CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

VII California: Development and Deployment— Lessons Learned

Jim Misesner, et al.

California PATH Research Report
UCB-ITS-PRR-2009-35

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation, and the United States Department of Transportation, Federal Highway Administration.

Final Report for Task Order 5217

August 2009

ISSN 1055-1425

CALIFORNIA PARTNERS FOR ADVANCED TRANSIT AND HIGHWAYS

Task Order 5217 Final Report

VII California: Development and Deployment
Lessons Learned

Authors

Jim Misener
Susan Dickey
Joel VanderWerf
Ashkan Shrafasaleh
Kang Li
Han-Shue Tan
Meng Li
Zhi-jun Zou
Fanping Bu
Ching-Ling Huang
Guan Xu
Steven Shladover
Tom Kuhn
Matt Barth
Michael Todd
Wei-Bin Zhang

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Karl Wunderlich of Noblis for providing the TCA software and advice about its use. We also wish to thank our PATH colleagues Kun Zhou and Yoa Mao for their support and discussion. The authors would like to thank Toyota ITC and Volkswagen of America ERL colleagues for their significant inputs in developing this curve over-speed warning system in the VII-CA project. The authors are also grateful to NAVTEQ for providing digital maps.

The authors would like to thank Miguel Egea, James Lao, Lindy Cabugao and numerous other Caltrans District 4 personnel who helped us with installation, repairs, troubleshooting and understanding traffic signal controllers. Special thanks go to Jim Arnold with FHWA for consistently providing updated information on the status, cost, and implementation details associated with the HA-NDGPS.

Additionally, the authors would like to acknowledge San Francisco Department of Parking and Traffic for the use of their facilities. We would also like to thank Bart Duncil for work in designing the gumstix assemblies and initial testing of the system, Bernard Liang for his work developing GPS utilities, Shel Leader for organizing the work with the San Francisco roadside equipment and for setting up vehicle equipment and antennas, and Thang Lian and Tim Duff for many hours spent driving around the block.

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. Finally, the mention of commercial products, their sources, or their uses in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products.

ABSTRACT

This PATH Research Report covers the (Vehicle-Infrastructure Integration) VII California Development and Deployment (Task Order 5217) efforts from October 2005 – December 2007. Because TO 5217 is followed by the continuation TO 6127, it is a compendium of very applications-oriented research to date as well as a final report to TO 5217.

It is organized to impart the very specific and generally very pragmatic implementation details first, beginning with an introduction (Section 1), description of VII hardware, general network and installation (Section 2), then progressing to a more detailed description of the network and operating software (Section 3) and finally to applications in development and prospective applications (Section 4). Because it is not yet a final report but rather a 'research in progress' report, it does not comprehensively address every task in the VII California family of task orders; specifically, several of the applications described in Section 4 are represented as works in progress (as they are at this writing), and the on-ramp metering study is not yet reported. Again, the purpose, of this report is to provide the reader with an understanding of how VII California was implemented.

KEYWORDS

Vehicle Infrastructure Integration, Curve Overspeed, Traffic Probe Data, HA-NDGPS, Roadside Equipment, Onboard Equipment, DSRC, RSSI

EXECUTIVE SUMMARY

This PATH Research Report covers the (Vehicle-Infrastructure Integration) VII California Development and Deployment (Task Order 5217) efforts from October 2005 – December 2007. Because TO 5217 is followed by the continuation TO 6127, it is a compendium of very applications-oriented research to date as well as a final report to TO 5217.

The report is organized to impart the very specific and generally very pragmatic implementation details first, beginning with an introduction (Section 1), description of VII hardware, general network and installation (Section 2), then progressing to a more detailed description of the network and operating software (Section 3) and finally to applications in development and prospective applications (Section 4). Because it is not yet a final report but rather a 'research in progress' report, it does not comprehensively address every task in the VII California family of task orders; specifically, several of the applications described in Section 4 are represented as works in progress (as they are at this writing).

Moreover, the report addresses a series of lessons learned as results of encountering issues and complications, and provides future insights to overcome to these impasses. These insights will potentially provide invaluable knowledge for future implementation in California and throughout the nation.

As highlighted in Section 2, the installation of roadside equipment (RSE) requires a large degree of cooperation between various institutions. Furthermore, cooperation between these institutions must also extend to maintenance and updating hardware.

In Section 3, the report discusses in extensive detail the network and operating software that accompanies the VII California testbed. Various prototype operating software are used in monitoring and troubleshooting the RSEs without the physical intervention of the operator. The report lists several constraints arise from networking problems and discuss several working solutions. Networking problems have been encountered and largely addressed during the course of work; this experience engenders the expectation of more constraints and 'more of the same' but at a larger scale as work progresses to address security, scalability, and conformance to the emerging national VII standard.

Additionally, Section 3 also discusses the details in implementation. PATH has developed a middleware for the exchange of these messages between in-vehicle client software and back-end application servers.

Lastly, the report discusses current ongoing work on several applications of VII, described to further detail within the document:

1. Curve Overspeed Warning System

   The Curve Overspeed Warning System (COWS) is one of the two cooperative active safety applications of VII California. (The other such time-critical safety of life

applications that requires low latency, highly reliable and available vehicle-to-infrastructure and infrastructure-to-vehicle communication is Cooperative Intersection Collision Avoidance Systems (CICAS), not covered in this report.) The COWS system is aimed to integrate on-board sensors, digital map, GPS and the broadcasted information from the RSE to predict safety speed and provide speed advisory or warning messages to the driver.

2. Traffic Probe Data Processing

   The advent of VII offers the opportunity to revolutionize transportation data collection by eventually enabling virtually all road vehicles to serve as traffic data probes, collecting information about their motions and transmitting it to roadside receivers. This has profound implications for the future of transportation planning and operations by providing most, if not all, of the data they have dreamt of having, and probably providing much more data than they have ever had to manage before. The primary challenge could quickly shift from determining how to extract as much information as possible from limited available sources of data to determining how to efficiently identify and extract the most relevant and important information from an avalanche of largely redundant data.

3. Application of VII Data on Real-time Arterial Performance

   The VII communication between cars and roads could benefit both parties and could be very powerful, as it would enable the full vision of intelligent transportation system (ITS). For example, the information transmitted from the roadside to the vehicle could warn a driver that it is not safe now to enter an intersection. On the other hand, vehicles could serve as data collectors and anonymously transmit traffic and road condition information from all major roads which are equipped with such communication infrastructures. Such information would significantly improve the quality and quantity of traffic data, which in turn will help transportation agencies and all other stakeholders to better understand, plan, and manage the transportation system.

4. Intersection Application

   Traffic signal timing is a mature and sophisticated study in its own right. Many existing traffic controllers already run software that supports communication protocols, such as the National Transportation Communications for ITS Protocol (NTCIP) [17] and California's Assembly Bill 3418 (AB3418) protocol, designed to communicate signal phase and loop detector information to traffic management centers. However, these protocols were designed for monitoring and control of actuated and coordinated intersection timings, from a central location over a modem, at a time granularity of seconds. They were not designed to notify vehicles of real-time signal phase count-down, needed for the signal violation countermeasures over a high bandwidth DSRC connection. Traffic signal controller protocol messages also contain information that it is difficult to interpret without access to the timing plan for the intersection developed by the maintaining entity (state or local traffic operations), and do not contain geographical/mapping information needed by vehicles in order to recognize what part of

the signal phase information applies to the vehicle's approach to the intersection. Nevertheless, the experiments in this application have successfully reconstructed sufficient information to broadcast the signal phase count-down needed for prototype vehicle applications.

5. High Accuracy-Nationwide Differential Global Position System

   The U.S. Department of Transportation and U.S. Coast Guard in conjunction with the Interagency GPS Executive Board are currently developing a High Accuracy-Nationwide Differential Global Positioning System (HA-NDGPS), targeted at a variety of transportation-related applications. The HA-NDGPS program provides the capability to broadcast corrections to Global Positioning System (GPS) receivers over long ranges to achieve a better than 10 centimeter (cm) (95 percent of the time) accuracy throughout the coverage area. HA-NDGPS is currently undergoing a research and development phase with a future implementation pending U.S. congressional budget approvals. Implementation of this technology will assist in providing advanced safety features for transportation, including lane departure warning, intersection collision warnings, and railroad track defect alerts. It also could be used for economic enhancements such as precision container tracking and automated highway lane striping.

Again, all work described in this TO 5217 final report is 'work in progress' and therefore it is not a finalized product; that will be done as a result of TO 6127. That being said, this report serves independent value in providing the reader an idea of why VII California was implemented and some of the challenges faced and overcome. Thus, it gives the reader the sense of distance in which VII California has covered and how far it will need to travel to reach its destination, and potentially find an extension of a new road that would lead into a greater development.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS i
ABSTRACT iii
KEYWORDS iii
Vehicle Infrastructure Integration, Curve Overpseed, Traffic Probe Data, HA-NDGPS, Roadside Equipment, Onboard Equipment, DSRC, RSSI iii
EXECUTIVE SUMMARY v
TABLE OF CONTENT ix
LIST OF FIGURES AND TABLES xvii
1. INTRODUCTION 1
2. VII CALIFORNIA TESTBED DESIGN 5
2.1 ELEMENTAL BUILDING BLOCK WITHIN NETWORK: ROADSIDE EQUIPMENT 5
2.2 NETWORK: TRANSPORT LAYER 7
2.3 TESTBED – INSTALLATION, OPERATIONS, AND MAINTENANCE ISSUES 9
2.3.1 RSE Installation 9
2.3.1.1 Institutional Issues 9
2.3.2.2 Hardware Issues 9
2.3.3 Installation Cost 11
2.3.4 Maintenance Issues 12
2.3.4.1 Hardware Maintenance 12
2.3.4.2 Institutional Considerations 14
2.3.4.3 Network Operations and Maintenance 14
2.4 IMPLICATIONS FOR THE FUTURE 14
2.4.1 Hardware 14
2.4.2 Backhaul Options 15
2.4.3 Institutional Issues 15
2.4.4 Summary of Key Lessons Learned 15
2.5 FUTURE DEVELOPMENTS 16
3. PROTOCOLS AND PERFORMANCE 18
3.1 RSE MONITOR 18
3.1.1 Introduction 18
3.1.2 Variables 18
3.1.3 Corrective actions 19
3.1.4 Web Interfaces 19
3.1.4.1 Current reports 20
3.1.4.2 Detail pages 21
3.1.4.3 Full report 22
3.1.4.4 Time-history graphs 22
3.1.4.5 Map of testbed with status indicators 27
3.1.4.6 Technical details 27
3.2 NETWORK: MIDDLEWARE FOR VII-CA MESSAGES 28
3.3 EVALUATION OF DSRC COMMUNICATION PERFORMANCE USING LOW COST PROTOTYPE SYSTEMS WITH MULTIPLE RADIOS PER VEHICLE 30

3.3.1 Introduction 30
3.3.2 Computer and Radio Equipment 31
3.3.3 Data Acquisition Software 32
3.3.4 Baseline Testing 33
3.3.5 Effects of Channel Switching 37
3.3.6 Mobile Radios 47
3.3.7 Urban Intersection Testing 49
3.3.7.1 Packet length and interarrival time 49
3.3.7.2 Data rate and queuing 54
3.3.8 Conclusion 57
4. APPLICATIONS 58
4.1 CURVE OVERSPEED WARNING SYSTEM (COWS) APPLICATION PART I: DYNAMIC ROAD CURVE RECONSTRUCTION USING DIGITAL MAP 58
4.1.1 Introduction 59
4.1.2 Digital Map as a Virtual Sensor for Advanced Driver Assistance System 61
4.1.2.1 Overview of Digital Map and ADAS 61
4.1.2.2 Architecture and Advantages of the ADAS and COWS Enhanced by VII 62
4.1.3 Dynamic Road/Lane Curve Reconstruction Using Digital Map Data 63
4.1.3.1 Conventional Approach ─ Spline 64
4.1.3.2 New Approach ─ Data Pre-filtering and Parametric Curve Fitting 65
4.1.3.2.1 Data Pre-Filtering by Circle Center Search (CCS) and Circle Selection (CS) Algorithms 67
4.1.3.2.2 Parametric Curve-Fitting 68
4.1.4 Curve Overspeed Warning Experiment 70
4.1.5 Conclusion and Future Work 72
4.2 TRAFFIC PROBE DATA PROCESSING FOR FULL-SCALE VII DEPLOYMENT 74
4.2.1 Introduction 74
4.2.2 Background 75
4.2.3 Simulation Model 76
4.2.4. Baseline Probe Data Sampling Approach 77
4.2.5 Probe Data Aggregation Issues 82
4.2.6. Probe Data Aggregation for Binary Data (Windshield Wiper Status Example) 85
4.2.7 Probe Data Aggregation for Location-Critical Event (Dropped Load Example). 85
4.2.8. Probe Data Aggregation for Traffic Speed Estimation 87
4.2.9 Key Findings and Future Direction 92
4.3 APPLICATION OF VII DATA ON REAL-TIME ARTERIAL PERFORMANCE MEASUREMENTS 94
4.3.1 Introduction 94
4.3.2 Data Characteristics 95
4.3.2.1 VII Probe Data 95
4.3.2.2 Traffic Data 97
4.3.2 Models Formulation 98
4.3.2.1 VII Probe Data (VPD) Model 98
4.3.2.2 Model for Downstream Section 98
4.3.2.3 Model for Upstream Section 100

4.3.2.4 Point-Based Detection (PBD) Model 101
4.3.2.5 Fusion Model 104
4.3.3 Model application 105
4.3.3.1 Simulation-Based Testbed 105
4.3.3.2 Preparation of Fusion Model 106
4.3.3.2 Application Results 106
4.3.4 Discussion 107
4.3.5 Conclusion and Future Direction 108
4.4 INTERSECTION SIGNAL CONTROLLERS AND VEHICLE-INFRASTRUCTURE INTEGRATION: PUTTING SAFE INTERSECTIONS TO PRACTICE 110
4.4.2 Signal Phase and Timing Acquisition 110
4.4.2.1 NTCIP 111
4.4.2.2 California AB3418 Protocol 111
4.4.2.3 Non-invasive Current Sniffer 111
4.4.3 Interface to Dedicated Short Range Communication 113
4.4.4 Communications Latency for Signal Phase and Timing Information 115
4.4.5 Communications Latency for Signal Phase and Timing Information 116
4.5 HIGH ACCURACY-NATIONWIDE DIFFERENTIAL GLOBAL POSITION SYSTEM (HA-NDGPS) 118
4.5.1 Introduction 118
4.5.2 Background 119
4.5.2.1 Vehicle-Infrastructure Integration (VII) Program 119
4.5.2.2 Global Positioning System 119
4.5.2.3 Differential GPS and Carrier-Phase GPS 121
4.5.2.3.1. Real-Time Differential Corrections 122
4.5.2.3.2. Wide Area Augmentation System (WAAS) 123
4.5.2.3.3 Real Time Kinematic (RTK) and CORS 124
4.5.2.3.4 United States Coast Guard National Differential GPS System 126
4.5.3 Differential GPS Evaluation 129
4.5.3.1 High-Accuracy Nationwide DGPS (HA-NDGPS) 130
4.5.3.2 WAAS Accuracy 134
4.5.3.3 Accuracy of RTK and Other Systems 134
4.5.4 DGPS and VII CALIFORNIA Integration Architectures 136
4.5.4.1 VII CALIFORNIA Architecture 136
4.5.4.2 Potential Implementation Methodologies 139
4.5.4.2.1 HA-NDGPS Options 140
4.5.4.2.2 WAAS Integration Architecture with VII California 141
4.5.4.2.3 Integrating the CORS Network with VII California 142
4.5.5 Recommendations 145
4.5.5.1 HA-NDGPS Recommended Architectures 145
4.5.5.2. WAAS Recommended Architecture 147
4.5.5.3 Other DGPS Architectures 148
4.5.5.4 RSE Integrated Recommended Architecture 149
4.5.5.5 Implementation Cost and Implementation Summary 150
4.5.6 SUMMARY 151

4.5.7. NEXT STEPS 152
5. CONCLUSION 154
APPENDIX: THROUGHPUT PERFORMANCE MEASUREMENTS INVOLVING DSRC WIRELESS COMMUNICATION CHANNELS 156
A.1 Introduction 156
A.2 UDP Throughput Tests 156
A.3 Comparison of UDP Throughput 157
A.4 Sample Iperf Commands 158
A.5 Probe client and server components 158
A.6 Monitor component 158
A.7 OBU proxy component support for HTTP over DSRC 159
A.8 Process controller (procon) component 159
A.9 Signage injector component 159
A.10 Denso telnet utility 159
A.11 Probe shell 160
A.12 "Remote" mode for control of running components 160
A.13 Dependency-graph based system administration 160
REFERENCES: 162

LIST OF FIGURES AND TABLES

FIGURE 1.1 A Map of the VII California Network (November, 2007) 3
FIGURE 2.1 Roadside and Backhaul Communications Architecture. This allows data to move from the vehicle, through the roadside infrastructure and to the VII California network. 5
FIGURE 2.2 Schematic of VII California RSE Components 6
Figure 2.3 Left and Center: VII California RSE Mounted in Type 332 Cabinet on Caltrans Right of Way; Right: DSRC Radio Enclosure and Antenna Mounted on Ramp Meter Signal Pole at Freeway Onramp. 7
TABLE 2.1 Costs ($) associated with VII California RSE installation. 12
TABLE 2.2 15
Table 3.1 VII Monitor Report 20
Figure 3.1 Example detail page, California at El Camino 21
Figure 3.2 Example full report 22
Figure 3.3 Graph of Packet Loss 23
Figure 3.4 Temperature Graph 24
Figure 3.5 Graph of ssh responsiveness 24
Figure 3.6 Resource graph of an RSE 25
Figure 3.7 Graph of Memory Usage 26
Figure 3.8 Testbed map 27
Figure 3.9 Architecture of Monitor System 28
Figure 3.10 Assembly of three gumstix netduo expansion boards. 31
Figure 3.11 Antenna placements with 8 active OBUs on one vehicle 32
Figure 3.12 Baseline operation, 1480 byte data per packet, 10 millisecond interval between sends, different pairwise runs of the four stations used for desk checking 34
Figure 3.13 Two runs showing interference between two pairwise communication streams causing slightly increased trip time (1480 byte data per packet, 10 millisecond interval between sends on both streams) 35
Figure 3.14 Large increase in trip time due to interference between a higher rate pairwise communications stream (1480 byte packets at 4 ms interval between sends) and a lower rate stream (1480 bye packets at 10 ms interval between sends). Note that the scale for trip time is in seconds; the lower graph shows sequence numbers. 36
Table 3.2 Parameters used for tests of frequent and in frequent channel switching. 37
Figure 3.15 Baseline case: no channel switching, light load. 38
Figure 3.16 Infrequent channel switching, messages sent at 10 Hz. (t1); each channel switch is indicated by a vertical line, at that point the channel changes 39
Figure 3.17 Close-ups, in case of infrequent channel switching (t1) 40
Figure 3.18 In this case, sw2, the send rate is half that of sw1, and round trip time remains low. 41
Figure 3.19 In this case, sw3, the send rate is twice that of sw1, and but the delivery of packets, as shown by sequence numbers remains good and the round trip time remains low. 42
Figure 3.20 Case sw4, with increased overall load and increased channel switching. 43
Figure 3.21 Close up of switching and sequence number loss for case sw4. 44
Figure 3.22 Case sw5 has increased loading and likewise increased latency and packet loss compared to case sw4. 45

Figure 3.23 Close-up of case sw5 shows that the switching times for the different processors have spread still more, and the latency has increased to almost 25 ms. 46
Figure 3.24 Eight radios, each sending 200 byte packets at 10 ms intervals, with background channel switching twice a second, from van parked within range of Richmond Field Station RSU. 47
Figure 3.25 Eight radios, each sending 200 byte packets at 10 ms intervals, with background channel switching twice a second, during a short ride on a test track at Richmond Field Station. 48
Figure 3.26 Location of RSU at 6th and Brannan in San Francisco. 49
Figure 3.27 Plot of average of remote and local RSSI by GPS locationfor all received packets in runs at 6th and Brannan. 50
Figure 3.28 Probability of successful round trip transmission, as a function of the interval between packet sends, for 1, 2, 4 and nodes, with varying packet sizes of 200, 400 and 800. 52
Figure 3.29 Time (in seconds) until the first 8000 bytes has been transmitted, as a function of the interval between packet sends, for 1, 2, 4 and nodes, with varying packet sizes of 200, 400 and 800 for tests at 6th and Brannan. 53
Figure 3.30 Location of RSU at 3rd and Townsend in San Francisco, and direction of test runs. 54
Figure 3.31 Transaction delay for 8000 byte transaction as a function of data rate, for 9 nodes, packet size 200 bytes, with various testing routes, 10 millisecond interval between packet sends. 55
Figure 3.32 Transaction jitter (standard deviation of delay) as a function of data rate, for 9 nodes, packet size 200 bytes, with various testing routes, 10 millisecond interval between packet sends 55
Figure 3.33 Transaction delay for 8000 byte transaction as a function of data rate, for 9 nodes, packet size 1400 bytes, with various testing routes, 10 millisecond interval between packet sends. 56
Figure 3.34 Transaction jitter (standard deviation of delay) as a function of data rate, for 9 nodes, packet size 1400 bytes, with various testing routes, 10 millisecond interval between packet sends. 56
Figure 3.35 Illustration of regions of performance for successful transmission 57
Figure 3.36 Illustration of regions of performance for transaction times. 57
Figure 4.1 Functional Components of a COWS system 60
Figure 4.2 Concept of ADAS Enhanced by VII 62
Figure 4.3 (a) Road curve fitted by splines (b) Radius of curvature derived from splines 64
Figure 4.4 Concept of Circle Center Search algorithm 66
Figure 4.5 Proposed road curve reconstruction method 66
Figure 4.6 Exit ramp from northbound I-580 to Bayview Ave. (Google Map) 67
Figure 4.7 (a) Preliminary Curvature Estimates by CCS Algorithm (b) Preliminary Estimates of Curvature Radii by CCS Algorithm 68
Figure 4.8 (a) Optimized estimates of road curvatures through parametric curve-fitting (b) Optimized estimates of curvature radii 69
Figure 4.9 On ramp from Marsh Rd. to southbound US 101 (Google map) 69
Figure 4.10 (a) Preliminary curvature estimates by CCS algorithm (b) Preliminary estimates of curvature radii by CCS algorithm 70

Figure 4.11 (a) Optimized road curvature estimates through curve-fitting (b) Optimized curvature radius estimates 70
Figure 4.12 Example of COWS Field Test at the RFS Test Track 71
Figure 4.13 Vehicle speed and GPS condition when approaching the curve (when the GPS is in outage, the resultant HDOP is set to zero by the GPS receiver) 72
FIGURE 4.15 Case Study Network for Probe Vehicle Simulations. 76
FIGURE 4.16 Snapshot Sampling at a Busy Intersection. 78
FIGURE 4.17 Examples of Delays between Probe Snapshot Sampling and Message Uploading to RSE. 79
FIGURE 4.18 Cumulative Distribution of Snapshot Latencies for Events Just Within and Just Without the RSE Range 81
FIGURE 4.19 Snapshot Latencies for Alternative Broadcast Protocols. 82
Table 4.1 Probe Data Aggregation Approaches, Based on Data Types and Applications 83
FIGURE 4.20 Probe Sampling of Wiper Status to Detect Rain at Five RSE Locations. 86
FIGURE 4.21 Average Speeds Derived from Five Minutes of Probe Snapshots. 88
FIGURE 4.22 Effects of Probe Snapshot Sampling on Speed Estimates at 50-100 m from Intersection. 89
FIGURE 4.23 Aggregation of complete vehicle trajectories at 20% market penetration 90
FIGURE 4.24 Aggregation of probe snapshots at 20% market penetration 90
FIGURE 4.25 Raw simulation speed profile averaged over 2 s and 10 s intervals and 20% market penetration averaged over 10 s 91
FIGURE 4.26 Simulation snapshot data averaged over 2 s and 10 s intervals and 20% market penetration averaged over 10 s. 92
FIGURE 4.27 VII Probe Data Collection Process 96
FIGURE 4.28 Typical snapshots distribution near around downstream section 99
FIGURE 4.29 Performance of the Cruise Speed Estimation Model 102
FIGURE 4.30 Delay Calculation Model with Typical Semi-actuated Control Layout 104
FIGURE 4.31 The Structure for the Neural Network Fusion Model 105
FIGURE 4.32 Simulation network and RSEs deployment 106
FIGURE 4.33 Model Performance Comparisons for a NB Link 106
TABLE 4.2 Model Performance Comparison 107
FIGURE 4.34 Installation and operation of “clamp-on” signal current sniffer at VII California intersection. 112
FIGURE 4.36 Modular design of signal phase acquisition and broadcast system 113
FIGURE 4.37 Types of information that must be specified to configure signal phase and timing broadcast. 114
FIGURE 4.38 Difference in transition time between AB3418 and sniffer acquisition, run simultaneously 116
Table 4.3 Summary of GPS Error Sources. [Trimble, 2006]. 122
Table 4.4 Summary of GPS signal correction methods. [Magellan, 2006]. 123
Figure 4.39 Recently installed WAAS Ground Uplink Facility in Napa California [SatNav,2005]. 124
Figure 4.40 USCG NDGPS broadcast site [FHWA, 2005]. 126
Table 4.5 Status of California NDGPS broadcast stations [USCG NAVCEN, 2006]. 127
Figure 4.41 Planned and Operating NDGPS sites [Wolfe et al., 2003]. 128

Figure 4.42 Projected NDGPS coverage [Wolfe et al., 2003]. 128
Figure 4.43 HA-NDGPS broadcast station hardware architecture [USCG NAVCEN, 2001]. .. 131
Figure 4.44 Hagerstown, HA-NDGPS broadcast station rack and receiver antennas [FHWA,2002]. 131
Figure 4.45. HA-NDGPS evaluation of compaction of RTCM 18/19 message types [FHWA,2005]. 132
Figure 4.46 HA-NDGPS real time rover data compared with post processed carrier phase [FHWA, 2002]. 133
Figure 4.47 Independent accuracy test of WAAS capable receiver [Wilson, 2006]. 134
Figure 4.48 HA-NDGPS rover results [FHWA, 2005]..... 136
Figure 4.49 Positioning error for uncorrected (green) and corrected (black) local differential GPS system [Du et al., 2006]. 136
Figure 4.50 VII California Field Operational Test Location [VII CALIFORNIA, 2006a]. 138
Figure 4.51 RSE cabinet and wireless communications antenna installation [VII CALIFORNIA, 2006b, c]. 139
Figure 4.52 VII CALIFORNIA proposed RSE architecture [VII CALIFORNIA, 2006a]. 139
Figure 4.53 WAAS base stations showing Fremont as the closest to the Northern VII California Field Operational Test [SatNav, 2003]. 142
Figure 4.54. California CORS stations with RCTM 2.3 carrier phase corrections [CSRC, 2006]. 143
Figure 4.55 Real time carrier phase DGPS base station utilized for RTK applications [CSRC,2006]..... 144
Figure. 4.56 Proposed WAAS architecture for potential VII CALIFORNIA implementation. . 147
Figure 4.57 Proposed RTK architecture for potential VII CALIFORNIA implementation. 148
Figure 4.58 Proposed RSE integrated DGPS architecture for potential VII CALIFORNIA implementation. 149

1. INTRODUCTION

This PATH Research Report covers the (Vehicle-Infrastructure Integration) VII California Development and Deployment (Task Orders 5217 and 6217) efforts from October 2005 – December 2007. It is not a final report (although it is expected that elements from this research report will be repeated in the final report); rather, it is an organization and report of collective and very applications-oriented research to date.

It is organized to impart the very specific and generally very pragmatic implementation details first, beginning with an introduction (Section 1), description of VII hardware, general network and installation (Section 2), then progressing to a more detailed description of the network and operating software (Section 3) and finally to applications in development and prospective applications (Section 4). Because it is not yet a final report but rather a 'research in progress' report, it does not comprehensively address every task in the VII California family of task orders; specifically, several of the applications described in Section 4 are represented as works in progress (as they are at this writing), and the on-ramp metering study is not yet reported. ***Again, the purpose, of this report is to provide the reader with an understanding of how VII California was implemented.***

Let us begin with the question, "Why VII California was implemented?" It is important to begin with an understanding of a significant impetus: the national VII "movement". The VII system in the United States would be enabled by roadside wireless hotspots generated by Dedicated Short Range Communication (DSRC) transceivers operating within 75 MHz of free, FCC-licensed bandwidth near 5.9 GHz [1]. These transceivers are incorporated in Roadside Equipment (RSE), which are connected by edge backhaul communications into a network architecture that addresses security, privacy and other design considerations (FHWA, 2005). Applications, standards and architecture are under active development. The goal of these efforts is a roadside-based network delivering low-latency, highly reliable data communications to support safety and mobility services to users. The coverage would be extensive: the proposed VII system has the potential to cover all urban roads within 2-minute travel times, 70% of all signalized intersections in 454 urban areas and up to 15 new million vehicles per year would be DSRC- and therefore VII-equipped (Cops, 2006).

The VII initiative is intended to lead to nationwide deployment of cooperating vehicle and infrastructure devices, producing an integrated transportation data network of unprecedented scope and complexity. The fully deployed system would include onboard equipment (OBE) installed on new vehicle manufactured after a specified date and roadside equipment (RSE) installed at all signalized intersections in major urban areas, at primary intersections in other areas, at all highway interchanges and along major intercity and many rural highways (representing about 240,000 RSE locations nationwide). This network would make it possible to

collect transportation operations data of great breadth and depth, to implement sophisticated traffic safety systems, and to manage traffic flows with previously unthinkable precision.

To assess the feasibility and desirability of VII, many technical, economic and institutional questions need to be answered. Considering the magnitude of the commitments that will be needed by a wide range of public and private sector organizations in order for VII to be deployed, these questions must be answered very convincingly before deployment can proceed. The process of developing the answers is just beginning, and is likely to require considerable time and effort.

Lessons for deployment might be learned from regional efforts, notably in Florida, Michigan and California (Florida DOT, 2004, Michigan DOT, 2007, Misener and Shladover, 2006) and most certainly in the forthcoming US DOT VII Proof of Concept experiments (Henry, 2007). Certainly, California is at the nexus of VII need and innovation. The same reasons to create a national VII program are acutely recognized in California: an obligation to better manage the safety and productivity of our highway system, and the understanding that a fusion of public, automotive and other private sector innovations can make this a reality.

However, California and regional stakeholders have specific transportation infrastructure, operating policies and needs than what may be universally addressed with a national VII program. These needs have led to a partnership between Caltrans and the San Francisco Bay Area Metropolitan Transportation Commission. Caltrans and MTC are addressing these needs with a multiyear effort to develop, demonstrate and deploy VII in a key corridor in Northern California. This corridor is comprised of roughly ten-mile segments of two routes North of Palo Alto and South of the San Francisco Airport. It encompasses two highways: State Route 82 and US 101. We note that this selection addresses both a high-volume freeway (US 101) and a major arterial complete with signalized intersections (SR 82), shown in Figure 1.1:

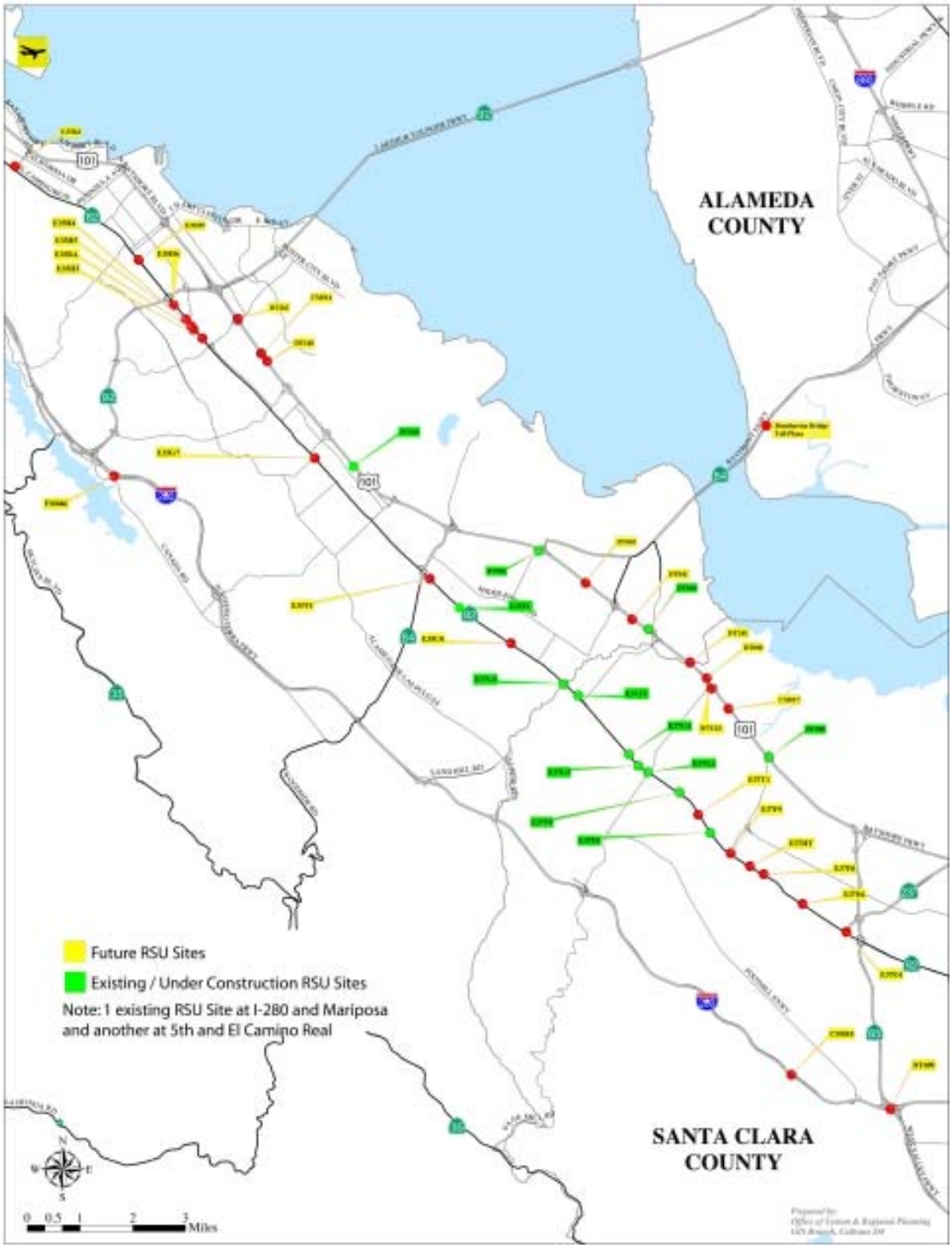


FIGURE 1.1 A Map of the VII California Network (November, 2007)

Overall, Caltrans and MTC aim to:

- Evaluate exemplar public use cases from which we can generalize VII feasibility;
- Evaluate institutional, policy and public benefit issues;

- Explore wireless communication deployment issues and options;
- Resolve key technical issues involving implementation and operation;
- Assess implementations of the VII infrastructure, architecture and operations; and
- Support private sector evaluation interests.

In support of *VII California*, PATH has conducted this project, VII California Development and Demonstration Deployment (VIIC D^3 or simply D^3) where the testbed was built and operated and where deployment and applications issues are investigated. Why? Because some of the technical questions can be addressed through analysis and simulation studies, while others need field testing under realistic conditions. Some of the most challenging questions, which will require the largest investments of time and effort, involve the scalability of system performance in advancing from small numbers of OBEs and RSEs to large numbers of OBEs for each RSE and then to large numbers of RSEs as well. Other questions involve the practical aspects of implementing RSEs in the real world of roadway operations, which can only be answered by starting to install and operate those RSEs and learning about the physical and institutional constraints, the failure modes and the maintenance needs that are encountered there. These have important ramifications for the design of the VII technology and for the long-term costs of deploying, operating and maintaining it in the field.

This report describes the lessons learned in the all phases of the PATH effort.

2. VII CALIFORNIA TESTBED DESIGN

2.1 ELEMENTAL BUILDING BLOCK WITHIN NETWORK: ROADSIDE EQUIPMENT

The hardware experience accrued in developing the VII California testbed focuses on the RSE and its connection to the existing Caltrans roadside appurtenances and their function. The RSE is the basic infrastructure building block for VII. Each RSE serves as an in-the-field gateway between locally transmitted vehicle data and the roadside communications infrastructure. At top level, an RSE is a computer with a radio transceiver and an antenna. The computer must have sufficient processing and storage capability to run a gateway application between its two network interfaces, DSRC and back-haul (to the Traffic Management Center and other servers), and to run additional local safety processor software, with data filtering, buffering, aggregation, and formatting, as needed, and illustrated in Figure 2.1.

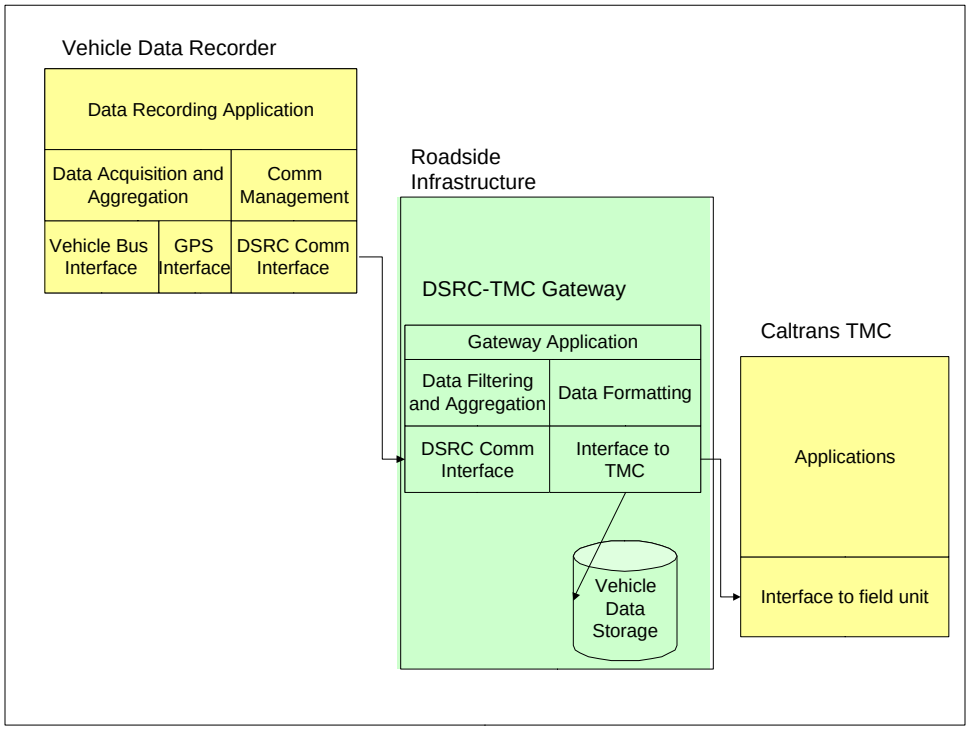


FIGURE 2.1 Roadside and Backhaul Communications Architecture. This allows data to move from the vehicle, through the roadside infrastructure and to the VII California network.

The RSE is therefore between the vehicle and the roadside backhaul network (which might already exist), providing the necessary interface. Physically, it could sit in Type 332 cabinet, and it would consist of a wireless transceiver (with antenna) atop the cabinet and a computer within the cabinet, with connections.

The initial VII California computer was designed to explore a variety of applications and therefore consists of a PC-104 stack with processor and a PCMCIA card with connection to the

Wireless Access to Vehicular Environments (WAVE) DSRC radio and antenna. To fit within the Type 332 cabinet, form factor is low: approximately 6" x 6" x 12". Power is standard 120V AC, with low current draw. A schematic of the VII California connected RSE, residing within a National Electric Manufacturer's Association (NEMA) environmentally rated traffic controller cabinet, is shown in Figure 2.2.

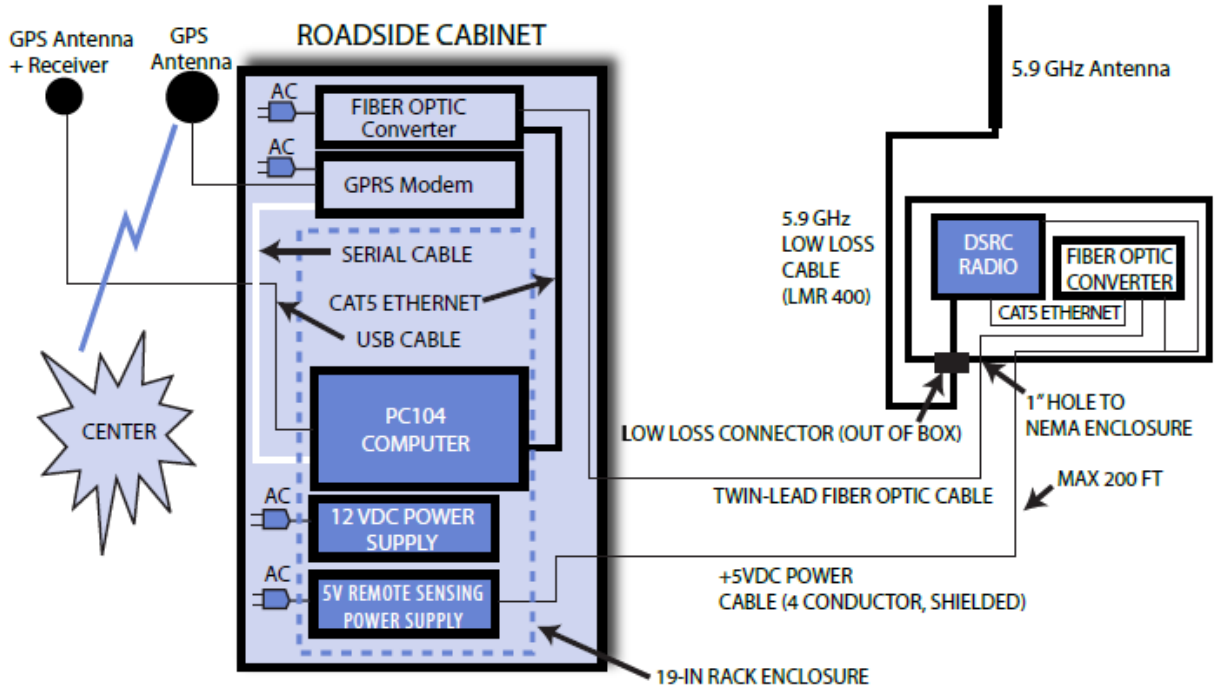


FIGURE 2.2 Schematic of VII California RSE Components

The hardware is split into two main clusters. One resides in the roadside cabinet and the second in a separate weatherproof enclosure. This design minimizes the length of the 5.9 GHz antenna cable to maximize signal strength. Some installations require a separation of up to 200 feet between the existing roadside cabinet and the location where the 5.9 GHz antenna needs to be in order to provide coverage of the approaching roadways. Figures 2.3 illustrate a typical RSE installation.

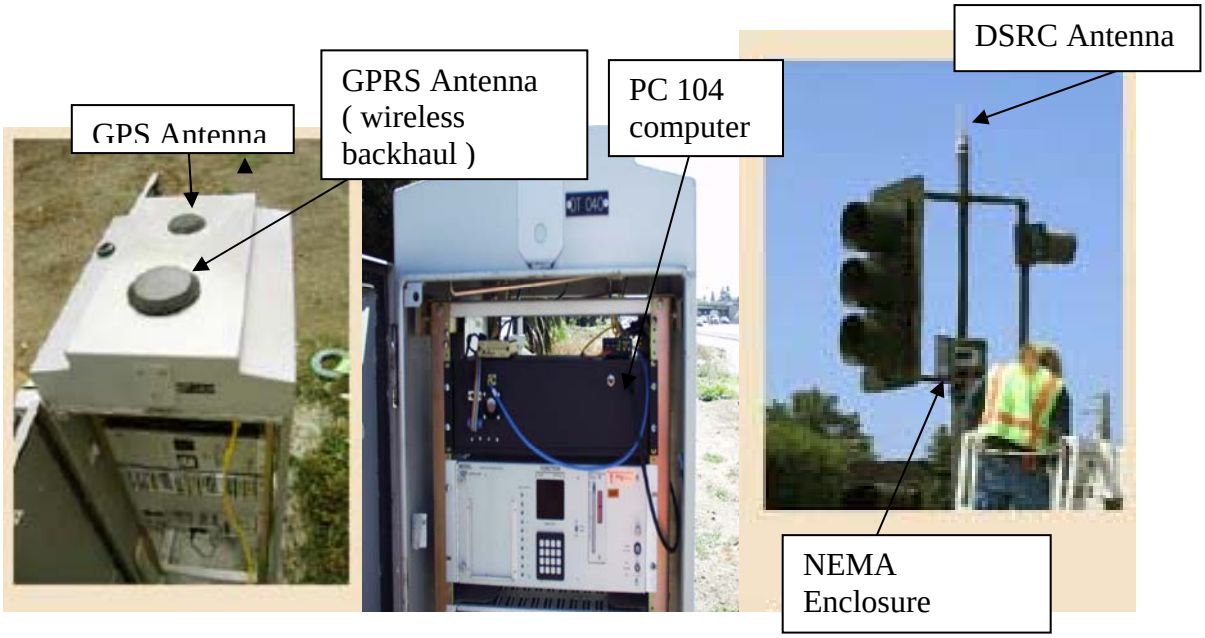


Figure 2.3 Left and Center: VII California RSE Mounted in Type 332 Cabinet on Caltrans Right of Way; Right: DSRC Radio Enclosure and Antenna Mounted on Ramp Meter Signal Pole at Freeway Onramp.

2.2 NETWORK: TRANSPORT LAYER

A major goal of the VII California testbed is to support an initial set of applications, including probe vehicles, public information providers, and bidirectional non-public services (such as navigation and tolling). The testbed has been used as a development laboratory, with experiments and small-scale demonstrations of these applications conducted with public and vehicle manufacturer partners, starting with the 2005 ITS World Congress and continuing through the TRB ITS, Vehicle-Highway Automation and Traffic Signal Systems committee meetings and to the VII Coalition and RITA Administrator demonstrations in 2007. Heavy use of this testbed has exposed a number of issues for future VII implementations and has motivated development of a message transport layer built on top of IP networking, a set of VII application servers, in-vehicle libraries, and administration tools. The characteristics of this software are best examined against the background of the constraints on the project:

1. The compressed development schedule has required the use of prototype-quality hardware. The Denso DSRC Wireless Access for Vehicular Environment (WAVE) *(1)* Radio Modules used initially in the RSE is particularly limited: they cannot transmit packets to more than four hosts on their wired side and they cannot assign dynamic IP addresses to vehicles. Firmware for several components (including DSRC radios and cellular modems) evolved to new generations during the project.
2. The backhaul network is heterogeneous, ranging from cellular modem to dedicated T1 landline. Hence, not all sites can support all services. Furthermore, this diversity

3. The inherent unreliability of mobile wireless networks requires a degree of robustness. Packet loss is likely in an environment with nodes traveling at high speeds with limited range, line-of-sight occlusions, and RF interference. No application can expect to have long-lasting, consistently available connections.
4. Application messages can be long. Probe messages can be up to 64 KB with multiple snapshots.
5. The transport layer should present VII California application code with an interface for communication between vehicles and servers that is message oriented (like UDP, User Datagram Protocol) but has inherent quality of service controls (to some extent like TCP, Transmission Control Protocol). It must be somewhat reliable, but not at the cost of excessive use of the channel.
6. Information from the vehicle must be kept as private as possible.

Within these constraints, the VII California transport layer succeeds in several ways:

1. The first packet from the vehicle to the server is a data packet; there is no delay for dynamic IP address assignment (DHCP) association, internet transmission control protocol (TCP) session initialization, or other handshaking. The reduction in delay increases the chance that a return message (if any) will be received before the vehicle goes out of range.
2. Message reception probability is robust in the face of short-term failures. The transport layer can detect the loss of a fraction of the packets in a message and retransmit the missing fragments. If packet loss is not extensive, this will quickly and efficiently complete the message transmission. Unlike TCP, if no fragments arrive, no bandwidth is wasted resending them, and no bandwidth is wasted on acknowledgment packets.
3. The suite includes a proxy which runs on the vehicle side to insulate vehicle code from the complexity of the wireless user datagram protocol (UDP protocol) and provide it with a TCP socket interface for sending and receiving messages.
4. On the server side, similarly, the interface to the transport layer is a TCP socket, with multiple vehicle sessions multiplexed in one TCP session using transaction IDs.
5. Message addressing is encapsulated across the wireless link to overcome the Denso limitations.
6. No identifying information from the vehicle (outside of the application payload) propagates farther than the roadside or is stored anywhere.

Certain aspects of a complete VII architecture are not fully addressed by this prototype software, including security, scalability, and conformance to emerging national VII standards (IEEE, 2007; SAE, 2006; Kurihara, 2003).

2.3 TESTBED – INSTALLATION, OPERATIONS, AND MAINTENANCE ISSUES

2.3.1 RSE Installation

2.3.1.1 Institutional Issues

Like any other new addition to existing systems, the inclusion and acceptance of RSEs as a standard traffic engineering device presents a series of challenges. The main challenge is to get the local agencies to buy into its usefulness for their constituencies. In the case of highways, the potential sites are normally the property of state DOTs and this single ownership makes it easier for approval purposes. In case of intersections, the owner may be the city, county, or State DOT, making the approval process more involved.

The original design of the RSE and any changes thereafter need to be approved by the infrastructure owner prior to any installation. This includes the approval from the owner agency's electrical, structural, and maintenance groups. The design should conform to specifications used by owner agency.

For the electrical group, the power requirements, type of components, and weather and temperature resistance are important issues. For the structural group, the size and weight of the components as well as the manner by which they are attached to the signal poles and mast arms are important. For Maintenance, the ease of installation and follow-up maintenance are most important.

In our case, all of our installations are on Caltrans rights of way and attached to Caltrans infrastructure. Our design and installations have complied with Caltrans Standard Specification (California DOT, 2006) and Caltrans Standard Plans (California DOT, 2006). Also, since Caltrans maintenance crews have done the installations, everything has conformed to Caltrans safety and standard practices.

Our original design and follow-up changes to it have been rigorously evaluated by Caltrans electrical and structural engineers. There is no specific limit to the size and weight of the equipment that is attached to signal poles and other infrastructure parts, but the structural engineers review each proposed equipment installation based on the weight and size of the equipment and its exact attachment spot.

2.3.2.2 Installation Process – Hardware Issues

The installation process consisted of surveying potential sites, pre-installation infrastructure work if needed, physical installation, and on-the-spot testing and troubleshooting upon completion of the installation.

2.3.2.2.1 Surveying potential sites

Multiple survey trips were taken by engineers and researchers and the infrastructure owner (Caltrans) to identify potential sites for RSE installations from among every intersection and controller cabinet along the freeways and arterials of our corridor. This surveying effort included

examining:

Line of sight issues and potential placements for DSRC antennae: DSRC radio antennas are specified to have a range of up to one mile in every direction. It is necessary for them to have a clear line of sight with the OBEs in the vicinity of the RSE. They also must be placed a minimum of 14 ft above the road surface in order to clear the top of large trucks and transit buses.

For installations along the freeways, the selected site might already have an ideal location for the DSRC antenna. For example, it might be placed atop the freeway ramp meter signal poles or mast arms, the nearby over-crossing structure, or another infrastructure feature that is at least 14 ft above the road surface and can be used for this purpose. If not, a pole that is at least 14 ft high can be mounted next to the freeway controller cabinet, and the DSRC antenna can be attached atop it.

For intersection installations, it is imperative that the DSRC antenna be placed at a position where it has a clear line of sight to all intersection approaches relevant to the intended intersection safety applications (Cohen, et al, 2007). In many intersections, this antenna can be affixed to the mast arms or atop signal poles. The DSRC antenna cable must not be too long, or the signal strength losses will not be acceptable, and may obviate the nominal (6 or 9 dB) gain of the antenna. In our VII California design and for the type of cable we have selected, the maximum DSRC antenna cable length is 20 ft. The low loss cables we typically use have attenuation (loss) numbers that are in range of 10.8 to 26.4 dB per 100 foot length at 5.8 GHz.

Placement of RSE**:** The exact placement of each component of the RSE needs to be planned, including the selection of wire routes to components. One major factor to be considered is how much space is available inside the conduits for RSE wires. Many existing conduits are already full of cables and cannot be used, and some may also be full of water and mud.

Communication cables are prohibited from running inside the power conduits due to safety concerns and in accordance with safety codes. The National Electrical Code (National Electrical Code, 2002) prohibits installing power and broadband conductors in the same conduit. A few locations have dedicated communication conduits that can be utilized for communication cables. The enclosures that are exposed to the environment and are attached to the poles need to be NEMA specified, and should be as small and light weight as possible to minimize concerns of the structural engineers about their weight and seismic, wind and snow loads.

Space availability in controller cabinet**:** Controller cabinets provide a safe, weatherproof place for RSE components as well as power. Space availability within the cabinet - or within the Uninterrupted Power Supply (UPS) cabinets that are attached to the fully-packed controller cabinets to provide extra space- needs to be determined prior to installation of the RSE. In freeway installations, there is usually enough space in the controller cabinet. In intersection installations, cabinet space availability is a more serious concern. The simple solution to address this issue is to add an UPS cabinet to the controller cabinet.

Investigation of backhaul: For backhaul, availability of an existing telephone line in the cabinet as well as the proximity of "phone demarcation boxes" to the RSE site should be determined

prior to the installation. In some locations, such as cabinets for Field Master controllers along major city streets or arterials and closed circuit TV installations along the freeways, phone lines are already connected to the controller in the cabinet. If the telephone landline option is available, then a T1 line provides more upload and download capacity than a dial-up line. If the landline does not exist or is too far away from the cabinets, then wireless options should be investigated. Even though the VII California testbed is located in a heavily urbanized area in one of the most technologically sophisticated regions in the country, only about 10% of the potential RSE locations that were investigated were already equipped with landline connections.

Use of bucket trucks: In many instances with the VII California RSE design, shown in Figure 2.3, the installation requires one or two bucket trucks, which adds to the cost and potential traffic disruption of the installation. Prior to the installation of any RSE, the use and manner by which these trucks will be used need to be carefully planned. For freeway installations, if a lane closure of the ramp or the mainline is needed, then this closure should be approved by the agency in charge. For intersection installations, the date and time of installation should be chosen to have a minimum effect on the local traffic. In any event, these issues highlight the kind of cost, resource allocation, traffic disruption and agency approval attention that will be involved in a large-scale RSE deployment.

2.3.2.2.2 Pre-installation infrastructure work

In some cases, it is necessary to perform pre-installation infrastructure work on the chosen site. This work could include a variety of improvements to the existing infrastructure to get it prepared for the actual installation, such as the erection of a pole, installation of pull boxes and conduits, or running cables through the existing conduits to make sure there is enough space within them.

2.3.2.2.3 On-the-spot testing and troubleshooting on the day of installation

At the conclusion of the physical installation, a complete test needs to be performed to make sure the system is up and running. This is usually done via a laptop computer that simulates an OBE by sending and receiving data from the RSE. If the system is not working, then the trouble shooting should begin immediately, while the installation crew and their tools, equipment, and vehicles are available to solve the problem.

2.3.3 Installation Cost

Table 2.1 provides actual costs for a typical installation of RSE using current-generation equipment along Caltrans right of way. The use of MCNU RSUs (the latest generation of DSRC radios) has been approved for installations that will be used during the Proof of Concept (POC) phase of the National VII program, but the purchase cost of the MCNU is not included in Table 1. The costs in Table 1are associated with installations where the controller is the legacy and widely-used 170 model.

TABLE 2.1 Costs ($) associated with VII California RSE installation.

Item	Parts ($)	Labor (man-hours)	Labor cost ($)
DSRC/WAVE Antenna	100		
DSRC/WAVE Antenna Mounts	50	12	600
DSRC/WAVE cable (~20 ft)	60		
Base Plate for MCNU and NEMA Box	200	28	1,680
GPS (unit plus antenna)	500		
GPS Mounts	75	12	600
Fiber Cable (up to 200 ft)	200		
Fiber Converter (4 per site if backhaul is wireless)	250		
Fiber Connectors (8 per site if backhaul is wireless)	120		
Power Supply and Its Base and Circuit Breaker Switch	130		
Uninterrupted Power supply	150		
Signal sensing circuit or sniffer including processing and data transfer modules	2000	80	4,000
Software configuration and testing to match phase timing of each site		80	4,000
Site Survey		2 - 4	160 - 320
Installation (maintenance at $30/hr)		25 – 50	750 – 1,500
Installation (engineering at $50/hr)		12 – 20	600 -1,000
TOTAL	**3865**	**251 – 268**	**12,390 – 13,700**

This shows a parts cost of ~$4,000 per RSE installation, exclusive of the cost of the DSRC radio itself, plus an additional ~$12,500 to ~$14,000 in labor costs for these early installations. Each installation takes about a week of preparations and one full day to physically install and test an RSE by a crew of three to four maintenance workers and two engineering staff. In the long term, when RSE installations are done in large volume, the parts costs are unlikely to decline significantly since these are already relatively common parts rather than custom-built equipment (except, of course, for the DSRC radio). The labor costs could be expected to decline significantly, except for the unavoidable installation and testing labor.

2.3.4 Maintenance Issues

2.3.4.1 Hardware Maintenance

After more than a year and half since the installation of the first RSE, the research team has gained valuable insights about maintaining RSEs in the real world. Lessens learned are described below.

Reliable and current information is essential to maintenance. To assist in rapidly acquiring and

assimilating information about our testbed, we developed a set of software components that run constantly on the RSE hosts and on a designated server. The RSE host components measure the status of the RSE and send reports to the server. The reports are processed into web pages, maps, and time-history graphs, which can help diagnose (and even predict) failures. The information includes:

- Temperature, fan speed, and voltage in the RSE hosts.
- Backhaul and DSRC network data rates -- actual traffic and maximum capacity.
- Packet loss on network segments.
- Host computer status, such as CPU and disk usage.
- Status of devices connected to the host computer.
- VII application status and performance (number of messages, message queue lengths).
- Configuration settings of various components.

The specific maintenance issues that we have encountered so far include:

- Thermal issues related to the fiber converters have been a source of some breakdowns of RSEs. As shown on Figure 2.2, fiber cables are used to communicate between the PC-104 which is placed inside the controller cabinet and DSRC radio in a separate NEMA box. The fiber converters are needed at both ends of the fiber cable. A clear correlation between high ambient temperature and the breakdown of these converters has not been established, but their higher rate of breakdowns in the hotter months of the year points to this conclusion. In recent months, most of the original fiber converters were replaced with another brand that is more heat resistant, but unfortunately the performance of these converters has not been completely satisfactory either. A third brand of fiber converters with higher heat resistance capacity is under evaluation by the research team for future installations, but the higher heat tolerance incurs higher cost as well. The temperature requirements of Caltrans (-40 to +70 C) exceed the specifications of most commercially available electronics, including the fiber converters.

- The most unusual and least preventable failures were caused by animals and nature. Rodents can damage underground cables including the very sensitive fiber cables in the pull boxes. To prevent these animals from doing their damage, the bottom of pull boxes needs to be covered by concrete. This can easily be done for new pull boxes but not for the existing ones. Also, if water and mud get into the conduits, they can damage the underground cables.

- Another source of failure is the interference and unfamiliarity of the local maintenance crew who may unintentionally cause damage. For example, they can damage the very delicate fiber by stretching it while working on another cable underground or while working on RSE equipment in the controller cabinet.

- One type of cellular modem failed due to firmware that had not been updated. It is important to be aware of updates from manufacturers and the problems that they solve.

Clearly, the long term cost of maintenance is difficult to determine at this point, while the technology is still maturing.

2.3.4.2 Institutional Considerations

The cooperation of the owner agency; in our case Caltrans District 4 Maintenance, is essential to get any down RSE back up and running in a reasonable time. The access to the right of way, controller cabinet and infrastructure is only possible if the owner agency's representatives are present. Thus no repairs or regular maintenance can take place without their direct support.

2.3.4.3 Network Operations and Maintenance

The network designed for the initial stages of the project has continued to work well, within the constraints initially identified. Some additional constraints arose later:

1. Some cellular network providers timed out an inactive TCP session after a short period, and data transmitted after this time is simply lost with no error reported. To fix this in Linux, we send a heartbeat packet every 45 seconds.
2. Firewall issues can appear to be network connectivity problems. The issues need to be revisited when new hosts are added or networking equipment is swapped.
3. Updating many (even 12) RSU hosts with new software versions can be time consuming, so we automated this process.

2.4 IMPLICATIONS FOR THE FUTURE

2.4.1 Hardware

From what we have learned thus far, the design of the RSE must be compliant with all national and local standard specifications and requirements. The design must be practical and the overall system must be very reliable. The design should compensate for infrastructure limitations. It should take into account the ease of installation and later on the ease of access to components for repairs and routine maintenance. Another factor is to design for simplicity so that maintenance crews can repair and maintain it with less time and effort. If the repair or regular maintenance requires lane closures, it is best that this task be completed in a minimum period of time and with minimal danger for the crew. The main objective is to put the maintenance crew in hanging buckets as infrequently as possible.

If the system is reliable; there is less need to repair the RSE. If the owner/operator of the RSEs finds it troublesome to keep up with its repair and maintenance, and if the data obtained from the RSEs are not reliable and data flow is not continuous, then RSEs become a burden and lose their utility.

It is best that the system be designed in such a way that repairs and regular updates for software in particular can be done remotely. A local wireless connection between a laptop computer and the RSE site can also be used to make rapid on-site software updates, provided that security can be maintained.

2.4.2 Backhaul Options

Since the installation of our first RSE, we have tried a variety of backhaul options, some using landlines like T1 and some using wireless. The availability and dependability of each backhaul option is a local issue. The fixed and monthly cost of service is a function of the bandwidth and the competition among the service providers. Another local issue is the pre and post installation level of service that one might get from the service provider. Table 2 shows telecommunication transmission costs and bandwidth for the backhaul alternatives considered by VII California. It is intended to serve as a snapshot of currently available backhaul alternatives.

TABLE 2.2

Telecom Backhaul Transmission Costs for VII California

				Bandwidth, kbps		Costs	
	Telco Media	Provider	Interface	Up	Down	Installation	Monthly Fee
wireless	EDGE/Raven	Cingular	Serial	130	70-135	$475	$50.99
	EDGE/Aircard860	Cingular	PCMCIA Card	130	70-135	$160	$50.99
	HSDPA/Aircard860	Cingular	PCMCIA Card	80	400 to 700	$160	$50.99
	HSDPA/Inmotion	Cingular	Ethernet, USB	80	400 to 700	$2,000	$50.99
	EVDO/Raven	Verizon	Ethernet, USB	80	400 to 700	$1,065	$61.00
	HSDPA/Digi	Cingular	Ethernet, USB	90	1,000	$565	$50.99
landline	DSL	AT&T	DSU/CSU	128	384	$221	$40.00
	DS1	AT&T	DSU/CSU	1,540	1,540	$1842	$232.20 + $5 for each mile from TMC Host
	DS1 w/DS3 backhaul	AT&T	DSU/CSU	1,540	1,540	$1,025	$143.60 + $5 for each mile from DS3
	DS3	AT&T	DSU/CSU	43,120	43,120	$1,390	$2025 + $35 for each mile from TMC Host
	ATM 3MG facility	AT&T	ATM Router	3,080	3,080	$900	$945.00
	T1	SpeakEasy	Ethernet	1,540	1,540	$0	$459.00

Note: The above installation and monthly charges are subject to change and could be different in other parts of the U.S.

2.4.3 Institutional Issues

The RSE owner could be either State DOT, county, or city. These diverse infrastructure owners, each possibly with different specifications and standards, can present a challenge for future installations. Interestingly, some intersections belong to the State DOT but are maintained by county or city. For these locations, both agencies need to accept the design, installation method, and maintenance of the RSE. A robust design that can satisfy the specifications and meet the standards of these diverse owners can be the answer.

