Title

Safetrip-21: Connected Traveler

Authors

Sengupta, Raja
Misener, Jim
Ahern, Katherine
et al.

Publication Date

2010-03-01

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

Safetrip-21: Connected Traveler

Raja Sengupta, et al.

California PATH Research Report
UCB-ITS-PRR-2010-23

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 6615

April 2010

ISSN 1055-1425

SAFETRIP-21: CONNECTED TRAVELER

Task Order 6615

FINAL REPORT

Raja Sengupta, Principal Investigator
Jim Misener, Project Manager

Katherine Ahern
Ching-Yao Chan
Somak Datta Gupta
Jerry Jariyasunant
Jing-Quan Li
Christopher Long
Eric Mai
Christian Manasseh
Christopher Nowakowski
Jessical O'Connell
Shahram Rezai
Michael Steelhorst
Liping Zhang
Wei-Bin Zhang
Kun Zhou
Xeusong Zhou

University of California, Berkeley
1357 S. 46th Street, Bldg. 190; Richmond, CA 94804-4648

March 2010

Contents

List of Figures:

1 Executive Summary

The US DOT RITA Volpe Center entered into a cooperative agreement with the California Department of Transportation (Caltrans) to establish the inaugural SafeTrip 21 field test site in the San Francisco Bay area [named Connected Traveler]. Specifically, the site encompasses I-880 from Oakland to San Jose on the east bay and from San Jose to just south of the San Francisco International Airport, along U.S. 101 and California State Route (SR) 82. The site includes the SR-84 Dumbarton Bridge toll crossing, which links I-880 and U.S. 101.

Caltrans's partners include the Metropolitan Transportation Commission, the University of California-Berkeley's California Partners for Advanced Transit and Highways (PATH), the California Center for Innovative Transportation, Nokia, Inc., NAVTEQ, Santa Clara Valley Transportation Authority, and Nissan. The cost of the $12.4 million field test was funded by private and public sector partners, including the USDOT contribution, equally. (Public Roads, September/October 2008)

The Networked Traveler is part of the Connected Traveler component of the U.S. DOT (Research and Innovative Technologies Administration, RITA) SafeTrip-21 initiative. SafeTrip-21 aims to expand and accelerate the U.S. DOT IntelliDrive initiative, and SafeTrip-21 builds upon research into the use of sophisticated information, navigation, and communications technologies to further national transportation goals. The Connected Traveler effort contains the several California Department of Transportation (Caltrans)-led projects under the SafeTrip-21. The Networked Traveler component is one of the partner projects, and roughly described, it has three components:

1. A technology exploration which acknowledges that a 3G wireless- and WiFi-connected smartphone is the currently available means to connect travelers and mobility- and safety-related *information-driven* and *personalized* information. This technology exploration was manifested in a multi-partner, multi-application transit bus where safety, mobility, transit- and road-systems operations transportation services was delivered. This effort was initiated in 2008, and it culminated in a demonstration of these applications, delivered in a New York City transit bus application. **It is primarily this – the vision, creation, development and demonstration – that this nomination is focused.**

2. The 2008 planning (and in 2009, delivery) of a field test enabling evaluation of the validity of the hypothesis that safety can be effected by providing situational awareness over a longer foresighted horizon, e.g., "slow traffic ahead" via consumer-held smartphones. As a second component of this field test is to use the same smartphone application and arterial travel data, plus real time transit schedule information, to capture the effect and desirability of smartphone-"pushed" data on transit users. Both portions will occur in the San Francisco Bay Area.

3. The 2008 planning (and again in 2009, delivery) of a field test where parking availability information and transit travel time for rail (Caltrain) and bus (SamTrans), traffic conditions, and travel time information for freeways and major arterial highways was delivered in real-time, to support travelers' trip decisions and to encourage the use of alternative modes of

transportation. The goal is to promote efficient use of the existing transportation infrastructure, and in turn, reduce congestion and associated vehicle emissions.

Component 1 (SafeTrip-21 Networked Traveler Development and Demonstration) has opened the gates on this project for Components 2 (Safety and Mobility Field Test and Evaluation) and 3 (Parking Information and Mode Switch Field Tests and Evaluation). More importantly, the Networked Traveler Development and Demonstration has opened the world's eyes on how [vision] x [near term technology] x [implementation] provides a product which in the near- to mid-term could transform how travelers receive information and, ultimately, behave.

Indeed, the Networked Traveler project is aimed at bringing information to the traveler, and then to the transportation system *now.* While it primarily addresses the need for transformational change in safety and mobility services desired by systems operators and travelers alike, it bootstraps off the observation that the web is here, that mobile and connected consumer electronic devices are here, and the dire need for safety and mobility information is evident.

A major element to the work was the Networked Traveler World Congress Development and Demonstration, given in New York City in November, 2008. To conceive and deliver the demonstration resulted in development of three main services. "Tell me about my trip" assists trip planning with traffic information, transit connections, and driving choices for an eco-route and a fastest route. "Tell me about my route" can provide travelers with real-time road-safety conditions, real-time traffic and parking conditions, schedule-driven transit information, real-time GPS-based transit status, road signage. "Watch out for me!" includes services such as the pedestrian-to-vehicle safety alert, the vehicle-to-pedestrian safety alert, road-to-vehicle road safety information, road hazard alerts, and work zone alerts.

Within this framework, a plethora of smartphone-delivered services were provided: Pedestrian alerts allowed slow-moving pedestrians to signal drivers to watch out; a work zone alert signaled phones on the demo bus to slow down for its approach to the cone area; centralized, real-time transit information helped virtual commuters meet their trains and buses on time; and smart parking simplified a modal switch by helping drivers find and reserve available parking in real-time. Other items in the demo included updated travel time estimates, next-stop alerts, and a hydrocarbon-savings calculator for transit riders and speed zone and signal priority alerts for drivers.

The objective was to show what could happen in a Networked Traveler ethos, then put it on the road to illustrate and punctuate the reality. Another objective is that the research team associated with Networked Traveler was able to learn significantly and the mid-term result of bringing selected applications into a field evaluation was brought closer to reality.

2 Background and Introduction

This research agreement addresses the field test elements of the Caltrans-US DOT cooperative agreement for SafeTrip-21 research award and therefore builds upon the technical foundation – using smartphones to deliver critical safety and mobility information.

2.1 Urgency, Payoff Potential, and Implementation

The urgency and payoff potential are linked and therefore high: the proposed research combines near term enabling situational awareness and transit applications from public, private and academic stakeholders. The research holds promise to combine the resources and commitment of the public sector, with private sector innovation, to provide multiple means to deliver safety and mobility applications for travelers on arterials and freeways alike for system management and traveler information, for multiple modes (transit, as well as passenger vehicles), and for multiple applications (to include commercial). It does not place all resources or "bet" on DSRC (or cellular or WiFi); it enables all. Importantly, it addresses safety improvement goals enabled by 'connectivity' with the advent of ubiquitous wireless 'smartphones' and therefore pervasive information-rich communications.

In light of how leading high-tech companies do business today, it is clear that fast-prototyping and an evolutionary approach are replacing central planning and long development cycles. It is the program team's belief that key safety goals can be readily explored and a rich dataset provided for independent evaluation. The concept of delivering to drivers multiple safety services via GPS-equipped cell phones carries important public benefits of utmost relevance to Caltrans. Additionally, the concept of providing transit data for traffic, and for delivering to other travelers, notably transit riders, connectivity information by smartphones, is also of importance and relevance to Caltrans.

2.2 Research Objective

The primary field test component will evaluate the hypothesis, "Smartphones can help people drive more safely." Thus, the primary deliverable of was a dataset enabling evaluation of the validity of the hypothesis. The dataset will record the experience of drivers using the connected traveler field test system. The field test system was comprised of a:

- software client downloaded by the driver onto the driver's smartphone,
- server that will support the client with information,
- roadside sensors that will provide information to the server, and
- networking services to connect the driver's smartphone to the server.

We expect a field test participant to provide his or her own phone and data plan in anticipation that the system will deliver sufficiently high value. We aim to design the field test system so that the driver perceives it to frequently deliver value. This is because, in the absence of cash incentives, drivers who do not perceive value or who are annoyed by the client are likely to react by turning it off or uninstalling it.

Likewise, to minimize distribution costs, we will implement a distribution strategy based on the perceived value of the field test system and will work to acquire participants from specific groups or associations.

Based on the field test hypothesis and the objective to minimize distribution costs by offering sufficient perceived value to drivers and media, we have conceived the following service package to provide the field test system to each participant:

- An origin-to-destination routing service. This was multi-criteria and based on that demonstrated by us at the ITS World Congress in New York. The user was able to customize routing based on the desire to travel with reduced emissions, reduced trip time, or reduced variance.
- Augmentation of this service by a "smart push" functionality able to actively monitor conditions on the route and pro-act to offer the driver alternative routes if conditions change significantly. Current systems require the traveler to "pull" information, whereas smartphones can enable the "smart push."
- Bundle the prime objective of the field test, the safety component, into the smart push. This module will provide the driver with situational awareness of approaching road hazards, such as slow traffic queues, incidents, and non-recurrent congestion. Recurrent congestion is covered by the routing service

The three services are provided via a single software download to be executed by the participant from the connected traveler website.

2.3 Research Plan and Deliverables

The field test was executed in two phases, resulting in the schedule and milestones in the table below. Phase 1 (Connected Traveler Testing) is the primary focus of this project and will directly address the safety field test, in addition to providing the SafeTrip-21 transit elements. Phase 2 (Exploratory Testing) focuses on the 'next step' services, conducting the pedestrian safety and eco-driving aspects of SafeTrip-21, with smaller scale data collection. Because of the exploratory nature of Phase 2, while the field testing are goals, the detailed definition is contingent on success of the development and experimental installation. This is in contrast to Phase 1, which is the primary focus and with explicit by-task deliverables.

Table 2-1: Task Listing

Task No.	Task	Start Date	End Date	Milestone
Phase 1, Task 1 (1.1)	Phase 1: Safety System Development	January 1, 2009	March 15, 2009	Phase 1 launch
1.2	Phase 1: Outreach	January 1, 2009	March 15, 2009	Phase 1 launch
1.3	Phase 1: Safety Benefit Feasibility Assessment	January 1, 2009	November 30, 2009	
1.4	Phase 1: Execute Safety Field Test	March 15, 2009	November 30, 2009	Phase 1 dataset
1.5	Phase 1: Transit System Development	January 1, 2009	March 15, 2009	Phase 1 launch
1.6	Phase 1: Transit Benefit	January 1,	November 30,	

	Feasibility Assessment	2009	2009	
1.7	Phase 1: Execute Transit Field Test	March 15, 2009	November 30, 2009	Phase 1 dataset
Phase 2	Phase 2: System Development and Field Tests	March 15, 2009	November 30, 2009	Phase 2 launch
Task 3	Project Management and Reporting	January 1, 2009	December 31, 2009	Reports / Final Report

3 Safety System Development

3.1 Experimental Design

3.1.1 Background

A California PATH research team is preparing the deployment of safety applications under a Networked Traveler project, in conjunction with the US DOT Safe Trip 21 efforts. This document describes the premise, hypotheses, rationale, and the approach for conducting field experiments, with the goal of assessing the validity and usability of the proposed safety applications.

3.1.2 Premise

If drivers are better informed of the traffic conditions in their driving environment, they was more aware and better prepared to take actions to avoid hazardous situations. Within the scope of this study, we focus on the "soft" safety applications that have a relatively longer time window to provide drivers with alerts. A "hard" safety application, by comparison, requires an immediate action. For example, a system that warns drivers of imminent freeway front-end collisions with another car in front will need to take effect within 1-5 seconds. On the other hand, one example of the "soft" safety applications that we propose to offer is the situational awareness alert that can be effective in a 10-60 second time window. For example, in a situation where drivers cannot see the slow or stopped traffic beyond a curved roadway ahead, at a distance of 1 mile or less before the driver reaches at location, we can issue an alert to the driver especially when the driver's subject vehicle is traveling significantly faster than the traffic queue ahead.

3.1.3 Safety Applications

A suite of applications are being considered and to be deployed for the planned field experiments. Most of these applications are designed for freeway driving conditions, which are the primary test targets in this stage of the Networked Traveler project. The core list of safety applications are given in the diagram below, and they are explained summarized as follows:

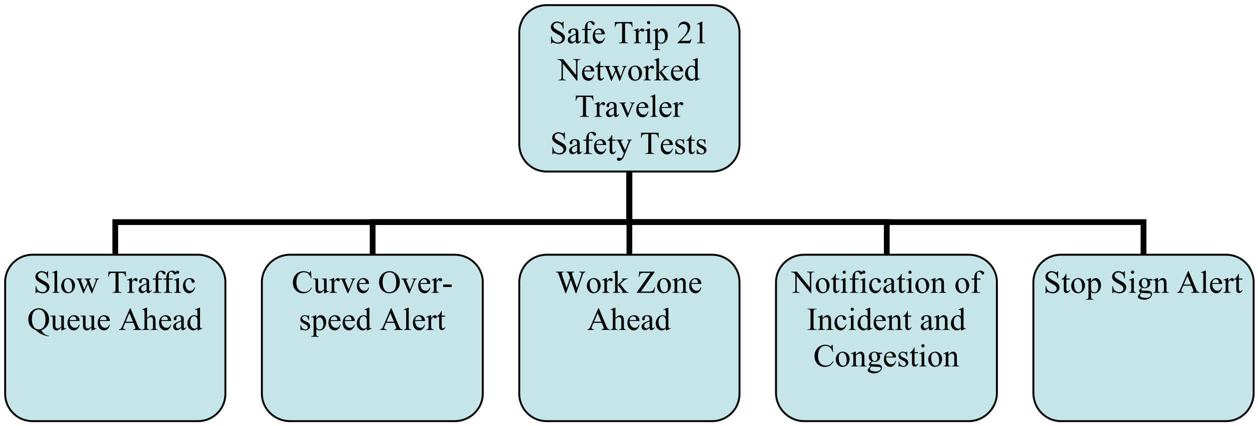


Figure 3-1: Safe-Trip 21 Safety Applications

3.1.3.1 <u>Situational Awareness of Slow Traffic Queues</u>

The first application is a situational-awareness function that provides alerts to drivers on driving conditions within a relatively short-latency time horizon, in the range of 10-60 seconds. For this application, we focus on specific highway locations where past traffic and crash patterns indicate that a Connected-Traveler system can offer useful alerts so that drivers can take tactical actions to avoid crashes, or in essence increase their 'safety alert time horizon' to beyond the just several seconds based on driver reaction or provided by active safety systems. In these situations, by offering a timely "slow or stopped traffic ahead" message the system can effectively inform drivers of roadway hazards to reduce chances of crashes and realize safety benefits.

The research team has investigated the collision data for the last ten years and identified some potential sites for this application. Based on the results of initial screening, a preliminary field survey was carried out in recent weeks. A list of candidate locations, their attributes and maps are given in Appendix A.

3.1.3.2 <u>Curve Over-speed Alert</u>

The second application is a curve over-speed warning, which can also be considered part of the situational-awareness function in Application One above. For this application, we focus on specific highway locations where users can benefit by being reminded to reduce speed as they approach certain roadway segments. Some of these segments are off-ramps from freeways, including a few sites identified in Appendix A. In these situations, by offering a timely "slow down for sharp curves ahead" message the system can effectively inform drivers of roadway hazards to reduce chances of crashes and realize safety benefits.

3.1.3.3 <u>Work Zone Alert</u>

Another application is a work-zone-ahead advisory, which can be applicable in a relatively short-time horizon such as Application One and Two above, just prior to a user reaches work zone areas. This application can also be provided in a relatively longer time frame, for example 5-20 minutes before the expected user route passes through ongoing work zones. In the former case, users can benefit by being reminded to reduce speed and be cautious as they approach work zones. In the latter case, the users can make strategic decisions to change routes to avoid passing through a restricted or congested area. In these situations, by offering a timely "work zone ahead" message the system can effectively inform drivers of roadway hazards to reduce chances of crashes and realize safety benefits.

3.1.3.4 Notification of Incidents and Non-recurrent Congestion Ahead

One other application is the notification of incidents and non-recurrent congestion on the road ahead to the drivers based on real-time traffic data. The major premise of this application of "Incident and Congestion Alert" is as follows:

(1) First of all, drivers can stay informed of roadway traffic conditions, which allow them to make trip planning choices. The information was "pushed" to users based on their current positions and travel routes.
(2) Secondly, in congested areas on highways, various hazardous scenarios may develop and lead to increased likelihood of collisions. With the proposed application, an earlier notification alert to the drivers, in the range of 2-30 minutes can offer drivers opportunities to take tactical and strategic actions, including:
    - Reducing speeds with increased awareness of slow traffic ahead
    - Choosing to change routes to avoid further trip delays
    - Changing transportation modes by switching to transit or travel plans, with benefits of reducing traffic demands on flow-stressed or incident-impaired roadway segments

3.1.3.5 Stop Sign Alert

Another application is a stop-sign advisory, which can be applicable for off-ramp or local street locations, just prior to a user reaches a stop sign. This application relies on the availability of stop sign database. In these situations, by offering a timely "stop sign ahead" message the system can effectively inform drivers of roadway hazards to reduce chances of crashes and realize safety benefits.

3.1.4 Safety Applications Test Hypotheses

For each of the applications that are to be deployed for the field tests, the hypothesized outcome and the expected user responses and the safety impacts are described and listed in Table 3-1.

Table 3-1: Safety Application and Test Hypotheses

Application	Applicable Situations	Hypothesized Outcome	Expected Driver Response and Safety Effects
Slow Traffic Ahead	• Traffic queues ahead of curved roadway with limited visibility • Collisions are caused by vehicles approaching too fast toward the end of	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to slow down in advance • Drivers can make cautious approach before reaching end of queue by receiving alerts • Drivers benefit by an elevated	• Drivers provide favorable assessment of the alerts • Drivers respond positively and noticeably to the alerts • Drivers reduce speed earlier or more

	slow traffic queue	sense of safety and comfort from the alerts even if the drivers can take timely, evasive actions without the alerts	significantly with the alerts than without the alerts • Drivers make lane change maneuvers to other lanes to avoid slow queue
Slow Traffic Ahead	• Off-ramp queue buildup and spillover into freeway • Collisions are caused by vehicles approaching too fast toward the end of slow traffic queue	• Same as above	• Same as above
Slow Traffic Ahead	• Severe traffic weaving section • Collisions are caused by vehicles sideswiping or rear-ending other vehicles are making moving across lanes in the weaving section	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to make cautious approach before reaching the weaving section • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the drivers can take timely, evasive actions without the alerts	• Same as above
Curve Over-speed Warning	• Curved roadway with tight turns, including off-ramps • Collisions are caused by vehicles moving too fast into the curve	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to make cautious approach before reaching the curve • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the drivers can take timely, evasive actions without the alerts	• Same as above
Work Zone Ahead	• Reduced-speed segments due to work zone • Road reconfigurations due to work zone • Some collisions are caused by vehicles moving too fast into slow traffic in work zone • Other collisions are caused by changes in traffic patterns due to work zone configurations	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to make cautious approach before reaching work zone • Drivers benefit by choosing alternative routes if alerts are given early enough for them to exit and to avoid problematic areas • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the drivers can take timely, evasive actions without the alerts	• Drivers provide favorable assessment of the alerts • Drivers respond positively and noticeably to the alerts • Drivers reduce speed earlier or more significantly with the alerts than without the alerts • Drivers make lane change maneuvers to other lanes to avoid work zone • Drivers change routes to avoid work zone

Notification of Incident On Route	• Potential slow traffic or congestion induced by incidents • Some collisions are caused by vehicles moving too fast into slow traffic in incident-induced congested areas • Other collisions are caused by stressed traffic conditions • Other collisions are caused by changes in lane configurations or traffic patterns	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to make cautious approach before incident segments • Drivers benefit by choosing alternative routes if alerts are given early enough for them to exit and to avoid problematic areas • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the drivers can take suitable actions without the alerts	• Drivers provide favorable assessment of the alerts • Drivers respond positively and noticeably to the alerts • Drivers reduce speed earlier or more significantly with the alerts than without the alerts • Drivers make lane change maneuvers to other lanes to avoid incident areas • Drivers change routes to avoid incident areas
Notification of Non-recurrent (unexpected) Congestion On Route	• Congestion caused by all probable causes unexpected by users • Some collisions are caused by vehicles moving too fast into congested areas • Other collisions are caused by stressed traffic conditions • Other collisions are caused by changes in lane configurations or traffic patterns	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to make cautious approach before congested segments • Drivers benefit by choosing alternative routes if alerts are given early enough for them to exit and to avoid problematic areas • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the drivers can take suitable actions without the alerts	• Drivers will prefer to receive alerts about non-recurrent congestion only • Recurrent congestion if issued frequently will appear to be nuisance therefore result in negative feedback • Drivers provide favorable assessment of the alerts • Drivers respond positively and noticeably to the alerts • Drivers reduce speed earlier or more significantly with the alerts than without the alerts • Drivers make lane change maneuvers to other lanes to avoid congested areas • Drivers change routes to avoid congested areas
Stop Sign Alert	• Stop sign at off ramps or on local streets • Collisions are caused by vehicles moving too fast into stop signs	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to make cautious approach before stop signs • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the	• Drivers provide favorable assessment of the alerts • Drivers respond positively and noticeably to the alerts • Drivers reduce speed earlier or more

		drivers can take suitable actions without the alerts	significantly with the alerts than without the alerts

There are a multitude of common elements given in the table above. They can be reorganized into the charts below. The first chart is a diagram showing the suite of applications.

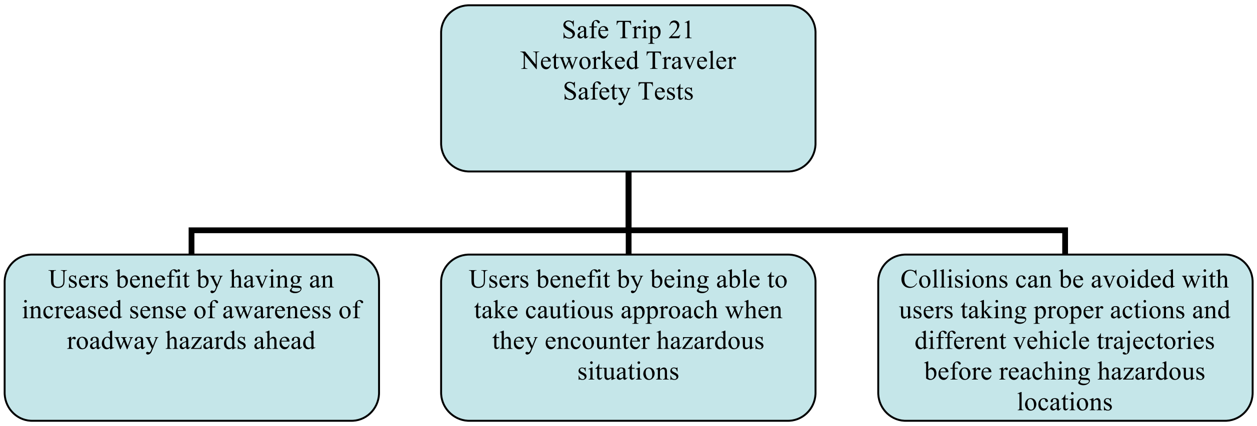


Figure 3-2: Hypothesized Expected Outcomes of Safe-Trip 21 Safety Applications

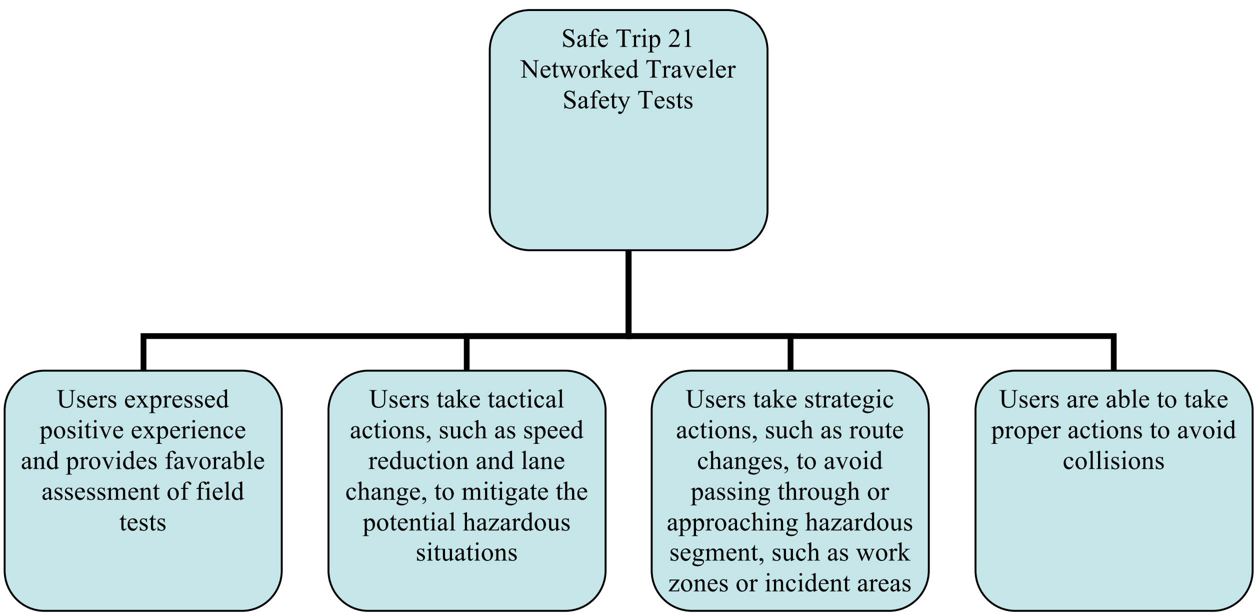


Figure 3-3: Expected User Responses and Safety Impacts of Safe-Trip 21 Safety Applications

3.1.5 Validation of Test Hypotheses

3.1.5.1 <u>Safety Application and Expected Test Outcome</u>

The previous section provides an overall description of the test hypotheses of the expected outcome. In the course of the field tests, it is expected that user experience was evaluated and field data was collected to explore the user needs and preferences as well as design validity and shortcomings of the safety applications. Therefore, experimental design of the field tests should emphasize on the observations of user response, qualitatively and quantitatively, to establish the foundation for making such assessment.

The diagram below shows the expected cause-and-effect sequence in the process of experimental designs. The top layer is the planned safety field tests. The second layer is the applications to be deployed. The third layer is expected user actions, if safety alerts are effective. The fourth layer is the expected observable outcome, as a result of the safety field tests.

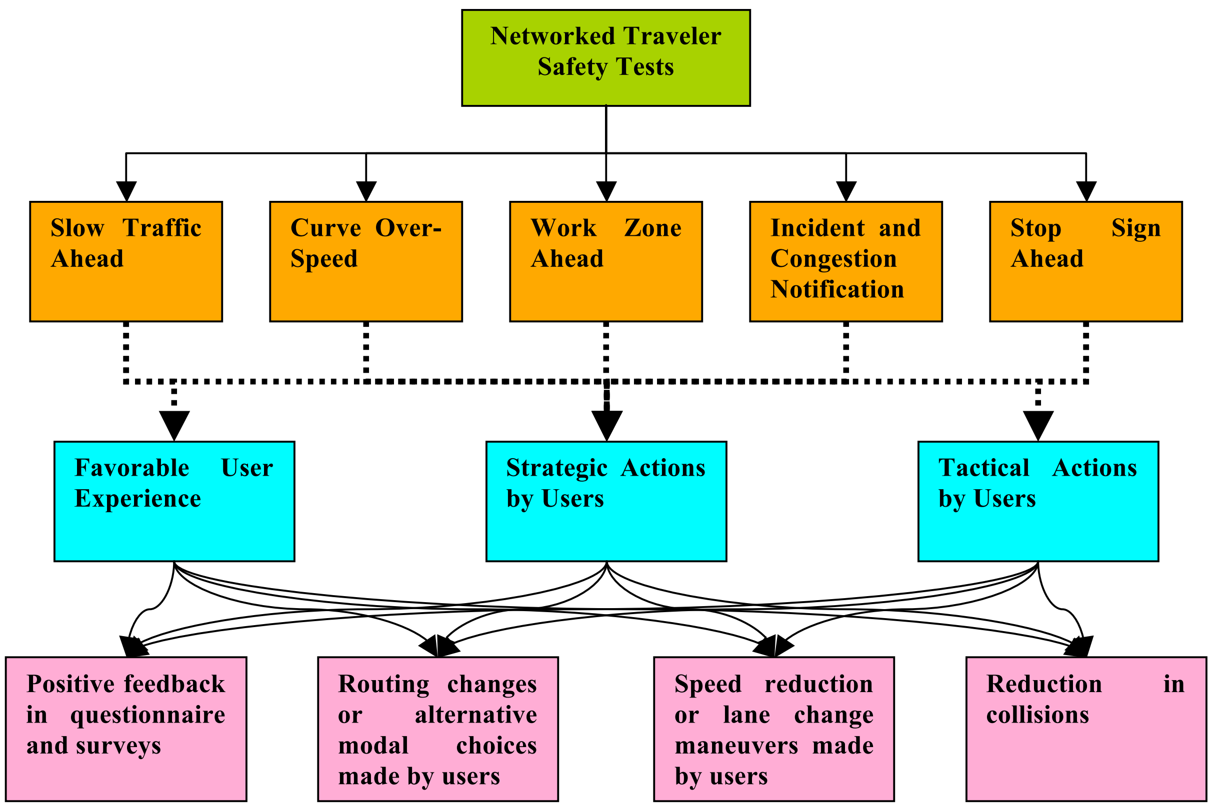


Figure 3-4: Safe-Trip 21 Safety Applications and Observable Outcome

3.1.5.2 __Functional Processes and Operating Constraints__

Before discussing further the necessary data collection and the availability of observable data to validate the test outcome, it is critical to point out the constraints of the data acquisition process within the context of the planned field tests.

3.1.5.2.1 Functional Process of Safety Applications

The Networked Traveler project is based on the concept of utilizing existing technologies that are currently available to consumers, such as GPS-enabled smart phones and personal digital appliance. The communication links may include 3G, Wi-Fi, and DSRC. The current set up envisioned for the planned tests can be described as follows:

(1) Users register for the service.
(2) Users enter origin and destination for the intended trip.
(3) Users activate services by choice and by preference.
(4) Once activated, the user's route and personal preference information are communicated to the system server.
(5) System server receives traffic speed and incident information for the whole test area (San Francisco Bay Area similarly covered by the 511 services).
(6) System server offers routing suggestions to users and monitor user positions by receiving users' current positions periodically.
(7) System maintains and updates traffic, incident and relevant information for the entire test bed area and executes safety algorithms for individual users based on their current positions and speeds.
(8) If any condition on a user's route warrants the issuance of safety alerts, the system server issues and sends an alert to the users; alerts are presented to users in auditory forms while routing information is available in both visual and auditory forms.

3.1.5.2.2 Constraints in Data Acquisitions

The following constraints in both quality and quantity of data acquisition should be noted:

(1) System server received traffic and incident information from other sources (Traffic.com and SpeedInfo); therefore the update rate and the quality of traffic data are not within the control of the application developers.
(2) Traffic information are only available for designated segments and specific locations in the test bed area, therefore the measurements of background traffic settings are not available continuously in time or in space.
(3) In this pilot test, only 13 locations are considered for "slow queue ahead" applications. Even though sensors was installed to capture the traffic conditions at these locations, they can mostly provide traffic speed measurements for a specific site or a short segment for the selected locations, and may not entirely reflect the actual representation of traffic parameters for ideal and robust processing and execution of alert generation algorithms.
(4) Users may or may not activate safety functions even if they volunteer and register for the services.
(5) Users may choose to activate selective functions of their preferences and thus only a subset of functional results and associated data are available for individual users.
(6) User positions are only available through the GPS coordinate on user's devices, and therefore the accuracy and availability of user trajectory (speed and position) depend on the GPS units and the settings of driving environment, which may vary significantly along users' routes.
(7) There is no measurement about the movements of surrounding vehicles, which may have impacts on the driver actions when alerts are issued. Therefore, the causal effects of driver actions in response to alerts may not be determined.
(8) There is only limited information about the traffic conditions at test sites. For example, in the "slow-queue-ahead" scenarios, only the average speed of existing traffic queue is available to be used as a criterion in alert-generation algorithms. Therefore, insufficient

description of the actual traffic settings may not explain whether drivers can positively respond to the alerts.

(9) There are multiple variables involved in the causation of collisions, including driver (fatigue, distraction, incorrect judgment, etc.), environment (visibility, roadway surface conditions, weather, etc.). A significance depth and broad spectrum of these related parameters are not available for assessment for this pilot test. This, it should be reiterated that this pilot test is only targeting and designing for a limited subset of situational awareness scenarios.

(10) The pilot test is further constrained by the fact that alerts are communicated to users through user interface implemented within the capabilities and flexibility allowed by the phone-based devices.

3.1.5.2.3 Measure of Effectiveness

The planned field tests of the suite of safety applications is on a limited scale, and considered a pilot test as it is to be carried out within a limited period of performance (for the year of 2009) and scope. However, it is still important to establish the framework and methodology to conduct the system assessment toward the end of the pilot test so that effectiveness and usefulness of safety applications can be properly measured.

Having discussed the limitations of the planned tests in the previous section, it can be further explored about the data to be collected and how they can be analyzed and investigated to assess the effectiveness of the suggested applications. The following table illustrates how a matrix of measure of effectiveness can be constructed.

Table 3-2: Anticipated Test Outcomes and Measure of Effectiveness

Expected Test Outcome and Driver Responses	Measures of Effectiveness (MOE)	Parameters and Variables to Assess MOE
Public awareness of safety campaign	• Spectrum of project partnerships	• List of partners in project • Scope of participation by partners • List of participating organizations outside of project team
	• Scope of community participation	• Number of participating users • Number of data samples collected in field tests • Percentage of positive feedback by users
	• Outreach efforts	• Sessions of activity reports held in public forums and conferences • Technical papers presented • Reports of media events

Favorable user experience and positive user feedback	• Willingness to participate and to maintain continual use of applications	• Number of participating users • Periods of active usage • Continuity and frequency in activating applications • Percentage of positive feedback by users
	• User feedback to surveys and questionnaire on - Function usefulness - Function acceptability - Timeliness of alerts - User interface friendliness	• User answers in surveys and questionnaires (to be detailed and designed later)
Strategic actions (routing changes or modal choices) by users	• Correctness and reliability of alert generation	• Numbers of alerts generated • Percentage of valid alerts • Conditions of false alerts (time of day, incident type, congestion status, travel time)
	• Timeliness in alert issuance	• Total latency in alert reception time versus design point in algorithms • Conditions of time lags (time of day, incident type, congestion status, travel time) • User feedback in real time or in surveys
	• User routing changes after alerts are issued • User choices of modal changes after alerts are issued	• Route records captured before and after alert issuance • User inputs or responses upon reception of alerts • Percentage of user responses • Percentage of confirmed and verified user responses - based on user real-time feedback - based on user surveys • Conditions of alert issuance (time of day, incident type, congestion status, travel time)
	• Timeliness in user response actions	• Noticeable trajectory changes (to be analyzed and defined later according to

		field GPS data resolution and fidelity) in time versus alert issuance point • Driver reaction time comparison to total latency in alert generation • Conditions of alert issuance (time of day, incident type, congestion status, travel time)
Tactical actions (speed reduction and/or lane change maneuvers)	• Correctness and reliability of alert generation	• Numbers of alerts generated • Percentage of valid alerts • Conditions of false alerts (time of day, speed differential, congestion status)
	• Timeliness in alert issuance	• Total latency in alert reception time versus design point in algorithms • Conditions of time lags (time of day, speed differential, congestion status, incident type) • User feedback in real time or in surveys
	• User speed reduction after alerts are issued	• Speed change maneuvers captured before and after alert issuance, for at least a 60-second time window before and after • Speed variations in time segments within the data window • Magnitude of speed change in various time segments • Percentage of speed reduction responses • Percentage of confirmed and verified user responses - based on user real-time feedback - based on user surveys • Conditions of alert issuance (time of day, congestion status, travel time, incident type)
	• User lane change after alerts are issued	• Lane change maneuvers captured before and after alert issuance, for at least a 60-second time window before and after • Trajectory variations in time segments within the data window • Percentage of lane change responses • Percentage of confirmed and verified

		user responses - based on user real-time feedback - based on user surveys • Conditions of alert issuance (time of day, congestion status, travel time, incident type)
	• Timeliness in user response actions	• Noticeable trajectory changes (to be analyzed and defined later according to field GPS data resolution and fidelity) in time versus alert issuance point • Driver reaction time comparison to total latency in alert generation • Conditions of alert issuance (time of day, congestion status, travel time, incident type)
Collisions avoided and reduced	Reduction in • Collision frequency (per counts of collisions) for individual site or collision types • Collision rate (per vehicle-mile traveled) for individual sites or crash types	• Collision reports • Cumulative counts of collision numbers at specific sites and particular types • Realistically difficult to observe in short time spans due to data stability and non-deterministic causal factors

3.1.6 Data Collection and analysis

3.1.6.1 Application Field Test Schedule

Table 3-3 provides a list of milestones and targeted applications for this phase of Safe Trip 21 field deployment tests.

