, 2008, the first New York test took place with a single car to debug the VTLs. Another test took place on September 23, 2008. A field test with 20 cars took place on September 24, 2008. Tests were completed of the live feed and of the VTL settings. The system was set up for VTL data to be fed into the schema and to be archived on September 25, 2008. A post mortem was performed on the collected data. Test failures were analyzed to fix communication flaws in the implementation.²⁹

²⁹ The unveiling of the real time system at the ITS World Congress will be described in the forthcoming report for Mobile Millennium.

8.3 Mobile Millennium Arterial Modeling – Initial Stages

Data for the model. The initial efforts for arterial modeling assumed that VTL-based data would be available from a small fraction of vehicles traveling on the network (about 5% to 10% of the flow). In addition, there were two sources of historical data. The first was from the Mobile Millennium databases, and the second was from Navteq. Navteq provided traffic patterns available for time of day (TOD) and day of week (DOW) at the granularity of one TMC (or about one block).

Goals and assumptions. The goal of these initial efforts was to estimate in real time, both travel times and congestion levels at the granularity of a single link. Data would be received from Nokia / Navteq, and processed in real time. A future tool for visualization of the results was assumed.

Challenges due to variability. Estimation of traffic state on arterials turned out to be extremely difficult. Substantial variability in traffic conditions were observed even for vehicles on the same network at the same time. Variability due to random events such as pedestrians, bad parking, delivery trucks, and construction posed significant challenges. In addition, a flow model such as that used for highways was not directly usable due to complications arising from the influence of traffic signals.

Lack of previous research. A review of the literature revealed a dearth of previous work on arterial traffic state estimation using floating vehicle data at low penetration rates. In contrast, prior studies focused on the use of loop detector data with high penetration rates of probes.

Discretization mismatch. One additional challenge was that input data was VTL-based, while the intended output was link-based. This mismatch caused additional difficulties.

Initial approaches. Although statistical approaches were initially considered, they suffered from complexity and computational intensity. For our application, real-time estimation was a hard requirement. Furthermore, the impending deadline for the Mobile Millennium field operational test exerted additional pressures. For these reasons, a simple moving-average filter was chosen.

Moving average filter with historical data

Step 1: Transform VTL observations into link travel times. Recall that the input data is VTL-based. These VTL observations must be transformed into link travel times. The difficulty in this process is that a pair of travel times measured between the same two points on the network may differ greatly if offset by several seconds. More time is spent at the downstream end of a link if a vehicle stops for a traffic signal.³⁰

Step 2: Update the link travel time estimate. Consider discrete estimation times to be denoted t . For each link in the network, observations at time t consist of elements $y_1 \dots y_n \in Y^t$. Each observation is a travel time for the entire link, and is made available between times t and $t - \Delta_t$, where Δ_t is the estimation window. The travel time estimate TT^t at time t is given recursively by

$$TT^t = \alpha TT^{t-1} + (1 - \alpha) \frac{\sum_{i=1}^n y_i}{n}$$

The parameter α optimizes the weight of previous estimates with respect to real time data. If for example, few real time data are available, then α should be large.³¹

Step 3: Incorporate historical data into the estimate. This process is performed heuristically based on the number of real time measurements. If this number exceeds a threshold, then only real time data is used. If the number is beneath the threshold, then an average is calculated where the weight of the historical value increases as the number of measurements decreases. For this purpose, historical data from both Navteq patterns and from previous data stored on the Mobile Millennium servers are used. If no real time data is available for a particular link, then the historical average is used in its place. The historical model is also used to compute a confidence interval for the estimates produced in step 2, above.

³⁰ Strategies to split measured travel times across links will be explained in the forthcoming report for Mobile Millennium. For now, consider that links are split up into segments. Segments begin and end at VTL locations or at link boundaries. If a pair of VTL measurements is fully contained within one segment, then the assignment is easy. However, when the VTL pair spans two links, a heuristic method is used based on the observation that delays are typically experienced at the downstream end of a link (at a traffic signal). We assume that the time spent on the second link was in free flow and calculate the travel time accordingly. The remaining time is allocated to the first link.

³¹ Substantial efforts were made to optimize α and to learn it in real time, details of these methods will be addressed in the forthcoming report for Mobile Millennium.

Additional measures

Confidence value on the estimates. The confidence value is an integer that takes one of three possible values where 1 indicates high confidence, and 3 indicates low confidence. This confidence value is computed from a combination of three different factors:

- Variability in the estimates. If estimates fluctuate wildly, the confidence value is reduced.
- Availability of real time data. The more data, the higher the confidence.
- Agreement with historical data. When little real time data is available, the confidence value depends on how close the estimate is to that from the historical model. The greater the agreement, the higher the confidence.

Congestion indicator. In addition to link travel time, and confidence, a congestion indicator is computed. The congestion indicator c^t is a real value, greater than one, and computed as

$$c^t = \frac{TT^t}{TT_{ff}}$$

where TT_{ff} is the free-flow travel time on the link. The free flow travel time is computed in accordance with the properties of the underlying network. Details will be furnished in the forthcoming report for Mobile Millennium.

Example of the calculation of free flow travel time.

The purpose of this section is to demonstrate that something as fundamental as determining a link travel time in free flow conditions is not as trivial as it might seem. Consider that this particular task is not appropriate for a human being to perform by hand for each link of a city network. In contrast, this task must be automated for the map data that happens to be available. In our case, the following procedures were implemented:

- Compute the travel time if traveling at the speed limit.
- Search for the presence of a traffic signal at the downstream end of the link in question.
- If a signal is present, search for the kind of signal (stop sign or traffic light).
- If the traffic signal is a traffic light, then search for tabular values of the signal parameters (cycle and red time), if this data is available.³²
- If no traffic signal information is available, choose reasonable values for the signal parameters depending on the importance of the road and the shape of the intersection.
- Compute an expected delay under free flow conditions. Even if there is only one vehicle on the road, the vehicle experiences delay due to the presence of a signal. This delay is composed of:
 - Deceleration time τ_d
 - Stopping time τ_s
 - Acceleration time τ_a
- Basic mechanics provides approximate deceleration and acceleration times for typical rates of deceleration and acceleration. The stopping time is equal to a couple of seconds for stop signs and is computed from the traffic parameters for a traffic light.
- For a vehicle that stops at a light, expected delay is $R/2$ where R is the duration of the red time (assuming uniform arrivals).
- For a vehicle traveling on a link with a traffic light, the probability of stopping under free flow conditions is R/C where C is the duration of the cycle. The total expected delay for a vehicle traveling on a link with a traffic signal is

$$\tau = \frac{R}{C} \left(\tau_d + \frac{R}{2} + \tau_a \right)$$

- For a vehicle traveling on a link with a stop sign, the expected delay under free flow conditions is given by

$$\tau = \tau_d + \tau_s + \tau_a$$

- Add the expected delay τ to the travel time experienced when driving at the speed limit to compute the free flow travel time.

³² This data was available for New-York City, from a difficult-to-read table. It took three undergraduate man-months to put the data in a useable format.