2.4.4 Summary of Key Lessons Learned

The experience of implementing the first RSEs in the Bay Area has provided some important lessons that should be kept in mind when planning for the national VII deployment and estimating its costs:

- Even though we were working in the high-density and technology-intensive Silicon Valley region of California, the large majority (almost 90%) of the candidate locations (major intersections and freeway interchanges) for VII testbed RSE installation were not already equipped with landline backhaul connectivity. This meant that backhaul connectivity had to be installed as part of the process of installing the RSE, incurring significant up-front costs. Nationwide VII implementation is therefore likely to require significant investments in backhaul communications.

- The selection of backhaul technology for the RSE installations had to be made on a site-by-site basis, depending heavily on local conditions (availability of technologies close to the host controller cabinet). This means that full-scale VII deployment is likely to have to depend on a heterogeneous mixture of backhaul technologies rather than only one or two preferred technologies.

- The cost and performance of backhaul communications (see Table 2) have been more challenging than anticipated at the onset of VII California. This implies that the VII architecture design should seriously consider how best to minimize the backhaul burden from the RSEs in order to contain backhaul capital and operating costs and minimize the losses of data associated with interruptions of the backhaul communication link. These considerations point toward distribution of intelligence throughout the VII network, including some data processing and storage resources at each RSE.

- Over time, multiple versions of RSEs may enter the market. It is imperative that they all work together transparently for the equipped vehicles. This will encourage not only the vehicle manufacturers and after market suppliers but also the agency owners to be more willing to invest in VII.

2.5 FUTURE DEVELOPMENTS

The VII California testbed is expanding, with a network of DSRC roadside units (12 at this writing and plans for up to 40), applications with on-board equipment (on light duty and transit vehicles), and a backhaul network (with heterogeneous backhaul links including T1 landline, 3G modem and soon coming online, WiMAX). Strategic outreach efforts are underway to bring onboard a variety of public and private entities that can benefit from and provide benefits to this project. Development and operation of this testbed has provided a wealth of technical and institutional knowledge in several domains: hardware and software on one hand; and installation, maintenance and operations on the other. Issues raised and overcome include:

- agency cooperation by different divisions that were not involved with the initial decision to conceive this project,
- installation, maintenance and operation under practical operational constraints (standards of practice and codes,
- use of legacy systems on the roadside and
- to transport data via heterogeneous means to ‘back office’ operations).

Significant network and architecture issues remain. The VII California implementation is admittedly experimental and designed to expedite applications development and experiments, but the questions asked and solutions implemented with regard to architecture and 'distribution of intelligence' to the roadside are important to consider as part of the larger VII picture.

These lessons learned will be brought to more mature practice as the VII California testbed transitions from an experiment of VII applied to State and regional needs toward informing Federal-level decisions. This transition is institutional, as VII California may become a development test environment to support the US DOT VII Proof of Concept experiments and to develop important applications, e.g., bridge tolling. Through this transition, testbed plans will transform into evolving design standards, and the cross-fertilization between what has been practical and already implemented versus what is conceived by the US DOT VII program will be an interesting convergence. Throughout this experience, the lessons learned by the VII California program should be carefully considered in VII and in any widespread deployment of networked technology in transportation.

3. PROTOCOLS AND PERFORMANCE

3.1 RSE MONITOR

3.1.1 Introduction

The RSE monitor software automates the tasks of monitoring the hardware and software installed on the roadside. The primary function of the monitor is to detect events and to record the status of hardware and software components and to periodically send reports to a central server. Reports are aggregated on the server and processed into useful forms, such as web pages, maps, and time-history graphs. The secondary function is to repair error conditions when it is possible to do so without operator intervention.

The intended users and use cases of the system are:

1. Administrators: detect, diagnose, and predict failures.
2. Testbed users: before driving out on a test run, determine which parts of the testbed are likely to be able to support the testing they have planned.

3.1.2 Variables

The monitor system collects and reports the following time-varying data:

Network

backhaul packet loss	from PATH server to RSE computer (avg/max/min/dev)
backhaul ssh delay	from PATH server to RSE computer[1]
traffic	data rates, broken down by DSRC/backhaul, in/out bound
capacity	backhaul capacity, broken down by in/out bound
beacons received	how many and from which RSE sites[2]
signage	contents of recent signage, travel time, and other public broadcasts
signal phase	contents of recent broadcasts of signal state

Denso Wave Radio Module (WRM)

configuration	channel, power level, data rate, etc.
status	WRM status fields
packet loss	between pc104 and WRM

Software

CPU usage	broken down by process
memory usage	broken down by process
configuration	VII-CA software configuration
status	recent errors, process uptime

1 *ssh* is "secure shell", the primary tool for administration of remote systems.
2 This information is useful in detecting overlap between nearby DSRC antennae.

diagnostics	monitoring system diagnostics
VII-CA applications	recent VII-CA message activity

Hardware (pc104)

voltage	on-board voltage sensor
fan speed	on-board fan speed sensor
temperature	3 on-board temperature sensors
clock skew	between RSE computer and PATH server
disk usage	
system uptime	

3.1.3 Corrective actions

The monitor system can correct some of the problems that it detects.

Signal current Indicator (sniffer)	Restart sniffer software if phase data is not detected or is obviously incorrect.
VII-CA software	Restart VII-CA programs if they have crashed or become unresponsive.
Denso WRM	Send "wakeup" packets to WRM if WRM incorrectly detects its host interface.
	Send channel/power packets to WRM to correct settings.

3.1.4 Web Interfaces

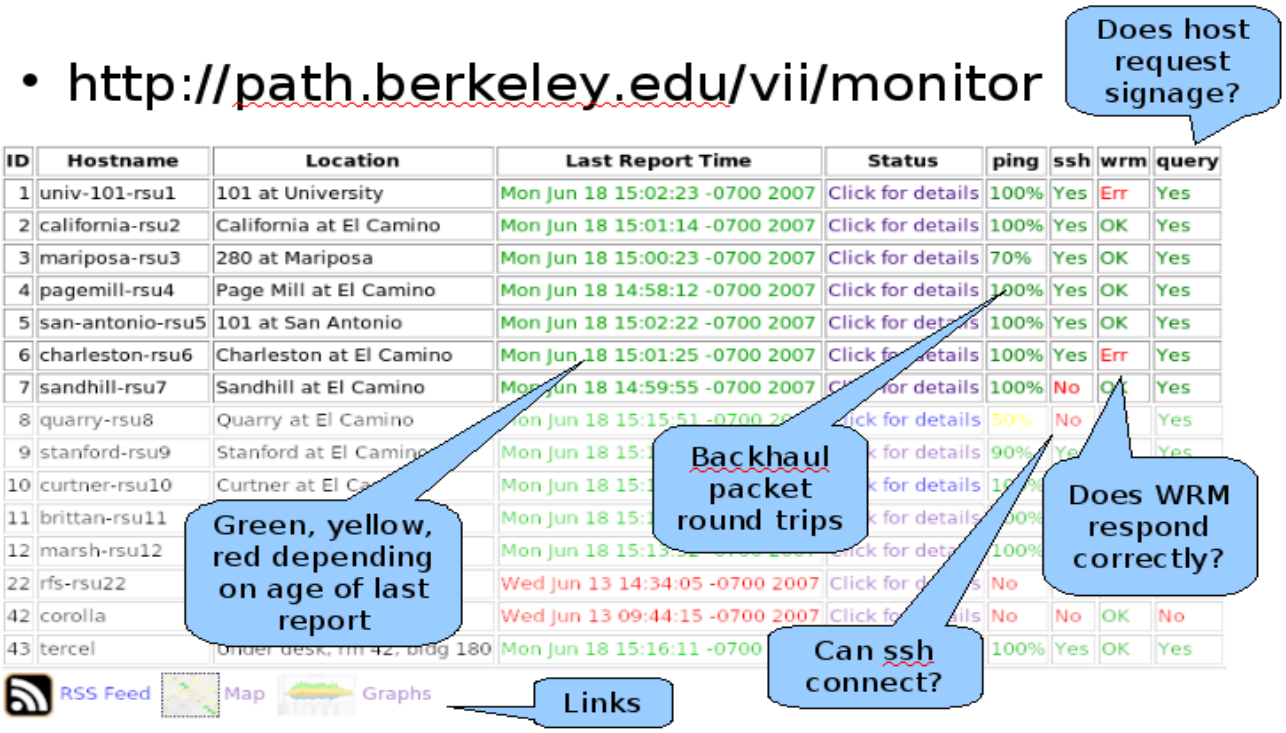
Access to the monitor data is through several web interfaces, available at http://path.berkeley.edu/vii/monitor.

Current reports	Show the most recently received data in a summary table and in detail pages.
Time-history graphs	Show history of variables over hours, days, or weeks.
Google map	Quick overview of testbed status, overlaid on Google map.

3.1.4.1 Current reports

The current reports are summarized in a table like Table 3.1:

Table 3.1 VII Monitor Report



ID	Hostname	Location	Last Report Time	Status	ping	ssh	wrm	query
1	univ-101-rsu1	101 at University	Mon Jun 18 15:02:23 -0700 2007	Click for details	100%	Yes	Err	Yes
2	california-rsu2	California at El Camino	Mon Jun 18 15:01:14 -0700 2007	Click for details	100%	Yes	OK	Yes
3	mariposa-rsu3	280 at Mariposa	Mon Jun 18 15:00:23 -0700 2007	Click for details	70%	Yes	OK	Yes
4	pagemill-rsu4	Page Mill at El Camino	Mon Jun 18 14:58:12 -0700 2007	Click for details	100%	Yes	OK	Yes
5	san-antonio-rsu5	101 at San Antonio	Mon Jun 18 15:02:22 -0700 2007	Click for details	100%	Yes	OK	Yes
6	charleston-rsu6	Charleston at El Camino	Mon Jun 18 15:01:25 -0700 2007	Click for details	100%	Yes	Err	Yes
7	sandhill-rsu7	Sandhill at El Camino	Mon Jun 18 14:59:55 -0700 2007	Click for details	100%	No	OK	Yes
8	quarry-rsu8	Quarry at El Camino	Mon Jun 18 15:15:51 -0700 2007	Click for details	50%	No		Yes
9	stanford-rsu9	Stanford at El Camino	Mon Jun 18 15:1	k for details	90%	Ye		Yes
10	curtner-rsu10	Curtner at El Ca	Mon Jun 18 15:1	k for details	10%			
11	brittan-rsu11		Mon Jun 18 15:1	k for details	90%			
12	marsh-rsu12		Mon Jun 18 15:13	Click for deta	100%			
22	rfs-rsu22		Wed Jun 13 14:34:05 -0700 2007	Click for d	No			
42	corolla		Wed Jun 13 09:44:15 -0700 2007	Click fo	No	No	OK	No
43	tercel	under desk, rm 42, bldg 180	Mon Jun 18 15:16:11 -0700		100%	Yes	OK	Yes

3.1.4.2 Detail pages

Each site has a page, to which the summary table links, showing more of the variables collected for that site, plus a link to a page with the full report:

Status of RSU 2, California at El Camino

Last report was **Tue Oct 16 14:44:24 PDT 2007**.

Full report

Report index

Variable	Value
Backhaul IP Address	64.81.31.227
WRM Address	192.168.1.148
WRM Channel	Radio Frequency: 5890 MHz (IEEE 178)
WRM Data Rate	6
WRM Power	Transmit Power: 8 dBm; Current Transmit Output Power 16.0 dBm

Recent Signage Broadcasts

First Observed	Last Observed	Message Text
Fri Oct 12 13:03:53 PDT 2007	Tue Oct 16 14:39:03 PDT 2007	Speed limit for CA-82 is 35 miles per hour

Recent Signal Messages

Last signal report was **Tue Oct 16 14:43:59 PDT 2007**. Status: **No errors**

Time Sent	Message Data
14:43:36.887	1: R: 21.1 2: G: 11.6 3: R: 37.4 4: R: 53.8 5: R: 75.8 6: G: 31.8 7: R: 63.8 8: R: 38.2
14:43:37.087	1: R: 20.9 2: G: 11.4 3: R: 37.2 4: R: 53.6 5: R: 75.6 6: G: 31.6 7: R: 63.6 8: R: 38.0
14:43:37.287	1: R: 20.7 2: G: 11.2 3: R: 37.0 4: R: 53.4 5: R: 75.4 6: G: 31.4 7: R: 63.4 8: R: 37.8

Figure 3.1 Example detail page, California at El Camino

3.1.4.3 Full report

The full report is the entire database contents for a particular RSE site.

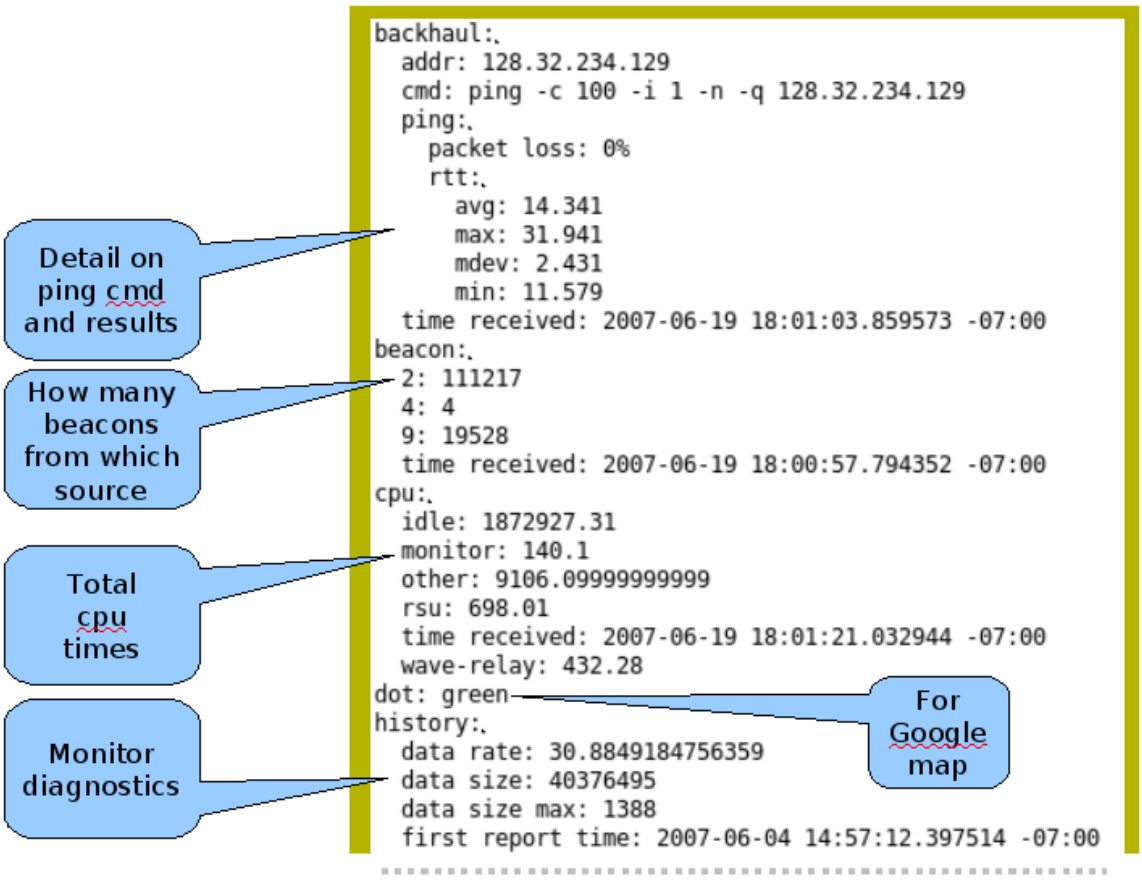


Figure 3.2 Example full report

3.1.4.4 Time-history graphs

The history database and graphs managed using the RRDTool, which has become a standard for system administration. RRD stands for Round Robin Database; "round robin" means that the database stores only the last N data points, where N is a fixed number. RRDTool also aggregates the data (average, min, max) and differentiates (converts byte counts to rates, for example). Graphs are generated based on user input, such as site ID, variable name, time window, and averaging period. The following kinds of graphs are currently available:

1. SSH responding for all sites
2. SSH response time for one site
3. Round-trip ping rate for all sites over 24 hours
4. Round-trip ping rate for selected sites and duration

5. WRM ping packet loss for selected sites and duration
6. WRM ping packet loss for selected site, duration, and averaging period
7. Temperature inside pc104 boxes
8. Fan speed inside pc104 boxes
9. Voltage inside pc104 boxes
10. Memory usage
11. CPU usage
12. Process uptime
13. Network usage, DSRC and backhaul
14. Network usage, DSRC only
15. Network usage, backhaul only
16. Backhaul network capacity

The following graph shows packet loss on the fiber-optic cable between the host computer and the WRM[3]:

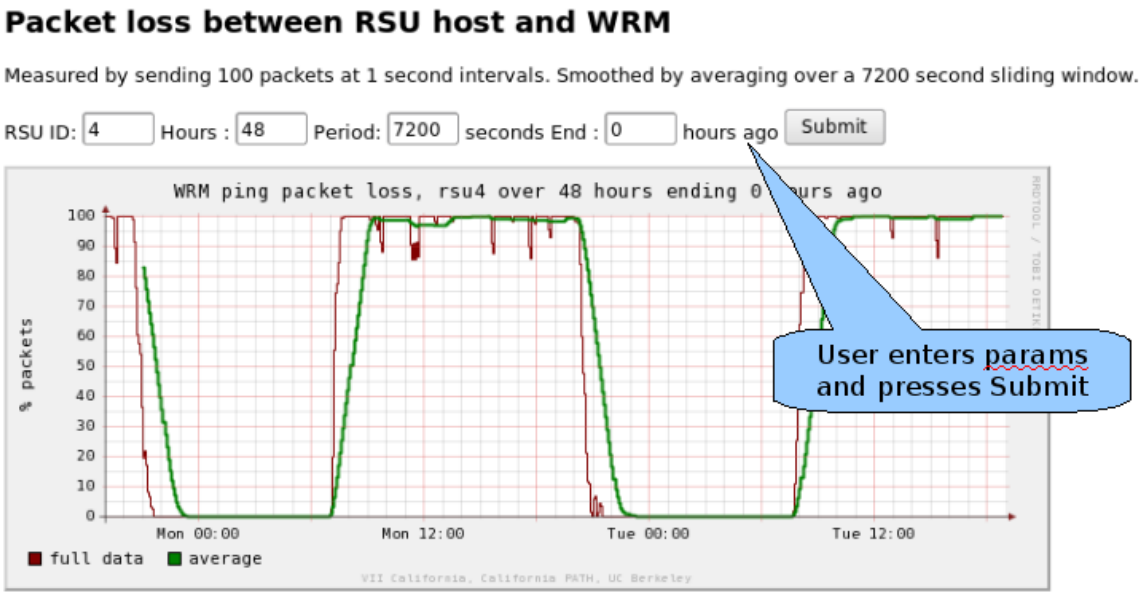


Figure 3.3 Graph of Packet Loss

3 This failure mode is still not completely understood, even though it has a sharply defined diurnal cycle. It may be temperature related, but there is apparently no simple cause-effect relation.

Another graph used to diagnose this situation was the temperature graph:

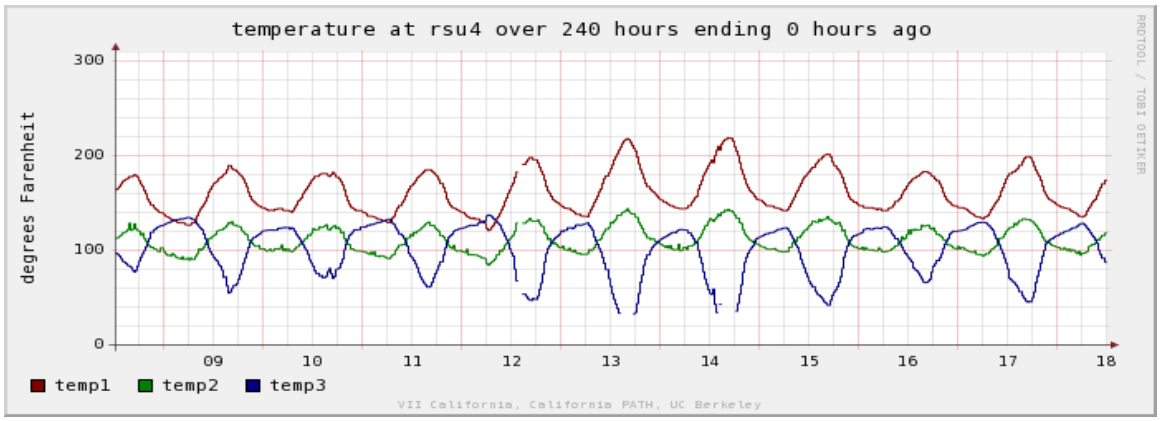


Figure 3.4 Temperature Graph

There are also summary graphs that show some aspect of the status of the whole system. The following graph shows the periods during which sites were not accessible using ssh:

Figure 3.5 Graph of ssh responsiveness

Resource usage graphs are essential to checking the health of a system[4]. The following graph shows four variables at one site: outgoing and incoming data rates on the DSRC and backhaul networks. The graph also shows how some of the image generation capabilities of RRDTool can enhance the visualization of complex information:

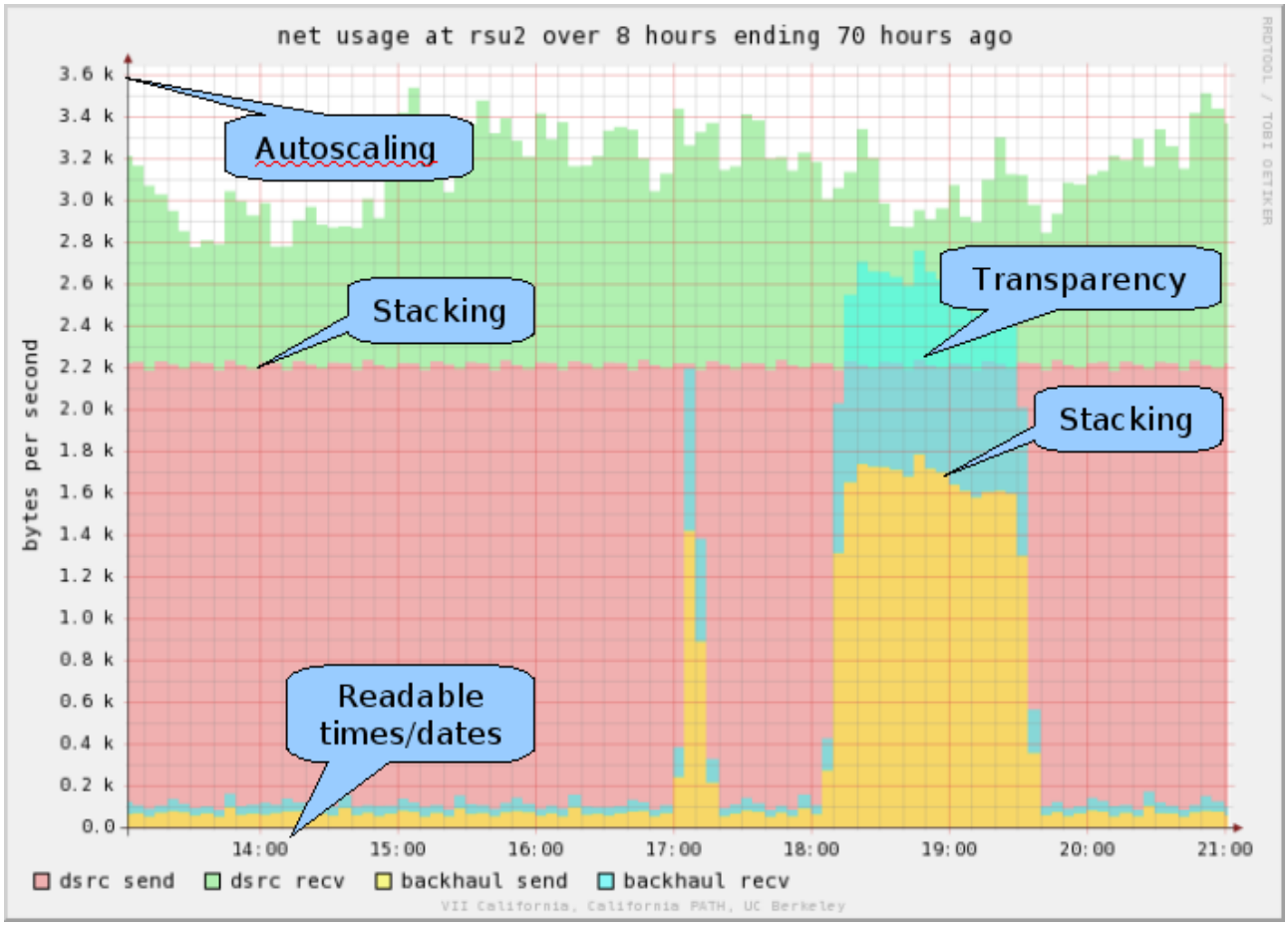


Figure 3.6 Resource graph of an RSE

4 A recent intrusion in one of the RSE hosts was detected using the memory and CPU resource graphs.

The following graph was used to detect and diagnose a software bug that caused unexpectedly high memory usage:

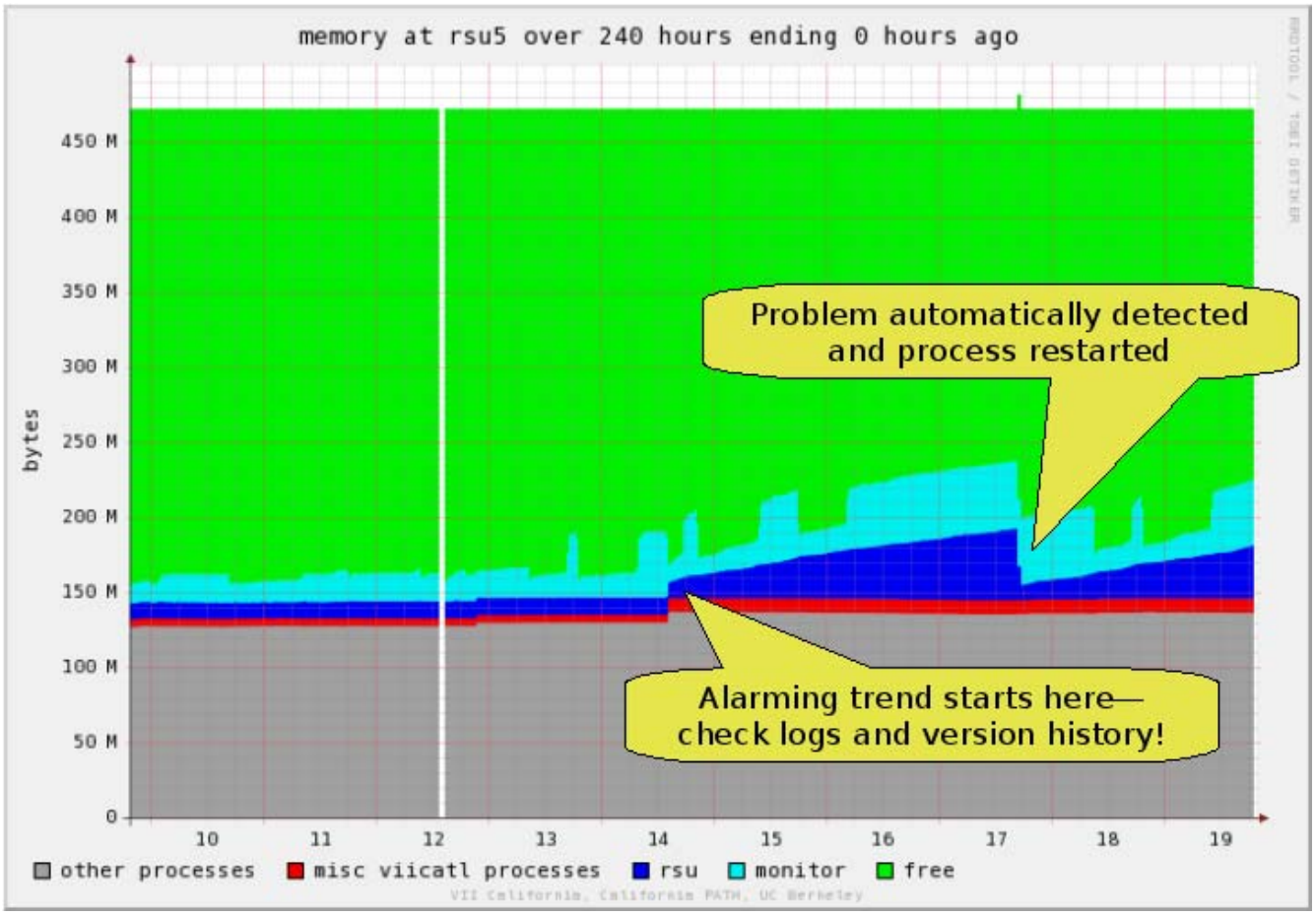


Figure 3.7 Graph of Memory Usage

3.1.4.5 Map of testbed with status indicators

The map is an overlay of a Google map with data from recent reports. The color of the dot is green if no serious problems are detected, and red otherwise.

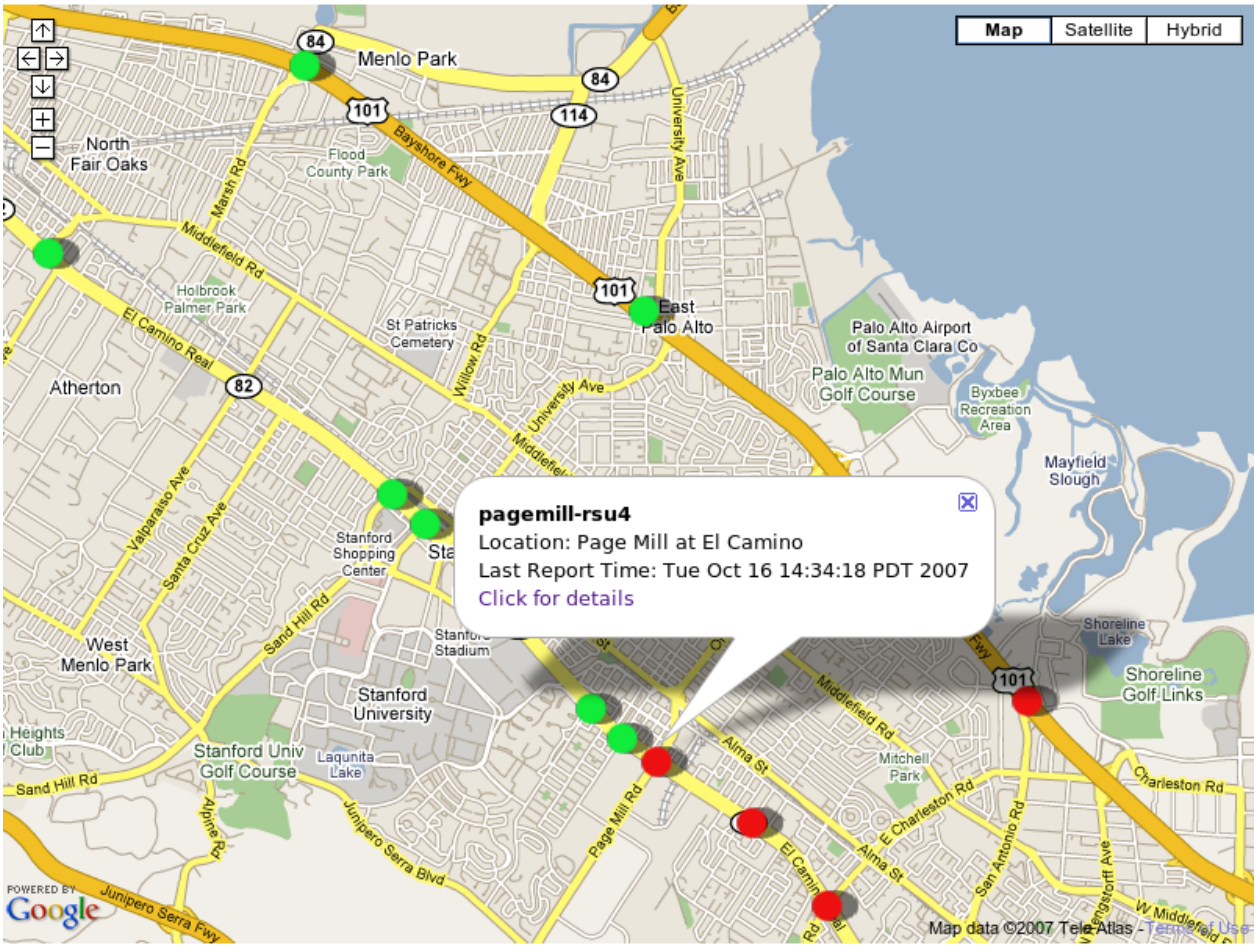


Figure 3.8 Testbed map

3.1.4.6 Technical details

The architecture of the monitor system involves two kinds of components:

- A *monitor client*, which runs on each RSU host device.
- A *monitor server*, which runs on a host somewhere on the Internet.

These programs are installed and running on the active field RSE and testing RSE and on the path.berkeley.edu server. The client software is installed on new RSE hosts as they become available.

Reports are sent from client to server periodically (typically, at 5 minute intervals). Data is sent over UDP, rather than TCP, because packets lost due to network problems are less important than newer packets containing newly sampled data. The monitor client takes care not to use too much

bandwidth, so that VII applications can proceed with minimal interference.

The architecture is shown in the following figure:

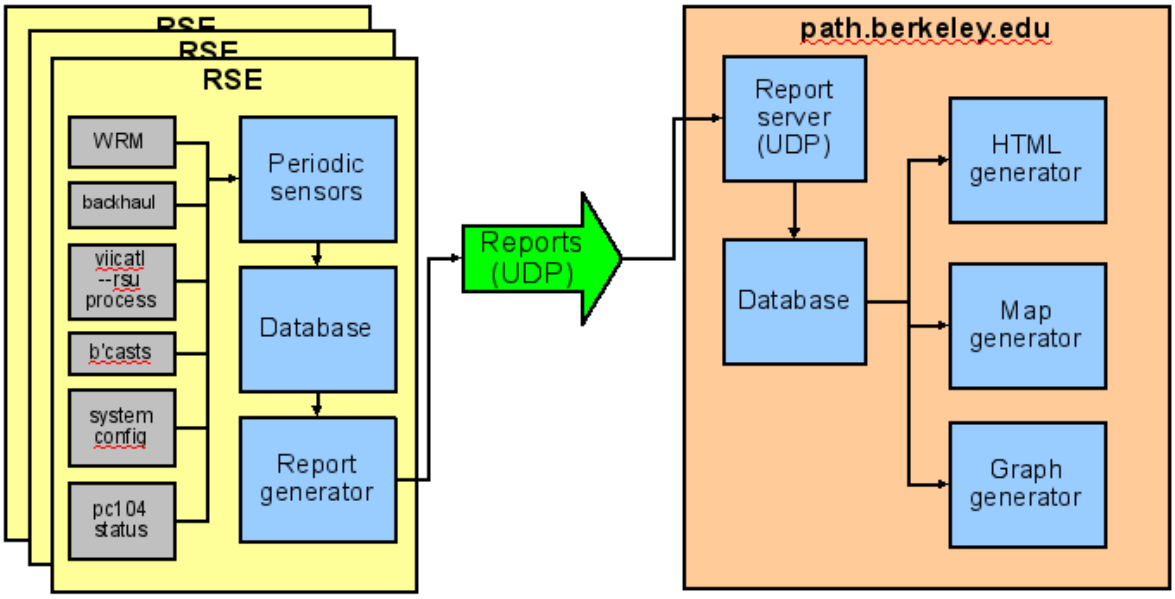


Figure 3.9 Architecture of Monitor System

3.2 NETWORK: MIDDLEWARE FOR VII-CA MESSAGES

A major goal of the VII California testbed is to support a diverse initial set of applications, including probe, public information, and non-public services (such as navigation and tolling). The VII-CA partners, led by PB Farradyne in 2005, defined an initial message set to support these applications, and PATH developed a middleware for the exchange of these messages between in-vehicle client software and back-end application servers.

The characteristics of this middleware are best examined against the background of the constraints on the project:

1. The compressed development schedule has required the use of prototype-quality hardware. Firmware for several components (including the Denso WAVE Radios Modules and the cellular modems) evolved to new generations during the project. Most components have bugs and limitations. The Denso radios are particularly limited: they cannot transmit packets to more than four hosts on their wired side and they cannot assign dynamic IP addresses to vehicles. This means that they cannot support TCP-based networking between vehicles and servers. Also, the Denso radios do not fully support the WAVE standards; they only support IP networking.
2. The backhaul network is heterogeneous, ranging from cellular modem to dedicated T1 land-line. Hence, not all sites can support all services. Furthermore, this diversity requires built-in limits on roadside-to-server communication, such as buffering, prioritizing, and possibly probe data aggregation.
3. The inherent unreliability of mobile wireless networks requires a degree of robustness.

Packet loss is likely in an environment with nodes traveling at high speeds with limited range, line-of-sight occlusions, and RF interference. No application can expect to have long-lasting, consistently available connections.

4. Application messages can be long. In the initial VII-CA message set, probe messages can be up to 64 KB with multiple snapshots. Media file downloads could be several megabytes.
5. The middleware should present VII California application programs with an interface for communication between vehicles and servers that is message oriented (like UDP, User Datagram Protocol) but has inherent quality of service controls (to some extent like TCP, Transmission Control Protocol). It must be somewhat reliable, but not at the cost of excessive use of the channel or complexity from the point of view of the application.
6. Information from the vehicle must be kept as private as possible.

Within these constraints, the VII California middleware succeeds in several ways:

1. The first packet from the vehicle to the server is a data packet; there is no delay for dynamic IP address assignment, TCP session initialization, or other handshaking. The reduction in delay increases the chance that a return message (if any) will be received before the vehicle goes out of range.
2. Message reception probability is robust in the face of short-term failures. The middleware can detect the loss of a fraction of the packets in a message and retransmit the missing fragments. If packet loss is not extensive, this will quickly and efficiently complete the message transmission. Unlike TCP, if no fragments arrive (likely because nodes are out of range or occluded), no bandwidth is wasted resending them, and no bandwidth is wasted on acknowledgment packets.
3. The suite includes a proxy which runs on the vehicle side to insulate vehicle code from the complexity of the wireless user datagram protocol (UDP protocol) and provide it with a TCP socket interface for sending and receiving messages.
4. On the server side, similarly, the interface to the middleware is a TCP socket, with multiple vehicle sessions multiplexed in one TCP session using transaction IDs.
5. Message addressing is encapsulated across the wireless link to overcome the Denso limitations.
6. No identifying information from the vehicle (outside of the application payload) propagates farther than the roadside or is stored anywhere.
7. The middleware supports general HTTP access (though it was not required for the initial set of applications), so that web-based applications can run from within vehicles, regardless of the Denso limitations.

Certain aspects of a complete VII architecture are not fully addressed by this prototype software, including security, scalability, and conformance to emerging national VII standards (Henry, 2007; FHWA, 2005; Cops, 2006).

3.3 EVALUATION OF DSRC COMMUNICATION PERFORMANCE USING LOW COST PROTOTYPE SYSTEMS WITH MULTIPLE RADIOS PER VEHICLE

3.3.1 Introduction

The adoption of the IEEE 1609 Dedicated Short Range Communication (DSRC) / Wireless Access in a Vehicular Environment (WAVE) family of standards for trial use in 2006 (USDOT, 2006) has given the public sector and the auto industry a well-worked-out framework for conducting proof of concept testing. However, there are many performance issues that require further research to ensure a safe and robust highway communication system. Simulation work such as that in (Yin, et al, 2004) and radio prototype and application feasibility development, as has been done by the Vehicle Safety Communications Consortium (Krishnan, 2006) and by exhibitors at the Innovative Mobility Showcase at the 2005 ITS America World Congress, among others, provides a solid basis for continued development, but changes from technology specifications in the current trial use standards may be required before full deployments (Misener and Shladover, 2006; Larson and McKeever, 2006).

Performance under congested traffic conditions remains an open issue. The density of vehicle communications on crowded freeways and busy intersections is high compared to the typical office WiFi hotspot, and furthermore highway communications are expected to include an unusually large number of broadcasts. Analysis and simulations cannot accommodate the full complexity of practical systems with large numbers of vehicles in close proximity. Overall performance depends not only on physical layer channel effects and on the performance of backoff at the 802.11 MAC layer, but on operating system and application structuring and scheduling of messages. Some standard 802.11 techniques for congestion mitigation, such as the Point Coordination Function, are not possible with vehicular systems because the set-up time is too long to be effective in a situation where many transmitters and receivers are entering and leaving the network in a short period of time. More research is needed to characterize the magnitude of congestion problems and the total amount of communication possible using real hardware in the field and to investigate methods of congestion mitigation.

In addition to the problem with the density of communications, safety applications have critical latency requirements. Communication measurements tools are often only concerned with the aggregate data transfer rates and error percentages, and do not measure the latency of individual messages. However, in DSRC systems the latency may be the limiting parameter on the data transfer rate of the system, since the system load cannot be allowed to rise to a rate that will delay crucial safety communications.

The work described in this section represents first steps in using inexpensive computer hardware with prototype radios to explore performance issues for multiple OBEs approaching a RSE. A major motivation for this research is the scenario described in the proceedings of the December 2004 DSRC workshop (Cash, 2004) in which many vehicles come in range of a transaction service advertised by a Road Side Unit simultaneously and immediately switch to a service channel and begin broadcasting.

The next section (3.3.2) in this report provides a description of the hardware used for testing. Following that, Section 3.3.3 gives a description of the testing software. Section 3.3.4 describes baseline testing done in the lab to validate the software and hardware as tools for measuring latency and to characterize the performance of the radios under test in conditions where all the radios were in range.. Section 3.3.5 describes some preliminary tests with channel switching and Section 3.3.6 presents data with mobile radios at Richmond Field Station. Section 3.3.6 also details experiments that were carried out at two intersections in San Francisco to emulate transaction processing on the service channel and investigate performance in an urban setting. Section 3.3.7 presents conclusions and plans for future research.

3.3.2 Computer and Radio Equipment

The equipment used for our tests takes advantage of the availability of inexpensive miniature microprocessor portable assemblies. Four assemblies were constructed; each containing three 400 MHz gumstix expansion boards with two Ethernet ports (Gumstix Inc, 2007) (see Figure 3.10). Funding for the equipment was provided by ARINC Corporation, as part of their work on DSRC for USDOT. ARINC also supplied 12 Mobile Mark magnetic mount antennas specially constructed for the 5.9 GHz frequency.

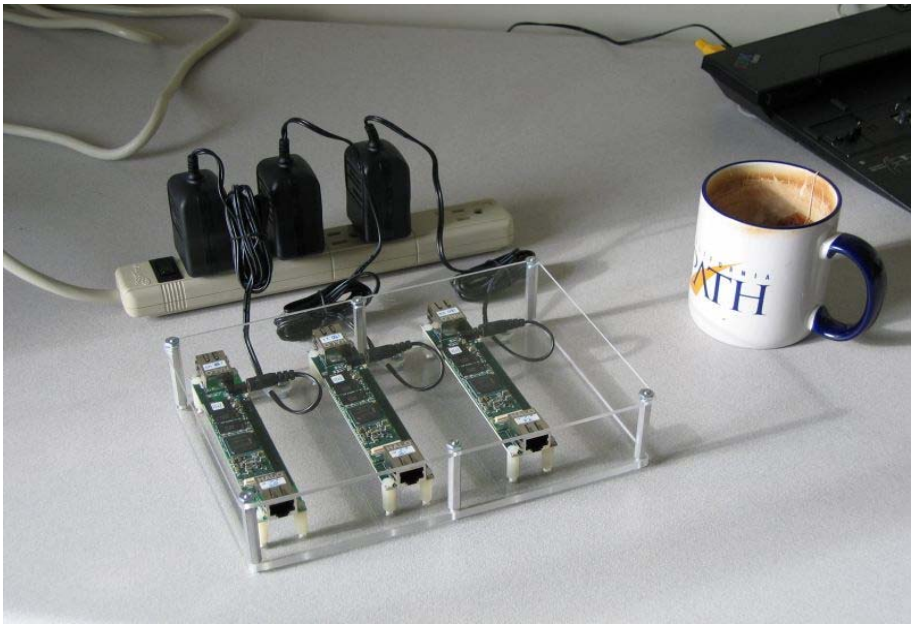


Figure 3.10 Assembly of three gumstix netduo expansion boards.

Each of the 12 individual computers runs an embedded Linux system and has two Ethernet interfaces. On one Ethernet interface it is paired with its own Denso WAVE Radio Module and rooftop antenna; on the other it can be connected to a host computer for data storage.

The assemblies were designed for flexible configuration. All the assemblies can be connected via a router and deployed in one vehicle for tests to an RSU, or individual assemblies can be placed

in different vehicles for vehicle to vehicle research or to test multiple approach paths to an RSU simultaneously. See Figure 3.11 for the test configuration used.

For most of the tests described in this section, the gumstix processors are connected through the router to a laptop, which provides a networked file system where trace data from the test run can be stored. In some tests, the laptop is also connecting through a USB port to a GPS unit, so that location can be saved as well. All processors are synchronized through the router.

Figure 3.11 Antenna placements with 8 active OBUs on one vehicle

3.3.3 Data Acquisition Software

There are issues with measuring latency, the delay in receiving a message, as well as overall rate. While many utilities, such as the ubiquitous ping and the widely used iperf, exist to give estimates of the performance of a network link, more thorough tracing is required in the case of DSRC. Because the characteristics of the link change as stations come closer together and move apart, average data transfer and round trip time measurements do not provide enough information to diagnose and elucidate link breakdown conditions. We developed programs for this project that allowed us to trace round-trip time on a message by message basis, accounting for clock skew in the process.

The basic programs developed une34 this effort included a sender, a receiver and a monitor process. The monitor process runs on the same system as the sender and records the round trip information when a return packet is received from a receiver. Two additional convenience programs evaluate the clock skew between a system and its network neighbor, and set the clock based on a message received from the clock skew program.

These processes use a library API for communicating with the Denso WRMs that implements the low-level processing of the IP headers and raw IP packets that is required in order to record Received Signal Strength Indication values for each packet.

The processes can all be configured for different power, channel and rate settings of the Denso WRM. The send process can additionally be configured with the number of bytes per packet, the time interval between sends in milliseconds, and the total number of sends in a run.

Data recorded by the monitor process at the originating sender for each round-trip packet were: time that the monitor process recorded the data; destination (RSU) IP; sequence number; number of bytes in packet; time packet was originally sent; time packet was received at remote system ;time return packet was received by monitor program; estimated one-way transmission time; average one-way transmission times over course of run; estimated absolute value of clock skew between OBU and RSU;RSSI at remote host ; and RSSI read at sender

Clocks on hosts and gumstix were synchronized at the start of the experiment, and post processing was used to interpolate GPS data and assign values for UTC time, latitude and longitude to all received packets.

3.3.4 Baseline Testing

As part of our base line testing, we first needed to determine that we were able to measure communication delays, and delays due to interference in communication. There was a question at the start as to whether operating system scheduling delays would make it difficult to measure communications effects. Figure 3.12 below shows that the basic trip time between two systems in good communication, with no competition from other stations for the airwaves, is typically 2-3 milliseconds. While the actual time transferring the symbols on the radio medium is much less than this, this includes the time buffering and copying the message on transmission and reception, and scheduling the application processes that are the source and destination of the message. Figure 3.13 shows the situation when two different stations are sending messages pairwise at a moderate rate. Competition for the airwaves between the two pairwise streams causes the round trip time to rise to between 3-5 milliseconds. Figure 3.14 shows how a high rate of transfer between two computers can cause greatly increased delay on a pairwise stream between two other computers communicating at a more moderate rate. Note that the trip time scale on this graph is in seconds, not in milliseconds as for the previous trip time graphs. See on the lower graph in Figure 3.14 that on the higher rate stream packets were actually dropped, due to exceeding the retry limit, between sequence numbers 1454 and 1746.

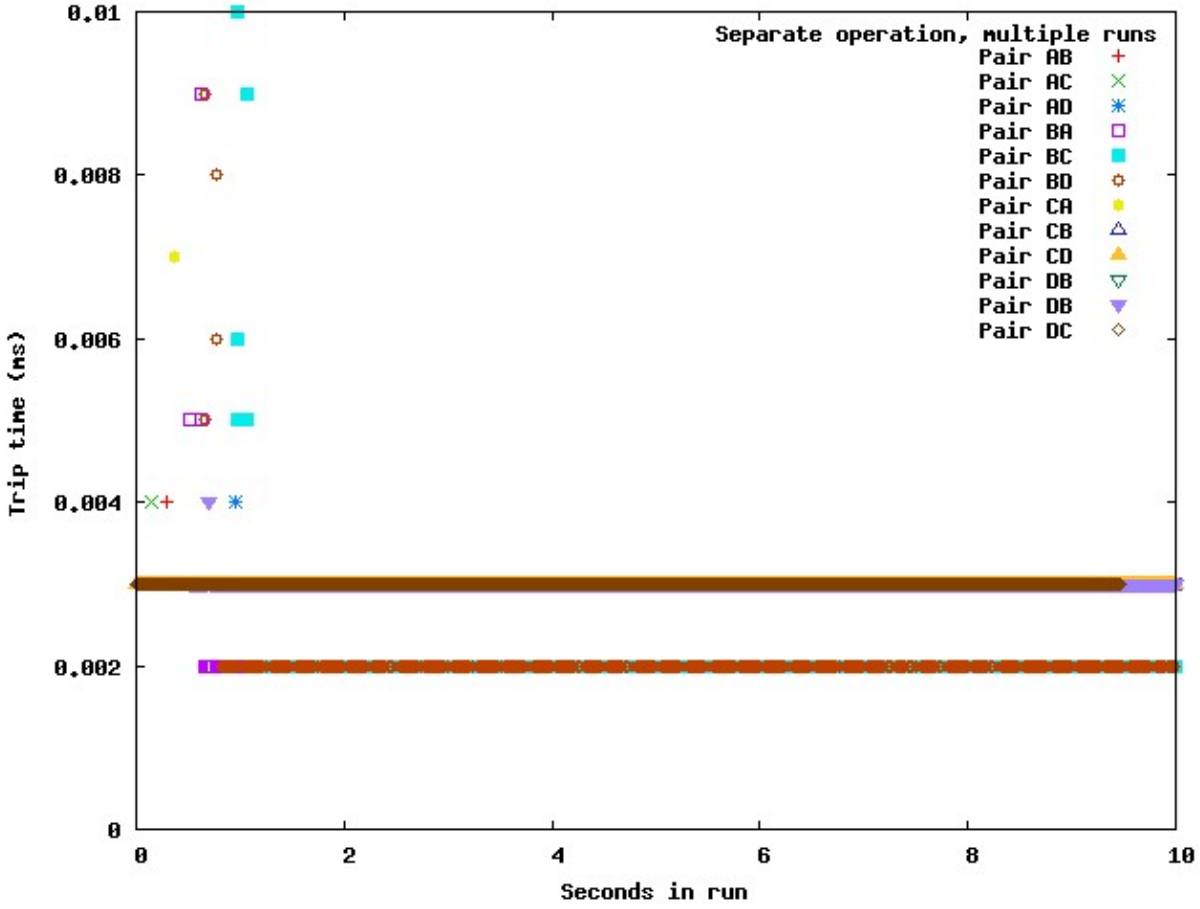


Figure 3.12 Baseline operation, 1480 byte data per packet, 10 millisecond interval between sends, different pairwise runs of the four stations used for desk checking

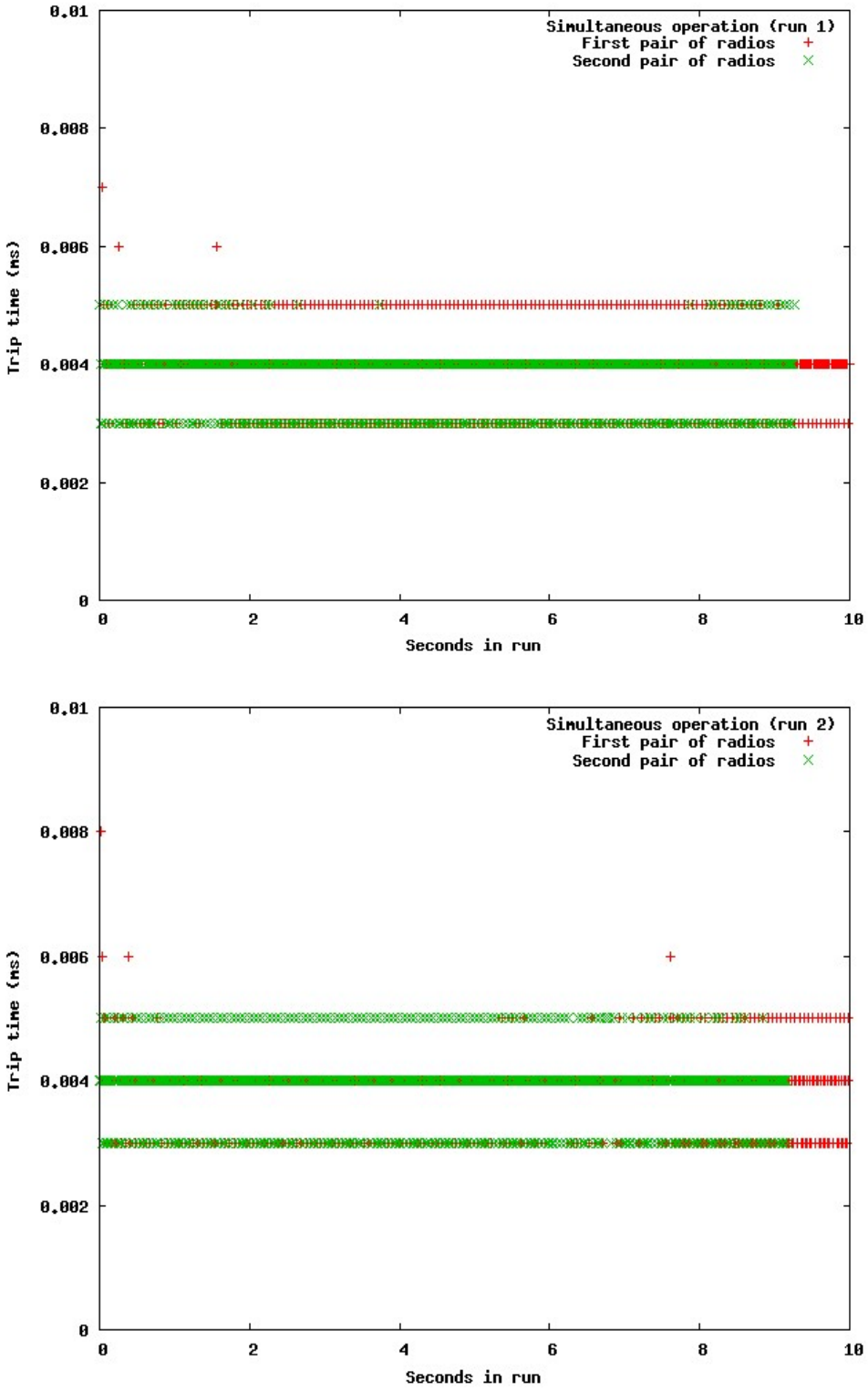


Figure 3.13 Two runs showing interference between two pairwise communication streams causing slightly increased trip time (1480 byte data per packet, 10 millisecond interval between sends on both streams)

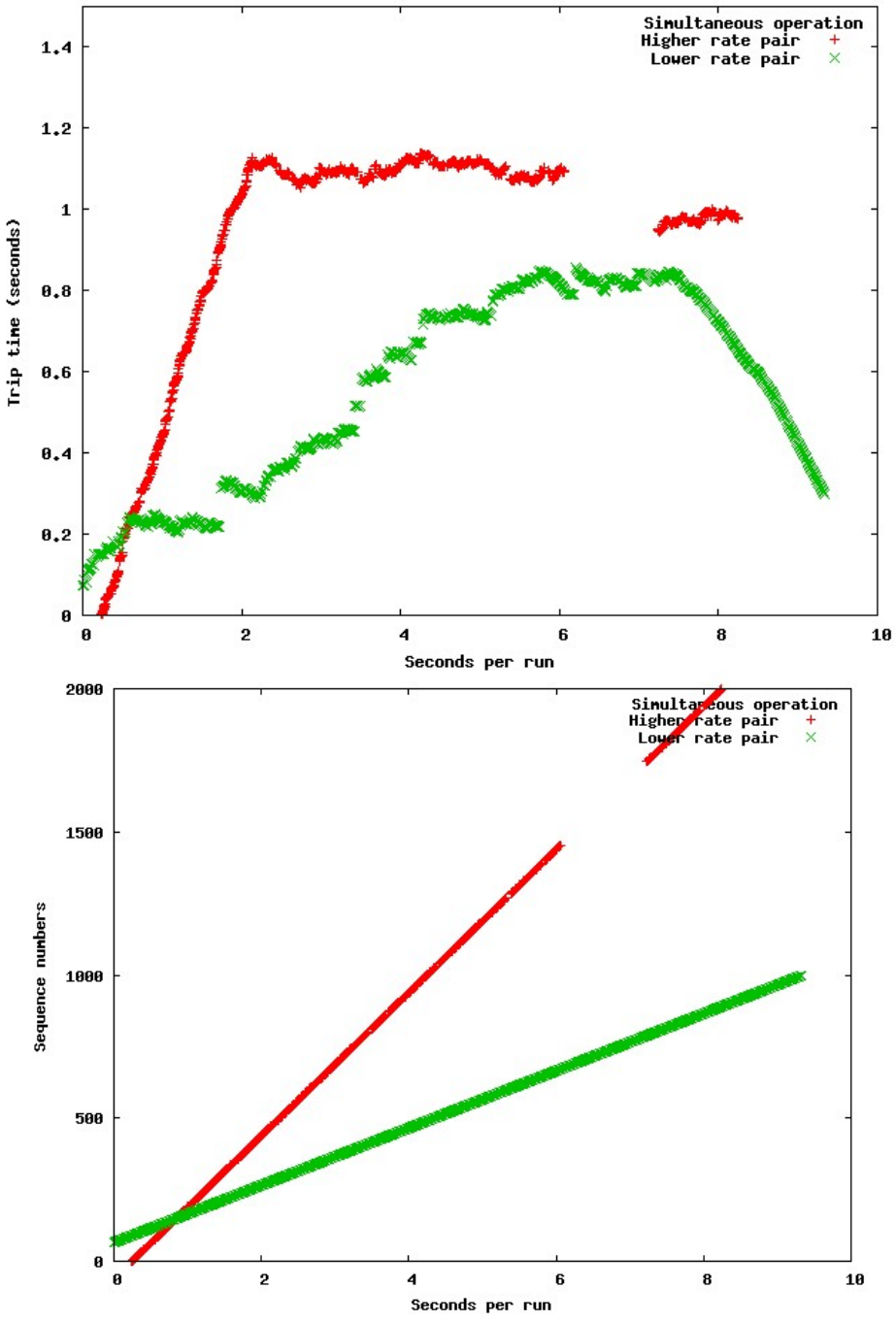


Figure 3.14 Large increase in trip time due to interference between a higher rate pairwise communications stream (1480 byte packets at 4 ms interval between sends) and a lower rate stream (1480 bye packets at 10 ms interval between sends). Note that the scale for trip time is in seconds; the lower graph shows sequence numbers.

3.3.5 Effects of Channel Switching

The prototype Denso WAVE Radio Modules that we were using for this preliminary testing were not designed to switch channels fast enough in order to allow a 50 ms control channel period and a 50 ms service channel period, as specified in the IEEE 1609.4 standard. However, we were interested in seeing how well communication could be maintained across channel switches even under adverse circumstances. We also wanted to test how well the gumstix processors were synchronized across the router. We found that, while for infrequent channel switching it was possible to continue communications with only a small percentage of lost packets, when channel switching was more frequent it was impossible with our current equipment to maintain synchronization well enough to avoid significant losses.

We began by performing a set of tests, with stationary equipment in close range. Five gumstixs acted as "vehicles" sending to a sixth used as the "road side unit" or receiver. Channels were switched between Channels 172 and 178 by a background process running on the host laptop that sends a message through the router telling each gumstix processor when it is time to switch channels. The gumstix processor writes a record to the shared file system of the time of the switch. Clocks on all gumstix processors are synchronized before the start of testing. Parameters are shown in Table 3.2 below.

Table 3.2 Parameters used for tests of frequent and in frequent channel switching.

Testname	Switch interval (between channel changes)	Packet Interval (between sends)	Packet size (data bytes)	Load at Receiver (bytes/sec from all senders)
sw0	No change	100ms	200	10K
sw1	5 sec	100ms	200	10K
sw2	5 sec	200ms	200	5K
sw3	5 sec	50ms	200	20K
sw4	100ms	25ms	200	40K
sw5	100ms	10ms	200	100K

Figure 3.15 below shows the baseline case with no channel switching. Note that in these graphs the sequence numbers for the different processors have had a constant added so that they appear in different Y ranges; in fact the sequence numbers begin at 1 for every invocation of the program.

Figure 3.16, with case sw1 indicates that infrequent channel switching causes little or no degradation in performance. Figure 3.17 shows that for this run the channel changes on the different gumstix processors are clustered within about 10-15 ms of each other, with delays through the router from the host laptop where the channel switch directive is initiated and scheduling delay on the individual processors adding up to this amount. Figures 3.18 and 3.19, cases sw2 and sw3, show no appreciable difference from case sw1. However, for case sw4, as shown in Figure 3.20 the increased overall load and increased channel switching. have caused some dropped packets (evidenced by missing sequence numbers) and some round trip times over

5ms. Notice in the close-ups in Figure 3.21 how for some processors, the time they are on a channel is skewed enough away from the time of the receiver that they may successfully send only one or two messages while they are connected on the same channel. Figures 3.22 and 3.23 show that these effects are more pronounced with the larger load. Overall system load on the router may also be contributing to the greater skew on channel switch times, when sw5 is compared to sw4.