Table 3-3: ST-21 Networked Traveler Field Test Initial Milestones

Milestone Date	Rollout Functionality	Precipitating Events
15 March	'Slow traffic ahead' from NAVTEQ Traffic / 511 Note: *Maximum leverage of World Congress Trip Planner (PATH, Univ. of Utah, NAVTEQ components)*	100s drivers recruited – 1. If expedited CHPS approval not completed: pilot subjects (100s) 2. If expedited CHPS approval completed: SFTMA, AAA, Stanford Commuter Club, MTC

		(listed in order of probability)
15 April	Add: SpeedInfo and ID'd high concentration locations	100s drivers recruited – 1. Expedited CPHS approval expected. 2. SpeedInfo completes their already-committed support.
15 June	Add: Feedback and learning	100s drivers recruited – Full-featured.

3.1.6.2 __User Recruiting__

We will recruit users in three phases, several hundred per phase (early Spring, late Spring, early Summer), by working with management from four organizations:

- SF Transportation Management Association
- Metropolitan Transportation Commission
- AAA (Northern California Automobile Association) and
- Stanford Commuter Club (given in order of priority in recruitment).

Each user was required to have a cell phone and an unlimited data plan. We will also recruit drivers sign up ad hoc to our webpage and wish to download the client application. The targeted number of users is in the range of 500-1000 at this planning stage. However, it should be noted that since this project offers no monetary compensation for participants, other than the incentive of providing a potentially useful applications for interested users, the actual number of users is difficult to predict and control.

3.1.6.3 __Data Collection Period__

As outlined in the application rollout schedule and milestones above, the safety applications was made available in late spring. Corresponding to this schedule, the data collection was implemented in several stages:

3.1.6.3.1 Quantitative Data

(1) Collection of baseline data

After the initial installation and activation of field test functions, typical routes and user experience was captured to establish the baseline data without the activation of safety alerts. This step is taken to ensure the applicability of GPS data in the functional algorithms with virtual alerts generated along the typical routes taken by each user. The baseline data was used in later stage of data analysis to examine the probable effects of safety applications after they are activated.

For users who take regular commute routes or follow consistent driving patterns, the four-week period should provide a sample of approximate 20 data points for each user at specific sites, such as those intended for Application “Slow traffic ahead” or “Curve over-speed alert”. For other applications, such as notification of incident or congestion or work zone, the number of data samples could be 3-10 times greater, depending on the actual traffic conditions. Note, however, the alerts may not be generated each time the users pass through the designated locations or where incidents or work zones are present, therefore, the sample sizes may be reduced substantially. For those occasions when no alerts are warranted, the driving records was kept as the baseline data for normal driving conditions.

This period of baseline data verification is expected to continue for 4 weeks, or until a valid sample of data points are established.

(2) Collection of user data in response to safety applications
After the initial validation period, the data collection will continue as long as the user opts to activate the functions in his driving routines. If a user signs up in the early stage of field test, the data collection period can go on for 5-6 months before the conclusion of the field tests. If the data collection is continuous and un-interrupted, then the numbers of data samples are expected to be 5-6 greater than the validation period.

For users who pass through the specific sites daily, this will means an approximate sample of 100 data points for each user. For non-location specific applications, the sample size can be several times greater. Note, however, the alerts may not be generated each time the users pass through the designated locations or where incidents or work zones are present, therefore, the sample sizes may be reduced substantially. For those occasions when no alerts are warranted, the driving records was kept as the baseline data for normal driving conditions.

3.1.6.3.2 User Survey and Questionnaire

(1) User information at registration
All users are required to register when they sign up for the application services. In this registration process, certain questions about the users was posed. Answers to some questions are required, and others are optional. For example, to assess the coverage of user base, the driving distance and zip codes for origins and destinations of regular routes was useful information to have in this registration process. The detailed form of questions was provided later.

(2) On-Line Feedback
Users was given the option of providing anytime feedback on problems encountered in the use of the applications as well as desired changes or suggestions on the applications that are offered.

(3) Mid-term Survey
Three months into the initial use of the field tests, each user was required to go through a web-based survey. This survey was an initial assessment of user experience on the safety applications.

(4) Final Survey
One month before the project is concluded, users was asked to go through another survey. This was another milestone to assess the user experience as well as to observe any noticeable changes

in user experience after exposure to the applications for an extended period. After the final survey, unless the user opts to discontinue the service, data will continue to be collected, which may be valuable for later evaluation of the field tests.

3.1.6.4 ***Types of Quantitative Data***

The system server, where the monitoring of traffic conditions and alert generation algorithms are executed, will continuously capture available data streams to facilitate later data analysis. The types of data that can potentially be collected include the following:

- Time of alerts issued
- Traffic and incident Status that trigger the activation of alerts, such as
    - Traffic speed variations in space and in time at the point of alert generation
    - Incident type and severity
    - Congestion status or travel time through congested segments
    - Distance from user location to incident location
- Contents of alerts presented to users
- Driver response to alerts
    - Real-time voice commands response to alerts
    - Delayed online submission of user feedback
- Trajectory data (GPS) of users before, during, and after alerts, from which additional data may be derived:
    - Post-alert braking or lane-changing responses
    - Speed variations in space and in time before, during, and after alert reception

3.1.6.5 ***Types of Qualitative Data***

Qualitative data to assess user subjective experience of the applications was collected through surveys and online feedback. The types of data that can potentially be collected include the following, but the exact form and questions of survey was developed later:

- Overall impression of applications
    - Usefulness
    - Timeliness
    - Reliability
    - Issues or problems in using applications
- Driver background information
    - Age
    - Gender
    - Familiarity or experience with smart phones
    - Driving distance daily or weekly
- Driver experience with specific applications
    - Number of alerts received daily or weekly
    - Perceived value of individual applications
    - Specific problems encountered with individual applications

3.1.6.6 ***Data Analysis***

The purpose of data collection and analysis is several-fold:

(1) To assess the effects of intended applications on users,
(2) To provide supporting evidence in determining the extent of success in project objectives, and
(3) To explore the weakness and shortcomings of implemented functions for future improvements

3.1.6.6.1 Quantitative Data Analysis

The quantitative data captured during the field tests was grouped and categorized to allow the determination of user responses.

The assessment should be performed for both the individual application and system as a whole. The elements of data analysis should include:

- Number of alerts generated
- Distribution of alert generation with traffic conditions:
    - Time of day
    - Speed differentials
    - Congested state (estimated travel time delays)
    - User speed
    - Speed differential
    - Distance of user location versus incident location
- Correlation of alert provision with user responses:
    - User baseline or normal driving conditions
    - Variations in driver trajectory
    - Statistical verification of behavior changes by comparison of before and after user experience
    - Exploration of functional forms representation of selective driver responses in terms of meaningful explanatory variables
    - Determination of critical variables on driver behavioral changes

3.1.6.6.2 User Response Data

In order to verify the statistical significance of data representation, several critical data elements must first be scrutinized.

(1) Validity of using GPS data for monitoring user trajectory

The GPS data is phone based and not vehicle mounted, therefore the trajectory traces expressed by the GPS unit do not fully reflect the vehicle actions. The data gathering process is further complicated by the resolution and accuracy of GPS data. Thus, several steps can be taken to evaluate the usage of such data:

- Preliminary laboratory and field tests can be performed to collect sample data sets for initial evaluation
- If necessary, filtering and data processing techniques can be applied to improve the usability of such data.
- A large sample of data sets wascome available after the first roll-out of applications. At that stage, certain roadway segments (tunnels or valleys) of the highway network was

discovered to have data issues. These data sets can be excluded from data analysis.

(2) Use of GPS data and user actions of tactical maneuvers such as speed change and lane change in response to alerts

Depending on the conditions of alert issuance, the users may or may not take immediately noticeable actions. For short-time frame alerts such as slow traffic queue ahead, the user response is unknown but expected within 60 seconds. For other applications, the time horizon is much longer, and the Therefore, one approach for dissecting the user follow-up actions is as follows:

- Continuously monitor the user trajectory for 60 seconds prior to the alert and right after the alert
- Divide the observation time window into multiple time segments
- Computer the speed differential or lane-change maneuvers within each time segments
- For evaluation of all users,
    - Computer the average and standard deviations of speed changes in each time segment
    - These variations are then compared to the baseline of all users
    - Statistical significance tests can be performed to determine if the variations in trajectory is meaningful at different time windows after the issuance of alerts
- For evaluation of individual users, if sufficient data samples exists,
    - Computer the average and standard deviations of speed changes in each time segment
    - These variations are then compared to the baseline of individual users
    - Statistical significance tests can be performed to determine if the variations in trajectory is meaningful at different time windows after the issuance of alerts

(3) User response evolution over time

At this planning stage, it is difficult to estimate whether sufficient data points was collected for individual users to evaluate their progressive acceptance and thus differences in response to the alerts. If such data become available for selective users, it was possible to examine their response in different stages over time.

3.2 Estimation of Sample Size

3.2.1 Problem Description

Sample size, in the context of this report, refers to the number of observations that are targeted for the evaluation of safety field experiments to be conducted for the Networked Traveler project. For this study, we are interested in assessing the outcome of experiments measured by the responses of users in reaction to the safety alerts. The required number of samples that allow statistically meaningful evaluation of experimental outcome depends on a number of parameters, including population size, confidence levels, statistical power desired, and the distribution of outcome variables.

3.2.2 Population Size

In the intended experiments, the population size means the total number of users who are exposed to the traffic conditions during their travel on a test site where there exists a traffic situation warranted for the issuance of alerts. An exemplar calculation of the population size is suggested as follows.

In the networked Traveler experiments, the safety alerts are applicable to a fraction of the driving population passing through a specific test site during the daily rush hours. The capacity of a freeway lane during the peak hours is normally between 1,500 and 2,000. Let us further assume that only 10% of the driving public is exposed to situations that warrant the issuance of alerts. This will amount to approximately 500 to 1,000 cases for each day or 10,000 to 20,000 per month. Over a 9-month test period, the total population size was in the range of 100,000 to 200,000. Our intent is then to define the necessary sample size or number of observations that was sufficient for us to assess the outcome of the safety application for this particular test site.

If the safety application is applicable to a larger fraction of the driving public, thent eh population size was much greater. For example, for all moving traffic a few miles upstream of a congested area or a work zone, the situation was generally applicable to traffic across multiple lanes.

As was seen in later discussions, the sample size becomes stabilized and does not change significantly when the population size increases to a relatively large number. For the illustration of sample size calculation, a large number was used in this document to derive a conservative estimation of sample size.

3.2.3 Experimental Outcome Variables and Their Distribution

The sample size calculation is a function of the types of variables and associated distribution of these variables. For example, we was interested in whether users will respond positively to a survey of the usefulness of an alert. The sample size needed for a dichotomous outcome (yes or no response) will depend on the expected distribution or ratio of positive and negative answers and the statistical parameters that are chosen. Examples are provided in a section below.
If the experimental outcome is a continuous variable, then similarly the sample size was affected strongly by the distributions of such variables. For example, after a safety alert is issued, we was interested in whether the user takes actions to reduce speed within a defined time window. For such evaluation, we was comparing the "before" and "after" behaviors of the users under similar situations. The "before" data are from the baseline data that are collected without the activation of alerts, while the "after" data are cases when alerts are provided to users. The change in speed will vary by individual users as well as by the corresponding traffic conditions. If the expected change in speed is a random sample with independent observations, then it can be assumed that a large population of such samples will approach a normal distribution. For the examples of sample calculation given in a section below, a set of normalized parameters, including the mean difference between before and after and the standard deviation, are used to show how samples sizes vary according to the chosen statistical parameters.

3.2.4 Sample Size and User Base

Multiple safety applications, including work zone, incident notification and curve over-speeding, are planned in the field experiments. The observations or the samples for specific applications should be categorized and separated for the purpose of evaluation as it can be expected that user responses to different applications can vary. For the alert-generation scenarios that are applicable to a wide area or a large of sites on different freeways, such as incident or work-zone notification, it can be expected that the required sample sizes can be reasonably achieved. Therefore, the challenge lies in the collection of sufficient observations for more restrictive, site-specific cases that are applicable to particular traffic conditions.

The current projected number of users to be recruited for the field experiments is in the range of several hundreds to one thousand. For most scenarios as illustrated in the section below, the user base should meet the targets of required sample sizes. It was noted, however, that there is no direct association of user pool size and the number of observations because the issuance of alerts depend on the traffic conditions at the time when the users pass through the test sites. Therefore, it remains to be seen how significantly large the number of observations can be collected for a selective subset of the intended applications.

Expected User Response and Safety Effects	**Assumptions**	**Sample Size**
1.Drivers provide favorable assessment of the alerts	1.Dichotomous (Yes/No) Outcome 2. Margin of error =5 % 3.confidence level =95 % 4. Population size=200000 5. Response distribution=50 %	385
2. Drivers respond positively and noticeably to the alerts	1.Dichotomous (Yes/No) Outcome 2. Margin of error =5 % 3.confidence level =95 % 4. Population size=200000 5. Response distribution=50 %	385
	1.Continuous Outcome(reaction to notice and response (two different outcomes)) 2.α (Type I error probability)= 0.1,0.05,0.1,0.001 3. δ(the mean difference of the pairs)=0.5 4. Power= 0.8 5.σ (standard deviation)=3	alpha power 0.6 0.7 **0.8** 0.9 --- 0.1 131 171 225 310 **0.05** 179 224 **286** 381 0.01 292 349 426 540 0.001 458 529 622 759
3. Drivers reduce speed earlier or more significantly with the alerts than without the alerts	1.Continuous Outcome 2.α (Type I error probability)= 0.1,0.05,0.1,0.001 3. δ(the mean difference of the pairs)=0.5 4. Power= 0.8 5.σ (standard deviation)=3	alpha power 0.6 0.7 **0.8** 0.9 --- 0.1 131 171 225 310 **0.05** 179 224 **286** 381 0.01 292 349 426 540 0.001 458 529 622 759
	1.Dichotomous(Yes/No) Outcome 2. margin of error =5 % 3.confidence level =95 % 4. population size=200000 5. response distribution=50 %	385
4. Drivers make lane change	1.Dichotomous(Yes/No) Outcome 2. margin of error =5 % 3.confidence level =95 %	385

3.2.5 Sample Size Table

The table above provides exemplar sample size calculations. Further explanations of parameters are provided in sections below.

3.2.6 Explanation of Parameters

3.2.6.1 __Dichotomous Outcome__

a) **The margin of error** is the amount of error that one can tolerate. If 90% of respondents answer *yes*, while 10% answer *no*, one may be able to tolerate a larger amount of error than if the respondents are split 50-50 or 45-55.

Lower margin of error requires a larger sample size.

b) **The confidence level** is the amount of uncertainty you can tolerate. Suppose that we have 20 yes-no questions in a survey. With a confidence level of 95%, we would expect that for one of the questions (1 in 20), the percentage of people who answer *yes* would be more than the margin of error away from the true answer. The true answer is the percentage we would get if we exhaustively interviewed everyone.

Higher confidence level requires a larger sample size.

3.2.6.2 __Continuous Outcome__

a) Description: Pair wise analysis is when we do two measurements on a single sample and then compare the outcome of the two measurements. Mostly a time factor is involved, a measurement is done, something "happens", an "intervention" for example, after which the measurement is done again. In this case, the "*before and after alert*" speed measurements are compared.

b) **Distribution**: In the case of speed means or averages, the speed for each individual before alert is subtracted from the speed measurement after the alert is issued. These differences are for all individuals added together producing a mean difference **(δ)** with an associated standard deviation (**σ**). Our null hypothesis is that the average is zero; overall (in net terms) the respondents did not change the speed. The sample size calculated is the sample size required to detect a postulated net change over all individuals. The expected mean difference, or net change, for all individuals and the associated standard deviation are assumed as given in the table. A sample size for the paired t-test is then calculated.

3.2.7 Technical Basis of Equations used

For calculating the sample size for dichotomous outcome, the following equation is used:
The sample size n and margin of error E are given by

$$x = Z_{(c/100)}^2 r(100-r)$$

$$n = {}^{Nx}/_{((N-1)E^2 + x)}$$

$$E = \text{Sqrt}[{}^{(N-n)x}/_{n(N-1)}]$$

where N is the population size, r is the fraction of responses that we are interested in, and $Z_{(c/100)}$ is the critical value for the confidence level c.

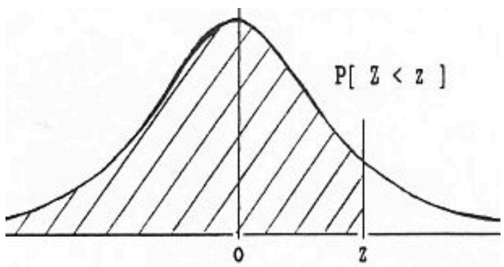
The assumptions made while using the above equation are:

1. The number of positive responses follows a normal distribution.
2. The above calculation assumes that you have more than about 30 samples.
3. The population size is assumed to very large.

A similar statistical example has been explained below.

Suppose that we was required to construct a 95% confidence interval for the proportion so that it will have a 5% length of confidence interval. What sample size would be required?
A conservative estimate of the standard error of the sample proportion is $\sqrt{0.5*0.5/n}$ .Since the length of the confidence interval should be 5 %, the 95 % confidence form for the proportion would take the form

$$\textit{Sample proportion} \pm 2.5\%$$



As the confidence interval is 95 %, the area required (in the above diagram) .95+.025= 0.975. Using Normal Distribution Tables, the corresponding z-value is found as 1.96.

Let $\hat{p}$ be the estimator of proportion. Then,

$(\hat{p} -1.96*\sqrt{\frac{.5}{n}} , \hat{p} + 1.96*\sqrt{\frac{.5}{n}}$) is the 95% confidence interval. Hence it follows that,
1.96*(Estimate of Standard Error) =0.025

Or equivalently

$$n= \left[1.96*0.5/.025\right]^2=1537$$

The sample size calculation above assumes that the sample size is small relative to the population size. If, however, we would like to incorporate a finite population correction adjustment, then we would have to incorporate the following adjustment:

To determine the sample size needed for a simple random sample, obtain half the harmonic mean of the population size and sample size calculated for a random sample with replacement.

The required sample size of 1537 can be adjusted by incorporating the population size. For a village of 10,000 residents, we would have to obtain a sample of size:

$$= \frac{1}{\frac{1}{10000} + \frac{1}{1537}} = 1332$$

For a town of 100,000 residents, the required sample size is:

$$= \frac{1}{\frac{1}{100000} + \frac{1}{1537}} = 1514$$

While for a country of 70,000,000 residents, the required sample size would be:

$$= \frac{1}{\frac{1}{7000000} + \frac{1}{1537}} = 1537$$

This is same as that obtained for sampling with replacement. This numerical result explains why nationwide polls typically use only 1500 to 2000 respondents. The important point to note is that large population size has virtually no effect on the choice of the sample size.

3.2.8 Variation of Sample Size with respect to Different Parameters

*3.2.8.1 **Dichotomous Outcome***

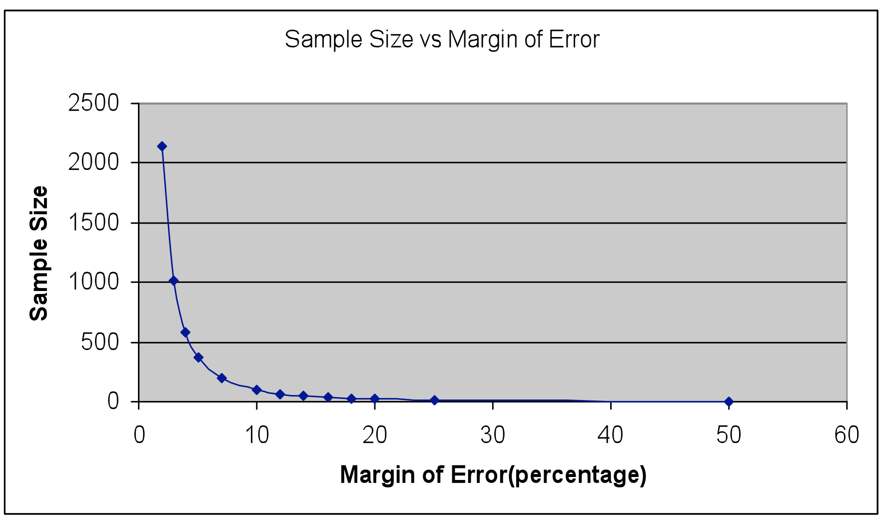


Figure 3-5: Sample Size versus Margin of Error

This graph shows that sample size is more sensitive to the margin of error when the margin of error is in the range of 0-5 %. It also indicates that choosing margin of error greater than 10 % does not give a conservative sample size. In order to have a conservative and large sample size, the margin of error should be in the range of 0-5 %.

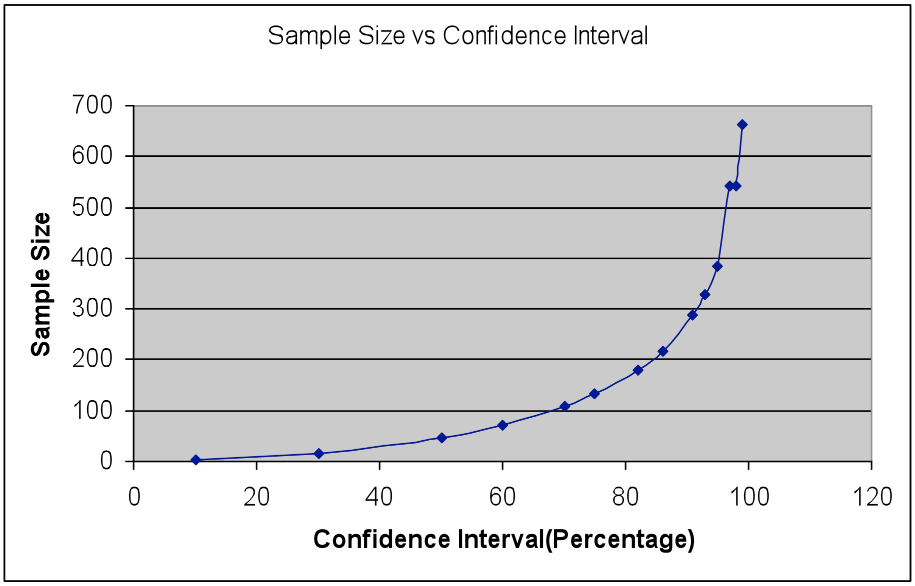


Figure 3-6: Sample Size versus Confidence Level

The graph above indicates that 90-100% confidence interval yields a conservative sample size range.

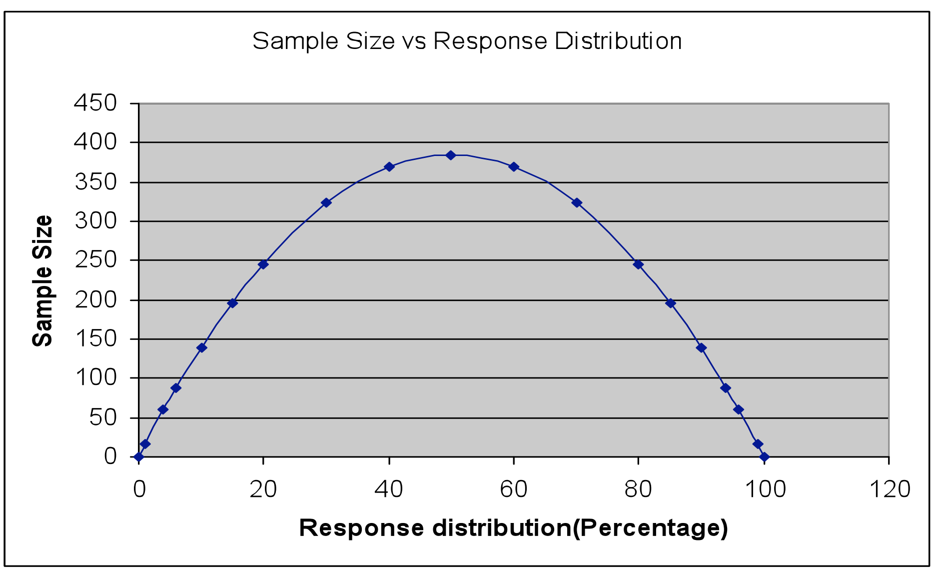


Figure 3-7: Sample Size versus Response Distribution

The graph above shows response distribution of 50% yields the most conservative sample size. Therefore, a response distribution of 50% is assumed when actual response distribution is unknown.

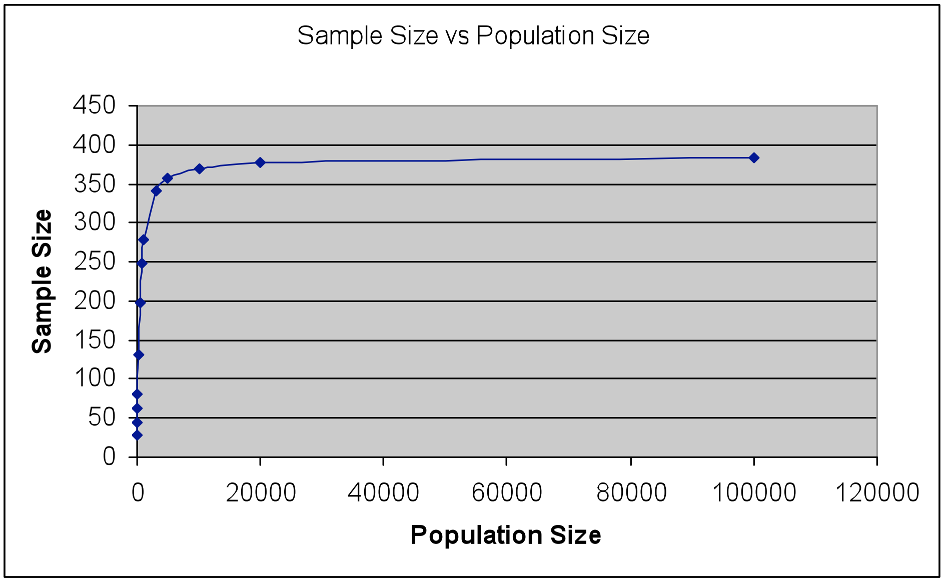


Figure 3-8: Sample Size versus Population Size

The graph above illustrates a fact that by increasing the population size greater than 10000 does not produce significant change in the sample size. Therefore, it is reasonable to assume a large population size when the actual population size is unknown.

In the present case, population size means the total number of drivers who are exposed to the same traffic conditions during their travel on a specific site where there exists a traffic situation calling for alerts. So, the population size is a fraction of AADT on a specific site.

3.2.8.2 __Continuous Outcome__

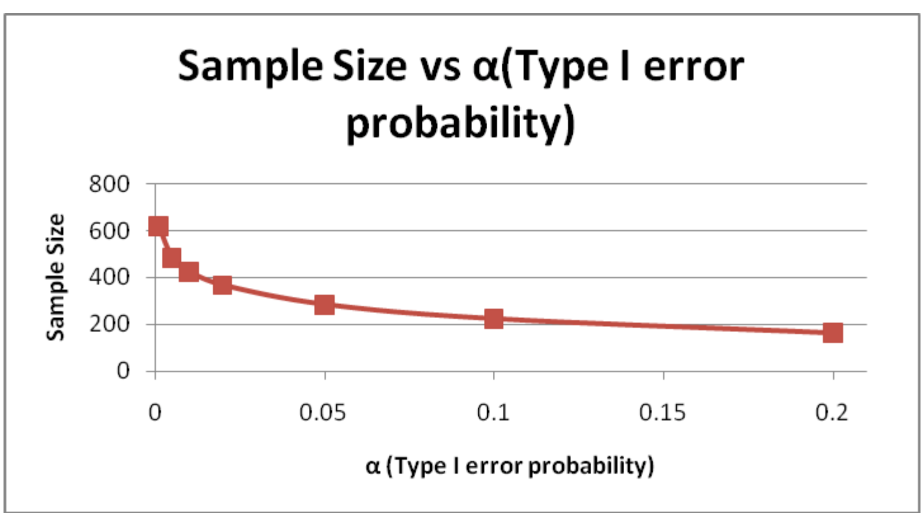


Figure 3-9: Sample Size versus α (Type I Error Probability)

The graph above explains that the sample size decreases as α increases. α indicates the probability of rejecting a null hypothesis when it is actually true. Plainly speaking, it occurs when we are observing a difference when in truth there is none. Therefore, it is obvious we require large sample size to observe a small α.

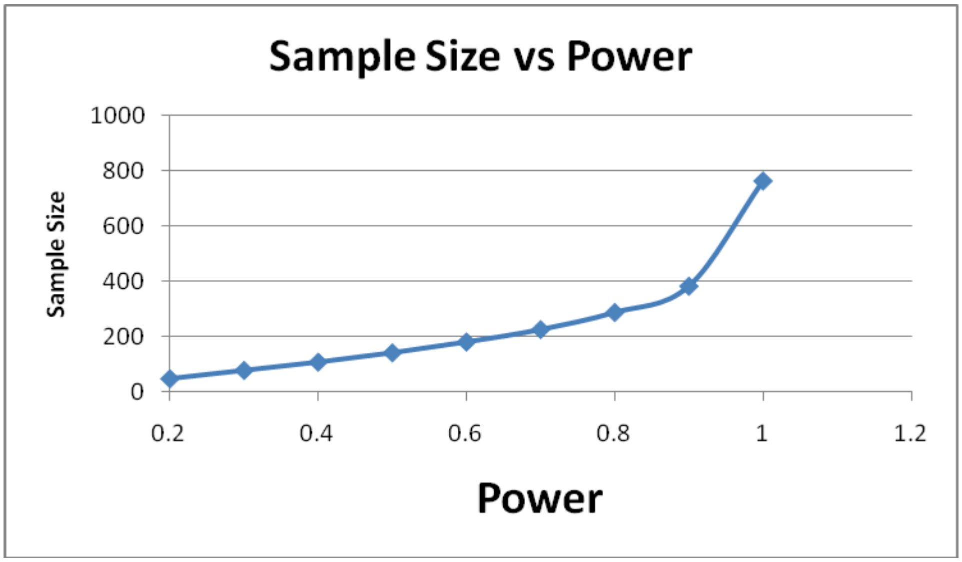


Figure 3-10: Sample Size versus Power (1-β)

The **power** of a statistical test is the probability that the test will reject a false null hypothesis (that it will not make a Type II error). As power increases, the chances of a Type II error decrease. The probability of a Type II error is referred to as the false negative rate (β). Therefore power is equal to $1 - \beta$.

Therefore, higher power requires larger sample size. This is evident from the graph above.

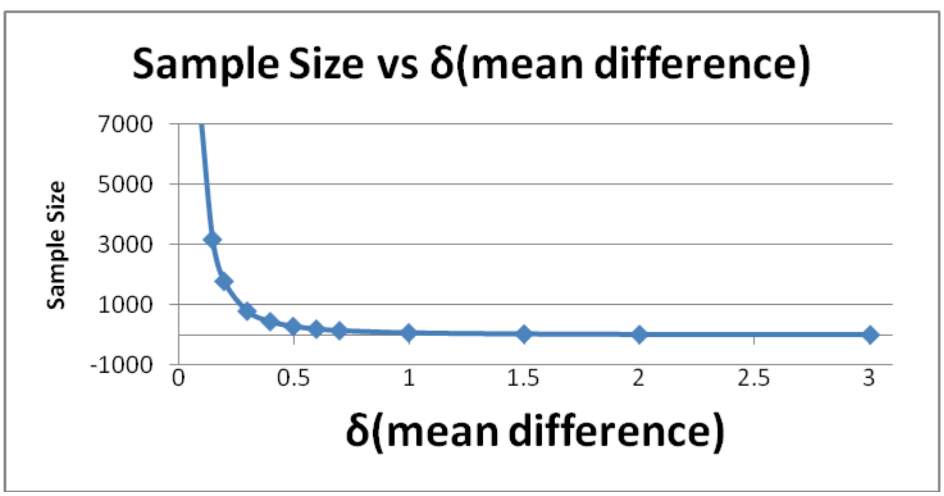


Figure 3-11: Sample Size versus δ (Mean Difference of the Pairs)

The graph above indicates that in order to observe a small mean difference between the pairs, we require a large sample size.

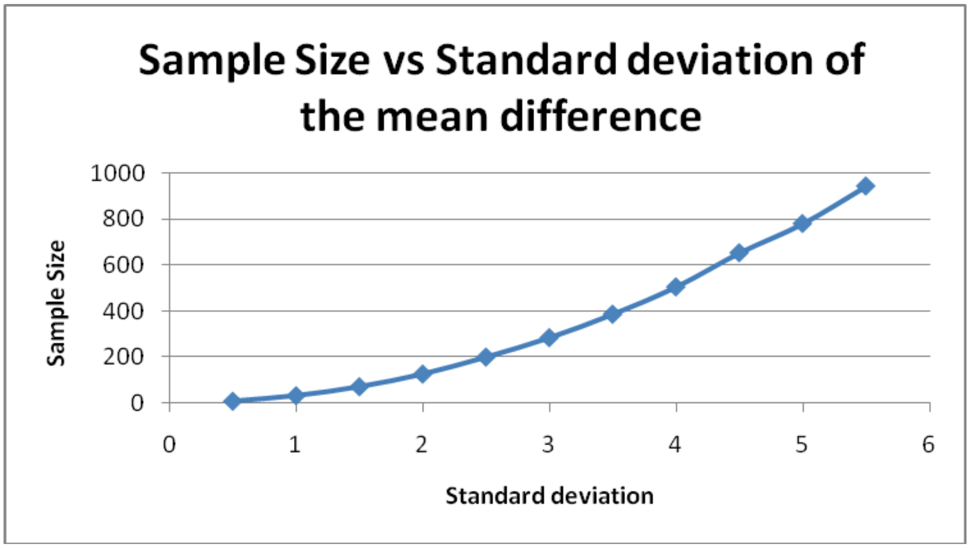


Figure 3-12: Sample Size versus σ (Standard Deviation)

Standard deviation is a simple measure of the variability or dispersion of a data set. A large standard deviation indicates that the data points are far from the mean and a small standard deviation indicates that they are clustered closely around the mean.

The graph above explains that a smaller sample size is required in order to maintain a smaller standard deviation.

3.3 System Architecture

Multiple information systems was integrated with middleware software providing standardized interfaces that was able to deliver information through a variety of mobile devices (see Figure 3-13).

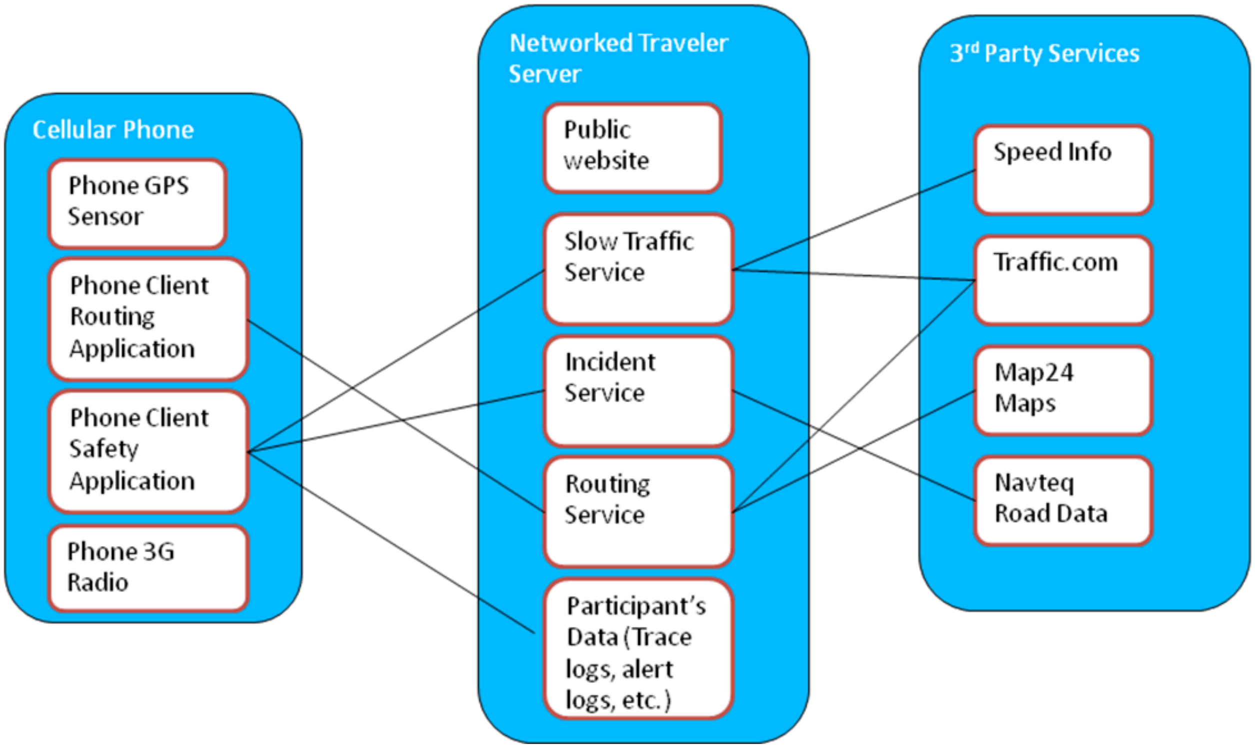


Figure 3-13: Networked Traveler System Architecture

The system architecture is formed of the client, residing on the cellular phone, the Networked Traveler server, hosting several services to enable the required functionality listed above, and 3rd party services that provide the real-time or semi-real-time data that is needed for the functionality.

Data exchange between these components is achieved through several mechanisms. At the lower level the 3rd party components interact with the Networked Traveler Server through a backhaul internet connection having the 3rd party servers as one end node on the Internet and the Networked Traveler as the other node on the Internet. Communication between the Networked Traveler Server and the Cellular Phone is achieved also over Internet communication protocols enabled by a cellular data plan provided by any of the Cellular Network companies. At a higher level of abstraction, data between the 3rd party servers and the Network Traveler Server is exchanged either through XML, JSON or some proprietary protocol developed by the 3rd party. Data exchange between the Networked Traveler Server and the Cellular Phone always occurs through the JSON standard communication over HTTP.

3.3.1 Networked Traveler Server

The Networked Traveler Server serves two main functions: it fuses together the data that comes from 3rd parties to host for the cellular applications, and it hosts the Networked Traveler website which serves as the first entry point to the user's experience of Networked Traveler and as a user profile management center later on during the field test.

3.3.2 Cellular Phone

The cellular phone is the interface that provides the Networked Traveler services to the user and collects user feedback.

On the client, two main functions exist (See Figure 3-14): Routing and Safety. The Routing function solicits information from the user regarding their desired destination and route preference. The Saftey function provides safety messages (in-line with the functionalities described above) related to the route defined by the routing function.

The Routing function converses with the Routing Service on the Networked Traveler Server to obtain the route from the current location of the cell phone to the desired destination. The Routing function also updates its route every set interval related to how the cell phone moves. This ensures that the client application is constantly aware of where the user is headed.

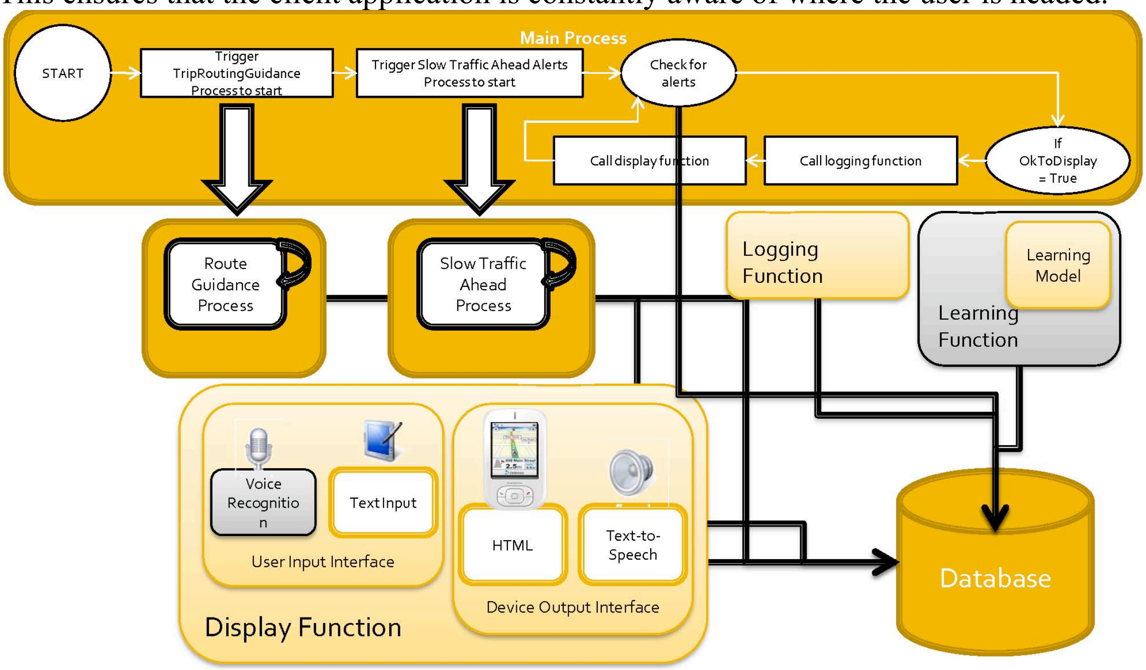


Figure 3-14: Cellular Phone Client Architecture

The Safety function runs two main processes. The first process communicates with the Networked Traveler Server to obtain the most recent safety information regarding the route the user is on. This information is cached on the client and refreshed every 30 sec. This mechanism ensures that any communication timeouts along the way would not result in system crashes. The second process runs every 2 sec and checks the current location of the client by requesting the GPS coordinates from the cellular device and comparing that with the safety information retrieved by the first process. In case of proximity of any of the safety information, and alert is displayed and voiced to the user.

The proximity to safety information is different for each safety service. For the case of Slow Traffic Ahead, the current cell phone location is compared to a "trigger point" location on its route. The "trigger point" is a point defined 60 sec upstream the slowed traffic location (as retrieved from SpeedInfo sensors). If the client is within 1000 feet of the trigger location and approaching at a speed 15 mph above the speed reported at the SpeedInfo location and alert is issued.

Figure 3-15 shows the full Networked Traveler user experience starting from registering on the Networked Traveler website, downloading the client application to their cellular phone and using the application while driving.

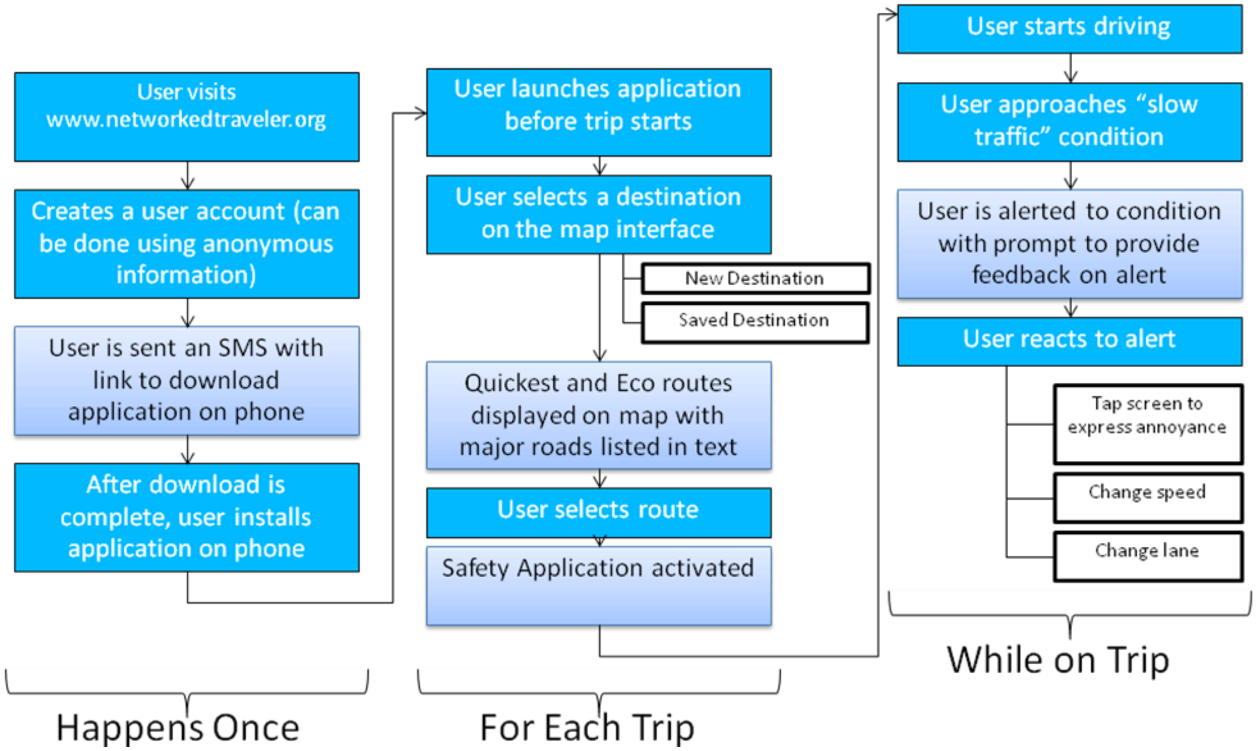


Figure 3-15: Networked Traveler User Experience

3.4 Services

3.4.1 Routing

*3.4.1.1 **Background***

With significant failure rates, existing in-pavement and road-side traffic sensors are typically located on a small subset of freeway links, and accurate travel time and traffic flow information on ramps and arterial corridors are critically needed but very costly to collect. In past several years, in-car navigation and cell phone systems using the Global Positioning System (GPS) technology have matured into a rapidly growing industry and its penetration rate in the U.S. is expected to exceed 9% in 2008. Recently, a new generation of commercial navigation system has been successfully developed to provide two-way connectivity through a built-in Wi-Fi or cellular connection, which allows a network of equipped drivers to anonymously share their speeds and locations, obtain up-to-date traffic flow information, and more importantly make smart route decisions.

The new generation of automobile navigation devices presents a data rich environment for regional traveler information systems to accurately measure route-based travel times and network-wide traffic flow distribution and evolution. It also offers an effective mechanism for traffic management centers (TMCs) to balance traffic on freeway and arterial corridors by delivering precise en-route diversion guidance. On the other hand, utilizing mobile traffic probe data, especially in their early deployment stage, could be constrained by low market penetration rates, leading to small data samples in statistical inference and thus high variances in travel time and network flow estimates. On the other hand, without fully integrating traffic management strategies from TMCs, independent real-time drivers acting non-cooperatively might also affect and even worsen the traffic conditions. By designing and implementing a mobile probe-based traffic monitoring and information provision prototype system, this research aims to (1) provide a web-based traffic visualization interface to end users (i.e. transportation planners and travelers) and (2) demonstrate potential benefits in increasing arterial street traffic observability and eventually improving system-wide traffic conditions.

The objective of this task is to provide a prototype web-based traffic information provision system to California PATH so as to receive input and feedback from the related users (travelers and transportation planners) regarding its interface and architecture. This task consists of two subtasks.

3.4.1.2 Setup Web Browser

Web Browser
Please use Firefox 2 as your web browser.
Firefox 2 can be downloaded from HTTP://www.mozilla.com/en-US/firefox/all-older.html
HTTP Address
HTTP://128.32.129.90/map/mapverajax1.8.html

3.4.1.3 Network Data Coverage

The New York network in the routing engine currently covers Brooklyn, Queens, Manhattan, Bronx and Staten Island (as highlighted in Figure 3-16). Paths cannot be found for origins and destinations outside of this area.
Total mileage of road covered: 3,388 miles, 33,518 nodes, 55,123 links
The web-based Map 24 Interface is provided by NAVTEQ.

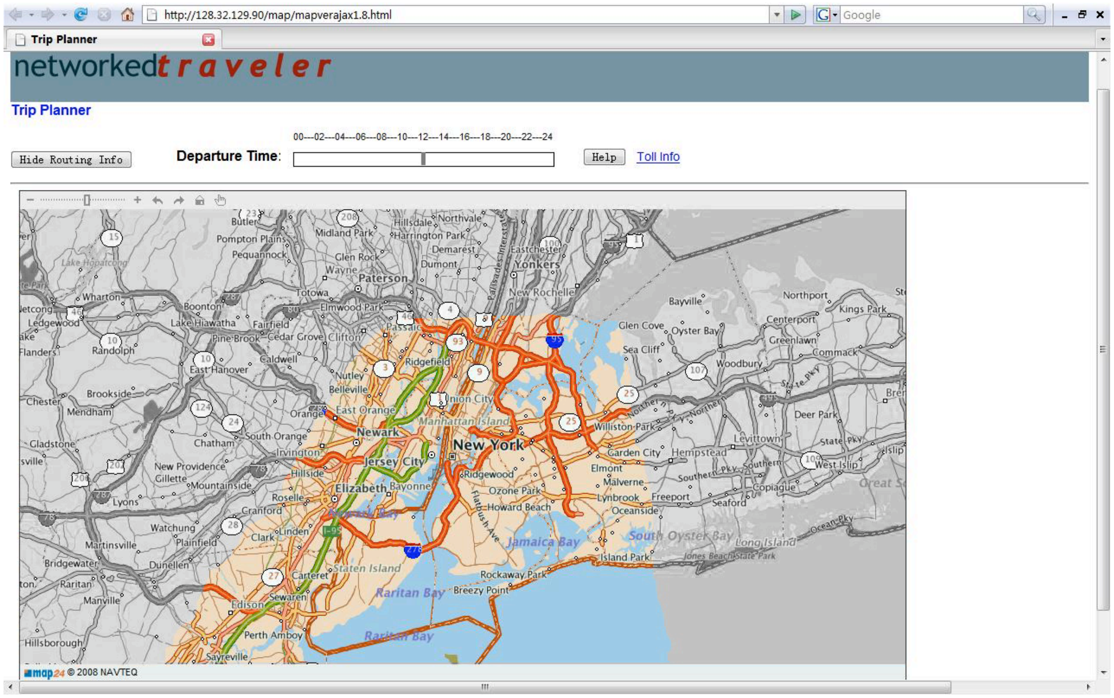


Figure 3-16: New York network in the routing engine

Find Routes: Define Origin and Destination: To find routes, first left-click the mouse on the map to define an origin point, and then left-click to define a destination point. The optimal routes for up to seven criteria was shown on the map, as shown in Figure 3-17.

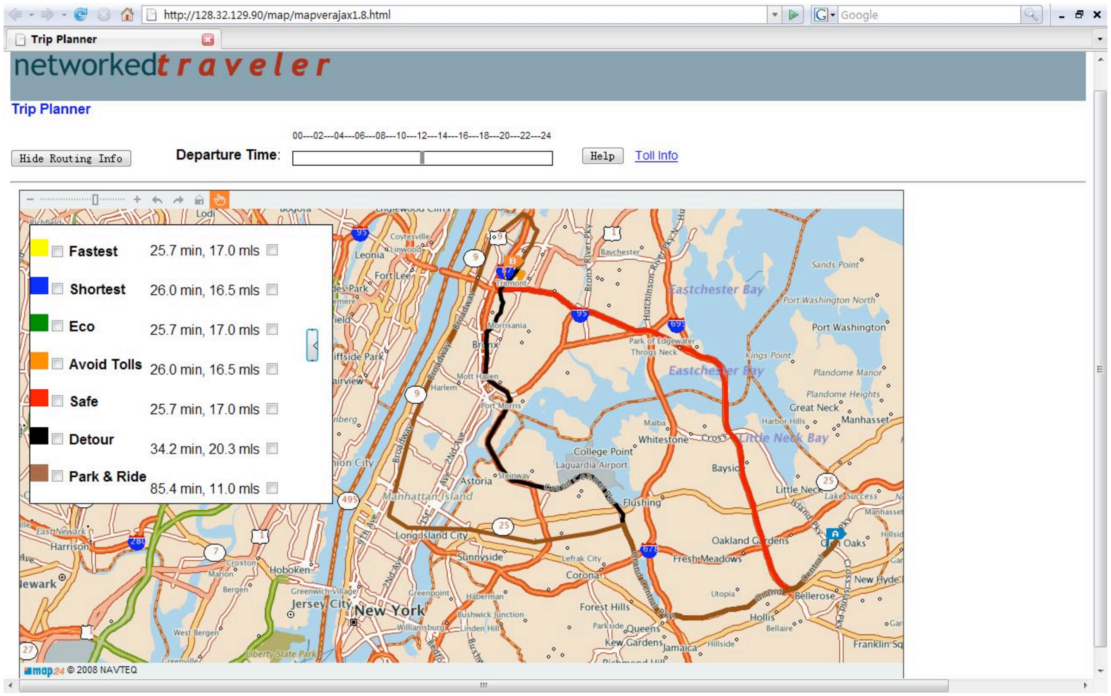


Figure 3-17: Optimal routes based on up to seven criteria

For each path, the corresponding travel time and travel distance are shown on the right.

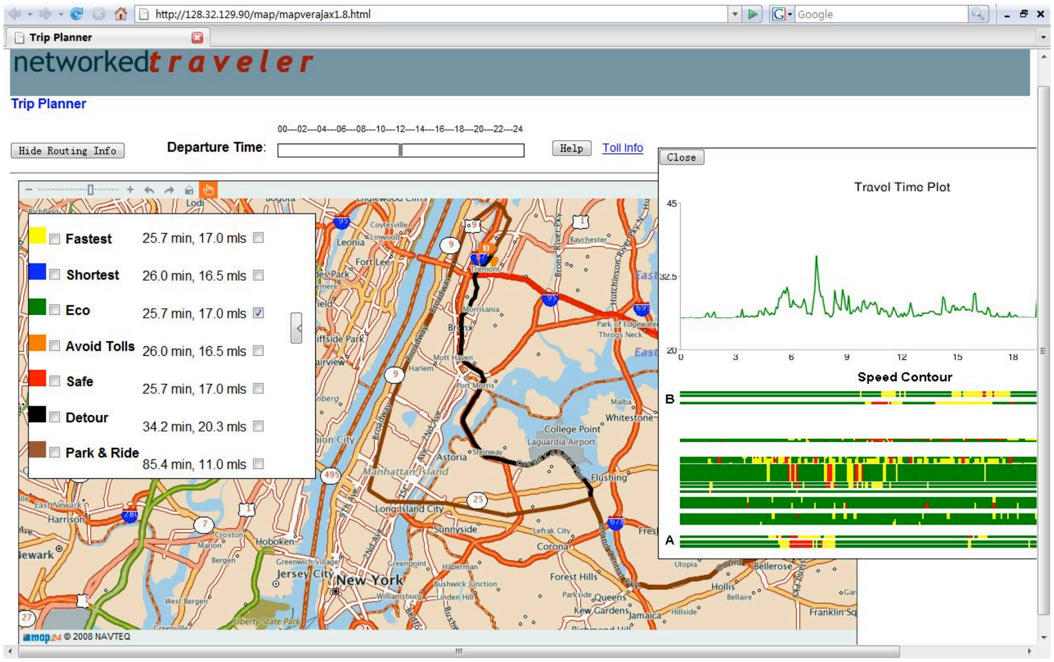
3.4.1.4 Understand Different Criteria Used in Route Finding

Fastest path (Yellow): the smallest travel time
Shortest path (Blue): the shortest travel distance
Eco path (Green): the average speed is close to eco-driving speed (50-60 mph)
Avoid toll path (Orange): no toll along the route or the lowest tolling fee
Safety path (Red): the smallest probability of seeing/being involved in a traffic incident during the whole trip
Detour (Black): highest travel time reliability
Park & Ride path (Brown): use park & ride intermodal option

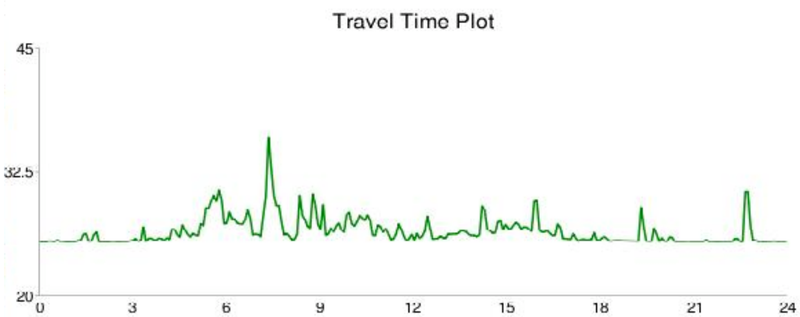
Remarks:
1) Different optimization criteria could lead to the same path.
2) Historical and live traffic data come from Traffic.com (NAVTEQ)
3) Transit and tolling data come from MTA.

3.4.1.5 Check Dynamic Traffic along the Route (through Check Boxes)

There are two check boxes associated with each path.
Click the left check box to highlight the selected path on the map,
Click on the right check box to show the time-dependent travel time profile and speed contour along the selected path.

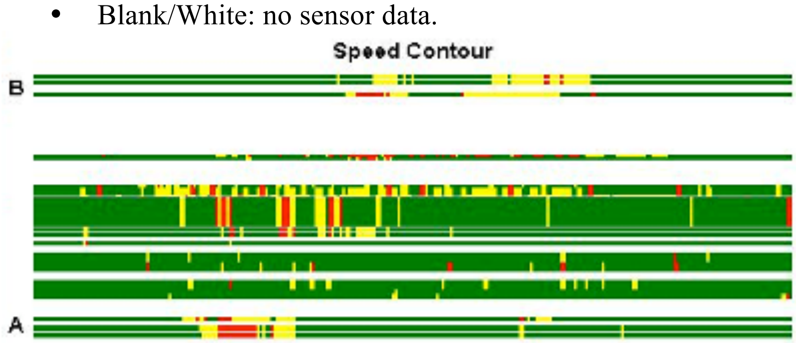


Travel Time Profile: The time-dependent travel time profile shows the travel time at different departure times of day.



Traffic conditions along the route: The speed contour shows the traffic speed data collected from road sensors along the selected path.
The X axis represents time of day.
The Y axis represents space along the path from A to B.
Color legend:

- Red: highly congested.
- Yellow: relatively congested.
- Green: free-flow.
- Blank/White: no sensor data.



3.4.2 Slow Traffic Ahead Alert

3.4.2.1 Introduction

The primary safety-related service of the safety application is Slow Traffic Ahead Alert for drivers. Networked Traveler's client application runs on the phone and executes the following major tasks:

- Reads GPS information (position, speed, heading angle, etc) of the phone.
- Uploads the GPS samples to Networked Traveler's server every few seconds (22 sec, in the current version) via cellular data communications network.
- Server searches around location of the vehicle and if found relevant sends back to the Client, a set of data related to slow traffic ahead alerts.
- Client uses the server data and if conditions satisfied, issues a slow traffic ahead alert to the driver. The alert is both audio and visual and can be also delivered through Bluetooth.

3.4.2.2 Alert Logic

The alert logic is described below and also illustrated in Figure 3-18.
Here are some definitions related to the alert logic:

- **Subject vehicle:** Vehicle that its driver is going to receive slow traffic ahead alert
- **Alert location:** Location ahead of the subject vehicle where traffic is slow
- **Trigger location:** Represented by GPS lat, long, and heading; about one mile (60 seconds of free flow speed) before the alert location. Alert is issued at or within 500 ft of the trigger location.

Suppose:

- V_s = Speed of the subject vehicle
- V_f = Speed of the vehicles at the alert location

Alert is issued if all conditions below are satisfied:

1) $Vf \leq 50$ mph
2) $Vs - Vf \geq 15$ mph
3) Distance between trigger location and subject vehicle location $\leq$ 500 ft and
4) Difference between trigger location's heading and vehicle's heading $\leq$ 50 deg

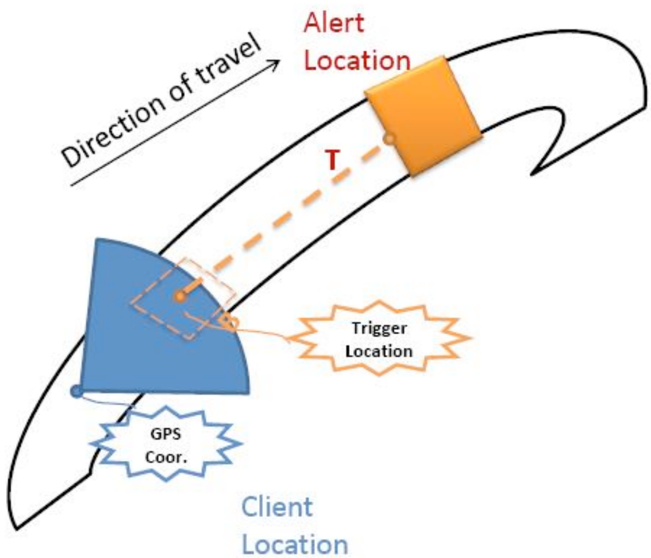


Figure 3-18: Slow traffic ahead alert logic

3.4.2.3 __Coverage Area__

We use two sources of traffic data: Navteq Traffic.com and SpeedInfo. In total, we have more than 2000 trigger points in Bay Area, which they cover all major freeways.

3.4.2.4 __Alert image__

Figure 3-19 shows the alert. The images take over entire screen of the phone. User may tap anywhere on the screen to provide real-time feedback for the alert. As text below the slow traffic ahead sign mentions, tapping the screen means user has found the alert to be not useful or irrelevant.

Figure 3-19: Image of the slow traffic ahead alert

3.5 Website

From mid-February to July 2009, NT-Safety developed the following web application to support Phase I of the Networked Traveler - Safety launch. The basic functionality of the website was to allow users to create a web-based user account which would:

1) Facilitate software download to the mobile phone,
2) Ensure users accept the Networked Traveler terms and conditions,
3) Visually identify specific alerts given to specific users, to enable qualitative feedback on the alerts.

Figure 3-20 shows screen shot of the account creation page. The first challenge was to provide users a way to download the application to his or her mobile phone. We created a simple form for users to fill out, including a unique username and password. Once a user had provided us with his or her username, password, mobile phone number and carrier, we sent a text message to the phone with a download link.

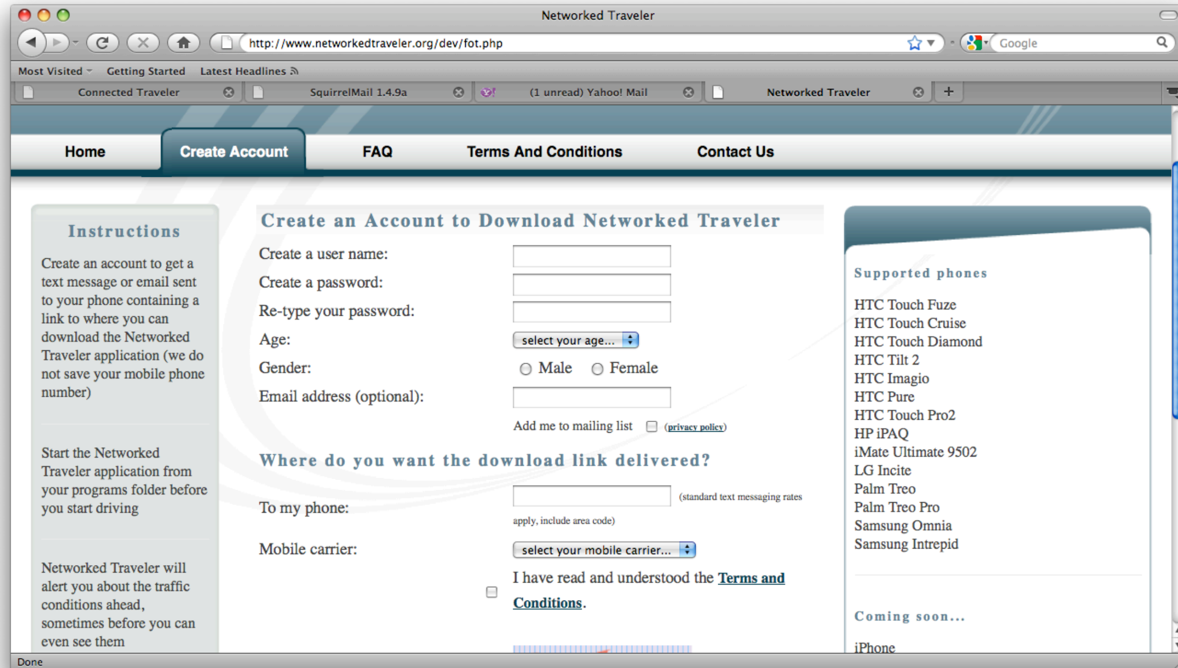


Figure 3-20: Account creation page of the Networked Traveler website

Visual design was taken and adapted from the GEMS efforts demonstrated at ITS World Congress in November 2008. The basic account creation was written in HTML and PHP, with a MySQL database capturing the basic user data. The database was also adapted from tables created in support of the GEMS web services efforts.

The user table was expanded after July, but most of the fields are as follows:

Field	Type
uid	int(11)
login	varchar(20)
password	varchar(32)
appCategories	varchar(600)
terms	varchar(12)
gender	varchar(3)
age	varchar(10)

Though only a preliminary survey was written and deployed by July, it provided the technological framework to collect the qualitative feedback of users. This supports the following evaluation objectives set out in the SafeTrip-21 California Connected Traveler Evaluation Plan:

1) Analyze the perceived timeliness, accuracy, and usefulness of safety alerts.
2) Explore the user-perceived benefits of the safety alerts.

The preliminary survey was an adaptation of questions posed in the experimental design document dated Feb 8, 2009. The innovation was using Google Maps to visually identify the alerts, and then tying alert-specific identifying data (time, date, and user i.d.) to survey responses. Figure 3-21 shows an example of it. This enables a finer-grained and richer collection of data than surveys which are not associated with particular events, though this method presents unique challenges, too. The survey content was modified after July.

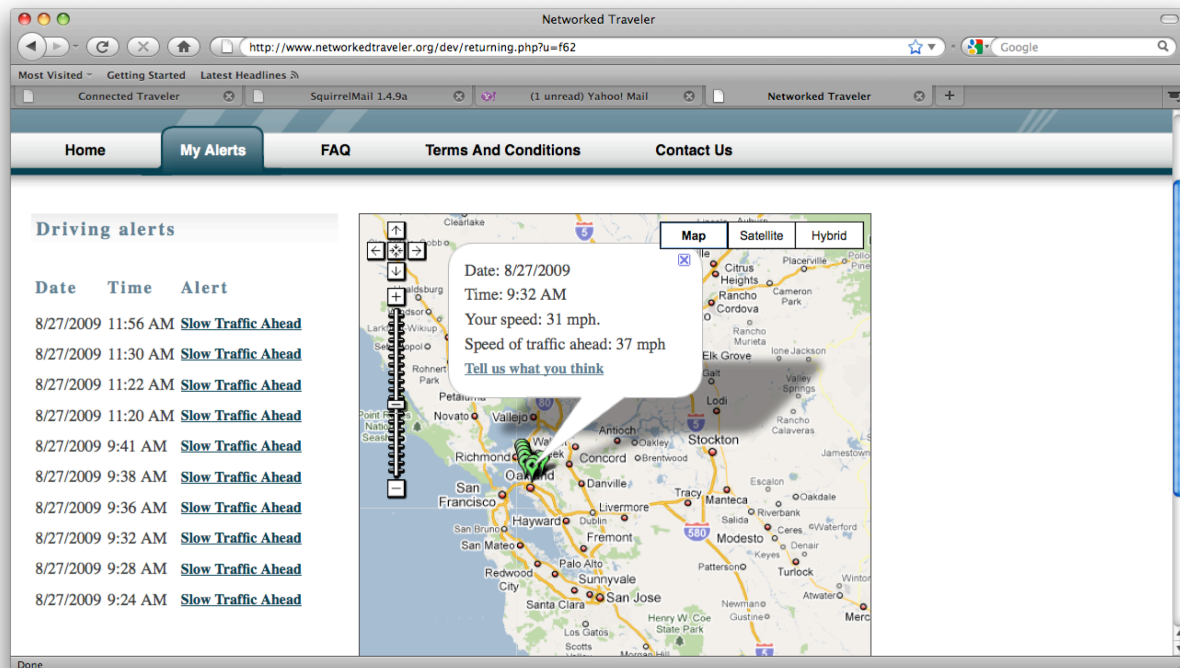


Figure 3-21: Google Maps visualization of the alerts, with the informational popup

Although Google Analytics was added after July to collect usage statistics, page views, completed downloads, etc., the basic site for which to collect these statistics was built in the time period we're discussing. The demographic survey to collect basic user demographic information immediately after account creation was also implemented after July.

3.6 Outreach

Currently, the research team of the Networked Traveler Project is in active discussions with the Metropolitan Transportation Commission (MTC) and several associations of drivers (e.g., CSAA and the Transportation Management Association of San Francisco) within the San Francisco Bay

Area Region. Ideally, the field tests of this second service were conducted through full integration of the traffic advisory incident data available from MTC's 511 system. In any event, as an important task initiated at the outset of this project and conducted in parallel to Task 1, we aim to find up to 1000 drivers already owning appropriate GPS-equipped smartphones and unlimited data plans.

3.7 Safety Benefit Feasibility Assessment

This section provides a description of the premise, hypotheses, rationale, and the approach for conducting safety field tests, with the goal of assessing the validity and usability of the proposed safety applications.