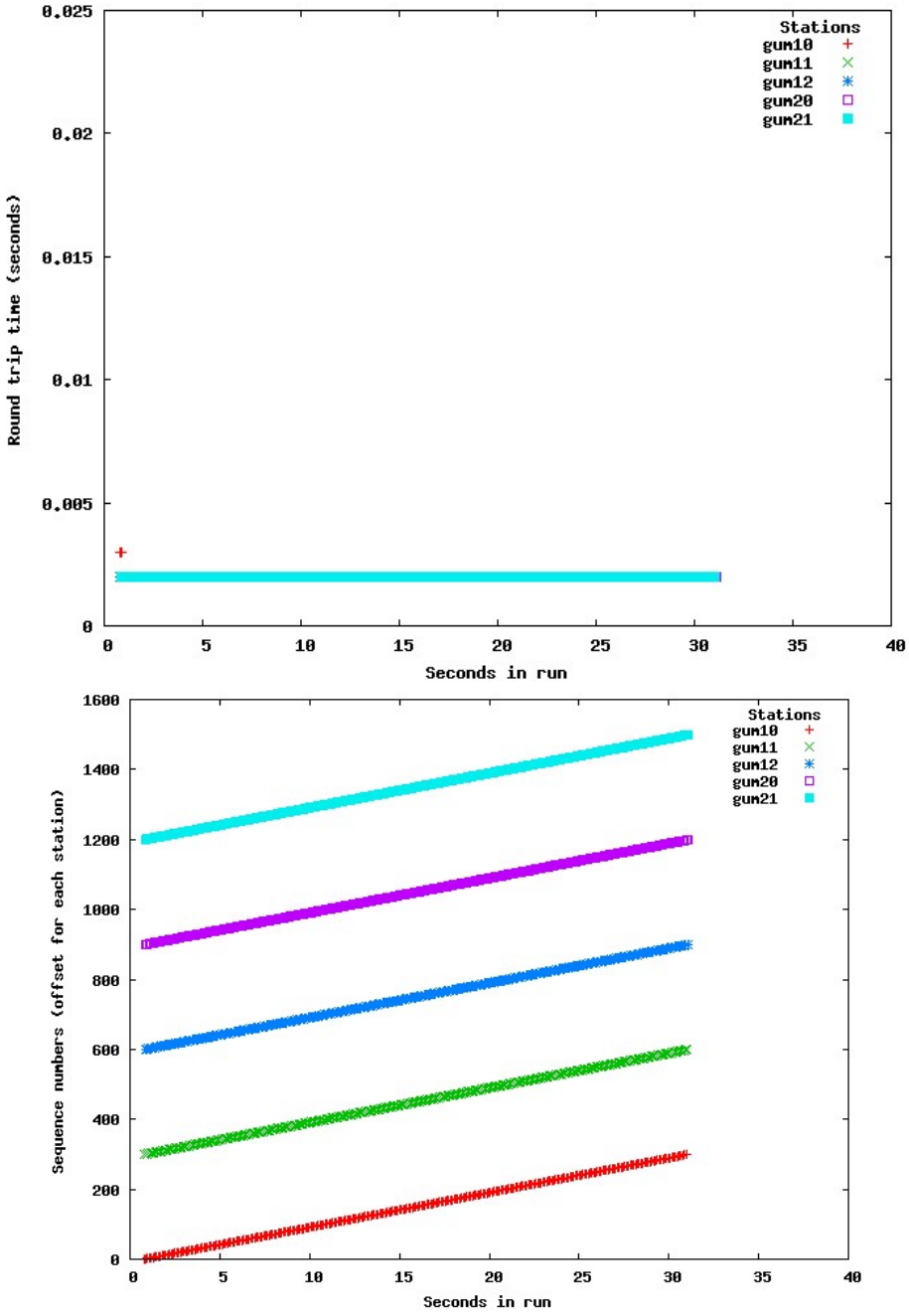


Figure 3.15 Baseline case: no channel switching, light load.

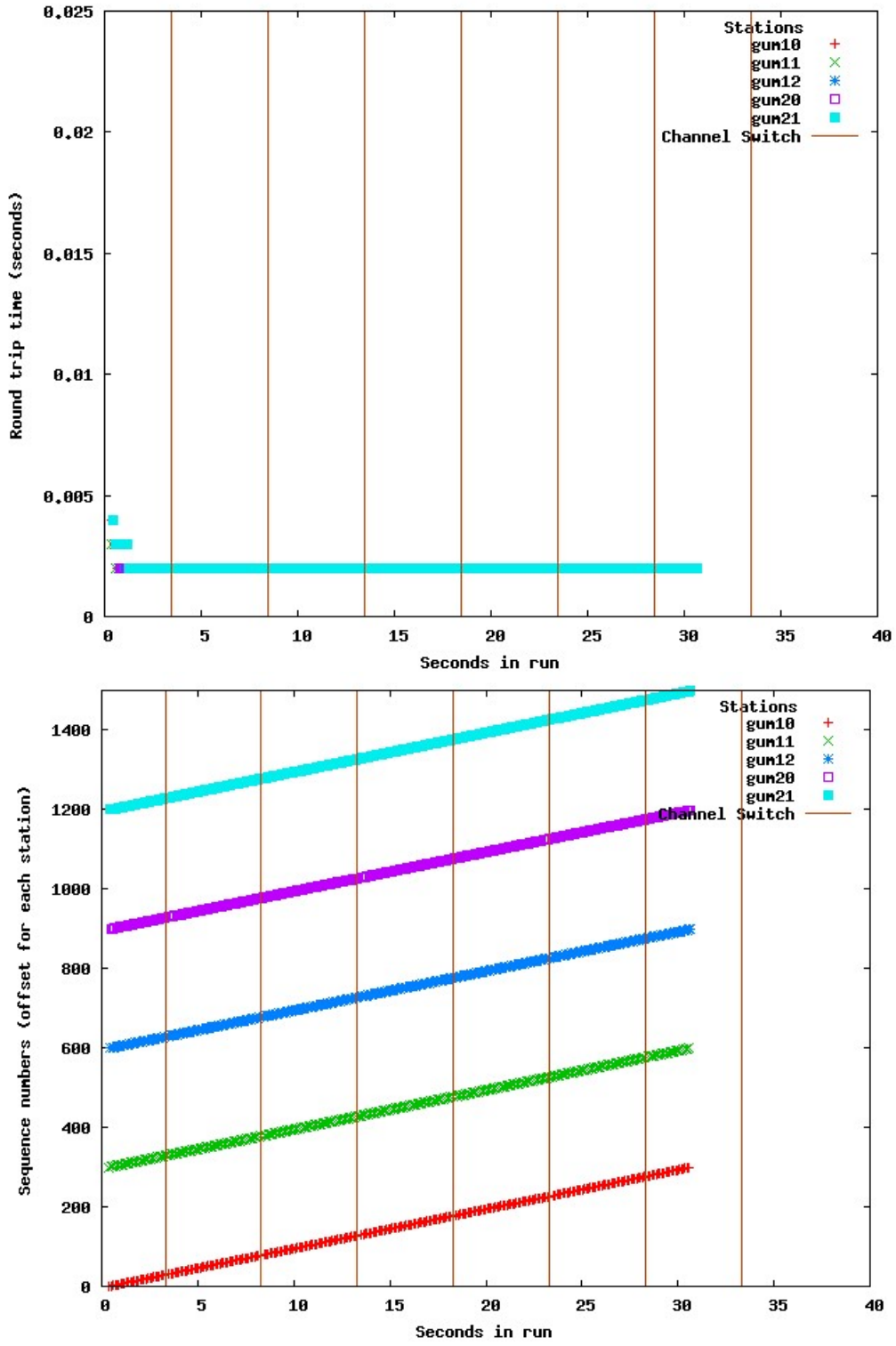


Figure 3.16 Infrequent channel switching, messages sent at 10 Hz. (t1); each channel switch is indicated by a vertical line, at that point the channel changes

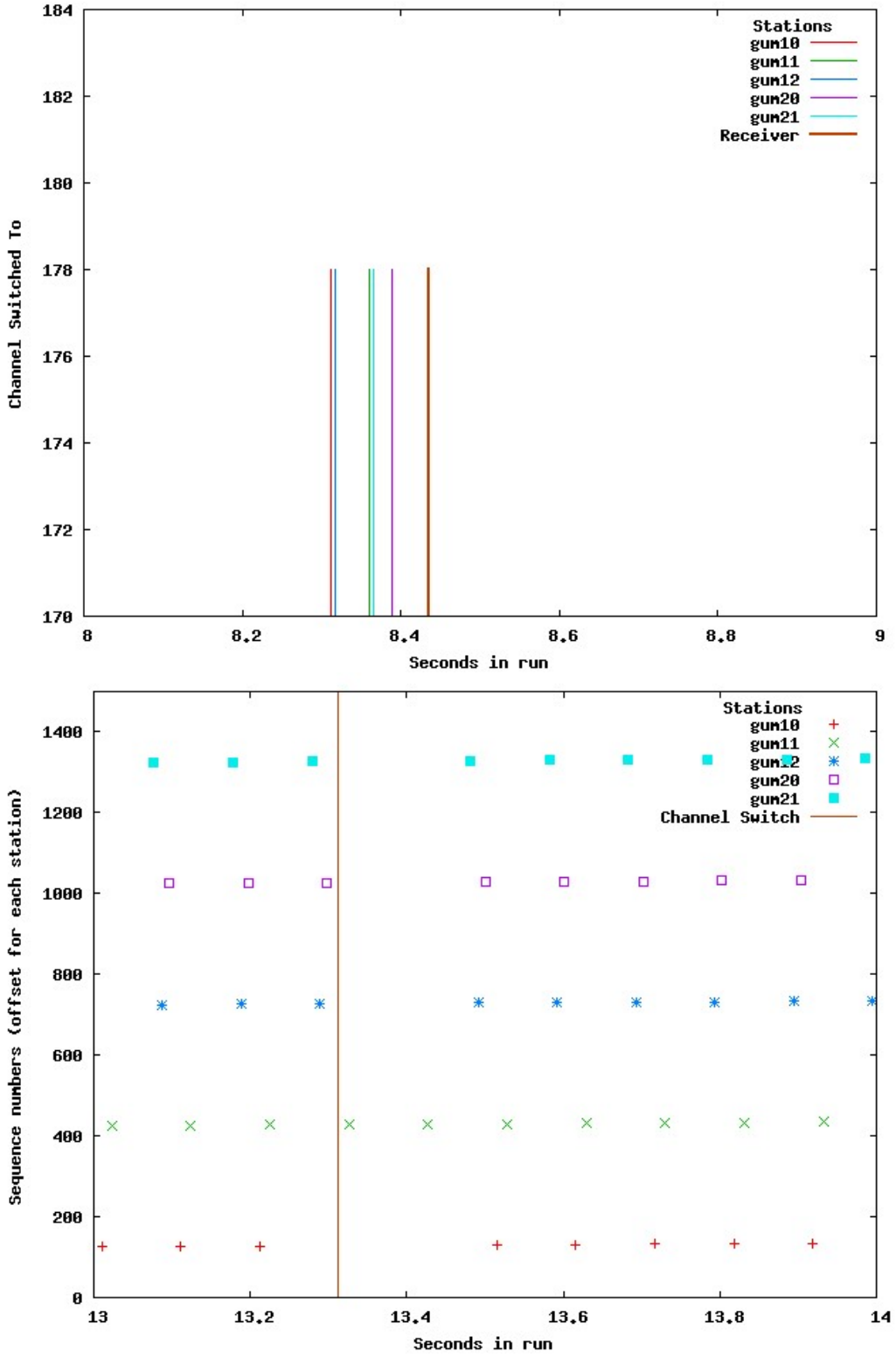


Figure 3.17 Close-ups, in case of infrequent channel switching (t1).

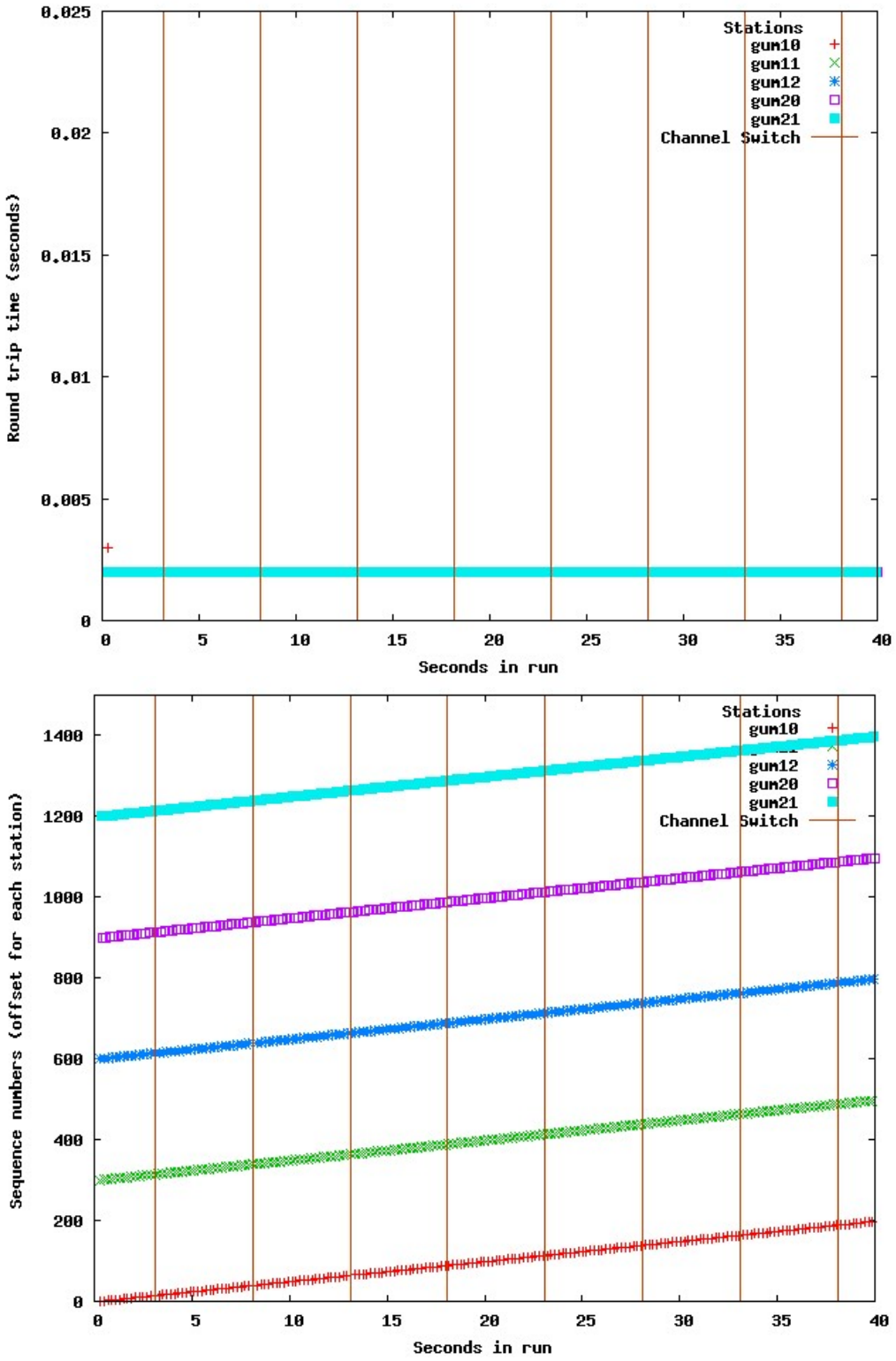


Figure 3.18 In this case, sw2, the send rate is half that of sw1, and round trip time remains low.

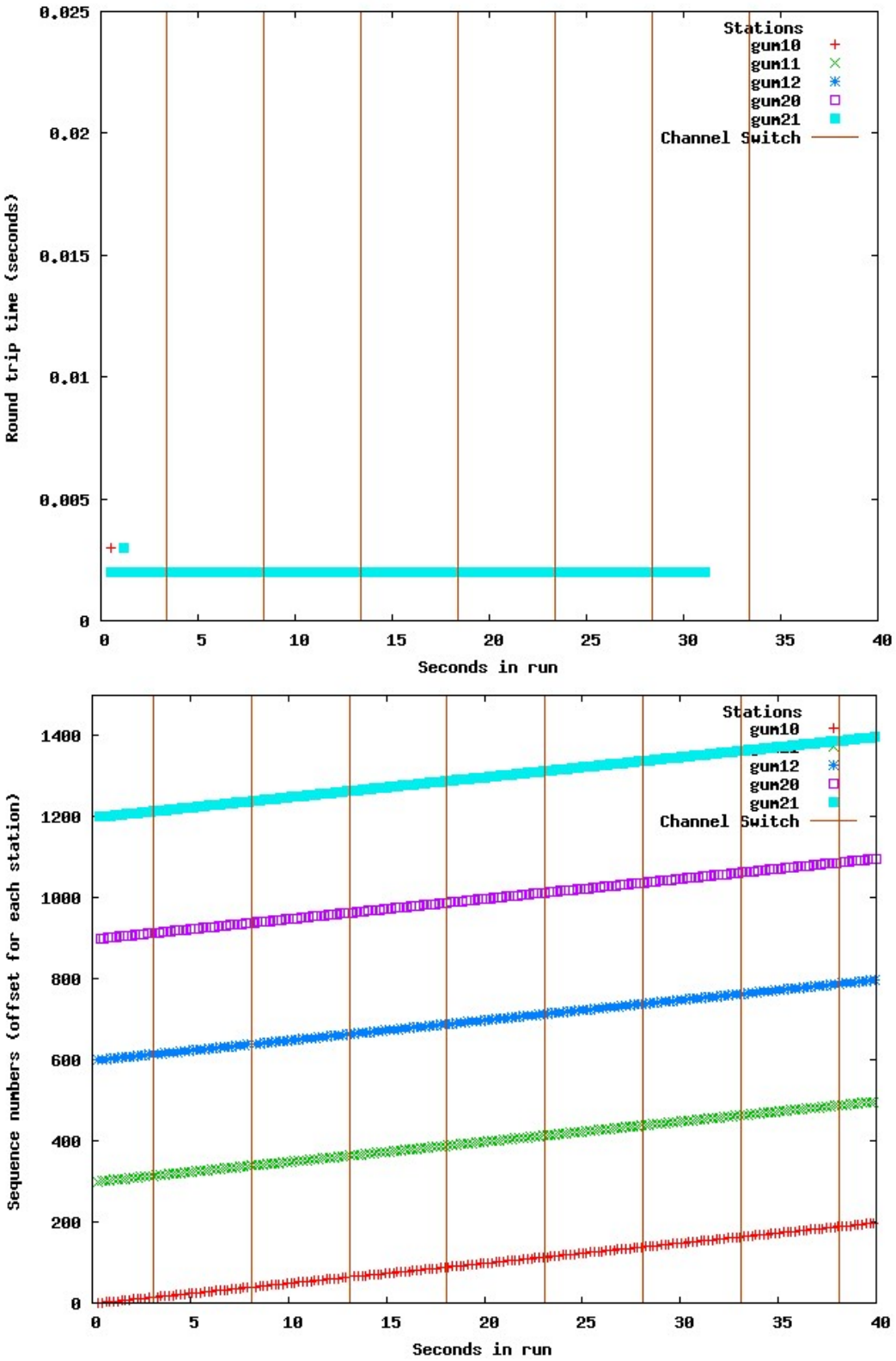


Figure 3.19 In this case, sw3, the send rate is twice that of sw1, and but the delivery of packets, as shown by sequence numbers remains good and the round trip time remains low.

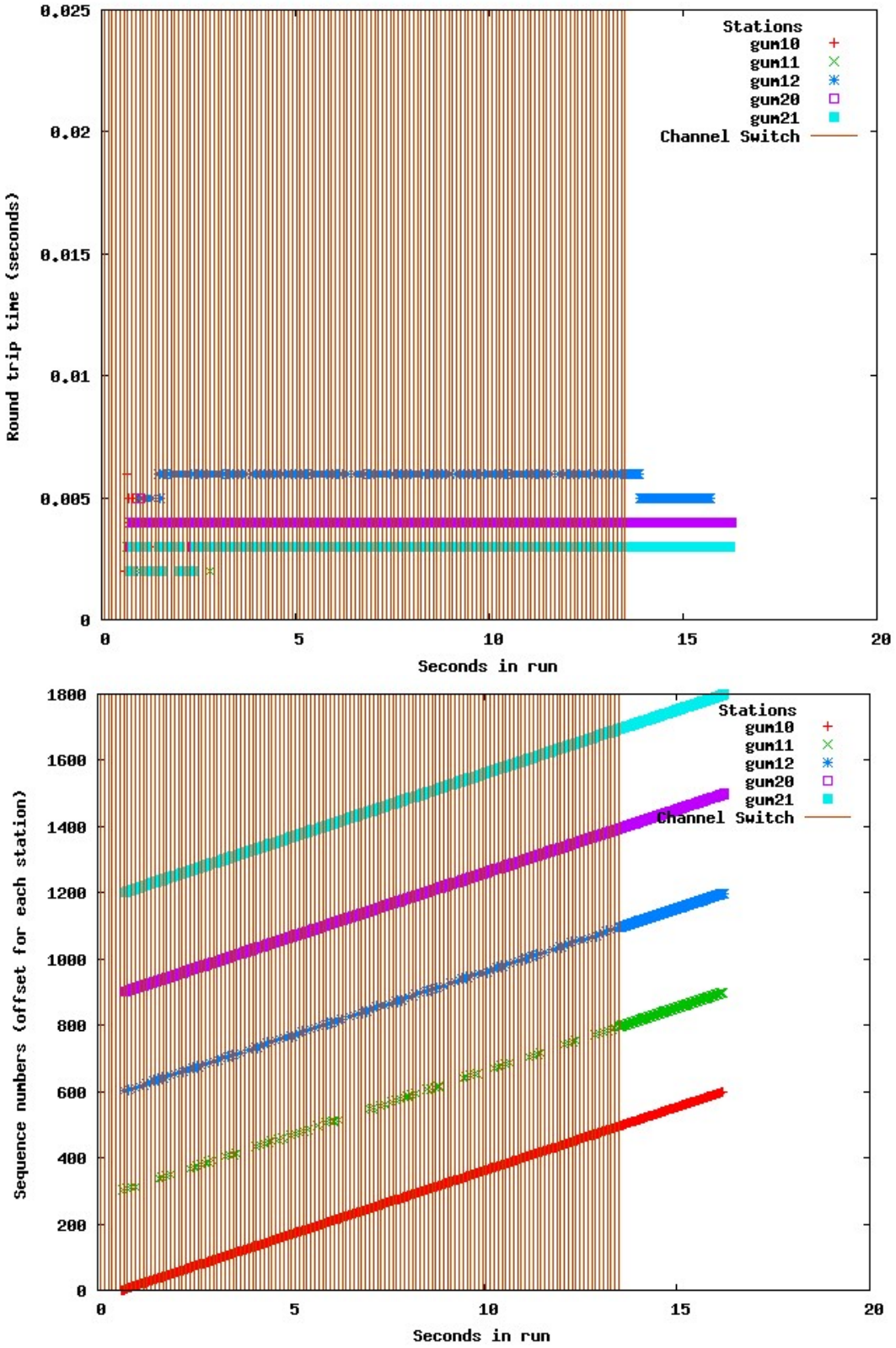


Figure 3.20 Case sw4, with increased overall load and increased channel switching.

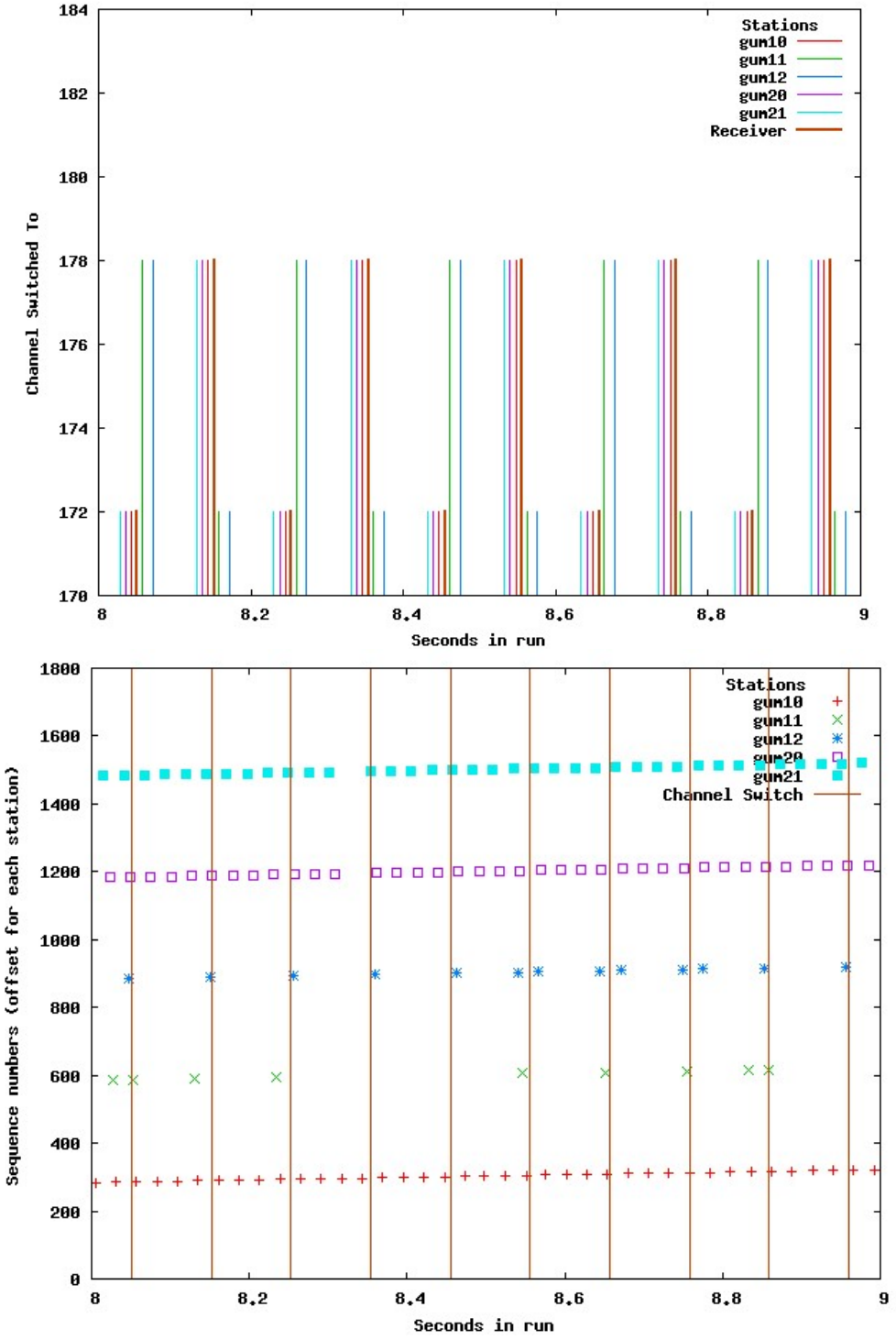


Figure 3.21 Close up of switching and sequence number loss for case sw4.

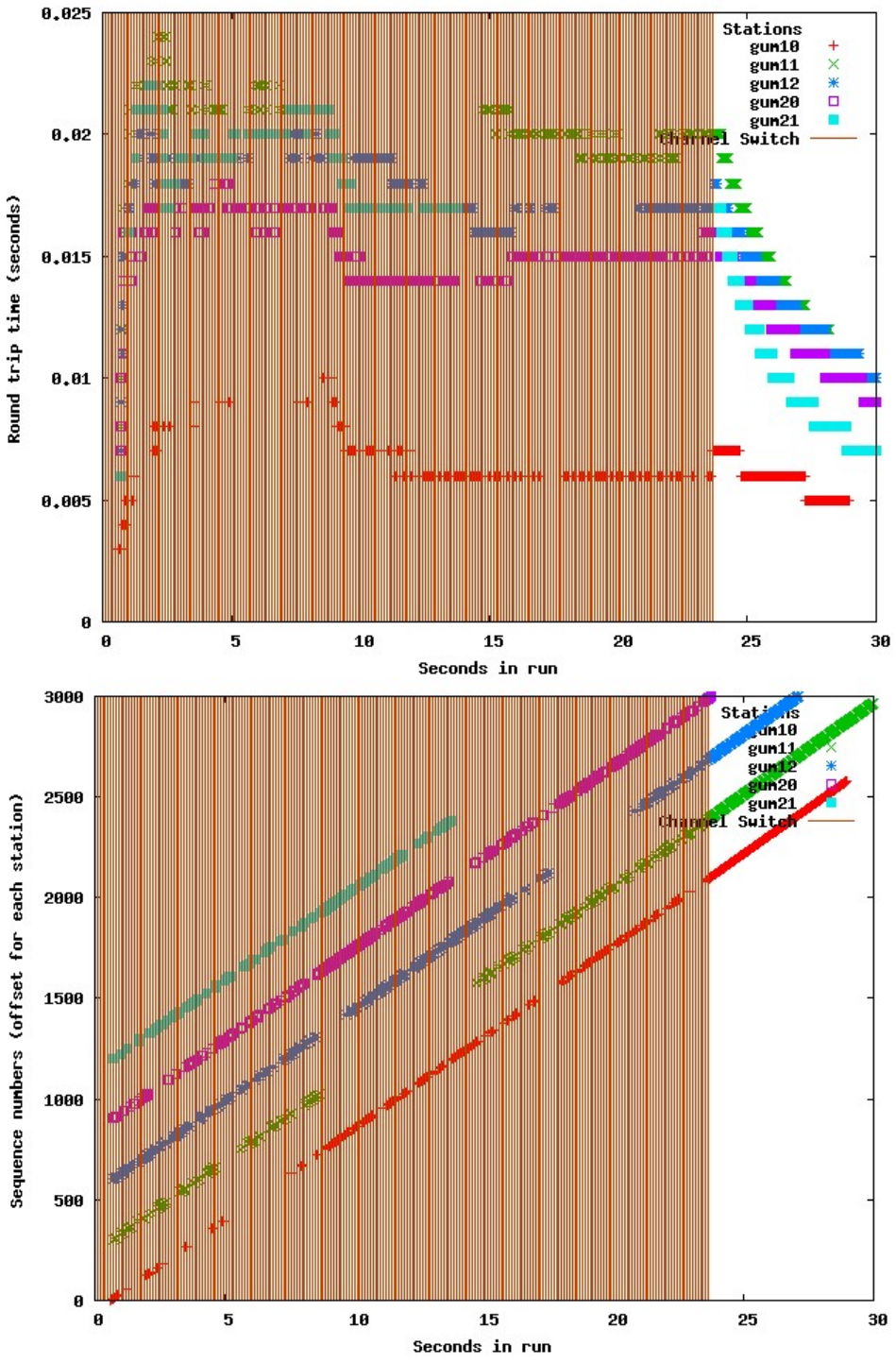


Figure 3.22 Case sw5 has increased loading and likewise increased latency and packet loss compared to case sw4.

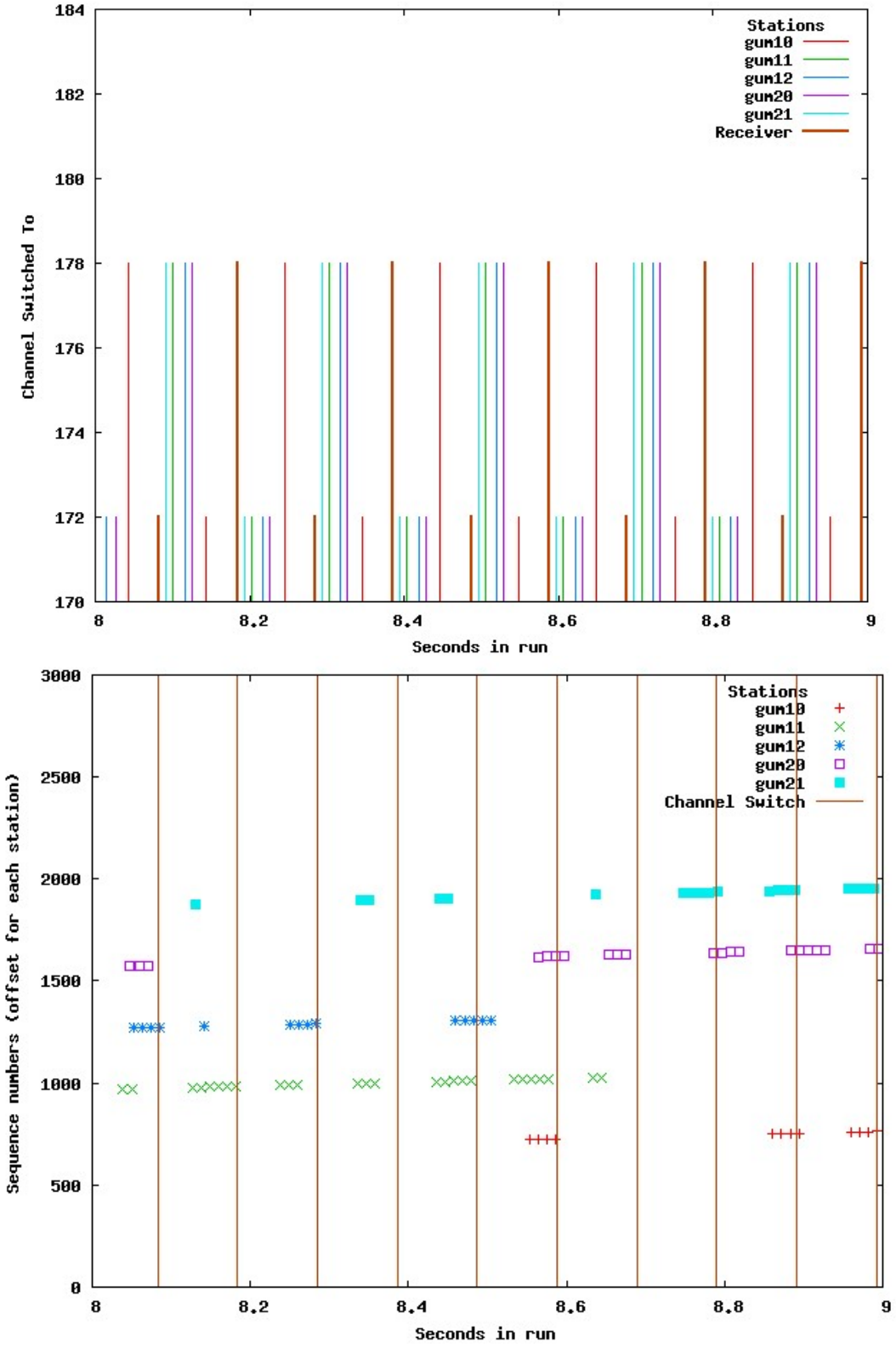


Figure 3.23 Close-up of case sw5 shows that the switching times for the different processors have spread still more, and the latency has increased to almost 25 ms.

3.3.6 Mobile Radios

As further validation of the test set-up, we carried out experiments with 8 radios on a single van at Richmond Field Station. We ran a variety of different loads and packet sizes, in a stable scenario where the van was parked at about 100 feet from the RSU, and in a moving scenario in which it drove a short distance along the test track. More of that data will be presented in a separate report, but as an example we show in Figures 3.24 and 3.25 two runs under the same conditions, except that one was moving and the other parked. In both cases the background channel switching is causing some disturbance. In the case of the moving vehicle the round trip time is higher almost from the start, and gets higher still as the vehicle moves out of range. Note for comparison with the next section that the overall load on the receiver is 160Kbytes/second.

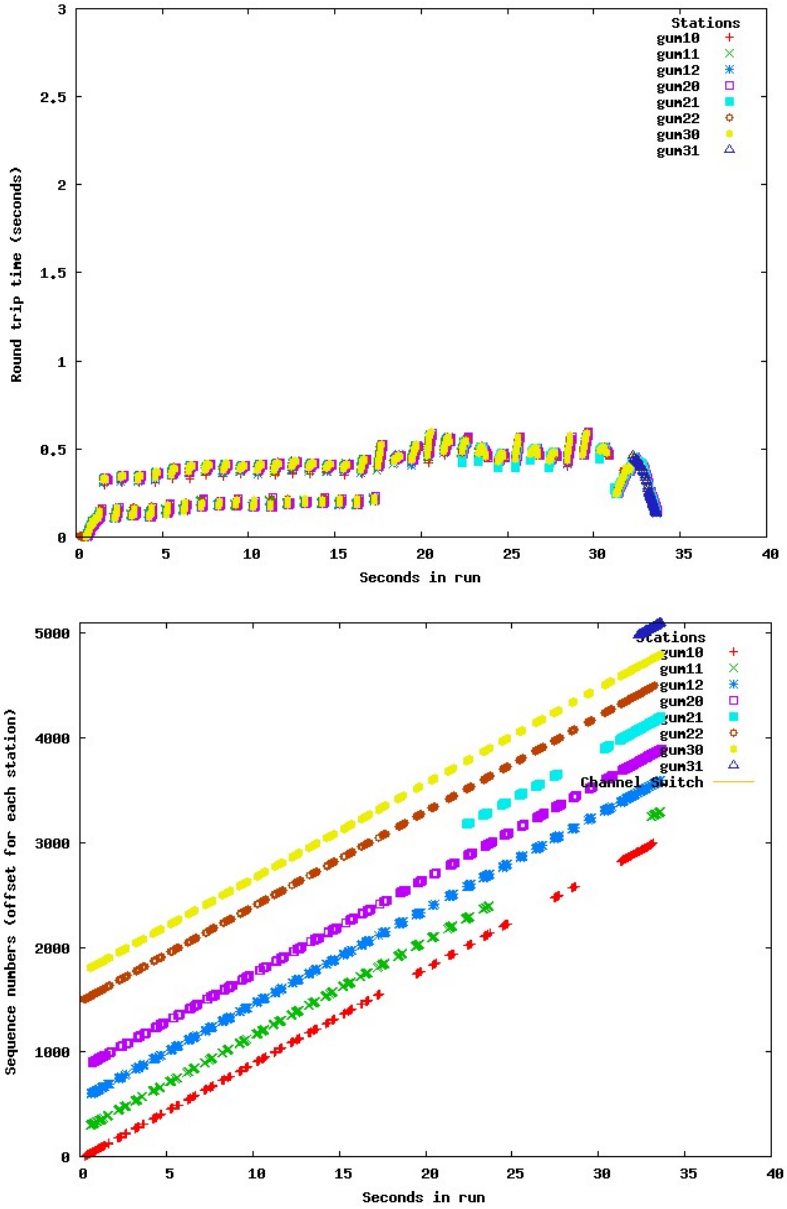


Figure 3.24 Eight radios, each sending 200 byte packets at 10 ms intervals, with background channel switching twice a second, from van parked within range of Richmond Field Station RSU.

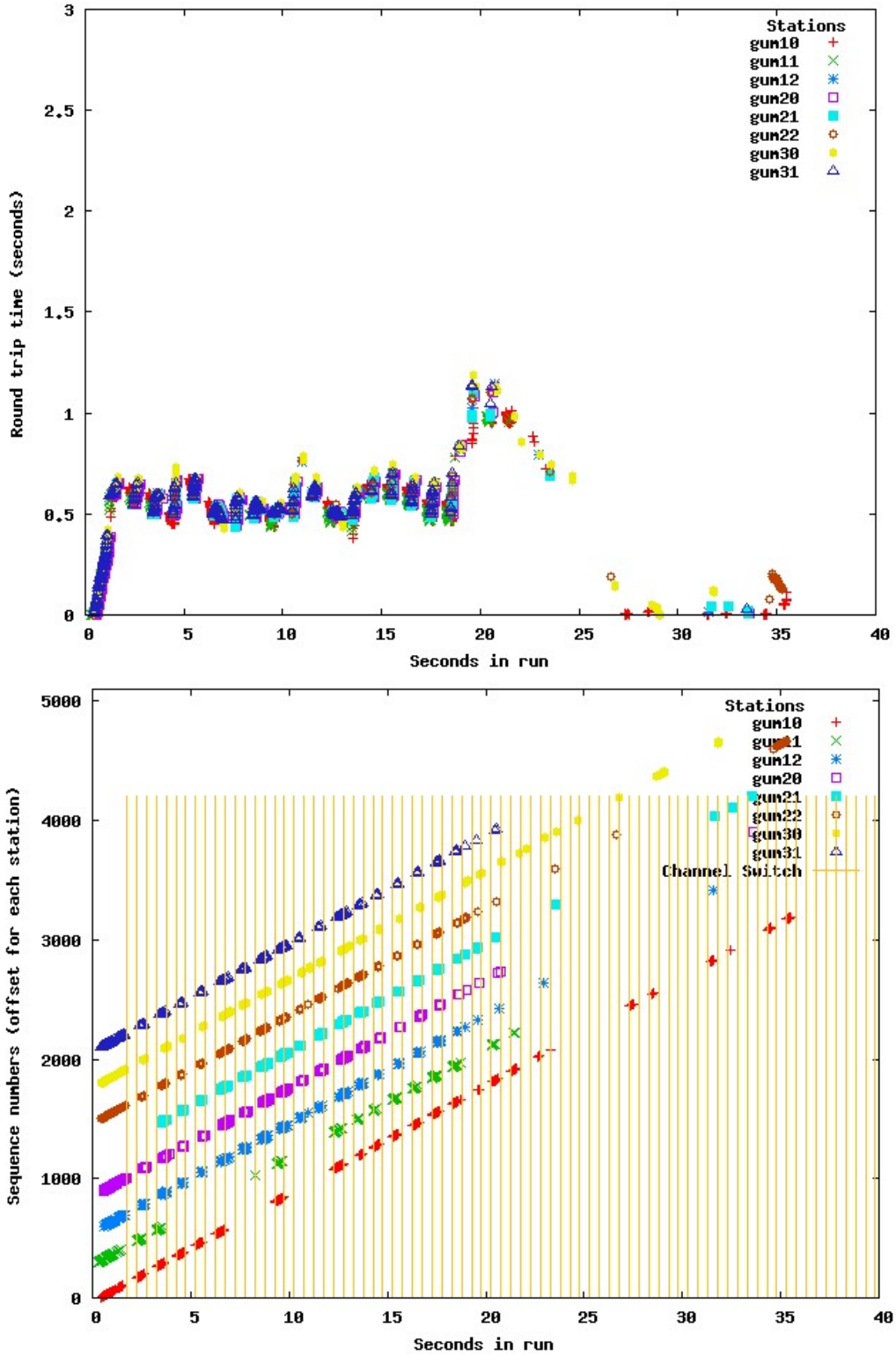


Figure 3.25 Eight radios, each sending 200 byte packets at 10 ms intervals, with background channel switching twice a second, during a short ride on a test track at Richmond Field Station.

3.3.7 Urban Intersection Testing

3.3.7.1 Packet length and interarrival time

Figure 3.26 Location of RSU at 6th and Brannan in San Francisco.

The data in this section were taken with the RSU that was installed at 6th and Brannan in San Francisco as part of the Innovative Mobility Showcase for the ITSA World Congress in the fall of 2005 (ITSA, 2005). The Denso WRMs for these RSUs were mounted on the signal arm, see the location in Fig 3.26. The wired interface of the WRM was connected by fiber to a PC104 in the traffic signal controller cabinet, running Linux. The software in the PC104 during this test was source identical with the data acquisition software running on the gumstix computers in the vehicle. These tests were conducted on September 24, 2006.

For all runs, data gathering started at the intersection with 5th and Brannan and continued for two minutes, while driving straight along Brannan to 7th street, turning right, coming back along Bryant, and then turning right at 5th. See illustration with dotted line in Fig. 3.26 of active part of the run.

There were three types of run, with low, medium and high overall load. For each set, the intervals between message sends were adjusted to maintain a constant load of messages at the RSU. In the first case the load was 40K bytes per second, in the second it was 100K, in the third 200K. For all runs on this day of testing the data rate setting in the Denso WRMs was set at 6Mbs.

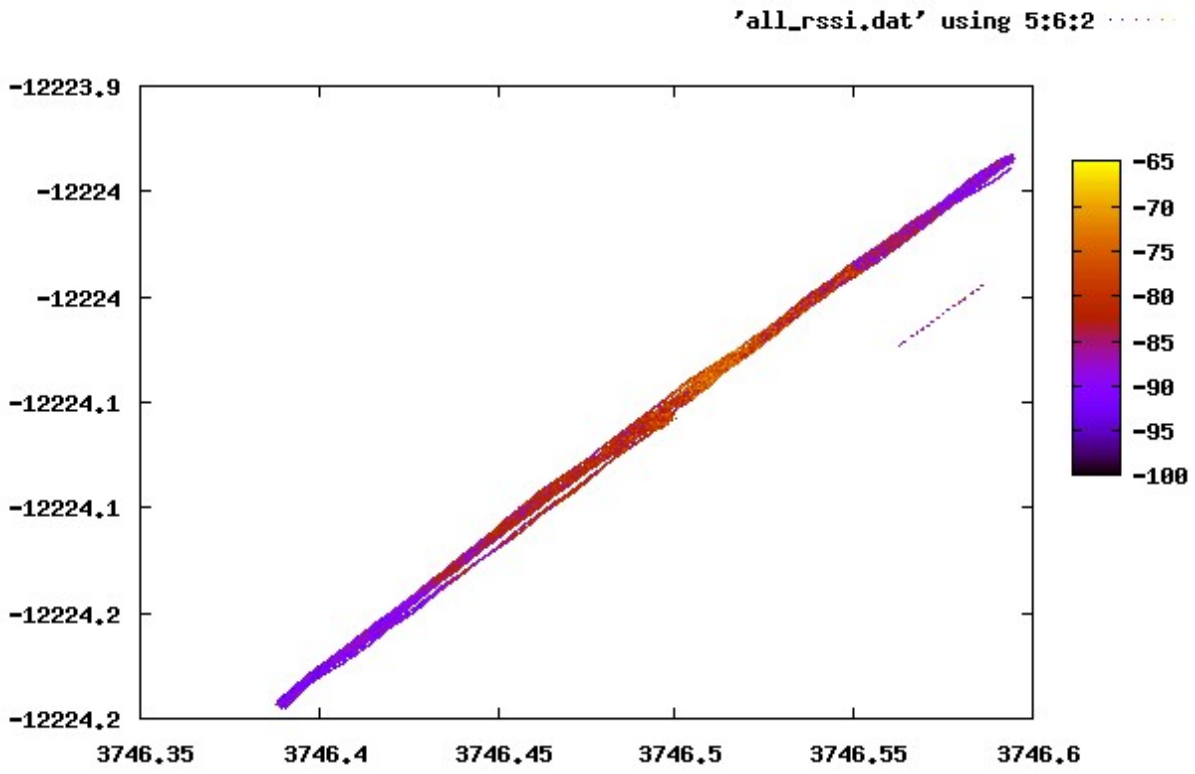


Figure 3.27 Plot of average of remote and local RSSI by GPS location for all received packets in runs at 6th and Brannan.

This data included two fields for local and remote RSSI, measured in dBm. RSSI values for all the runs are plotted together in against latitude and longitude in Fig. 3.27.

Variable parameters for each run were: Denso WAVE Radio Module rate setting, in Mbits/second; number of senders in run; packet size in bytes for this run; number of milliseconds in interval between packets; number of sends in the test (in some cases continuous and halted by user when out of range).

To approximate a model of transaction processing in which the traffic is predominantly generated on a service channel in response to short broadcasts on a control channel received, post-processing created the following summary data for each run: run type identifier string ; B total load for run (in kilobytes); R rate setting of Denso WRM in Mbs; S number of senders; M bytes per packet; I time interval between packet sends;•N number of sends in test; successful sends in test; average RSSI (over round trip);•start time (time of first reception);message time (time of reception for 8000 bytes -- line 8000/P);end time (time of last reception);•message period (message time - start time); connected period (end time - start time);minimum distance

from RSE; maximum distance from RSE; average distance from RSE; success rate (line count / possible transmissions; between start time and end time, i.e. while in range);•gumstix processor ID number.

In our figures we have graphed two measures of communications performance with respect to interval between sends. In Fig. 3.28 the success rates for each run are used as a measure of probabilities of successful round trip transmissions. In Fig. 3.29, the time required for successful reception of 8000 bytes is used as a measure of transaction time. Reading the graphs, for each line with a fixed packet size, the values on the left represent higher overall loads than the ones at the right. Due to testing time limitations, the set of run types with the highest load was completed only for 2 nodes and 8 nodes.

Looking at the highest load values for 2 nodes and 8 nodes, in these tests, for the same overall load, the 400 and 800 byte packet size tests (with smaller send intervals) have better performance in terms of success rate of packet transmission than the 200 byte packet size. While this is interesting, clearly much more testing is required to characterize the regions of operation for this effect.

Note that even with 8 nodes, although the probability of individual packet success is low for smaller sending intervals, because of collisions, at the overall loading used in these tests the time until an 8000 byte transaction is completed, from the time of first connection, stays under 2.5 seconds.

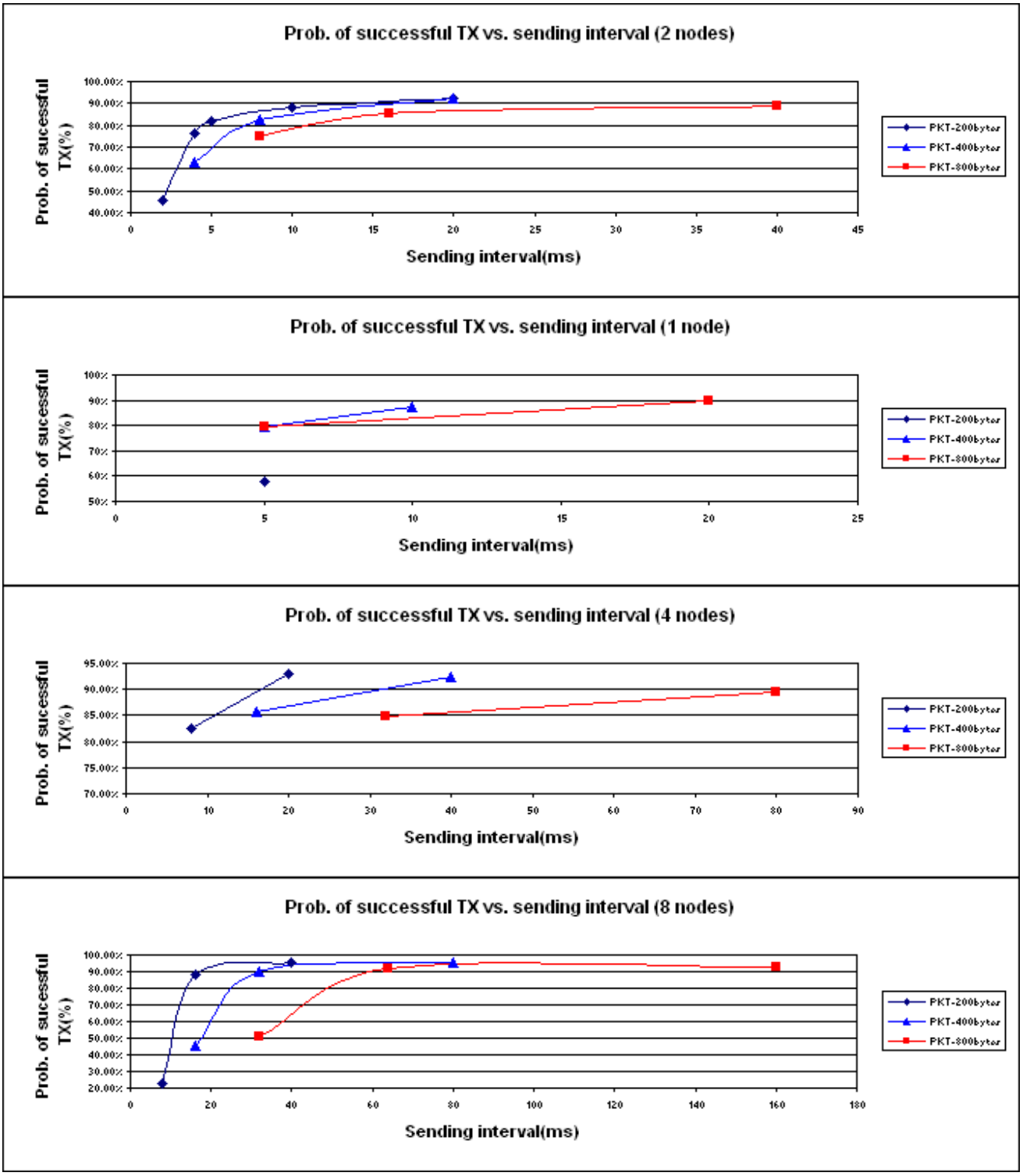


Figure 3.28 Probability of successful round trip transmission, as a function of the interval between packet sends, for 1, 2, 4 and nodes, with varying packet sizes of 200, 400 and 800.

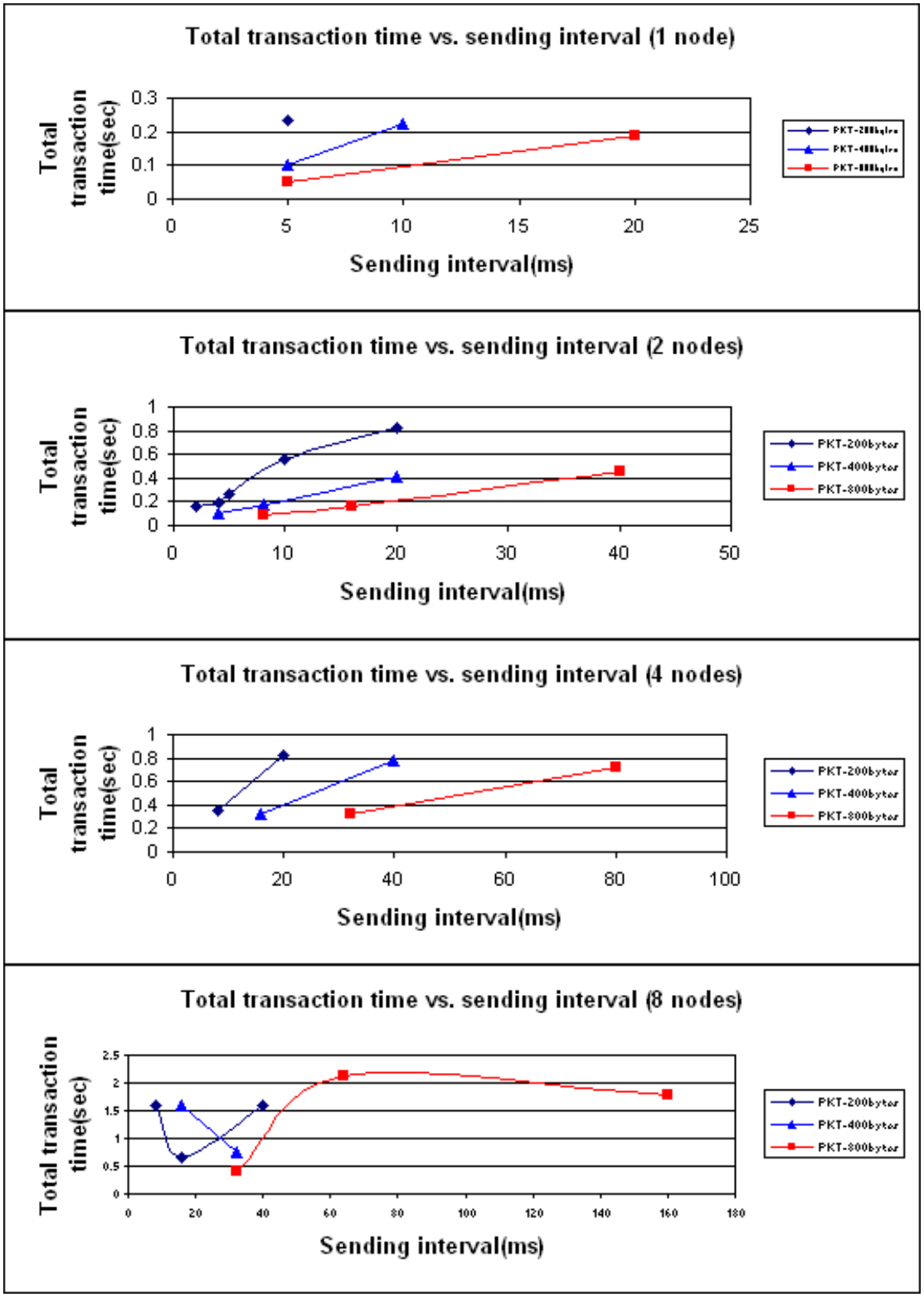


Figure 3.29 Time (in seconds) until the first 8000 bytes has been transmitted, as a function of the interval between packet sends, for 1, 2, 4 and nodes, with varying packet sizes of 200, 400 and 800 for tests at 6th and Brannan.

3.3.7.2 Data rate and queuing

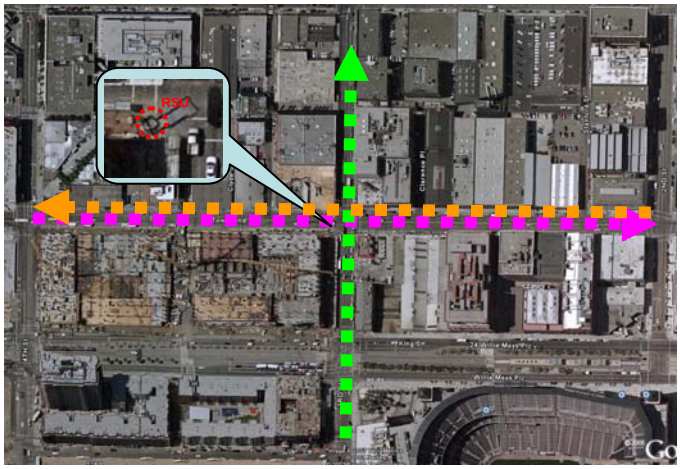


Figure 3.30 Location of RSU at 3rd and Townsend in San Francisco, and direction of test runs.

The data described in this section was taken with the 3rd and Townsend RSU in San Francisco (see Figure 3.30), another RSU that had been installed as part of the Innovative Mobility Showcase. Data was taken on June 15, 2006, just before the PC104 computer connected to the Denso WAVE Radio Module was removed because space in the traffic signal control cabinet was required for other uses.

Three different approaches were made to the RSU, from 2nd and Townsend past the RSU at 3rd to 4th and Townsend, from 4th and Townsend past the RSU to 2nd and Townsend, and from 3rd and Berry traveling on 3rd past the RSU. Additional data was taken while sitting still and with a background process switching channels. The channel switching data is not analyzed in this paper. No RSSI data was taken with this run; the code to record this data had not yet been integrated into the testing code.

For all approaches, packets were sent every 10 milliseconds, with either a short (200 bytes) or a long (1400 bytes) packet length. This gives an overall load, for 9 gumstix senders of 1.44 Mbps for the smaller packet size, and 10.08 Mbps for the larger packet size. In the presence of lost messages due to marginal RSSI or contention, backoff appeared to cause the transmission rate of the receiver to be lowered to the point where queues saturated the receiver, so that delays of several seconds for round trips were observed on some runs. On other runs, some of the senders seemed to either be completely blocked out or have dropped out for some other reason, so for these runs there were fewer than nine vehicles.

Fig. 3.31, 3.32, 3.33 and 3.34 show the transaction delay and the standard deviation of this delay as a function of the data rate setting. The most striking feature of the data is the high standard deviations and the very poor performance on the run on 3rd Street and the run from Townsend and 2nd to Townsend and 4th compared to the run in the other direction on Townsend and the run that was taken from a car parked in 3rd Street in close range to the RSU. In the presence of lost messages due to marginal RSSI or contention, the transmission rate of the receiver appeared to be lowered to the point where queues saturated at the receiver, so that delays of several seconds for round trips were observed on some runs. On other runs, some of the senders seemed

to either be completely blocked out or have dropped out for some other reason, so for these runs there were fewer than nine vehicles.

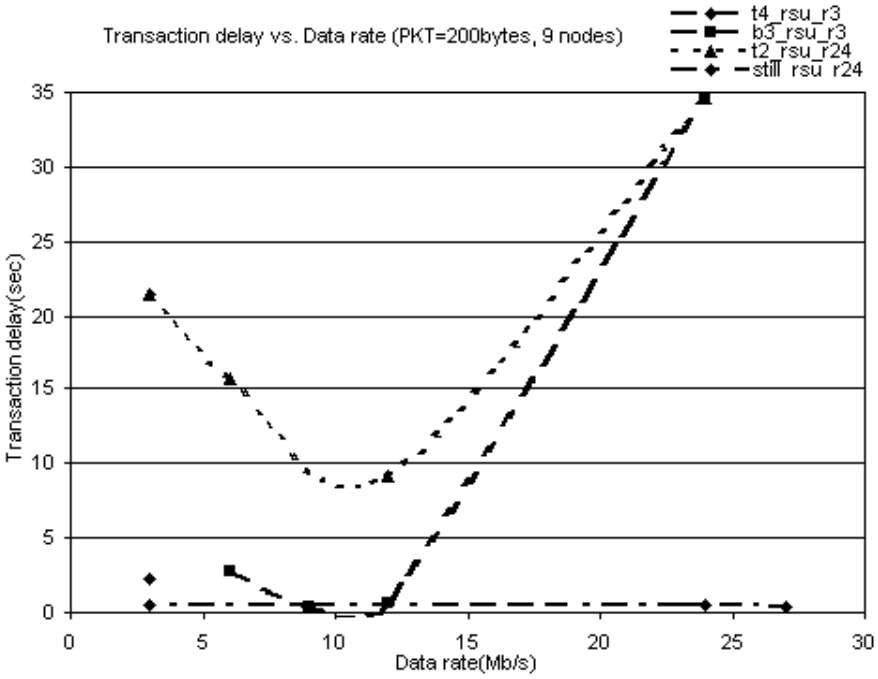


Figure 3.31 Transaction delay for 8000 byte transaction as a function of data rate, for 9 nodes, packet size 200 bytes, with various testing routes, 10 millisecond interval between packet sends.

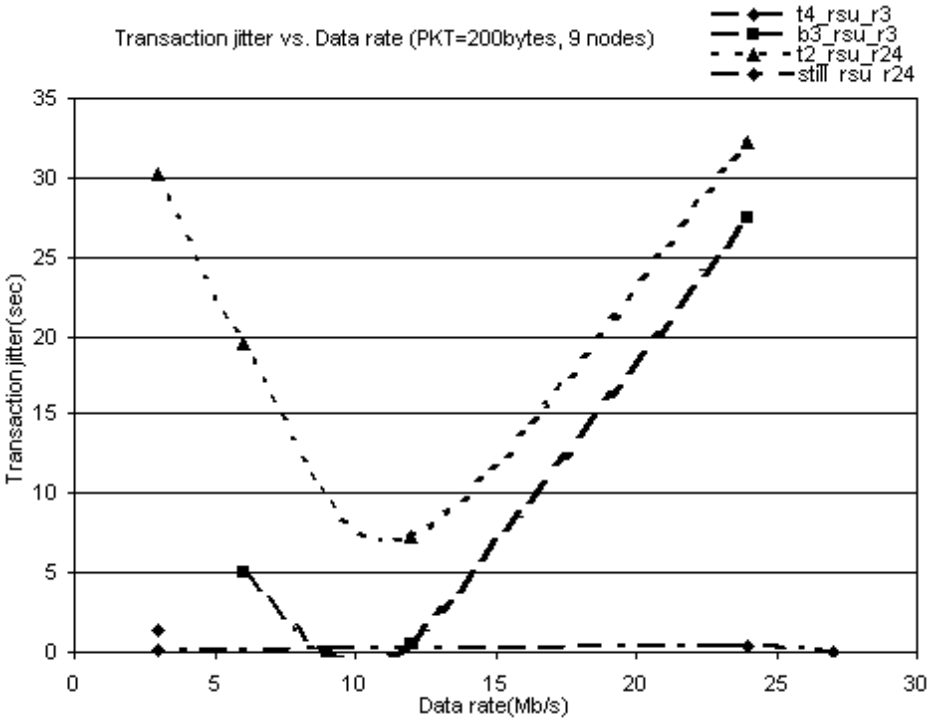


Figure 3.32 Transaction jitter (standard deviation of delay) as a function of data rate, for 9 nodes,

packet size 200 bytes, with various testing routes, 10 millisecond interval between packet sends**.**

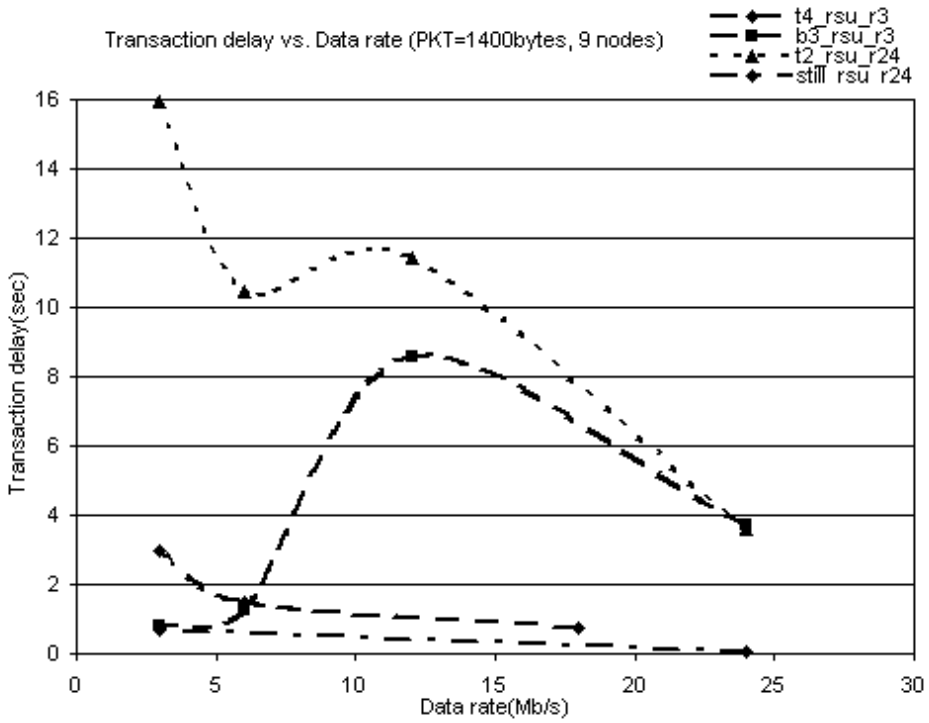


Figure 3.33 Transaction delay for 8000 byte transaction as a function of data rate, for 9 nodes, packet size 1400 bytes, with various testing routes, 10 millisecond interval between packet sends.