3.7.1 Premise

If drivers are better informed of the traffic conditions in their driving environment, they was more aware and better prepared to take actions to avoid hazardous situations. Within the scope of this study, we focus on the "soft" safety applications that have a relatively longer time window to provide drivers with alerts. A "hard" safety application, by comparison, requires an immediate action. For example, a system that warns drivers of imminent freeway front-end collisions with another car in front will need to take effect within 1-5 seconds. On the other hand, one example of the "soft" safety applications that we propose to offer is the situational awareness alert that can be effective in a 10-60 second time window. For example, in a situation where drivers cannot see the slow or stopped traffic beyond a curved roadway ahead, at a distance of 1 mile or so before the driver reaches at location, we can issue an alert to the driver especially when the driver's subject vehicle is traveling significantly faster than the traffic queue ahead.

3.7.2 Applications

A suite of applications are being considered and to be tested for the planned field experiments. Most of these applications are designed for freeway driving conditions, which are the primary test targets in this stage of the Networked Traveler project. The core list of safety applications are given in Figure 3-22, and they are explained and summarized as follows:

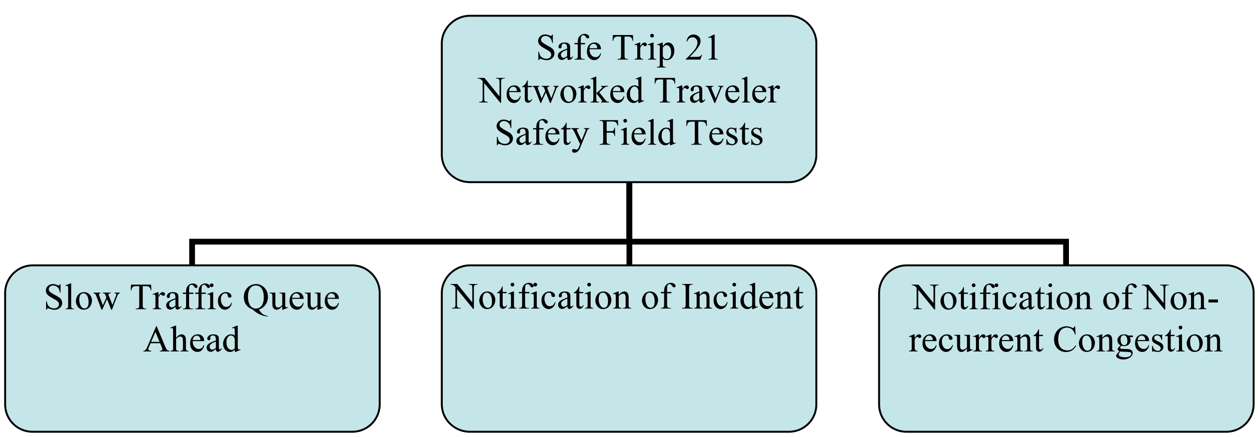


Figure 3-22: Safe-Trip 21 Safety Applications

3.7.2.1 __Situational Awareness of Slow Traffic Queues__

The first application is a situational-awareness function that provides alerts to drivers on driving conditions within a relatively short-latency time horizon, in the range of 10-60 seconds. For this application, we focus on specific highway locations where past traffic and crash patterns indicate that a Networked Traveler system can offer useful alerts so that drivers can take tactical actions to avoid crashes, or in essence increase their 'safety alert time horizon' to beyond the just several seconds based on driver reaction or provided by active safety systems. In these situations, by offering a timely "slow or stopped traffic ahead" message the system can effectively inform drivers of roadway hazards to reduce chances of crashes and realize safety benefits.

To seek the test sites that may offer the most significant benefits for crash reduction, the research team investigated the collision data for the last ten years for the greater San Francisco Bay Area and identified sites where specific crash patterns might warrant the deployment of such applications. A list of candidate locations, their attributes and maps are given in the following sub-section.

Subsequently, it is reasoned that the "slow traffic ahead" application can be extended to a much broader network of highway spots where traffic speed can be detected. If the application user's vehicle speed is significantly higher than the speed of traffic stream ahead, then an alert may be warranted. Therefore, through the collaboration with Metropolitan Traffic Commission (MTC) and two companies that provide real-time traffic information, SpeedInfo and Traffic.com, a network with more than 2,000 nodes of traffic speed information is incorporated into a greater area test bed to be included in the field tests of safety applications. Exemplar illustrations in a sub-section below illustrate the distribution of these nodes in the Bay Area highway network.

3.7.2.1.1 Study Sites with specific patterns of crash concentration that warrant situational awareness alerts

This section contains the list (Table 3-4) and maps of candidate sites currently being considered for Safety Application #1.
Note: Nomenclature in the tables below:

- B: Northbound; SB: Southbound; EB: Eastbound; WB: Westbound
- PM: Post Mile as defined in California Highway Database on a certain route within a county

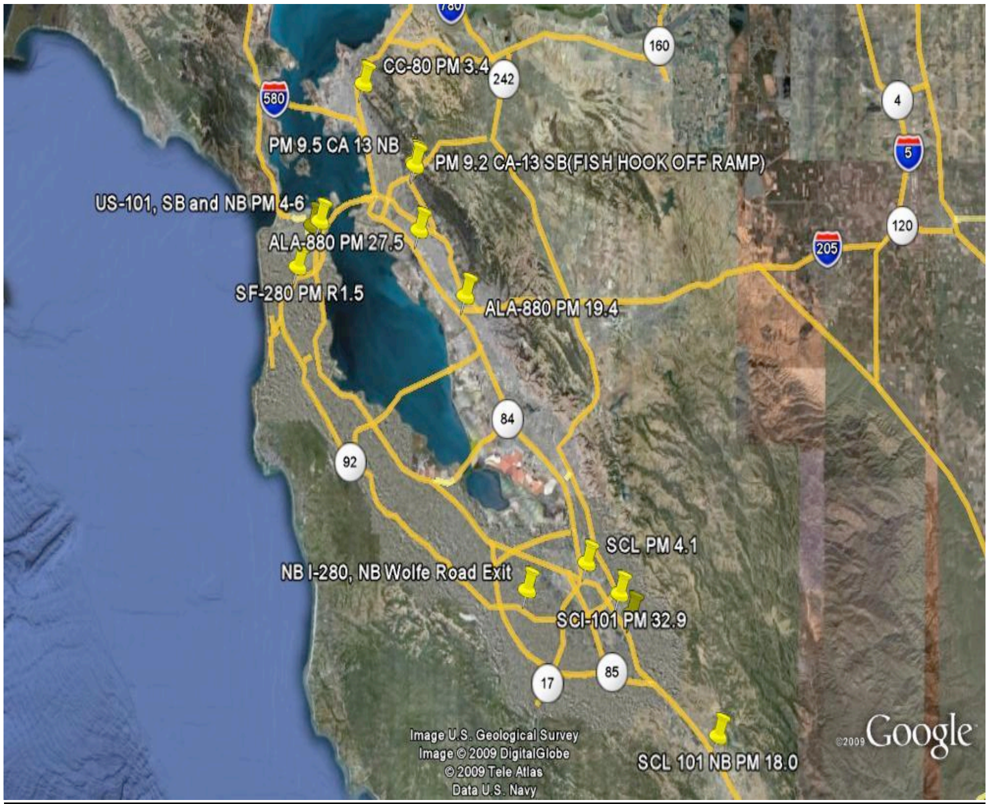


Figure 3-23: Area Wide Map of Potential Study Sites Locations

Table 3-4: Candidate Sites for Field Tests of Safety Application "Slow Traffic Ahead"

Site No.	Site Location	Site Characteristics	Type of Situation Awareness
1	Alameda County SR-13, NB PM 9.5	• Limited line of site to sections ahead of curve • Slow queue ahead due to off-ramp backup into mainline	• Slow traffic ahead on right lane due to bottleneck • Slow traffic after curve
2	Alameda County SR-13, SB PM 9.25 Broadway Terrace Off-Ramp	• Severe fish-hoop off-ramp • Combined with a stop and traffic light at end of ramp	• Curve over-speed • Queue at off- ramp
3	Alameda County I-880, SB PM 27.5	• Combined vertical and horizontal curves • A merge of high-street entry ramp and mainline prior to curve • Congestion in rush hours	• Slow traffic after curve
4	Alameda County I-880, NB	• Off-tramp bottleneck with backup into right lane of	• Slow traffic ahead on right lane due to

	PM 19.4 Connector to SR-238	mainline traffic	bottleneck
5	San Francisco County I-280, NB PM 1.5 Geneva Off-Ramp	• Off-tramp bottleneck with backup into right lane of mainline traffic	• Slow traffic ahead on right lane due to bottleneck
6	SR101 NB, PM 4.2, Mission & Dubose off Ramp	• Traffic backing up in the off ramp. • Congestion	• Traffic Weaving • Slow traffic due to congestion
7	San Francisco County US-101, SB and NB PM 4-6 (Hospital Curve)	• Combined vertical and horizontal curves • Congestion bottleneck • Traffic Weaving with on- and off-ramp traffic	• Slow queue ahead • Traffic Weaving
8	Santa Clara County I-880, NB PM 4.2 Connector to US-101 NB	• Frequent collisions with K-rail barrier on left side	• Off-ramp curve
9	Santa Clara County US-101, SB PM 32.9 Segment between Tully Road and Story Road	• Considerable traffic weaving between two exits in congestion periods	• Traffic Weaving
10	Santa Clara County US-101, NB PM 18.0 Near Cochrane Interchange	• Near 60% of collisions on left lane • Near Transition of 3-Lane into 4-lane segment with HOV lane on left • On-ramp and off-ramp nearby	• Lane Transition Traffic Weaving
11	Santa Clara County US-101, SB Tully Road EB Exit	• Speeding a major factor (more than 75%) • Rear-end Collision dominate (near 90%) • More than 85% ramp collisions at ramp exit and cross street	• Slow queue ahead • Ramp over-speeding
12	Santa Clara County I-280, NB Wolfe Road Exit	• Limited line of sight to ramp queue from mainline • Signal controlled cross street at ramp exit • More than 90% ramp	• Slow queue ahead • Ramp over-speeding

		collisions at ramp exit and cross street	
13	Contra Costa County I-80, EB PM 3.4 Segment between San Pablo Ave and Solano Ave Exits	• Combined vertical and horizontal curves • Congestion bottleneck	• Slow traffic ahead due to bottleneck • Traffic Weaving

3.7.2.1.2 Network of Traffic Speed Measurements Nodes Included for Safety Alert Generation

The current version of Networked Traveler utilizes an area-wide real-time traffic data map that includes more than 2,000 locations in the San Francisco Bay Area. Data feed comes from SpeedInfo and Traffic.com and processed into a Networked Traveler server and is used to update all calculations. Figure 3-24 and Figure 3-25 show exemplar displays of locations where traffic data are available on freeways near San Francisco.

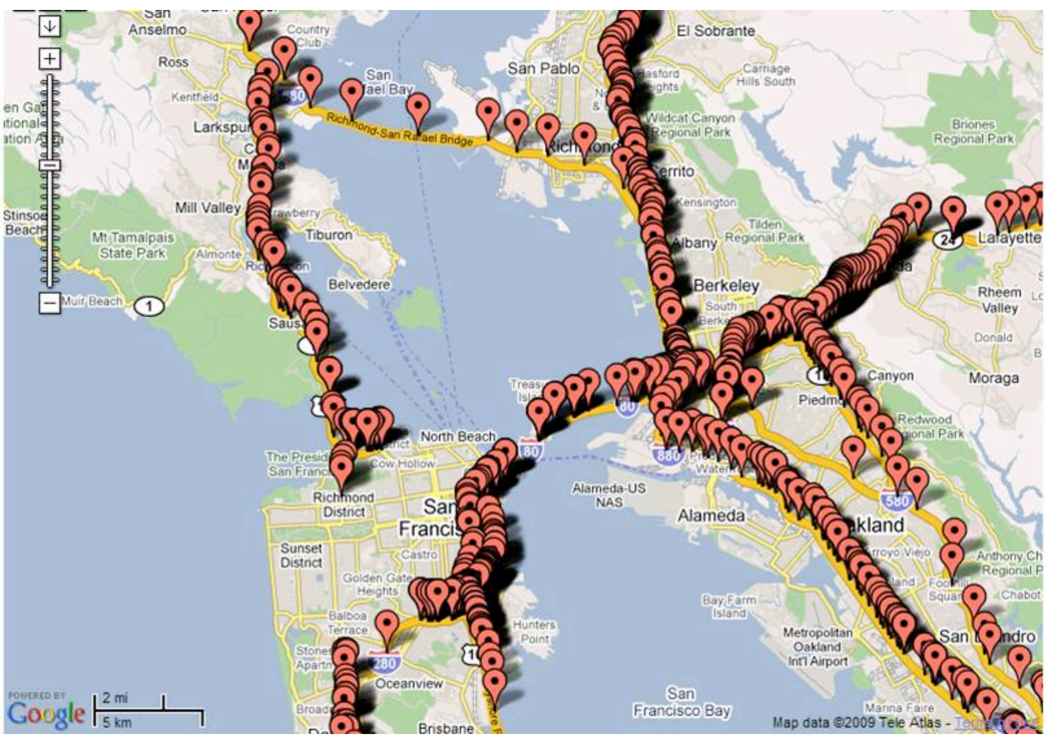


Figure 3-24: An Illustrative Map of Mapped Locations with Real-Time Traffic Data Feed

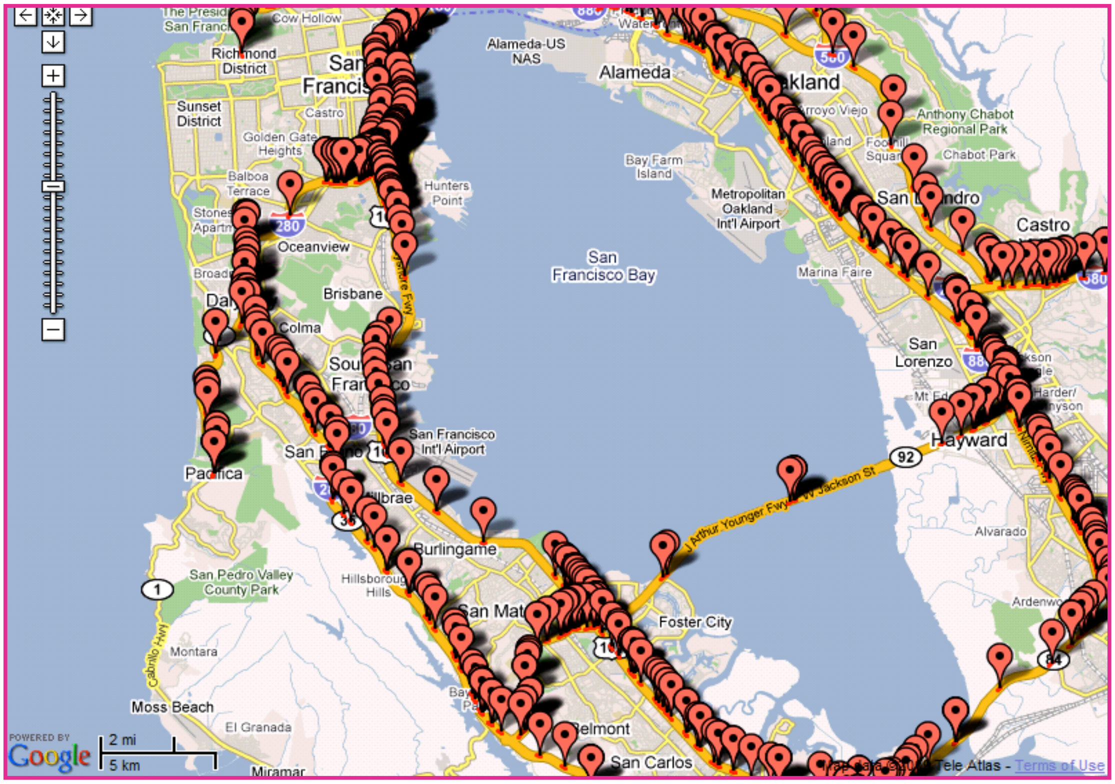


Figure 3-25: An Illustrative Map of Mapped Locations with Real-Time Traffic Data Feed

3.7.3 Notification of Incidents and Non-recurrent Congestion Ahead

One other application is the notification of incidents and non-recurrent congestion on the road ahead to the drivers based on real-time traffic data. The major premise of this application of "Incident and Congestion Alert" is as follows:

(1) First of all, drivers can stay informed of roadway traffic conditions, which allow them to make trip planning choices. The information was "pushed" to users based on their current positions and travel routes.

(2) Secondly, in congested areas on highways, various hazardous scenarios may develop and lead to increased likelihood of collisions. With the proposed application, an earlier notification alert to the drivers, in the range of 2-30 minutes can offer drivers opportunities to take tactical and strategic actions, including

- reducing speeds with increased awareness of slow traffic ahead
- choosing to change routes to avoid further trip delays
- changing transportation modes by switching to transit or travel plans, with benefits of reducing traffic demands on flow-stressed or incident-impaired roadway segments can effectively inform drivers of roadway hazards to reduce chances of crashes and realize safety benefits.

3.7.4 Safety Applications Test Hypotheses

For each of the applications that are to be deployed for the field tests, the hypothesized outcome and the expected user responses and the safety impacts are described and listed in Table 3-5.

Table 3-5: Safety Application and Test Hypotheses

Application	Applicable Situations	Hypothesized Outcome	Expected Driver Response and Safety Effects
Slow Traffic Ahead	• Traffic queues ahead of curved roadway with limited visibility • Collisions are caused by vehicles approaching too fast toward the end of slow traffic queue	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to slow down in advance • Drivers can make cautious approach before reaching end of queue by receiving alerts • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the drivers can take timely, evasive actions without the alerts	• Drivers provide favorable assessment of the alerts • Drivers respond positively and noticeably to the alerts • Drivers reduce speed earlier or more significantly with the alerts than without the alerts • Drivers make lane change maneuvers to other lanes to avoid slow queue
Slow Traffic Ahead	• Off-ramp queue buildup and spillover into freeway • Collisions are caused by vehicles approaching too fast toward the end of slow traffic queue	• Same as above	• Same as above
Slow Traffic Ahead	• Severe traffic weaving section • Collisions are caused by vehicles sideswiping or rear-ending other vehicles are making moving across lanes in the weaving section	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to make cautious approach before reaching the weaving section • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the drivers can take timely, evasive actions without the alerts	• Same as above
Notification of Incident On Route	• Potential slow traffic or congestion induced by incidents • Some collisions are caused by vehicles moving too fast into slow traffic in incident-induced	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to make cautious approach before incident segments • Drivers benefit by choosing alternative routes if alerts are given early enough for them to	• Drivers provide favorable assessment of the alerts • Drivers respond positively and noticeably to the alerts • Drivers reduce speed earlier or more

	congested areas • Other collisions are caused by stressed traffic conditions • Other collisions are caused by changes in lane configurations or traffic patterns	exit and to avoid problematic areas • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the drivers can take suitable actions without the alerts	significantly with the alerts than without the alerts • Drivers make lane change maneuvers to other lanes to avoid incident areas • Drivers change routes to avoid incident areas
Notification of Non-recurrent (unexpected) Congestion On Route	• Congestion caused by all probable causes unexpected by users • Some collisions are caused by vehicles moving too fast into congested areas • Other collisions are caused by stressed traffic conditions • Other collisions are caused by changes in lane configurations or traffic patterns	• Collisions can be avoided if drivers are given alerts in a time frame that allows them to make cautious approach before congested segments • Drivers benefit by choosing alternative routes if alerts are given early enough for them to exit and to avoid problematic areas • Drivers benefit by an elevated sense of safety and comfort from the alerts even if the drivers can take suitable actions without the alerts	• Drivers will prefer to receive alerts about non-recurrent congestion only • Recurrent congestion if issued frequently will appear to be nuisance therefore result in negative feedback • Drivers provide favorable assessment of the alerts • Drivers respond positively and noticeably to the alerts • Drivers reduce speed earlier or more significantly with the alerts than without the alerts • Drivers make lane change maneuvers to other lanes to avoid congested areas • Drivers change routes to avoid congested areas

There are a multitude of common elements given in the table above. They can be reorganized into the charts below.

Figure 3-26 is a diagram showing the safety applications may provide benefits in several manners:

- Users benefit by having an increased sense of awareness of roadway hazards ahead
- Users benefit by being able to take cautious approach when hazardous situations are encountered
- Collisions can be avoided with users taking proper actions and alter trajectories before reaching hazardous situations

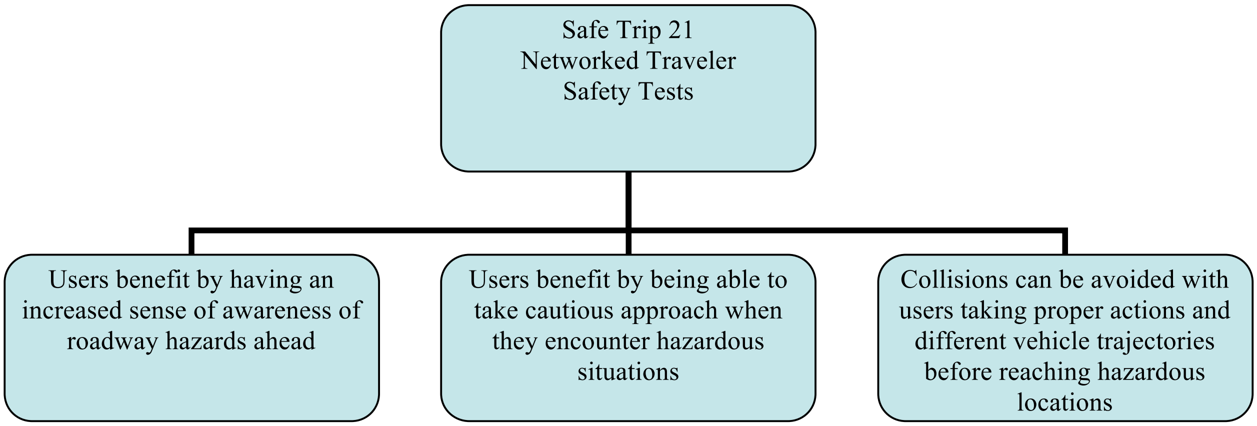


Figure 3-26: Hypothesized Expected Outcomes of Safe-Trip 21 Safety Applications

In order to assess the benefits that are received by users, the hypotheses need to be verified through the actions or responses from the users, including:

- Users express positive experience and provide favorable feedback of field tests
- Users take tactical actions, such as speed reduction or lane change, to mitigate the potential hazardous situations
- Users take strategic actions, such as route changes, to avoid passing through or approaching hazardous locations, such as incident areas.
- Users are able to take timely actions to avoid crashes

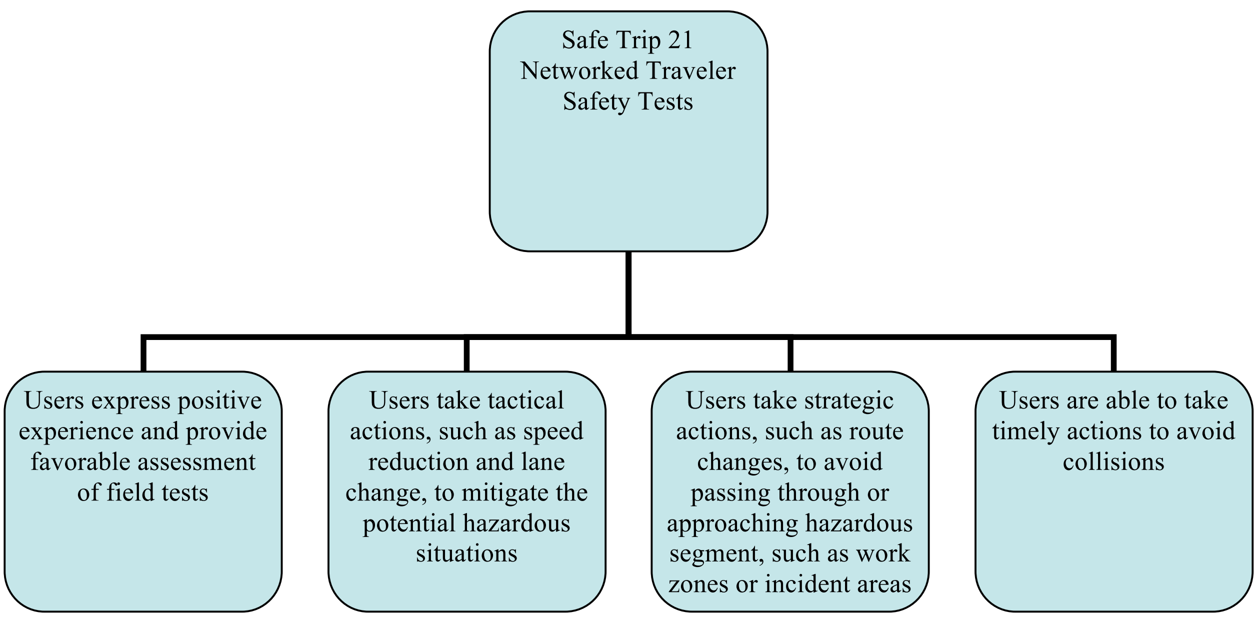


Figure 3-27: Expected User Responses and Safety Impacts of Safe-Trip 21 Safety Applications

3.7.5 Validation of Test Hypotheses

3.7.5.1 <u>Safety Application and Expected Test Outcome</u>

The previous section provides an overall description of the test hypotheses of the expected outcome. In the course of the field tests, it is expected that user experience was evaluated and field data was collected to explore the user needs and preferences as well as design validity and shortcomings of the safety applications. Therefore, experimental design of the field tests should emphasize on the observations of user response, qualitatively and quantitatively, to establish the foundation for making such assessment.

The diagram below, Figure 3-28, shows the expected cause-and-effect sequence in the process of experimental designs. The top layer is the planned safety field tests. The second layer is the applications to be deployed. The third layer is expected user actions, if safety alerts are effective. The fourth layer is the expected observable outcome, as a result of the safety field tests.

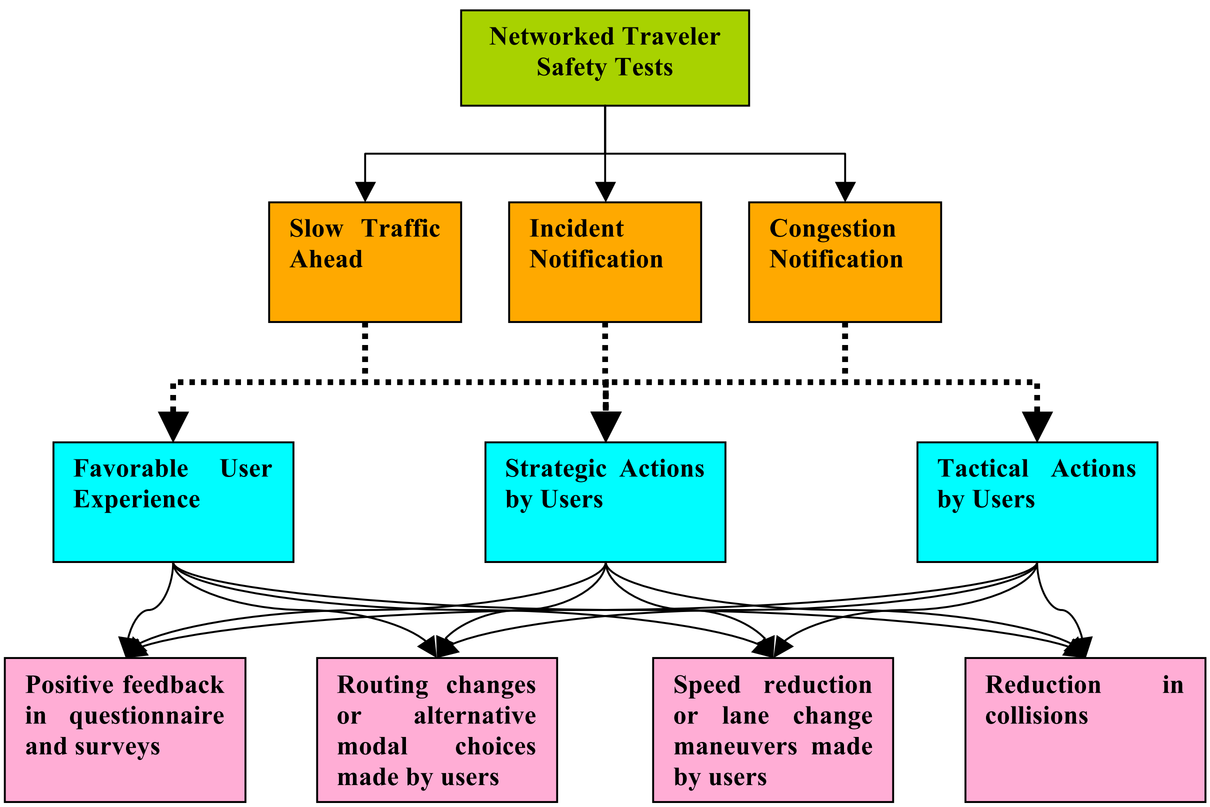


Figure 3-28: Safe-Trip 21 Safety Applications and Observable Outcome

3.7.5.2 <u>Data Collection Constraints</u>

While the availability of observable data to validate the test outcome is essential for benefit assessment, it is critical to point out the constraints of the data acquisition process within the context of the planned field tests.

The following constraints in both quality and quantify of data acquisition should be noted:

(11) System server received traffic and incident information from other sources (Traffic.com and SpeedInfo); therefore the update rate and the quality of traffic data are not within the control of the application developers.

(12) Traffic information are only available for specific locations in the test bed area, therefore the measurements of overall corridor traffic settings are not available continuously in time or in space.

(13) Even though sensors were installed to capture the traffic conditions at these locations, they can mostly provide traffic speed measurements for a specific site or a short segment for the selected locations. It does not provide the information about the ending locations of traffic queues and may not entirely reflect the actual representation of traffic parameters for ideal and robust processing and execution of alert generation algorithms.

(14) Users may or may not activate safety functions even if they volunteer and register for the services.

(15) Users may choose to activate selective functions of their preferences and thus only a subset of functional results and associated data are available for individual users.

(16) User positions are only available through the GPS coordinate on user's devices, and therefore the accuracy and availability of user trajectory (speed and position) depend on the GPS units and the settings of driving environment, which may vary significantly along users' routes.

(17) There is no measured information about the movements of surrounding vehicles, which may have impacts on the driver actions when alerts are issued. Therefore, the causal effects of driver actions in response to alerts may not be determined.

(18) There is only limited information about the traffic conditions at test sites. For example, in the "slow-queue-ahead" scenarios, only the average speed of existing traffic queue is available to be used as a criterion in alert-generation algorithms. Therefore, insufficient description of the actual traffic settings may not explain whether drivers can positively respond to the alerts.

(19) There are multiple variables involved in the causation of collisions, including driver (fatigue, distraction, incorrect judgment, etc.), environment (visibility, roadway surface conditions, weather, etc.). A significance depth and broad spectrum of these related parameters are not available for assessment for this pilot test. This, it should be reiterated that this pilot test is only targeting and designing for a limited subset of situational awareness scenarios.

(20) The pilot test is further constrained by the fact that alerts are communicated to users through user interface implemented within the capabilities and flexibility allowed by the phone-based devices.

3.7.5.3 Measure of Effectiveness

Having discussed the limitations of the planned tests in the previous section, it can be further explored about the data to be collected and how they can be analyzed and investigated to assess the effectiveness of the suggested applications. The following table illustrates how a matrix of measure of effectiveness can be constructed.

Table 3-6: Anticipated Test Outcomes and Measure of Effectiveness

Expected Test Outcome and Driver Responses	Measures of Effectiveness (MOE)	Parameters and Variables to Assess MOE
Public awareness of safety campaign	• Spectrum of project partnerships	• List of partners in project • Scope of participation by partners • List of participating organizations outside of project team
	• Scope of community participation	• Number of participating users • Number of data samples collected in field tests • Percentage of positive feedback by users
	• Outreach efforts	• Sessions of activity reports held in public forums and conferences • Technical papers presented • Reports of media events
Favorable user experience and positive user feedback	• Willingness to participate and to maintain continual use of applications	• Number of participating users • Periods of active usage • Continuity and frequency in activating applications • Percentage of positive feedback by users
	• User feedback to surveys and questionnaire on - Function usefulness - Function acceptability - Timeliness of alerts - User interface friendliness	• User answers in surveys and questionnaires (to be detailed and designed later)

Strategic actions (routing changes or modal choices) by users	• Correctness and reliability of alert generation	• Numbers of alerts generated • Percentage of valid alerts • Conditions of false alerts (time of day, incident type, congestion status, travel time)
	• Timeliness in alert issuance	• Total latency in alert reception time versus design point in algorithms • Conditions of time lags (time of day, incident type, congestion status, travel time) • User feedback in real time or in surveys
	• User routing changes after alerts are issued • User choices of modal changes after alerts are issued	• Route records captured before and after alert issuance • User inputs or responses upon reception of alerts • Percentage of user responses • Percentage of confirmed and verified user responses - based on user real-time feedback - based on user surveys • Conditions of alert issuance (time of day, incident type, congestion status, travel time)
	• Timeliness in user response actions	• Noticeable trajectory changes (to be analyzed and defined later according to field GPS data resolution and fidelity) in time versus alert issuance point • Driver reaction time comparison to total latency in alert generation • Conditions of alert issuance (time of day, incident type, congestion status, travel time)
Tactical actions (speed reduction and/or lane change maneuvers)	• Correctness and reliability of alert generation	• Numbers of alerts generated • Percentage of valid alerts • Conditions of false alerts (time of day, speed differential, congestion status)
	• Timeliness in alert issuance	• Total latency in alert reception time versus design point in algorithms • Conditions of time lags (time of day, speed differential, congestion status, incident type)

		• User feedback in real time or in surveys
	• User speed reduction after alerts are issued	• Speed change maneuvers captured before and after alert issuance, for at least a 60-second time window before and after • Speed variations in time segments within the data window • Magnitude of speed change in various time segments • Percentage of speed reduction responses • Percentage of confirmed and verified user responses - based on user real-time feedback - based on user surveys • Conditions of alert issuance (time of day, congestion status, travel time, incident type)
	• User lane change after alerts are issued	• Lane change maneuvers captured before and after alert issuance, for at least a 60-second time window before and after • Trajectory variations in time segments within the data window • Percentage of lane change responses • Percentage of confirmed and verified user responses - based on user real-time feedback - based on user surveys • Conditions of alert issuance (time of day, congestion status, travel time, incident type)
	• Timeliness in user response actions	• Noticeable trajectory changes (to be analyzed and defined later according to field GPS data resolution and fidelity) in time versus alert issuance point • Driver reaction time comparison to total latency in alert generation • Conditions of alert issuance (time of day, congestion status, travel time, incident type)
Collisions avoided	Reduction in	• Collision reports

and reduced	• Collision frequency (per counts of collisions) for individual site or collision types • Collision rate (per vehicle-mile traveled) for individual sites or crash types	• Cumulative counts of collision numbers at specific sites and particular types • Realistically difficult to observe in short time spans due to data stability and non-deterministic causal factors

3.8 Execute Safety Field Test

3.8.1 Client Software Distribution

In this task, working with our outreach ('driver') organizations using the plan developed in Task 2, we will provide, via a website, our Connected Traveler software client and download instructions developed in Task 1 to up to 1000 drivers.

3.8.2 Data Collection and Release

One major task within the field tests was to collect data on the alert situations and user feedback to enable further refinement of system design and to lay the foundation for a larger scale deployment.

The types of data that can potentially be collected include the following:

- Time of alert issued
- Traffic and Incident Status that trigger the activation of alerts
- Content of alerts presented to users
- Location of alert received (GPS coordinates)
- Snapshots of GPS traces prior and after alerts (if measurable via smartphone GPS traces)
- Post-alert braking and lane-changing response (if measurable via smartphone accelerometer)
- Non real-time user feedback, from web-based or paper-based surveys

The actual volume of data to be collected in the field tests depends on the size of the user pool and the number of test sites. The following is an estimated scope of data that can be expected:

- For Application One (Situational Awareness of Slowed Traffic Queues), the messages was only issued when users pass through the designated test sites. As a result, a relatively low number of alerts are expected. The number of reported data instances for Application One will probably be in the order of hundreds within a test period of 3-6 months. The actual volume will depend on the travel routes of users and the placement of test sites.
- For Application Two (Notification of Incident and Congestion), it will not be unreasonable to expect one or more alerts issued to each active (traveling) user every day. The number of reported data instances for Application Two can be expected to on the same order of magnitude as the number of users.

With data collected from field experiments, two primary categories of user evaluation can be conducted:

- Subjective and qualitative response and inputs from users, including:

    - Usefulness
    - Timeliness
    - Interface friendliness
- Quantitative analysis of system functionalities, including:
    - Spatial differences between hazard spots and alert reception spots
    - Time latency in alert generation and transmission
    - False-positive and false-negative alert rates, and the causal factors of these system errors
    - Reliability and variability of phone-based GPS data for the use of intended applications
    - Relative user acceptability of different interface design
    - Correlation of hazard situations and traffic conditions to system performance and user acceptability

4 Mobile Probe-Based Traffic Monitoring and Information Provision Systems

4.1 System Architecture

- Provide traffic operators with **more information** for network-wide and path-level decision support
- Provide travelers with **more options** that can avoid traffic jams and reduce commuting delays through routes, departure times, and mode changes
- Encourage both transportation system users and managers to make better **informed decisions**

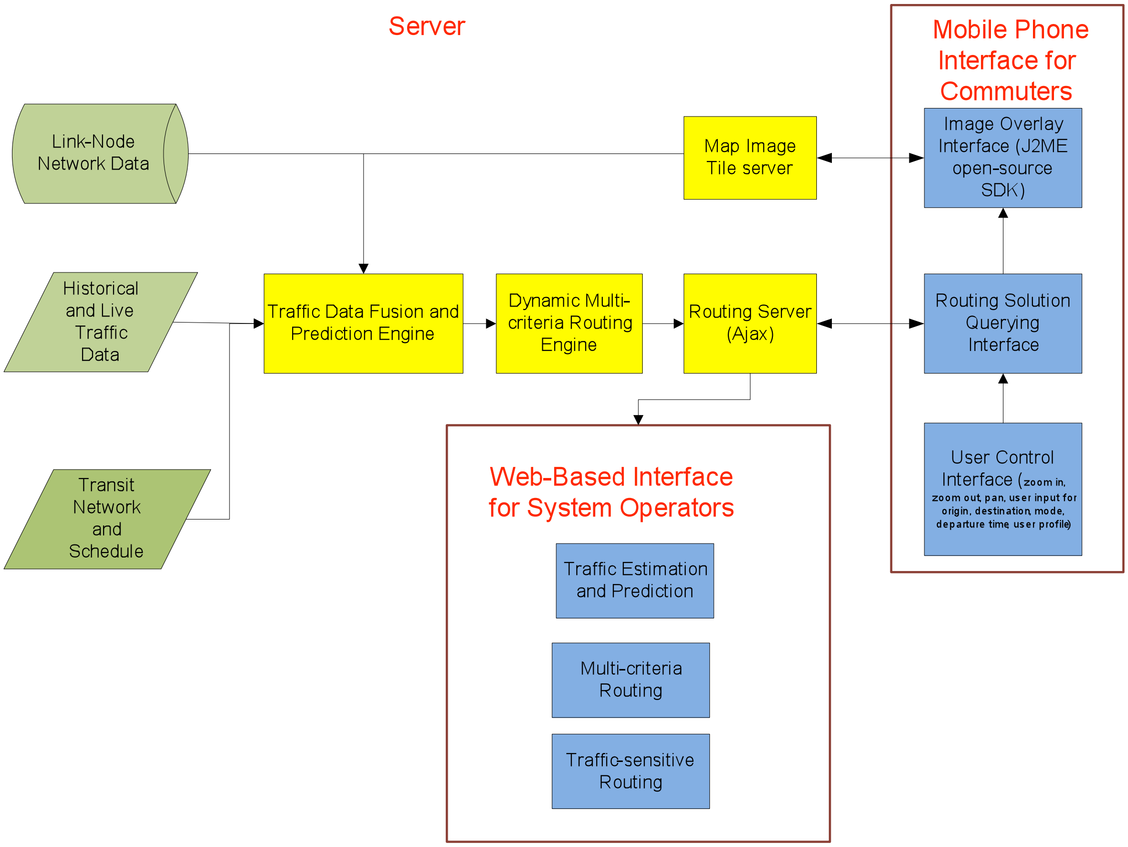


Figure 4-1: System architecture of mobile Probe-Based Traffic Monitoring and Information Provision Systems

4.2 Traffic Congestion Information Dissemination

4.2.1 Multi-criteria Routes in New York City

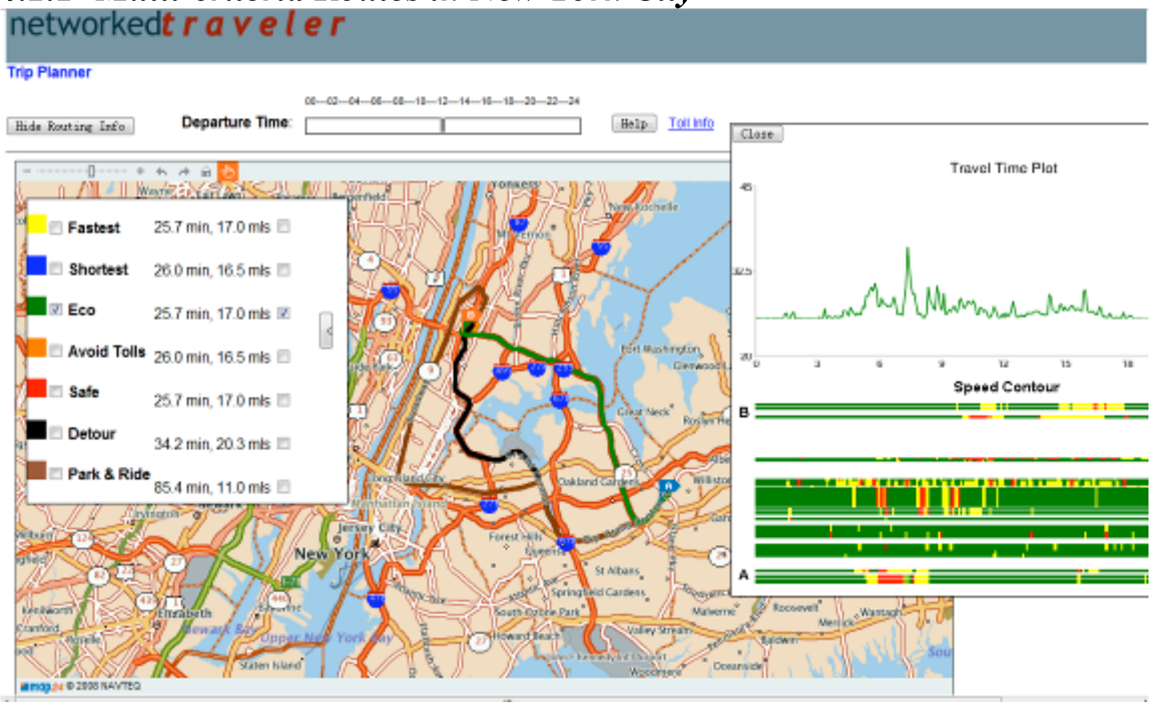


Figure 4-2: Multi-criteria Routes in New York City

Remarks:
Traffic estimation and prediction algorithm uses real-time traffic data from NAVTEQ (Traffic.com)

4.2.2 Park and Ride Route (Brown) in Salt Lake City

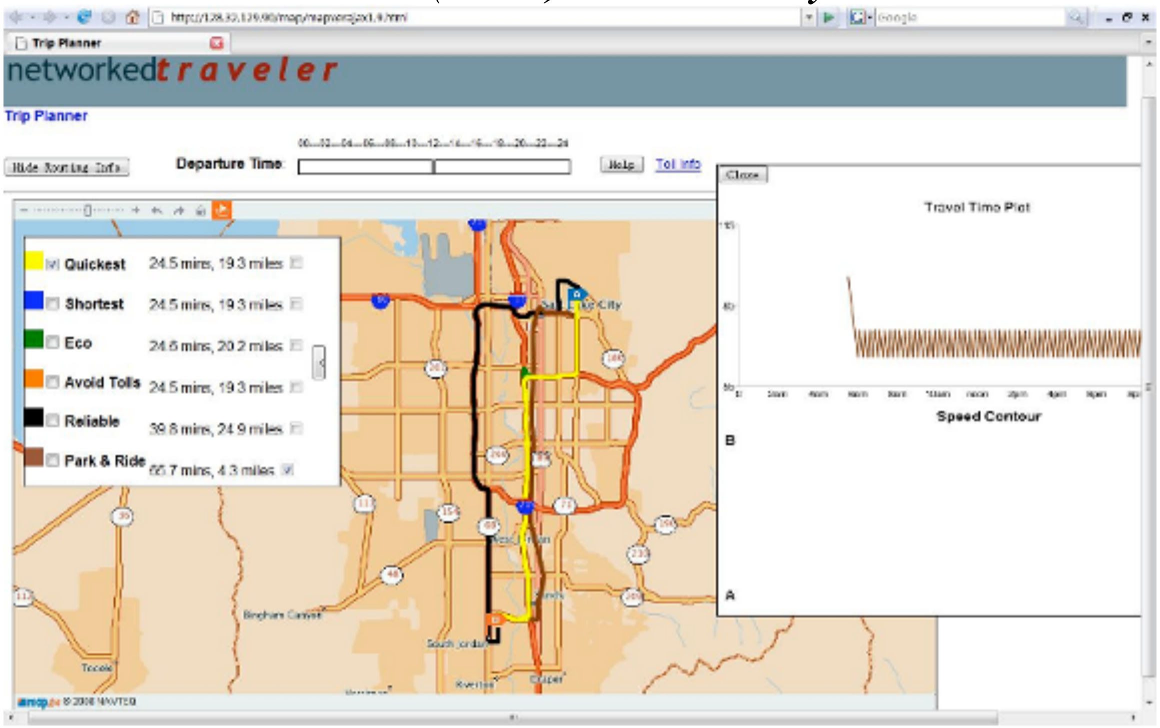


Figure 4-3: Park and Ride Route (Brown) in Salt Lake City

Remarks:
The zig-zap time-dependent travel time profile reflects the different waiting times due to transit system headway.

4.3 Mobile Phone-Based System Implementation

4.3.1 System Setup

http://128.32.129.90/map/mapverajax1.5.htm

1. Use Firefox 2, or login the PATH routing server remotely (Joel knows the user name and password). (Firefox 3 has some compatibility problems)
2. Turn on Nokia N95, Go to the embedded web browser, input the web address, wait for a couple of minutes to load the web page, save the address as a bookmark for quick access next time.

Remarks: If the Map24 API problem can be resolved on Pharos phones before this Wednesday, we will also demonstrate Pharos phone-based navigation application.

4.3.1.1 ***Introduction and Motivation***

Currently, there are a wide variety of traffic services, for example, Google Map, 511.org and GPS navigation devices such as Garmin and TomTom.

4.3.1.2 ***What is additional traffic information commuters still want?***

(1) More data coverage on arterial streets (in additional to only freeway data provided by Google Map and 511.org)
(2) Dynamic routes that can give me alternative routes based on prevailing traffic conditions at different times of day (e.g. 8AM vs. 5PM)
(3) Reliable routes that can help me arrive the destination on time (to catch a meeting), even the travel distance is relatively longer
(4) Safe routes that can reduce the probabilities of having an incidents (for a teen driver or a truck driver)

4.3.1.3 ***How to deliver information to end-users (travelers)?***

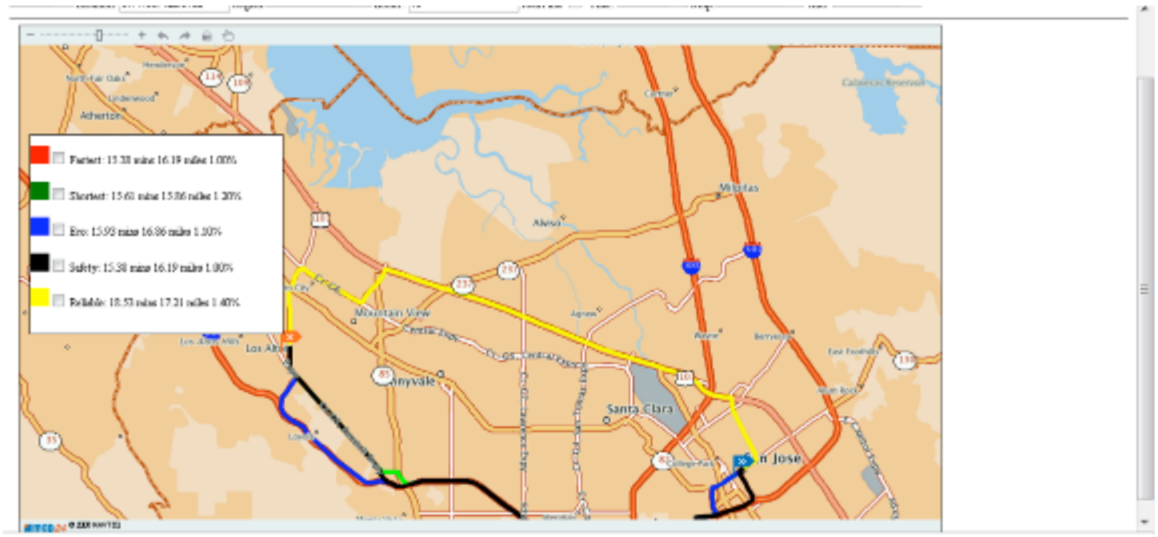
(1) Internet web browser for pre-trip planning
(2) GPS enabled cell phones, such as Nokia N95, iPhone, Pharos Phone, for on-route navigation

4.3.1.4 ***What are our traveler information services?***

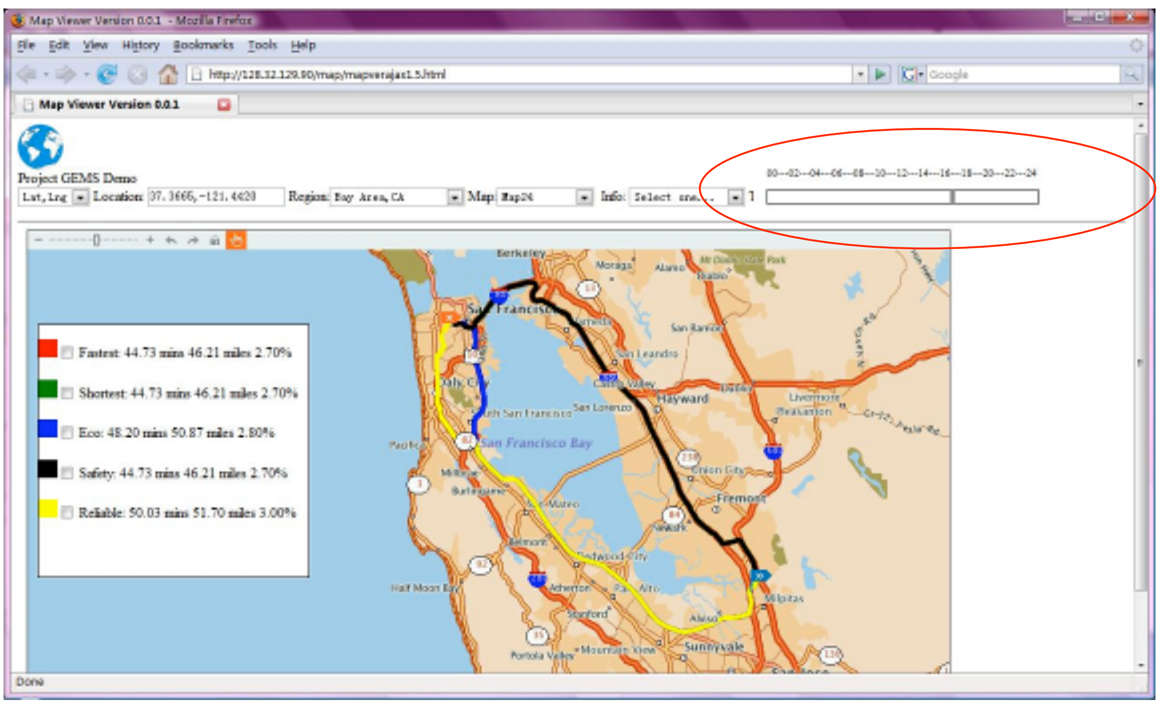
Provide different routes according to different travelers' criteria.

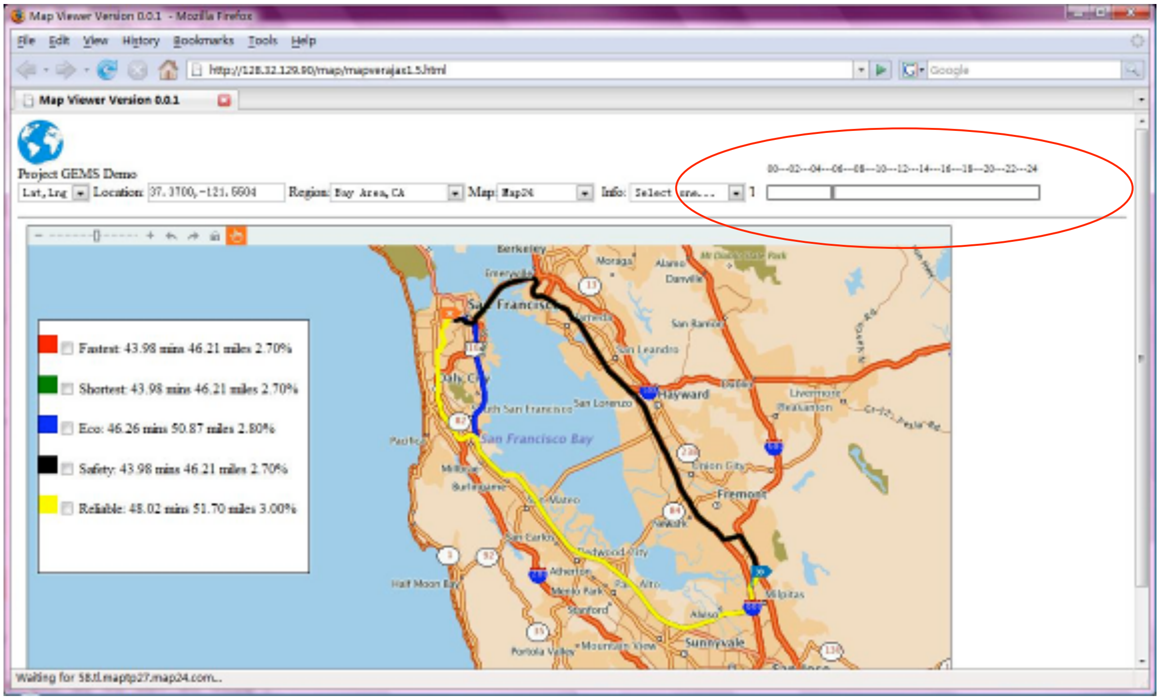
Storylines:
Support we are leaving from the San Jose Airport, then we need to go to the VII California headquarter at Los Altos to attend a meeting
>> select origin at San Jose with one mouse click, destination at Los Altos with another mouse click.

Instead of providing the single shortest path from Google Map that minimizes travel distance), the GEMS routing server also provides
dynamic least travel time path (according to real-time conditions)



>> Select the current time (or your preferred departure time) at the time slide bar, generate different paths





Notes: With limited traffic data, it's really hard to find out a situation with route change. But we can see the change on travel time.

- reliable paths (provide better travel time reliability that can help a user reach the destination on time),

- safe routes (minimizes the probabilities of getting involved in a car incident.)

>> The safe route minimizes the probabilities of having incidents (i.e. get involved in an incident) along your trip **(during a year).**

4.4 GPS Data Mining System for Safety-related Benefit Analysis

4.4.1 Introduction

Population growth and economic development lead to increasing demand for travel and pose mobility challenges on capacity-limited transportation networks. The time it takes to travel from one place to another is among the many activities that fill the busy schedules, Travel time is one of the most important factors in,the personal route choice decision, in conjunction to fuel costs and safety concerns. There are a number of factors that have an impact on trip times, to name a few, link lengths, road classes and related speed limits. In addition, some roads are more likely to have incidents, congestion, and/or construction than others. Some roads are busier during some times of the day or week than at other times. Therefore, collecting dynamic and up-to-date information about traffic patterns, and further estimating and predicting travel times along different routes accurately has been a challenging issue in mitigating traffic congestion in metropolitan areas.

There are currently a number of sensor technologies available to measure traffic conditions, including point sensors, automatic vehicle identification readers and the Global Positioning System (GPS). With significant failure rates, existing in-pavement and road-side traffic sensors are typically located on a small subset of freeway links, and accurate travel time and traffic flow information on ramps and arterial corridors are critically needed but very costly to collect. With the number of GPS enabled devices available today, cellular phones in particular, collecting and analyzing the GPS data that such devices receive provides for a more precise estimate of the actual traffic conditions on all the links throughout a trip. With two-way connectivity through a built-in Wi-Fi or cellular connection, emerging navigation systems allow a network of equipped drivers to anonymously share their speeds and locations, obtain up-to-date traffic flow information, and more importantly make smart route decisions..

Map matching is one of core data processing algorithms in this process, which uses latitude, longitude, and bearing of a probe vehicle to search nearby links (roads) and further determining which one the vehicle. In some cases, there are multiple possibilities for map-matching results. Thus, there are a number of different methods used to find the closest or most probable match. Quddus et al. categorized **current map matching algorithms for transportation applications** as: geometric, topological, probabilistic, and other advanced techniques. In particular, the geometric algorithms use the geometry of the links only; topological algorithms use the geometry of the link as well as the connectivity of the network. In the probabilistic approach, an error region is first used to determine matches, and the topology is then used when multiple links or segments of links lie within the error region created. Advanced algorithms include Kalman Filtering, Bayesian interference, Belief Theory, and fuzzy logic. In this study, we use the probabilistic approach by using an error region, and the topology of the road network to determine the correct path taken at an intersection.

What are our advantages compared to their algorithm?

The importance of offline map-matching algorithms
Difference between real-time and offline map-matching:
Data availability,
Look-up window,
Quality requirements.

Innovation:
1) Grid-based index system: pre-processing

4.4.2 Problem statement

- Input
    - Network data: Node, link, shape point
    - Raw GPS data sequence
        - Location (lng/lat), point speed, bearing, time stamp
- Output
    - Matched road links
    - Link travel time

    - Partial trace handing (for partially traveled roads)
- Data quality and confidence evaluation

4.4.3 Road Network Decomposition and Grid Construction:

The network structure used in this study consists of nodes, links, and segments, where links are made up of segments and connected by nodes. Other useful link data includes: the length, the speed limit, and the travel direction. For this study, we use a high quality road network that is provided by NAVTEQ. Due to the amount of data included in a high quality network, it is helpful to have a method for decomposing a regional map to different subareas.

The network is divided up in such a way that the network data needed for the map matching process is easy to acquire and that the set of data acquired is not too much larger than the set of data needed. This task is accomplished by setting up a grid system. The grid system is both efficient and easy to use. Access to it requires a simple conversion of the GPS point's latitude and longitude values to u and v values. The u and v values are the index of the grid cell in which the GPS point lies. With this index, we are able to get the contents of the grid cell (a list of nearby links) from a HashMap.

The grid used in this study is a basic, 2-dimensional grid. The grid covers a rectangular area that is between the minimum and maximum latitude and longitude values. These values represent the area covered by the network. The values are then converted to u and v coordinates using basic mathematics. For example, the pair (minimum latitude, minimum longitude) is converted to the u, v pair (0,0). The maximum latitude and longitude are converted to (gridwidth, gridheight). The values between the maximum and minimum values are linearly interpolated.

The user is able to specify the height and width of the grid. The height and width are chosen based on what works the best. In this study, for example, a 512 x 512 grid works great, a 32 x 32 grid performs fairly, and a 16x16 grid is much less efficient.

Next, the links are added to the appropriate cells in the grid. This is done by adding the link id of each link to the cells within the error range of each of its segments. An error range is necessary in order to overlap the areas covered by the cells. A segment that lies close to the border of the cell is then visible to the GPS point that lies close to the border of a neighboring cell. Adding the segment four times, as shown in Figure 4-4, covers the error range of the segment.

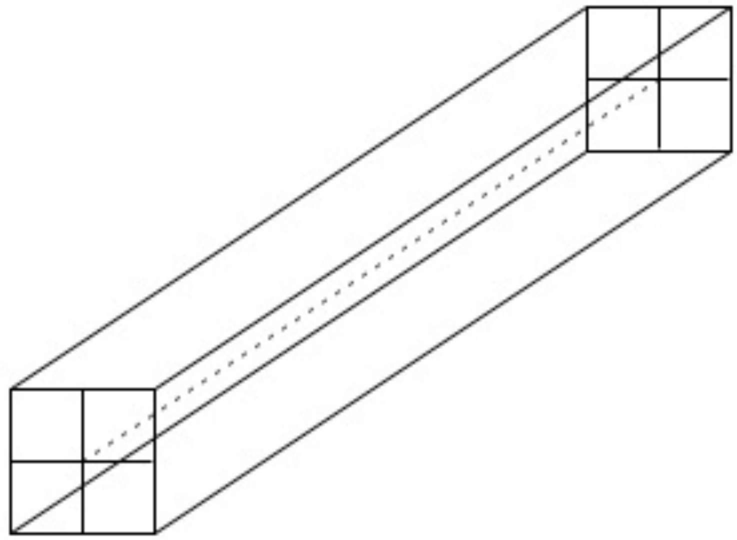


Figure 4-4: Link segment as a dashed line as well as the four parallel segments surrounding the dashed line.

Begin adding a link segment to the grid by determining the u and v values of each of the endpoints. Then, start at one point and trace the segment through the cells to the other point. To do this, begin at one of the endpoints, determine which side of the cell the segment crosses, and then step into the next cell. Continue doing this until the cell of the other endpoint is reached, see Figure 4-5.

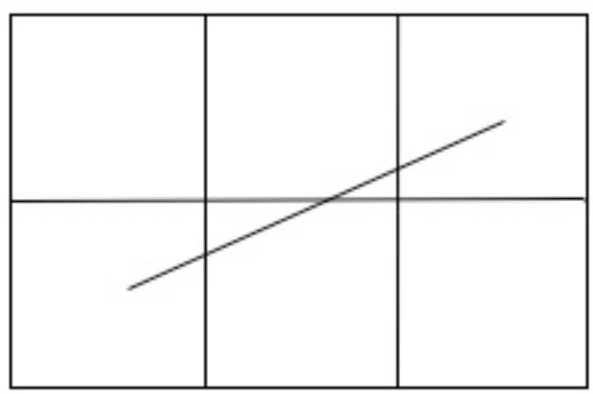


Figure 4-5: Representation of a link segment and the grid cells it lies within.

4.4.4 Map Matching:

4.4.4.1 __Offline Map Matching__

The probabilistic map matching algorithm used in this study is described below. The algorithm is used as an offline process in order to guarantee high quality results. Using an offline process rather than an online process allow for the option of looking at sufficient data either before or after the current GPS data in order to more correctly determine the match. The window of data used is adjustable and varies depending on the confidence of the match.

4.4.5 The Initial State

In the beginning, the map matcher is in a lost state. That is, the map matcher doesn't have a set of possible links that the sending device could be on. Therefore, the set of possible links consists of all of the links in the network. The latitude and longitude values of the GPS points are converted to grid coordinates and used to access the grid. The contents of the grid cell that a GPS point lies within become our set of possible links. Of those links, the set of links that match the GPS data (are within a certain error threshold in distance and angle) are each assigned to a path.

The path structure is used to keep track of where the vehicle has been, the possible next links, and the time the vehicle spent on the link. The path is a linked list where each next (or child link) contains a pointer to its previous (or parent) link. Define a time-out period (e.g. 10 second) for canceling a path.

The initial state is the state at the beginning of the execution, as well as every time the GPS points fail to match any link for 10 seconds or more (at which time the map matcher is considered lost). If there is a gap in the data that is longer than 10 seconds, the paths are removed and the map matcher returns to this state. In this state, there are no remaining paths and it is necessary to retrieve the possible links from the grid.

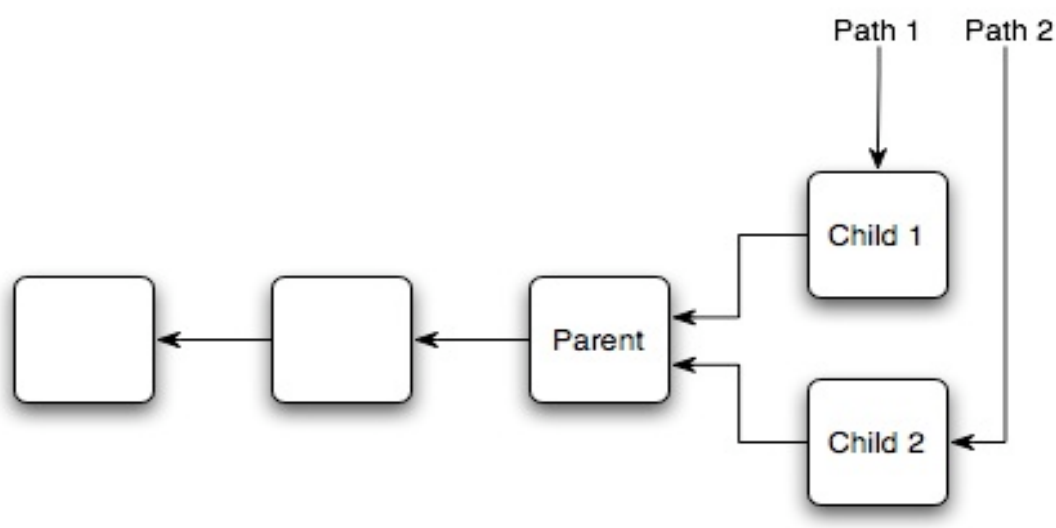


Figure 4-6: An example of the path representation. Each child points to its parent.

4.4.6 Intersections

4.4.6.1 Topology

The map matcher is not lost when there are possible paths for the GPS data to map to. At the end of each path there is a parent and a child. The parent is the link that the GPS data has matched most recently. The child is one of the links that is connected to the parent by the intersection. When the data matches a child link, the child then becomes the parent and new children are added.

To add new children: retrieve the contents of the grid cell that contains the intersection node and add the links that are connected by the intersection. This is done by comparing the to-node of the parent link to the from-node of the child link. Thus, the geometry of the road is preserved.

4.4.6.2 Closest Match

The GPS data always matches the closer of the parent and the child. The distance and the angle are used to measure the closeness of the links. When the links are similar in angle, the distance is the major factor in determining the closeness. When the links are very different in angle, the angle has a higher weight. This method is also used when determining which segment of the link is closest.

4.4.7 Cycles

A cycle occurs when two different parents have the same child. This is when two different possible paths connect at an intersection. When the GPS data matches a link that comes out of that intersection, it becomes hard to determine which of the paths the vehicle actually took. A closeness measure can be taken to determine the most probable path, but the confidence becomes lower than 100%.

The map matcher removes cycles. The travel time of the links on a cycle is not 100% certain, so the data isn't needed and the appropriate links are removed from the path. The two paths are joined.

4.4.8 Travel Time

A path is committed, or a travel time is written to the database when there is only one possible path. In this way, the travel time was calculated only for the correct path. The path keeps track of the time that travel on each link begins. The times are linearly interpolated based on the along-link distance between the projected positions of the current GPS point and the previous one. This is incorporated in order to accommodate cell phone data.

Each link is committed (the travel time is calculated and written to the database) when it is no longer to parent or the child of the path and there is no other conflicting path.

4.4.9 Cellular Phones:

*4.4.9.1 **<u>Sampling Interval and Data Quality</u>***

The quality of data produced by a dedicated GPS device is quite good. The small errors that exist within the data are accounted for within the map matcher, with not too much trouble. In this study, the data is received at a rate of 1 sample per second. However, when receiving data from a cellular phone, a sampling rate of 1 per second is quite high. A sampling rate of 1 per 5 seconds is more feasible, but will require some extra support from the map matcher.

One benefit to having a sampling interval of 5 seconds is that the path taken is still discernable, so you can achieve the same results with less processing of data. Even so, there are some complications that are introduced by this interval. First, travel times either need to be interpolated, or accepted as accurate within 5 seconds. Second, although most links in the network take more than 5 seconds to travel, some links take less. When considering the geometry, it is important to allow for the possibility of skipped links. More specifically, a driver could travel a link in less than 5 seconds and have no GPS data sent while on that link. The result is that the link is skipped.

*4.4.9.2 **<u>Adjustments</u>***

Without allowing for skipped links, the GPS data is matched to a path by determining whether the point is closer to the current, or parent, link, or any of the next, or child, links (the links whose from node is the parent link's to node). There is one path for each pair of parent and child links. For example, if the current GPS point matches one parent link and there are three child links, then there are three paths. Although there are three possible paths, as long as the GPS point matches the parent link (which is common among the three paths) there is still only one path that is being matched.

To account for skipped links, it is necessary to look ahead a little bit further in the geometry. One way to do this is to add next links or grandchildren to the child links. It is more complicated and tedious to always add grandchildren and find the closest of three links. A solution to this is to try to match the GPS point to the parent and child links and add grandchildren only when necessary. Please refer to the path algorithm diagram in the appendix.

4.4.10 Allowing for Multiple Devices:

To avoid mixing the data from the different devices, each data provider is assigned a unique identification number. Data from the different devices may come simultaneously. The Map Matcher stores the data according to its identification number.

4.4.11 Future Work:

There are a number of situations that would cause a car to move very slowly or stop along a road. Among the situations are: cars troubles, stoplights, stop signs, pull over, congestion, and incidents. For some of these situations, such as congestion and incidents, the computed travel-time is very valuable. In other situations, such as pulling to the side of the street, the travel-time should not be used. There is much work that can be done in distinguishing the difference between the situations. For example, using an error radius of 25 meters makes it difficult do detect whether or not the car is on the road, or pulled over to the side. Distinguishing the difference between the situations will improve the quality of the calculated travel times. Another way to improve the travel time in these situations is to remove the outliers.

For links that are parallel and close together (within the error range of each other), the travel times given by the current implementation aren't computed if the link traveled can't be determined with 100% confidence. For such links, it is possible by design that the travel time is never computed. If one link is a lot closer to the given GPS points than the other, there is a high probability that the closer link is the correct link. Considering this situation, the current map matcher can be improved by assigning a confidence level to the travel times and using the high confidence once when there are no full confidence (100%) travel times available.

For the GPS points that are not matched to a link that has a valid travel time, the closest link is assigned. An improvement to this design is to assign it to the highest probable path. This eliminates the case where one GPS point would matches one link, a second GPS point matches another link, and a third GPS point matches the same link as the first, where the first second and third GPS points are ordered by the time at which they are received.