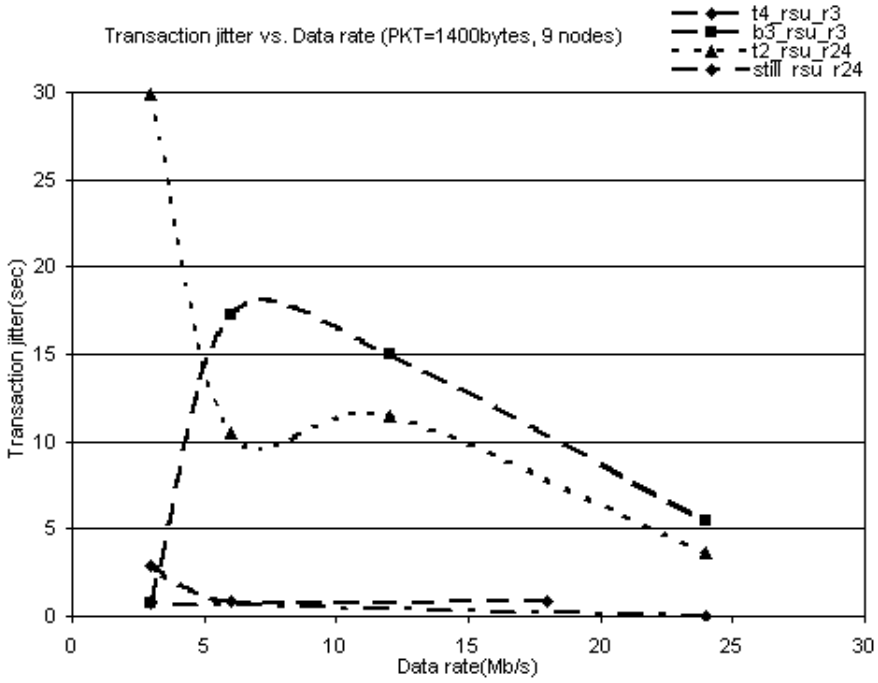


Figure 3.34 Transaction jitter (standard deviation of delay) as a function of data rate, for 9 nodes, packet size 1400 bytes, with various testing routes, 10 millisecond interval between packet sends.

3.3.8 Conclusion

The differing performance measurements we have obtained so far indicate the wide variations in performance that can occur with the same installed equipment, depending on variations in sending interval, packet size, and detailed geographical characteristics of the approach. In many of these runs, we see evidence of two regions of performance with respect to the sending interval (see Figure 3.35 for an illustration for probability of success transmission, and Figure 3.36 for transaction time). In region R1, the probability of a successful round trip transmission increases as the sending interval increases, because there are fewer collisions. In region R2, the probability of a successful round trip transmission appears to approach a limit. Since the sending interval is already long enough for contention resolution, the performance approaches a bound of *p* which is decided by PHY layer performance. With respect to transaction times, too low a probability of success will cause delays if the sending interval is too short, but over the optimal sending interval the transaction time will increase because of wasted time between packets. The points of optimality in Figures 3.35 and 3.36 need not be the same, and a great deal of testing, load characterization and site characterization will be required to determine these intervals in practice. It is necessary to be able to integrate data acquired at all layers of the network stack for complete understanding. We hope to use this data acquisition software and equipment to carry out further vehicle to roadside and vehicle to vehicle tests at VII California RSUs, in conjunction with more physical layer testing, and to development a methodology for site characterization that will allows us to predict the performance of communication with RSUs.

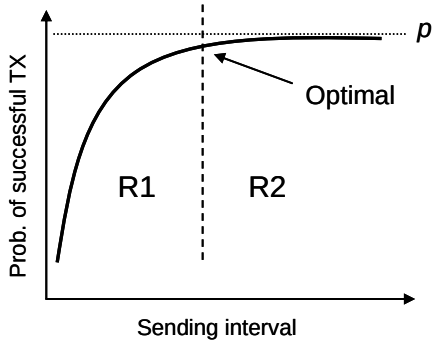


Figure 3.35 Illustration of regions of performance for successful transmission.

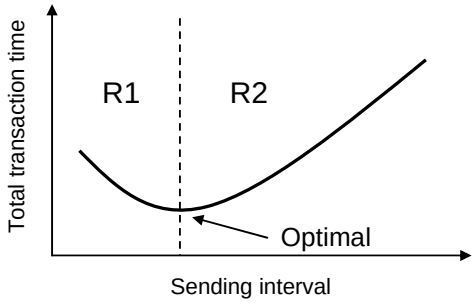


Figure 3.36 Illustration of regions of performance for transaction times.

4. APPLICATIONS

This chapter will discuss several applications that are currently being developed in association with the VII California effort. In addition, each application discussed will provide the current progress and a promising future outlook. The first section will discuss in details the Curve Overspeed Warning System (COWS), and then will proceed to discuss in details of additional applications organized in different sections including Traffic Probe Data Processing, Real-Time Arterial Performance Data, Intersection Safety, and lastly, High-Accuracy Nationwide Differential Global Positioning System (HAND-GPS).

4.1 CURVE OVERSPEED WARNING SYSTEM (COWS) APPLICATION PART I: DYNAMIC ROAD CURVE RECONSTRUCTION USING DIGITAL MAP

The Curve Overspeed Warning System (COWS) is one of the two cooperative active safety applications of VII California. (The other such time-critical safety of life applications that requires low latency, highly reliable and available vehicle-to-infrastructure and infrastructure-to-vehicle communication is Cooperative Intersection Collision Avoidance Systems (CICAS), not covered in this report.) The COWS system is aimed to integrate on-board sensors, digital map, GPS and the broadcasted information from the RSE to predict safety speed and provide speed advisory or warning messages to the driver.

To gain general applicability on production vehicles, the current system design adopted in this project utilizes a minimum set of common mobile sensors to achieve the required COWS functionalities. The prototype components include wheel speed sensors, yaw rate sensor, GPS with/without differential correction and a digital map. These four items would be coupled with a processor, communication devices, and a human-machine interface. The current research in VII California follows five main objectives:

1. Improve the integration of a digital map, GPS and mobile sensors for the COWS application
2. Conduct map-based road modeling for real-time road geometry estimation and prediction
3. Conduct dynamic lane-level vehicle positioning and detection with digital map and GPS or DGPS
4. Enhance map update by vehicle-infrastructure (V-I) communication
5. Enhance cooperative safety with V-I communication by estimating road surface condition and providing dynamic information to RSE

This section describes the preliminary results with respect to Objectives 1 and 2 above. It details the development procedure of a new method for dynamic reconstructing of road curvatures using digital map as well as demonstrating its effective in a prototype COWS system implemented in a PATH vehicle. While digital map has been applied to many Advanced Driver Assistance System (ADAS) applications, one critical attribute – road/lane curvature is not available in the existing digital maps. Curvature derivation using *Splines* is a well-known method in the computer graphics modeling community; however it was also found that the direct application of the

Spline method is often not appropriate for real-time safety applications due to the resultant discontinuities in the curvature estimates and the lack of robustness against map data errors [1].

This report presents a method to reconstruct road curvature attributes in real-time using existing digital map data based on the proposed Circle Center Search (CCS) and Circle Selection (CS) algorithms. Simulation and experimental results on a prototype system demonstrated that the proposed method can deliver curvature estimates that meet the desired accuracy[5] for a typical Curve Over-Speed Warning system.

4.1.1 Introduction

In-vehicle navigation and telematics systems which utilize GPS, digital map databases and wireless communication to receive external information, are becoming popular in recent years. The increased use of the in-vehicle navigation systems has also motivated the exploration of extracting road information from the digital map databases for other in-vehicle applications. In particular, many safety-related systems are candidates benefiting from the use of digital map information. While ADAS are becoming more and more popular today, map data and positioning information of navigation systems are being developed to improve driver assistance functions, which facilitates the development of Predictive Information and Assistance functions. It is believed that there is a significant potential for the use of a digital map and the vehicle's position to predict the road geometry and to track related attributes ahead of the vehicle. Moreover, ADAS-Applications can benefit from this potential, and new functionality may likely be enabled. A curve warning system, an embodiment of this concept, based on a navigation system and enhanced by VII is one of such application example.

Nowadays many driver safety assistance and stability control systems, such as Lane Departure Warning System (LDWS), Adaptive Cruise Control (ACC), Electric Stability Program (ESP), and Active Rollover Prevention (ARP), are often equipped on modern vehicles. These systems typically react to an event that has already occurred. The idea of dynamically extracting incoming road attributes from a digital map database and treating the map as a virtual sensor, and integrating such a sensor into the vehicle positioning system in order to improve the ability of predicting an impending safety event becomes very appealing. Efforts have been made to realize this concept in the ITS industry such as in the US EDMap project and in the European IN-ARTE, NextMAP, ActMAP, and PReVENT projects [CAMP, 2004; Tango and Saroldi, 2001; Pandazais, 2002; Otto, et al, 2003; Wevers and Blervaque, 2004].

The concept of "using the digital map as a virtual sensor" has been studied in some of the aforementioned projects in the sense of a "horizon-predictive sensor" to provide hot spot warnings to a driver; it has also been applied to enhance the GPS/INS-based positioning system for lane-identification purpose in this research project. This report presents a Curve Overspeed Warning System, which was developed under this concept. Figure 4.1 shows the block diagram of this COWS, which exploits digital map data in the following functional units: Map-Matching, Attribute Provider, Position Enhancement, and Safety Application modules.

[5] Reference [1] reports that a 10% curvature error is allowed for the near-term (after 2005) Curve Speed Assistance (CSA) (warning) application. The resulting vehicle speed error is approximately 4%, making the error within a 5 mph bound, which is acceptable.

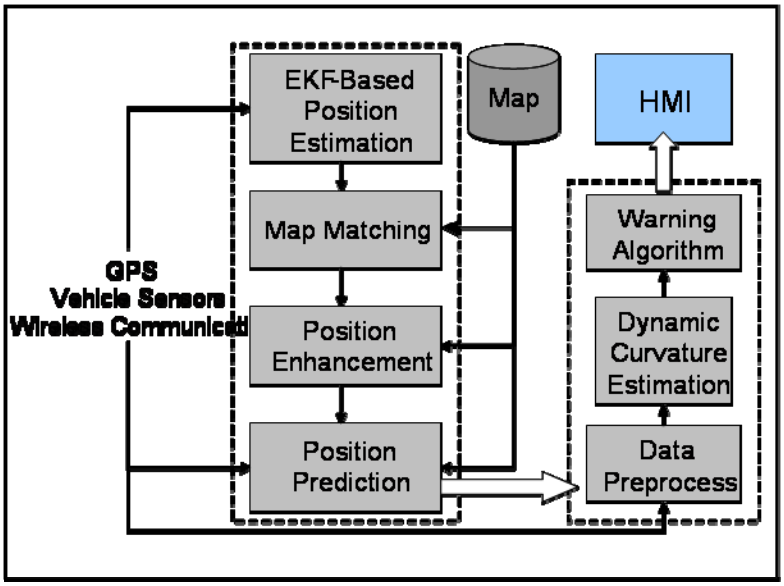


Figure 4.1 Functional Components of a COWS system

This report focuses on several prerequisites of the COWS: reconstructing dynamic road curve using digital map for Curve Speed Assistant[6] (CSA) system as well as assessing its applicability in COWS. Road or lane curvature, an important parameter for computing safe speed on a curve, is typically not available as an attribute in the existing digital maps. This attribute can be either dynamically estimated in real-time or extracted from a set of pre-processed data saved in the database[7]. However, practical considerations, such as data storage limitation, map update frequency, accommodation of complex or dynamic road attributes[8] (e.g. super-elevation, friction, and road grade) motivate the researchers to seek out methods that provide reliable road/lane curvature estimates based on the existing digital map data. Curvature derived from splines was used in EDMap project [CAMP, 2004], however, it was concluded that "systematic problems"[9] exist that prevent the curve warning application from directly using curvature derived from the splines," and "more work must be done to improve the methods used for representing curvature to enable the curve warning functionality." To provide more accurate real-time curvature estimates, a new approach based on locating the centers of curve sections, a Circle Center Search (CCS) algorithm, was developed and validated through simulations and experiments. The resultant curvature estimate errors, varying from 5% to 10% depending on the map data accuracy, meets curvature accuracy requirements (10% for CSA-W and 5% for CSA-C) suggested by [CAMP, 2004].

[6] Curve Speed Assistant includes Warning and Control functions (CSA-W and CSA-C) according to [1]. VII-based COWS is endeavored to be an enhanced CSA-W.

[7] Reference [1] presents two methods for curvature estimation: map-based and yaw-rate approaches. The former derived road curvature from the cubic b-splines, which is a representation method describing road and lane geometry in EDMap. The latter utilized yaw-rate and vehicle speed to estimate road curvature for CSA performance evaluation.

[8] Reference [7] , i.e. the AASHTO Green Book provides a minimum radius calculation formula, which is a function of superelevation, side friction, target vehicle speed and gravity.

[9] Reference [1] indicates several systematic problems, such as missed or false detection of curves, which are mainly due to discontinuity in the curvature data derived from splines.

The outlines of this report are as follows. Section 4.1.2 describes typical map-based ADAS systems, in particular the COWS, as well as the framework of an enhanced ADAS system using vehicle infrastructure integration. Section 4.1.3 proposes a new method for curvature estimation based on the existing digital map and addresses the related design issues. Comparisons between the conventional spline approach and the proposed method are also provided. Section 4.1.4 shows the preliminary experimental results conducted at the Richmond Field Station for a prototype Curve Overspeed Warning System using a PATH vehicle. Conclusions are made in Section 4.1.+5.

4.1.2 Digital Map as a Virtual Sensor for Advanced Driver Assistance System

4.1.2.1 Overview of Digital Map and ADAS

Over the years, the specifications of a digital map have gradually evolved to fulfill the requirements of various emerging in-vehicle applications, such as GPS-based navigations and ADAS applications. A typical digital map is a geographical database containing road geometrical and attributive data. Based on a specific road network model, the geometric relationship and road features can be described based on certain rules. Geographic Data File (GDF) map is currently an international ISO standard for navigation maps. It is both a data model and an exchange format for digital maps that can accommodate different map contents and data formats defined by individual map makers [Wevers, and Blervaque, 2004)]. Hence, different navigation systems can interpret all digital maps following this standard.

A modern navigation system typically has two main components: a map database and a GPS-based positioning unit that can be integrated with “dead reckoning” sensors such as gyroscope and odometer. The calculated vehicle position can be “projected” to the most likely point on the map and be associated with a road segment. Through this “map-matching” technique, the relevant road attributes and information can be extracted from the database. Given the destination provided by the driver and the continuously updated actual positions, the navigation system guides a driver to his destination through the best route.

In contrast to the relatively simple structure of a navigation system, ADAS applications often require two additional subsystems, ADAS Horizon Provider (AHP) and ADAS Horizon Reconstructor (AHR), to extract and process data for the ADAS computation [Wevers, and Blervaque, 2004]. The processed ADAS specific map data (in both of geometrical and attributive contents) is called the ADAS horizon (AH), which is the extracted map data around the current position as well as in the “look-forward” direction. The “low-level” data extraction and aggregation into AH (2D) is executed by AHP based on the map-matched position received from the positioning unit. AHR receives AH (2D) from AHP and transmits only the incremental updates of AH (1D) to the ADAS application, i.e. AHR functions as a “filter” to accommodate bandwidth constraints in the application side. This interface between AHR and the application is referred to as the “high-level interface” since it can be developed as an internal API of an application.

Note that the correctness of the AHP and AHR functions hinges on accurate map-matched positions, especially for the high-accuracy ADAS applications such as Lane Following Assistant, Forward Collision Warning, and Curve Speed Warning or Control. Both the vehicle positioning

system and the digital map may need to be enhanced [CAMP, 2004] for such ADAS applications. In addition to accuracy, the map database needs to be updated frequently to guarantee reliable up-to-date road information. As a consequence, an easily accessible and cost-effective map updating method needs to be provided for maintaining such map-based ADAS applications. VII is one such application candidate [Misener and Shladover, 2006] and its concept of operation is described below.

4.1.2.2 Architecture and Advantages of the ADAS and COWS Enhanced by VII

Figure 4.2 shows the main components of this new ADAS system enhanced by the vehicle and infrastructure integration in this report. This system integrates (D)GPS, vehicular sensors, wireless communication (e.g. Dedicated Short Range Communication: DSRC), and digital application map to perform safety-related applications such as curve overspeed warning, stop sign warning and junction/intersection speed warning. The application map is developed based on the commercial GDF map provided by NAVTEQ. Since most current digital maps are "road-level" maps, detailed road/lane attributes, such as number of lane, lane width, and stop sign location, are not available but are required for many safety applications. Therefore, this application map was generated by extracting relevant road geometrical and attributive information from the existing GDF map and adding the aforementioned detailed attributes to the new database.

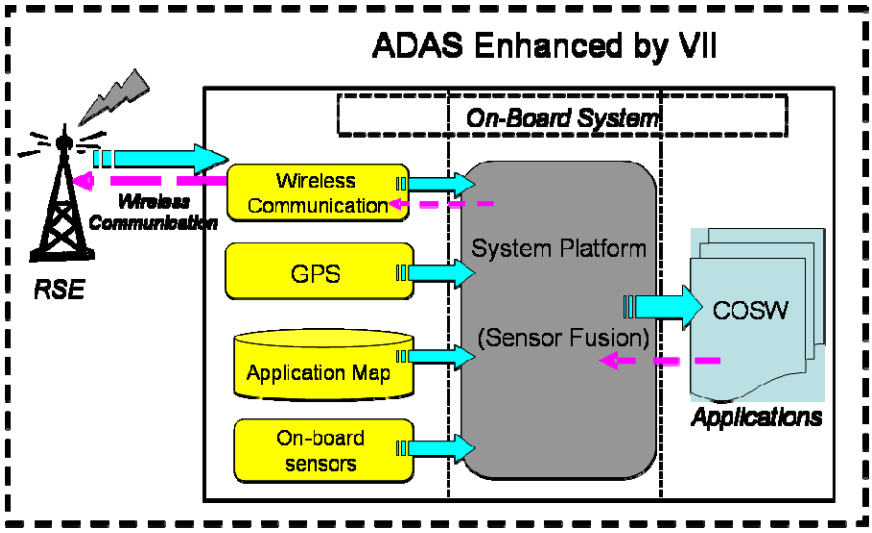


Figure 4.2 Concept of ADAS Enhanced by VII

The on-board system platform consists of the following functional modules: EKF-based GPS/INS positioning unit, Map-Matching processor, Attribute Provider, Position Enhancement unit, Safety Application module, and the Wireless Communication unit. The advantages of this system design are described as follows.

The system can not only operate in a stand-alone fashion using on-board sensors and digital map, but it can also operate in a cooperative enhancement manner through VII. When there are RSE)

nearby[10], the equipped vehicle can exchange data with RSE via V-I wireless communication, e.g. DSRC. Detailed or dynamic road attributes such as super-elevation, grade, friction, traffic signal at ramp end (metering) etc., as well as event-based messages such as traffic accident, lane closures, detour, or construction zone, can be transmitted from the RSE to the vehicle. In addition, it is possible to provide GPS correction signal and map update service to the on-board system through VII [CAMP, 2005]. On the other hand, the equipped vehicle can transmit detected or computed data such as speed advisory, air bag activation and ABS activation, to the RSE, so that RSE can inform other nearby drivers of an incident ahead.

The system also employs a positioning enhancement unit to improve the map-matched position for lane identification using a "road-level" positioning unit. As opposed to the "lane-level" vehicle-map positioning system, which relies on measurements from the lane-level positioning unit (e.g., a RTK-DGPS integrated with a MEMS IMU described in [CAMP, 2004]), the proposed system has advantages in terms of robustness[11] and cost-effectiveness.

The current and future road/lane curvature can be estimated in real-time based on the road "node" positions, and lane attributes, i.e. lane width and number of lane. The details will be described in the next section.

4.1.3 Dynamic Road/Lane Curve Reconstruction Using Digital Map Data

This section focuses on one technical issue regarding road/lane curvature estimation using map data in certain specific ADAS applications such as Curve Over-Speed Warning: estimating in real time the radius of curvature, curvature direction, and the position of the curvature center. More complex road attributes such as super-elevation, grade and side friction, are expected to be taken into account for further curvature refinement[12].

Since the road or lane curvature is not an available attribute in the existing commercial digital map, this important attribute should either be estimated in real-time, or be created off-line and saved in advance before operation. However, if a reliable method utilizing the existing map data can be developed to estimate the road/lane curvature in real-time, significant data storage space can be saved.

The following sub-sections first explain the feasibility of the curvature derivation based on splines, a well-known mathematical tool in the computer graphics modeling community. Then a nonlinear filtering approach using CCS and CS algorithms will be proposed and compared with the splines approach.

---

[10] The California Partners for Advanced Transit and Highways (located at University of California, Berkeley) is working with Caltrans on the infrastructure (roadside unit) implementation and vehicle-infrastructure messaging and communication of the VII data, whereas Telvent Farradyne is working with MTC on backhaul communications and collection, processing and archiving of data at the center.

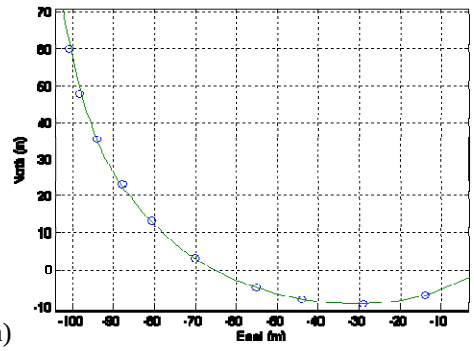
[11] [1] reports that "..., many of the lane-level applications were not able to gracefully degrade from WHICHLANE to WHICHROAD functionalities. Whenever poor positioning would cause the vehicle to not be matched to a lane, the map would report it "off road" Once the vehicle was determined to be off road, the ADASRP (Advanced Driver Assistance System Research Platform) map access software would stop to output mapplets (CAN data messages) and the lane-level applications were forced to abruptly cease operation."

[12] In this report, those complex road attributes are not taken into consideration for curvature estimation.

4.1.3.1 Conventional Approach – Spline

Spline is a common tool for road geometry representation. A spline typically refers to a wide range of functions defined piecewise by polynomials that are used in applications requiring data interpolation and/or smoothing. Using the spline format, it is possible to interpolate road position coordinates at any point along the spline based on the discrete map positions called nodes. Road attributes, such as headings, tangents, and curvatures can also be derived using the resulting splines. It is therefore possible to deliver the road geometry information to the ADAS applications in the spline format. However, curvatures derived from splines may not be directly used by the ADAS applications as found in [CAMP, 2004].

As shown in Figure 4.3(a), the curve represents the road center line of an exit ramp from northbound highway I-580 to Bayview Avenue. It is constructed by the cubic b-spline based on the "node" positions. This curve seems to be well-described by the spline function, y = f(x). One can derive the road curvature by the following formula using this spline function: $K(x) = \frac{1}{R(x)} = \frac{\left|\frac{d^2y}{dx^2}\right|}{\left[1+(\frac{dy}{dx})^2\right]^{3/2}}$ where K(x) and R(x) are the resulting curvature and radius, respectively. Figure 4.2(b) shows the estimates of the curvature radius by this approach.

(a) 

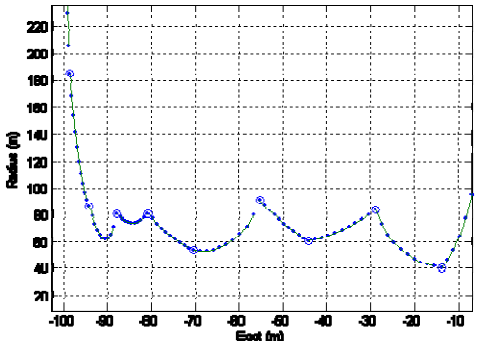
(b) 

Figure 4.3 (a) Road curve fitted by splines (b) Radius of curvature derived from splines

Two main problems can be observed from this example. First, errors of the curvature estimates at the curve ends are large, which is due to the insufficient positional constraints. This type of errors would appear frequently especially when such data are delivered by the "incremental" AH. Second, the variance of curvature estimates is also large, and it can be detrimental to the performance of CSA as it may cause large undulation in the resultant safety speeds.

In addition, curve-fitting by splines cannot guarantee "robustness" against the positioning inaccuracy because of its intrinsic property[13]. Unfortunately, the map data has relatively high positional inaccuracy due to the sensor noise and the accumulated errors in the repositioning process [CAMP, 2004]. Road information derived from splines may be mathematically correct but not necessarily physically justified[14].

[13] An example of the effect of subtle road position variations on the curvature is shown in Figure B-22 of reference [1].

[14] Reference [1] indicates that "*Due to the sparse nature of the shape points, much of the road shape must be mathematically re-constructed between the shape points to ensure fair curvature*"," it was expected that much of the spline's deviation from the road center would occur

4.1.3.2 New Approach – Data Pre-filtering and Parametric Curve Fitting

The new approach presented in this report includes two parts: data pre-filtering and parametric curve fitting. As discussed in the last section, curvature derivation from splines, a non-parametric cure-fitting approach, may not be able to be directly used by the ADAS applications. The main difficulty lies in the discontinuity and high variance in the curvature estimates. Although a possible remedy to the variance problem is adding some sort of filtering to the estimates from splines, the discontinuity in the curvature estimates would be more difficult to resolve in that the discontinuity could be a real road feature or it could be a result of positional inaccuracy. While an arbitrary filtering process may lead to missed detection of a curve, an un-filtered discontinuity due to "noise" could also cause false detection[15].

The above two problems raise some fundamental questions: "Is a non-parametric curve-fitting method effective for curvature derivation?" And "Is it possible that the shape of a road can be modeled in such a way that the resultant road curvature can be correctly estimated in a piece-by-piece fashion using this specific model assumption[16]?" In addition, for time-critical applications, computation efficiency is another important issue. Considering all the above factors, a parametric curve-fitting approach that captures the "dynamics" of road curvature as well as rejects the "noise" from the road data is explored.

Figure 4.4 depicts the concept of this new approach. As shown in this figure, a vehicle is traversing a curve with speed V (at C.G.). This curve consists of several nodes, N1, N2 and N3, recorded in the digital map. For simplicity, the steady state curvature is of our interests[17], i.e. R1=R2=R3, C is the center of curvature, and R_v represents the turning radius of the vehicle. Assuming further that this vehicle is in quasi steady state, namely, the magnitude of V is constant and the center of turning radius ideally matches the center of road curvature C. Therefore, one can easily computed the radius of road curvature and curvature center based on the node positions N1, N2 and N3 regardless of positional inaccuracy (Appendix 1). The safety speed can then be predicted based on the estimated curvature and the desired lateral acceleration before entering this curve.

As a matter of fact, drivers normally approximate the road curvature ahead based on the visible lane stripes and regulate the vehicle trajectory accordingly. The main idea of this new approach is to model the road shape from a driver's perspective. Since a vehicle at a fixed steering angle and constant speed follows a constant arc, a road curve can then be consisted of one or several arcs, and each arc (a portion of the circumference of a circle) has a constant curvature. For a straight road, the arc has an infinity radius of curvature. If the curvature of each arc can be identified, for example, by fitting the circle function as in Equation (1), to the node positions of arc, the road curvatures can then be obtained piece by piece. Hence, the road curvature

---

between the shape points" and "*Database road center link geometry definitions and connectivity definitions impose limitations in modeling actual road center shape while maintaining fair curvature.*"

[15] Reference [1] reports 13.0% missed detection rate and 26.1% false detection rate for map-based curve detection.

[16] Note that road geometry is delivered from AHR to applications in an incremental-AH manner as described in section 2.

[17] The road curve (curvature) design typically needs to consider the following parameters: desired target speed, maximum superelevation expected and the maximum expected road friction according to [7]. For safety and comfort, a drastic change of curvature is generally not allowed in the curve design except the connection point with another road.

estimation is rendered to be a "circle search" problem, which is realized by the Circle Center Search algorithm presented below in Figure 4.4.

However, position inaccuracy may still impede the direct use of the curvature estimates by the CCS algorithm. Therefore, the CCS algorithm is used together with the Circle Selection algorithm for data pre-filtering, i.e. curve-detection and curve-point selection. Based on the preliminary curvature estimates including curvature centers and radii of curvatures, a parametric curve fitting is employed to optimize the curvature estimates using a circle function. Figure 4.5 shows the overall processes of the proposed curve reconstruction method.

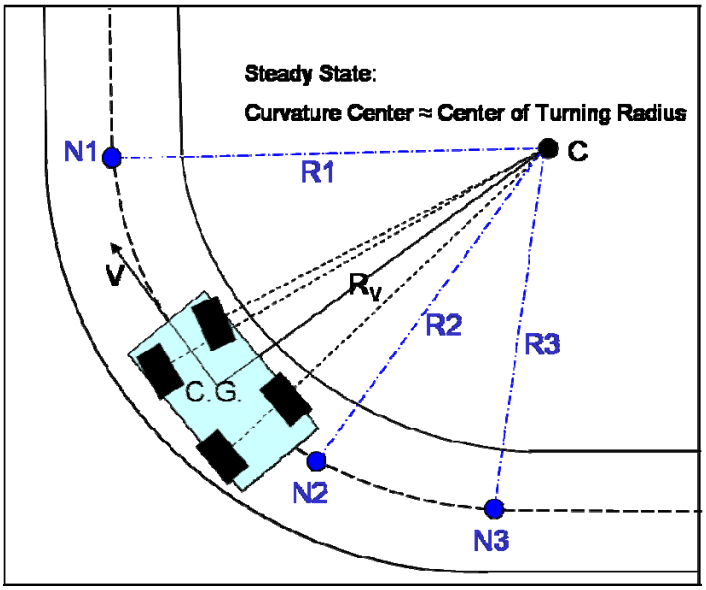


Figure 4.4 Concept of Circle Center Search algorithm

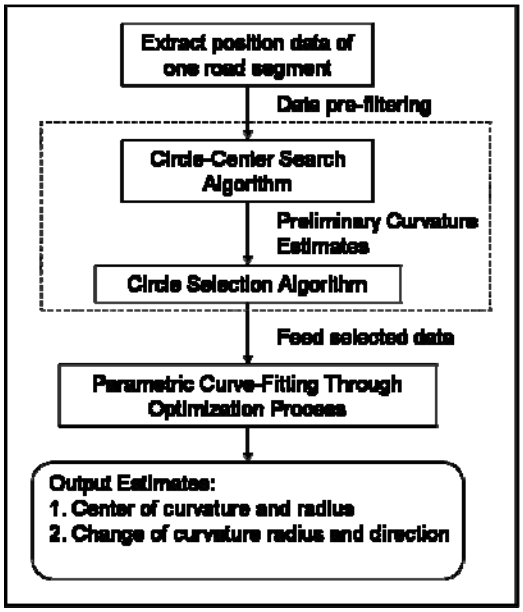


Figure 4.5 Proposed road curve reconstruction method

4.1.3.2.1 Data Pre-Filtering by Circle Center Search (CCS) and Circle Selection (CS) Algorithms

Figure 4.6 shows the satellite picture (from Google map) of the same ramp presented in section 3.1. This curve can be roughly divided into two arcs with different radii of curvatures. The radius of initial portion is around 63 meters, and the curvature radii of the latter portion are greater than 90 meters. Figure 4.7(a) shows the preliminary curvature estimates by CCS algorithm. Blue stars are the so-called "node" data in a commercial digital map, and they are located on the road center line. The first node at entrance is circled in read. Red stars are the preliminary curvature center estimates, and each one is computed based on three consecutive nodes (Appendix 1). The resultant estimates of curvature radii are shown in Figure 4.7(b).

Several interesting phenomena can be observed from Figure 4.7(a) and 4.7(b): (1) the first six curvature center estimates are distributed in a relatively small region, i.e. they appear to converge to the true curvature center, (2) the last six curvature center estimates appear to diverge from the average position of the first six ones, (3) the first six curvature radius estimates vary around the average value of 65 meters within the 10 meter bound, (4) the seventh to ninth radius estimates appear to have a different average value of 95 meters, and (5) the radius estimates jump from 96 meters to 136 meters in the last portion. (Note that the last three radii are identical due to using the same three nodes in CCS computation. If the nodes of the straight road in connection with the curve end are used, the curvature radius estimates will diverge.)

In this example, the whole ramp can be divided into two parts. The first curve consists of the first to the sixth nodes, and the second curve is between the sixth and the last nodes. Since the ramp end is connected to a straight road, curvature of the second curve is not stationary. However, the errors appear to be within an acceptable range for application use after further process by parametric curve fitting discussed below.

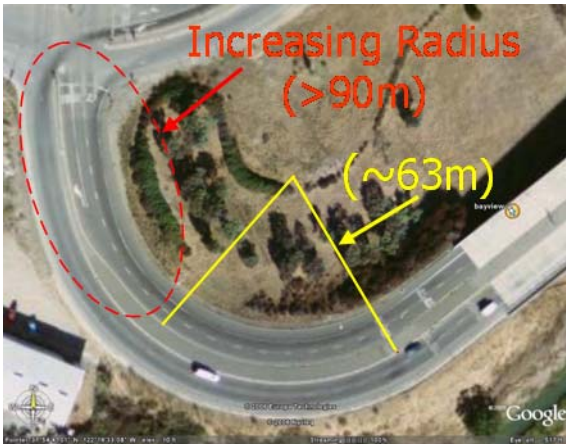
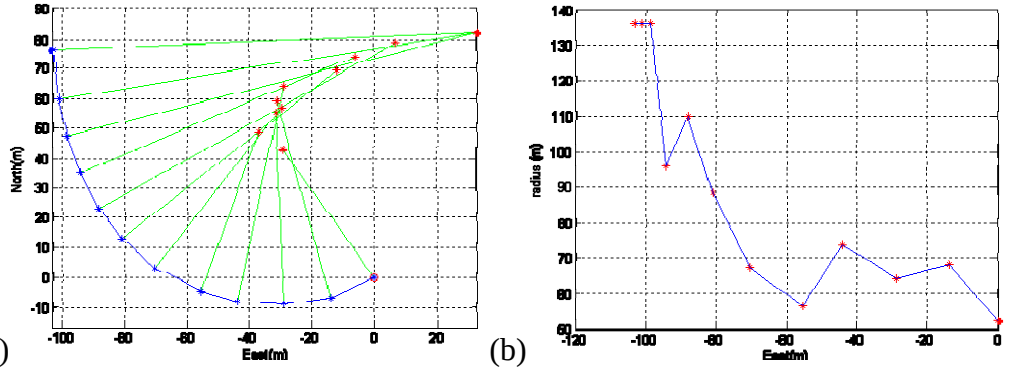


Figure 4.6 Exit ramp from northbound I-580 to Bayview Ave. (Google Map)



(a) (b)

Figure 4.7 (a) Preliminary Curvature Estimates by CCS Algorithm (b) Preliminary Estimates of Curvature Radii by CCS Algorithm

The main idea of Circle Selection algorithm is to group proper nodes based on the preliminary curvature radius estimates by CCS algorithm. By detecting relative large difference between two consecutive radius estimates, a discontinuity of road curvature and the connection node between two curves can be identified. The thresholds for detecting curvature discontinuity can be defined in terms of absolute radius difference or relative radius difference depending on (1) the desired sensitivity determined by the application need and (2) the map quality in terms of "relative accuracy"[18] as well as the number of nodes on a curve. In general, thresholds defined in terms of "relative difference" can better fit high sensitivity demanding application using high quality map[19].

In summary, the CCS and CS algorithms function as a nonlinear filter, which can be used to detect a road curve and distinguish road curvature differences. However, for more accurate curvature estimates, further parametric curve-fitting is required as described below.

4.1.3.2.2 Parametric Curve-Fitting

Based on the previous data pre-filtering results, the selected node positions (x_i,y_i), i = 1~n, of each portion of curve are used in the curve-fitting. The objective function *J* for the curve-fitting is defined in Equation (1):

$$J = \sqrt{(x-x_0)^2 + (y-y_0)^2 - R^2} \qquad (1)$$

By minimizing this objective function using (x_i,y_i), the position of curvature center (x_0,y_0) and radius of curvature *R* can be estimated.

Figure 4.8(a) shows the optimized curvature estimates for the same ramp discussed in Section 4.1.3.2.1. It shows clearly that this ramp can be divided into two curves as suggested earlier in Figure 4.6. Although the last node, which is a connection point with a straight road, is taken for

[18] Typically, the higher relative accuracy of the map can guarantee better curve shape fitting to the actual road shape.

[19] Several mock ramps and curves were created at Richmond Field Station in this study. The ramps have relatively smaller curve radii in the range of 10~35 meters compared with real ramps. In addition, an experimental digital map (lane-level) was created for Richmond Field Station. The simulation results show that thresholds defined in terms of relative difference are more appropriate for the created smaller ramps.

the second curve-fitting, it does not noticeably affect the fitting result. The enhanced curve radius estimates are quite accurate as shown in Figure 4.8(b).

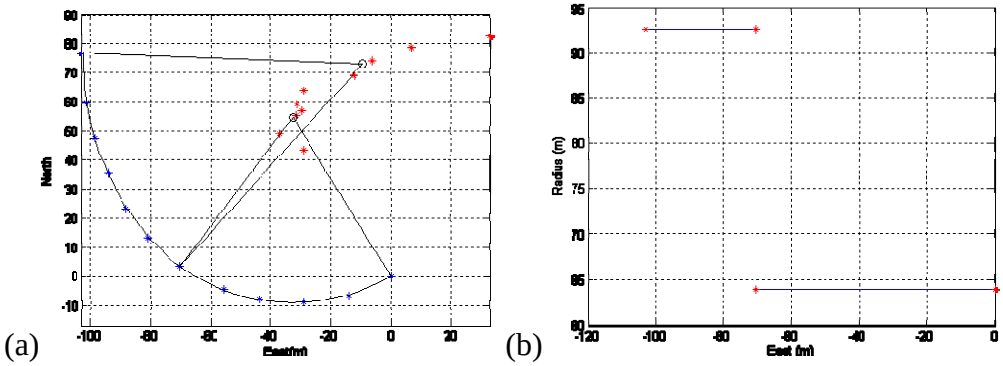
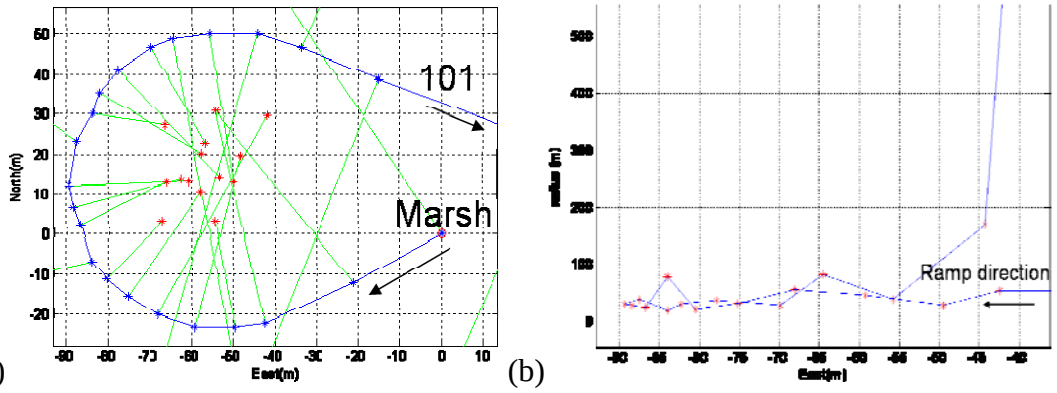


Figure 4.8 (a) Optimized estimates of road curvatures through parametric curve-fitting (b) Optimized estimates of curvature radii

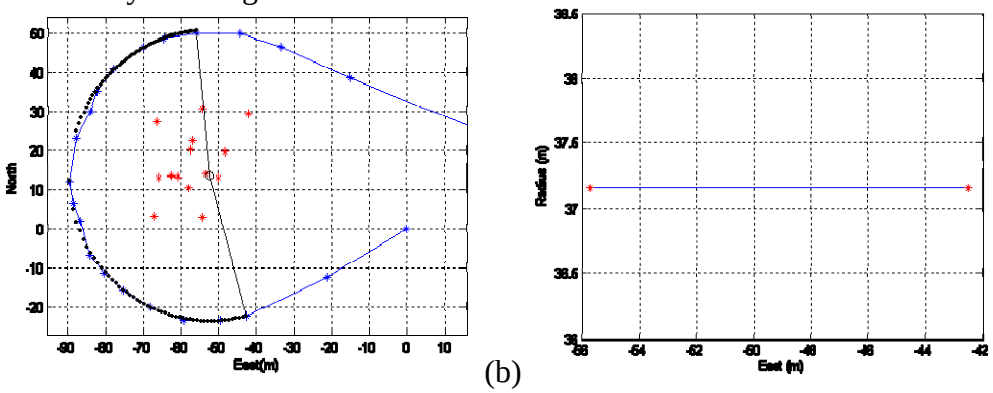
Another example is carried out for an on-ramp from Marsh Road to southbound US 101 as shown in Figure 4.9. This example shows a more complete ramp that can be divided into three parts. The first and the last portions are straight roads, and the intermediate part is one curve with radius of curvature of about 37~38 meters. Figure 4.10(a) and Figure 4.10(b) show the preliminary curvature estimates by CCS algorithm using the map data from the existing commercial digital map. The curvature center estimates of the intermediate portion are distributed within the ramp except two nodes due to the undulation of the curve shape. The result also reflects that the map quality is an important factor for the threshold design of the CS algorithm. Figure 4.10(b) shows that the curvature radius estimates appear to converge in the direction from ramp entrance to the intermediate portion and diverge in the direction toward the ramp end. The optimized curvature estimates are shown in Figure 4.11(a) and Figure 4.11(b).

Figure 4.9 On ramp from Marsh Rd. to southbound US 101 (Google map)



(a) (b)

Figure 4.10 (a) Preliminary curvature estimates by CCS algorithm (b) Preliminary estimates of curvature radii by CCS algorithm



(a) (b)

Figure 4.11 (a) Optimized road curvature estimates through curve-fitting (b) Optimized curvature radius estimates

4.1.4 Curve Overspeed Warning Experiment

This section presents a preliminary result of a prototype curve overspeed warning application using an experimental ADAS system at the California PATH Program. Currently, the on-board experimental system integrates GPS measurements, available vehicular sensors (e.g., odometer, gyroscope, and accelerometer), and digital map for safety applications in an autonomous mode. Incorporating the vehicle-infrastructure communication into this system platform for further performance enhancement is one of the on-going tasks in the project and left for future discussion.

The proposed dynamic curvature estimation algorithm was implemented in the experimental system and tested in real-time at the Richmond Field Station test tracks. Figure 4.12 shows one test result on a curve in one of the test track where GPS satellites are not available due to the surrounding tall trees. The safe speed for negotiating this curve is 9.3 m/s as computed by the COWS algorithm in real-time. The experiment showed that the COWS system did provide appropriate warnings when it detected the vehicle speed was unsafe for an upcoming curve based on both the positioning system and the map information. Figure 4.13 shows the corresponding vehicle speeds, number of satellites used in position computation, and the HDOP of GPS. When approaching this curve, GPS speed was constant due to satellite signal outage. The wheel speed

was higher than the safety speed between 163~ 165.3 seconds, and the speed warning was activated. After the vehicle was slowed down, the warning was automatically turned off by the system.

One can also see that the satellite signal outages did adversely affect the EKF-based GPS/INS position accuracy. A position enhancement method through sensor fusion of digital map was developed to address this issue; however the development is not yet finished and therefore will be left for the future report. Figure 4.12 does show that the enhanced EKF-based positions have the potential to achieve the desired "lane-level" positioning requirement.

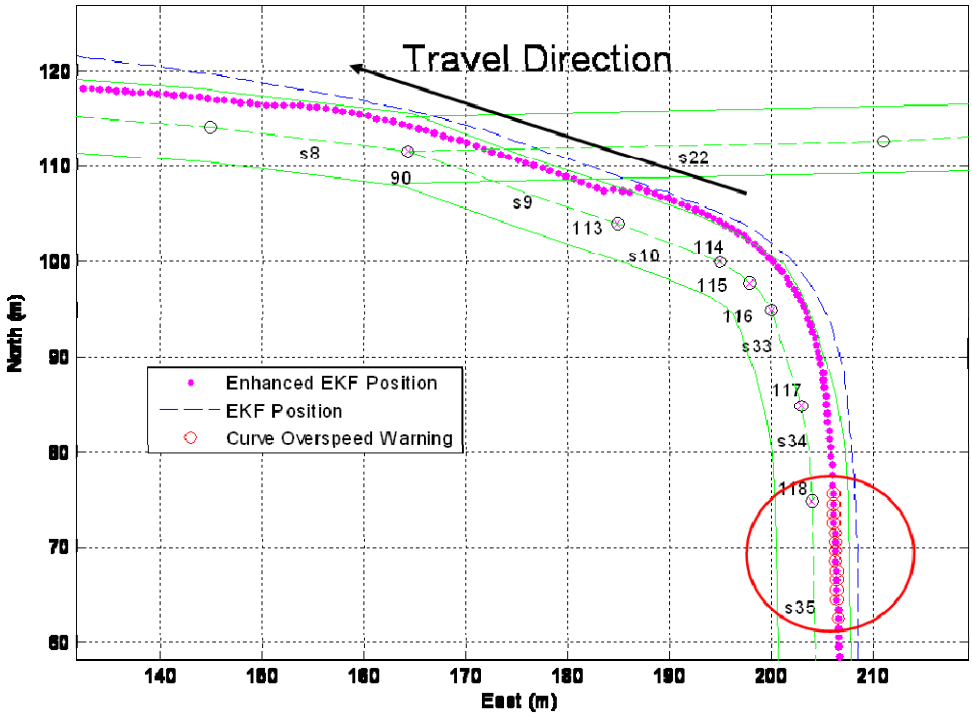


Figure 4.12 Example of COWS Field Test at the RFS Test Track

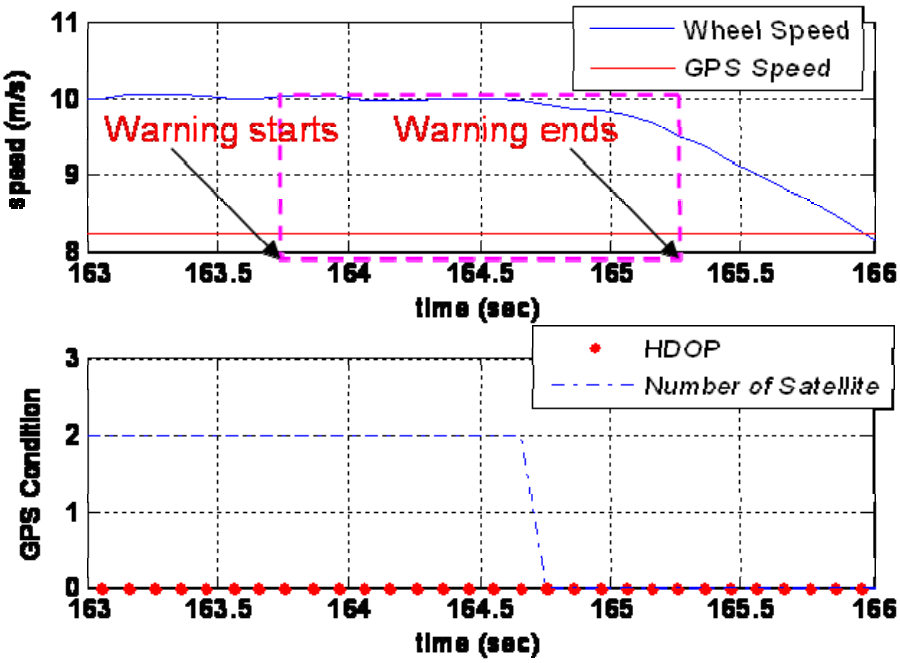


Figure 4.13 Vehicle speed and GPS condition when approaching the curve (when the GPS is in outage, the resultant HDOP is set to zero by the GPS receiver)

4.1.5 Conclusion and Future Work

This report describes the general framework of a Curve Overspeed Warning System and presents a dynamic road curve reconstruction method using the existing map database information for this specific application. The curve attributes computed by this approach appear to be better than those from the conventional spline approach.

This new approach following a typical stochastic estimation approach by combining a data pre-filtering and parametric curve-fitting function under the assumption that the characteristics of road curve can be modeled in terms of various connecting "circles.". Under this assumption, any given road curve consists of a number of connecting arcs where each arc corresponds to a specific circle. If each of these circles can be identified, the correct curvatures along this curve can therefore be derived. In order to deal with the positioning inaccuracy in the map data, the parametric curve fitting is conducted based on this circle model. Data pre-filtering (that detects curvature changes as well) selects specific "nodes" of each identified circle, which is conducted by the proposed Circle Center Search and Circle Selection algorithms, while curve-fitting is utilized to remove the variance in the curvature estimates.

Preliminary experimental results demonstrated that this prototype COWS achieved the basic performance requirements for a purely map-based COWS without involving vehicle-infrastructure communication. To achieve the optimal system performance, the enhanced map data and dynamic road information will be incorporated into the COWS computation through VII.

The future work to accomplish the VII COWS is summarized below.

1. Develop a robust position enhancement module based on current scheme, so that the GPS-based positioning unit can still achieve "lane-identification" accuracy as GPS differential correction is not available.
2. Develop the on-board wireless communication module.
3. Develop the DSRC map information message format.
4. Incorporate the enhanced map data and dynamic road information transmitted from the RSE into the COWS algorithms.
5. Develop the required software and hardware for vehicle-infrastructure communication.
6. Develop the vehicle-to-infrastructure information feedback mechanism for event, hazard or accident detection.

4.2 TRAFFIC PROBE DATA PROCESSING FOR FULL-SCALE VII DEPLOYMENT

4.2.1 Introduction

Transportation planning and operations have been severely constrained by a paucity of data ever since they have been studied seriously. Researchers, planners, system designers and system operators have always complained that they do not have data of sufficient quantity or quality to fully understand how well the transportation system is performing. This is because very detailed transportation data collection has been too complicated, costly and intrusive on people's privacy until now.

The advent of VII offers the opportunity to revolutionize transportation data collection by eventually enabling virtually all road vehicles to serve as traffic data probes, collecting information about their motions and transmitting it to roadside receivers. This has profound implications for the future of transportation planning and operations by providing most, if not all, of the data they have dreamt of having, and probably providing much more data than they have ever had to manage before. The primary challenge could quickly shift from determining how to extract as much information as possible from limited available sources of data to determining how to efficiently identify and extract the most relevant and important information from an avalanche of largely redundant data.

Existing approaches to transportation planning and operations have of necessity been designed around the limitations of the available data. When those data limitations vanish, it should become possible to develop better approaches to delivering both today's transportation services and completely new services that would not have been feasible before. Considerable research will be needed on these issues in support of the VII business case, based on the benefits that could be gained from these yet-to-be-determined uses of the new VII data.

The main focus of attention in this section is on use of VII-equipped vehicles to collect traffic probe data that could be used in support of transportation planning or operations. These data are expected to include measurements of vehicle location and speed, as well as vehicle status indications such as braking and steering actions, suspension motion, ABS and traction control activation, and windshield wiper, turn signal and headlight usage. With suitable processing, these data can be turned into real-time and archival indications of traffic speed and volume, travel times, incidents, and weather and road surface conditions.

Following background information about probe data collection and VII in general, the simulation environment that serves as the basis for evaluation of alternatives is described. This is followed by a review of the baseline probe data sampling protocols that are currently being assumed for the VII program, and their limitations. The rest of the section focuses on issues of probe data aggregation, with case studies for three types of probe data of progressively increasing complexity. The section concludes with a summary of the most important findings with respect to using probe data with VII and recommendations for further work on the topic.

4.2.2 Background

There is already a substantial literature about use of vehicles as traffic data probes. Because of the cost and institutional complexity of equipping vehicles as probes, prior work has tended to focus on use of special-purpose, heavily utilized fleet vehicles such as transit buses [Bertini and Tantiyanugulchai, 2003; Hall and Vyas, 2000] taxis [Sun, et al 2005], or other public service vehicles [Odawara, 2004], or vehicles from one manufacturer seeking to attain a competitive advantage with information available only to their customers [Hauschild, 2005; Tsuge and Arai, 2004]. In these cases, the probe vehicles represented at best only a small fraction of the total vehicle population and the special-purpose fleet vehicles were driven differently from typical passenger cars (many more stops, for example). These considerations were dominant in the design of the probe data processing algorithms, which renders these algorithms less useful for the VII scenario of essentially ubiquitous probe vehicles.

Some papers have focused on use of traffic data probes to detect freeway incidents, while others have focused on arterial incidents and some have addressed both [Srinivasan and Jovanis, 1996], but in all these cases the frequency of probe samples has been low (up to a few vehicles per minute), so the minimum sample aggregation intervals have had to be long. Freeway incident detection times were estimated, based on a traffic simulation, to range from 12 minutes at a 5% market penetration of probe vehicles to 3-4 minutes at 50% market penetration, with misclassification percentages ranging from 15% to 18% [Cheu and Lee, 2002]. A real-world experiment using electronic toll collection tags to identify freeway incidents showed performance comparable to that of conventional infrastructure-based traffic detectors [Mouskos, et al, 1999]. A similar experiment showed how to use toll tags to estimate freeway travel times, with minimum aggregation intervals of five minutes because of the limited density of equipped vehicles [Park, 2002]. This approach should be scalable for higher density applications with shorter aggregation intervals.

The more challenging application of probe vehicles to identifying arterial incidents and travel times was addressed in the ADVANCE project [Thomas and Hafeez, 1998; Saricks, et al, 1997; Sethi, et al, 1995], using a combination of traffic simulation and field experiments with up to 15 vehicles driving pre-specified routes and seeking out incidents. This research proposed identifying incidents based on differences in link travel times [Thomas and Hafeez, 1998], and then showed that travel time estimates could be obtained using as few as 3 vehicles within a 5 minute sampling period [Saricks, et al, 1997]. The ratio of link travel time to the “normal” travel time for that link was found to be most effective at discriminating incidents, and the smaller the number of probe samples available, the higher would be the threshold ratio needed to declare an incident [Sethi, et al, 1995].

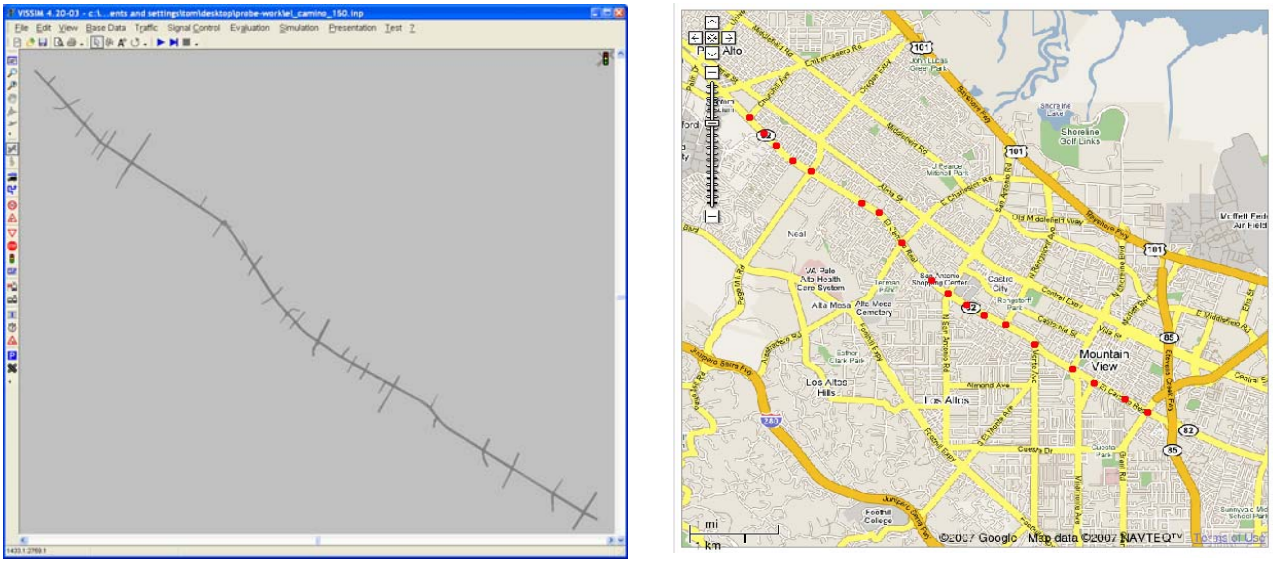
The VII program has been developing through a collaboration among the U.S. DOT, AASHTO and several state DOTs, and a consortium of automotive OEMs. At the time of writing this paper, little information has been released in public about the design and architecture of the VII information system, beyond a preliminary system architecture [FHWA, 2007]. The probe data sampling and encoding approach has been developing through a standardization process of the Society of Automotive Engineers, under SAE J2735 [SAE, 2007]. Some of the reasoning behind the choices embedded in J2735 has been described in a technical report [ITS-JPO, 2006]. More

recently, Noblis has been testing the J2735 probe protocols using both sample field test data and traffic microsimulations (Noblis, 2007).

The current research is being conducted as part of the VII California program, under the sponsorship of the California Department of Transportation, which wants to understand the implications of VII deployment in California and the contributions that a mature VII system could make toward solving California's transportation problems. The original reasoning behind exploring efficient probe data aggregation approaches was explained in [Shladover, 2006], and the first stages of research on simulating traffic data probe alternatives were presented in [Shladover and Khun, 2007]. This report builds on the work in [Shladover and Khun, 2007] and extends it in breadth and depth.

4.2.3 Simulation Model

The performance of the traffic probe data collection system is depends heavily on the number of vehicles communicating probe data to the roadside. Since it is neither technically nor economically feasible to equip a large number of probe vehicles at the current stage of the VII program, the assessment of probe data collection approaches must be done using simulations of traffic. The traffic microsimulation approach provides the additional advantage that the simulation outputs represent an authoritative "truth model" reference of the complete vehicle motion trajectories that the probe data is meant to represent. This means that the comparisons of the probe sample outputs with the simulation outputs give a useful indication of how accurately the probe sampling and data processing approach is representing "reality".



(a) VISSIM Network Representation

(b) RSE Locations (red dots) Simulated Along El Camino Real

FIGURE 4.15 Case Study Network for Probe Vehicle Simulations.

The traffic simulation (using a VISSIM model) represents about 10 km of El Camino Real (State Route 82) in the cities of Palo Alto and Mountain View, CA, with 25 signalized intersections modeled, as shown in Figure 4.15. This corridor includes part of the VII California testbed, which is already equipped with several roadside communication hotspots (RoadSide Equipment, RSE) for testing probe data uploading. The traffic can be simulated at several levels of congestion, and for each simulation the vehicle trajectories are logged as outputs. The model is run and data logged with a time step of 0.1 second to retain fidelity of motion, and then the data are down-sampled at 1 second intervals for further analysis. The simulation was set to model a high level of traffic congestion for the analyses in this report in order to maximize the data flow volumes.

4.2.4. Baseline Probe Data Sampling Approach

The baseline assumptions for probe vehicle data sampling follow directly from the SAE J2735 draft [15] and are based on the concept of "snapshots", which are the periodic samples of the state of the vehicle. This represents the approach being followed in the national VII program

The key rules associated with the recording and transmission of these snapshots are:

- Snapshots are generated periodically while the vehicle is driving, and for special events, stops and starts
  - At 6 s intervals for speeds < 20 mph
  - At 20 s intervals for speeds > 60 mph
  - For special events – vehicle status changes (TCS, ABS activation)
  - At stops (no movement for 5 s, no other stops within 15 s)
  - At starts (speed exceeds 10 mph)
- Snapshots are stored in a buffer that can hold 32 snapshots (default minimum value)
- Priority for snapshot transmission is: events, then stops and starts, then periodic samples
- FIFO priority for snapshot buffer, then snapshots are overwritten when buffer capacity is exceeded
- Association of snapshots from the same vehicle is restricted (Greater of 120 s or 1 km limit en route, no snapshots in first 500 m of trip) to protect privacy
- A probe message upload from vehicle to local roadside equipment (RSE) may contain up to 4 snapshots
- Transmission to each RSE occurs only once, when the vehicle first comes within 400 ft (122 m) range of the RSE (no subsequent transmission until reaching the next RSE). This transmission may include multiple messages to transmit all snapshots stored in the buffer.

These rules have been designed based on a series of compromises, to economize on data traffic under general traffic conditions, not necessarily the most congested conditions, and to support applications that are not severely time sensitive.

For purposes of simulation and evaluation, a post-processing program represents these probe data collection strategies, operating on the simulated vehicle trajectories in the same way that the traffic probe data would be sampled in the field. The post-processor, known as the Trajectory Conversion Application (TCA), was originally developed by Noblis [Noblis, 2007] for use in the national VII program, and was generously made available for use in this project. The TCA models both probe data generation in the vehicle and transmission to roadside equipment (RSE).

The VISSIM simulation and TCA post-processor have been used to test the performance of these baseline rules in the test corridor, and the initial results of that testing were presented in [Shladover and Khun, 2007]. Some of those results are described here to provide background information needed to establish the foundation for the new results reported in Sections 4.2.5 - 4.2.7. Figure 4.16 illustrates the snapshot samples generated at a busy intersection near the middle of the corridor, superimposed on continuous trajectory plots representing the complete VISSIM simulation outputs.

The simulation data of Figure 4.16 show that the basic snapshot generation rules appear to be providing reasonably complete sampling of the vehicle trajectories while they are moving and while stopped at the traffic signal. The vehicle trajectories can clearly be re-created based on the rate of sample generation incorporated here, even with the complications created by the traffic signal. However, the dark square symbols represent snapshots that were not transmitted from the vehicles to the roadside because the vehicles turned and left the network before encountering the next RSE or the snapshot buffer overflowed.