4.5 Technical Support for System Field-testing and Deployment

Web Service Performance Test

- **Response time**: How fast the web service is running for normal requests
- **Load test**: How the web service performs in a high traffic condition (maximum loading condition)
- **Stress test**: How the web service responses in an over-loaded environment
- **Virtual User (VU):** Used to simulate the clients of the web service, usually works iteratively to simulate continuous requests.
- **Transaction:** A response the virtual user received from server side

System Capacity

- User sends request every 30 seconds
- Assume the user requests come in as Poisson distribution
- System capacity = (# of request handled per second) x (30 seconds)

Oracle Testing Application Testing Suite

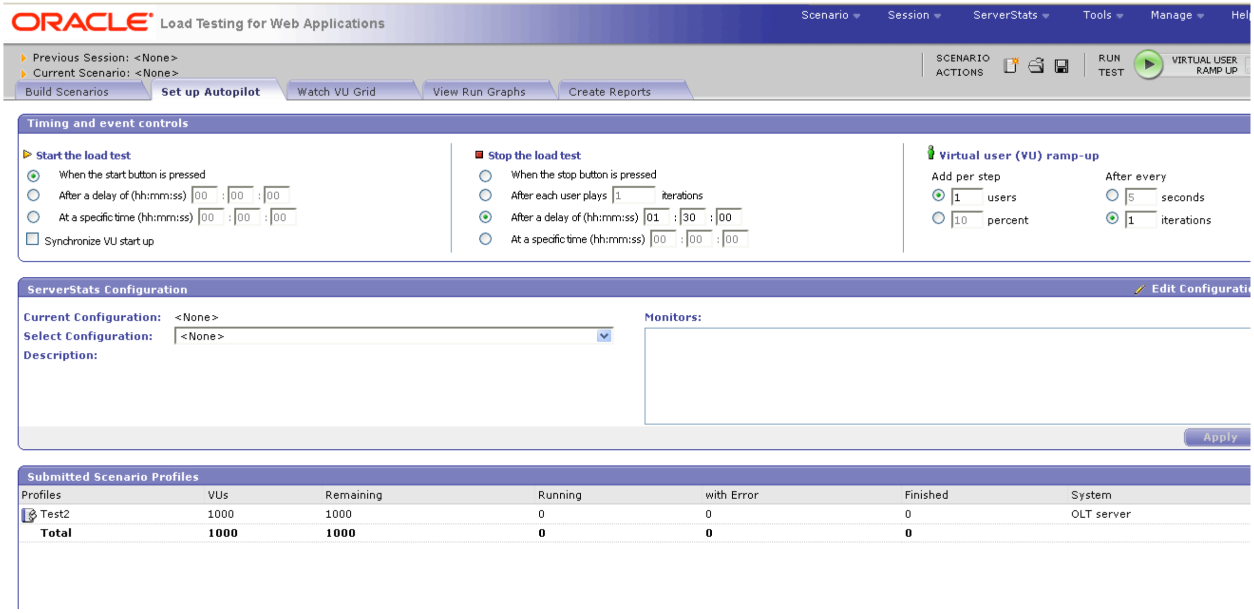


Figure 4-7: Routing engine web Service performance test

Routing Service

- Testing Environment:
    - CPU: Intel Core 2 Due 1.8 GHz
    - Memory: 2 GB
- Testing Scenario:
    - 90 VUs
    - 2 sec iteration delay
- Testing Results:
    - # of request can be handled: 25 – 30 per second
    - System capacity: 750 – 900 users

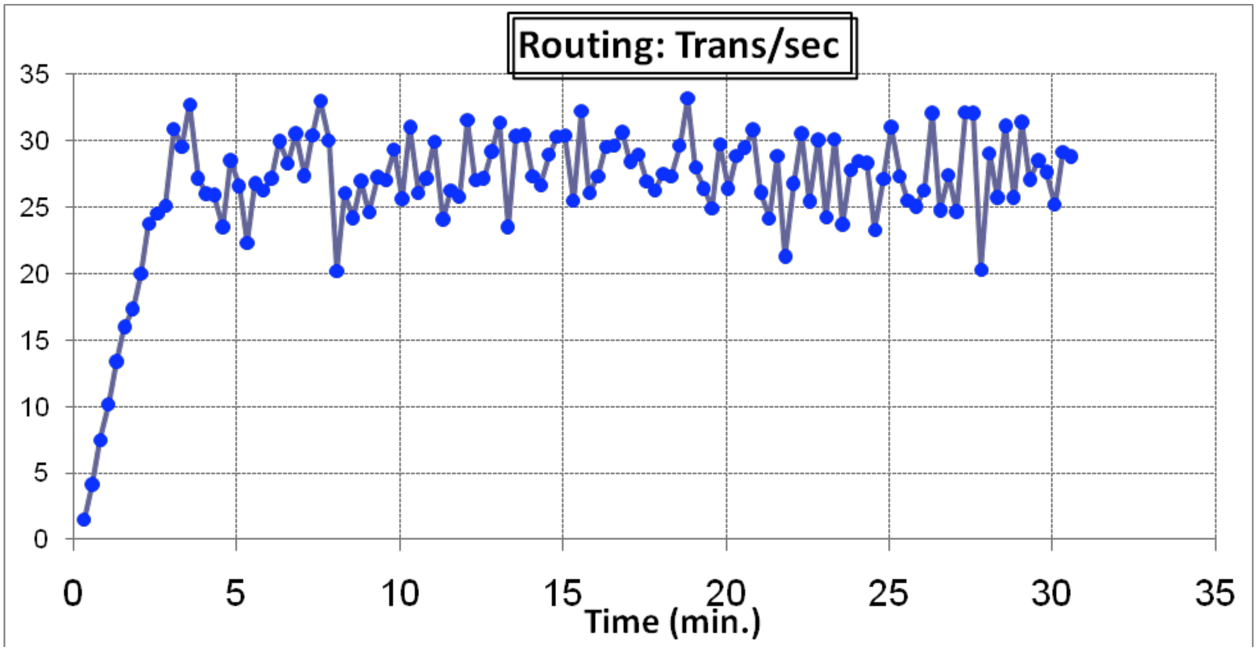


Figure 4-8: Routing engine performance results, number of transitions per second

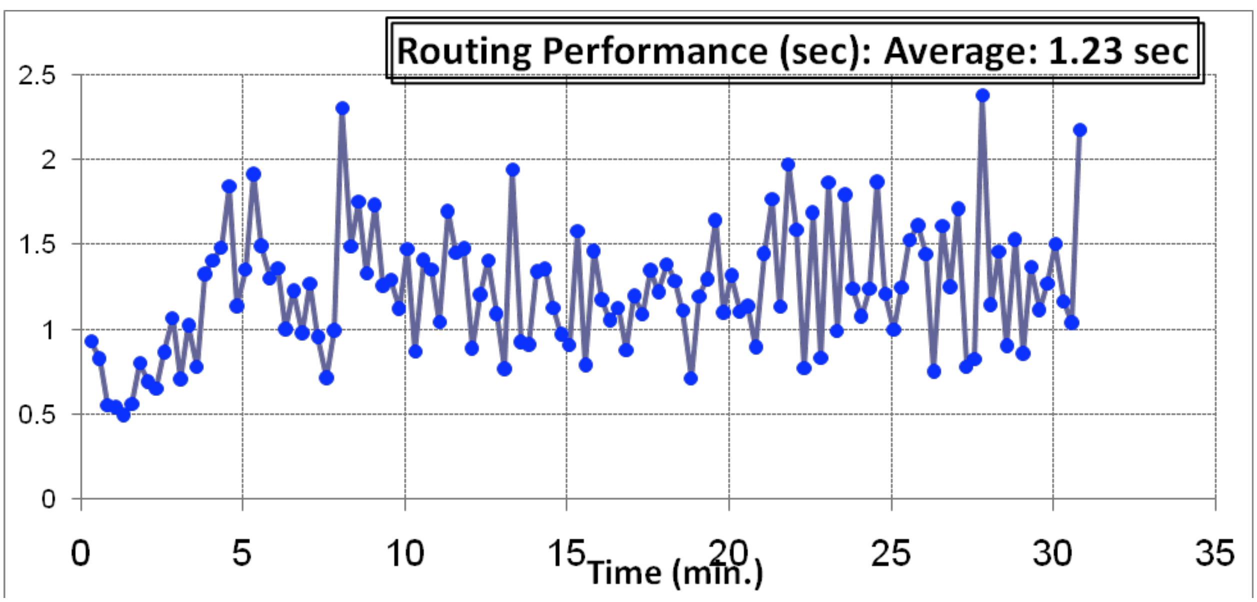


Figure 4-9: Routing engine performance results, responding Time

Over-loaded System Performance

- With over-loaded request traffic, the responding time increases dramatically

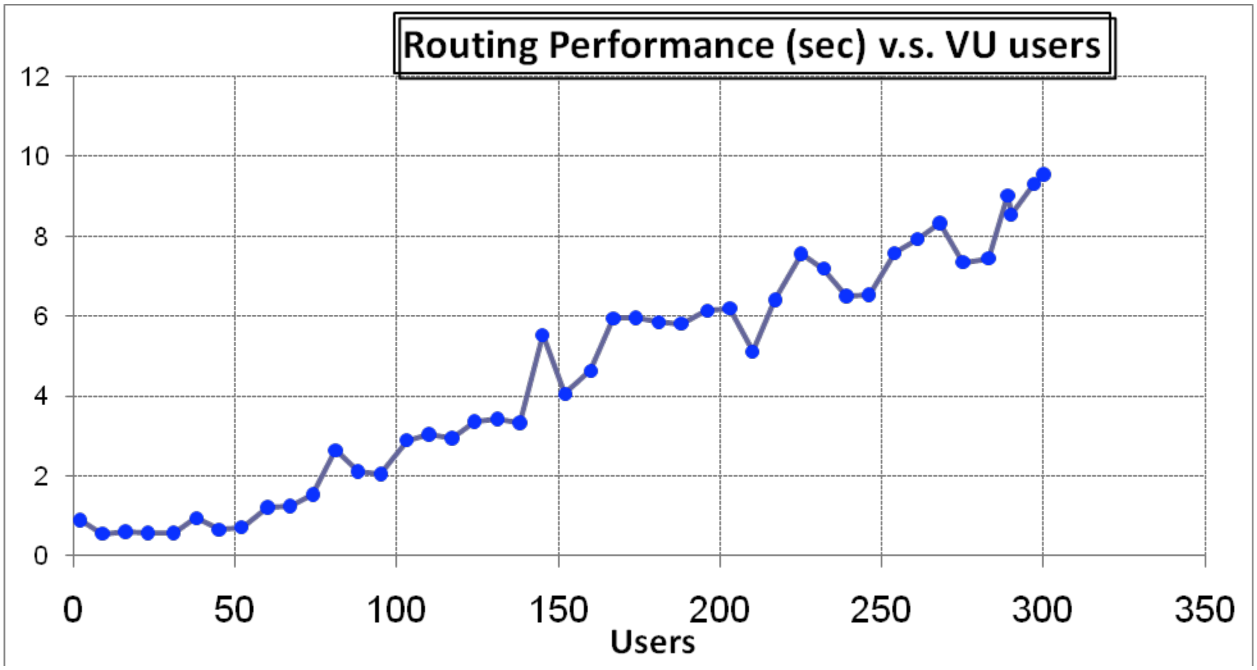


Figure 4-10: Routing engine over-load performance results, response Time

Map-matching Service

- Testing Environment:
    - CPU: Intel Core 2 Due 1.8 GHz
    - Memory: 2 GB
- Testing Scenario:
    - 60 VUs
    - 2 sec iteration delay
- Testing Results:
    - # of request can be handled: 22 – 24 per second
    - System capacity: 660 – 720 users

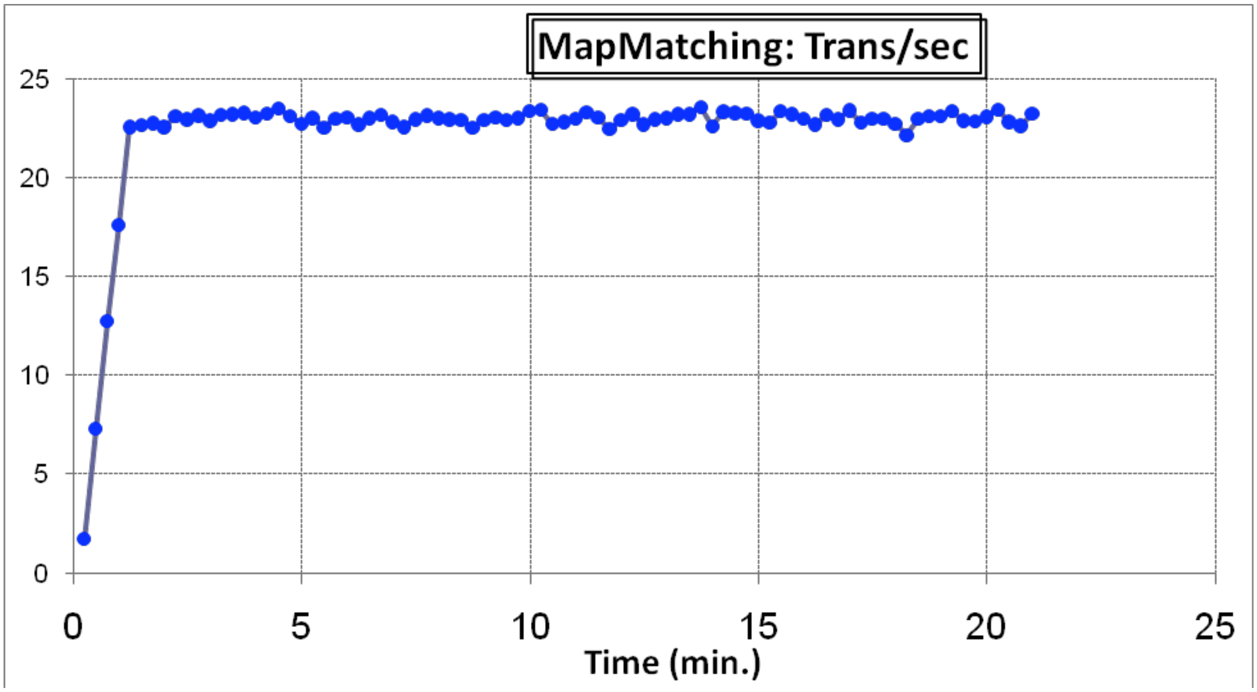


Figure 4-11: Map matching engine performance results, number of transitions per sec

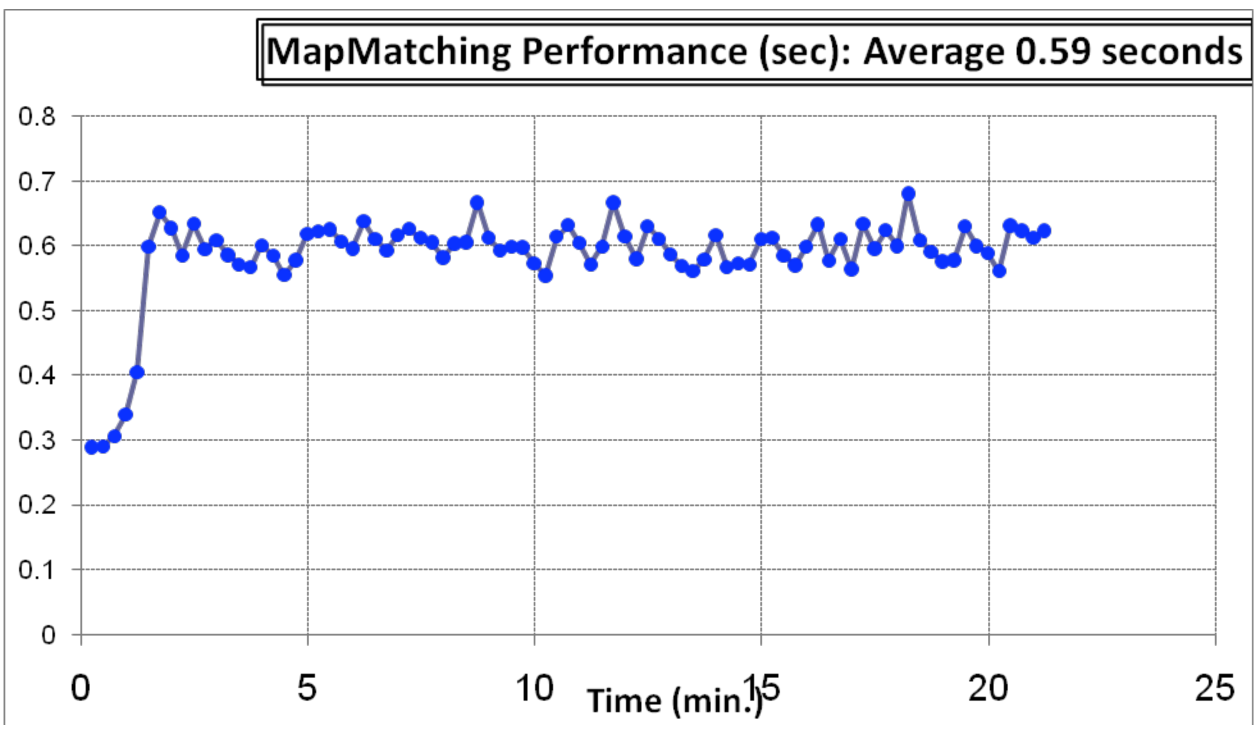


Figure 4-12 Map matching engine performance results, response time

5 Transit System Development

5.1 Experimental Design

5.1.1 Background

A California PATH research team is preparing the deployment of transit applications under a Networked Traveler project, in conjunction with the US DOT SafeTrip-21 efforts. This document describes the premise, hypotheses, rationale, and the approach for conducting field experiments, with the goal of assessing the validity and usability of the proposed transit applications.

5.1.2 Premise

(1) If travelers are better informed of the travel options in real-time, including driving, transit and mixed mode, they was more likely taking transit.

Transit has become increasingly viable for travelers as a result of gas price hike. APTA reported a 4.36 % increase in ridership nationwide in 2008 compared with a year ago due to the gas price hike. For similar reasons, the ridership increase for rail is about 12%. The ridership data provided by transit agencies in the Bay Area is consistent with national statistics and, promisingly, ridership continues to grow despite the fact that gas prices have become lower. While the gas price is the key factor in causing a mode shift for many, the fact that riders stay with the transit mode indicates that most travelers may not know their transit mode option as an alternative before being ‘triggered’ (in this case, by the pocketbook) for mode transfer. Once triggered, these former drivers have stayed with the transit. It appears that the mode shift ‘experiment’ due to gas price hikes is due to knowledge and information about the transit alternatives. It is therefore hypothesized that this is such knowledge will help attract riders and that transit can be a viable, realistic mode for commute corridors where travel time for transit is competitive.

(2) If transit buses can be used as probes, travel time information for freeways and arterials can be achieved.

Transit buses frequently travel on freeways and arterials, offering potential as probe vehicles for travel time estimation and incident detection. Once proven, buses as probes can be feed into the Bay Area’s 511 traveler information system to provide both the real-time freeway traffic condition information and additional predictive travel time information for local roadways and expressways. It is recognized that buses may not operate consistently with traffic. Buses travel on HOV lanes on freeways, where available, and stop at bus stops. ‘Filtering’ algorithms need to be developed to remove the travel pattern specific or to associate bus travel pattern with general traffic therefore to achieve a good estimation of traffic condition and travel time. Under this task, these filtering algorithms was developed and field tested.

5.1.3 Transit Applications

Two transit applications was evaluated in the field, including bus as probes and dynamic transit information.

Dynamic *en route* transit information: Dynamic *en route* transit information, when presented together with highway condition information (e.g., congestion state, travel times), may be very useful for travelers as the make mode choice decisions. In this task, we will use real-time Automatic Vehicle Location (AVL) data collected from transit buses/trains and provide through mobile devices real-time *en route* transit information on bus arrival, connection information to other links and trip time.

Buses as probes: Bus probes offer potential data gathering and operational benefits to use of automobiles as probes, because all trip data collected on transit vehicles can be utilized in their entirety. Buses operate along fixed routes, and there is no privacy concern with revealing bus location and ODs. Moreover, significant numbers of transit buses in metropolitan areas are instrumented with GPS and communication systems. For these reasons, there is significant value in using the GPS instrumented transit buses as probes to provide traffic and travel time information.

5.1.4 Transit Applications Test Hypotheses

For each of the applications that are to be tested in the field, the hypothesized outcomes, expected benefits and anticipated user responses are summarized in Table 5-1.

Table 5-1: Transit Application and Test Hypotheses

Application	**Applicable Situations**	**Hypothesized Outcome**	**Expected Benefits to and Responses from Travelers**
En-route traveler information	• Traveler catches a train/bus at a station	• Travelers benefit from the real-time bus arrival information such that he/she can catch the next bus/train with certainty	• Travelers provide favorable assessment of the next bus/train information • Travelers reduce wait time at stations • Travelers avoid unnecessary rushing • Traveler stress is reduced
En-route traveler information	• Traveler needs to alight at his/her destination station	• Travelers was informed the destination station , so that he/she can alight the bus/train stop	• Travelers provide favorable assessment of the destination bus/train information • Travelers avoid anxiety caused by unfamiliarity of the bus stops • Travelers avoid missing the destination stop

			• Travelers avoid rushing at the destination stop • Travelers benefit by avoiding unecessary wait time at stations
En-route traveler information	• Traveler needs to transfer to a different bus/train	• Travelers was informed the destination stop so that he/she can be prepared to alight • Travelers was informed the arrival time of the transferring bus/train so that he/she can catch the transferring bus/train with certainty	• Same as above ("traveler catches a train", "traveler needs to alight")
En-route traveler information	• Traveler needs to obtain the estimated travel time	• Travelers was informed the travel time so that he/she can know it with certainty	• Travelers provide favorable assessment of the destination bus/train information • Travelers can inform relevant parties of their arrival time • Travelers avoid missing their appointments
Bus as probes	• Transit bus provide travel time and congestion level on freeway	• Buses probe data are used to provide transit travel time, which in turn was provided to travelers through the en-route information described above • Transportation agencies such as Caltrans or MTC become interested in this data and decides to fuse the buses as probes data with other data to improve richness and accuracy of travel time	• Travelers provide favorable assessment of the destination bus/train information • Travelers will trust the travel time prediction • Also see above
Bus as probes	• Transit bus provide travel time and congestion level on arterials	• Buses probe data are used to provide transit travel time, which in turn was provided to travelers through the en-route information described above • Transportation agencies such as Caltrans or MTC wascome interested in such data and decide to useit as arterial traffic data that they currently don't have	• See above

The common elements given in Table 5-1 are reorganized into the chart below to capture the suite of applications .

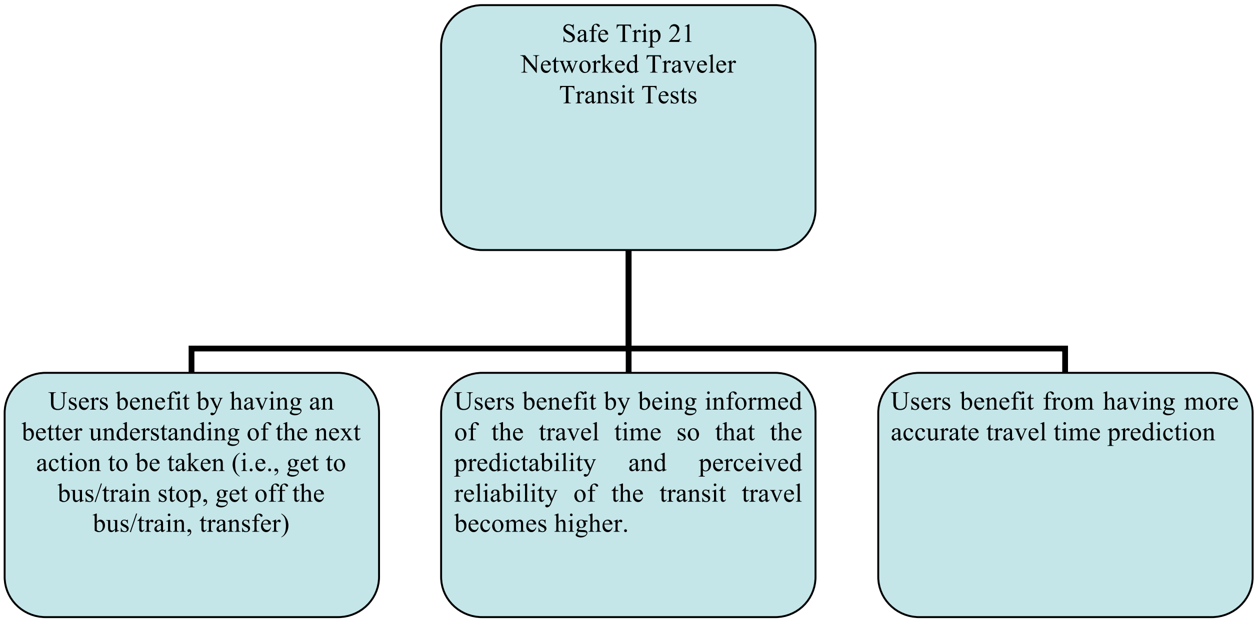


Figure 5-1: Hypothesized Expected Outcomes of Safe-Trip 21 Transit Applications

5.1.5 Validation of Test Hypotheses

5.1.5.1 <u>Transit Application and Expected Test Outcome</u>

Similar to the safety experimental design, it is expected that the user experience for transit applications was evaluated and field data was collected to explore the user needs and preferences. The applications will also serve as design validation and to identify shortcomings of the transit applications. Therefore, the experimental design of the field tests emphasize the quantitative and qualitative observations of user response.

The observable outcome for buses as probes was validated through limited tests using instrumented vehicles to show the accuracy of the travel time estimation. The expected cause-and-effect of transit information on travelers' behavior is rather difficult to directly measure. However, the methodology was to assess user's attitude toward the information that was provided, e.g.,., what information is used more often and by the most people.

5.1.5.2 <u>Functional Processes and Operating Constraints</u>

Before discussing further the necessary data collection and availability of observable data to validate the test outcome, it is critical to point out the constraints of the data acquisition process within the context of the planned field tests.

5.1.5.2.1 Functional Process of Transit Applications

The Networked Traveler project is based on the concept of using technologies that are currently available to consumers, such as GPS-enabled smartphones. The communication links available on smartphones include 3G and Wi-Fi. The set up envisioned for the planned tests can be described as follows:

(1) Users register for the service.
(2) Users enter origin and destination for the intended trip.
(3) Users activate services by choice and by preference.
(4) Once activated, the user's route and personal preference information are communicated to the system server.
(5) The system server receives AVL data from SamTrans and VTA for buses running along El Camino Real and Caltrain.
(6) The system estimates travel time on freeways and along El Camino Real based on selected bus AVL data that are instrumented with GPS loggers (with more frequent position polling).
(7) The system server provides trip time estimation, next bus stop, connection information, etc. to users and monitor user positions by receiving users' current positions periodically.

5.1.5.2.2 Constraints in Data Acquisition

The following constraints in both quality and quantity of data acquisition should be noted:

(1) The system server receives transit AVL data from SamTrans and VTA, which poll the buses every 60-120 seconds. The methods of data polling and acquit ions being may pose additional delays; therefore the update rate and the quality of transit data are not within the control of the application developers.
(2) User positions are only available through the GPS coordinates on user's devices, and therefore the accuracy and availability of user trajectory (speed and position) depends on the GPS units and environmental conditions, e.g., occlusions to GPS satellites, multipath interference, which may vary along bus routes.
(3) Transit information are only available for designated segments and specific locations in the test bed area; therefore, the measurements of background traffic are not always available to the user.
(4) Users may or may not activate transit functions even if they volunteer and register for the services.
(5) Users may choose to activate selective functions of their preferences and thus only a subset of functional results and associated data may be available from individual users.
(6) The travel information provided by the system is based on the OD inputted by the user. Hence, variations in the route actually taken by the traveler will cause perceived errors in travel information.
(7) The field test is further constrained by the fact that travel information are communicated to users through user interface implemented within the capabilities and flexibility allowed by phone-based devices.

5.1.5.3 Measure of Effectiveness

The planned field tests of the transit applications are also on a limited scale, and within a limited period of performance and scope. In order to establish framework and methodology to conduct the system assessment such that effectiveness and usefulness of transit applications can be measured, the following table illustrates how a matrix of measure of effectiveness are constructed.

Table 5-2: Anticipated Test Outcomes and Measure of Effectiveness

Expected Test Outcome and Driver Responses	Measures of Effectiveness (MOE)	Parameters and Variables to Assess MOE
Public awareness of transit applications	• Spectrum of project partnerships	• List of partners in project • Scope of participation by partners • List of participating organizations outside of project team
	• Scope of community participation	• Number of participating users • Number of data samples collected in field tests • Percentage of positive feedback by users
	• Outreach efforts	• Sessions of activity reports held in public forums and conferences • Technical papers presented • Reports of media events
Favorable user experience and positive user feedback	• Willingness to participate and to maintain continual use of applications	• Number of participating users • Periods of active usage • Continuity and frequency in activating applications • Percentage of positive feedback by users
	• User feedback to surveys and questionnaire on - Function usefulness - Function acceptability - Timeliness of alerts - User interface friendliness	• User answers in surveys and questionnaires (to be detailed and designed later)

Strategic actions (routing changes or modal choices) by users	• Correctness and reliability of real-time en-route information	• Accuracy of the estimation of the information • Latency of the information
	• Usefulness of information	• Frequency of the usage of information per trip • Frequency of information usage over time
Ability of buses as probes to support Networked Traveler applications	Contributions to • Freeway travel time measurements • Arterial travel time measurements	• Estimation accuracy of travel time • Timeliness of the travel time estimation • Quality of estimation of incident reporting

5.1.6 Data collection and analysis

5.1.6.1 Application Field Test Schedule

Table 5-3 provides a list of milestones and targeted applications for this phase of Safe Trip 21 field evaluation tests.

Table 5-3: ST-21 Networked Traveler Field Test Initial Milestones

Milestone Date	*Rollout Functionality*	*Precipitating Events*
June 30th 2009	Dynamic en route transit information, Bus-as-probes system	(1) Integration with Networked Traveler Safety applications (2) SamTrans approves Caltrain instrumentation; (3) MTC provides transit information; (4) SamTrans provides access to AVL/C system;
Nov 30th, 2009	Field testing	CPHS approval 100s drivers co-recruited with Networked Traveler Safety applications
Dec 31st 2009	Dataset and report	

5.1.6.2 User Recruiting

We will recruit users in three phases, early Spring, late Spring, early Summer, by working with management from four organizations:

- SF Transportation Management Association
- Metropolitan Transportation Commission
- AAA (Northern California Automobile Association) and
- SamTrans
- Santa Clara Valley Transportation Authority
- Stanford Commuter Club (given in order of priority in recruitment)

Each user was required to have a cell phone and an unlimited data plan.

5.1.6.2.1 Targeted Users

For the transit information application, the targeted users are those who commute using transit along US-101 corridor, including the main transit routes of VTA 522 BRT, SamTrans 390 or 391 along SR82 (a.k.a. El Camino Real), or alternatively, Caltrain. We target as our users commuters (who use transit on daily basis) to obtain more samples during the test period. For the travelers commuting between San Francisco and different cities among San Mateo County and Santa Clara County, we focus on those commuters that travel northbound to San Francisco County / City in the morning (for work). This is based on the fact that there are more transit riders northbound in the morning. (According to a Caltrain annual report, 60% of the riders on board in the morning peak are traveling northbound).

Based on the year 2000 San Francisco Bay Area census data, Table 5-4 provides the percentage distribution among travel modes to work in San Francisco from each of the six major cities/areas of San Mateo County.

Table 5-4: Percentage Distribution of Travel Modes to Work in San Francisco from San Mateo County

	Daly City-Pacifica	Brisbane-San Bruno	Millbrae-Burlingame	San Mateo-Coastside	RWC-San Carlos	Menlo Park-E. Palo Alto
Car, truck, van	74.7	81.6	81.5	78.6	78.0	81.0
Bus	9.9	6.0	3.2	7.1	3.2	2.6
Streetcar	0.3	0.6	0.5	0.0	0.0	0.0
Subway/elevated	12.8	8.2	4.5	3.7	1.2	1.7
Railroad	0.7	2.0	10.2	10.0	17.6	13.8
Motorcycle	0.3	0.7	0.0	0.3	0.0	0.0
Walked	0.5	0.4	0.2	0.3	0.0	0.0
Other	0.8	0.6	0.0	0.0	0.0	0.9

The size of the sample from which Table 4 was developed is 3,512 (from a total Census Sample of 5% of the residence) representing approximately 70,200 daily commuters to SF City and County. We do not currently have enough granularity in the data to route the commuters are taking to work, so we do not know how much percentage of the transit riders are actually using the routes we plan to provide information for. The bus routes (SamTrans 390/391, VTA 522), however, are major routes serving the corridor, with high ridership. These routes also have (several) direct transfers to Caltrain. They are assumed to be favorable choices for the transit riders. We also included static transit information from other transit routes connecting major cities to Caltrain stations.

In Table 5-5, we list the population size of the users who we have targeted to recruit. The calculation is based on the year 2000 San Francisco Bay Area Census data and also the year 2008 Caltrain annual report. Note that the Caltrain report does not include destinations, so the number of train riders are northbound only and, not necessarily those with San Francisco destinations. This is still suitable set of riders who we will test, as the final destination does not necessarily to be San Francisco. Additionally, the software was designed to accommodate this variety in destination cities.

Table 5-5: User Recruitment: Cities and Sample Population Sizes

Commuter group	**Commuters that frequently ride trains in morning peak**		**Commuters that take bus**	
destination of the commute	San Francisco		San Francisco	
Will recruit users from these cities:	Redwood City, Menlo Park, East Palo Alto and Palo Alto		Millbrae, San Mateo, Redwood City	
Number of the population of users that match the condition in the candidate cities:	Millbrae-Burlingame	~300	Millbrae-Burlingame	~800
	Redwood City San Carlos	~500	San Mateo-Coastside	~400
	Menlo Park East Palo Alto	~200	Redwood City San Carlos	~100
	Palo Alto	~ 500		

5.1.6.2.2 Sample Size Estimation for Transit Information Application

A sample is defined as one usage of the applications of one occurrence of the en route transit information provided to the user. For different tests, the sample size requirement would be different. Also there are two dimensions for the sample size, the number of participants in the test and the number of repeated experiments per participant, which are calculated below. Note that the samples to be collected per participant may vary depending on the duration of his/her participation and pattern of usage of the application, therefore the sample size calculation is based on minimum total samples needed to constitute a valid test.

The required sample size for the experimental tests are calculated based on the type of the survey, margin of error, confidence level and the candidate population size. The variation of the required sample size versus the increasing number of population size (when expected number of samples per participant varies) is plotted in Figure 5-2. Note that in this sample size calculation, we have assumed that the samples are independently distributed. While when multiple samples are collected from one participant, the results could be correlated. We anticipated to use only samples from one participant when there are differences in the sample collection conditions, such as the location when the alert is given (which station) and the route. This way the correlation among the samples from one participant can become lower.

Table 5-6: Sample Sizes for Applications

Expected Traveler Response	Sample Definition	Assumptions	Sample Size
1. Traveler provides positive response to: En route transit information- Alert of user need to get off at next bus stop or train stop	A sample is one set of yes / no feedbacks from one user, to the survey questions such as usefulness, interface friendliness, data accuracy (about predicted time to transfer stop, etc), timeliness (time the alert is given), etc.	1.Dichotomous (Yes/No) Outcome 2. Margin of error =5 % 3.confidence level =95 % 4. Population size= ~2,800 (number of samples per participant, assumed to range from 1 to 10) 5. Response distribution=50% (50% is the worst case which requires the most samples)	338 ~ 379 (or with a sample size of 100, the margin of error would be 9.6%~9.78%)
2. Traveler provides positive feedback to: En route transit information: Alert of transfer to next bus / train route	A sample is one set of yes / no feedbacks from one user, to the survey questions such as usefulness, interface friendliness, data accuracy (about predicted time of next train / bus arrival, when available), etc.		
3. Traveler provides positive feedback to : Total Trip time estimation	A sample is one set of yes/no feedbacks from a user to the survey questions about the accuracy and usefulness of this information after a whole trip is finished.		

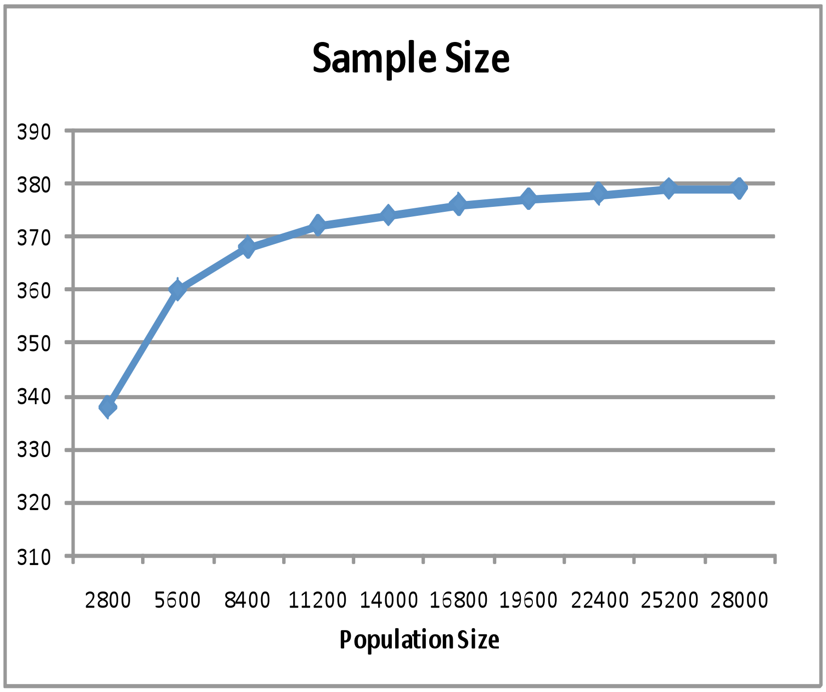


Figure 5-2: Sample Size Requirements for Different Population Sizes

5.1.6.3 Data Collection Period

As outlined in the application rollout schedule and milestones above, the transit applications was made available in June. Corresponding to this schedule, the data collection was implemented in several stages:

5.1.6.3.1 Quantitative Data

Collection of user data in response to transit applications: After the initial validation period, the data collection will continue as long as the user opts to activate the functions in his/her transit routines. If a user signs up in the early stage of field test, the data collection period can go on for 5-6 months before the conclusion of the field tests. The samples per trip vary depending on the trip length. Data samples was collected at each transit station arrival and when the traveler's acknowledgement of receiving the information.

5.1.6.3.2 User Survey and Questionnaire

User survey forms was provided in conjunction of the Networked Traveler survey form by adding a specific transit section.

(1) User information at registration:

All users are required to register when they sign up for the application services. In this registration process, certain questions about the users was posed. Answers to some questions are required, and others are optional. For example, to assess the coverage of user base, the driving distance and zip codes for origins and destinations of regular routes was useful information to have in this registration process. The detailed form of questions was provided later.

(2) On-Line Feedback:

Users was given the option of providing anytime feedback on problems encountered in the use of the applications as well as desired changes or suggestions on the applications that are offered.

(3) Survey:

One month before the project is concluded, users was asked to go through another survey. This was another milestone to assess the user experience as well as to observe any noticeable changes in user experience after exposure to the applications for an extended period. After the final survey, unless the user opts to discontinue the service, data will continue to be collected, which may be valuable for later evaluation of the field tests.

5.1.6.4 Types of Quantitative Data

The system server, where the bus movements monitoring and travel time generation algorithms are executed, will continuously capture available data streams to facilitate data analysis. The types of data that was collected include the following:

- Travel time estimation
  - Bus / Train Travel time estimation using SamTrans and VTA bus AVL data
  - Travel time estimation using buses as probes
- User acknowledgements
  - Push button of activation of the information
  - Selection of the information type (e.g., trip time)
- Content of information presented to users
- Trajectory data (GPS) of users at the time when information is triggered, relative to bus stop positions.

5.1.6.5 Types of Qualitative Data

Qualitative data to assess the user subjective experience of the applications was collected through surveys and online feedback. The types of data that was collected include the following, but the exact form and questions of survey was developed later:

- Overall impression of applications
  - Usefulness
  - Timeliness
  - Reliability
  - Issues or problems in using applications
- Traveler background information
  - Age
  - Gender
  - Familiarity or experience with smart phones
  - Driving distance daily or weekly
- Traveler experience with specific applications

   - How often traveler use the transit daily or weekly
   - How frequent traveler receive the information
   - Which information is most useful
   - Specific problems encountered with individual applications
   - Recommended changes

5.1.6.6 __Data Analysis__

The purpose of data collection and analysis is several-fold:

(1) To assess the effects of intended applications on users,
(2) To provide supporting evidence in determining the extent of success in project objectives, and
(3) To explore the weaknesses and shortcomings of implemented functions for future improvements

5.1.6.6.1 Quantitative Data Analysis

The quantitative data captured during the field tests was grouped and categorized to allow the determination of user responses.

The assessment should be performed for both the individual application and system as a whole. The elements of data analysis should include:

- Number of information-receiving actions
   - User initiation of the information
   - Acknowledgment of the of the receiving information (after information is triggered by the system)
   - Requests of additional information (after the initial information is received)
- Distribution of information generation:
   - On bus
   - On train
- Correlation of information provision with trip milestones:
   - Distance of user location to next bus stop location
   - At different stage of the trip

5.1.6.6.2 User Response Data

In order to verify the statistical significance of data representation, several critical data elements must first be scrutinized:

(1) Validity of using GPS data for monitoring user trajectory for non driving scenarios

The GPS data is phone based, therefore the trajectory traces expressed by the GPS unit do not fully reflect the vehicle actions. The data gathering process is further complicated by the resolution and accuracy of GPS data. Thus, several steps can be taken to evaluate the usage of such data:

- Preliminary laboratory and field tests can be performed to collect sample data sets for initial evaluation
- If necessary, filtering and data processing techniques can be applied to improve the usability of such data.

(2) Use GPS coordinate to determine how the information is being utilized:

- Design the interface in such a way that the riders are given options to make acknowledge of the information being
- Record the GPS location after the user acknowledge the information being received
- Correlate the acknowledgment location relative to bus/train stations
- For evaluation of all users,
    - How many users activate the service prior to the trip
    - How often the information is being used (i.e., receiving information is acknowledged)
    - Under what circumstances the information is used (e.g., at stations, on buses)
- For evaluation of individual users
    - Derive the time and location information being used
    - These variations are used to correlate with station location to see where the information is being used
    - Analysis was conducted to make inference what information is mostly used and therefore most useful

(3) User response evolution over time:

To the extent possible within the sample set, examine the pattern of information being used in different stages over time.

5.2 System Architecture

5.2.1 Architecture overview

The Connected Traveler (CT) transit system itself also needs to fit in a bigger general architecture, so that other tasks such as transit operation management, transit planning and transit maintenance management can also be integrated into a whole ITS system. Figure 5-3 is such a layered transit ITS architecture. The dynamic passenger information (DPI) system architecture described in this report is designed based on this architecture.

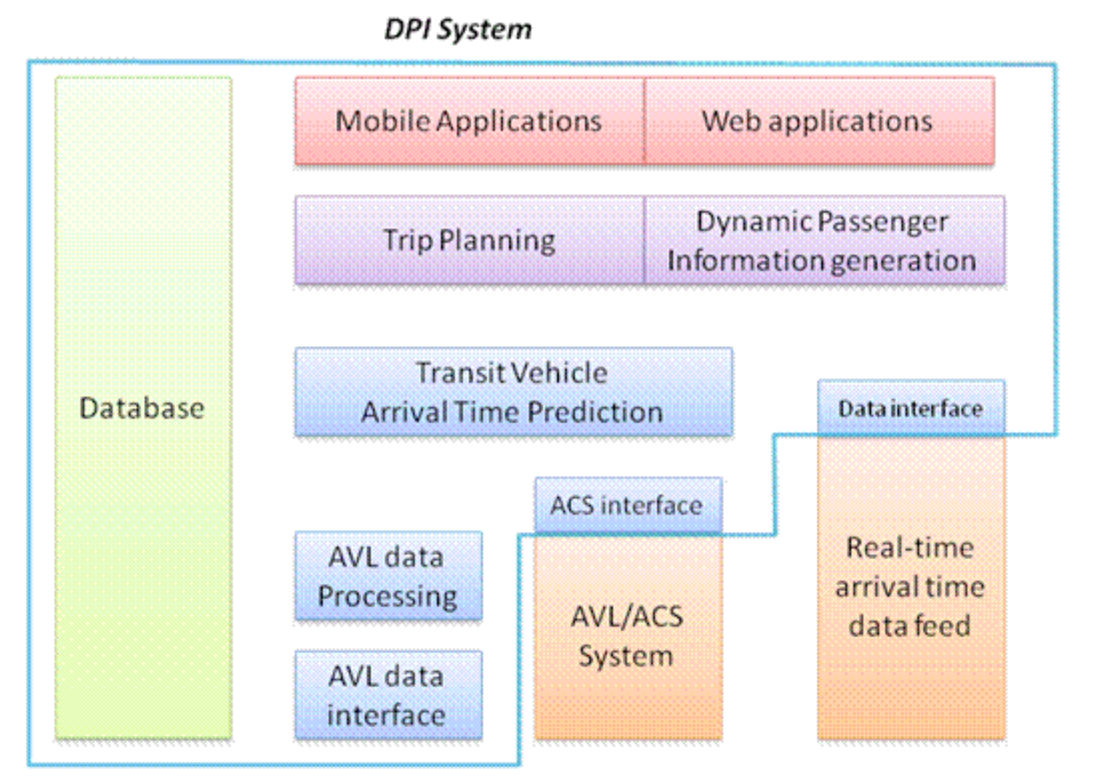


Figure 5-3: System architecture of the DPI system

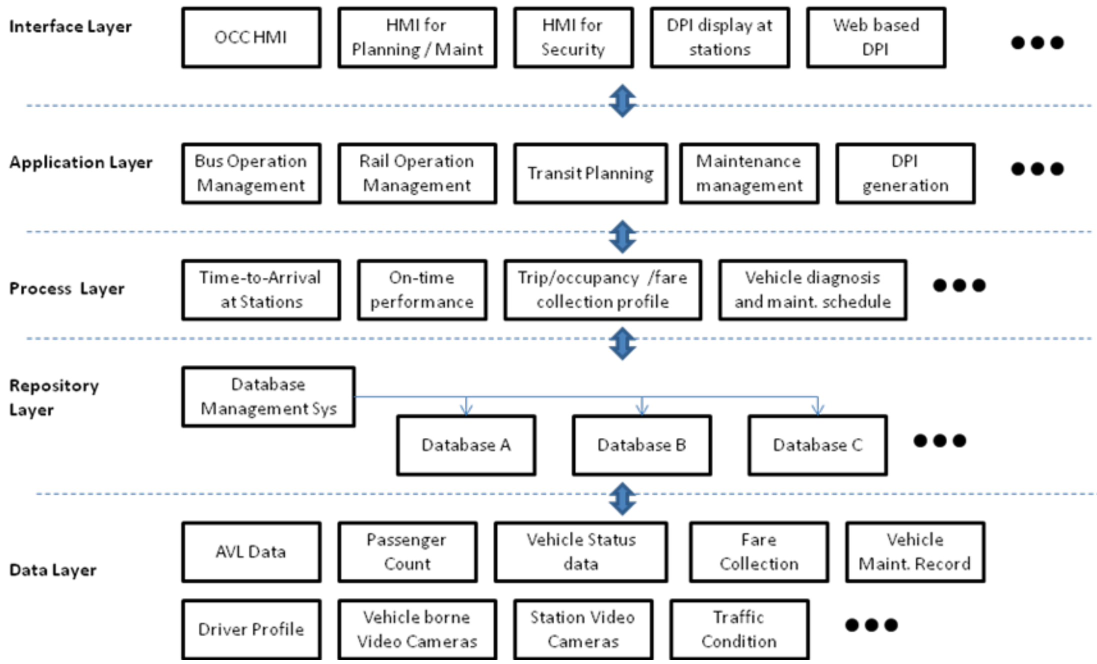


Figure 5-4: A generalized transit ITS system architecture

5.2.2 System components

The CT transit system composes of the following typical components (Figure 5-5):

- AVL system which obtains transit vehicle's location, using GPS devices;
- (wireless) communication system which provides two way linkage between a center and transit vehicles (buses or trains);
- database system which archives static transit schedule and route information, real-time AVL data as well as the generated DPI information data;
- bus-as-a-probe center that aggregates the AVL data from the buses and generates real-time arterial traffic information update. The real-time update is fed into the bus arrival time prediction, and can also be displayed at transit operation centers;
- central processor which aggregates the data and generates estimated time to arrival (ETA) for the buses / trains, generates the DPI information for various information processes (for personal information, bus stop, etc) and optimizes the routes for trip planning;
- Transit server that provides services for information in various formats and via different media.

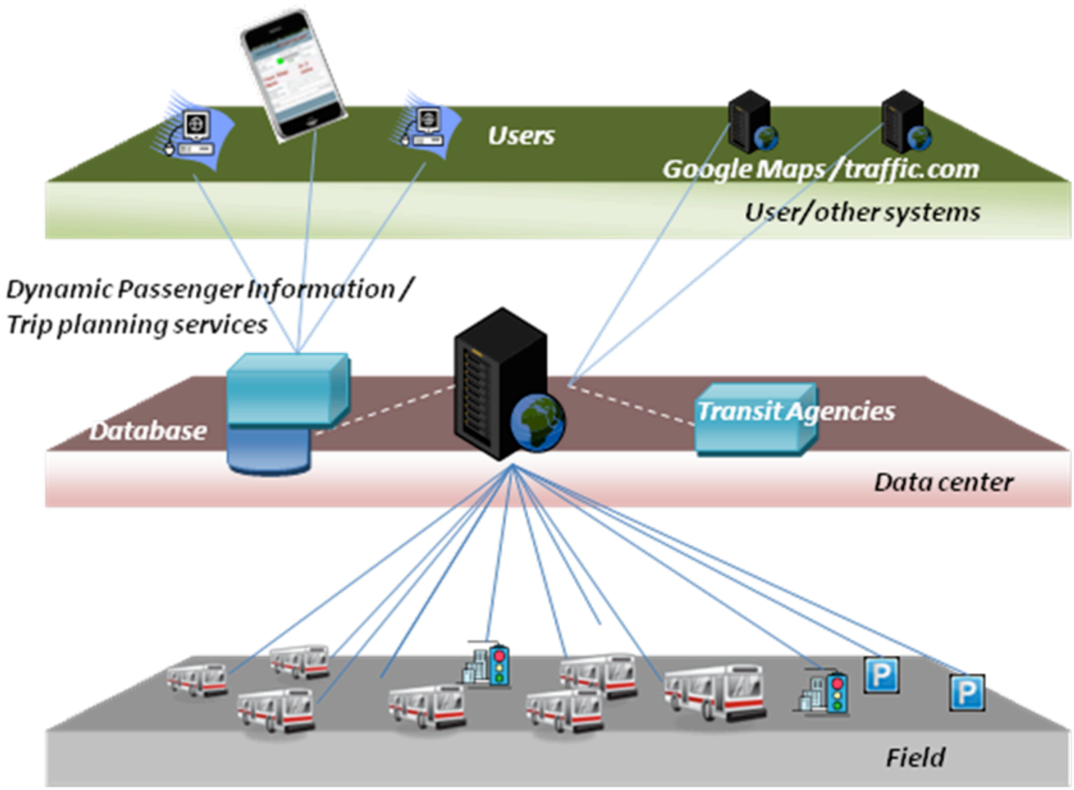


Figure 5-5: Connected Traveler Transit System Components

In summary, the system is designed to have a scalable architecture to meet different needs from existing systems of transit agencies. The design of the system architecture supports several different scenarios, is scalable, and takes advantage of existing AVL/ACS system as well as the existing real-time information system to provide a highly flexible solution for DPI system.

5.2.3 Transit data elements and related systems

The transit system and corresponding systems of CT transit are listed in the following table.

Application	Brief Description	Associated Systems
Passenger Information	Dynamically provide next bus and other location-based services to the passenger	Traveler Information System, Parking Management (PRK) GIS and AVL
Traffic probe	Traffic data collection using bus as probe	AVL System, GIS
Transit Signal Priority	Transit Signal Priority Process	AVL Data Repository , GIS

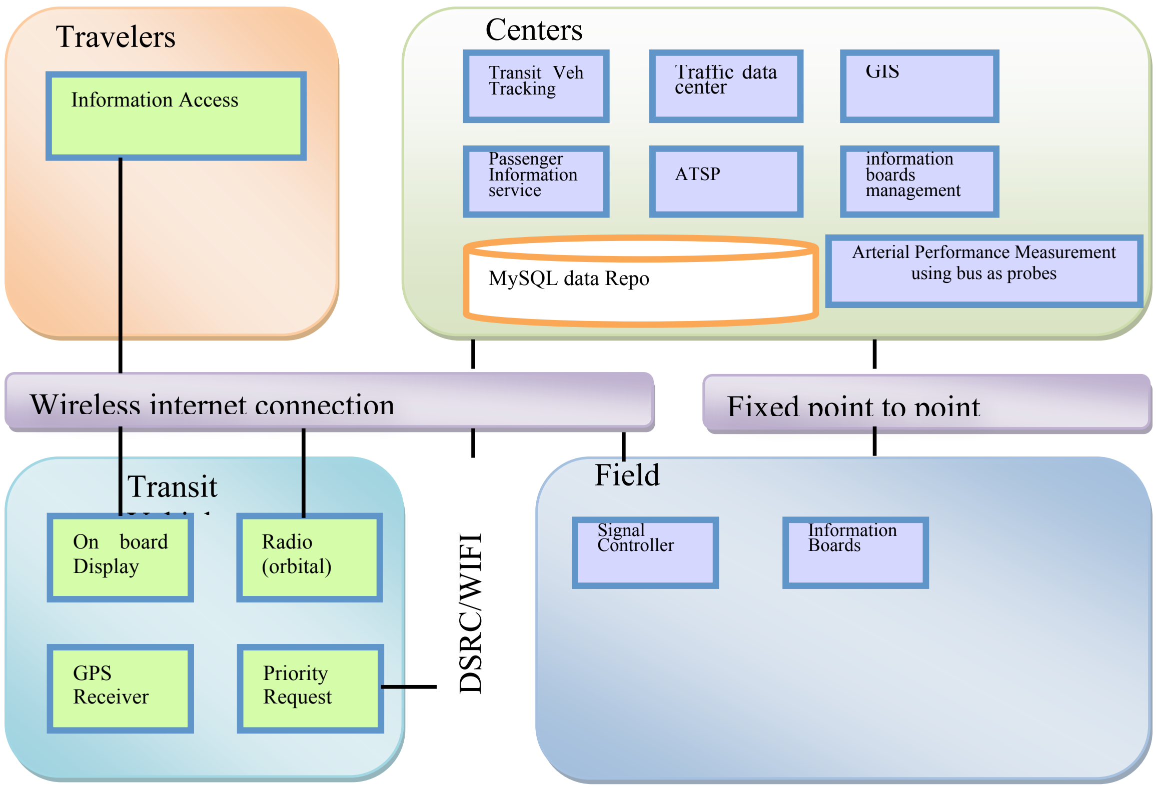


Figure 5-6: System Components of the transit system

The real-time arterial performance measurement system (APeMS) is a standalone subsystem within the NT transit system. Figure 5-7 presents the functional architecture of APeMS. The central MySQL database stores the second-by-second BRT bus data together with signal status data. Once a probe bus is detected stopping at a bus stop, its travel time from the previous bus stop will go through three filter programs to squeeze the bus stop effects, the cruise speed difference, and the signal waiting time. The residual time is bus queuing delay that is assumed to be the same as traffic queuing delay. The average arterial travel time is the queuing delay plus the free flow travel time and the average signal waiting time for other traffic.

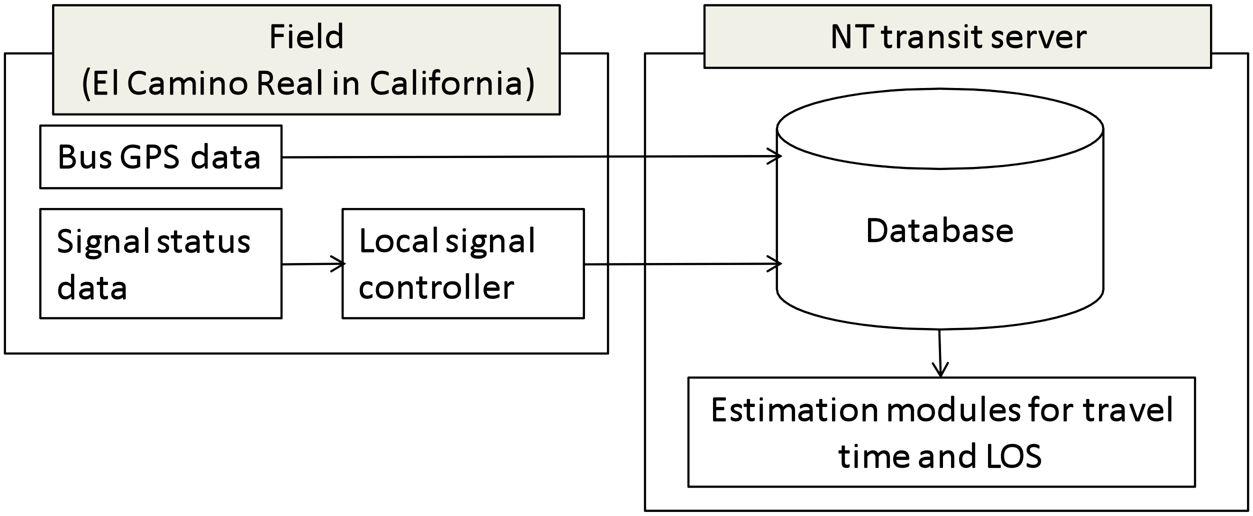


Figure 5-7: Architecture of real-time arterial performance measurement system (APeMS)

5.3 Services

The purpose of NT Transit services is to provide travelers more reliable information in real-time about their transit trips. The timely and accurate information improves the transit riders' perception of the transit service and thereby attracting "choice riders".

At this stage, the project focuses on U.S. 101 corridor, including the following major transit routes:

1) Santa Clara Valley Transportation Authority (VTA) Bus Rapid Transit (BRT) route 522, between Eastridge Transit Center at San Jose and Palo Alto Transit Center,
2) Caltrain rail services between Gilroy/San Jose and San Francisco. **Error! Reference source not found.** geographically illustrated the above referenced transit routes along HW-101 corridor.

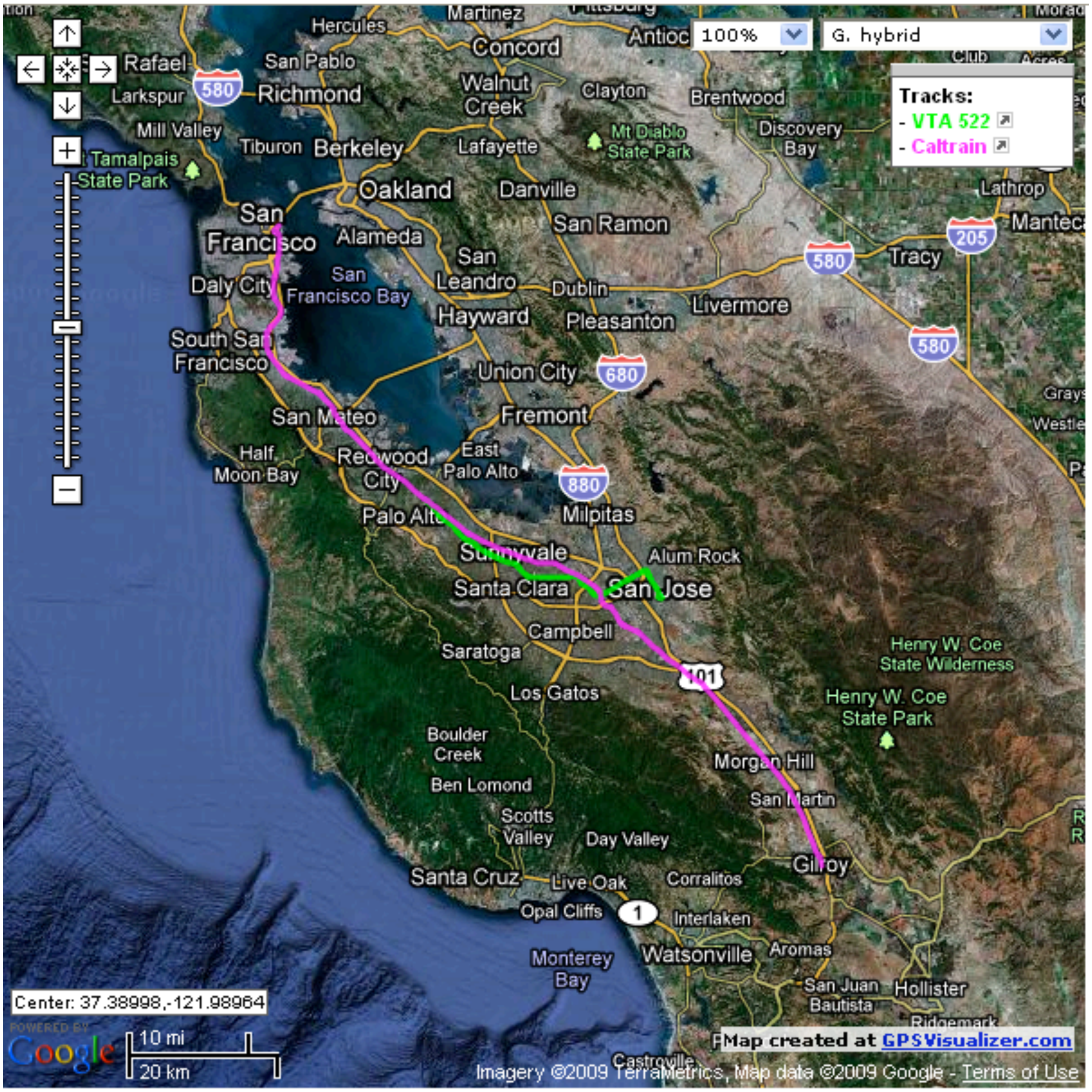


Figure 5-8: The NT Transit Bus /Train Routes Along US101 Corridor

5.3.1 Reliable Real-Time Transit Information Based on AVL System

5.3.1.1 Review of Real-Time Transit Information System Based on AVL

To obtain the real-time GPS locations of the bus fleet and the locomotives for the selected routes, we have the GPS devices with wireless modem installed on the 19 buses of the BRT 522 dedicated fleet and the 29 locomotives of Caltrain.

The AVL devices are sending out data to the NT transit server second by second. Each device has a device ID that is associated with a vehicle ID (thus a route). The NT transit system processes the real-time AVL data and generates the predictive arrival time at downstream bus stops.

There have been a few passenger information systems deployed (or to be deployed) in the San Francisco Bay area, including the SamTrans and VTA. Both SamTrans and VTA have deployed advanced communication system (ACS) on their bus fleets. The ACS is utilized to track the location of each bus. Each bus is equipped with a GPS receiver that allows the on-board Advanced Mobile Data Terminal (AMDT) to determine its current real-time location and schedule adherence. This information is transmitted via Ultra High Frequency (UHF) radio through repeater sites to ACS servers at Operations Control Center (OCC). The ACS server polls each bus approximately every 1 minute to 2 minutes for its location status. The AMDT also provides two-way text messaging capability between the bus and OCC. The AMDT has a small liquid crystal screen used to display simple text messages (such as bus detours or service interruptions) in addition to time and schedule information for the bus driver. In year 2009, SamTrans started its dynamic passenger information (DPI) project, which via using the ACS system provides real-time bus arrivals at the bus stops.
The NT transit system tries to explore the possibility of better transit service beyond what the current ACS/DPI system can provide (1) better data quality with the ubiquitous wireless wideband connectivity which enables much higher frequency in the AVL data update; and more important based on the more frequent AVL data and a more reliable predictive bus arrival time which is generated with full understanding of the transit routes, the real-time bus location, the arterial traffic and the signalized intersection delays; and (2) personalized passenger information that is pushed to the smart phones to the travelers based on the itinerary. The NT transit server generates the information of "your bus", "your train" and "your stop" in addition to the traditional transit information that is only for a certain station.

5.3.1.2 Understanding the Transit Operations

There are two types of transit stops (bus stops or train stations): stops and time-points. The difference between them is that, at time-points, a transit vehicle can arrive before - but not leave earlier than (so called time-point holding) - the stated time as indicated in the route schedule.

Table below lists the number of stops, number of time-points and route length for the 4 above referenced transit routes. Most of Caltrain's rail services are operating between San Francisco and San Jose, therefore the information for Caltrain is provided both from- and to- Gilroy and San Jose. As a BRT line, VTA Rapid 522 has much less stops, compared with normal transit routes.

Transit Route	Length (miles)	Direction	Origin	Destination	Total No. of Transit Stops	No. of Time-Points
VTA 522	25.8	WB	Eastridge T.C.	Palo Alto T.C.	30	13
		EB	Palo Alto T.C.	Eastridge T.C.	30	13
Caltrain	77.2 (47.5)	NB	Gilroy (San Jose)	San Francisco	31 (25)	31 (25)
		SB	San Francisco	Gilroy (San Jose)	31 (25)	31 (25)

Transit agencies operate multiple services along a one-way transit route, with different origin-destination (O-D) pairs. For example, the origin point for weekday Caltrain northbound trips can start at Gilroy Station, Tamien Station and San Jose Diridon Station.
SamTrans routes 390 and 391 and Caltrain provide schedule-based transit services, where point-point holding discipline is applied. Although VTA has published schedule for Rapid 522, Rapid

522 buses will travel as fast as traffic and signals allow, meaning that buses will not sit idle at time-points when ahead of route schedule. Therefore, Rapid 522 is more like a headway-based service.

Buses share the roadways with general traffic. In the design of route schedule, the expected route travel time (for example, the 85-percentile traffic travel time) is combined with the slack time, lead to schedule stability. If the slack time is insufficient, transit vehicles are unlikely to catch the schedule when falling behind, thereby downgrading the service reliability. On the other hand, large slack time reduces service frequency and increases transit waiting time and travel time.

Figure 5-9 and Figure 5-10 clearly show that route schedules match with the traffic patterns. The scheduled travel time is smaller in the early morning and evening, when traffic is lighter, and is larger during the rush hours.

Caltrans' rail service, although also following the time-point discipline, is different than bus service as trains do not share roadway with traffic, thus schedule having little correlation with traffic patterns. **Error! Reference source not found.** below shows the scheduled travel times between San Jose Diridon Station and San Francisco Station for northbound and southbound baby bullet trains, which are consistently at about 60 minutes.

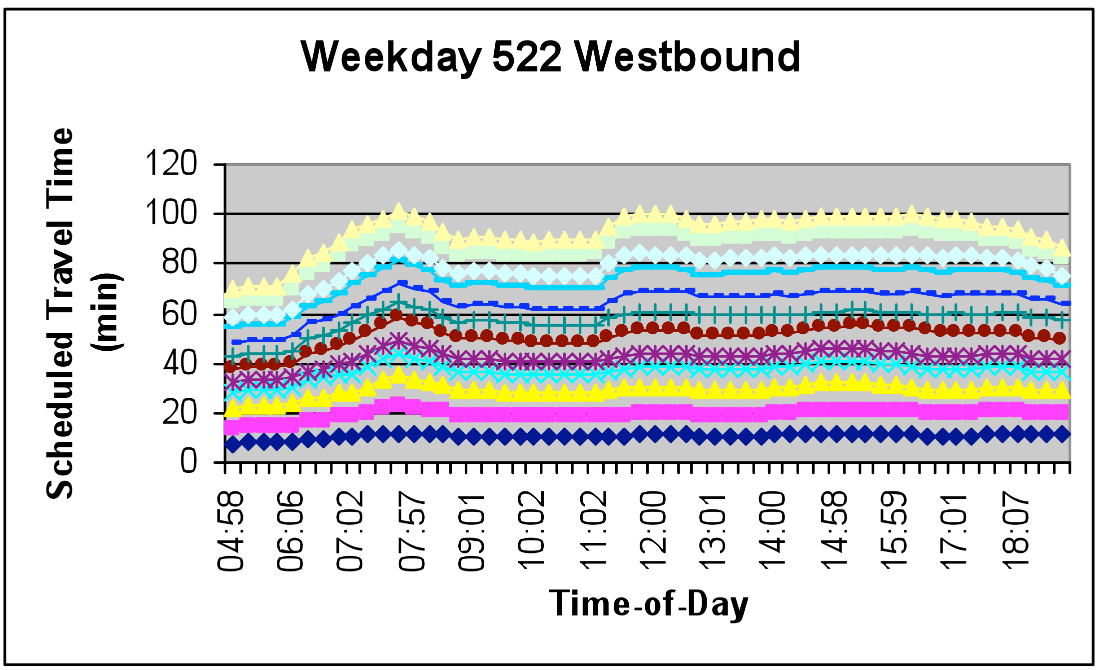


Figure 5-9: Scheduled Travel Time to Time-Points (VTA Rapid 522 WB Weekday Trips)

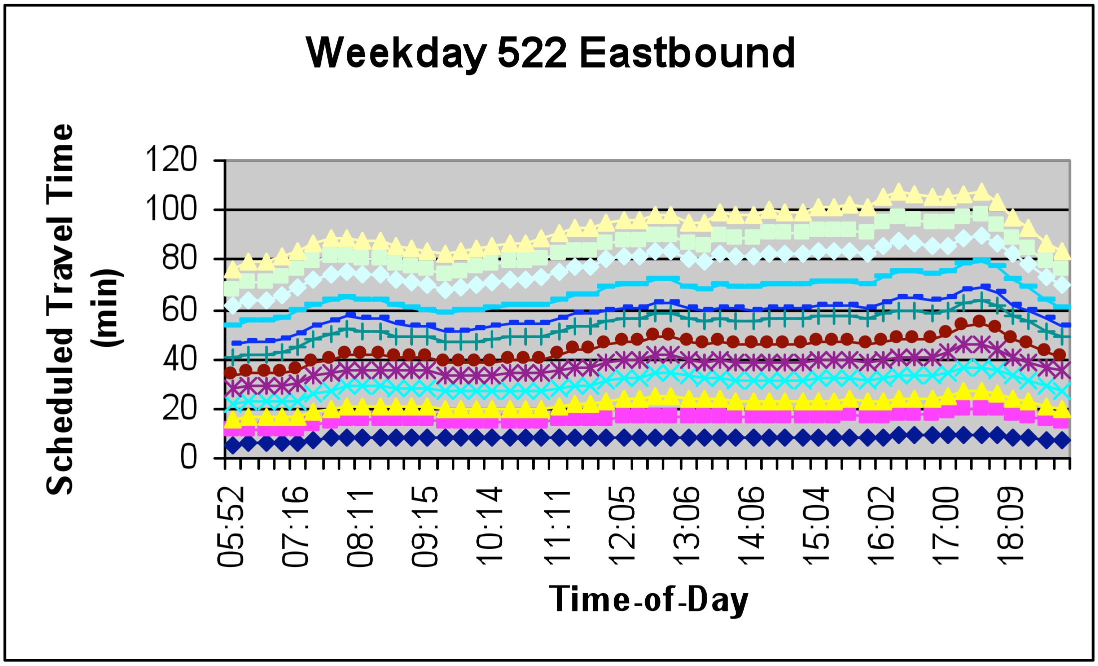


Figure 5-10: Scheduled Travel Time to Time-Points (VTA Rapid 522 EB Weekday Trips)

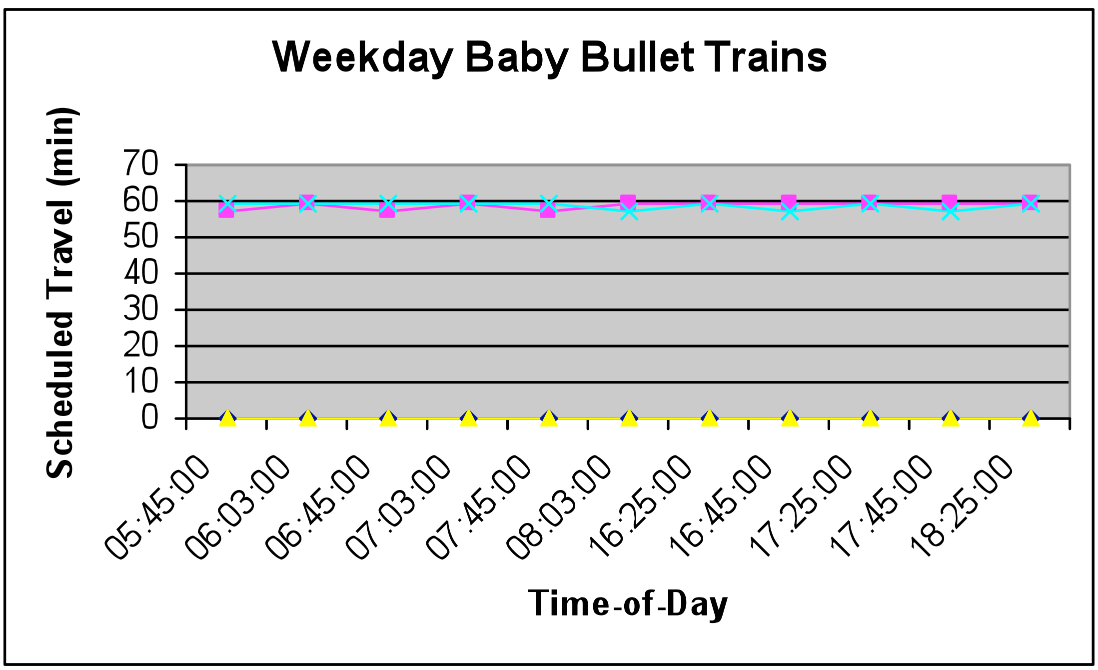


Figure 5-11: Scheduled Travel Time between San Jose and San Francisco

(Caltrain Weekday Baby Bullet Train Trips: -x- southbound, -■- northbound)

When utilizing AVL for DPI features, the impacts of AVL polling rate (between 1 to 2 minutes) need assessment. Actually when the AVL polling rate is low (such as 1-2 minutes polling as in the current ACS system) there is a crucial problem in providing travelers accurate transit information. Before receiving the next bus location update, the system may indicate the bus is approaching or waiting at a stop, while in fact the bus has already left the stop.