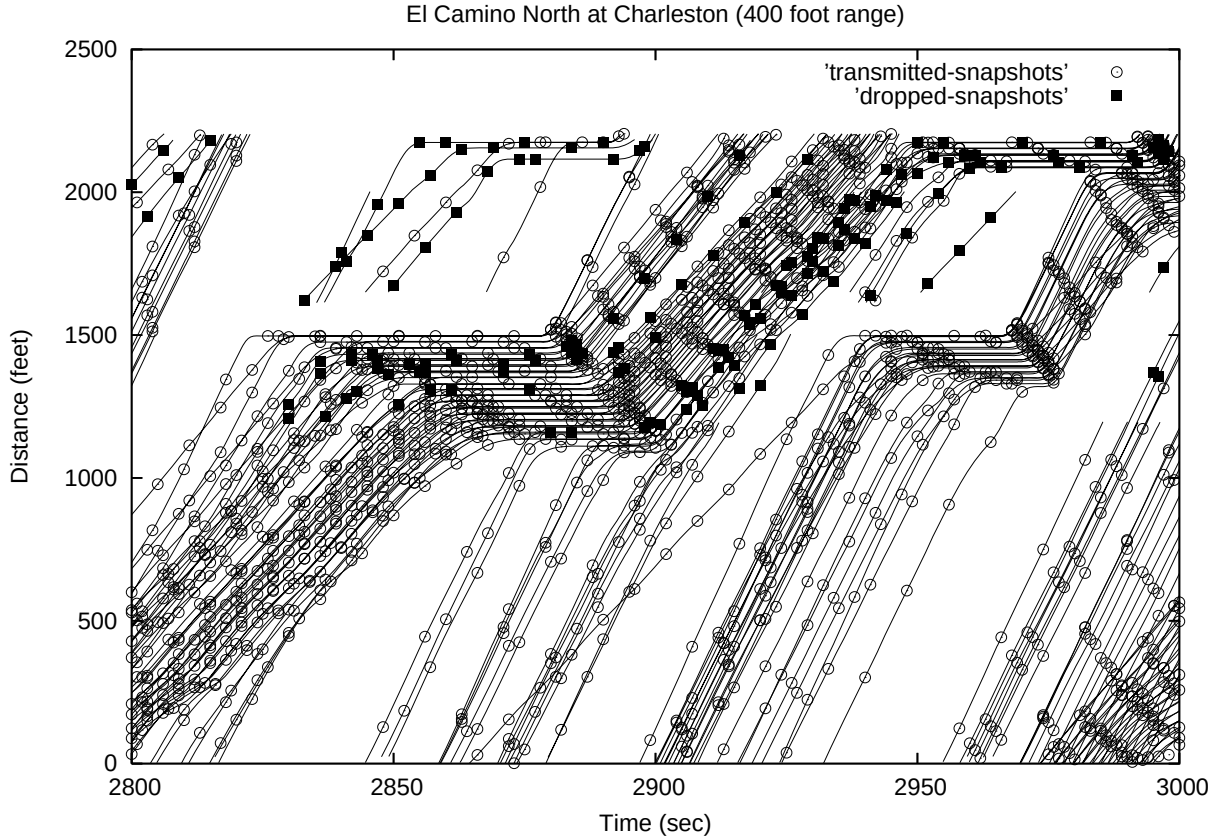


FIGURE 4.16 Snapshot Sampling at a Busy Intersection.

A more serious limitation of the baseline probe data management strategy is the latency introduced by prohibiting the vehicle from transmitting probe messages more than once per RSE. As shown in the simulation data of Figure 4.17, this means that after a vehicle has done its initial probe data upload (nominally 400 ft. from the RSE), it is not permitted to upload any more probe data until it encounters the next RSE, which could be one or more intersections away. The travel

time for the vehicle to reach that next RSE introduces a significant latency, illustrated in Figure 4.17 for a small sub-sample of the vehicle trajectories shown in Figure 4.16. The "+" signs indicate when the probe snapshots are actually uploaded, when the vehicle approaches within 400 ft. of the next RSE. The intersection located at about 2200 ft. in Figure 4.17 is not equipped with an RSE, so the snapshots accumulated after the vehicles pass the 1200 ft. location cannot be uploaded until the vehicles are within 400 ft. of the following intersection (at times of about 3000 and 3030 s on the plot).

The statistics of this data latency are discouraging for real-time applications, with a 50%ile delay of about one minute, and some samples delayed as much as two minutes. More than 20% of the snapshots are never transmitted, primarily because their vehicles turn onto side streets and leave the network without encountering another RSE. In a fully deployed VII network, the intersections on these side streets would also be equipped with RSEs so these data would not all be lost.

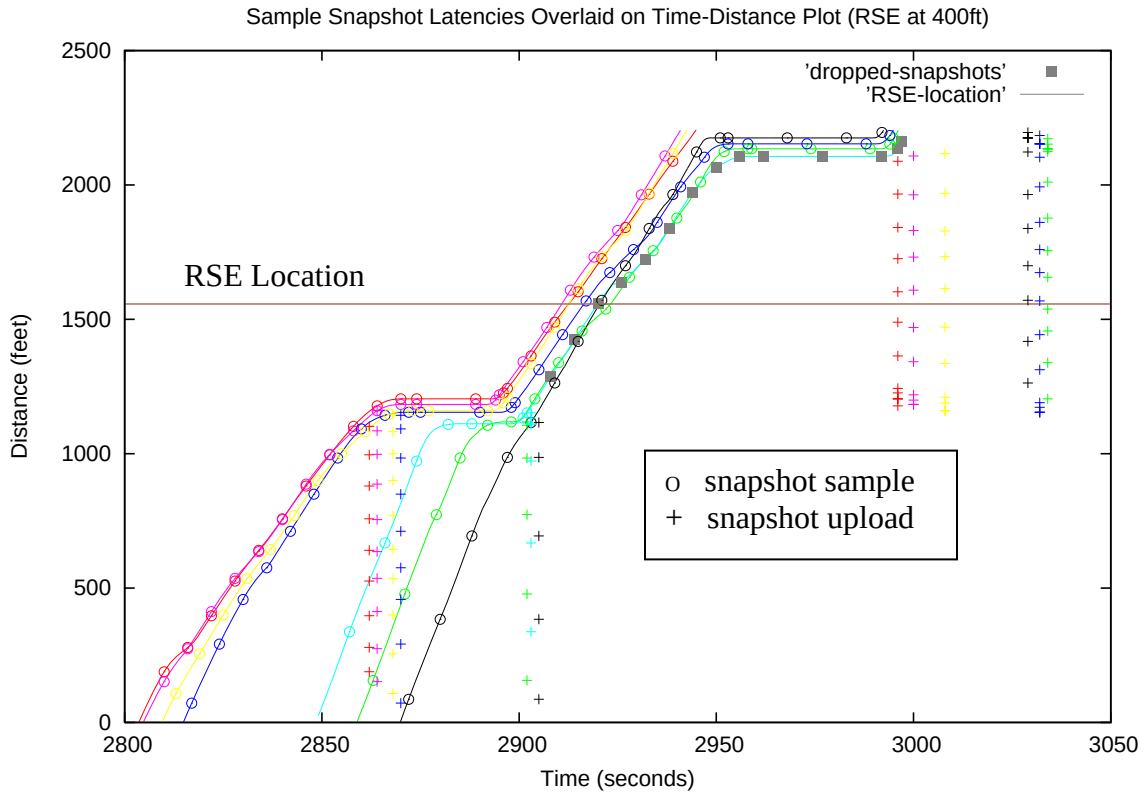


FIGURE 4.17 Examples of Delays between Probe Snapshot Sampling and Message Uploading to RSE.

Although a snapshot latency of one or two minutes would not raise serious concerns for applications that involve off-line post processing of probe data for planning purposes, it could limit the applicability of the probe data for time-critical real-time applications. Since high-fidelity traffic probe data have not generally been available in the past, there is little if any literature that specifies the latencies that would be acceptable. However, the range of acceptable latency can be associated with the time scales on which the underlying processes change:

- Probe detection of weather conditions (precipitation, road icing): These vary slowly in both space and time, so that detection of changes within one or two minutes should be adequate.
- Detection of incident or traffic disturbance (emergency braking or swerving, activation of ABS or traction control): Although these processes happen within a fraction of a second, the responses by TMC operators (dispatching emergency response or maintenance teams, notifying approaching drivers of a problem) are likely to take a minute or two. This means that latencies of tens of seconds will not make a serious difference in the overall response time and should be acceptable.
- Detection of network or link travel time and speed information: Although these characteristics may change in a few seconds as a result of an incident, the processes that use this information (regional traffic management and traveler information systems) influence decisions that take place over time scales of minutes. In this case, latencies of tens of seconds should be acceptable because they will not noticeably delay these processes.
- Probe detection of microscopic traffic conditions for dynamic adaptive traffic signal control: This appears to be the most demanding application because it requires knowledge of vehicle locations and speeds with high spatial and temporal resolution in order to enable traffic signals to respond to individual vehicle motions. Latency of no more than 2 seconds is likely to be needed so that signals can minimize un-used green time and vehicle stoppages.

The baseline probe data management rules should be modified to enable the probe data to be used for high-performance real time traffic control applications, for example. Modifications to the message upload protocol were tested in simulation to identify their influence on snapshot latency. These results are very sensitive to whether the snapshot is generated just before the vehicle reaches the range when it is allowed to upload to the RSE or whether it is generated just beyond that range. Figure 4.18 shows the cumulative distributions of snapshot latencies for these two extremes, with the red curve showing relatively low and consistent latencies for snapshots of a modeled event that occurs just before the vehicle reaches the RSE range, and the green curve showing much longer and inconsistent latencies for snapshots of a modeled event that occurs right after the vehicle reaches the RSE. In the latter case, the snapshots cannot be uploaded until the vehicle reaches the next RSE, which may be one or more intersections away, and for the simulated case, nearly 20% of the snapshots are lost because the vehicles turn off the route before reaching the next RSE.

Figure 4.19 shows how the snapshot latency throughout the simulated network changes as the probe message management broadcast rules vary, with several sample cases shown (broadcasting when buffer accumulates 2 or 4 snapshots, or after 10 or 30 seconds, and the default case). Since the maximum probe message size is 4 snapshots, the 4 snapshot triggering rule would not reduce the efficiency of message transmittal at all. However, because there is a significant message transmittal overhead, there will be an efficiency penalty incurred to achieve the lower latency of the 2-snapshot buffer upload rule. Locations that are applying dynamic adaptive traffic signal control are likely to require the most frequent probe message broadcasts (2-snapshot buffer accumulation before broadcasting again, after arriving within range of RSE). In locations where

the probe data are used for incident detection and link travel time and speed information, where latencies of a few tens of seconds should be acceptable, it should be possible to get by with probe broadcasts at 30 second intervals or waiting for 4 snapshots to accumulate in the buffer.

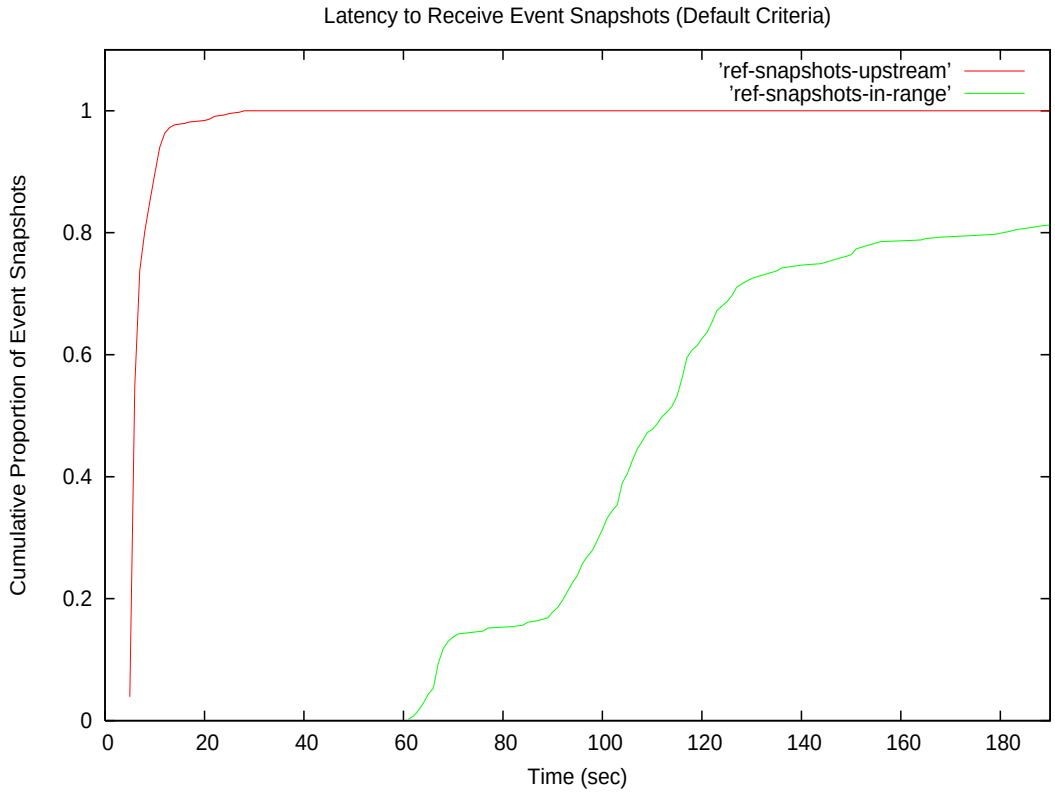


FIGURE 4.18 Cumulative Distribution of Snapshot Latencies for Events Just Within and Just Without the RSE Range

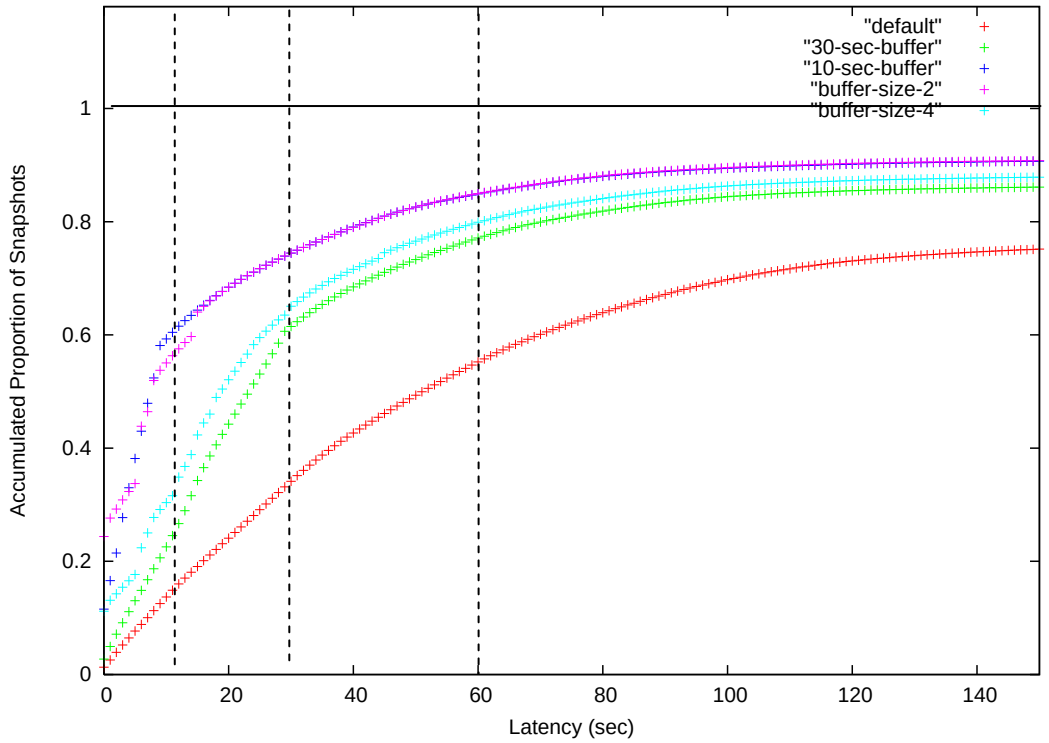


FIGURE 4.19 Snapshot Latencies for Alternative Broadcast Protocols.

The analyses reported here are based on 100% market penetration of vehicles equipped as traffic data probes, which is the best case condition. During the period of VII deployment the market penetration will be lower than this, which means that more intensive sampling of vehicle snapshot data is likely to be needed to achieve comparable performance. This should not pose technical problems for the capacity of the system because of the smaller number of communicating vehicles, but the system design and standards need to be adjusted to ensure that this is permitted, so that the benefits of traffic probe data can be gained as early as possible.

4.2.5 Probe Data Aggregation Issues

The architecture for the national VII program has assumed that **all raw probe vehicle data** received at each RSE would be forwarded on a backhaul communications link to a "Service Delivery Node" (SDN) and from there to an "Enterprise Network Operations Center" (ENOC), where it could be accessed by potential users ("subscribers"). Concerns about the potential inefficiencies of this approach [Shladover, 2006] provided the initial motivation for the current study. These concerns are primarily:

a. Data from all the vehicles traveling on the same roadway are highly correlated with each other and therefore redundant, so it should be possible to provide information of the same value by aggregating these data locally at the RSE, thereby saving on backhaul communication capacity and cost.
b. Providing modest computing resources locally at the RSE enables distributed data processing, which should be more cost effective than requiring a central processor to manage the raw data from all vehicles traveling within a region.

c. Providing aggregate probe data for each RSE greatly simplifies the data processing burden on the end-users who will be implementing traffic management or traveler information services based on the probe data.

The backhaul data burden (without any local data aggregation) was calculated for a sample busy intersection based on the simulation using the baseline rules. Under those rules, the minimum size for a probe data snapshot would be 27 bytes and the maximum would be 151 bytes (assuming a wider range of vehicle variables to be transmitted). The average rate of snapshot uploads was 880 per minute and the maximum was 2630 per minute (on a one-minute average) at the single RSE, based on one hour of simulated traffic. At the maximum snapshot upload rate, the data rate per RSE would range from 10 to 53 kbits/s, depending on the size of the individual snapshots. If one considers a large urban region equipped with thousands of RSEs, the volume of data that a probe data information service provider would have to manage and process for that region could be in the range of 10 to 100 Mbps.

If the raw data were aggregated locally at the RSE, the backhaul data rate could be reduced drastically and privacy could be enhanced because the raw probe data could be destroyed at the RSE and would not even be uploaded to the VII network. If each RSE had to separately aggregate traffic condition data for four roads approaching its intersection and for the environs of four adjoining RSEs, this would represent eight sets of data. For a conservative analysis, consider an RSE where the probe data applications need 31 vehicle attributes (the maximum permitted in an individual probe message in SAE J2735), with each represented by a two-byte identifier and two bytes for the data, the total data content would be 124 bytes per data set. With eight data sets, and assuming that the aggregated values could each be represented in 2 bytes, the total backhaul data rate would be 124 x 8 data sets = 992 bytes/s or 7936 bps (8 kbits/s) if every data element were to be updated once per second. This data traffic could be reduced significantly by encoding the data elements more efficiently than currently specified in SAE J2735 (binary data elements such as ABS or windshield wiper activation status only need one bit rather than the four bytes (32 bits) currently assigned to each) or by uploading them less frequently than once per second.

The probe data sampling and aggregation issues differ, depending on the type of data that are being created and how they are intended to be used. These are summarized in Table 4.1 below showing the range of characteristics that aggregation methods must be able to accommodate:

Table 4.1 Probe Data Aggregation Approaches, Based on Data Types and Applications

Data Type and Scope	**Application**	**Data Characteristics and Issues**
Binary, Slow changing, Wide area	**Estimating weather conditions from windshield wiper status**	**Simple aggregation of percentage of wipers on for all vehicles approaching RSE**
Binary, Fast changing,	**Emergency braking or activation of ABS or TCS to**	**Need to identify specific location of event, with enough samples to**

Localized	**indicate slippery road surface**	**provide high confidence**
Continuous, Localized	**Vehicles slowing down or swerving to avoid a dropped load**	**Need to identify specific location of event, based on a model to infer the traffic disturbance from vehicle speed and direction changes**
Continuous, Corridor	**Estimating network traffic speed information for traveler information and regional traffic condition monitoring**	**Need aggregation of measurements along corridor, with appropriate weighting to accurately estimate conditions**
Continuous, Localized, Fast changing	**Microscopic traffic speed and queuing information to support real-time adaptive signal control**	**Need to select increments of time and location to define "bins" for aggregating snapshot data**
Continuous, Corridor	**Travel time estimates by link, for trip planning and traveler information**	**Need to associate successive snapshots before aggregating, within restrictions imposed for privacy**

4.2.6. Probe Data Aggregation for Binary Data (Windshield Wiper Status Example)

The simplest probe data applications involve use of binary data elements transmitted from the vehicles to identify conditions that change gradually in space and time, such as use of the windshield wiper status to indicate whether it is raining. The wiper status is binary (or may extend up to two bits if we include intermittent, low and high speed settings as well as off). Weather conditions change relatively slowly and the geographical precision of weather condition information does not need to go below the RSE location. This means that the probe data processing can be extremely simple: count the number of vehicles broadcasting within range of the RSE with wipers on and off, and divide the number with wipers on by the total number of broadcasting vehicles to obtain the estimated percentage using their wipers. Because the weather changes are gradual, this information can be updated at intervals of minutes rather than seconds.

The VISSIM simulation and TCA post-processor were used to represent an artificial condition of rain falling on part of the simulated network, but only during 25 minutes of simulated traffic (from t = 3000 to 4500 seconds). Figure 4.20 shows how well the wiper status summary statistic was determined for sampling intervals of 1 and 5 minutes. Each plot shows the percentages of probe snapshots indicating activated wipers on vehicles approaching 5 RSE locations. For the RSEs in the center of the rainy zone, the percentages are clearly very high, while the RSEs at the boundary of the rainy zone only reach about 50% wiper activations. The data for the boundary RSEs are clearly noisier for the shorter sampling interval (fewer samples to average), but the detection of the weather condition change is clearly faster. This is a fundamental trade-off in designing probe sampling strategies.

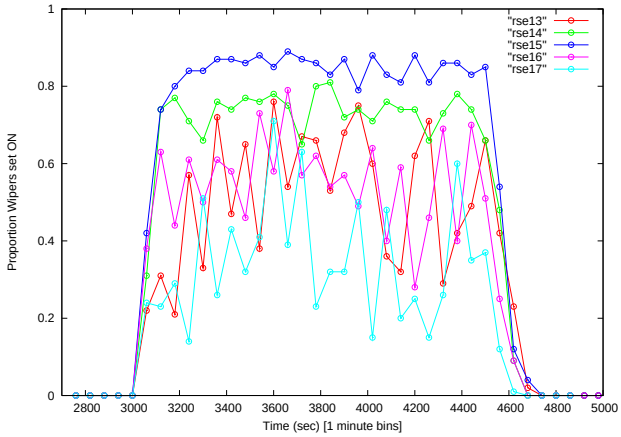
The results in Figure 4.20 show how wiper status at five RSE locations can be used to detect local rainy conditions at 100% and 10% market penetration of probe vehicles. When smaller percentages of the vehicle fleet are equipped, the detection is slower and less reliable. At the 10% market penetration, the 1-minute sampling is too sparse, producing some sample intervals indicating no rain even when there is rain present, but the 5-minute sampling appears to be adequate, provided that the threshold value of percentage of vehicles with wipers active is well chosen.

These analyses create five bins (one per RSE) at each time step and present one averaged value for each data point. At 100% market penetration, there are on average 610 samples supporting each 1 minute value, and more than 3080 for each 5 minute value. At 10% market penetration, there are on average 60 samples for each 1 minute value and 304 for each 5 minute value. Choosing the more stable 5 minute sampling interval, this indicates a data rate reduction of a factor of 300, even at the 10% market penetration level.

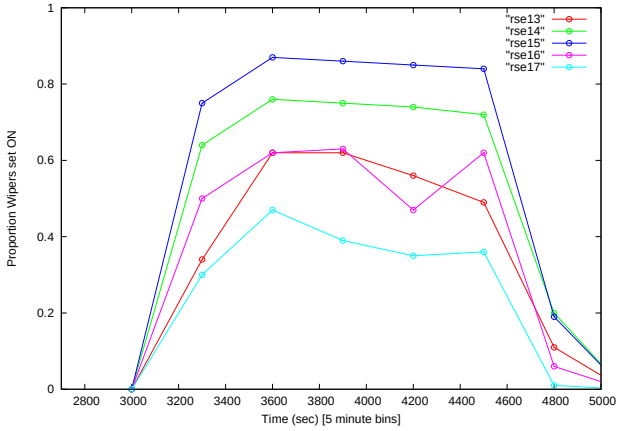
4.2.7 Probe Data Aggregation for Location-Critical Event (Dropped Load Example).

One of the most important potential applications of probe vehicle data is rapid and reliable detection of incidents that disrupt traffic flow. This is more challenging than the weather condition application because it requires more precise localization and a faster response time. The location of the incident should be identified with an accuracy of about 100 m (in the correct direction of travel!) so that incident response teams can be dispatched efficiently and other

travelers approaching the scene can be alerted about the location to avoid. In order to achieve prompt incident response and a satisfactory traveler information level of service, this information should be available within a few tens of seconds of the incident.

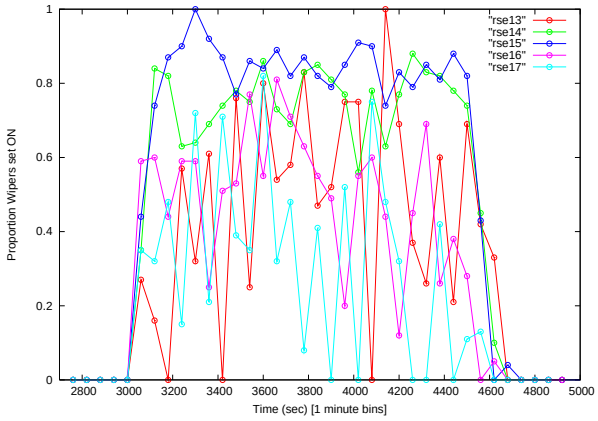


1-minute sampling interval

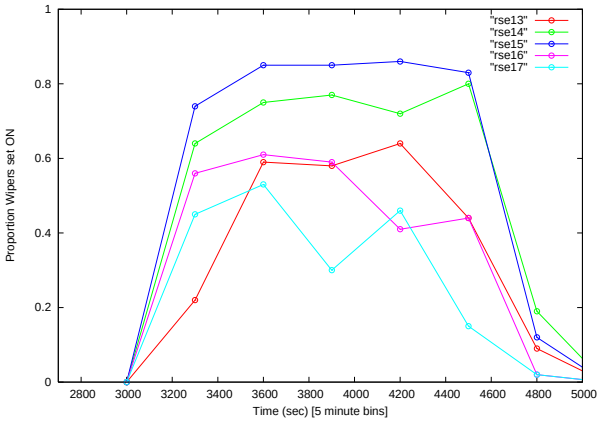


5-minute sampling interval

(a) 100% market penetration of probe vehicles



1-minute sampling interval



5-minute sampling interval

(b) 10% market penetration of probe vehicles

FIGURE 4.20 Probe Sampling of Wiper Status to Detect Rain at Five RSE Locations.

The VISSIM simulation is not designed to represent sudden traffic disturbances, so the traffic incident scenario was represented by a dropped load that could be detected by equipped vehicles driving over it and experiencing abnormal suspension deflections. This scenario was evaluated in terms of detection latency in our previous work [Shladover and Khun, 2007]. That work showed a large contrast in detection latency between disturbances located shortly before and shortly after vehicles entered the wireless range of an RSE. In the former case, the information would be broadcast to the RSE very quickly, but in the latter case the latency could be as long as one or two minutes because of the baseline probe management restriction on vehicles broadcasting after their initial encounter with the RSE.

The previous analysis of incident detection latency was extended, as shown in Figure 4.19, to consider the distribution of latencies throughout the entire corridor, accounting for incidents that could occur anywhere. The Figure 4.19 results show that adequate incident detection latencies should be achievable when probe broadcasts are permitted at 30 second intervals or when the buffer accumulates 4 snapshots.

4.2.8. Probe Data Aggregation for Traffic Speed Estimation

Most traffic management and traveler information applications need to know the speed and/or travel time associated with each section of roadway (network link). These are more complicated than the examples shown in Sections 4.2.6 and 4.2.7 because they involve continuous variables that must be sampled over the length of the roadway section, and it is important to know how the speed varies with both time and location. Travel time estimates introduce additional complications that can best be addressed through association of probe snapshots, but since that is a controversial issue for VII because of its privacy implications, primary attention here is focused on speed estimates.

Estimation of arterial traffic speeds from probe vehicles is more challenging than estimating freeway speeds because of the large disturbances introduced by traffic signal cycles. The speed-vs.-time profiles generated by the signal cycles must be captured accurately by the probe sampling process, particularly if the probe data are to be used to support improved traffic signal control strategies such as real-time adaptive signal control. The need to know traffic speed as a function of both location and time introduces the requirement to aggregate the probe data samples into "bins" defined in terms of location and time. With small bins in time, it should be possible to observe fast changes with little delay, and with small bins in space it should be possible to localize events and estimate queue lengths accurately. However, if the bins get too small there are not enough samples to support stable summary statistics. Since the real-time signal control system must be able to respond quickly to changes in local traffic conditions, the time bins were limited to 2 s and 4 s duration, while the spatial bins were varied from 20 m to 100 m at the 2 s duration.

The most demanding application for traffic speed data is real-time adaptive signal control because of the high resolution it needs in space and time in order to capture information about formation of queues and movements of platoons of vehicles. This application has been studied for 100% market penetration of equipped vehicles and a range of probe sampling strategies. Using simulation results from one link of the El Camino Real network, one hour's worth of snapshots was recorded and sorted into bins for aggregation (without concern for transmission latencies or which RSE might have received the snapshots). Even with the largest bin spatially (100 m) and 2 s sample interval, the mean number of snapshots per bin was only 2.25, and 36% of the bins contained no snapshots. Reducing the bin size to 50 m and doubling the sample interval to 4 s left the mean number of snapshots unchanged, but increased the number of empty bins to 41% of the total.

Based on simulations for a range of conditions, attention has been concentrated on roadway sections 50 m long, so that each block between successive intersections contains several such sections. Figure 4.21 shows traffic speed as a function of location (50 m increments) and time (2

s increments), in which each data point represents the average value of the snapshots within its bin, when sampled according to the SAE J2735 rules (but without regard to the latencies associated with when snapshots may be broadcast).

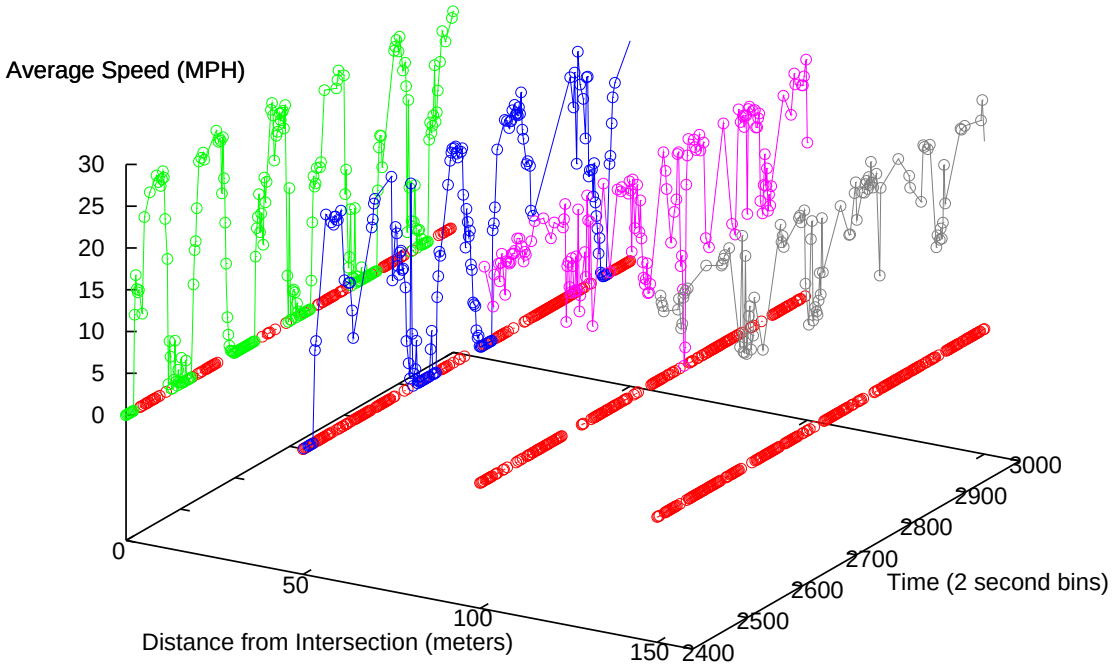


FIGURE 4.21 Average Speeds Derived from Five Minutes of Probe Snapshots**.**

The left-most set of data, for the 50 m closest to the traffic signal, clearly shows the speed profile associated with the signal cycles. These speed variations are gradually washed out and plots become noisier as we move further upstream from the signal (to the right on the plot). Even with 100% market penetration, the bins are sparsely populated, with an average of only 1.1 snapshots per bin, and 54% of the bins contain no snapshots (represented by red data points at the lower limit of the vertical axis of the plot).

Figure 4.22 shows a close-up view of the average values per bin of the probe snapshot data over two different sampling intervals (2 s and 10s) for the second 50 m roadway section (from 50 to 100 m from the signal), together with the comparable averages per bin of the baseline data representing the "truth model" of the VISSIM simulation (sampled at 1 s intervals) over a period of ten minutes. The bins without samples are again indicated by the data points at the bottom of the plots, where it is evident that the snapshot sampling (red and blue) has many more empty bins than the complete simulation output (gray). Note that there are extended periods between about 2500 and 2530 seconds and again between about 2820 and 2890 seconds when the snapshot data lack samples, seriously limiting the usability of these data for real-time signal control. These data indicate that the speed profiles are generally similar, but with a modest lag and less variability for the longer duration bins. The level of approximation of the 10-second

sampling interval is likely to be adequate for most traffic management and traveler information applications, except for real-time adaptive traffic signal control.

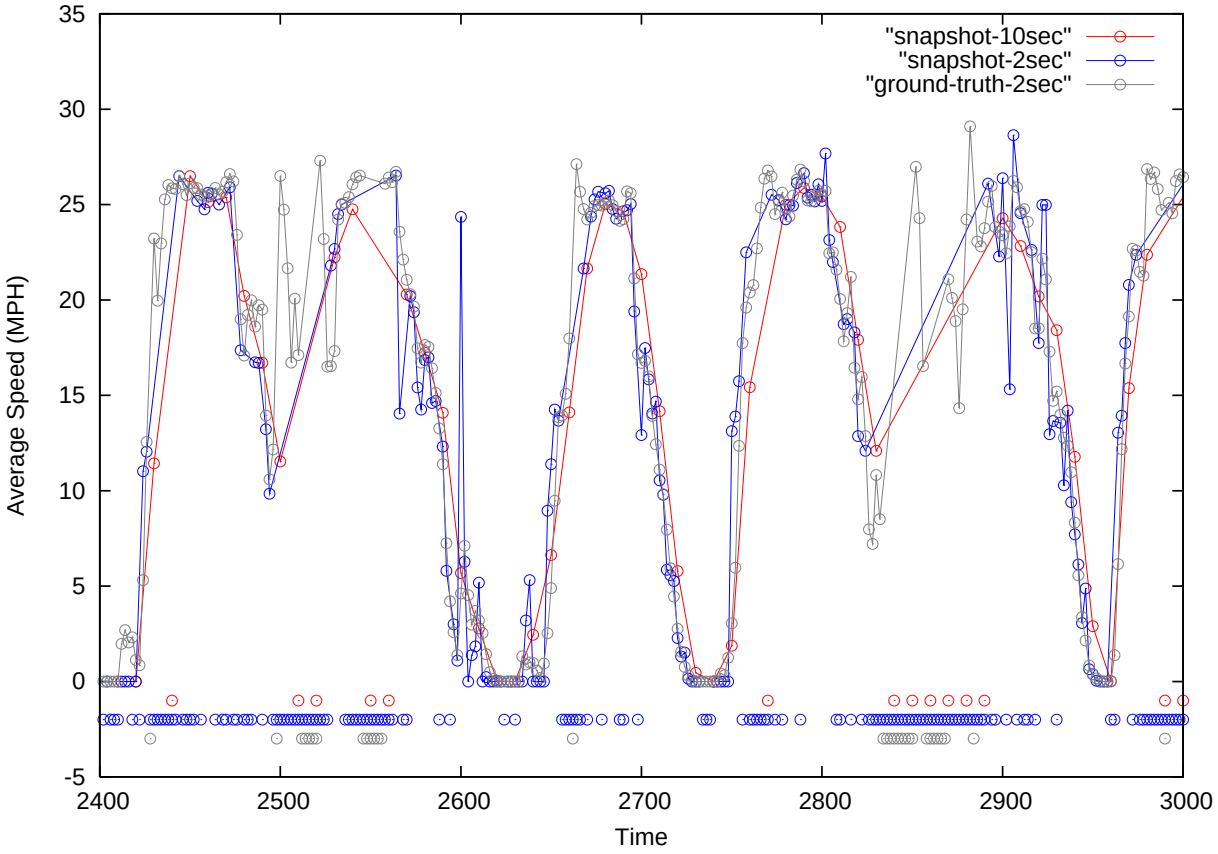


FIGURE 4.22 Effects of Probe Snapshot Sampling on Speed Estimates at 50-100 m from Intersection.

Aggregation of the speed profile samples shown here to produce overall link travel speed statistics must be done with great caution. Simple averaging of the samples produces biased speed estimates because the rate of sampling varies with speed, and the sampling rate is lower when vehicles are stopped than when they are moving. For the corridor that was simulated here, these differences were large enough to be of concern, leading to average probe snapshot speeds that were 16% higher than the baseline truth model.

The results in Figures 4.21 and 4.22 have been based on 100% market penetration of probe vehicles, but it is important to understand how much information will be available at lower market penetrations. Examples for 20% market penetration of probe vehicles are presented next, to show the effect that this has on the availability of speed profile data. Figure 4.23 shows the spatial and temporal aggregation of the complete vehicle trajectories, while Figure 4.24 shows the probe sample snapshots aggregated, both at the 20% market penetration.

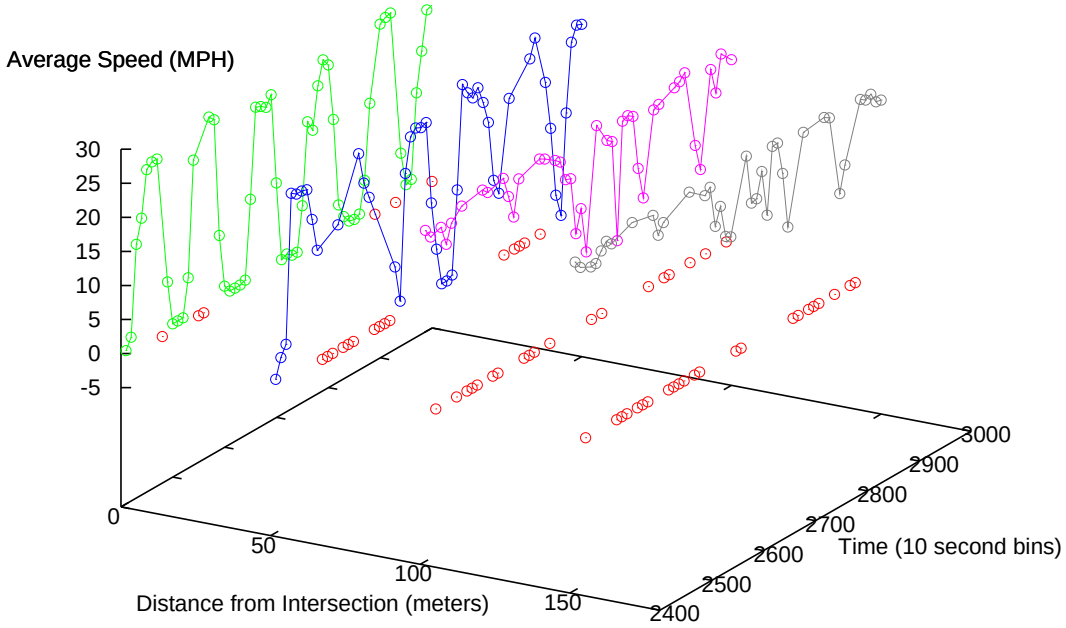


FIGURE 4.23 Aggregation of complete vehicle trajectories at 20% market penetration

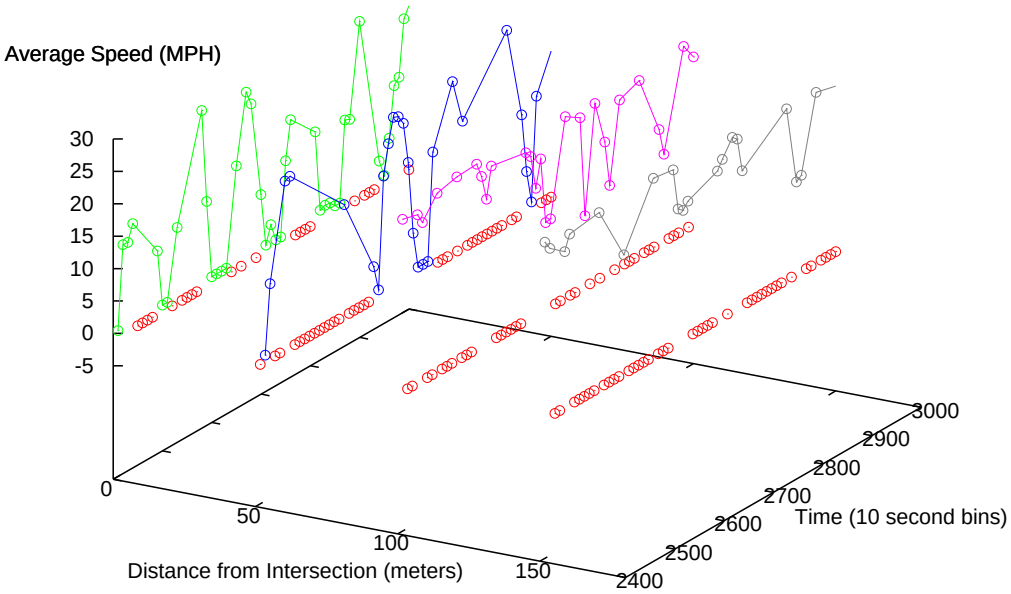


FIGURE 4.24 Aggregation of probe snapshots at 20% market penetration

Figure 4.24 should be compared with Figure 4.23 to see how much information has been lost by the factor of five reductions in the quantity of probe samples. Even in the 50 m bin closest to the intersection, the traffic signal cycle speed patterns have become crude, and they are virtually lost in the next 50 m bin. This calls into serious question the feasibility of using the probe data for real-time adaptive signal control at this level of market penetration.

Taking a closer look at the zone from 50 m to 100 m from the intersection, the effects of the 20% market penetration can be seen more dramatically by direct comparison with full market penetration sampling. Figures 4.25 and 4.26 show the extent to which the shapes of the speed profiles are distorted by the sparser availability of data for sampling. Figure 4.25 is based on the complete raw simulation data, without accounting for the probe snapshot sampling effects. Even at this level, it is evident that the shape of the speed profile is significantly misrepresented during the time from 2500 to 2650 seconds. When the probe snapshot sampling rules are applied, the amount of available raw data is significantly diminished and the quality of the results suffers considerably. Figure 4.26 shows serious losses of data from 2450 to 2650 and again from 2800 to 2900 seconds. This indicates the importance of adjusting the snapshot sampling rules at low market penetration, so that as much data as possible can be captured from the smaller number of equipped vehicles. When considering the 10 s aggregation bins for 20% market penetration, about 31% were empty for the baseline speed profile, but 54% were empty for snapshots following the rules in SAE J 2735, further indicating the importance of capturing as many snapshots as possible in order to obtain statistically valid samples.

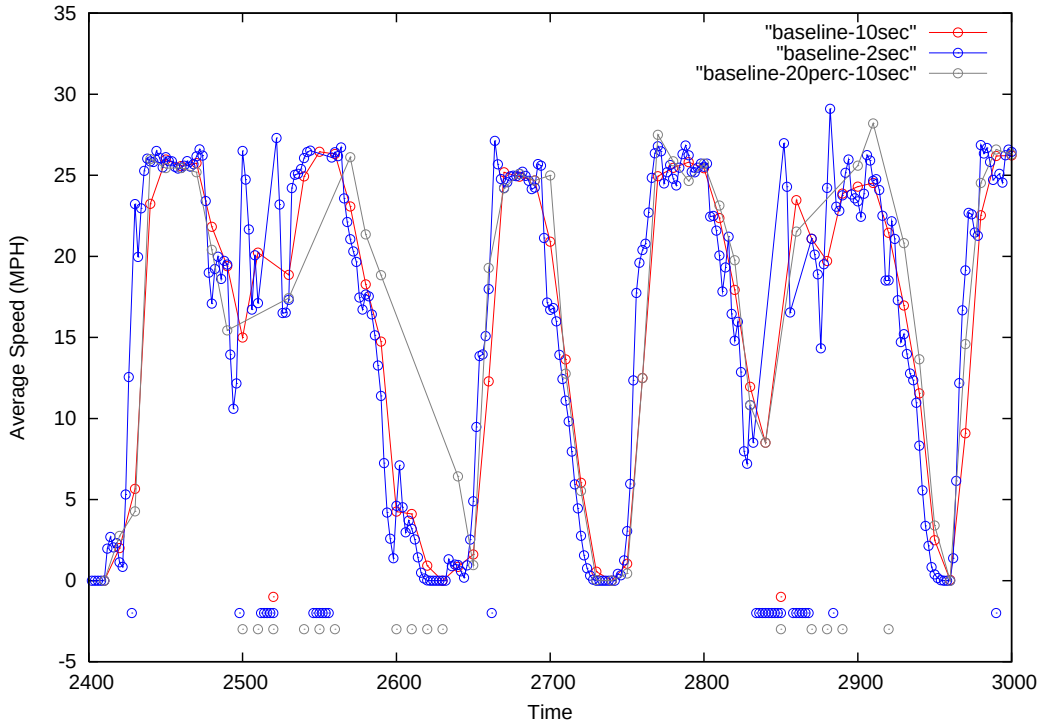


FIGURE 4.25 Raw simulation speed profile averaged over 2 s and 10 s intervals and 20% market penetration averaged over 10 s

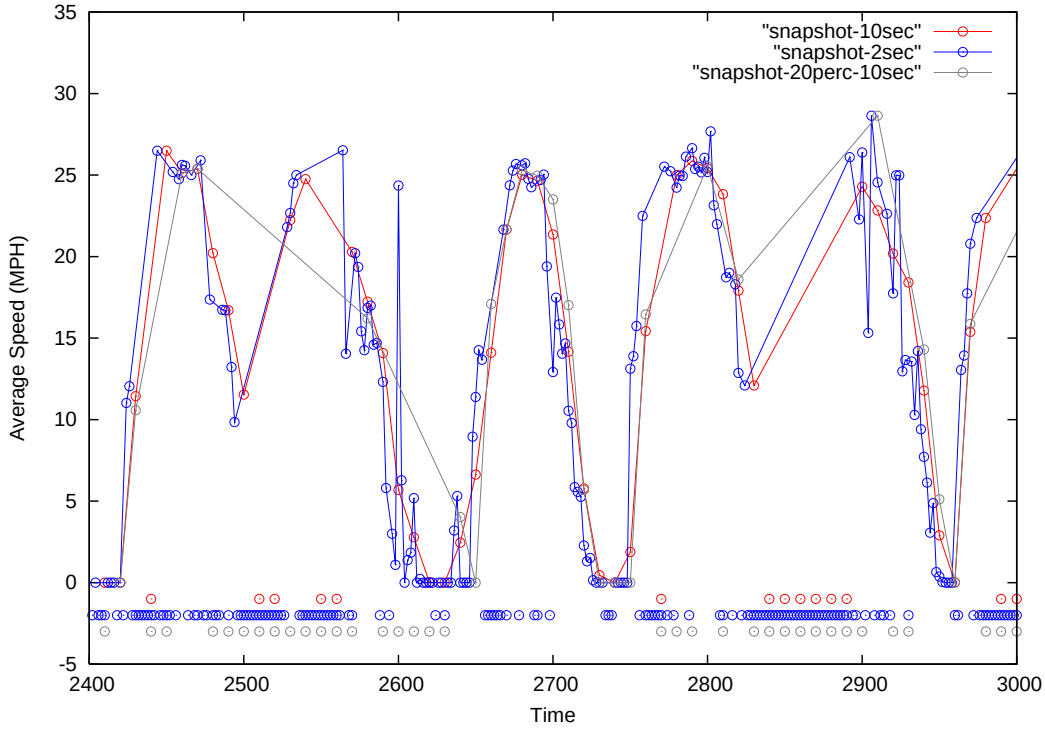


FIGURE 4.26 Simulation snapshot data averaged over 2 s and 10 s intervals and 20% market penetration averaged over 10 s.

4.2.9 Key Findings and Future Direction

Probe vehicle data from VII has been eagerly awaited in the transportation community because of its potential to fill substantial gaps in the available data about transportation operations. A simulation-based study of the probe sampling protocols that are under development for VII has given initial indications of the opportunities and limitations associated with VII probe data. Key findings from this work include:

- The default probe sampling protocols provide reasonable representations of traffic conditions under normal operating conditions, for applications that can tolerate latencies of one to two minutes.
- The restrictions on broadcasting probe messages after a vehicle has first encountered an RSE must be relaxed in order to support time-critical applications such as incident detection and real-time adaptive traffic signal control.
- Because of the high spatial and temporal resolution of the data needed by real-time adaptive traffic signal control, this application is not likely to become feasible until the market penetration of VII-capable vehicles is high.
- When market penetration of VII-capable vehicles is low, the probe sampling limits that were defined in order to economize on message traffic should be relaxed, since the total message traffic will still be low and it will be important to capture as many probe snapshots as possible.

VII California: Development and Deployment
Lessons Learned

Additional research is needed on many aspects of probe data sampling and processing, including:

- exploring the full range of market penetrations for all applications;
- defining probe message processing algorithms to extract network speed and travel time statistics efficiently and accurately;
- simulating probe data in more diverse roadway environments, including urban and rural freeways, dense urban street grid networks, and low-density rural areas;
- developing strategies for managing probe data at the boundaries of equipped regions, where vehicles have traveled for considerable time or distance since last encountering an RSE;
- developing algorithms for locally synthesizing probe data received at clusters of adjacent RSEs;
- developing more efficient representations of the data being acquired from probe vehicles.

4.3 APPLICATION OF VII DATA ON REAL-TIME ARTERIAL PERFORMANCE MEASUREMENTS

4.3.1 Introduction

The VII communication between cars and roads could benefit both parties and could be very powerful, as it would enable the full vision of intelligent transportation system (ITS). For example, the information transmitted from the roadside to the vehicle could warn a driver that it is not safe now to enter an intersection. On the other hand, vehicles could serve as data collectors and anonymously transmit traffic and road condition information from all major roads which are equipped with such communication infrastructures. Such information would significantly improve the quality and quantity of traffic data, which in turn will help transportation agencies and all other stakeholders to better understand, plan, and manage the transportation system.

The VII-based 'traffic detection' contains much richer sets of information than existing traffic detection means, along with greater accuracy. It is envisioned that the information provided by VII will enable a variety of advanced applications previously not possible using existing traffic detection means. Per National VII Coalition discussions in June 2005, advanced traveler information system (ATIS) is among the most prioritized and promising "Day 1" – or initial – VII applications. One of the most critical elements of ATIS is to provide reliable and accurate traffic information, such as travel time and average speed, for pre-trip route planners and en-route travelers. If the current and/or predictive travel time information could reach drivers through the V-I communication, "smart" route choices would be made at the system wide. Therefore, win-win situation can be realized as both drivers' satisfactions and the efficiency of the roadway system would be improved.

As the state of practice, travel time information estimation is based on either aggregate point-based measurement, i.e. volume, occupancy, and/or spot speed, or some individual section-based measurements, i.e. probe vehicle travel time, and/or average travel speed. However, both of the two popular techniques are constrained by some major limitations. For the point-based estimation technique, the Inductive Loop Detector (ILD) is the widely used detection technology. ILDs provide vehicle presence, traffic count and, when two consecutive ILDs are used, vehicle speed. Because tuning ILDs for optimal threshold is very time consuming, most existing ILDs over- or under-estimate traffic counts. The accuracy of the outputs from ILDs is low. According to some field experiments [Li, et al, 2007], the absolute relative error for fine tuned ILDs can be reduced to about 25%. This low accuracy may not cause major issues for close-loop actuated signal control, but it presents significant problems for most other applications such as traffic flow estimation, travel time estimation, and adaptive signal control. More importantly for an urban street network, the average travel time, which consists of link travel time and control delay at intersections, is interruptive rather than continuous. Therefore, the point-based measurement may not fully reflect the roadside situations at intersections. And the location of ILDs is important to help detect traffic conditions such as level of service and under- or oversaturated [Ayoub, 2007], although most ILDs are installed at 220 feet away from the intersection stopline. Again, the fixed location is designed for close-loop traffic signal control but unfavorable for the traffic information estimation.

While for the section-based measurement, the travel time estimation highly depends on the penetration rate of probe vehicles in the road network. Due to the dynamic and stochastic attributes of individual vehicles, the travel time has relative high variance-to-mean ratio (VMR). Limited number of probe vehicle trips alone cannot provide enough information for a statistically reasonable estimation. In view of the current limited penetration rate of equipped probe vehicles and/or the low quality transit information from their automatic vehicle location (AVL) systems, the coverage in most urban networks is far from sufficient [Xie, Cheu and Lee, 2004].

As an innovative section-based detection technique, VII is envisioned that the market penetration will keep growing. At some point, it will reach a desirable penetration level so that many of ITS applications will rely on its sound and comprehensive traffic information. But before the high penetration rate is achieved, the effects of VII may not be significant. In this case, VII-equipped vehicles will be used as vehicle probes. During this transition period, it would be invaluable if VII-based traffic detection can be used in combination with the point-based traffic detection means because the drawbacks of the two detection techniques are not overlapped. The fused traffic detection information can provide more comprehensive information and better accuracy than with each of data source alone and can begin to offer greater benefits. In fact, fused VII-based and traditional traffic detection approach could be a desirable solution even after the penetration of VII based vehicles becomes high because the information provided by the two detection means are complementary.

In this research, we attempted to develop a robust travel time estimation model which utilizes the information collected through the V-I communication in concert with those collected by conventional point-based traffic detectors. It is expected that the fusion model would fully take advantage of the merits of the two data collection means and improve the accuracy, robustness, and coverage of the estimation results which exclusively applies single data source.

The remainder of the report is organized as follows. The VII probe data and traffic data characteristics are first introduced. Then the developed estimation models which customized for such data sources are presented. The third section demonstrated the applications of the models on a simulation arterial. The model performance comparisons are also presented in the third section. The fourth section is a discussion of potential recommendation for the draft VII message process and standard. The last section concludes this report with recommendations for future research.

4.3.2 Data Characteristics

4.3.2.1 VII Probe Data

It is worth reiteration that in the VII system, OBE, DSRC, and RSE are the key elements involved in probe data collection. The OBEs are the vehicle side of the VII system, while RSEs may be mounted at interchanges, intersections and other locations providing the interface to vehicles within their range. At intervals, the OBEs gather and store vehicle data, which typically includes vehicle trajectory data (speed, position, etc.), vehicle device status data (turn signal, brake, etc), event-driven data (antilock brake, airbag activation, etc.), as a series snapshots in OBEs' memory buffers. When the probe vehicle is within the range of RSE, all the stored snapshots will be sent to RSE via DSRC communication. Then, the OBE buffer will clear off and begin to store new snapshots data until the vehicle meets next RSE. Through a certain

communications network, the collected probe data will be further transmitted from RSE to various data user [4]. The VII probe data collection process is illustrated in FIGURE 4.27.

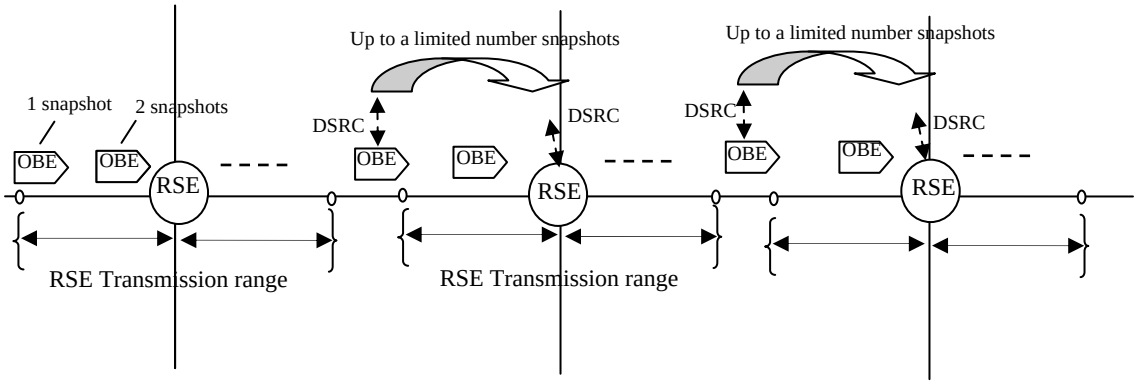


FIGURE 4.27 VII Probe Data Collection Process

Currently, the VII probe message process is still under development. The corresponding standards are yet to be finalized. The Society of Automotive Engineers (SAE) J2735 standard issued in March 2007 [SAE, 2007], is probably the most recent draft version, in which, some basic concepts, parameters regarding traffic probe message are defined. If not specifically denoted, the parameters' default values given in this report are accordant with SAE J2735.

VII probe vehicles store two types of snapshots, periodic snapshots and event snapshots. Periodic snapshots are generated at specific time intervals. The interval between two consecutive snapshots is determined by the speed of the probe vehicle. Logically, the fast the vehicle is traveling the longer interval between snapshots and vice versa. According to J2735, the default values of low speed and high speed thresholds are 20 mph and 60 mph respectively, the corresponding short interval and long interval are 6 and 20 seconds. Event snapshots are triggered when specific vehicle events occur, including the application of anti-lock brakes, dynamic traction control, or other infrequent but important events. In addition, vehicle stops and starts are also treated as triggered events. A 'start' is defined as when the vehicle speed exceeds a specified threshold (default is 10 mph), a 'stop' is defined as when vehicle does not move forward for a specific 'stop' time (default 5) and no other stops have occurred within the 'last stop' specific time (default 15). The 'last stop' specific time threshold is intended to prevent multiple snapshots when vehicles are creeping forward. No snapshots are taken between a 'stop' and 'start'.

Snapshots association is to identify the snapshots generated by the same VII probe vehicle at different time points, which can be used to trace vehicle trajectory. However, due to privacy consideration, VII snapshots of the same vehicle do not have the same ID all through its entire path. Instead, a temporary ID is generated periodically and persists within a limited distance or a duration time since its generation, which ever comes later. According to J2735, the default values of the distance and duration limitation are 1 kilometer and 120 seconds respectively.

The maximum number of snapshots stored on the OBE is determined by the buffer storage capacity (default 30). When the buffer is full and a new snapshot is created, a previous snapshot will be deleted based on specific retention priority rules. By default, the event snapshots has

highest retention priority, stop and start snapshots are at the second priority level, periodic snapshots are at the lowest level. Among snapshots at the same priority level, the most recent snapshot has highest priority.

When a VII probe vehicle is within range of an RSE (default 400 feet), it transmits all snapshots stored on the OBE. After that, by default, the vehicle will not send any more report to RSE when it is in range of the RSE.

VII probe data, as a section-based data, can be considered as a better data source for travel time estimation compared with point-based data. However, based on the above discussion, VII probe data have some obvious characteristics which may influence our travel time estimation.

VII probe vehicles generate snapshots at discontinuous time points with intervals ranging from 4 seconds to 20 seconds, trajectory data between intervals will not be collected and thus may cause travel time estimation error. The association of snapshots data at different position is the direct and accurate way to measure the travel time over a specific road segment. VII probe data may provide a certain level of association between different snapshots. However, as the unique vehicle temporary ID can only 'live' over a limited distance or time period, data associations are restricted within a certain travel distance. In city network, considering 1 km and 2 minutes as default boundary, data association typically may be feasible over 2-3 blocks.

Since the maximum snapshot number is limited by buffer capacity, some snapshots may be overwritten by newer snapshots, thus the trajectory data over some specific segment may not be available. Basically, snapshots data space coverage is primarily determined by buffer memory capacity, RSE distribution and snapshots retention priority management as well. Since the VII probe vehicles generate new snapshots after entering a current RSE and then send out all stored snapshots until meet next RSE. Thus, there will be a time lag between the data generated and received time. This time lag is primarily determined by the distance between RSEs and the probe vehicle speed. If assuming the distance between 2 RSEs is 600 meters, vehicle average speed is 35 km/h, then the maximum lag time is about 1 minutes, which is from the generation time of the first snapshot at last RSE to the transmission time at next RSE. Some other factors, such as GPS position error, mapping error, communication delay, may also have impacts on VII probe data quality.

4.3.2.2 Traffic Data

The traffic data consists of two parts: static inputs and dynamic inputs. The static inputs include historical cruise speeds, saturation flows, geometric info such as link lengths and detector locations, and traffic signal timing parameters. The dynamic input consists of traffic signal status, ILDs counts and ILDs occupancies. All the dynamic data are time stamped for synchronization purpose and in a high resolution with 1~5 seconds updating rate. With the deployment of communication links and ITS technologies, such dynamic data are more readily available. For example, Parsons Traffic and Transit Laboratory (Parsons T^2 Lab) at University of California is now collecting second-by-second traffic signal status, ILDs counts and occupancies in the real time from more than 50 intersections along the El Camino Real Corridor at Santa Clara County, California and the San Pablo Corridor at Alameda County, California [Li, et al, 2007].

4.3.2 Models Formulation

4.3.2.1 VII Probe Data (VPD) Model

VPD model is to estimate average link travel time based on the VII snapshots data. A link is defined as a road segment from upstream intersection stopline to the downstream intersection stopline. From the statistic point of view, the best estimate of the link travel time is the mean of the difference between the link check-out and check-in times. However, as aforementioned, VII probe vehicles may change their temporary IDs at any point of a link. Thus some of the collected check-out snapshots cannot be associated with check-in snapshots because of the unmatched vehicle IDs. All of such snapshots have to be discarded. Therefore, the method does not fully utilize the collected VII data which might lead to a high estimation error, particularly when the penetration rate is low and the link is long.

From this point of view, the VPD model has been developed in order to make good use of not only the associated link travel times but also the snapshot data which cannot be associated. Because the probability of vehicle ID change is spatially uniform, thus the shorter the observation section has lower percentage of trips which cannot be associated. Considering an urban street link can be divided into a cruise travel section and an intersection impacted section, we divides each link into an upstream section and a downstream section accordingly, as illustrated in the middle of FIGURE 4.28. The downstream section refers to the approaching segment close to the downstream intersection, while the upstream section is the rest of the link farther from downstream intersection.

Different strategies regarding the associated and unassociated data are applied to the two sections. Typically, the upstream section is longer and has smaller space speed variations comparing with the downstream section. Thus, in addition to the associated trips, other snapshots data can also be used to estimate the average space speed over the upstream section. At this point, to some longer link, the upstream section may necessarily be further divided in order to more precisely reflect the space speed variation spatially. While in the downstream section, the associated trip data is more appropriate because average space speeds estimated from VII snapshots might not be able to capture the speed variations. In principle, the separation point of the two sections should be the boundary of the intersection impact area, so the maximum back-of-queue is a good candidate. The point-based detection can help update this boundary for each estimation interval. For simplification, 100 meters (328 feet) will be applied in the latter model application section.

4.3.2.2 Model for Downstream Section

Snapshots association, trajectory interpolation and travel time aggregation are the three steps for the downstream section travel time estimation. Snapshots association is to identify snapshots near the boundaries of the downstream section. Then, the kinematics model is applied to estimate the exact check-in and check-out time at the boundaries. Finally, the travel time can be calculated and aggregated for statistic analysis.

FIGURE 4.28 shows a typical snapshots distribution near and within the downstream section. Typically, there are four snapshots need to be identified, which are Snapshot Before Start point (SBS), Snapshot After Start point (SAS), Snapshot Before End point (SBE) and Snapshot After End point (SAE). An effective sample must at least have one of the SBS and SAS, in the mean

time, the SAE must be available. After that, based on SBS and/or SAS, the arrival time at the start point can be estimated. Similarly, based on SBE and SAE, the arrival time at the end point can also be deduced.

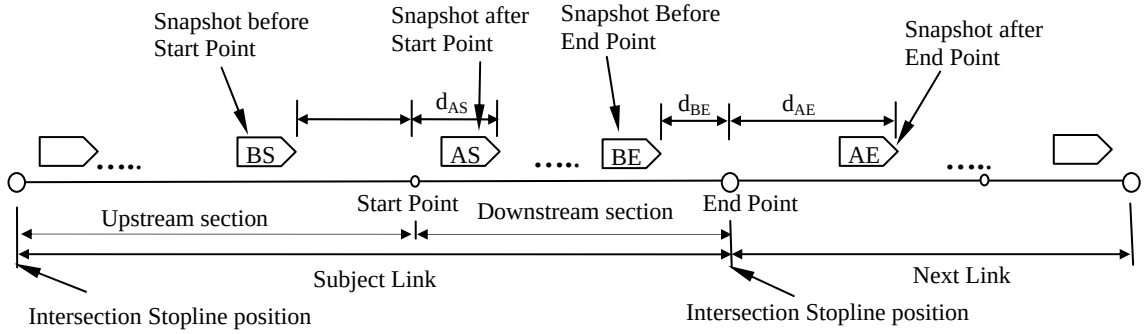


FIGURE 4.28 Typical snapshots distribution near around downstream section

Since the associated snapshots may not be exactly generated at the start and end point, trajectory interpolation is needed to estimate the arrival time at start and end point. For the interpolation, there are various scenarios with regard to the availability of SBS and SAS.

Both SBS and SAS are available
The arrival time at the start point (SP) can be estimated using equation (1) and (2). The reason why the acceleration rate is estimated using distance measurement is because the instant time measurements for event snapshots are not as accurate as distance measurements.

$$\hat{a}_{BS\to AS} = \frac{v_{AS}^2 - v_{BS}^2}{2(d_{AS} + d_{BS})} \quad (1)$$

$$\hat{t}_{SP} = t_{AS} - \frac{1}{\hat{a}_{BS\to AS}}(v_{AS} - \sqrt{v_{AS}^2 - 2\times \hat{a}_{BS\to AS}\times d_{AS}}) \quad (2)$$

where $\hat{a}_{BS\to AS}$ = the average acceleration between the measured Before-Start (BS) point to the After-Start (AS) point; v_{AS} , v_{BS} , t_{AS} = measured snapshot speeds and times; d_{AS} , d_{BS} = the distance from SP to AS and BS; and $\hat{t}_{SP}$ = the estimated arrival time at SP.

Because SAS is within the queue range, some SAS are possibly stop event snapshots. A "Stop" snapshot is generated when the vehicle have been in still for a threshold time, with default 5 seconds. Thus, equation (2) is adjusted to equation (3). Similar strategy will be applied if any of the snapshots are "stop" events.

$$\hat{t}_{SP} = t_{AS} - 5 - \frac{1}{\hat{a}_{BS\to AS}}(v_{AS} - \sqrt{v_{AS}^2 - 2\times \hat{a}_{BS\to AS}\times d_{AS}}) \quad (3)$$

Only SBS available
When no snapshot after SP is found other than SAE, it actually implies this probe vehicle passed stopline without a stop. Thus, SBS is also SBE. $\hat{t}_{SP}$ can be estimated by using equation (4) and (5).

$$\hat{a}_{BS\to AS} = \hat{a}_{BS\to AE} = \frac{v_{AE}^2 - v_{BS}^2}{2(d_{AE\to SP} + d_{BS})} \quad (4)$$

$$\hat{t}_{SP} = t_{BS} + \frac{1}{\hat{a}_{BS\to AS}}(\sqrt{v_{BS}^2 - 2\times \hat{a}_{BS\to AS}\times d_{BS}} - v_{BS}) \quad (5)$$

where $d_{AE\to SP}$ = the distance from AE to SP.

Only SAS available

When no other snapshot can be referred before AS, $a_{BS\to AS}$ may be assumed to the same as $a_{AS\to BE}$, or $a_{AS\to AE}$ if SBE is the same as SAS. The adjust equation is similar with equation (4) and (5).

However, such algorithm does not work when SAS is a 'Stop' event. Then the vehicle is assumed to make a smooth stop with an acceleration rate $\bar{a}_{SS}$ and stops at AS. Consequently, t_{SP} can be estimated by applying equation (6).

$$\hat{t}_{SP} = t_{AS} - 5 - \sqrt{\frac{2\times d_{AS}}{\bar{a}_{SS}}} \quad (6)$$

Both SBE and SAE are available

Without considering the overlaps with SBS and SAS, SBE and SAE can always be found. The arrival time at the end point t_{EP} can be interpolated in the similar way as equation (1) and (2). However, if SBE is a "Stop" event, and SAE is a "Start" event, the probe vehicle is assumed to speed up with a constant acceleration $a_{BE\to AE}$ from BE to AE. Then t_{EP} can be estimated using equation (7) and (8).

$$\hat{a}_{BE\to AE} = \frac{v_{AE}^2}{2(d_{BE} + d_{AE})} \quad (7)$$

$$\hat{t}_{EP} = t_{AE} - \frac{1}{\hat{a}_{BE\to AE}}(v_{AE} - \sqrt{v_{AE}^2 - 2\times \hat{a}_{BE\to AE}\times d_{AE}}) \quad (8)$$

With $\hat{t}_{SP}$ and $\hat{t}_{EP}$, the average travel time in the downstream section during the period from *T* to *T'* can be estimated using equation (9).

$$\bar{t}_{T\to T'}^{downstream} = \frac{\sum_{i=1}^{N(T\to T')}(\hat{t}_{EP}^i - \hat{t}_{SP}^i)}{N(T\to T')} \quad (9)$$

where $N(T\to T')$ = the total number of associated VII probe vehicle trips in the downtown section during the period from *T* to *T'*.

4.3.2.3 Model for Upstream Section

The estimation of upstream section travel time $\bar{t}_{T\to T'}^{upstream}$ can be separated into three steps:

1) For the associated trips, the downstream section model can be applied. With the reasonable number of associated trips, the average speed approach below will not be applied.

2) On some portions of the upstream section, there might be only some discrete snapshots. Those data cannot be associated together to form vehicle trajectories. The average speed of these isolated data will be used as the average travel speed on corresponding portions. It is also very likely that there may be no VII vehicle snapshot data coverage at all for some portions of the upstream section. An average speed will be computed from all the VII snapshot data and will be used as the average travel speed for the portions without VII data coverage.
3) Travel time of different portions of upstream section will be assembled together to form the travel time of whole upstream section.

4.3.2.4 Point-Based Detection (PBD) Model

The point-based detection model is different from the aforementioned VII model which separate link travel time geometrically into separate travel times on the upstream and the downstream section respectively. This model separates link travel time by the control impacts, thus the link travel time $\bar{t}_{m,j}^{link}$ is modeled as equation (10) which is the sum of the link cruise time and the control delay introduced by the traffic signal. The analytical model estimates the link travel time for each lane group and each signal cycle.

The average link cruise speed $\bar{v}_{m,j}^{cruise}$ is calculated by using the speed limit V^{limit} and the average probe vehicle speed $\bar{v}_{m,j}^{probe}$. A hybrid model (11) using the probe vehicle penetration rate $\rho_{m,j}^{probe}$ with the power of α as the weighting factor has been developed to balance the credibility of probe vehicle measurement. FIGURE 4.29 shows the performance of a simple hybrid model with $\alpha = 1/6$.

$$\bar{t}_{m,j}^{link} = \frac{L}{\bar{v}_{m,j}^{cruise}} + \bar{d}_j^m \tag{10}$$

$$\bar{v}_{m,j}^{cruise} = V^{limit} + (\rho_{m,j}^{probe})^{\alpha} \times (\bar{v}_{m,j}^{probe} - V^{limit}) \tag{11}$$

$$\rho_{m,j}^{probe} = \frac{N_{m,j}^{probe}}{N_{m,j}^{ILD}} \tag{12}$$

where $\bar{t}_{m,j}^{link}$ = the average link travel time (sec) for lane group *m* and signal cycle *j*; *L*= the link length; $\bar{v}_{m,j}^{cruise}$ = the average link cruise speed (meter/sec) for lane group *m* and signal cycle *j*; $\bar{d}_j^m$ = the average control delay for lane group *m* and signal cycle *j*; V^{limit} = the link speed limit (meter/sec); $\rho_{m,j}^{probe}$ = the probe vehicle penetration rate (%) for lane group *m* and signal cycle *j*; α = weighting parameter; $\bar{v}_{m,j}^{probe}$ = the average probe vehicle link travel speed for lane group *m* and signal cycle *j*; $N_{m,j}^{probe}$ and $N_{m,j}^{ILD}$ = the number of detected probe vehicles and the number of vehicles detected by ILDs for lane group *m* and signal cycle *j*.

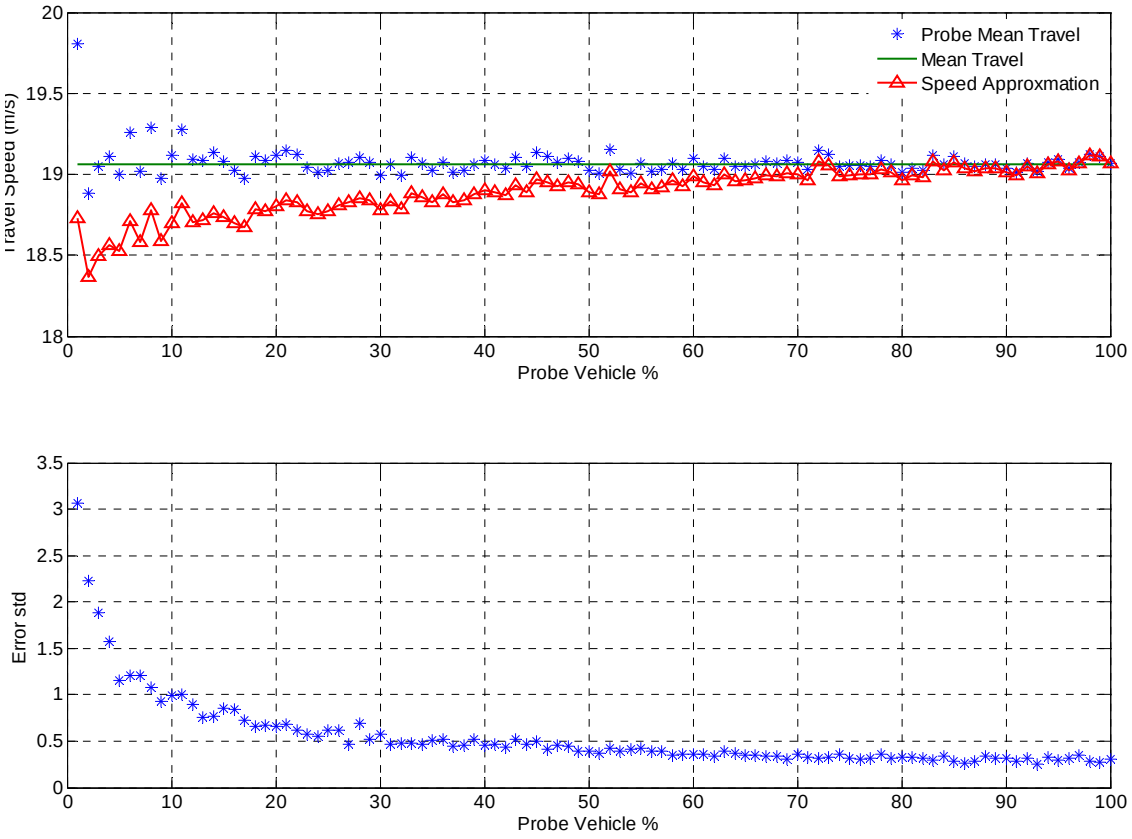


FIGURE 4.29 Performance of the Cruise Speed Estimation Model

For the ideal case where both arrival detection and departure detection are present and appropriately located, the departure N-T curve and the arrival N-T curve can be readily obtained. The average control delay can be estimated by using cumulative plots [Dion, Rahka, Kang, 2004]. A virtual arrival curve is defined by shifting the actual arrival curve towards the location of departure detectors by the free flow travel time from the arrival detectors to the departure detectors. Then the total control delay can be estimated by calculating the area of the region between the virtual arrival curve and the departure curve. The realization of the trajectory and delay for each individual vehicle is not needed when the average control delay is the primary goal. Therefore for such scenario with ideal detection setup, equation (13) can be applied to calculate average control delay for lane group *m* and time period from *0* to *T*.

$$\bar{d}^{m}_{0\to T} = \frac{\int_{t=0}^{T} D^{m}(t)dt - \int_{t=0}^{T} VA^{m}(t)dt}{\sum_{t=0}^{T} D^{m}(t)} \qquad (13)$$

where $\bar{d}^{m}_{0\to T}$ = the average control delay (sec/veh) for lane group *m* and time period from *0* to *T*; $D^{m}(t)$ = the instant departure count at time *t* for lane group *m*; and $VA^{m}(t)$ = the instant virtual arrival count at time *t* for lane group *m*.

However, most exiting signalized intersections do not have the ideal detection layout. For example, the most popular scenario which is under semi-actuated coordination only has arrival detectors for the main street through directions (sync NEMA movement 2 and 6) and only has departure detectors for the non-sync movements [California DOT, 1983], as shown in the bottom part of FIGURE .30. Consequently, none of the movements can directly apply equation (1). Alternatively, approximate approaches were developed to estimate the average control delay.

As illustrated in FIGURE 4.30, cycle j for the sync movement 2 or 6 started at $c^{j-1}+1$. The green started at g^j. We assumed that the queues would be released with the saturation flow rate μ after the signal light changes from red to green plus a lost time interval $g^j + T^{Lost}$. The lost time aims to account for the additional delay due to driver reaction time and vehicle acceleration constraints. At the advance loops which is L =220 feet away from the stopline, arrival count is detected until the queue length is longer than L. When queues extend over the advance loops t_{LQ}^j, the observed occupancies at advance loops will be a high value for a fraction of the cycle. The detected curve after t_{LQ}^j will be the departure curve of queues rather than the actual arrival curve. Therefore, we assumed traffic arrivals are uniformly distributed with the average arrival flow rate λ after t_{LQ}^j until the end of the cycle c^j. t_{LQ}^j for such a long queue is detected by a near zero flow and a near one occupancy for a critical time T^{LQ}.

Similarly with the aforementioned ideal case, a virtual arrival curve is defined by shifting the actual arrival curve towards the location of departure detectors by the free flow travel time from the arrival detectors to the departure detectors T^{FF}. Then the total control delay can be estimated by calculating the shadowed area. And the queue clearance time t_{CQ}^j can be calculated by the assumed departure curve and the virtual arrival curve. Consequently, for under-saturated scenario, the average control delay for lane group *m* can be calculated using equation (14). For the over-saturated scenario when $t_{CQ}^j > c^j$, an initial queue and the corresponded extra delay will be calculated for the next cycle.

$$\bar{d}_j = \frac{\frac{\mu}{2}[(t_{CQ}^j - (g^j + T^{Lost})] \times [(t_{CQ}^j + (g^j + T^{Lost}) - 2(c^{j-1}+1)] - \sum_{t=c^{j-1}+1}^{t_{CQ}^j} A(t - T^{FF}) \times (t - c^{j-1})}{\sum_{t=c^{j-1}+1}^{c^j} A(t - T^{FF})} \quad (14)$$

where $\bar{d}_j$ = the average control delay (sec/veh) for cycle *j*; μ = the saturation flow rate; t_{CQ}^j = the clearance queue time for cycle *j*; g^j = the green start time for cycle *j*; T^{Lost} = the lost time; c^j = the end time for cycle *j*; and $A(t)$ =the instant arrival count at advance loops at time *t*; T^{FF} = the free flow travel time from advance loop to the stopline.

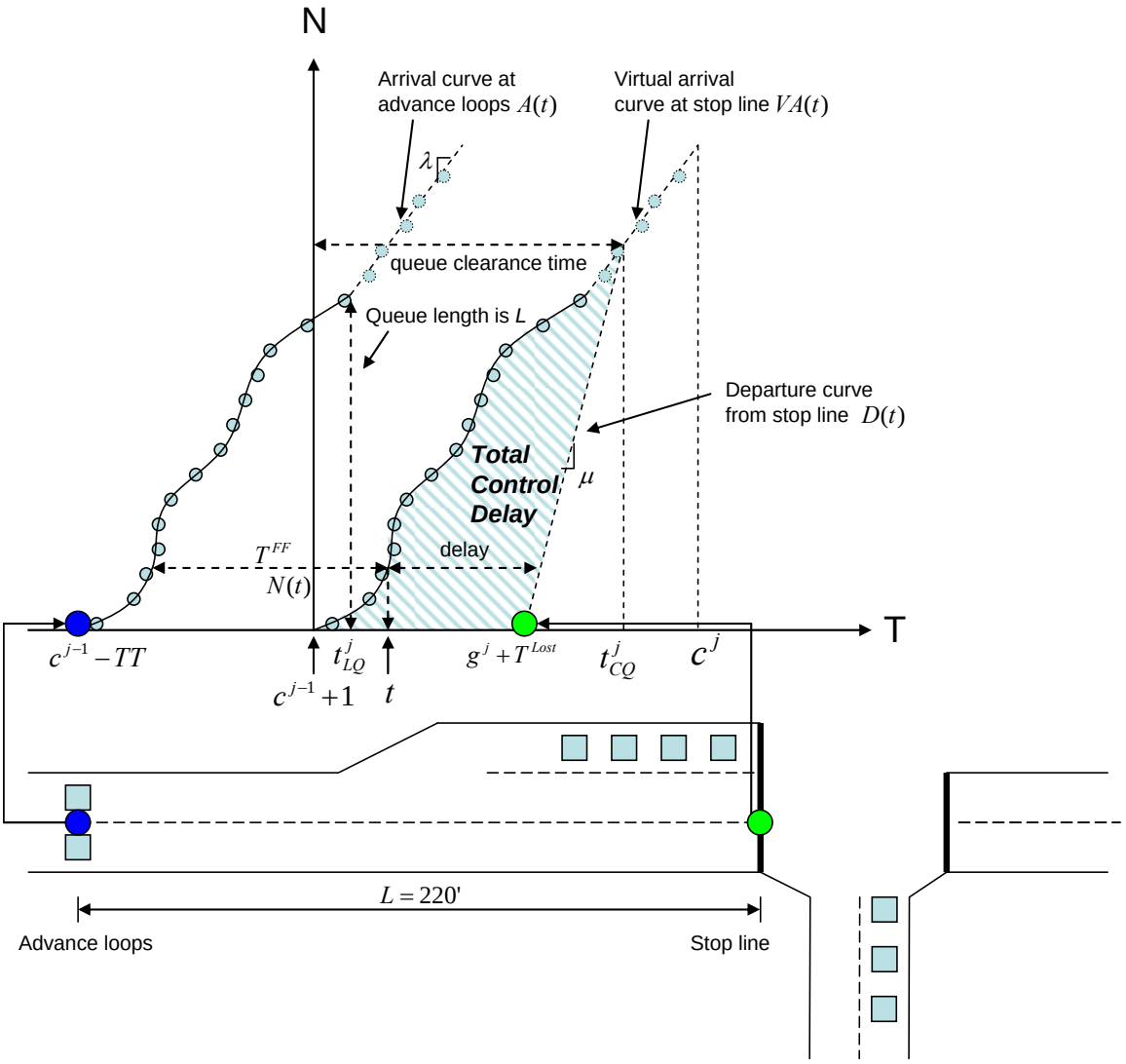


FIGURE 4.30 Delay Calculation Model with Typical Semi-actuated Control Layout

4.3.2.5 Fusion Model

Although data from fixed loop detector and VII probe vehicle cover different aspects of the arterial traffic information, a mathematical model which explicitly links these two data source is not readily available. Therefore, a "black box" approach, i.e. a multi-layer neural network approach which has the advantage to fuse two data sources without exploiting explicit mathematical relationship between two data sources, is adopted in this paper. FIGURE 4.31 shows the structure of the multi-layer neural network fusion module. The input is composed of estimation from the aforementioned PBD model and the VPD model together with VII probe vehicle penetration rate. Following the 2n-1 rule where n is the number of input nodes, the hidden layer is chosen to include 5 nodes with Tan-Sigmoid transfer function at first. The results from initial tuning suggest that 7 nodes provide better fusion results. The output layer has one node with linear transfer function which is the fused travel time estimation.

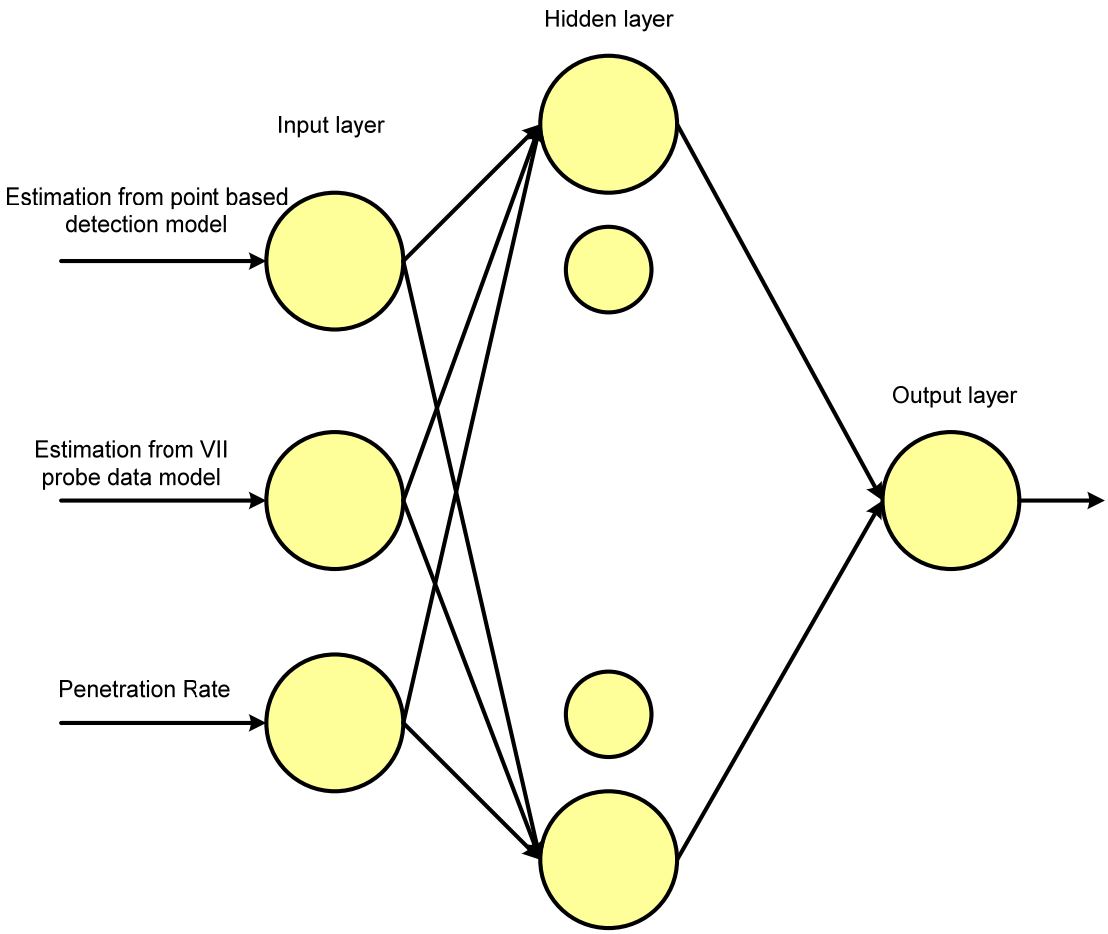


FIGURE 4.31 The Structure for the Neural Network Fusion Model

4.3.3 Model application

4.3.3.1 Simulation-Based Testbed

As shown in FIGURE 4.32, a stretch of El Camino Real arterial in Palo Alto, California with six intersections and 1-mile-long was selected as the testbed of our model. Under the ongoing VII California project, some prototype RSEs have been installed and tested for the testbed. We constructed the arterial using the popular microscopic simulation model VISSIM. Three RSEs are considered at Dinah, San Antonio, and Jordan respectively, among which Dinah is the northmost intersection, Jordan is the southmost intersection, San Antonio is in the middle and also the busiest intersection within the stretch.

VISSIM can record second by second vehicle trajectory data, which were used as the ground truth. Then the TCA (Trajectory Conversion Algorithm) tool was used to get the snapshots data out of the whole second by second vehicle trajectory data. The TCA algorithm emulates the VII probe message process accordance with SAE J2735 standard. Also, from VISSIM simulation,

second by second detectors data can be recorded. With the snapshots data and detectors data, the proposed VPD model and PBD model can be applied.

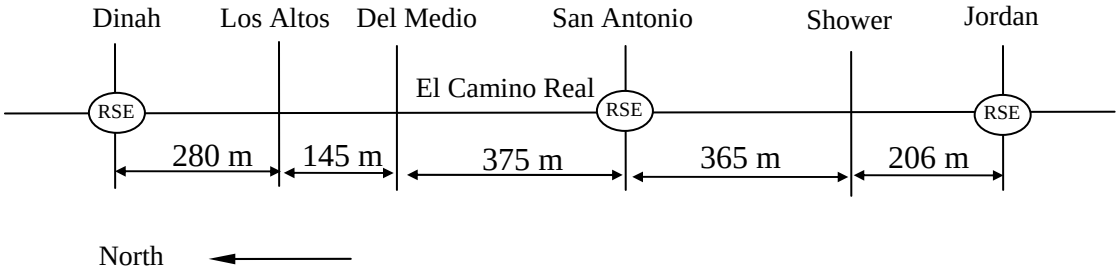


FIGURE 4.32 Simulation network and RSEs deployment

4.3.3.2 Preparation of Fusion Model

Levenberg-Marquardt backpropagation algorithm is adopted for the neural network training. MSE (mean square error) is used as the error criterion. From 432 samples, 259 samples are used for the training and the remaining samples are divided equally between validations and testing. The training stopped after 123 epochs and the results are presented in the next section.

4.3.3.2 Application Results

FIGURE 4.33 shows the comparison estimation results for the north bound link between Shower and San Antonio.

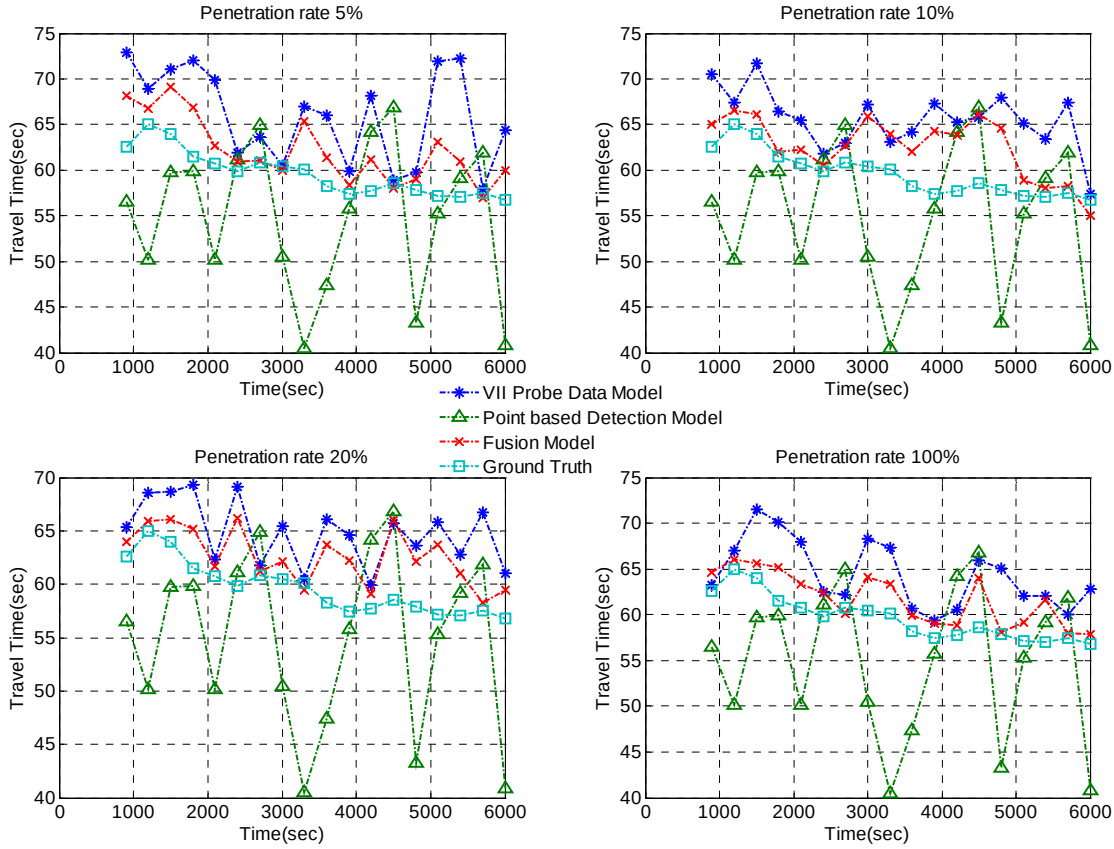


FIGURE 4.33 Model Performance Comparisons for a NB Link

Estimation results from VPD model, PBD model and the fusion model with different penetration levels are plotted with the ground truth data. Generally speaking, VPD model does perform better than the PBD model due to the more accurate information. Fusion model estimation achieves the best results among the three. With the increase of penetration rate, the estimation results are also improved for both VPD model and the fusion model.

To further illustrate the effectiveness, TABLE list the mean absolute percent error (MAPE) for both link level and route level estimation and with different VII probe vehicle penetration rates. Again, similar trend can be observed.

TABLE 4.2 Model Performance Comparison
(Mean Absolute Percent Error %)

	Link Travel Time				**Arterial Travel Time**			
Penetration Rate	5%	10%	20%	100%	5%	10%	20%	100%
	Northbound							
PBD Model	10.1	10.1	10.1	10.1	13.9	13.9	13.9	13.9
VPD Model	13.6	9.1	8.9	7.6	13.5	9.7	9.0	8.1
Fusion Model	**8.1**	**7.5**	**7.3**	**5.4**	**9.1**	**8.0**	**7.8**	**6.1**
	Southbound							
PBD Model	10.2	10.2	10.2	10.2	11.7	11.7	11.7	11.7
VPD Model	12.5	8.7	8.5	7.1	13.1	9.1	8.5	7.8
Fusion Model	**7.2**	**7.0**	**7.0**	**5.2**	**8.5**	**7.9**	**7.2**	**5.9**

4.3.4 Discussion

As discussed in the data characteristics section, different options of the parameters in VII message process may result in different properties in data association, data spatial coverage and data latency, thus may have different influences on our travel time estimation accuracy. In this paper, only default values are used to test the travel time estimation models. Thus, the sensitive analysis of various parameters options on travel time estimation need to be further studied. However, there are still some preliminary understandings or concerns can be gain from this study:

Snapshots association is a key factor influencing the VPD model. For the link travel time estimation, it's preferred to have associated snapshots to cover the whole link, which can be used as another input the fusion model. However, the percentage of associated probe trips depends on the link length. Typically in the U.S., the distance between two urban street intersections is 80~280 meters (260~900 feet) [Ben-Joseph and Szold, 2005], it means for the worst scenario more than a quarter of the probe vehicle trips are disjoined within a link. In order to guarantee there is a link trip between any two ID changes, the minimum ID persistence distance is should be at least a double of a block distance which is 560 meters.

According to the "Start" event definition, vehicles accelerating to 10 MPH generate such snapshots. In our study, we assume the vehicle accelerate with a constant rate. However, it might not be eligible for the normal creeping situation. Therefore, for the VII application in urban street network, we suggest a real "Start" event snapshot with zero threshold speed to be made.

In the simulated testbed, the 375-meter-long link is longest. The maximum number of generated snapshots is 11. Since the congestion level is moderate, thus 15-20 maximum snapshots number in a 500 m length link can be reasonably assumed in the worse situation. Therefore, the optimal spacing between RSEs without losing any snapshot coverage should be 1KM given the maximum buffer size 30 snapshots.

4.3.5 Conclusion and Future Direction

This report presents one promising application of Vehicle-Infrastructure Integration (VII) which is to measure arterial performance in real-time. It utilizes the information collected through the V-I communication in concert with those collected by conventional point-based detectors. In this paper, the average travel time is chosen as the major measure of effectiveness (MOE) for arterial traffic conditions. A VII probe data (VPD) based model has been developed and customized for the latest VII probe message standard. In parallel, a point-based detection (PBD) based model with real-time inputs from inductive loop detectors and traffic signal controller and partial VII inputs has been developed. Lastly, a neural network based fusion model has been developed and trained by using the data generated by the simulation tool VISSIM. A six-intersection arterial testbed has been chosen to evaluate the developed models. According to the simulation results, the mean absolute percent error (MAPE) for the PBD model is 13.9%. While the VPD model performs better when probe vehicle penetration rate is 5%. The fusion model improves the results by 35% under the same penetration rate. With the increase of penetration rate, the estimation results for the VPD model and the fusion model can be further improved. In conclusion, this report shows that the developed analytical models work pretty well and are able to produce accurate and reliable estimations along the testbed arterial. Moreover, the report proves that the VII can be a powerful enabler for a wide range of ITS applications.

Although the simulation evaluation results show impressive results, there are many potentials for the three models to improve. For the PBD model, the over-saturated situation and other ILDs layouts might need further development. In the VII public data item list [FHWA, 2005], vehicular status information, weather and road condition information can help improve the proposed model. For example, the vehicle type information can help on the physical queue length and queue discharging time estimation. Brake applied status can help on queue length estimation, particularly the case when the queue extends over the advance loops. Some other information such as exterior temperature, rain sensor state, sun sensor state, wiper system state, fog lamps state, and headlights state can help determine lighting condition and rain, ice, snow, and fog conditions. Consequently, such inputs can help the proposed PBD model adjust some system parameters such as link free flow speed, saturation flow, and lost time.

For the VPD model, some simple assumptions can be future refined by using other data sources for example the signal status information. For the fusion model, the associated link trip time could be used as the fourth input. Although "black box" fusion approach such as neural network

has its inherent advantage without exploiting explicit mathematical relationship among its data sources, the drawbacks of such approaches are also obvious. The training data should be representative to the application's extent. The lengthy training process may not converge and may have to be performed for different sites and conditions. Since its nodes and weights do not have any physical meanings, it is very hard to utilize our existing knowledge to help setup those neural networks. Therefore a more transparent modeling approach is always preferred.

4.4 INTERSECTION SIGNAL CONTROLLERS AND VEHICLE-INFRASTRUCTURE INTEGRATION: PUTTING SAFE INTERSECTIONS TO PRACTICE

4.4.1 Introduction

Two of the VII California intersection RSEs, at Page Mill Rd and SR-82 and at California Ave and SR-82, use the radio and processor architecture in section 3.1, and use a special "sniffer" circuit board as described in section 4.4.2.3, to get signal phase and timing information from a cabinet equipped with a California Type 170 traffic signal controller. A third RSE, which was installed at 5th and SR-82 for use by CICAS-V researchers, uses the TechnoCom Multiband Configurable Networking Unit (MCNU) [TechnoCom, 2007], which includes both a processor and a newer generation DSRC radio. The MCNU, which will be used for future VII California RSEs, is designed to mount on a pole, which allows short cabling to the radio antenna. Caltrans has installed a Type 2070 traffic signal controller in the cabinet at 5th and SR-82 so that the AB3418 serial protocol (described in Section 4.4.3) may be used to communicate with the MCNU processor. Serial-to-fiber conversion is done in the cabinet, and the signal phase and timing information is delivered to a network interface on the MCNU through optical fiber in a conduit.

4.4.2 Signal Phase and Timing Acquisition

Traffic signal timing is a mature and sophisticated study in its own right. Many existing traffic controllers already run software that supports communication protocols, such as the National Transportation Communications for ITS Protocol (NTCIP) [NTCIP, 2007] and California's Assembly Bill 3418 (AB3418) protocol, designed to communicate signal phase and loop detector information to traffic management centers.

However, these protocols were designed for monitoring and control of actuated and coordinated intersection timings, from a central location over a modem, at a time granularity of seconds. They were not designed to notify vehicles of real-time signal phase count-down, needed for the signal violation countermeasures described in Section 4.4.1, over a high bandwidth DSRC connection. Traffic signal controller protocol messages also contain information that it is difficult to interpret without access to the timing plan for the intersection developed by the maintaining entity (state or local traffic operations), and do not contain geographical/mapping information needed by vehicles in order to recognize what part of the signal phase information applies to the vehicle's approach to the intersection. In addition, many intersections contain legacy equipment that do not support external communication.

Nevertheless, in our experiments we have successfully reconstructed sufficient information to broadcast the signal phase count-down needed for prototype vehicle applications. In different installations we have used the NTCIP protocol, the AB3418 protocol and a special "sniffer" circuit that non-invasively detects signal phase without any interaction with the traffic signal controller.

4.4.2.1 NTCIP

Our first implementation of signal phase and timing acquisition used NTCIP data objects supported on the Siemens Eagle software for the Type 2070 controller. The Phase Status Group Management Information Base (MIB) objects defined in were supported by the Eagle software, and code was written to access this information. In an experimental intersection at our facility (Richmond Field Station, University of California, Berkeley), we integrated serial communication with the 2070 controller into the computer/sensor system used for the research described in [Misener and Shladover, 2006; Zennaro and Misener, 2003; Chan, et al, 2005; Ragland, et al, 2006]. By reading the red, yellow or green phase status from the NTCIP message from the 2070, and keeping a count-down based on the timing plan in for the Richmond Field Station's (non-actuated) test intersection, we were able to display the signal phase in real-time in an in-vehicle diagnostic program, and correlate information taken on driver behavior with the signal timing. In conjunction with an automobile manufacturer, we developed and demonstrated an early traffic signal violation warning in one of their vehicles. At this test intersection we have also been able to conduct experiments evaluating communication latency (described in Section 4.4.5) since we were also able to use the AB3418 and "sniffer" methods. However, so far we have not been able to work with an intersection in the field where NTCIP was supported on a traffic signal controller approved for use by the maintaining entity.

4.4.2.2 California AB3418 Protocol

The Traffic Signal Control Program (TSCP) for the Type 2070 traffic signal controller (originally written by the Los Angeles Department of Transportation and further developed by Caltrans) supports signal phase and timing acquisition using the *GetLongStatus8* message of the AB3418 extended (or CTNET) message set in a serial protocol that at the level of packet encoding is similar to NTCIP. The fields of the single *GetLongStatus8* message make available much of the information that can be accessed through the PhaseStatusGroup and VehicleDetectorGroup MIB objects in NTCIP, albeit with less generality and redundancy. A single byte *active_phase* field can be decoded to get signal phase status for up to 8 phases; additional information from the *interval* field can be used to distinguish the "green" part of "active" from "yellow" and "red clearance." This signal phase status can be used, as for NTCIP, in conjunction with the timing plan, to construct a phase countdown. The *interval* field also includes encoding of additional information about the segment of the phase, such as "Min Green" or "Max Gap", that may be useful in the future for providing more information to vehicle safety applications concerned with signal state.

The AB3418 is also available on some specially configured Type 170 controllers, and thus in California may be the most readily available means to acquire signal phase information at a Caltrans-maintained intersection. However, most Type 170s are not configured for AB3418, and 2070s are also not yet widely available at Caltrans intersections.

4.4.2.3 Non-invasive Current Sniffer

A path of least resistance to deploying the equipment needed to bring dynamic signal state advisories to DSRC radio may be to retrofit existing or legacy controllers in a "clip-on" fashion. Hence, an electronic signal output to indicate the traffic light signal state in a non-contact fashion was developed. It operates at no interference with the normal signal operation by "sniffing" via

a commercially-available and highly sensitive current sensing coil clamped around the insulated cladding of active signal circuit wires, “red”, “green” or “yellow”. There is therefore no need to remove the wire contact from the active circuit in order to thread it through the coil. The selected device is shown in Figure 4.34 at installation at a VII California intersection (specifically, SR-82 and Page Mill Rd in Palo Alto) and has been shown to work with the very low current draw used by the modern LED traffic signal bulbs.

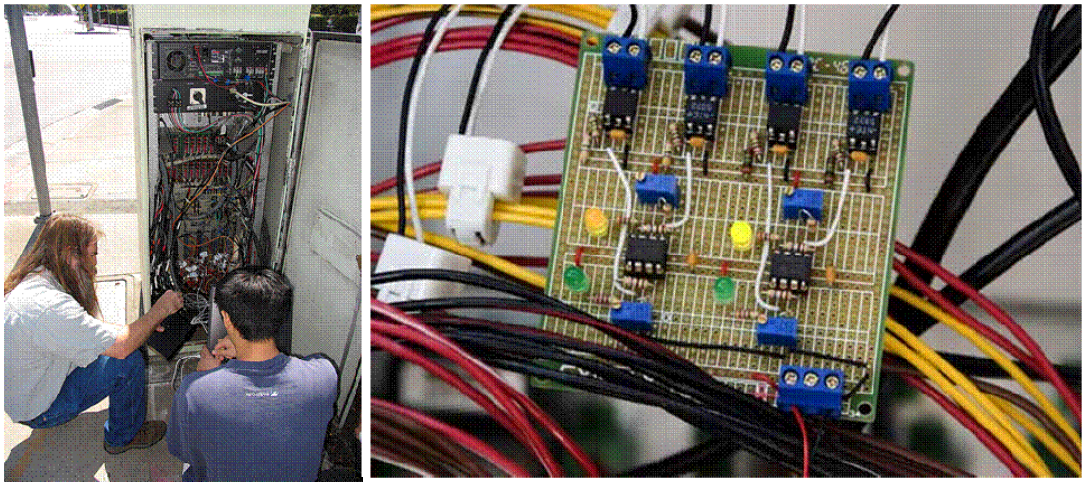


FIGURE 4.34 Installation and operation of “clamp-on” signal current sniffer at VII California intersection.

As shown in Figure 4.35, the basic sniffer design consists of the clamp-on coils, a full wave bridge rectifier to produce a dc voltage during signal bulb on state, and a comparator circuit. The comparator is a dual 8 pin dip package. This allows for inverting or noninverting outputs, and the open collector outputs are used to translate the output signal voltage to TTL levels for computer digital inputs or other voltage levels to trigger different devices. The comparator threshold is adjustable to allow for testing flexibility on a variety of intersection types. The output is converted to digital format, then interfaced to external devices USB port. Depending on how many signal wires are used and on the complexity of the intersection, the information sent may be just “green” information, that will be compared to the intersection’s timing plan to determine red and yellow countdown as well, or redundant information for the red and yellow phases may also be sent, which would theoretically allow software reading the information to “learn” the intersection’s timing plan.

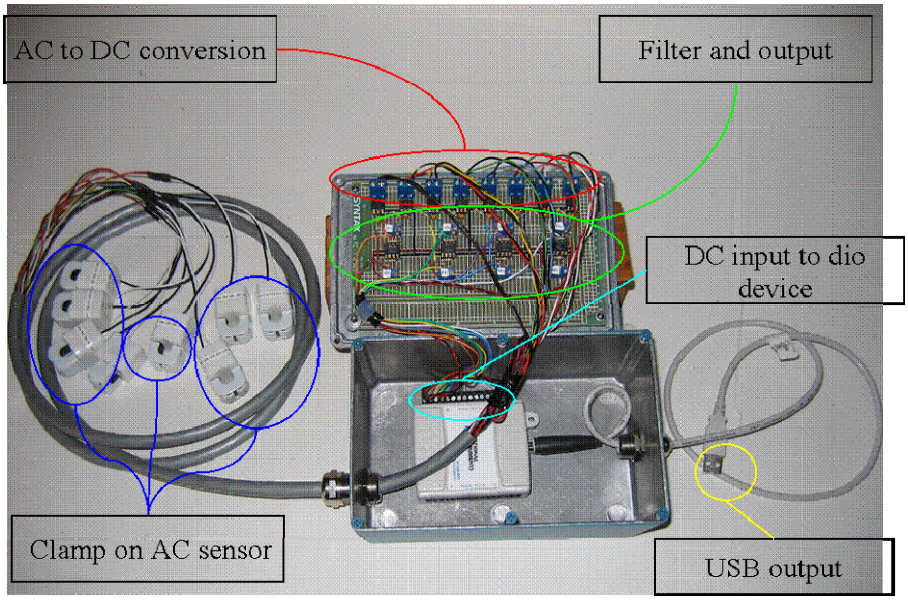


FIGURE 4.35 Depiction of "clamp-on" signal sniffer.

4.4.3 Interface to Dedicated Short Range Communication

All three methods of acquiring signal phase timing information described in Section 4.4.2 can be integrated with the same phase count-down and broadcast software. It is desirable to have well-specified interfaces between the signal phase acquisition, the phase countdown and the broadcast in order to isolate changes required in different installations. The modular system design and logical connections shown in Figure 4.36 have been used at all three of our VII California testbed intersections, two of which use the sniffer and one of which uses AB3418, as well as at the experimental intersection, where we have tested all three methods. In each case the most complicated part of the software is the phase countdown itself, where we are doing in effect a simplified emulation of the internal logic of the traffic signal controller in order to duplicate the countdown that the traffic signal controller is itself carrying out in order to set the signal phases.

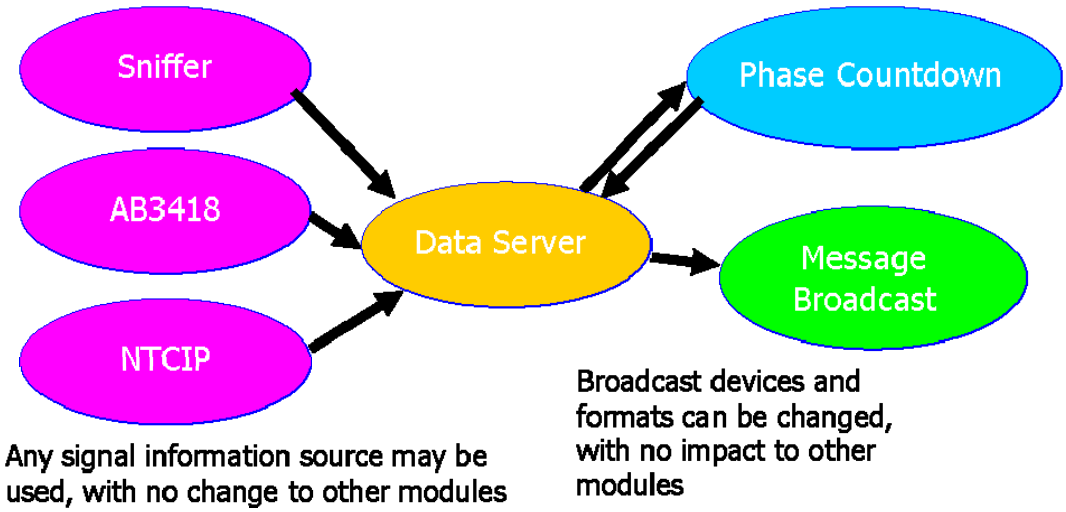


FIGURE 4.36 Modular design of signal phase acquisition and broadcast system

In order to do this, even in simplified form, we require configuration files for each intersection that contain much of the information that has been programmed by the maintaining entity into the traffic signal controller as the intersection's timing plan, as well as a specification of how the signal phase information is being acquired and delivered, and the relationship between the internal phase designations that are meaningful to the traffic signal controller and the geographic description of each approach to the intersection that is meaningful to the approaching vehicle.

We currently use three types of configuration files:

- A file that specifies what signal wires the sniffer is connected to, in terms of signal phase, (numbered as in the traffic signal controller logic), and type of light (red, yellow, green). This is not needed with AB3418 or NTCIP.
- A file that specifies the timing plan, giving the interval lengths assigned to each phase, and whether the system is fixed or actuated and/or coordinated.
- A file that gives the correspondence between each approach, defined by geographical information and used as the basic information element in the broadcast message, and the internal traffic signal phase that controls that approach. The geographical information specified in this file becomes part of the broadcast message and is used by the in-vehicle application to determine which signal phase and timing information is relevant.

Figure 4.37 shows diagrams of the phase plans of two of our intersections as an illustration of the different characteristics that need to be captured in the configuration files. Enough geographical information must be included, in the form of latitude and longitude information describing the lane leading up to a signal, for a vehicle to be able to distinguish what approach it is on to the signal and determine the relevant phase countdown.

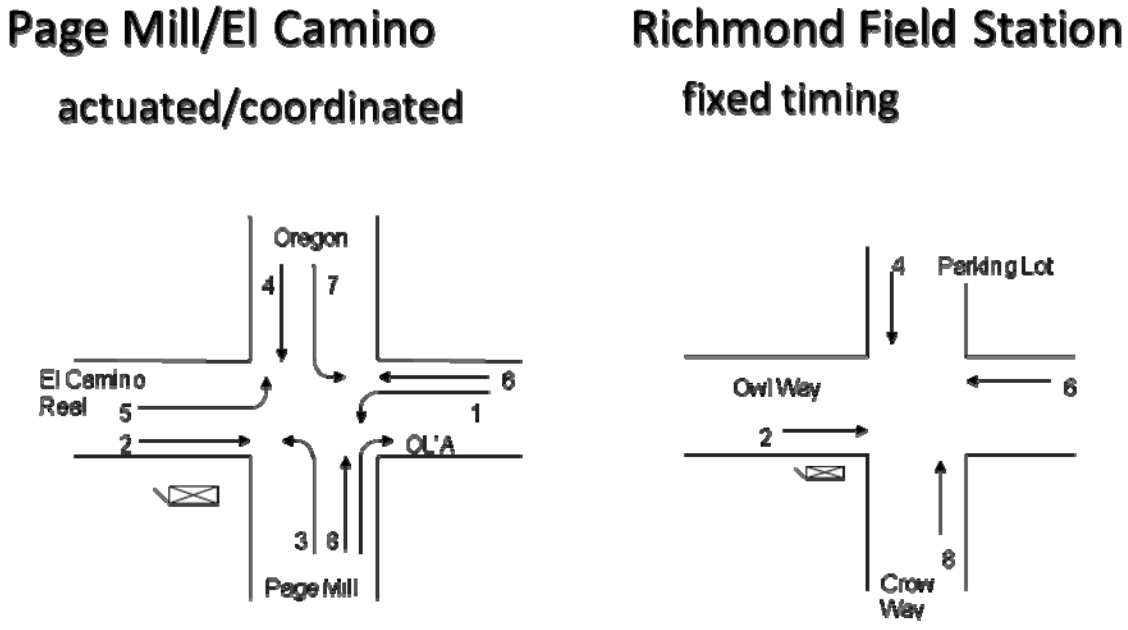


FIGURE 4.37 Types of information that must be specified to configure signal phase and timing broadcast.

4.4.4 Communications Latency for Signal Phase and Timing Information

A serious question about delivering signal phase and timing information to vehicles is whether it can be delivered with low enough latency. Fortunately, our experiments show that it is not difficult to deliver the phase transition information in well less than 100 ms from the time the transition occurred using current technology.

To measure this, since we did not have access to the internal delays and logic of the traffic signal controllers, we first measured the delays on our sniffer circuit with an oscilloscope. These delays are tuneable, by changing the triggering points of the comparator circuit, and we adjusted them until the delay for setting the digital output corresponding to the "on" transition of the green signal was reliably under a millisecond. We then measured the sample time of the digital I/O subsystem and found that it was reliably about 2 millisecond. Since the rise time for the signal driving the light LED was on this order, we then used the transition time for the on signal with the sniffer circuit as our ground truth. Note that this transition time will be later than the time when the activation signal leaves the traffic signal controller to begin to drive the load switches to turn on the light, but within a few milliseconds of the time when the light can be expected to actually change visually in response to the change in the AC signal.

For the AB3418 and NTCIP protocols, which are both query/response protocols, we adjusted the time between queries to be the minimum that the system would tolerate. For AB3418, this was a cycle of 50-60 milliseconds for the entire query response cycle; for NTCIP, this was a cycle of 75 milliseconds required between requests in order to avoid a fault.

The results we obtained were somewhat surprising, in that we had expected the AB3418 protocol to typically report a transition time later than the sniffer, but this was not the case, see Figure 4.38. We conjecture that the internal state machine in the traffic signal controller using AB3418 is setting up the values for the AB3418 message ahead of actually driving the signal, and this allows the AB3418 time to be as early as the sniffer's on occasion, despite it 50-60 millisecond query/response time. For the NTCIP implementation we have tested, on the other hand, the delay was much greater, over half a second. Further investigation is required to determine if there is an error in our NTCIP system set-up, or if it was not considered necessary to update the information for the NTCIP protocol at a fine time granularity.

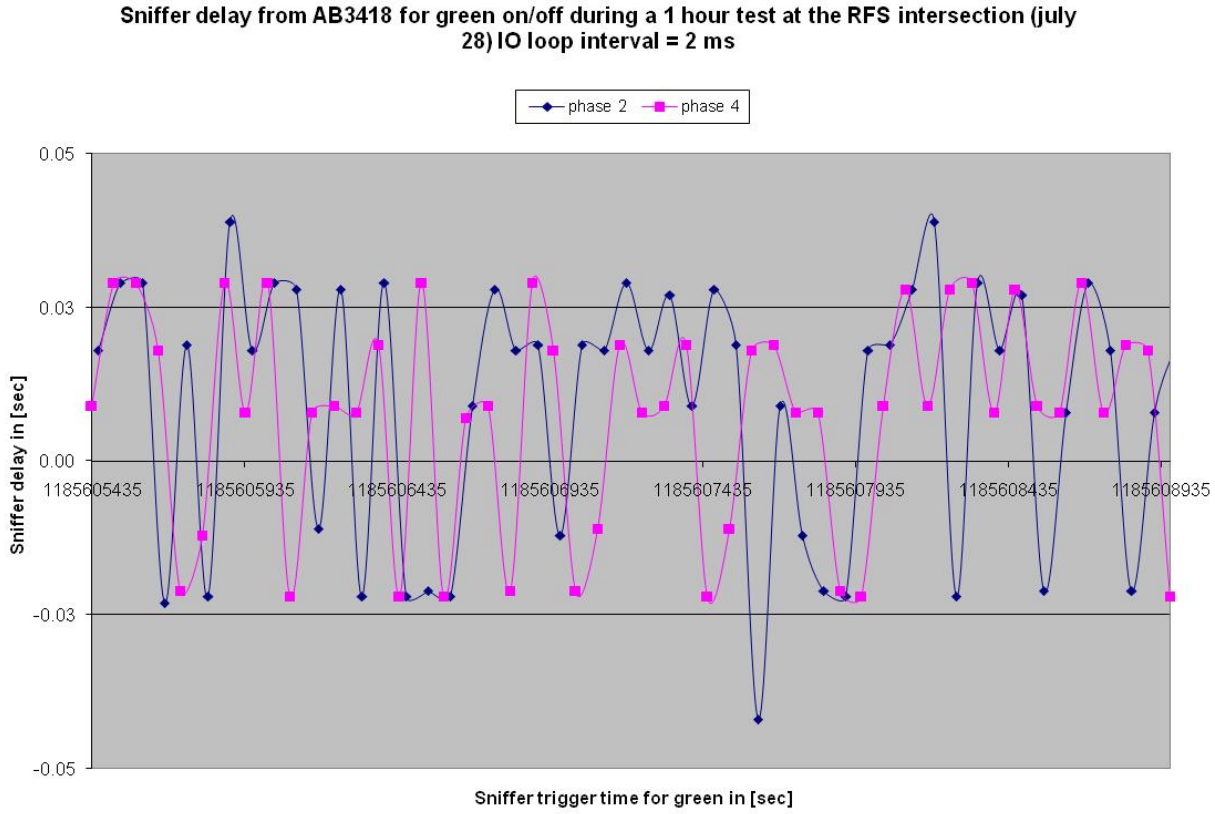


FIGURE 4.38 Difference in transition time between AB3418 and sniffer acquisition, run simultaneously

4.4.5 Communications Latency for Signal Phase and Timing Information

In summary, our efforts to obtain and broadcast up-to-date traffic signal phase and timing information using current traffic signal controllers, existing communications standards and moderately priced hardware have been successful. It is not difficult to reformat the information received from either the sniffer or AB3418 into the object structure defined by NTCIP, as we have done, and use that as the basic interface to the signal phase information. In the future it may be desirable to make the phase countdown program more sophisticated and exact for actuated intersections by including loop detector information. This also can be formatted into NTCIP data object definitions before passing on to the phase countdown module.

If exact count-down information is required for all phases, at some point the phase countdown logic would represent a complete duplication of the traffic signal controller logic. It might be better if the information could be sent directly from the controller. Indeed, this has already been done in; signal phase and timing information was sent out as an Ethernet broadcast message and forwarded to the Denso WAVE Radio Module as part the Innovative Mobility Showcase at the 2005 World Congress [Provenzano, 2005]. This seems the simplest solution when installing a new intersection, but given the large installed base of existing traffic signal controllers without communication capabilities, add-on devices to do signal phase broadcast and countdown may be the only alternative in many installations for some time. In addition, it may turn out after further

experimentation that less detailed information, providing exact countdowns only in certain situations of most importance to vehicle safety applications, may be sufficient.

Implications for the future where the intersection crash problem may be addressed through dynamic red light running countermeasures and other VII-enabled systems that "talk" (in a data sense) between the intersection and car are multi-fold. Clearly, institutional and related capital and operations costs not described in this report play a part. Additionally, human factors issues ranging from usability to specific design are important and also not described in this paper. However, what is clearly important and covered in this research are the software and hardware implementation and interface issues. Some of those issues – how to adapt to legacy systems, either in an analog fashion (via our "sniffer") or through software and existing signal controller communications protocols – may actually help circumvent some of the uncovered issues. Our research shows that indeed there can be deployment path to retrofit the heterogeneous and still "ain't broke" legacy systems and given its reliability, we may be able to deliver to the driver reliable and low latency VII- and DSRC-enabled signal timing information that addresses the purpose of this research, to save lives at intersections.

4.5 HIGH ACCURACY-NATIONWIDE DIFFERENTIAL GLOBAL POSITION SYSTEM (HA-NDGPS)

4.5.1 Introduction

The California PATH program and Caltrans have a number of research applications that can benefit from a highly accurate vehicle positioning implementation. The U.S. Department of Transportation and U.S. Coast Guard in conjunction with the Interagency GPS Executive Board are currently developing a High Accuracy-Nationwide Differential Global Positioning System (HA-NDGPS), targeted at a variety of transportation-related applications. The HA-NDGPS program provides the capability to broadcast corrections to Global Positioning System (GPS) receivers over long ranges to achieve a better than 10 centimeter (cm) (95 percent of the time) accuracy throughout the coverage area. HA-NDGPS is currently undergoing a research and development phase with a future implementation pending U.S. congressional budget approvals. The signal is available for test purposes from Hagerstown, MD, and Hawk Run, PA. Implementation of this technology will assist in providing advanced safety features for transportation, including lane departure warning, intersection collision warnings, and railroad track defect alerts. It also could be used for economic enhancements such as precision container tracking and automated highway lane striping.

In addition to the evaluation of HA-NDGPS for California Vehicle Infrastructure Initiative (VII) applications, several other differential GPS architectures have been reviewed within the context of this study. The scope of the study was expanded in response to concerns associated with funding shortfalls associated with the HA-NDPS implementation. While the original goals of the study have been maintained relative to evaluating HA-NDGPS for VII California implementations, additional DGPS architectures such as WAAS (Federal Aviation Administration Wide Area Augmentation System) and Real Time Kinematic (RTK) carrier-phase solutions have been comparatively evaluated.

In order to better understand its capabilities and how it can benefit the California PATH and Caltrans research program and in particular the VII California Program, a five-month evaluation has taken place consisting of four tasks: 1) Review of HA-NDGPS and differential GPS architectures; 2) Development of the HA-NDGPS Concept for VII Operations; 3) Development of HA-NDGPS and differential GPS requirements for VII California; and 4) Recommendations.

These tasks are presented in the following chapters:

- **Background** – details the development of the GPS system, differential GPS methodologies, and general operational characteristics;
- **Differential GPS System Evaluation** – evaluation summaries and performance review of HA-NDGPS, WAAS, RTK and other VII appropriate systems;
- **VII Proposed System Architectures** – proposed architectures and methodologies for VII California implementations;
- **Recommendations** – recommended solutions relative to VII California positioning requirements ; and,
- **Conclusions & Future Work** – summary of results, recommendations, and next steps.

4.5.2 Background

4.5.2.1 Vehicle-Infrastructure Integration (VII) Program

The VII initiative centers around establishing communications between individual vehicles as well as between vehicles and the infrastructure in order to enable a variety of systems to be developed that significantly improve roadway safety, operations, and maintenance. VII can also serve a variety of applications to support the needs of vehicle manufacturers and other private interests [USDOT, 2005]. The concept of VII is not new, the general idea has been around since the inception of ITS, where a variety of applications such as intersection collision avoidance and automated highway systems could best be accomplished through sophisticated vehicle-infrastructure and vehicle-vehicle communication [ITSA, 2005]. In the last two years, there has been significant interest in this communications capability and a cooperative effort has evolved involving State Departments of Transportation (DOT) through the American Association of State and Highway Transportation Officials (AASHTO), local government agencies, vehicle manufacturers, and the USDOT [USDOT, 2005] in the establishment of the national VII initiative [USDOT, 2006].

California has quickly become one of the leaders in embracing VII, with the formation of the VII California program [CAATS, 2006]. Several workshops and demos have been held including the recent ITS World Congress held in San Francisco in November 2005. At this demo, several different VII applications were demonstrated including using vehicles as traffic probes, transmitting travel times and incident data to vehicles, and the establishment of intelligent intersections. As part of this ITS World Congress demo, California has established a VII testbed in the San Francisco Bay area that will be utilized for a variety of experiments. Further, another VII testbed is currently being established in the San Diego area. It is clear that California is becoming one of the leaders in VII efforts. With the integration of vehicle and roadside communications, the determination of vehicle location has become a critical component of several VII applications. Global Positioning System (GPS) technology has been the dominated means for vehicle positioning for over a decade and will likely play an important role for many VII applications.

4.5.2.2 Global Positioning System

The Global Positioning System (GPS) is one of the most convenient and accurate methods for determining a vehicle's position within a global coordinate system [Farrell & Barth, 1999; Farrell & Givargis, 2000]. Utilization of a GPS receiver has become the most prominent method of determining the physical location of a vehicle. A detailed and concise description of GPS is provided in [Farrell & Barth, 1999]. The system is built around a set of 24 satellites (with additional, spares, replacements, upgrades) that orbit the Earth. The orbits are designed in a manner that allows the signals from at least four satellites to be received simultaneously at any point on the surface of the Earth (neglecting obstacles such as tunnels, urban canyons, dense forest etc.). A GPS receiver on the surface of the Earth utilizes the signals from at least four satellites to determine its own antenna position (x, y, z) according to various measurements of the pseudoranges between the satellites and the receiver antenna. The receiver measurements include various code-based pseudoranges and potentially carrier phase information. The standard deviation of uncorrected GPS position estimates is on the order of 10-20 meters. Increased accuracy better than 3 meters can be achieved through the use of code-range based differential

GPS (DGPS) techniques, such as those described in [Navstar, 1991; Kee, 1994; Blomenhofer, 1994; and Farrell & Barth, 1999].

Satellite transmission of orbital position (ephemeris), timing, and satellite status data to GPS receivers is known as 'code,' and measurements of the distance between the satellite and the receiver using this 'code' information are called code-based measurements. The GPS receivers utilize the ephemeris data to calculate the positions of each satellite. With the position of a satellite known, the distance between the receiver and the satellite (called the pseudorange) can be determined using the propagation delay of the signal. When the distance between a receiver and at least four distinct satellites is known, a position lock can be obtained [Bajikar et al., 1997]. Accurate calculation of the distance from the satellite to the receiver requires precise determination of the propagation delay. Each GPS satellite utilizes four atomic clocks whose precise average time is transmitted for purposes of time synchronization. The satellites also transmit distinct pseudo-random time-dependent binary signals that are compared with identical internally generated signals within GPS receivers. When the receiver correlates the satellite signal with its own internally generated pseudo-random signal, the receiver differentiates between when the satellite sent the signal and the time of reception to calculate the distance between itself and the satellite [Love, 1994].

Each GPS receiver attempting to calculate an accurate position must possess a clock that is correctly synchronized with the atomic satellite clocks. Since the cost of atomic clocks is prohibitive for integration in individual GPS receivers, alternative synchronization methods are utilized. The clocks of the GPS receivers are synchronized precisely with the satellite clocks by measuring propagation delays of four individual satellites. Typically, the GPS receiver's clock is still slightly offset after the process of synchronization (creating a very small clock error). This receiver clock offset is thus introduced as a fourth variable in the pseudorange equations and four satellites are typically utilized to solve for receiver position. When receiver clock offset is assumed to be reasonably small due to very accurate time synchronization, then precise three-dimensional position determination can be achieved.

Most low-cost GPS receivers are single frequency L1 (1575.42 MHz) C/A code sensors, which measure the distances between a receiver and satellites and estimate the receiver's antenna position based on these pseudorange measurements. Given the range standard deviation of common-mode noise on the order of 10 – 20 meters [Farrell & Barth, 1999] and non-common-mode noise at order of 0.1 to 4.0 meters [Farrell & Barth, 1999; Farrell & Givargis, 2000], a standalone GPS receiver has a horizontal standard deviation of position error around 20 meters. This accuracy is sufficient for the road-network-level navigation of route finding/planning and guidance, which usually positions the vehicle at street level with the assistance of map matching algorithms. Most of the current existing vehicle navigation systems utilize the GPS receivers that are in this category.

In terms of vehicle spatial positioning, a standard GPS receiver with the accuracy of 20 meters coupled with map matching has performed well over years in the route guiding navigation systems. More accurate vehicle navigation, such as differentiating the vehicle's lane, requires higher accuracy positioning capability. Positioning systems of centimeter-level accuracy for vehicle operational control, e.g., GPS coupled with Inertial Navigation Systems (INS), have been

tested over years and demonstrated sufficient performance in the control applications. However, they are currently too expensive for general navigation or lane-level VII applications. Improved accuracy can be obtained by using *differential* GPS techniques to remove the common-mode error from the measurements, making the standard deviation of the error small enough to satisfy the lane-level accuracy requirements. Increasing accuracy beyond the lane level requires additional corrections such as carrier phase corrections and/or dual channel corrections.

It is apparent that a low-cost positioning system capable of discriminating between lanes is needed for lane-based VII applications, requiring approximately 1 to 2 meter level accuracy. A DGPS system that uses carrier-phase based range observations can provide accuracy up to 1-3 centimeters, however these receivers require dual frequency reception and additional algorithms to solve integer ambiguity issues, resulting in significantly higher cost (see, e.g., [Navstar, 1991; Blomenhofer, 1994; Farrell & Givargis, 2000]). Vehicle applications requiring lane-level accuracy (2 meters) can potentially utilize a low-cost GPS receiver coupled with other on-board electronics and firmware to increase the accuracy. These lower cost configurations require the use of differential corrections originating from a base station in relatively close proximity.

4.5.2.3 Differential GPS and Carrier-Phase GPS

Several sources of error contribute to calculation inaccuracies of the pseudoranges, including satellite clock bias, atmospheric delay, ephemeris prediction data, and receiver tracking error noise [Bajikar et al., 1997]. These errors are frequently referred to as the 'common mode errors' because they are common to all receivers in a local (< 50 km) area. Because these errors are common to all receivers in a local area, they can be eliminated through differential GPS (DGPS). Utilizing a stationary base station receiver whose position has been accurately surveyed, common mode errors can be determined for the local area surrounding the base station. The methodology of transmitting these local errors to nearby mobile receivers (rovers) forms the basis of a differential GPS architecture. A breakdown of GPS and DGPS errors are provided in Table 4.2.

There are three primary types of DGPS position corrections that can be provided by the base station: code (or range-space), carrier-phase, and Doppler corrections. Within the context of this study, code and carrier-phase corrections are described. Carrier-phase corrections, which typically produce centimeter level accuracy, are traditionally achieved with post processing algorithms while code-based DGPS corrections are frequently utilized for real-time applications, resulting in 2-3 meter accuracy. However, it is possible to also transmit carrier-phase correction in real time, resulting in what is called Real Time Kinematic (RTK) positioning.

For code-based DGPS, the base station receives the satellite signals and calculates the common mode errors from each satellite, using the known accurate position of the base station. Once the common mode errors for a given time instant are known, they can be transmitted to all receivers within a local area. The rover receivers subtract these transmitted common mode errors for the satellites in its view to calculate its position, resulting in a typical accuracy of 2-3 meters. The accuracy is also influenced by the age of the correction, which is the time between the base station's calculation of the errors and the receiver's calculation of its position. To maintain the accuracy integrity, correction updates are typically needed at least every 10-15 seconds [FHWA, 2005; Barth et al., 2005].

The carrier frequency is used for transmitting the code-based signal from the satellite to the receiver. Carrier-phase DGPS functions through phase measurements of the carrier signal, detecting the change in the number of carrier cycles between the receiver and a given satellite. The total number of carrier cycles between the receiver and a given satellite multiplied by the carrier wavelength can provide a more accurate measurement of the range. An 'integer phase ambiguity' arises when the total number of carrier cycles is not known; only the change in this total number of cycles can be directly observed. If the code measurement can be made accurate to within a few meters, then there is only a few wavelengths of the carrier signal to consider to determine which cycle really marks the edge of the carrier frequency timing pulse. Resolving this carrier phase ambiguity for just a few cycles is a much more solvable problem and with increasing processing power and functionality, it's possible to make this kind of measurement in real time. In this method, the common mode errors are identical to the common mode errors encountered in the code-based measurements, but the non-common mode errors are approximately 100 times smaller than their corresponding errors in the code-based measurements. Eliminating the common mode errors with code and phase corrections, it is then possible to achieve centimeter-level accuracy using carrier-phase DGPS.

Table 4.3 Summary of GPS Error Sources. [Trimble, 2006].

Typical Error in Meters (per satellite)	**Standard GPS**	**Code-Phase Differential GPS**
Satellite Clocks	1.5	0
Orbit Errors	2.5	0
Ionosphere	5.0	0.4
Troposphere	0.5	0.2
Receiver Noise	0.3	0.3
Multipath	0.6	0.6

The process of transmitting carrier-phase corrections in real time from a localized based station to a rover receiver capable of processing carrier-phase corrections is commonly referred to as Real Time Kinematics (RTK). The RTK methodology is commonly utilized by the survey industry to achieve accurate geographic positions within a localized area. While the traditional survey techniques have typically required the rover receiver to be stationary for several minutes to achieve centimeter level accuracy, improved methods are allowing for real time, second-by-second, carrier-phase corrections. Table 2 summarizes the correction techniques and their accuracies.

4.5.2.3.1. Real-Time Differential Corrections

Several industry standard DGPS correction signals have been implemented for various differential services, including RTK. Differential correction formats were created for marine, aviation, navigation, and proprietary systems. The most frequently utilized correction streams that are non-proprietary are termed:

- *RTCM-104, and*
- *RTCA-159 (WAAS)*

Table 4.4 Summary of GPS signal correction methods. [Magellan, 2006].

Measurement Type	Real-time or Post-processing	System Type	Accuracy	Coverage Area
Code	Post-processing	Post-processed DGPS, post-processed LADGPS or post-processed WADGPS	from 1 m to ~10 m	Nationwide
Code	Real-time	DGPS, LADGPS, or WADGPS	from ~2 m to ~10 m	Nationwide
Carrier phase	Post-processing	Kinematic, rapid static or static	from < 1 cm to several cm	From several km to several 1000 km
Carrier phase	Real time	Real-Time Kinematic (RTK)	from < 1 cm to several cm	From several km to several 10 km

Many GPS receivers are "RTCM-capable" meaning they accept DGPS correction messages through a real-time data communication link (e.g., VHF or UHF radio, cellular telephone, FM radio sub-carrier, or satellite com link). The Radio Technical Commission for Maritime Services Special Committee 104 recommended standards for DGPS correction messages, which have become termed RTCM-104. Several versions exist which include the recent addition of message types 18/19/20/21 for carrier phase solutions such as RTK. For RTK applications, RTCM version 2.1 provides Type 18 (carrier phase raw data) or Type 20 (carrier phase corrections).
Radio Technology Committee for Aviation Special Committee 159 developed minimum standards that define the basis for FAA approval of equipment using GPS for aircraft navigation in the United States. The RTCA DO-229 document entitled "Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment" was prepared by Special Committee 159 in 1996. It contains the standards for airborne navigation equipment using GPS augmented by a Wide Area Augmentation System (WAAS).

As described previously, several proprietary differential data streams exist including systems such as Starfire, and Omnistar. While these systems may be adaptable for VII applications, it is presumed that an open architecture method is preferred for achieving the accuracy requirements desired. Similarly, numerous proprietary survey grade (<5cm) DGPS systems exist which could be modified to meet VII requirements, but have also received limited focus due to proprietary correction signals that would become an issue for VII implementation. While the GPS industry has integrated WAAS and RTCM (RTK capability) corrections into many receivers, there is a third DGPS correction architecture evaluated in this study, referred to the National Differential GPS (NDGPS) architecture. Greater detail is provided on WAAS, RTK, and NDGPS correction methodologies below.

4.5.2.3.2. Wide Area Augmentation System (WAAS)

The Federal Aviation Administration (FAA) developed a Wide Area Augmentation System to transmit satellite based DGPS correction signals for civil aviation. WAAS utilizes a network of

precisely-located ground reference stations that monitor GPS satellite signals. These stations are located throughout the continental United States, Hawaii, and Alaska, with additional stations being installed throughout North America. These stations collect and process GPS information and send this information to WAAS Ground Uplink Stations (GUS). The WAAS ground uplink stations develop a WAAS correction message that is sent to user receivers on the L1 transmission signal via geostationary satellites. The most recent GUS installed in California is located in Napa Valley and is shown in Figure 4.39.

Figure 4.39 Recently installed WAAS Ground Uplink Facility in Napa California [SatNav,2005].

Using WAAS, GPS signal accuracy is improved from 20 meters to approximately 3 meters in both the horizontal and vertical dimensions.

WAAS reached initial operational capability for aviation use on July 10, 2003 as the first of several Space-Based Augmentation Systems (SBAS) being developed throughout the world and is compatible with all other international satellite-based augmentation systems. Although the WAAS was designed for aviation users, it supports a wide variety of non-aviation uses including agriculture, surveying, recreation, and surface transportation. The WAAS signal has been available for non safety-of-life applications since August 24, 2000, and numerous manufacturers have developed WAAS-enabled GPS receivers for the consumer and OEM market.

The next phase of WAAS is referred to as the Global Navigation Satellite System Landing System (GLS) segment. The GLS phase of WAAS is scheduled to coincide with the operational capability of GPS modernization and is scheduled to be completed in 2013. GLS will utilize, and depend upon, improvements that the Department of Defense (DoD) will make as part of its GPS modernization program. GPS modernization will improve WAAS performance during periods of severe solar storm activity and provide additional security against interference to the GPS.

4.5.2.3.3 Real Time Kinematic (RTK) and CORS

RTK is a process where carrier-phase GPS signal corrections are transmitted in real time from a reference receiver at a known location to one or more remote rover receivers. The use of an RTK

capable GPS system can compensate for atmospheric delay, orbital errors and other variables in GPS geometry, increasing positioning accuracy up to and sometimes within a centimeter. Used by engineers, topographers, surveyors and other professionals, RTK is a technique employed in applications where high precision is necessary. RTK is used, not only as a precision positioning instrument, but also as a core for navigation systems or automatic machine guidance, in applications such as agriculture, civil engineering, and dredging. It provides advantages over other traditional positioning and tracking methods, increasing productivity and accuracy. Using the code and carrier phase of the GPS signals, RTK provides differential corrections to produce the most precise GPS positioning.

The RTK process begins with a preliminary integer ambiguity resolution. This is a crucial aspect of any kinematic system, particularly in real-time where the velocity of a rover receiver should not degrade either the achievable performance or the system's overall reliability. The base station is responsible for assembling the base station carrier phase correction message, which aids the receiver in solving the integer ambiguity. With integer ambiguity solved, the receiver is capable of centimeter level accuracy at frequencies of 1 Hz.

With increased prevalence of centimeter level applications, RTK correction messages are being integrated within a network of reference stations. The National Geodetic Survey (NGS), an office of the National Oceanic and Atmospheric Administration's (NOAA) National Ocean Service, coordinates two networks of Continuously Operating Reference Stations (CORS):

- the National CORS network; and,
- the Cooperative CORS network.

The national CORS network is operated, managed, and maintained through NGS agreements with site data available through NGS; while the Cooperative CORS is a network of independent equipment operators each responsible for their own data collection, storage, and transmission. Each CORS site provides GPS carrier phase and code range measurements in support of 3-dimensional positioning activities throughout the United States. Currently, only a small portion of the sites provides real time (1 Hz) correction data capable of supporting RTK applications.

The CORS system benefits from a multi-purpose cooperative structure involving many governments, academic, commercial and private organizations. New sites are evaluated for inclusion according to established specifications and criteria. Cooperative CORS data are available from the participating organization that operates the respective site. Specific to California, an additional collaborative effort has evolved to operate and maintain reference stations with real time correction capability.

The California Spatial Reference Center (CSRC) is operated through the Cecil H. and Ida M. Green Institute of Geophysics and Planetary Physics at UCSD's Scripps Institution of Oceanography. In partnership with surveyors, engineers, GIS professionals, the National Geodetic Survey, the California Department of Transportation, and the geophysics community, CSRC has developed a plan to establish and maintain a network of GPS control stations

necessary to meet the demands of government and private businesses for a reliable spatial reference system in California.

CSRC has had significant influence on the implementation, development, and operation of reference stations in and around Southern California. The implementation of these sites has predominantly been associated with geophysics (earth crust movements) and municipal survey requirements. The real time data requirements for these applications are compatible with RTK data requirements.

4.5.2.3.4 United States Coast Guard National Differential GPS System

Another system for providing differential corrections is the U.S. Coast Guard's Nation-wide Differential Global Positioning System (NDGPS). This service is a land-based differential correction system that receives and processes signals from orbiting GPS satellites, calculates corrections from known positions, and broadcasts these corrections via a Medium Frequency (307 kHz) beacon transmitter to DGPS users in the broadcast site's coverage area [Wolfe et al., 2000].

Figure 4.40 USCG NDGPS broadcast site [FHWA, 2005].

As the DGPS concept was being developed in the late 1980's, a variety of alternate methods to enhance the accuracy and integrity of GPS through differential correction were considered. The motivation that guided development of the USCG NDGPS service were redeployment of decommissioned radio beacon sites distributed throughout the United States, availability of radio beacon frequencies (285-325kHz), and the need for harbor navigational aides [USDOT, 2001]. NDGPS sites were engineered to broadcast applicable pseudorange corrections at 200 bits per second or less up to a range of 450 km. At the time, the selective availability (S/A) dithering error was the greatest error to correct, dwarfing other errors such as atmospheric, multipath, algorithmic, and noise errors. Correcting for pseudorange errors allowed DGPS receivers to meet the 10 meter accuracy requirement for positioning aids to navigation and for the 8-20 meter accuracy requirement for harbor and harbor approach (H/HA) [USDOT, 2001].

Selective Availability was eliminated for GPS broadcasts on May 2, 2000, in accordance with a Presidential Decision [USPDD, 1999]. With GPS unencumbered by S/A, GPS receiver accuracy achieved dramatic improvements. DGPS receivers also benefited by the removal of S/A by using algorithms to correct remaining GPS errors, achieving accuracies of 1-3 meters [USDOT, 2001]. Atmospheric errors from the effects of the ionosphere, as well as the wet and dry portions of the troposphere, are now the largest contributor of error to the GPS signal. While DGPS pseudorange corrections effectively minimize these errors at the DGPS reference location, spatial decorrelation degrades these corrections as the distance between the receiver and reference station increases. Modeling techniques using real time monitoring show significant potential to achieve sub-meter accuracies by supplementing the pseudorange corrections with improved atmospheric models [Gutman et al., 1999].

The NDGPS system is in the process of finalizing a nationwide network of DGPS broadcast sites to provide dual terrestrial coverage in the continental U.S. Spatial decorrelation effects are minimized by the utilization of these multiple reference stations. New generation DGPS receivers take advantage of this dense network to compare the signals of multiple broadcast sites to enhance positional accuracy. Data from multiple locations can be transferred to other locations to rebroadcast or be combined at a central location and then rebroadcasted [Last et al., 2002].

When the required coverage is met, DGPS users in the U.S. will possess signal availability of better than the required 99.9% (availability represents the percentage of time the DGPS signal is usable). The NDGPS system provides users with broadcast messages as defined by the Radio Technical Commission for Maritime Services (RTCM) and utilizes Reference Station Integrity Monitor (RSIM) messages for intra-system communication [RTCM, 2001a; RTCM, 2001b].

Figure 4.41 illustrates the planned and operating NDGPS sites in the U.S. Figure 4.42 shows the estimated coverage area for installed NDGPS sites as proposed in September 2002. Single coverage regions (gray areas) and regions where more than one signal is available (yellow areas) were identified using Millington's method for determining ground wave signal strength [Wolfe et al., 2003]. Table 4.5 gives the current status of California's NDGPS broadcast stations.

Table 4.5 Status of California NDGPS broadcast stations [USCG NAVCEN, 2006].

Site ID	Location	Status
899	Petaluma	Engineering Test Site
875	Essex	Planned Const - CY06
882	Lompoc	OPERATIONAL
764	Lincoln	OPERATIONAL
881	Point Loma	OPERATIONAL
878	Chico	OPERATIONAL
795	Bakersfield	OPERATIONAL
885	Cape Mendocino	OPERATIONAL
883	Pigeon Point	OPERATIONAL

Figure 4.41 Planned and Operating NDGPS sites [Wolfe et al., 2003].

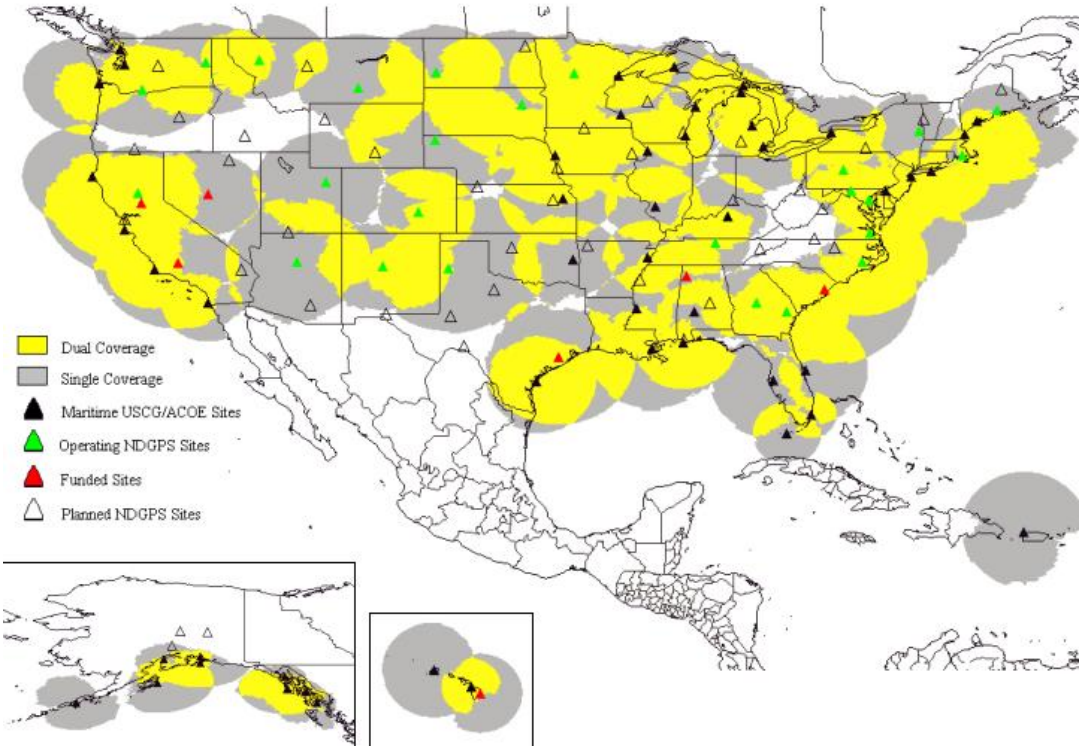


Figure 4.42 Projected NDGPS coverage [Wolfe et al., 2003].

The U.S. Coast Guard developed the original NDGPS network. The U.S. Army Corps of Engineers (USACOE) needed similar capability along inland waterways and, with the help of the USCG, established similar broadcast stations. USACOE later adopted the NDGPS concept and developed additional stations. The US Department of Transportation (USDOT) continues to install NDGPS stations at retired Air Force Ground Wave Emergency Network (GWEN) sites. When all proposed installations are complete, there will be approximately 137 similar stations broadcasting Course/Acquisition (C/A) code correctors to users. This will enable meter-level positioning with an associated level of integrity, according to the Federal Highway Administration [FHWA, 2005].

These beacons today broadcast single-frequency GPS code range correctors enabling few-meter level positioning and navigation along the U.S. coasts and inland waterways. An expected pseudorange value is computed for each GPS satellite in view at the NDGPS sites. These computed values are compared with the actual pseudorange measurements made at the sites; the difference is known as a pseudorange corrector. The assumption is that the errors are common to both the reference site and any user site. These correctors are placed in a formal bit stream with message header, message type, and with parity considerations and broadcast to users as a Type 9 RTCM message. This message is one of many broadcast as part of the RTCM-104 protocol developed by the Radio Technical Committee for Maritime Services (RTCM) [FHWA, 2005].

Based upon the RTCM-104 message, users are able to perform meter-level positioning in static or moving applications. This message requires approximately 660 bits to send C/A code pseudorange correctors for 12 satellites. It should be noted that correctors are sent because correctors are expected to require fewer bits than observations. Real Time Kinematic (RTK) techniques applied to this 200 baud system may not provide the accuracies of a traditional RTK application, but may show merit in achieving decimeter level accuracies at a limited range from the DGPS broadcast location.

The USCG and others are currently enhancing the NDGPS system, providing for higher accuracy in the sub-meter range. This effort is referred to as the High Accuracy NDGPS (HA-NDGPS) system, which is described in greater detail in the following chapter.

4.5.3 Differential GPS Evaluation

A thorough review of the GPS industry has provided several DGPS correction architectures which possess non-proprietary correction signals and meet the needs of potential VII California implementation strategies. These architectures include:

> **Wide Area Augmentation System (WAAS)** – FAA administered correction signals are broadcast via satellite to the receiver's L1 channel. Horizontal accuracy of 2-3 meters is achievable under ideal conditions.
> **HA-NDGPS** – USCG administered sites broadcast a correction signal on the 455 kHz band with modifications to the current NDGPS sites and beacon broadcasts. Sub-meter level accuracy has been achieved for rover applications. Further HA-NDGPS site deployment and system operation is dependent upon future federal funding.
> **Real Time Kinematic solutions** – Methodology of utilizing a locally placed base station with a communication link to one or more receiver units. Corrections transmitted

between the base station consist of code and carrier phase corrections. Horizontal accuracies ranging from centimeters to 2 meters is achievable depending upon configuration, signal, and operational characteristics.

Background has been provided in the previous chapter on the general GPS architectures and differential correction methodologies. The HA-NDGPS capability is an additional capability integrated within the previously described NDGPS.

4.5.3.1 High-Accuracy Nationwide DGPS (HA-NDGPS)

The U.S. Coast Guard has joined in a collaborative effort with the Federal Highway Administration and other government entities to engineer and deliver centimeter-enabling GPS base station data to a broad range of users by means of a low-cost enhancement to existing and planned NDGPS beacons [Remondi, 2003]. The prototype system utilizes a reference station data link that transmits compacted carrier phase data at 455 kHz. This link and compressed data format is robust enough to transmit sufficient data to users so they can produce centimeter-level solutions at rates up to 1Hz. This High Accuracy NDGPS concept has completed a prototype phase at the NDGPS sites in Hagerstown, MD, and Hawk Run, PA. Initial testing has been extremely promising, achieving excellent data link performance out to 45 km and acceptable performance out to 245 km [Remondi, 2003]. The message format contains a full suite of dual code and carrier measurements into less than 1000 data bits to be transmitted to users [Remondi, 2003]. Utilizing the HA-NDGPS system it is conceivable that the user would be able to reliably achieve 10-30 cm accuracy on the fringe of a typical beacon site's coverage area.

The RTCM committee developed two message type pairs—18 and 19, and 20 and 21—to meet the need of carrier phase corrections achieving sub-meter level accuracy. With this RTCM enhancement, centimeter-level survey and dynamic positioning became possible (based upon an open format) in support of surveying, transportation, and other industrial uses. The carrier phase correction pairs 18 and 19 require approximately 3000 bits to broadcast the full suite of dual code and carrier measurements required by high-precision users. The 3000 bit requirement is too large to fit within the bandwidth allotted to HA-NDGPS on the 455 kHz beacon signal and is therefore compacted [Wolfe et al., 2003].

Alternative formats exist, which are often proprietary, for sending code and carrier measurements or correctors over a link to a user. Other formats are currently under development. Relative to HA-NDGPS, it is very desirable to fit the entire GPS carrier and code observations for 12 satellites within 800 to 1000 bits per second. The HA-NDGPS architecture required a system tailored technique to reduce the message format size while retaining the relevant data.

Experience with Medium Frequency Data Links (such as 455 kHz) has shown that, time-wise, shorter messages provide substantial performance advantages. In cases in which pseudorange corrections are sent, corrections can be broadcast in messages that contain only subsets of satellites. For carrier phase observables, it may be preferable to package all the observables from a given epoch in a single message. To maximize code and carrier synergy, it also may be desirable to maintain the inclusion of code and carrier observables in a single message. The HA-NDGPS beacon signal utilizes a combination of these strategies to achieve a message format within the 1000 bit requirement.

Each HA-NDGPS capable station requires hardware modifications to process and broadcast the 455 kHz signal. The dual-frequency code and carrier message stream are sent to a modulator which is output to the HA-NDGPS transmitter. The output is then diplexed with the NDGPS broadcast to the user based receiver. Figure 4.43 displays a block diagram of the HA-NDGPS broadcast station architecture. As an example implementation, the Hagerstown HA-NDGPS broadcast station rack and receiver antennas are illustrated in Figure 4.44.

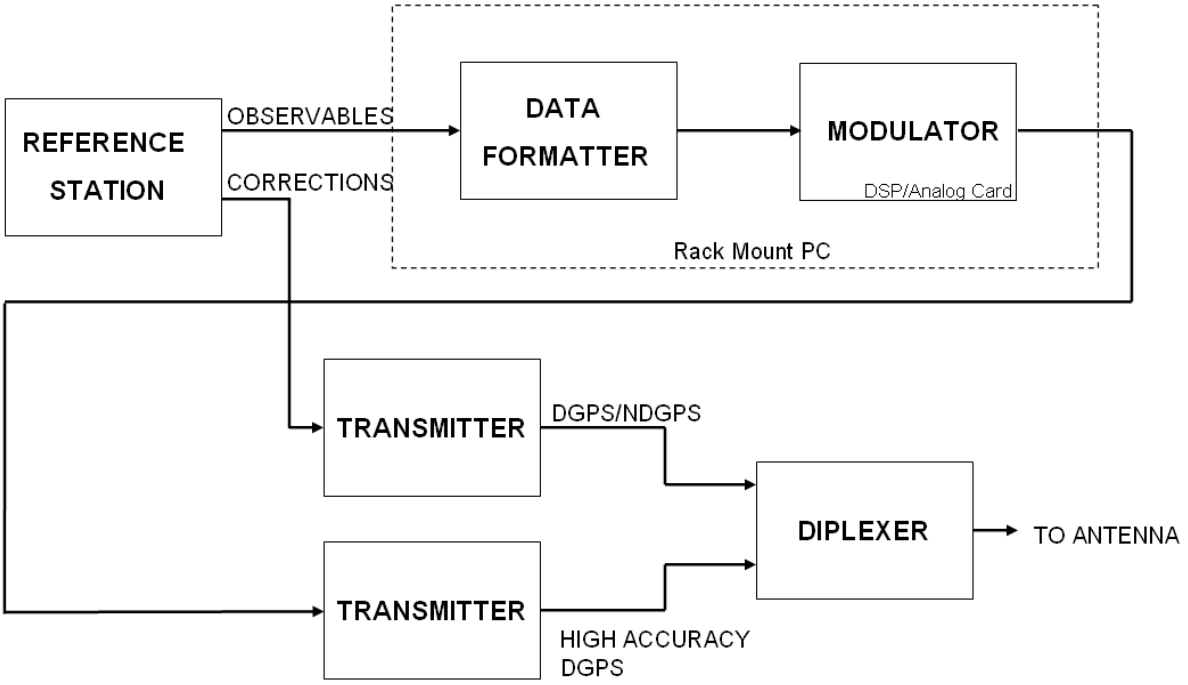


Figure 4.43 HA-NDGPS broadcast station hardware architecture [USCG NAVCEN, 2001].

Figure 4.44 Hagerstown, HA-NDGPS broadcast station rack and receiver antennas [FHWA,2002].

VII California: Development and Deployment
Lessons Learned

The Federal Highway Administration (FHWA) converted several traditional NDGPS sites to meet the high accuracy requirements. The modifications involved interfacing the modulator output to the HA-NDGPS transmitter, the output of which was combined (i.e., diplexed) with the current NDGPS broadcast signal for final broadcast over the air to the equipped user. Hardware requirements and station configuration are shown in Figure 4.46. It was necessary to fit the GPS measurements or correctors (12 satellites) within approximately 1000 bits. The end user's receiver is responsible for demodulating the beacon signal and translating the HA-NDGPS beacon message into RTCM 18 and 19 compatible formats.

The outputs of this module were studied in two different ways. In the first effort, RTCM 18 and 19 messages were converted into Receiver Independent Exchange (RINEX) format and compared with the original data before the RTCM 18 and 19 message formation. These data compared correctly [FHWA, 2005].

The results in Figure 4.45 show that RTCM 18/19 capable user GPS receiver equipment accepted these messages and performed RTK positioning. To test the capability somewhat further, the RTCM 18/19 stream was intentionally disconnected and reconnected. These breaks show up in the plot in Figure 4.45. The number of satellites (divided by 100) and the PDOP (Position Dilution Of Precision, divided by 10) are included in the plot.
A HA-NDGPS beacon station transmits omnidirectionally over an approximate circular area of radius 225 km or more.

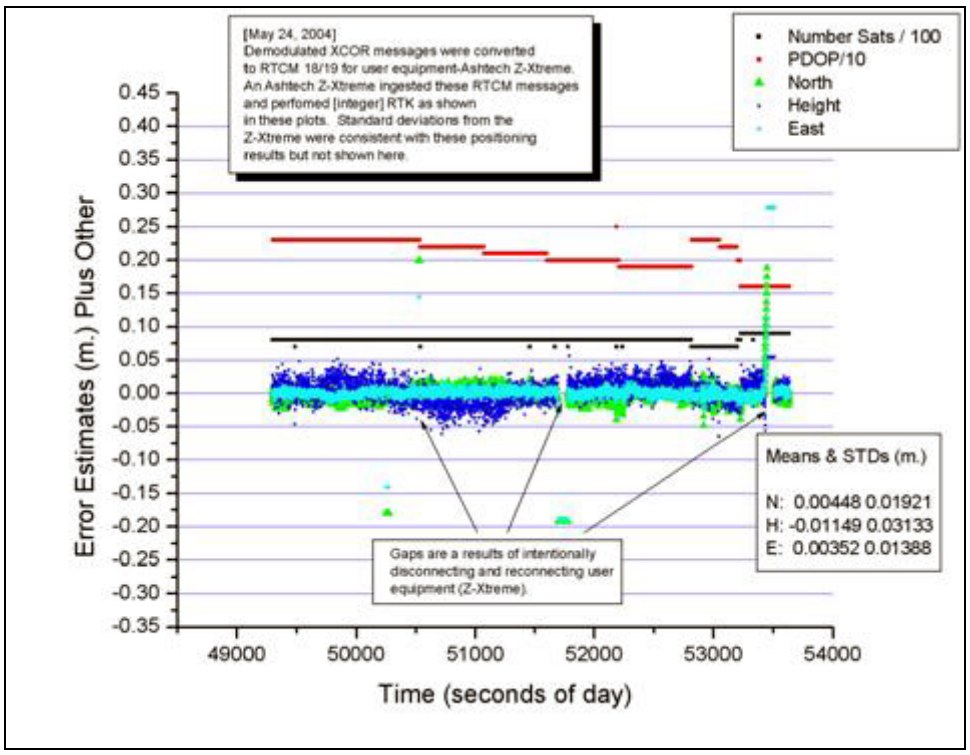


Figure 4.45. HA-NDGPS evaluation of compaction of RTCM 18/19 message types [FHWA,2005].

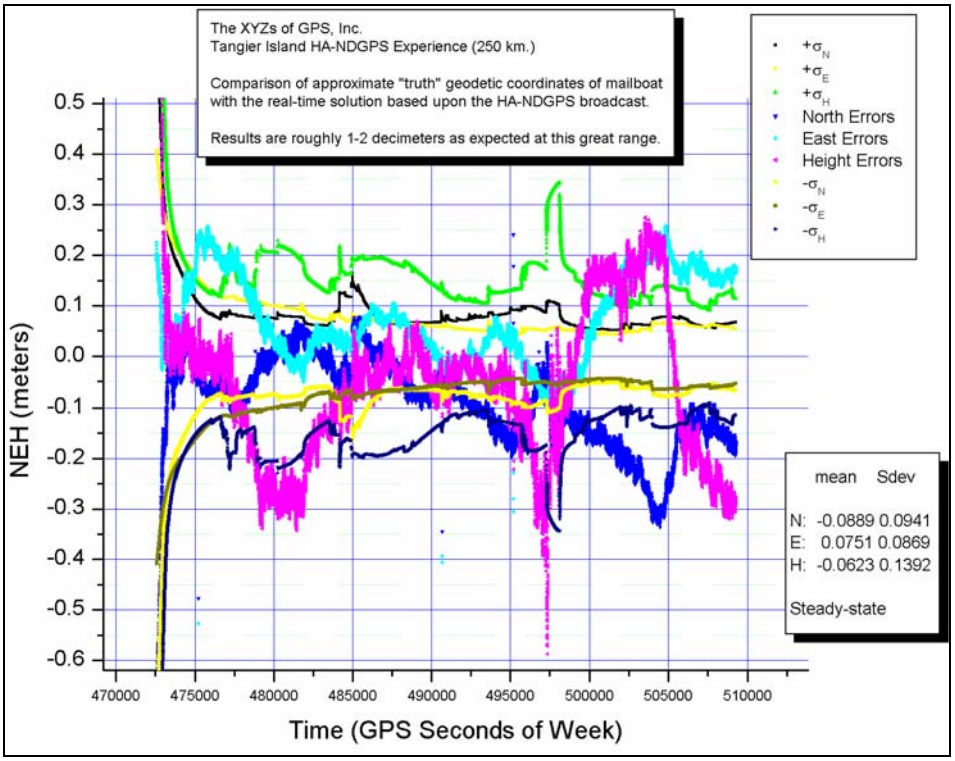


Figure 4.46 HA-NDGPS real time rover data compared with post processed carrier phase [FHWA, 2002].

Significant data manipulation is required to condense the correction data within a 1000 bit message. An existing Ashtech Z-12 GPS receiver provides data to the operational DGPS network. A second output port from the same receiver transmits 1 Hz dual-frequency GPS code and carrier measurements to a PC via an RS-232 port into a software program. This software compresses the data into the allocated bandwidth and passes the resulting "packet" to a modulator. The software's other primary role is to log the data in one or more formats: Raw Ashtech Port Output for possible future playback; compressed packets for possible future playback; or RINEX format for future processing. The software passes the data packets to the modulator, and the modulator readies the packet for transmission [FHWA, 2002].

The ability to squeeze dual- (and even triple-) frequency GPS code and carrier data and corrections into a small bandwidth is important because bandwidth is precious. It also is preferred to deliver the data on a carrier signal of lower frequency. A lower frequency signal becomes necessary to reach well beyond the line of sight. Depending on output power and height of the transmission tower, a transmitter might reach 75 km with a high-frequency line-of-sight signal, but in contrast, the same transmitter could reach 3-4 times farther with a lower frequency signal that can follow the ground and curvature of the earth. The difference between a high-frequency and a low-frequency signal can be a 10-fold factor in terms of the number of base stations required. There is yet another factor that is critical to robust delivery. Low-frequency signals are more useful in buildings, forests, valleys, etc., than are high-frequency line-of-sight signals.

4.5.3.2 WAAS Accuracy

The WAAS correction signal broadcasted from geostationary satellite(s) has been available since 2000 and obtained FAA approval in 2003. The FAA approval followed the rigorous testing completed by Raytheon to meet the FAA requirement of consistent 7 meter or better vertical and horizontal accuracy. Testing results demonstrated for air navigation purposes the system achieved better than 3 meter accuracy [SatNav, 2003]. It should be noted that these tests were for air navigation purposes where surface based obstacles (trees, buildings, mountains, urban canyons) were minimized.

Numerous ground based field experiments have demonstrated some interference issues associated with reception of the WAAS correction signal [Barth et al., 2005; Du, 2006]. Nevertheless, many ground based experiments have demonstrated 2 to 3 meter accuracy with 95% probability (see example in Figure 4.47). As with all differential correction methods, the distance relationship between the reference station and rover receiver plays a role in the accuracy of the correction signal. It is strongly recommended to perform field measurements for a regional application prior to assuming a specific accuracy from a WAAS broadcast.

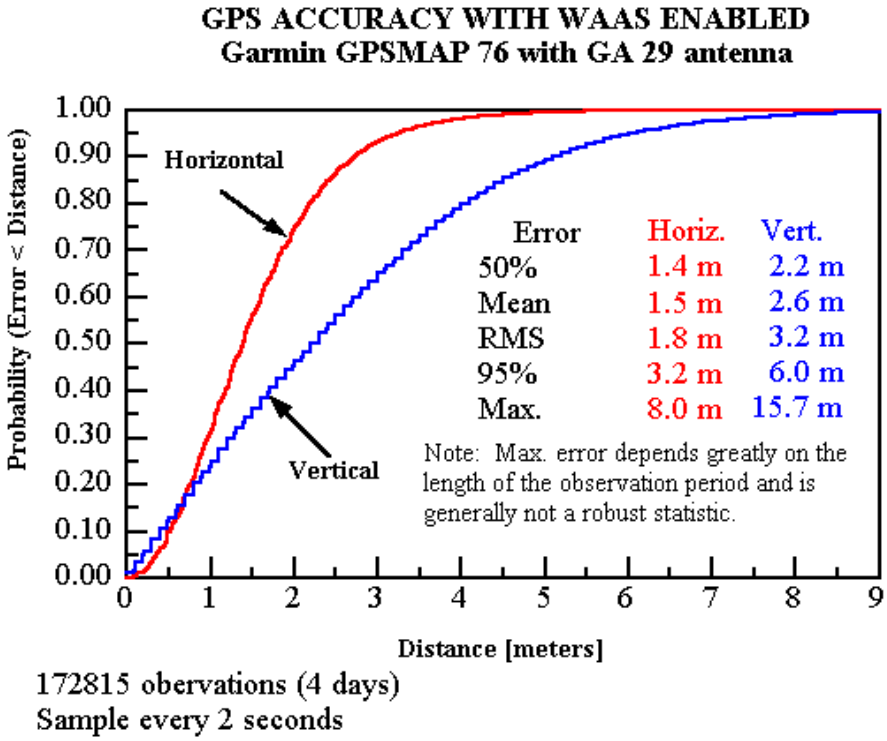


Figure 4.47 Independent accuracy test of WAAS capable receiver [Wilson, 2006].

4.5.3.3 Accuracy of RTK and Other Systems

As described earlier, a Real Time Kinematic solution is a process where GPS carrier-phase signal corrections are transmitted in real time from a reference receiver at a known location to one or more regional rover receivers (<50 km). The use of an RTK-capable GPS base station can compensate for atmospheric delay, orbital errors and other variables in GPS geometry, increasing

positioning accuracy up to within a centimeter using the carrier phase corrections. RTK is used not only as a precision positioning instrument, but also as a core for navigation systems or automatic machine guidance, in applications such as civil engineering and dredging.

The accuracy of ad hoc RTK systems has been shown to vary from millimeter to 2 meters depending upon the system configuration and correction message format. RTK systems utilizing the full carrier phase correction signal can expect to achieve accuracy of 20 centimeters or better. RTK systems which utilize a condensed or partial carrier correction message achieve accuracies of sub meter to 2 meter [FHWA, 2005].

The HA-NDGPS, considered as a RTK solution, utilizes a condensed RTCM 18/19 message while achieving accuracies better than a meter. Figure 4.48 shows HA-NDGPS rover results.

Several other systems exist that provide real-time differential corrections for obtaining high accuracy positioning results. One such example is the CORS network, as described in Section 4.5.2.3.3 and in Chapter 4.5.4. Real-time systems that send only code corrections can reliably achieve 2 meters accuracy, slightly better than what is capable from WAAS. The main advantage of these systems is that the accuracy is consistent (as long as the base station is within approximately 50 km) where the accuracy provided by WAAS changes with different locations in the U.S. An example implementation of this was the design of a Vehicle Lane Positioning System developed for Caltrans during the previous year [Barth et al., 2005]. This system has horizontal accuracy of 2 meters (as shown in Figure 4.51). Using this 2 meter accuracy and stochastic lane map matching algorithms, correct lane determination was achieved 97% of the time [Du et al., 2006].

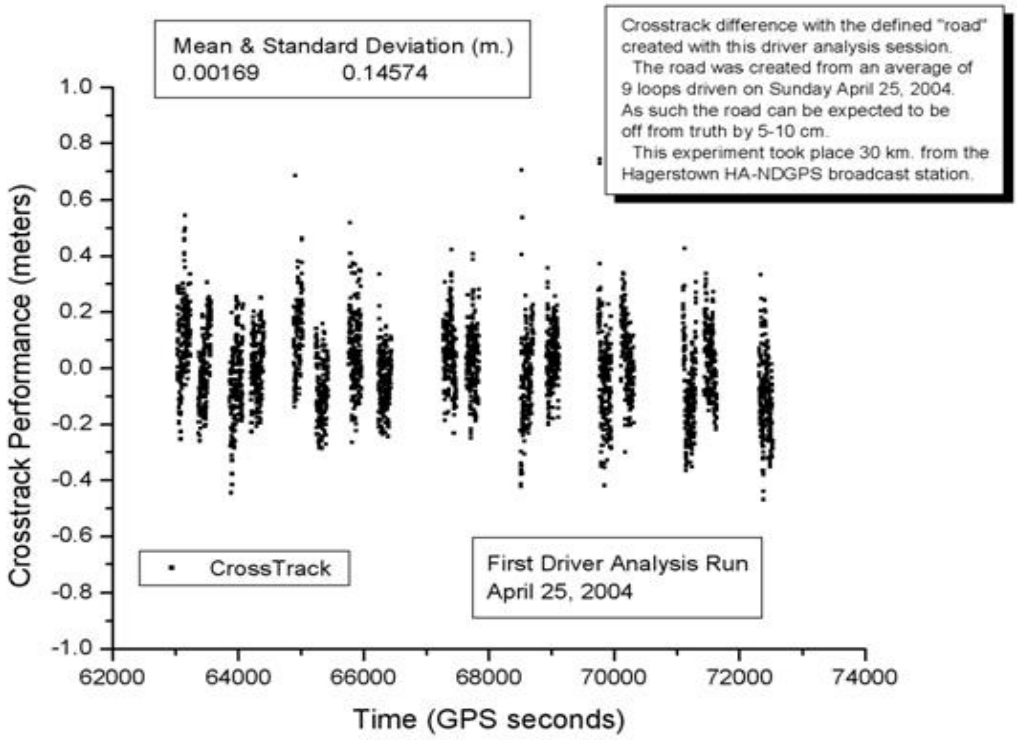


Figure 4.48 HA-NDGPS rover results [FHWA, 2005].

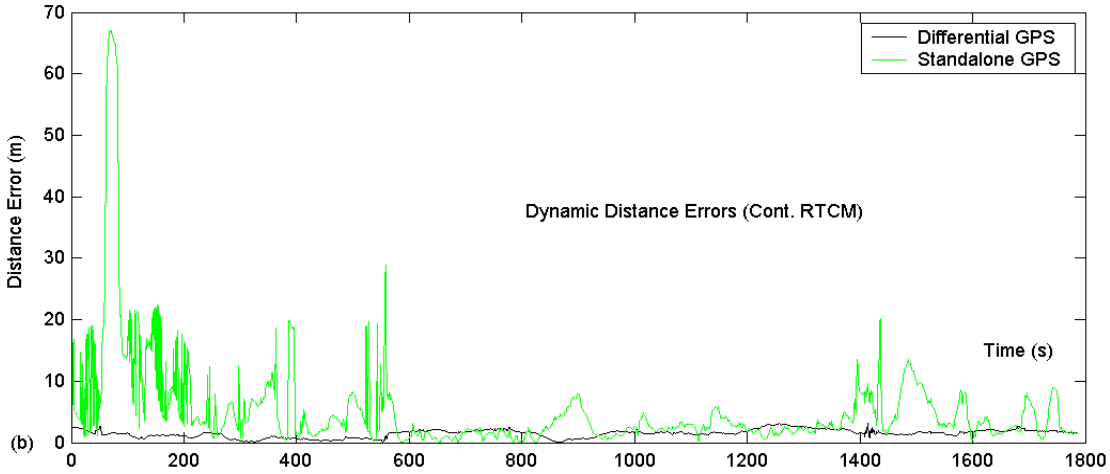


Figure 4.49 Positioning error for uncorrected (green) and corrected (black) local differential GPS system [Du et al., 2006].

4.5.4 DGPS and VII CALIFORNIA Integration Architectures

4.5.4.1 VII CALIFORNIA Architecture

The overall goal of the Vehicle-Infrastructure Integration (VII) initiative is to deploy and enable a communications infrastructure that supports vehicle-to-infrastructure, as well as vehicle-to-vehicle communications, for a variety of applications, with a major focus on improving vehicle safety and transportation operations. California is establishing a VII research and development program that includes testbed development at several locations in the state, on which it is possible to examine a variety of deployment options and potential applications.

The VII California program currently is focused on a three-year project to develop, demonstrate and deploy VII in a key corridor in Northern California. Current partners include Caltrans, MTC, and automobile manufacturers coordinating on the Northern California testbed deployment. The overall goals are being achieved in a three-phase development cycle, the first two phases leading to an initial demonstration over a limited set of roadside units (RSU) and culminating in a Field Operational Test (FOT). The FOT test area is shown in Figure 4.50.

The ITS test applications proposed within VII California necessitate accurate vehicle position of two meters or better for several applications. These applications must determine the specific lane a vehicle is traveling within and potentially provide greater detail relative to velocity and heading. As part of the VII FOT, the vehicles will already possess Dedicated Short Range Communication (DSRC) capabilities, allowing for the communications between the vehicle and the Road Side Equipment (RSE) placed at strategic intersections within the study area. VII applications that will benefit from lane-level positioning accuracy include:

- Traveler Information (standard GPS accuracy is most likely sufficient for this application, unless lane-level navigation is required)

- Lane/Road Departure Warning Systems (lane-level accuracy is desired but not required, since other sensors are used for detection)

- Curve Overspeed Warning Systems (standard GPS accuracy is most likely sufficient)

- Probe Vehicles Operations (positioning accuracy depends on specific application, but a larger number of applications are possible with lane-level positioning capability)

- Ramp Metering (standard GPS accuracy is most likely sufficient)

- Electronic Payment (standard GPS accuracy is most likely sufficient)

- Intersection Safety (lane-level positioning capability is necessary, possibly centimeter-level accuracy for control applications)

It is important to point out that two-meter positioning accuracy does not automatically guarantee the correct determination of lane position. A standard lane is on the order of 3 to 4 meters in width. Lane determination can be made using map matching techniques with accurate lane-center line data [Barth et al., 2005]. Further, stochastic estimation techniques can improve on the lane determination, as described in [Du & Barth, 2006].

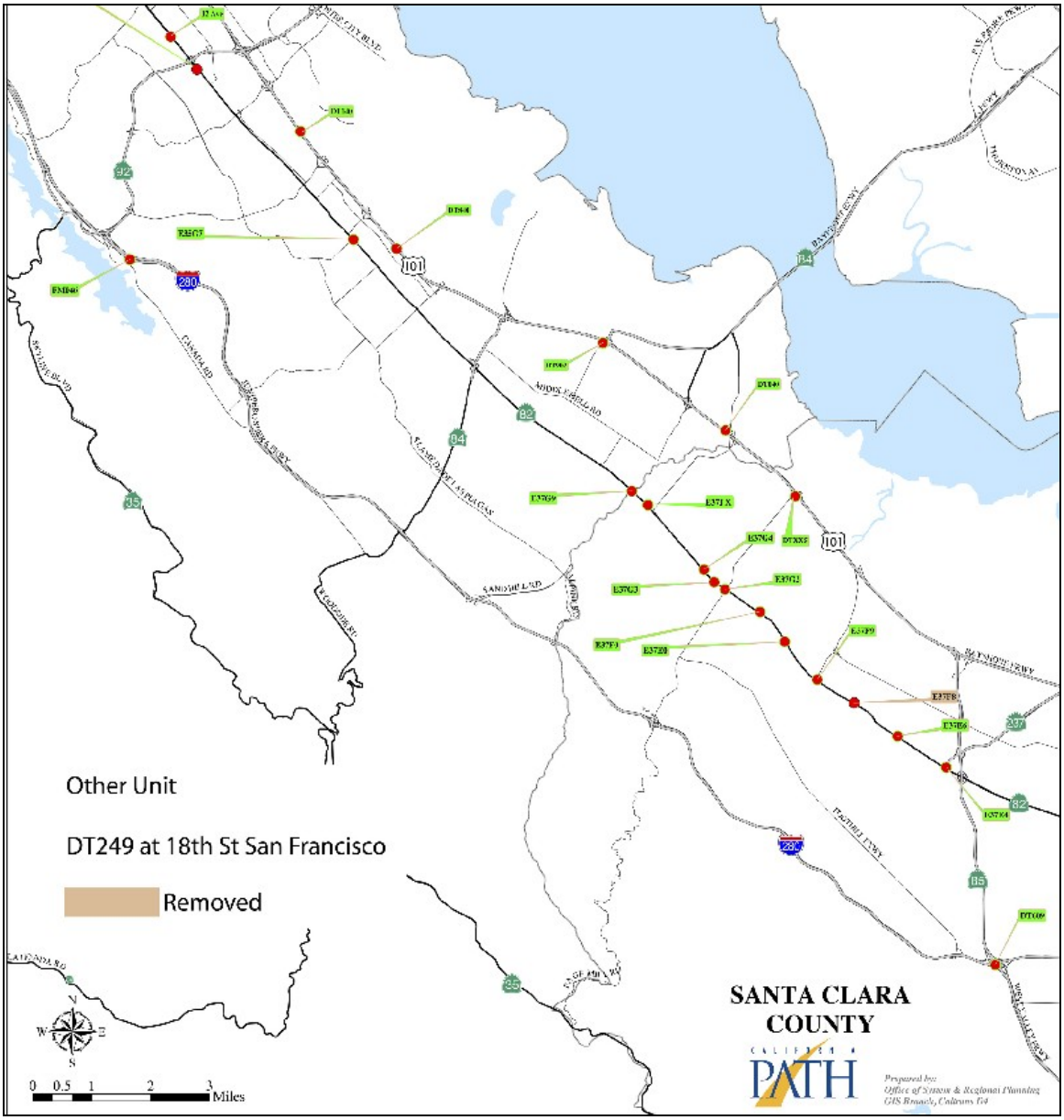


Figure 4.50 VII California Field Operational Test Location [VII CALIFORNIA, 2006a].

The RSE proposed installation configuration for the Northern California testbed is illustrated in Figure 4.51 with the system architecture shown in Figure 4.52. The RSE configuration utilizes a DSRC connection to communicate between the RSE and the vehicles within wireless communication range. The messaging between the DSRC and TCP/IP based RSE network is manage by a PC104 processor located within the RSE cabinet. TCP/IP communications throughout RSE network are achieved through GPRS, T1, or DSL network links.

Figure 4.51 RSE cabinet and wireless communications antenna installation [VII CALIFORNIA, 2006b, c].

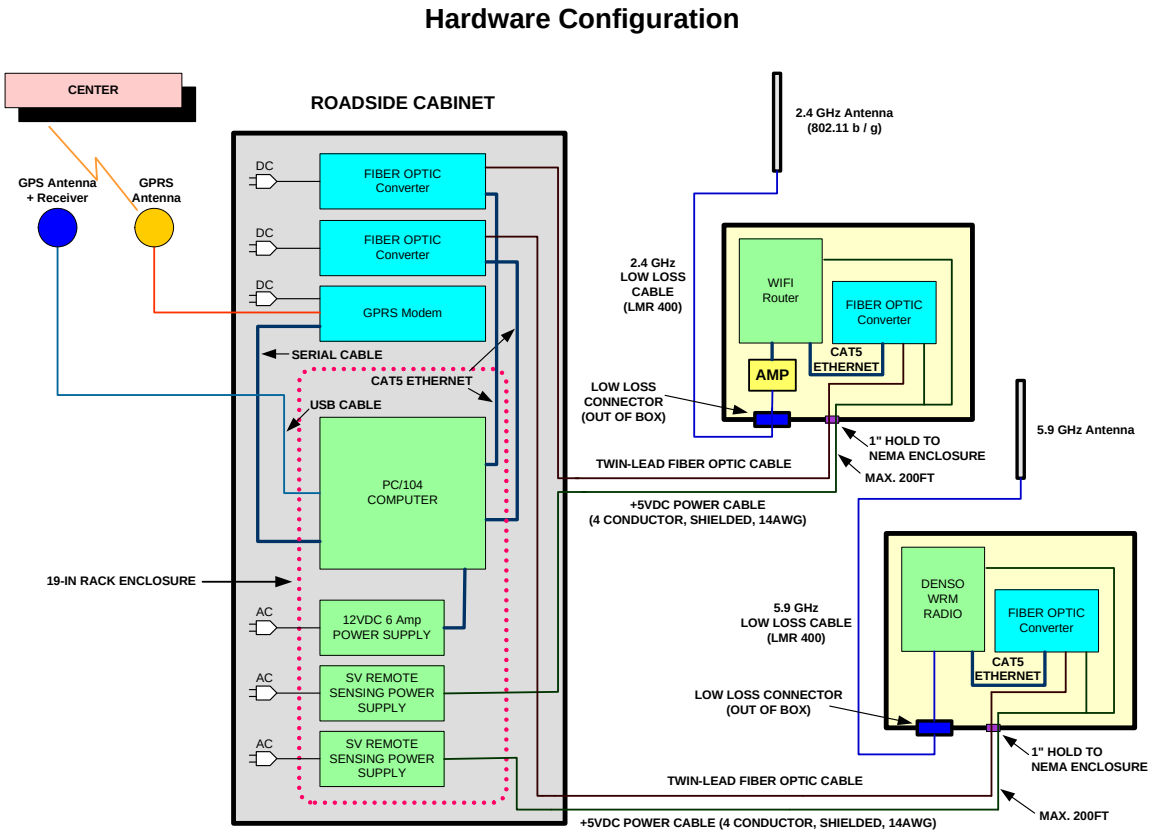


Figure 4.52 VII CALIFORNIA proposed RSE architecture [VII CALIFORNIA, 2006a].

4.5.4.2 Potential Implementation Methodologies

There are several potential methods for implementing 2-meter or better positioning accuracy in the VII CALIFORNIA system. One of the key recommendations of this report is to take

advantage of the communications infrastructure that is already being implemented as part of the VII CALIFORNIA program. This communications infrastructure can be used to receive differential corrections for the various DGPS architectures described earlier, when the vehicle is within range of a road-side unit. It is understood that the vehicles may not always be within communications range of the road-side unit network and therefore the differential correction method cannot rely on this system entirely. For the situation when the vehicles are out of range of the DSRC, other communications means can potentially be used, such as cellular technology.

4.5.4.2.1 HA-NDGPS Options

Several options exist for providing HA-NDGPS corrections to the vehicles:

High Accuracy NDGPS Station Beacon Upgrade with Vehicle Integrated Beacon Receiver:
This architecture requires that a local NDGPS station be upgraded with the HA-NDGPS hardware and software described in Section 4.5.3.1. The upgrade would include: a modulator, transmitter, diplexer, compression software, drivers, network interfaces and control algorithms. The approximate cost of upgrading a single NDGPS facility to perform High Accuracy beacon transmissions is $100,000 per site [USCG NAVCEN, 2001; Arnold, 2005]. Each vehicle would require a dual channel 455 kHz beacon capable receiver with carrier phase RTCM message type 18, 19 processing ability. These receivers are projected to cost anywhere from $3,500-$10,000 (based on an informal pricing survey). Vehicles equipped with the appropriate receiver within the range of the HA-NDGPS beacon signal would be capable of sub-meter (10-30cm) real-time GPS measurements independent of the VII RSE architecture and DSRC communications.

High Accuracy NDGPS Station Beacon Upgrade with RSE Integrated Beacon Receiver:
This architecture also requires that a local NDGPS station be upgraded with the HA-NDGPS hardware and software. Again, the improvements would include: a modulator, transmitter, diplexer, compression software, drivers, network interfaces and control algorithms, and the approximate cost of upgrading a single NDGPS facility to perform High Accuracy beacon transmissions is $100,000 per site. Each vehicle would require a differential capable receiver with carrier-phase RTCM message type 18, 19 processing ability. However, the receivers would not be required to be dual channel to receive the HA-NDGPS beacon signals directly. Instead, carrier phase corrections would be received through the RSE network. These vehicle receivers are projected to cost from $2,500-$8,000 each depending upon capability and manufacturer specifications. Vehicles equipped with these receivers would be able to achieve sub-meter level accuracy once carrier phase corrections were transmitted through the DSRC located at the RSE. A beacon capable HA-NDPGS receiver would be centrally located within the RSE network. Carrier phase corrections would be received through a single centrally located beacon receiver and transmitted through the RSE network. Each DSRC location would have the ability to broadcast the corrections to vehicles within its broadcast area. Individual vehicles would implement a low cost receiver and the RSE network would integrate a HA-NDGPS beacon receiver at a cost of $3,500-$10,000.

High Accuracy NDGPS Station IP Upgrade with RSE/DSRC:
This architecture requires that a local NDGPS station be upgraded with the ability to generate a HA-NDGPS differential correction message and transmit the message via TCP/IP to the RSE network. The improvements are primarily in software and minimal in networking hardware (i.e.,

communication software, drivers, network interfaces, and control algorithms). The approximate cost of upgrading a single NDGPS facility to perform High Accuracy network transmissions is approximately $15,000 per site [Arnold 2006]. Each vehicle would require a differential capable receiver with carrier phase RTCM message type 18, 19 processing ability. The receivers would not be required to receive the HA-NDGPS beacon signals directly. These receivers are projected to cost from $2,500-$8,000 each depending upon capability and manufacturer specifications. Vehicles equipped with carrier phase compatible receivers would be expected to achieve sub-meter level accuracy once carrier phase corrections were transmitted through the DSRC located at the RSE. Carrier phase corrections would be transmitted from the HA-NDGPS station to the RSE network via TCP/IP and transmitted to vehicles by DSRC. Each DSRC location would have the ability to broadcast the corrections to vehicles within its broadcast area. Individual vehicles would implement a low cost receiver and the RSE network would require a reliable and dedicated TCP/IP connection to the HA-NDGPS station generating the corrections.

4.5.4.2.2 WAAS Integration Architecture with VII California

FAA's Wide Area Augmentation System is based on a network of approximately 25 (currently adding 13 additional stations) ground reference stations that covers a very large service area in North America. Signals from GPS satellites are received by a network of ground reference stations. Each of these precisely surveyed reference stations receives GPS signals and determines the common mode errors. Each ground reference station is linked to form the WAAS network, and the ground reference stations relay the data to the wide area master station where information is computed. A correction message is prepared and uplinked to geosynchronous satellites via a ground uplink system. The message is then broadcast from the satellite on the same frequency as the standard GPS signals (L1, 1575.42MHz), which can then be received directly from GPS receivers on the ground.

Integration of the WAAS DGPS correction within the VII CALIFORNIA architecture would simply require the installation of WAAS-capable GPS receivers in each vehicle. There is already a wide array of WAAS-capable GPS receivers on the market, and as a result, the cost is quite low. It is important to note that several accuracy experiments have shown that WAAS-corrected DGPS provides 2-3 meter accuracy on average, however the consistency of signal reception and accuracy for ground based navigation applications can vary under different vehicle telemetry conditions and locations (see, e.g., [Du & Barth, 2005]). The WAAS correction signal would be received from one or two geostationary satellites covering the western continental U.S. The WAAS signal is potentially interrupted by dense vegetation, buildings, canyons, or mountains. The VII applications would require consistent accuracy under optimum WAAS conditions.

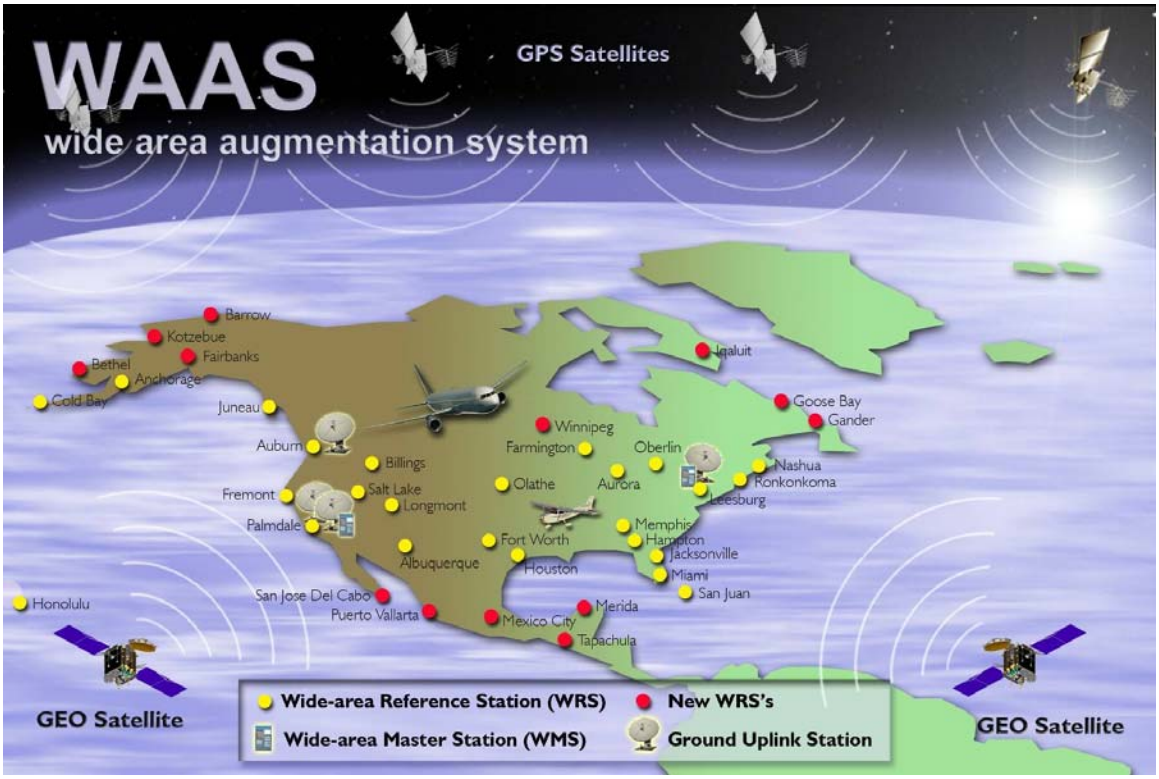


Figure 4.53 WAAS base stations showing Fremont as the closest to the Northern VII California Field Operational Test [SatNav, 2003].

4.5.4.2.3 Integrating the CORS Network with VII California

As described in Section 4.5.3.3., the California CORS (Constantly Operating Reference System) network maintained through the California Spatial Reference Center has had significant influence on the implementation, development, and operation of DGPS reference stations throughout California. A subset of the reference stations in Southern California transmit carrier-phase corrections suitable for RTK corrections, allowing for sub-meter accuracy. Figure 4.54 shows the California CORS base stations. This system is currently well developed throughout California for standard code-based corrections, providing approximately 2-meter accuracy. For carrier-phase corrections, the system is well developed for Southern California and less so for Northern California (with planned expansion).

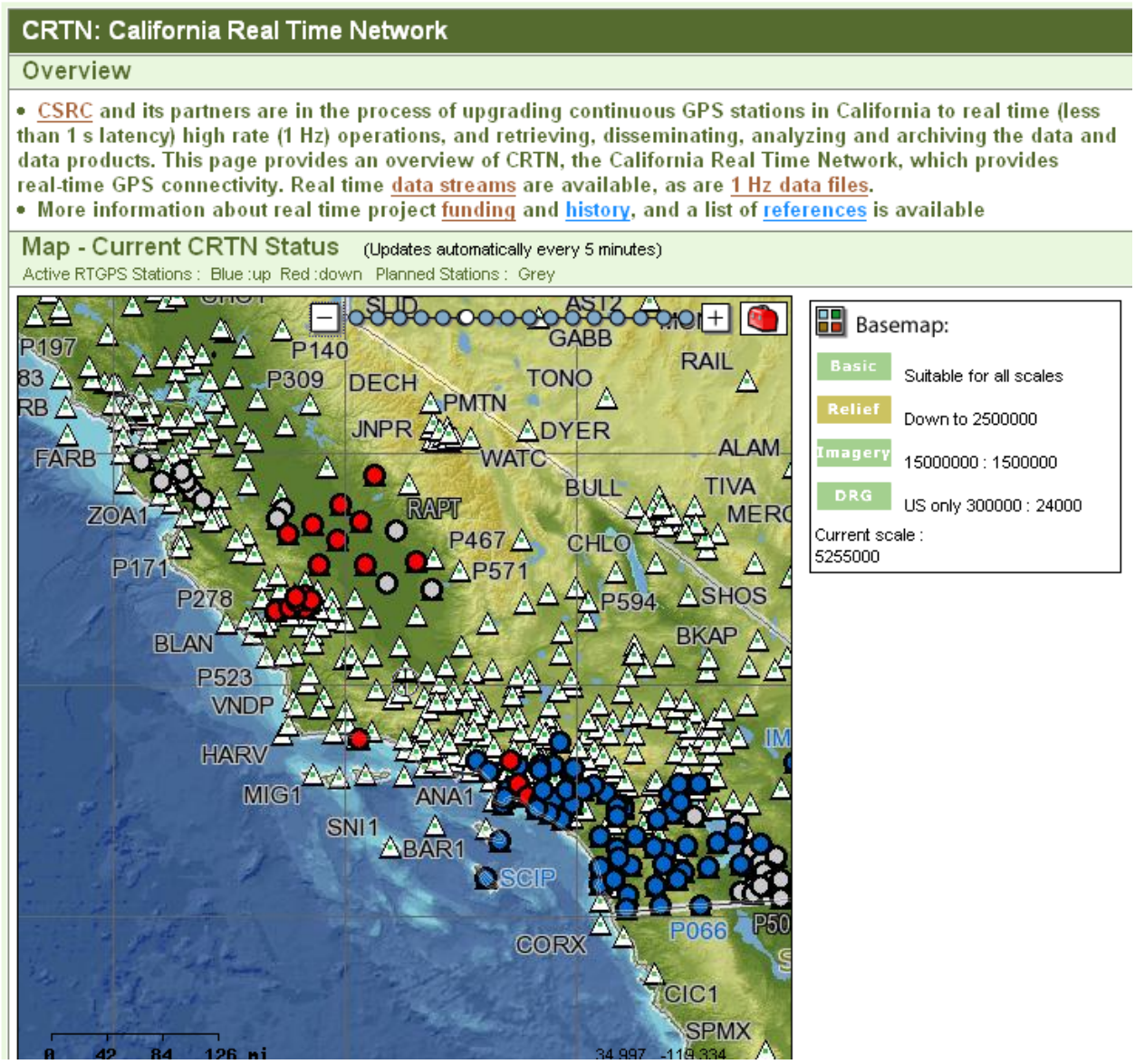


Figure 4.54. California CORS stations with RCTM 2.3 carrier phase corrections [CSRC, 2006].

In terms of the VII California requirements, the CORS system is capable of provide code-based corrections from all of their sites. These code-based corrections should provide consistent 2-meter accuracy for the vehicles that operate within 50 km of their established base stations. If sub-meter accuracy is needed, then carrier-phase corrections can be received from a subset of the CORS network. All of the corrections are provided in the standardized RTCM formats.

If the existing CORS network base stations do not provide sufficient coverage for the VII CALIFORNIA applications (e.g., if sub-meter accuracy is needed), it is relatively straightforward to establish another (carrier phase) base station in the proximity of the applications. This base station can run independently and potentially become part of the overall CORS network. The base station should be centrally located within the test region with less than a 50 km radius of VII operations. Additional base stations can be added as needed to provide

adequate coverage. The base station(s) would possess a TCP/IP connection with the RSE network. As described previously, the RSE could utilize the DSRC communications to transmit carrier phase corrections (RCTM 104 message type 18, 19) to vehicles operating within the program.

The on-board vehicle DSRC transceiver would pass the correction signals to the on-board GPS receiver. The receiver would calculate it position (2 meters for code-based corrections, sub-meter for carrier phase corrections) based on the received corrections. The corrected position could then be utilized on-board and even transmitted back to the RSE network.

This configuration would utilize on-board GPS receivers capable of processing carrier phase corrections. The price range of these receivers is $2,500-$8,000. If only 2-meter accuracy is required, then a GPS receiver would be significantly less, approximately $300 - $1000. If only 2-meter accuracy is required, then existing CORS base stations can be used. However, If sub-meter accuracy is desired, the system would require an initial base station capable of generating carrier phase corrections. The base station must also be capable of networked based transmission of corrections. The price of the base station would be approximately $3,000-$10,000. Figure 4.55 shows an example base station configuration implemented as part of the California CORS real time network.

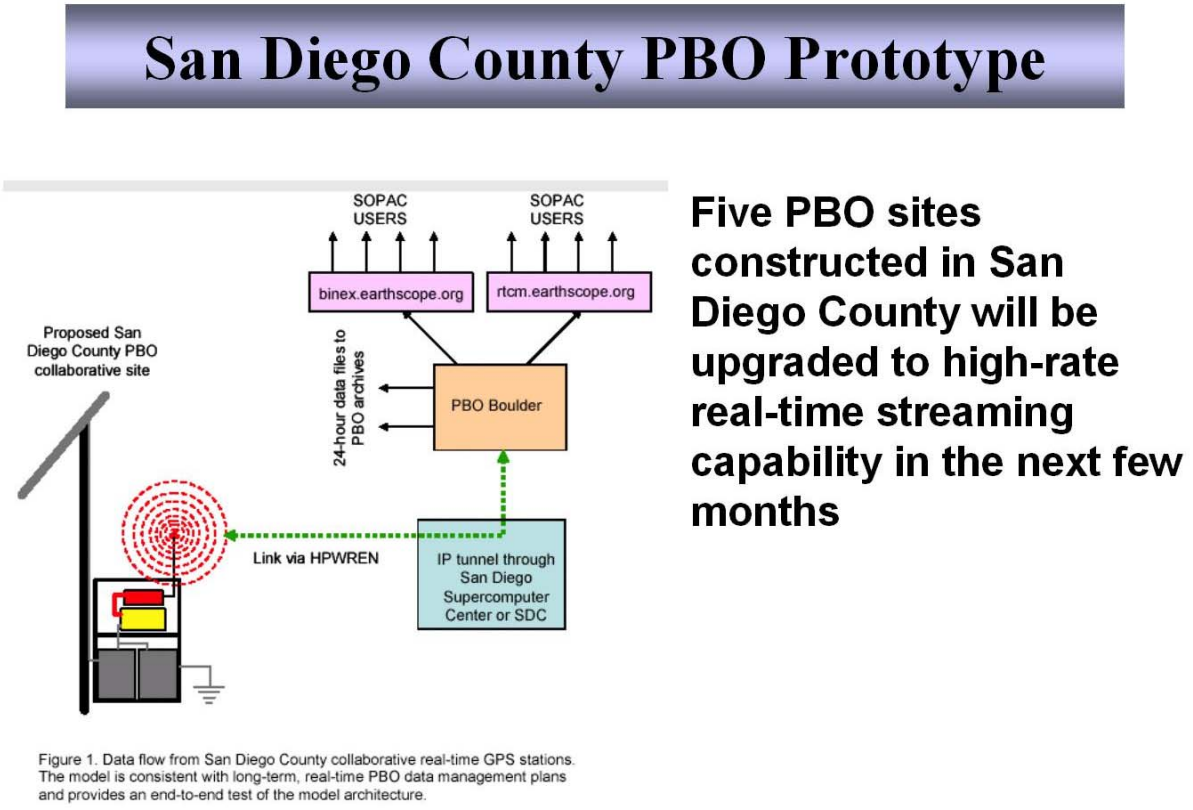


Figure 4.55 Real time carrier phase DGPS base station utilized for RTK applications [CSRC,2006].

4.5.5 Recommendations

This report thus far has described numerous differential GPS architectures which could potential satisfy the positioning needs of the VII California program. The HA-NDGPS architecture has been presented in detail with descriptions of technology, integration, and potential future system development. Since significant doubt has evolved associated with the future federal funding of the HA-NDGPS deployment, other alternatives have been extensively explored. We believe that the alternatives that have been explored also meet the accuracy and performance characteristics required by the current applications being developed for the VII California program (as discussed in the previous section). This section provides recommendations regarding various DGPS architectures that are most suitable for the VII California program, beginning first with HA-NDGPS, followed by the alternatives. A summary table is provided at the end of this section. One of the key components of these recommendations is to make use of the VII California communication architecture (where possible) to provide the correction signals to the vehicles. This potentially eliminates the need for having an additional communication system for providing the corrections directly to the DGPS receivers, thereby providing a significant cost savings. However, it is understood that the vehicles may not always be in direct communications with a road-side communications unit, therefore if lane-level positioning is required at all possible times, then a secondary communications methodology will be required.

Further, it is assumed that nearly all of the VII California applications require lane-level positioning accuracy. This lane-level positioning can easily be satisfied with sub-meter accuracy provided by carrier-phase DGPS techniques, including the HA-NDGPS system. However, DGPS receivers capable of processing carrier-phase signals are significantly more expensive than other (code-based) DGPS receivers. Nevertheless, approximately 2-meter accuracy can be achieved using code-based corrections, and along with appropriate lane centerline map matching algorithms and estimation techniques, it has been shown that lane-level positioning is possible [Du et al, 2006]. Therefore, a DGPS solution that provides consistent 2-meter accuracy can satisfy the needs of many VII California applications. If sub-meter accuracy is necessary for intersection safety and control and other applications, then carrier-phase DGPS techniques will indeed be necessary.

4.5.5.1 HA-NDGPS Recommended Architectures

The HA-NDGPS recommendations include two general architectures: 1) HA-NDGPS with radio beacon correction transmissions, and; 2) HA-NDGPS without radio beacon correction transmissions; instead corrections are provide through the VII communication network (Internet TCP/IP to RSEs then DSRC to the vehicles).

The HA-NDGPS utilizes carrier phase corrections originating from one or more base stations preferably located within 250km radius of the application, as described in previous sections. The typical method of transmitting the correction signal from the base station to a field receiver is utilizing a defined beacon signal at the 455kHz frequency. The receivers utilized for HA-NDGPS are designed with an additional communications receiver and software to process the correction signals. The specialized HA-NDGPS receivers are designed to process the system specific 1000-bit differential correction message. While the system is intended to primarily transmit corrections through the beacon methodology, there is also the opportunity to transmit corrections via the

Internet to the VII California network infrastructure. The details of both architectures and associated characteristics are presented.

The HA-NDGPS utilizing beacon correction transmissions would necessitate the implementation of two architecture components to successfully position with sub-meter level accuracy. First, a NDGPS base station should be geographically located within 250 km and possess the necessary upgrades to transmit the high accuracy signal on the 455kHz dedicated beacon signal. This upgrade has been estimated to cost approximately $100,000 per location [Arnold, 2006]. Second, the receiver expecting to receive the high accuracy signal must be manufactured with the proper hardware and software for process the signal. These receivers, currently being designed and manufactured, are expected to cost anywhere between $3,500 and $10,000 per unit (installed in each vehicle).

This HA-NDGPS beacon architecture consists of every vehicle requiring sub-meter positioning to be equipped with the DGPS integrated beacon receiver and that the regional NDGPS station(s) be upgraded with the high accuracy transmission capability. While the initial cost of the HA-NDGPS is high, the long-term costs to individual system users are significantly reduced. The high accuracy signal would be unrestricted and available to all individuals and entities possessing the proper receivers to interpret the beacon signal. Therefore the long term projected cost of an additional entity or vehicle migrating to HA-NDGPS may decrease to the $1,000-$2,000 range. Nonetheless, a HA-NDGPS receiver equipped to process the beacon signal must possess additional hardware, which will always increase costs compared to competing differential receivers.

NDGPS facilities upgraded to transmit the high accuracy signal typically integrate the 455 kHz beacon RF transmitters and associated hardware. Alternatively, the facility can be upgraded to transmit the high accuracy signal via the Internet. This data transmission can occur over a high bandwidth Internet link between the NDGPS base station and the VII California communications WAN. Since the VII California program proposes to have an independent communications infrastructure, high accuracy DGPS corrections can potentially be transmitted to the vehicles via the VII communications infrastructure. This proposed architecture utilizes the existing communications networks to reduce the hardware and infrastructure costs of the HA-NDGPS architecture proposed previously. The cost of upgrading a NDGPS facility to broadcast high accuracy corrections via IP is estimated to be approximately $15,000 without the beacon infrastructure. Additionally, once a high accuracy signal reaches the VII communications infrastructure and is transmitted to a vehicle, the vehicle hardware does not require the additional beacon receiver. The on-board GPS receiver would have to be coupled with the appropriate software to process the corrections in the high accuracy format. The receivers suited for this type of implementation are estimated to be $500-$1000 less than the receivers designed with the HA-NDGPS integrated receiver.

Another possibility is to NOT upgrade the NDGPS facilities to high-accuracy, but still have the vehicles receive code-based NDGPS corrections, providing 2-meter level accuracy. However, the correction signals must still be provided either through the radio beacon or through the Internet and DSRC via the RSE. This possibility would not require any base-station costs but each vehicle would require specific NDGPS receivers.

4.5.5.2. WAAS Recommended Architecture

The FAA WAAS DGPS correction architecture is perhaps the most prevalent of code-based DGPS corrections. As described in previous sections, the correction signal is transmitted on the L1 GPS frequency band and most GPS manufactures are already mass-producing low cost receivers (approx. $500) that integrate these WAAS corrections. While WAAS-capable receivers are frequently found to possess accuracies better than 3 meters, these results are often associated with consistent and reliable signal reception in addition to proximity to a WAAS base station. Additionally, GPS receiver manufactures possess varied software methodologies for receiving and processing the WAAS corrections.

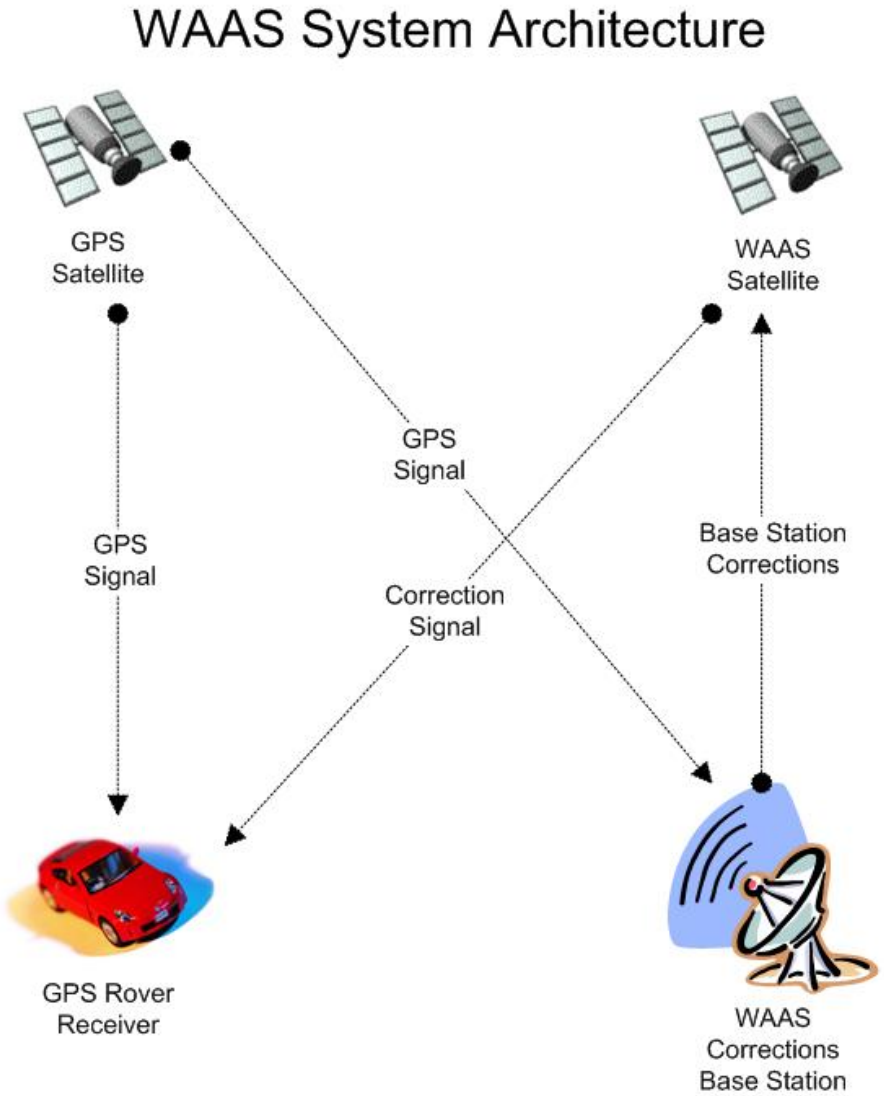


Figure. 4.56 Proposed WAAS architecture for potential VII CALIFORNIA implementation.

The integration of WAAS receivers within the VII infrastructure is straightforward, as illustrated in Figure 4.56. The signal is currently active and available with a numerous receivers suitable for integration within vehicles. While WAAS is the lowest cost and easiest to implement of the

proposed solutions, the accuracy of the corrected result is the most questionable for VII purposes (of those presented). The VII program requires consistent and reliable lane level accuracy of approximately 2 meters. While WAAS receivers certainly achieve this level of accuracy for stationary applications in certain areas, there has been some inconsistency observed for mobile applications. Therefore, while WAAS is very attractive relative to cost and implementation, it may not provide consistent lane-level accuracy and it is cautioned without thorough evaluation for accuracy for specific VII applications.

4.5.5.3 Other DGPS Architectures

Real Time Kinematic (RTK) systems were presented previously and shown to provide sub-meter accuracy for mobile applications. These systems also utilize carrier phase corrections, specialized software, and dedicated communications to achieve this level of performance. Most RTK receivers are intended to be paired with compatible base stations utilizing proprietary software. However, there is significant activity within California to implement the infrastructure to broadcast standardized (non-proprietary) RTK corrections at 1 Hz, as part of the CORS program. These efforts are primarily centered within Southern California and have yet to provide any real time carrier phase corrections to the area of initial interest (i.e., Northern California) to the VII California program [CSRC, 2006].

Similar to that described in Section 4.5.5.1, the integration of an RTK solution for VII California can take advantage of the VII communications infrastructure for broadcasting corrections to the vehicles' GPS receivers. The base station to obtain corrections can be centrally located for the VII effort and can be integrated with the California statewide effort if desired. Vehicles would require receivers capable of processing carrier phase corrections. These receivers are currently dual channel (L1, L2) receivers produced by numerous receiver manufactures in the price range of $2,500-$8,000.

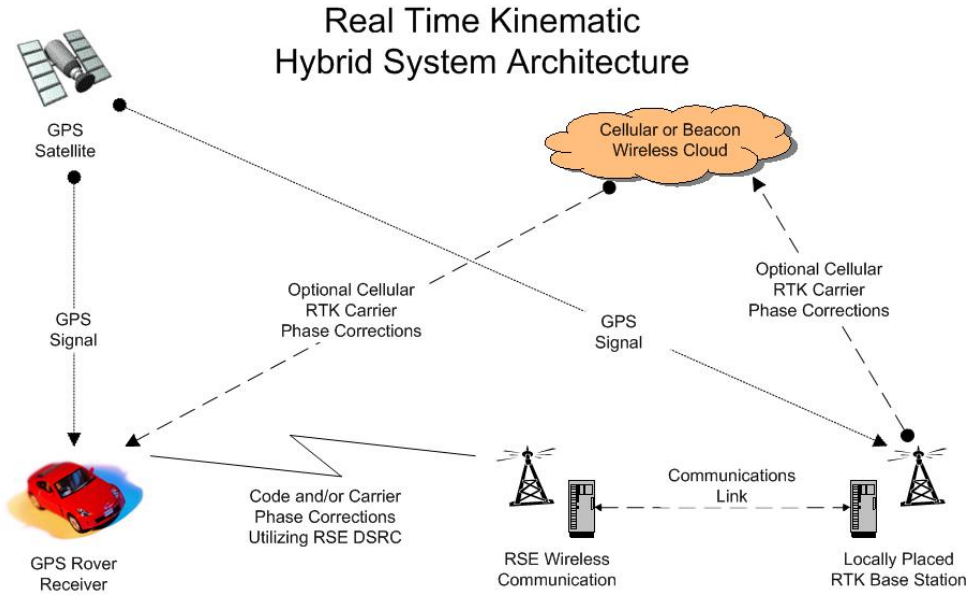


Figure 4.57 Proposed RTK architecture for potential VII CALIFORNIA implementation.

4.5.5.4 RSE Integrated Recommended Architecture

Again, if sub-meter level accuracy is not needed and lane-level positioning is the primary requirement, a DGPS system providing 2-meter level accuracy should be sufficient (coupled with appropriate lane centerline map matching estimation techniques). The same CORS system already developed throughout California can provide code-based corrections, allowing for inexpensive DGPS receivers to achieve 2-meter level accuracy reliably and consistently.

This method also utilizes a locally placed differential base station but only transmits RTCM pseudorange (code-based) corrections. The base station GPS receiver can have integrated multipath mitigation and employ strobe correlator digital signal processing to produce RTCM SC-104 (Radio Technical Marine Service) differential correction messages and raw data output of code at a typical update rate of 1 Hz. This methodology of transmitting corrections has shown improvement over WAAS for vehicle based differential corrections [Barth et al., 2005]. This overall architecture is illustrated in Figure 4.58. The correction signals would be provided directly via the VII California existing communication architecture.

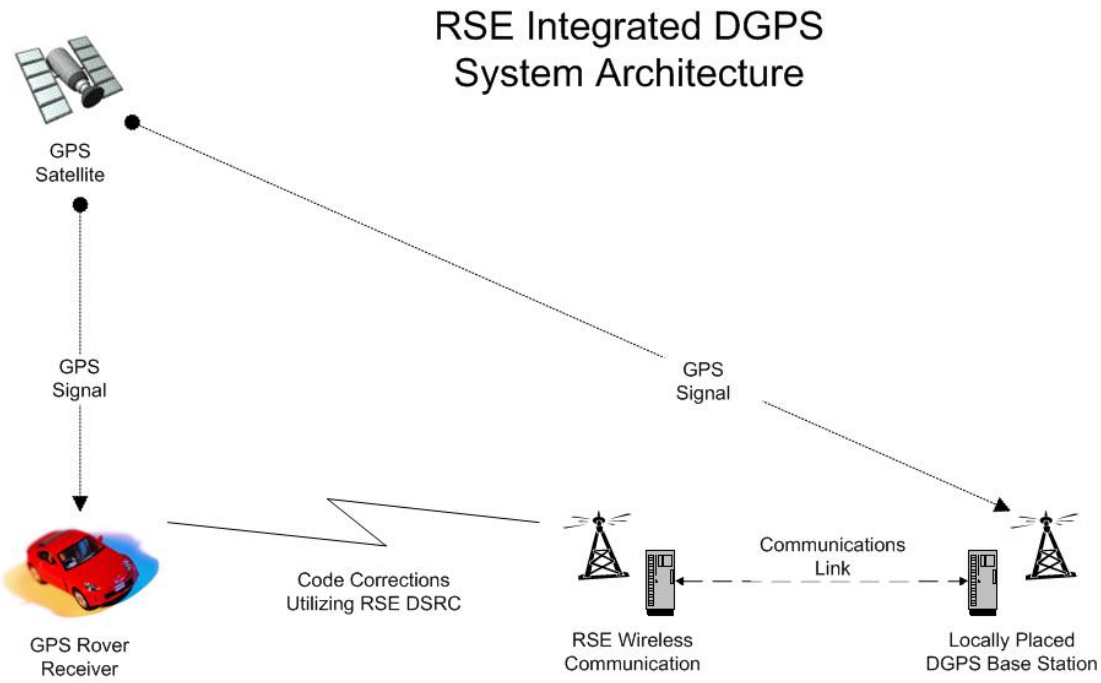


Figure 4.58 Proposed RSE integrated DGPS architecture for potential VII CALIFORNIA implementation.

This approach possesses several desirable characteristics:

- corrections can be sent in a standard RCTM format over existing VII dedicated communications;
- differential capable base station(s) can be placed at central locations depending on VII testbed sites;
- software and hardware are not proprietary;

- the architecture can be expanded to send carrier phase corrections (RCTM 19/20, CMR);
- low cost receivers can be utilized for 2-meter accuracy requirements while more expensive dual channel receivers are utilized for sub-meter level accuracy;
- the infrastructure can be integrated with existing CORS network, and
- the system is easily upgradeable as GPS technology continues to improve.

This general methodology has already been evaluated for numerous other applications (e.g., see [CSRC, 2006]).

4.5.5.5 Implementation Cost and Implementation Summary

The analysis and review of these various architectures has provided insight into numerous desirable qualities. While no single option is found to possess every desirable characteristic, this evaluation has clarified strengths and weaknesses among the options. These characteristics are summarized in the table below.

Table 4.6 Proposed VII CALIFORNIA DGPS implementation costs and comparison.

Architecture	Correction	Accuracy	Base Station	Base Station	GPS Receiver
HA-NDGPS beacon	Carrier phase via beacon	<1m	NDGPS conversion	Approx. $100,000	$3,500-$10,000
HA-NDGPS no beacon	Carrier phase via DSRC	<1m	NDGPS conversion	Approx. $15,000	$2,500-$8,000
WAAS	Pseudorange via satellite	~3m	exists	existing	$250-$1000
integrated base station-RSE	Pseudorange via DSRC	~2m	local	$3000-$10,000	$250-$1000
integrated CORS -RSE	Pseudorange via CORS	~2m	central	$0 - $10,000	$250-$1000
RTK – local wireless	Carrier phase via DSRC	<1m	local	$3000-$10,000	$2,500-$8,000
RTK – CORS integrated	Carrier phase via CORS	<1m	central	$0 - $10,000	$3,500-$10,000

Each of these options possesses varied levels of integration requirements for implementation within the VII California program. The HA-NDPGS options require the conversion of current NDGPS facilities with the integration of software and hardware. These station improvements require 2-6 months. The WAAS corrections require no additional infrastructure, but do require the appropriate receivers within the vehicles. Some RTK systems are commercially available and only require the appropriate ordering and installation delays. VII customized RTK systems require the integration of the base station (1-3 months) and the transmission protocols through the VII communication infrastructure. Similarly, the RSE integrated corrections utilized a centrally implemented base station (1-3 months) and transmission protocols to be integrated within the VII communication infrastructure.

While the WAAS architecture possesses the least cost and simplest integration, it has the least accuracy and lacks consistency below 3 meters. Utilizing RSE integrated corrections improves the accuracy slightly with improved consistency, but there are increased costs with base station installation and communications integration. RTK will provide optimum performance but possess high dual channel receiver costs. Finally, HA-NDGPS solutions will also provide optimum performance on a nationwide proposed system with high initial costs relative to base

station conversion. The HA-NDGPS dual channel receivers are also costly and include beacon receivers for beacon-based corrections.

4.5.6 SUMMARY

This study has evaluated numerous high accuracy differential GPS correction architectures suitable for VII California implementation. The California PATH program and Caltrans have a number of research applications that can benefit from these high accuracy vehicle positioning implementations. HA-NDGPS and several other differential GPS architectures have been reviewed relative to operational characteristics, integration requirements, relative cost, and potential positional accuracy. The scope of the study was expanded in response to concerns associated with funding shortfalls associated with the HA-NDPS implementation. While the original goals of the study have been maintained relative to evaluating HA-NDGPS for VII CALIFORNIA implementations, additional DGPS architectures such as WAAS and Real Time Kinematics (RTK) have been comparatively evaluated.

RTK systems were reviewed and shown to provide excellent accuracy (<10cm) for mobile applications. These systems utilize carrier phase corrections, specialized software (predominantly), and dedicated communications to achieve this level of performance. Most RTK receivers are intended to be paired with compatible base stations utilizing proprietary software.
If possible, the integration of an RTK solution within the VII program should take advantage of the VII DSRC communications infrastructure for broadcasting corrections to the vehicles' GPS receivers. The base station to obtain corrections can be centrally located for the VII effort and can be integrated with the California statewide CORS effort if desired. Vehicles would require a receiver capable of processing carrier phase corrections. These receivers are currently dual channel (L1, L2) receivers produced by numerous receiver manufactures in the price range of $2500-$8,000.

While the RTK solution utilizes dual channel carrier phase corrections to achieve reliable accuracy better than 10 cm, there is a single channel L1 pseudorange correction methodology similar in architecture, which achieves 2-meter accuracy reliably and consistently. This method also utilizes a locally placed differential base station but only transmits RTCM pseudorange (code based) corrections. This methodology of transmitting corrections has shown some improvement over WAAS for vehicle based differential corrections.

Integration of standard DGPS correction signals with the VII communication infrastructure possesses several desirable characteristics:

- corrections can be sent in a standard RTCM format that many GPS receivers are able to read;
- base station(s) can be placed centrally to the VII areas (e.g., within 50 km of vehicles);
- software and hardware are not proprietary;

- the architecture can be expanded to send carrier phase corrections (RTCM 19/20, CMR) for greater accuracy (< 10cm);

- low cost receivers can be utilized for 2m accuracy requirements while more expensive dual channel receivers are utilized for centimeter level accuracy;

- the infrastructure can also be integrated with CORS; and

- the system is easily upgradeable as GPS technology continues to improve.

While this hybrid RSE approach has not been developed commercially, the methodology and experimentation has been thoroughly evaluated in University research programs. Additionally, the utilization of RSE DSRC allows the gradual integration of multiple DGPS correction signals. Initial integration may consist of only code corrections while future implementations could add carrier phase corrections with only software changes to the implemented base station and RSE hardware.

4.5.7. NEXT STEPS

Several important future steps should take place prior to any DGPS positioning system is put into place for VII California. These steps include:

- ***Field Testing:*** field testing should be carried out to first integrate a (non-permanent) DGPS base station within the VII DSRC communications network to evaluate whether correction signals can be viably transmitted via the RSE equipment to the vehicles. As part of this field testing, bandwidth and latency calculations should be performed for correction signals that are transmitted over the VII communications infrastructure.

- ***Protocols and Data Handling:*** If it is deemed possible to integrate DGPS corrections with the VII communications infrastructure, then it will be necessary to adapt and develop protocols and DGPS data handling methods for code and carrier phase corrections with RSE;

- ***WAAS Evaluation:*** Further detailed comparisons should be carried out on vehicle positioning accuracy, with a focus on the WAAS architecture and the DGPS base station-RSE integrated approach. It may be that WAAS provides sufficient accuracy when the vehicle is not in direct communications with a roadside unit.

- ***Lane Determination Estimation Techniques:*** these techniques have been developed elsewhere and are described in the literature (e.g., the EDMAP program and [Barth et al., 2005]. However, these lane estimation techniques utilizing 2m positioning accuracy should be further evaluated to determine their suitability for VII California applications.

- ***Degradation Analysis:*** Thus far in this evaluation, the effects of GPS degradation and/or drop outs (tunnels, bridges, urban canyons) has not been examined. This is a crucial step that could have large implications on signal reacquisition and position degradation.

Application Testing: Once a suitable field testing system is developed and initial tested, other VII regional application testing should take place with code phase and carrier phase corrections methodologies.

5. CONCLUSION

This TO 5217 is intended to be an interim report on lessons learned from this figurative midpoint in our testbed deployment. Each section reports lessons learned in the particular coverage area. It is worth noting, however, some top-level lessons to be taken from the interim work: installation of RSE requires a large degree of cooperation between various institutions. Furthermore, cooperation between these institutions must also extend to maintenance and updating hardware. With regard to the to-be-deployed backhaul network, there needs to be additional work to address security, scalability, and conformance to the emerging national VII standard.

Additionally, several applications have 'snapshot' interim conclusions to report:

1. Curve Overspeed Warning System

   The COWS system with integrated on-board sensors, a digital map, GPS and the broadcasted information from the RSE to predict safety speed and provide speed advisory or warning messages to the driver seems implementable on at least an experimental scale.

2. Traffic Probe Data Processing

   The primary challenge with VII probe data processing will transform from determining how to extract as much information as possible from limited available sources of data to determining how to efficiently identify and extract the most relevant and important information from an avalanche of largely redundant data.

3. Application of VII Data on Real-time Arterial Performance

   Arterial information would significantly improve the quality and quantity of traffic data, which in turn will help transportation agencies and all other stakeholders to better understand, plan, and manage the transportation system.

4. Intersection Application

   Experiments with intersections have successfully reconstructed sufficient information to broadcast the signal phase count-down needed for prototype vehicle applications.

5. High Accuracy-Nationwide Differential Global Position System

   The 'jury's still out' with the viability of HA-NDGPS, and subsequent work under TO 6217 should clarify the viability of this.

APPENDIX: THROUGHPUT PERFORMANCE MEASUREMENTS INVOLVING DSRC WIRELESS COMMUNICATION CHANNELS

A.1 Introduction

In order to understand the throughput performance of Dedicated Short Range Communication (DSRC) wireless channels for automotive safety applications, a series of measurements were performed using Denso radio modules and the TechnoCom MCNUs wireless access point. This document explains a method to test the performance of wireless cards in general, provides a benchmark for future performance tests of DSRC wireless radios as well as sets an upper limit for the maximum transfer of information in a DSRC wireless communication channel.

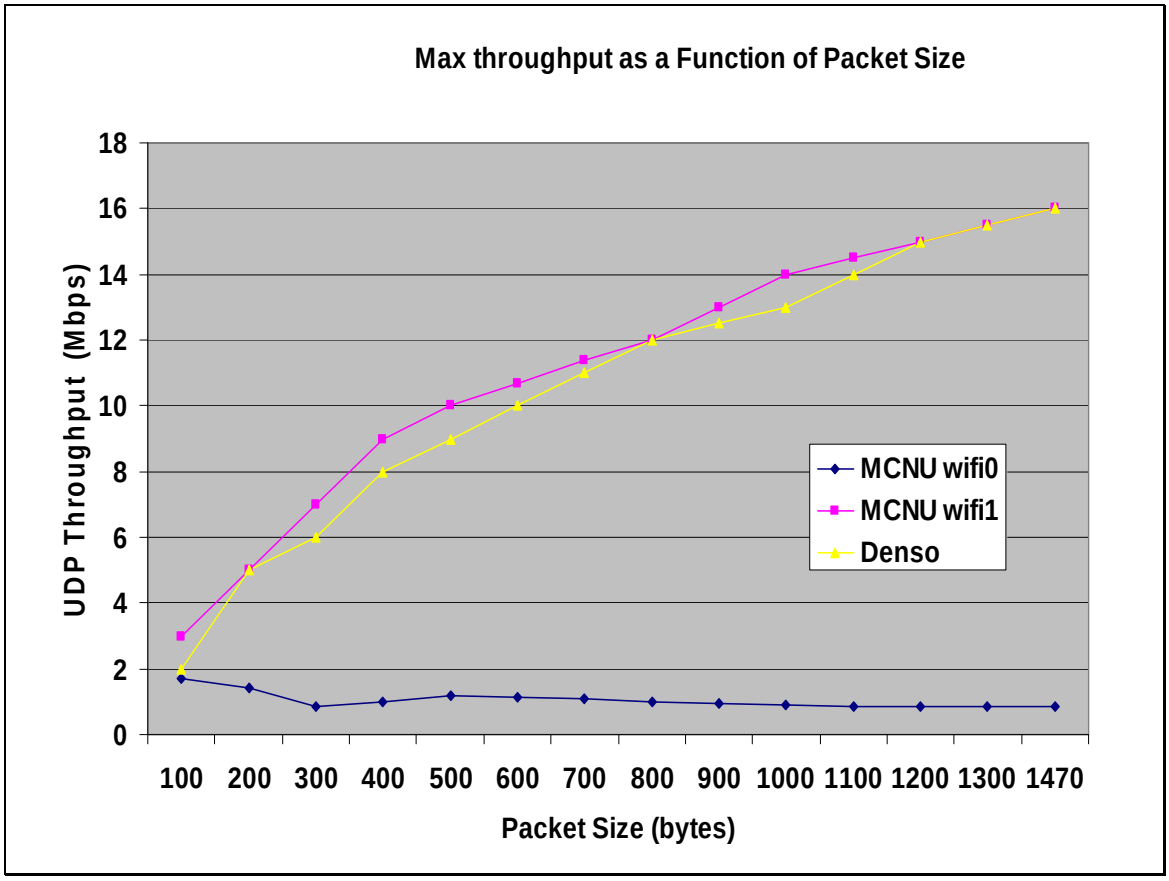
PATH at UC Berkeley is currently a beta-tester for the TechnoCom MCNUs. The MCNU has two wireless network cards identified as Wifi0 and Wifi1. The Denso box has one wireless radio. The maximum throughput (goodput) involving Denso boxes as well as the Technocom MCNU were measured using the free bandwidth network performance tool Iperf (version 2.0.2): http://dast.nlanr.net/Projects/Iperf/.

A.2 UDP Throughput Tests

Iperf was used to measure the maximum UDP packet throughput of the wireless cards as a function of packet size (100 to 1470 byte payload). The throughput is sometimes called goodput because it is the actual payload transfer rate which does not include the headers, preamble or checksum. The maximum throughput in indoor lab conditions for the Denso box and the MCNU WiFi1 wireless card were equivalent. However, the Technocom MCNU wireless card WiFi0 has a reduced throughput of about an order of magnitude (except for packet sizes around 100 bytes). According to TechnoCom, there was a hardware assembly problem involving the bus associated with wireless card WiFi0 (as of November 2007). The Wifi0 card problem should be resolved soon.

Iperf was run on two IBM laptops (windows XP service pack 2) connected to two Denso boxes with 0dB gain antennas approximately 1 foot apart. Transmit power levels from 5 dBm to 18 dBm were used. The TechnoCom MCNU ran a version 1.6rc1 image (Linux Ferora Core 4), and used TechnoCom's wireless drivers versions dot3-1.1.5-570 and dot4-1.1.5-570. Two different types of antennas 1 foot apart and power levels from 5 dBm to 18 dBm were used with no significant difference in performance. The ahdemo mode was used on the MCNU.

A.3 Comparison of UDP Throughput



The figure above compares the maximum UDP throughput of the MCNU wif0 and wifi1 card with the throughput of a Denso Box. The top curve (pink) is the results for the MCNU wifi1 wireless card. The second curve (yellow) is for the Denso box. The difference between the two curves is not statistically significant. The throughput of the Denso box and the MCNU wifi1 card are equivalent. The throughput for the DSRC wireless cards is approximately 2-3 Mbps for 100 byte packets and increases in an approximate linear fashion to approximately 16 Mbps for 1470 byte packets. This result provides a benchmark for future performance tests of DSRC radios.

Conventional 802.11a radios would have a maximum throughput of approximately 24 Mbps for 1470 byte packets. Since the DSRC wireless radio is a modified 802.11a radio with 10 MHz bandwidth (half of the bandwidth of an 802.11a wireless channel), we expect a somewhat lower throughput. For completeness, 802.11b cards have a maximum throughput of approximately 6 Mbps for 1470 byte packets and 802.11g cards have a maximum throughput of approximately 18 Mbps for 1470 byte packets. These values will vary for cards from different manufactures because the wireless card manufactures will use different backoff algorithms (without making it public information).

In an outdoor environment, we would expect the total throughput to decrease as a function of distance and topography. The transmitting wireless card will slow down in response to not receiving acknowledgement packets leading to lower total throughput. Therefore, these results should be considered an upper limit and the best you can expect for a DSRC wireless channel.

The bottom curve (blue) is for the wifi0 card in the MCNU and clearly demonstrates a reduced performance. Except for 100 byte packets, the MCNU throughput is approximately a factor of 10 lower than the Denso boxes and is essentially flat as a function of packet size. The normal behavior would be for the throughput to increase as the packet size increases. Since the preamble, header, and checksum overhead in the packet remains the same with different payload sizes, a larger payload will lead to higher total throughput in an uncompromised system.

A.4 Sample Iperf Commands

The command
iperf -c 192.168.9.1 -u -b 16M -t 10
on the client will attempt to send out 1470 byte UDP packets at 16 Mbps for 10 seconds to ip address 192.168.9.1

iperf -s -u -i 1 will print out results every 1 second on the server
the –l option will set the packet size.

The following error messages shows up in /var/log/messages on the server MCNU Linux box when we run iperf. over

Jan 1 00:06:54 mcnu kernel: wifi0: rx FIFO overrun; resetting
Jan 1 00:07:54 mcnu last message repeated 232 times

The wifi0 wireless card throughput is apparently restricted by a FIFO buffer overflow problem due to a problem with the bus.
This appendix summarizes developments since the 2005 report. The protocols and network architecture changed only slightly after the 2005 World Congress demo (one protocol change was the addition of the message priority field). There are, however, several new components and utilities, described in the following sections.

A.5 Probe client and server components

For the demos conducted in 2006-7, we developed a simple probe application whose in-vehicle component runs the without GPS or CAN-bus information, and whose server component processes the resulting probe messages into driving time messages to be broadcast to back to vehicles traveling the same route as the probe.

A.6 Monitor component

This component is covered in a separate section of the report.

A.7 OBU proxy component support for HTTP over DSRC

The OBU proxy component was originally developed in 2005 to mediate between application level messaging operations (send and receive) and the complexities of RSDUP (fragmentation, retransmission, repetitions, transaction identity). The intended use of the component was exchanging VII-CA messages with VII-CA servers, but after some adaptation it has found a use as a proxy for transporting HTTP requests to web servers and responses back to vehicles. Browser-based applications (interactive maps and navigation, for example) and by other HTTP-based programs now work in-vehicle, as long as it is in range of a RSE antenna. No changes are required in the web client or server software, except that they must be configured to use a HTTP proxy port.

The advantages of proxying HTTP through broadcast UDP, compared with standard TCP, are:

- No delay for dynamic IP assignment (DHCP).
- No delay for TCP setup between vehicle and RSE.
- Symbolic domain name lookups happen at the roadside, so there is no need for vehicle to perform this task before sending messages.
- Works with Denso WRMs.

A.8 Process controller (procon) component

This component keeps other components running, restarting them when necessary. This prevents occasional crashes from bringing the whole system down for more than a few seconds. This also simplifies system administration, because only one program, the procon, needs to be registered with the OS to be started at boot time and to be restarted periodically by the OS if not running.

A.9 Signage injector component

For ease of quickly preparing a demo, we developed a component with a command-line (telnet) interface for injecting signage into the system without writing a special program to talk to the signage server. The program lets you enter the RSU ID, expiration time, and all the standard fields that make up a VII-CA signage message. The message is sent to the server to be downloaded by the rsu component.

A.10 Denso telnet utility

This command-line utility is intended as a replacement for the Denso WRM's built-in telnet interface. It has the following functions:

- Scan for Denso WRMs on the LAN using raw IP packets (useful if the WRM's IP address is not known).
- Telnet into a detected WRM, automatically logging in by supplying the user name and password.
- Execute the full range of WRM commands available in the standard Denso telnet interface.
- Command line editing, history, and completion.
- Automatic detection and repair of WRM configuration problem.

- Automatic re-login after WRM reboot.
- Scripting.
- On-line help.

Note that Denso provides its own utility for performing these kinds of tasks. However, this utility has some disadvantages:

- it is a graphical, and so it cannot be used it remotely.
- it is MS Windows based, and it so cannot be used on a Linux host, such as a typical RSU host device.
- it is (apparently) not scriptable.

The VII-CA Denso telnet utility and some of the other utility modes discussed in this section are documented at:

http://path.berkeley.edu/vii/doc/html/cli.html

A.11 Probe shell

For testing probe display applications, we developed a command-line interface for generating probe messages and sending them to a probe server. The interface allows you to "drive" an imaginary vehicle around the testbed, changing its velocity and heading angle.

A.12 "Remote" mode for control of running components

Since early 2006, the VII-CA software has made it possible to query and reconfigure running programs using a "remote" mode. This mode lets you connect to the program (assuming it has been configured to allow this operation) and:

- view statistics, performance measurements, and the most recent error conditions
- inspect or change any variable in the program state
- load a bug fix into running program
- stop, pause, or restart the program
- subscribe to log messages
- issue any command the program understands
- check which processes are running
- start a process *only* if it is not already running
- encode or decode binary message formats

A.13 Dependency-graph based system administration

Most routine administration tasks performed on the VII-CA testbed no longer require manually logging into the RSE host computer (or server computer) and then editing configuration files, downloading software updates, and restarting processes. These tasks (along with the more repetitive software development tasks, such as packaging, regression testing, and documentation generation) are now automated using a program that understands the dependencies between tasks

and that can detect which tasks have been successfully completed. A software or configuration update can be deployed to the testbed with one command, which performs (in parallel for each destination) the update tasks in the correct order and without redundancy. Even installing VII-CA software on a new site is mostly automated.

REFERENCES:

1. *AASHTO GREEN BOOK - A Policy on Geometric Design of Highways and Streets*, American Association of State and Highway Transportation Officials, 2001.

2. Arnold J. (2006), FHWA, phone interview conducted on September 19, 2006.

3. Ayoub, N., *"Impacts of Oversaturation on Traffic Signal Detector Design and Operations"*, Operating Traffic Signal Systems in Oversaturated Conditions Workshop, Traffic Signal Committee (AHB25), Regional Transportation Systems Management and Operations Committee (AHB10), Washington, D.C., January, 21, 2007

4. Bajikar S., Gorjestani A., Simpkins P., and Donath M.(1997), "Evaluation of in-vehicle GPS-based lane position sensing for preventing road departure," *Proc. IEEE Conf. Intelligent Transportation Systems*, Boston, MA, Nov. 9–12, 1997, pp. 397-402.

5. Barth, M, J. Du, J. Masters (2005) "Differential GPS Evaluation for Enhanced Probe Vehicles" *Final Report submitted to the California Center for Innovative Transportation and the California PATH Program*, 93 p., 2005**.**

6. Ben-Joseph, E., Szold, T.S., "*Regulating Place: Standards and the Shaping of Urban America*. Routledge", 2005

7. Bertini, R., and Tantiyanugulchai, S. "Transit Buses as Traffic Probes", *TRR 1870*, 2003, pp. 35-45.

8. Blomenhofer H., et al. (1994), "Development of a real-time DGPS system in the centimeter range", *IEEE 1994 Position Location and Navigation Symposium*, Las Vegas, NV, USA, pp. 532-539.

9. California Department of Transportation, *"C-8 170 Signal Controller Manual"*, Sacramento, California 1983

10. California Alliance for Advanced Transportation Systems, "California Vehicle Infrastructure Integration Workshop", October 2005, Santa Clara, CA, see http://www.caats.org/events/vii/vii_workshop.pdf.

11. CAMP, *Enhanced Digital Mapping Project Final Report*, United States Department of Transportation, Washington, DC, 2004.

12. CAMP Vehicle Safety Communications Consortium*, Vehicle Safety Communications Project Task 3 Final Report - Identify Intelligent Vehicle Safety Applications Enabled by DSRC*, U.S. Department of Transportation, National Highway Traffic Safety Administration, 2005.

13. Cash, B. "DSRC Training Workshop", Annapolis, Maryland, Dec 7-8, 2004.

14. Chang, J., Cohen, D. Blincoe, L., Subramanian, R., and L. Lombardo (2007). "CICAS-V Research on Comprehensive Costs of Intersection Crashes" (Report No. 07-0016), National Highway Traffic Safety Administration, Washington D.C. Available from http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/esv/esv20/07-0016-O.pdf.

15. Cheu, R., H. Qi and D.-H. Lee, "Mobile Sensor and Sample-Based Algorithm for Freeway Incident Detection", *TRR 1811*, 2002, pp. 12- 20.

16. Cops, M. "VII Strategy for Safety and Mobility", *The Third ACM International Workshop on Vehicular Ad Hoc Networks (VANET 2006),* Los Angeles, California, September, 2006.

17. CSRC, California Spatial Reference Center, California Real Time Network, http://sopac.ucsd.edu/projects/realtime/ accessed November 2006.

18. Daganzo, C.F. Fundamentals of Transportation and Traffic Operations. Elsevier Science Ltd. 1997

19. Dion, F., Rakha, H., and Kang Y.S., *"Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections",* Transportation Research Part B 38 (2004) 99-122, Elsevier Science Ltd.

20. Du, J., J. Masters, M. Barth. "Next Generation Automated Vehicle Location Systems" in press, *IEEE Transactions on Intelligent Transportation Systems*, 2006.

21. Du, J. "Vehicle Lane Position Estimation for Link Level In-Vehicle Navigation and Automatic Vehicle Location" Ph.D. Dissertation, UC Riverside Electrical Engineering, December 2005.

22. Du, J., J. Masters, M. Barth. "Lane-Level Positioning for In-Vehicle Navigation and Automated Vehicle Location (AVL) Systems," *2004 IEEE Conference on Intelligent Transportation Systems*, Washington, D.C., October 2004.

23. Du, J. and M. Barth. "Bayesian Probalistic Vehicle Lane Matching for Link-Level In-Vehicle Navigation," *IEEE Intelligent Vehicle Conference*, Tokyo, Japan, June 2006.

24. Farrell J. and M. Barth. *The Global Positioning System and Inertial Navigation: Theory and Practice*, ISBN-0-07-022045-X, McGraw-Hill, January 1999.

25. Farrell J.and Givargis T. "Differential GPS Reference Station Algorithm: Design and Analysis", *IEEE Trans. Control Systems Tech*. V8 (3), pp 519-531, 2000.

26. FHWA, "High Accuracy-Nationwide Differential Global Positioning System Test and Analysis": *)"*, Federal Highway Administration, US Department of Transportation, Phase II Report Publication No. FHWA-HRT-05-034 July 2005.

27. FHWA, "Phase I High Accuracy-Nationwide Differential Global Positioning System Report", *)"*, Federal Highway Administration, US Department of Transportation, FHWA-RD-02-110, July 2002

28. FHWA, *"VII architecture and functional requirement (Version 1.1)"*, Federal Highway Administration, US Department of Transportation, Washington, D.C., July, 2005

29. FHWA, Vehicle Infrastructure Integration (VII) National System Requirements, *)"*, Federal Highway Administration, US Department of Transportation Version 1.2.1, February 2007.

30. FHWA. Vehicle Infrastructure Integration: All Architecture and Functional Requirements, Version 1.1, FHWA, ITS Joint Program Office, July 20th, 2005

31. Florida Department of Transportation, District 5, "Surface Transportation Security and Reliability Information System Model Deployment: iFlorida - Final Development Plan", Prepared for US DOT-FHWA, Cooperative Agreement Number DTFH61-03-H-00105, January 29th, 2004

32. Gumstix, Inc, "Products", http://www.gumstix.com/products.html, accessed in May 2007

33. Gutman, S., Wagoner S., Codrescu, M., and Fuller-Rowell, T; National Oceanographic and Atmospheric Administration; "The Use of Atmospheric Models to Improve Differential GPS Positioning Accuracy" 1999.

34. Hall, R. and Vyas, N. "Buses as a Traffic Probe", *TRR 1731*, 2000, pp. 96-103.

35. Hauschild, M. "Weather and Road Surface Information and Hazard Warnings: Data Content Acquisition Through Advanced Probe Vehicle Systems", *Twelfth World Congress on Intelligent Transport Systems*, San Francisco, November 2005, Paper 2337.

36. Henry, D. "VII Strategy for Safety and Mobility", *ITS Michigan Wireless Communications Seminar,* Brighton, Michigan, March, 2007

37. IEEE P1609.5, Guide for Wireless Access in Vehicular Environments (WAVE) –WAVE System Communications (2006)

38. ITS Joint Program Office, "Vehicle Infrastructure Integration (VII) Probe Message Processes, Version 1.0, April 11, 2006.

39. ITSA, Intelligent Transportation Society of America, "Primer on Vehicle-Infrastructure Integration", October, 2005, available at http://www.itsa.org/resources.nsf/Files/VIIWhitePapers/ITSAVIIPrimerFinal.pdf.

40. ITS America Library Resources, "Innovative Mobility Showcase – VII", 2005, http://www.itsa.org/IMS_VII.html

41. Kee C. and Parkinson B., "Wide Area Differential GPS as a Future Navigation System in the US," *IEEE 1994 Position Location and Navigation Symposium*, Las Vegas, NV, USA, pp. 788-795.

42. Krishnan, H. "Vehicle Safety Communications Project," CAMP Vehicle Safety Communications Consortium, February 15, 2006, http://www.sae.org/events/ads/krishnan.pdf.

43. Kurihara, T.S. DSRC Security Project, P1556, Presentation to DSRC Task Group within the IEEE 802.11 Working Group, September 2003

44. Last, David, Grant, Alan, Williams Alwyn, Ward, Nick, "Enhanced accuracy by regional operation of Europe's new radiobeacon differential system", *Proceedings of the Institute of Navigation,* ION-2002, Sept 2002.

45. Li, M., Wu, G., Li, Y., Bu, F., and Zhang, W.B., "*Active Signal Priority for Light-Rail Transit (LRT) at Grade-Crossings*", Transportation Research Record, Washington, D.C., May, 2007.

46. Li, M., Zhang, W.B., Zhou, K., Leung, K., and Sun, S., "*Parsons Traffic and Transit Laboratory (Parsons T² Lab)*", ITS World Congress, Beijing, China, October, 2007

47. Larson, G., McKeever, B. "VII California Testbed Development Plans", presented at Institute of Transportation Engineers Annual Meeting, August 2006, Milwaukee, Wisconsin, available at http://itsCalifornia.org/2006%20pdfs/NorCalVIITestBedWhitePaper060404.pdf

48. Love, A.W. "GPS, Atomic Clocks and Relativity, Potentials", IEEE Volume 13, Issue 2, April 1994 pp 11-15

49. Magellan, Differential Global Positioning System, http://professional.magellangps.com/en/products/aboutgps/dgps.asp, accessed October, 5 2006.

50. Michigan Department of Transportation. "VII Concept of Operations: VII Data Use Analysis and Processing", Mixon/Hill of Michigan, Inc. a Report prepared for Michigan Department of Transportation, June 2007

51. Misener, J.A., Shladover, S.E. "PATH Investigations in Vehicle-Roadside Cooperation and Safety: A Foundation for Safety and Vehicle-Infrastructure Integration Research," *Proceedings of the IEEE Intelligent Transportation Systems Conference*, Toronto, Canada, September 2006.

52. Mouskos, K., et.al., "Transportation Operations Coordinating Committee System for Managing Incidents and Traffic: Evaluation of the Incident Detection System", *TRR 1679*, 1999, pp. 50 – 57.

53. *National Electrical Code*, Section 725.55, 2002 Edition

54. Navstar, "Navstar GPS Interface Control Document", Report #ICD-GPS-200B-PR, Joint GPS Program Office, 1991.

55. Noblis, "Vehicle-Infrastructure Integration (VII) Probe Data Characteristics: Analyses of Three Field Data Sets", Draft Technical Report to USDOT, March 2007.

56. Odawara, Y., "Current Status and Future Prospects for the Utilization of Probe Data in Road Administration", *Eleventh World Congress on Intelligent Transport Systems*, Nagoya, October 2004, Paper 3313.

57. Otto, H.U, et al., *Actual and dynamic map for transport telematics applications*, IST-2001-34141 ActMAP project, deliverable D 3.1, NextMAP Consortium, 2003.

58. Pandazis, J.C. *Final Report*, IST-1999-11206 NextMAP project, deliverable D1, NextMAP Consortium, 2002.

59. Park, D., et.al., "Estimating Travel Time Summary Statistics of Larger Inervals from Smaller Intervals Without Storing Individual Data", *TRR 1804*, 2002, pp. 39 – 47.

60. Remondi, "Phase I High Accuracy-Nationwide Differential Global Positioning System Report", FHWA-RD-02-110, 2003.

61. RTCM, Radio Technical Commission for Maritime Services Special Committee No. 104, "RTCM Recommended Standards for Differential Navstar GPS Service, V2.3", RTCM Paper 136-2001/SC104-STD, August 20, 2001.

62. RTCM, Radio Technical Commission for Maritime Services Special Committee No. 104, "RTCM Recommended Standards for Differential NAVSTAR GPS Reference Stations and Integrity Monitors (RSIM), V1.1", RTCM Paper 12-2001/SC104-248, May, 2001.

63. Saricks, C, et.al., "Evaluating Effectiveness of Real-Time Advanced Traveler Information Systems Using a Small Test Vehicle Fleet", *TRR 1588*, 1997, pp. 41-48.

64. SatNav, "WAAS is Commissioned", *SatNav News*, Volume 21, November 2003, pp. 1

65. Sethi, V., et.al., "Arterial Incident Detection Using Fixed Detector and Probe Vehicle Data", *Transportation Research Part C, Volume 3C*, No. 2, April 1995, pp. 99 – 112.

66. Shladover, S., "Probe Vehicle Data Needs and Opportunities", *Thirteenth World Congress on Intelligent Transport Systems,* London, October 2006.

67. Shladover, S. and T. Kuhn, "Evaluation of Traffic Probe Data Processing Alternatives", Fourteenth World Congress on Intelligent Transport Systems, Beijing, October 2007.

68. Society of Automotive Engineers. SAE J2735 Dedicated Short Range Communications (DSRC) Message Set Dictionary, VIIC Document No. App190-01, December 2006

69. Society of Automotive Engineers, Surface Vehicle Recommended Practice (Draft) J2735 – Dedicated Short Range Communications (DSRC) Message Set Dictionary, March 21, 2007.

70. Srinivasan, K., and P. Jovanis, "Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network", *TRR 1537*, 1996, pp. 15-22.

71. *Standard Specifications,* State of California, Department of Transportation, May 2006

72. *Standard Plans,* State of California, Department of Transportation, May 2006

73. SubCarrier System Corp (SCSC), Society of Automotive Engineers (SAE), *"Draft SAE J2735 Dedicated Short Range Communications (DSRC) Message Set Dictionary"*, March, 2007.

74. Sun, J., H. Wen, J. Guo and X. Deng, "Study on the Data Fusion of Floating Car and Sensors Based on Fuzzy Theory", *Twelfth World Congress on Intelligent Transport Systems*, San Francisco, November 2005, Paper 3393.

75. Tango F., and A. Saroldi, *Final Project Report*, IN-ARTE Project TR 4014, deliverable D 1.1, IN-ARTE Consortium, 2001.

76. Thomas, N. and B. Hafeez, "Simulation of an Arterial Incident Environment with Probe Reporting Capability", *TRR 1644*, 1998, pp. 116-123.

77. Trimble, "Summary of GPS Error Sources"**,** http://www.trimble.com/gps/howgps-error2.shtml#3, accessed October 5, 2006.

78. Tsuge, M. and M. Arai, "Practical Use of Floating Car Data and Its Benefits", *Eleventh World Congress on Intelligent Transport Systems*, Nagoya, October 2004, Paper 3441.

79. USCG NAVCEN, "High Accuracy DGPS Program", Jim Radice 2001 www.navcen.uscg.gov/cgsic/meetings/summaryrpts/40thmeeting/06Radice40Pres.ppt, accessed October 5, 2006.

80. USCG NAVCEN, US Coast Guard Navigation Center http://www.navcen.uscg.gov/AccessASP/NDgpsReportAllSites.asp, accessed October 5, 2006.

81. U.S. Department of Transportation, "IEEE 1609 - Family of Standards for Wireless Access in Vehicular Environments (WAVE)", Intelligent Transportation Systems Standards Fact Sheet, Jan 9, 2006, http://www.standards.its.dot.gov/fact_sheet.asp?f=80.

82. USDOT, U.S. Department of Transportation and U.S. Department of Defense, "2001 Federal Radionavigation Plan", December, 2001.

83. USDOT, U.S. Department of Transportation, ITS Joint Program Office "Vehicle Infrastructure Integration (VII) Architecture and Functional Requirements, Version 1.1", Technical Report, July 2005.

84. USDOT, U.S. Department of Transportation, ITS Joint Program Office, Vehicle-Infrastructure Integration website: http://www.its.dot.gov/vii/vii_concept.htm, accessed July 2006.

85. USPDD, "Presidential Decision Directive NSTC-6" of March 28, 1999.

86. VII CALIFORNIA, "VII California Team Meeting", Oakland California March 1, 2006.

87. VII CALIFORNIA, "VII California Team Meeting", Richmond Field Station California June 8, 2006.

88. VII CALIFORNIA, "VII California Team Meeting", Oakland California November 2, 2006.

89. VII CALIFORNIA, VII California Program, see http://www.viiCalifornia.org.

90. Wevers, K., Blervaque, V. *MAPS & ADAS Safety enhanced digital maps and standard interface to ADAS*, Proceedings of the 11th World Congress on ITS, Nagoya, Japan, 2004.

91. Wilson, "David L. Wilson's GPS Accuracy Web Page" http://users.erols.com/dlwilson/gpswaas.htm, accessed October 5, 2006.

92. Wolfe, D., Judy, C., Haukkala, E., Godfrey, D, "Implementing and Engineering an NDGPS Network". *Proceedings of the Institute of Navigation*, ION-2000, Sept 2000.

93. Wolfe D.B, "Nationwide DGPS: 2003 and Beyond", *Proceedings of the Institute of Navigation*, 2003 National Technical Meeting Proceedings, January 22-24, 2003.

94. Xie, C., Cheu, R.L, and Lee, D.H., *"Improving Arterial Link Travel Time Estimation by Data Fusion"*, Proceedings of the 83rd Transportation Research Board Annual Meeting, Washington, D.C., 11-15 January, 2004.

95. Yin,J., ElBatt, T., Yeung, G., Ryu, B., Habermas, S., Krishnan, H., Talty, T. "Performance evaluation of safety applications over DSRC vehicular ad hoc networks," Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks, Philadelphia, PA, USA, pp 1 – 9, 2004.