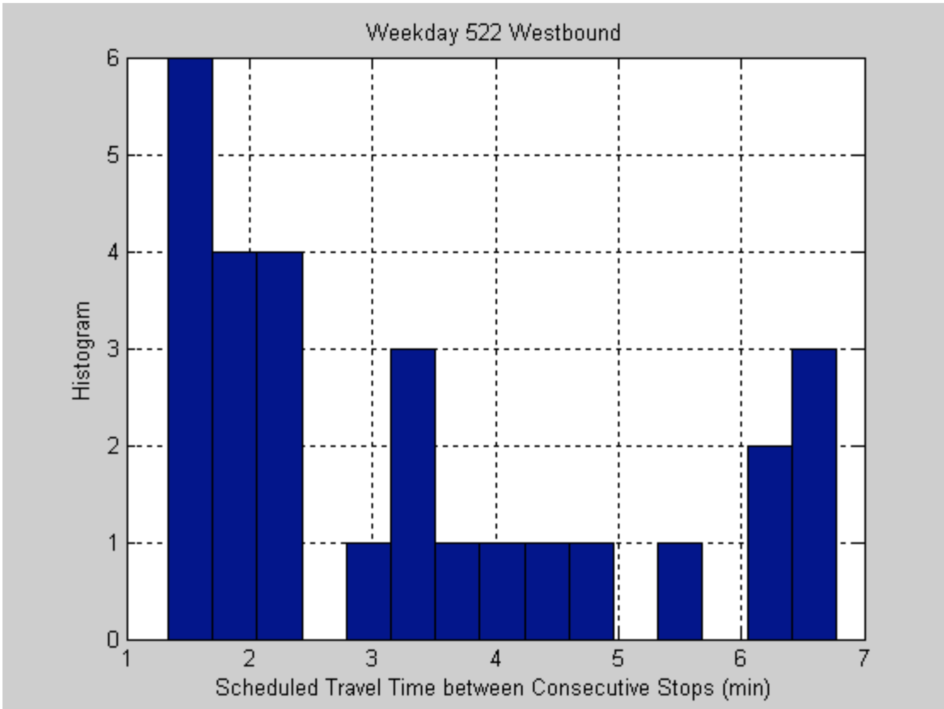


Figure 5-12: Scheduled Travel Time between Consecutive Stops

(VTA Rapid 522 WB Weekday Trips)

In summary, we have the following conclusions:

- Delivering dynamic passenger information to the riders via cell phone imposes critical requirement for the accuracy of the predictive arrival time information, which in turns requires the AVL sampling rate to be high enough that the current approaching stop of the bus / train could be timely captured;
- Understanding the schedule pattern and the time-point adherence performance is critical for better bus arrival time prediction;
- Traffic conditions have impact on the travel time of the bus travel time. Therefore the integration with the traffic prediction tool would help to further improve the prediction accuracy.

5.3.2 En route information: Waiting for Transit to Arrive

After a transit trip has been planned, the next step is generally for the rider to go to the bus or train stop and wait for the next bus or train to arrive. At this point in the trip, there are two very important things that a rider needs to know. First, the rider needs to know where the bus stop is, and second, the rider needs to know when the next bus was arriving. Furthermore, the immediate informational needs of the rider may vary and depend on whether or not the rider is currently walking to the bus stop or standing at the bus stop. While walking to the bus stop, the information is focused on the boarding stop name. If waiting at the bus stop /transfer point, the information includes the next two arrivals and the alighting stop as well.

Your Route Information		
Boarding at El Camino Real at S Fairoaks		
Route Dir	*Next Bus*	*ETA*
522 to Palo Alto Transit Center	11:58 AM	*** in 12 mins**
522 to Palo Alto Transit Center	12:13 PM	**in 27 mins**
**Real-time transit info*		

Figure 27: En-Route – Next Transit Vehicle Information

Bus Arrival	
El Camino Real at S Fairoaks	
522 to Palo Alto Transit Center	*** in 10 mins**
522 to Palo Alto Transit Center	**in 26 mins**
Your Stop Location	
El Camino at California	
**Real-time transit info*	

Figure 5-13: En-Route – Next Transit Vehicle Information While Waiting at the Bus Stop

When the bus /train is approaching the boarding stop, the user was alerted with a popup screen and an audio message. When the rider is detected to be waiting at the bus/train stop, the alert is issued when bus /train is coming within 1 to 2 minutes. If the user is walking to the bus /train stop, the alert time is calculated as the time when the rider could catch the bus if keep walking at average walking speed (2.72mph).

5.3.3 En route: Riding Transit

When a transit rider is riding a bus or train, the basic questions that the riders need answered center around, “Where am I?” and “When should I get off the bus?” The Networked Traveler transit application provides these information to the rider in real-time. First information is the ETA to your stop. The goal of the en route ETA to your transit stop screen is to provide the rider with information about where the bus is in relation to where the rider needs to get off the bus. Additionally, there is a lot of additional information that could be provided on the screen.

Figure 5-14: En Route – My Stop Information

When the bus /train is approaching the alighting stop of the rider, an alert was given to the user as a popup screen with an audio alert "Your stop next". This is based on the same assumption as the one for "bus/train coming alert" that the smart phone user might be using some other applications (such as email, music, etc) at the time and not paying attention to the transit app. In this case, the audio alert could be helpful to the rider to avoid missing the stop.

One important concern for a smart phone application is its battery consumption. The GPS based applications are especially power demanding due to the extra power consumption in the GPS chip. NT transit system is designed with a "selective GPS source (SGS)" technology, which does a probabilistic matching of the user GPS with the GPS data from all the relevant transit fleet. Server keeps receive GPS data from the smart phone until only one single bus /train is successfully matched with the user. Multiple buses / locomotives might be close by the user at certain time, especially at the bus /train station. So the SGS module on the transit server is designed to keep track of the hypothetical matches in parallel and makes the decision until only one pair persists.
By applying the SGS technology, the smart phone could smartly turn off the GPS when server knows that the user is currently on a bus / train soon after the user gets onboard. Thereafter the smart phone would receive real time update without need to send out the GPS to the server. As soon as the bus arrives at the transfer point of the rider, the server notifies the smart phone to turn on its GPS again.

5.3.4 Arterial Travel Time Estimation using Bus-as-probe

As real-time communication technologies develop, it becomes possible to collect real-time traffic data from urban arterials as well as freeways, and this data collection can be used in estimating travel times in urban arterials and freeways. This travel time information helps many travelers make their decision on trip start time, mode/route choice, etc. In addition, transportation operators can easily evaluate developed traffic control methods and transportation policies by using the estimated travel time. The aim of this research is to develop a robust methodology to measure arterial performance and estimate arterial travel time using bus probe data.

The Bureau of Public Roads *(5)* developed equation model to calculate in-link travel time according to the variation of demand/capacity ratio. In comparison, the model in Highway Capacity Manual *(6)* specifically considers the impacts by traffic signal control. However, these methods are static and based on historical data. Therefore, they are not appropriated for the real-time travel time estimation.

Turner et al. *(7)*, Frechette and Khan *(8)* and Zhang *(9)* studied statistical models to estimate the arterial travel time. They estimated relationships between link travel time and flow characteristics such as loop detector data, free flow speed, saturation flow and vehicles' spacing. These studies shows relatively good estimation compared with the previous studies. However, these studies are much site-specific, and their estimation methods are difficult to apply to other links.

Concerning these problems, many researchers focused on real-time travel time estimation using real-time loop detector data and signal data in microscopic scale. Skabardonis and Geroliminis *(10)* developed an analytical model based on kinematic wave theory. They used loop detector data and signal timing data as input data. Liu and MA *(11)* studied a virtual probe vehicle model by using loop detector and signal timing data. They introduced a virtual vehicle into arterial and calculated the vehicle's trajectory based on Newell's car following theory. Then, they estimated the arterial travel time using this trajectory. These methods using real-time loop detector and signal data have advantages in processing the real-time traffic data and showing the real-time travel time. However, these approaches rely on point-detection at location where loop detectors are installed and assume that loop detector data such as count, occupancy and speed represent flows characteristics at other locations on the link. Under this assumption, variations of flow characteristics along a road link cannot be captured.

To deal with this problem, probe vehicles in arterials are used to estimate the travel time. Hall and Vyas *(12)* compared bus probe data of the Orange County Transportation Authority with automobile trajectories and found that buses are likely to be delayed when automobiles have long delays. While the reverse situation is not always true. Bertini and Tantiyanugulchai *(13)*, Uno et al. *(14)* and Chakroborty and Kikuchi *(15)* developed travel time estimation model by estimating the relationship between automobiles and buses travel time after eliminating bus stop dwelling time. Recently, a fusion model using bus probe data and loop detector data was developed to support travel time estimation when there are no probe runs *(16)*. However, there is well-known limitation of these previous studies. Because there is only one bus probe at the scheduled time in case of using transit probes, it is hard to say that this bus probe represents all vehicles in arterials at same time period even after using methods developed in the previous studies. For example, the bus probe might get good signal coordination in the arterial, but the other vehicles do not. In this case, the estimated travel time is much faster than the average travel time of other vehicles in the arterial.

5.3.4.1 Data Characteristics and Methodology

Second-by-second global positioning system (GPS) data is available for the whole fleet of Rapid 522. As probe vehicles for the purpose of measuring arterial performance, BRT buses have prominent advantages over local buses because BRT service is meant to be the transit service which is as efficient as driving personal cars. First of all, the BRT bus runs more like other traffic than the local bus does. Figure 5-15 illustrates three typical trajectories for a BRT bus, a local bus, and a testing vehicle, respectively. All the three vehicles started within the same time window to cross the same segment of El Camino Real. As shown in the figure, the cruising speeds, which are the slopes of curves, are somewhat different among the three vehicles. The local bus has much lower cruising speed than the test vehicle and the BRT bus. In contrast, the cruising speed for the BRT bus is quite similar with that for the test vehicle. There are three main reasons. The first one is the advanced BRT vehicle allows them to accelerate rapidly and cruise with higher speed. The other reason is that BRT typically runs headway-based service, so it doesn't need to adapt their cruising speed and dwelling time to meet the schedule at pre-defined check-points.

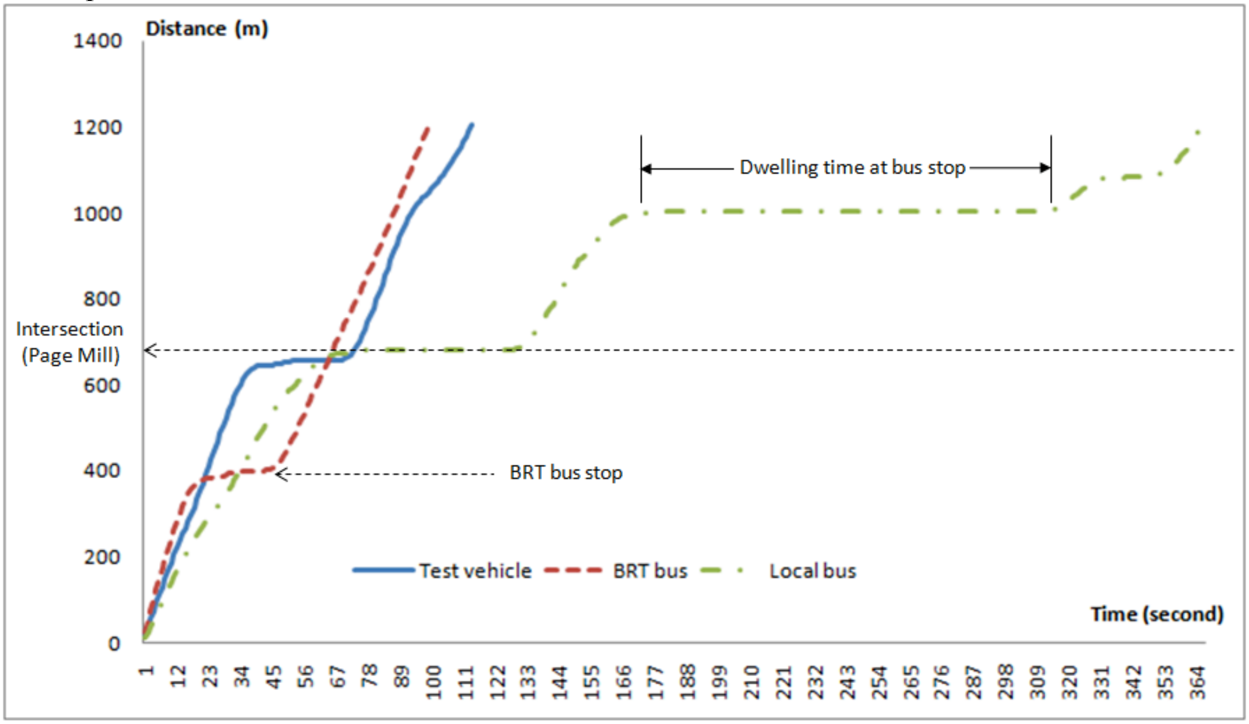


Figure 5-15: Trajectories of test vehicle, BRT and local buses

Second, BRT buses can flow more smoothly with other traffic because they don't have to stop as frequently as local buses do. For example, the bus stops for Rapid 522 are spaced approximately one-mile apart compared to stops spaced less than a quarter mile apart for local Line 22 serving the same route. Third, when BRT buses have to stop at bus stops and leave the major traffic platoon, transit signal priority (TSP) can help those buses to catch up with the traffic platoon. As shown in Figure 5-15, the local bus got the longest delay, while the test vehicle experienced shorter delay. The difference is partially due to the random arrival at the intersection. However, the reason why the BRT bus experienced almost zero delay is that this particular bus received prioritized treatment at the intersection. It is noted that BRT buses would need less TSP when flowing with major traffic platoons under coordination.

Although BRT buses have some advantages to be traffic probes, the travel time for BRT buses cannot be directly referred as the arterial travel time for general traffic. There are three main factors that cause the travel time difference between BRT buses and general traffic. They are: (1) bus stops effects; (2) cruise speed difference with general traffic; and (3) traffic signal effects and signal coordination. The concept of our methodology is to filter the bus trajectory by eliminating the bus stop effects and difference of cruise speed and replace the intersection delay with average delay for general traffic.

The major difference between bus travel time and other traffic travel time is the delay caused by the dwelling time at bus stops. The delay consists of three elements: stop time, deceleration delay and acceleration delay. To calculate these delays, it is the prerequisite to detect whether a bus halts at a bus stop. Buses sometimes skip bus stops if there is no passenger to board/alight at bus stops. We check the bus's speed near the bus stop.

After detecting bus's halting at a bus stop, the process to calculate deceleration and acceleration delay follows. To calculate these delays, we need to find out time points when a bus starts decelerating before halting at a bus stop and when it finishes accelerating after the departure from a bus stop. The GPS velocity from a running bus fluctuates due to 1) the dynamic of traffic situations on urban streets and 2) GPS noise. Kinematic model was built to detect the vehicle approaching bus stop behavior from the GPS trajectories.

The delays caused by bus stops are eliminated from the bus travel time by subtracting stop time, deceleration and acceleration delays from bus travel time.

Cruise speed

Cruise speed difference between bus and general traffic also contributes to their travel time difference. In almost all previous studies, researchers did not use BRT bus data but local bus data. Thus, they had to calculate the relationship between buses' and general traffic's cruise speed. In this study, however, we used BRT bus data. Because the density of BRT bus stops is low, BRT bus drivers tend to use inner lanes rather than a shoulder lane. This characteristic makes it possible for BRT buses to run with higher cruise speed in comparison with the local buses that mostly use the shoulder lane. When a BRT bus is not freely flowing due to congestions, other traffic typically slows down to run with a similar speed as the BRT bus does. Therefore, we can simply assume that if the BRT bus is not freely flowing, the velocity for other traffic is also the same as that for the BRT bus. In free flow condition, however, each mode might have different cruise speed.

We processed second-by-second velocity data from a BRT bus and compared with the data from a test vehicle. Both of the two vehicles were running along a 3-mile segment of El Camino Real during the same period and under free flow traffic condition. The cumulative distribution function (CDF) of the free flow speed is shown in Figure 5-16. The velocities while accelerating or decelerating closing to traffic signals and bus stops were excluded. As illustrated by the figure, the free flow speeds of both modes are very similar with each other. The average cruise speed for the test vehicle and the BRT bus are 17.5*m/s* and 17.6*m/s*, while their standard deviations are 1.37 *m/s* and 1.24 *m/s* respectively. Furthermore, the result of statistical test shows that there is no significant difference between both modes' free flow speed at 95% significant level. This relationship has also been verified by using other sample probe data. Thus, we do not have to estimate the cruise speed relationship between two modes when estimating the arterial travel time.

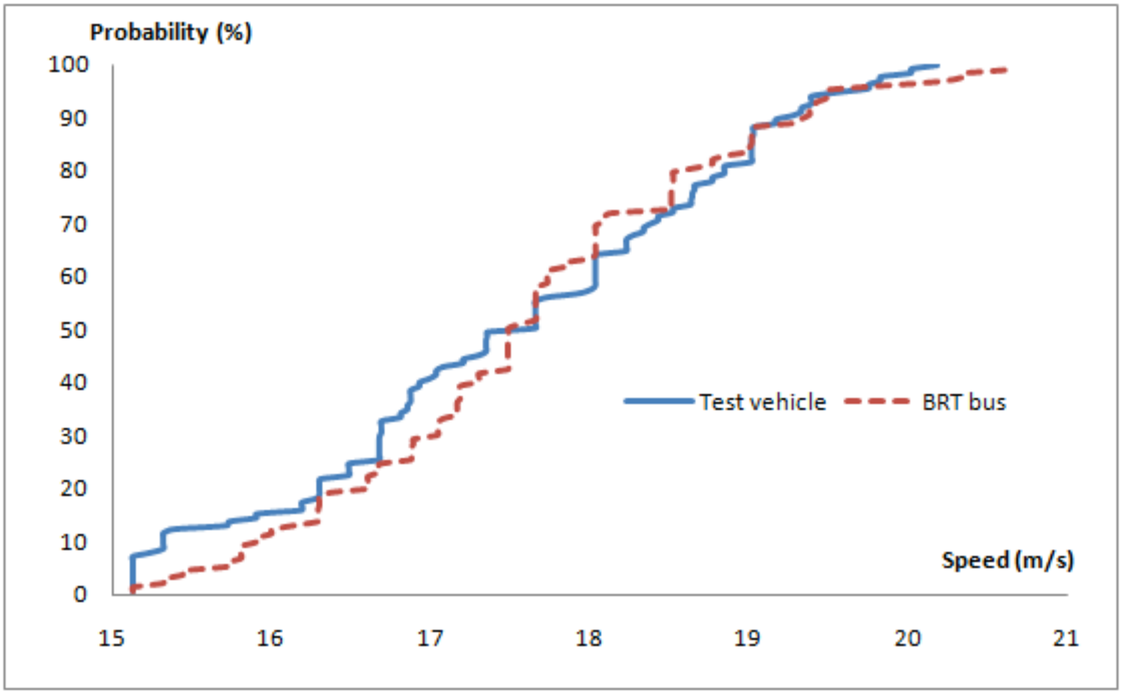


Figure 5-16: Cumulative distribution function (CDF) of cruise speeds for a BRT bus and a test vehicle

5.3.4.2 __Field Calibration of the Travel Time Estimation__

The bus-as-a-probe real-time arterial traffic data is fed into the bus ETA prediction model to make the bus arrival time prediction results more adaptive to the various traffic situations. Moreover, the trip planning tool in DPI system is using the estimated arterial performance information to provide dynamic planning results.

5.4 Outreach

Client Software Distribution

The client software for personalized transit information was distributed, together with the Connected Traveler software client, via a website, with the objective of targeting up to 1000 travelers.

Data Collection and Release

The transit probe data was collected at the Parsons Traffic and Transit Lab. The data will include bus id and AVL location each second. Processed data was archived in conjunction with the probe data and was made available for other Connected Traveler applications.

We will collect the usage data at the data server to record the way that personalized transit information is used. User feedback will also provide the basis for evaluating the effectiveness of the personalized transit information and was used for further refinement of system design and to lay the foundation for a larger scale deployment. The types of data that can potentially be collected include the following:

- Origins and Destinations inquiries
- Estimated vs. actual bus arrival at the stops of interest
- Estimated travel time vs. actual travel time
- Frequency of acquiring next stop or connection information and their relationship with transit on-time performance
- Traveler searching patterns
- How often a specific user accesses the personalized information
- User feedback, from web-based or paper-based surveys

5.5 Transit Benefit Feasibility Assessment

(3) If travelers are better informed of the travel options in real-time, including driving, transit and mixed mode, they was more likely taking transit.

Transit has become increasingly viable for travelers as a result of gas price hike. APTA reported a 4.36 % increase in ridership nationwide in 2008 compared with a year ago due to the gas price hike. For similar reasons, the ridership increase for rail is about 12%. The ridership data provided by transit agencies in the Bay Area is consistent with national statistics and, promisingly, ridership continues to grow despite the fact that gas prices have become lower. While the gas price is the key factor in causing a mode shift for many, the fact that riders stay with the transit mode indicates that most travelers may not know their transit mode option as an alternative before being 'triggered' (in this case, by the pocketbook) for mode transfer. Once triggered, these former drivers have stayed with the transit. It appears that the mode shift 'experiment' due to gas price hikes is due to knowledge and information about the transit alternatives. It is therefore hypothesized that this is such knowledge will help attract riders and that transit can be a viable, realistic mode for commute corridors where travel time for transit is competitive.

(4) If transit buses can be used as probes, travel time information for freeways and arterials can be achieved.

Transit buses frequently travel on freeways and arterials, offering potential as probe vehicles for travel time estimation and incident detection. Once proven, buses as probes can be feed into the Bay Area's 511 traveler information system to provide both the real-time freeway traffic condition information and additional predictive travel time information for local roadways and expressways. It is recognized that buses may not operate consistently with traffic. Buses travel on HOV lanes on freeways, where available, and stop at bus stops. 'Filtering' algorithms need to be developed to remove the travel pattern specific or to associate bus travel pattern with general traffic therefore to achieve a good estimation of traffic condition and travel time. Under this task, these filtering algorithms was developed and field tested.

5.6 Execute Transit Field Test

5.6.1 NT Transit Test Hypothesis

5.6.1.1 Test Hypothesis

For each of the applications that are to be tested in the field, the hypothesized outcomes, expected benefits and anticipated user responses are summarized in Table below.

Transit Application and Test Hypotheses

Application	Applicable Situations	Hypothesized Outcome	Expected Benefits to and Responses from Travelers
En-route traveler information	• Traveler catches a train/bus at a station	• Travelers benefit from the real-time bus arrival information such that he/she can catch the next bus/train with certainty	• Travelers provide favorable assessment of the next bus/train information • Travelers reduce wait time at stations • Travelers avoid unnecessary rushing • Traveler stress is reduced
En-route traveler information	• Traveler needs to alight at his/her destination station	• Travelers was informed the destination station , so that he/she can alight the bus/train stop	• Travelers provide favorable assessment of the destination bus/train information • Travelers avoid anxiety caused by unfamiliarity of the bus stops • Travelers avoid missing the destination stop • Travelers avoid rushing at the destination stop • Travelers benefit by avoiding unecessary wait time at stations
En-route traveler information	• Traveler needs to transfer to a different bus/train	• Travelers was informed the destination stop so that he/she can be prepared to alight • Travelers was informed the arrival time of the transferring bus/train so that he/she can catch the transferring bus/train with certainty	• Same as above ("traveler catches a train", "traveler needs to alight")
En-route traveler information	• Traveler needs to obtain the estimated travel time	• Travelers was informed the travel time so that he/she can know it with certainty	• Travelers provide favorable assessment of the destination bus/train information • Travelers can inform

			relevant parties of their arrival time • Travelers avoid missing their appointments
Bus as probes	• Transit bus provide travel time and congestion level on freeway	• Buses probe data are used to provide transit travel time, which in turn was provided to travelers through the en-route information described above • Transportation agencies such as Caltrans or MTC fuse the buses as probes data with other data to improve richness and accuracy of travel time	• Travelers provide favorable assessment of the destination bus/train information • Travelers will trust the travel time prediction • Also see above
Bus as probes	• Transit bus provide travel time and congestion level on arterials	• Buses probe data are used to provide transit travel time, which in turn was provided to travelers through the en-route information described above • Transportation agencies such as Caltrans or MTC will have arterial traffic data that they currently don't have	• See above

The common elements given in above table are reorganized into the chart below to capture the suite of applications .

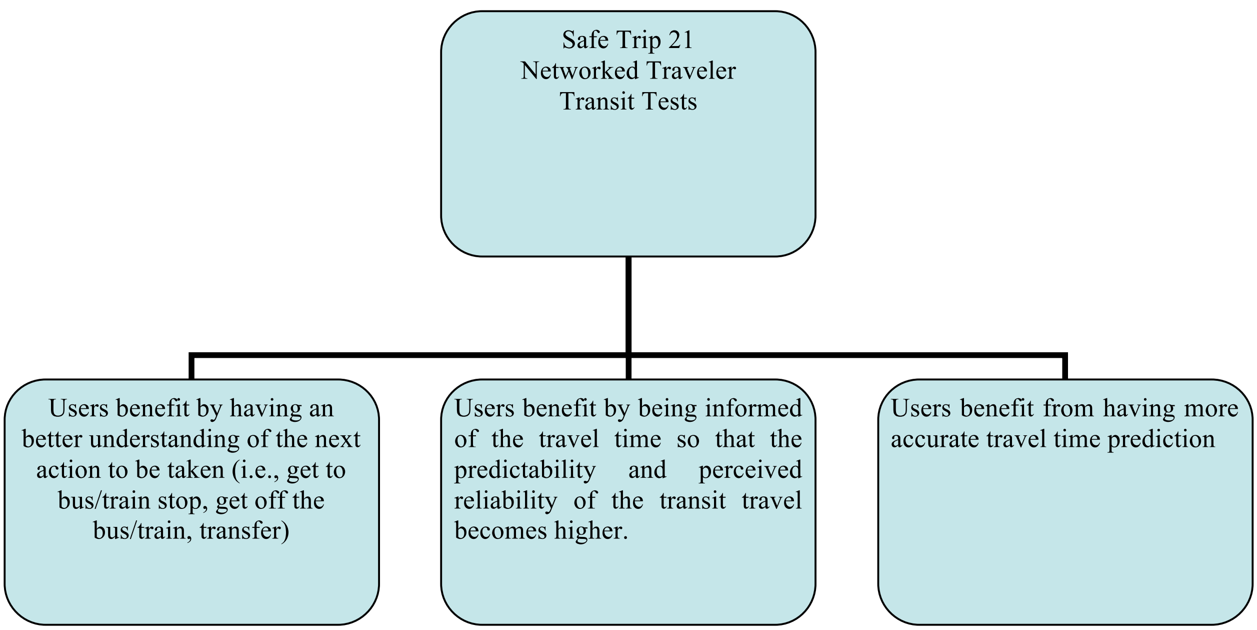


Figure 5-17: Hypothesized Expected Outcomes of Safe-Trip 21 Transit Applications

5.6.1.2 Validation of the Test Hypothesis

Similar to the safety experimental design, it is expected that the user experience for transit applications was evaluated and field data was collected to explore the user needs and preferences. The applications will also serve as design validation and to identify shortcomings of the transit applications. Therefore, the experimental design of the field tests emphasize the quantitative and qualitative observations of user response.

The observable outcome for buses as probes is rather straightforward, which was validated through other instrumented vehicles to show the accuracy of the travel time estimation. The expected cause-and-effect of transit information on travelers' behiavior is rather difficult to directly measure. However, the methodology was to assess user's attitude toward the information that was provided, e.g.,., what information is used more often and by the most people.
The Networked Traveler project is based on the concept of using technologies that are currently available to consumers, such as GPS-enabled smartphones. The communication links available on smartphones include 3G and Wi-Fi. The set up envisioned for the planned tests can be described as follows:

- Users register for the service.
- Users enter origin and destination for the intended trip.
- Users activate services by choice and by preference.
- Once activated, the user's route and personal preference information are communicated to the system server.
- The system server receives AVL data from SamTrans and VTA for buses running along El Camino Real and Caltrain.
- The system estimates travel time on freeways and along El Camino Real based on selected bus AVL data that are instrumented with GPS loggers (with more frequent position polling).
- The system server provides trip time estimation, next bus stop, connection information, etc. to users and monitor user positions by receiving users' current positions periodically.

The following constraints in both quality and quantify of data acquisition should be noted:

(1) The system server receives transit AVL data from SamTrans and VTA, which poll the buses every 60-120 seconds. The methods of data polling and acquisitions being may pose additional delays; therefore the update rate and the quality of transit data are not within the control of the application developers.
(2) User positions are only available through the GPS coordinates on user's devices, and therefore the accuracy and availability of user trajectory (speed and position) depends on the GPS units and environmental conditions, e.g., occlusions to GPS satellites, multipath interference, which may vary along bus routes.
(3) Transit information are only available for designated segments and specific locations in the test bed area; therefore, the measurements of background traffic are not always available to the user.
(4) Users may or may not activate transit functions even if they volunteer and register for the services.
(5) Users may choose to activate selective functions of their preferences and thus only a subset of functional results and associated data may be available from individual users.

(6) The travel information provided by the system is based on the OD inputted by the user. Hence, variations in the route actually taken by the traveler will cause perceived errors in travel information.

(7) The field test is further constrained by the fact that travel information are communicated to users through user interface implemented within the capabilities and flexibility allowed by phone-based devices.

The planned field tests of the transit applications is also on a limited scale, and within a limited period of performance and scope. However, it is still important to establish the framework and methodology to conduct the system assessment such that effectiveness and usefulness of transit applications can be measured. The following table illustrates how a matrix of measure of effectiveness can be constructed.

Anticipated Test Outcomes and Measure of Effectiveness

Expected Test Outcome and Driver Responses	**Measures of Effectiveness (MOE)**	**Parameters and Variables to Assess MOE**
Public awareness of transit applications	• Spectrum of project partnerships	• List of partners in project • Scope of participation by partners • List of participating organizations outside of project team
	• Scope of community participation	• Number of participating users • Number of data samples collected in field tests • Percentage of positive feedback by users
	• Outreach efforts	• Sessions of activity reports held in public forums and conferences • Technical papers presented • Reports of media events
Favorable user experience and positive user feedback	• Willingness to participate and to maintain continual use of applications	• Number of participating users • Periods of active usage • Continuity and frequency in activating applications • Percentage of positive feedback by users
	• User feedback to surveys and questionnaire on - Function usefulness - Function acceptability - Timeliness of alerts	• User answers in surveys and questionnaires (to be detailed and designed later)

	- User interface friendliness	
Strategic actions (routing changes or modal choices) by users	• Correctness and reliability of real-time en-route information	• Accuracy of the estimation of the information • Latency of the information
	• Usefulness of information	• Frequency of the usage of information per trip • Frequency of information usage over time
Ability of buses as probes to support Networked Traveler applications	Contributions to • Freeway travel time measurements • Arterial travel time measurements	• Estimation accuracy of travel time • Timeliness of the travel time estimation • Quality of estimation of incident reporting

5.6.2 System Evaluation of the NT Transit System

The purpose of the system evaluation is to test the NT transit system performance, and to measure the critical parameters of the system to understand whether the NT transit system works and how reliably it works and whether it can reliably support the NT transit services.

Therefore the system testing is carried out to measure the following list of MOEs:

- data accuracy (GPS, predicted arrival time, etc)
- communication delay / outage
- Other critical measures which indicate whether the system is working properly: (rate of the giving out alerts correctly, rate of the user getting real-time updates vs. schedule based updates).

5.6.2.1 __AVL Performance__

The AVL system is a core component of the system and the performance of which has a great impact on the overall performance of the NT transit services. Therefore the AVL performance need to be evaluated to make sure the quality of the AVL data can meet the requirements of the NT transit services.

The list of measurements for the AVL system is listed below:

- iDEN Service Availability
  - Percentage of Package Loss and Outage due to network connection issues
- iDEN network data communication Latency
  - End to End Latency
- AVL data Updating Rate

- o Consecutive GPS update rate at data server

- The statistics of eight cell phones and one data center over 10 days were averaged to form the following performance indexes as shown in table below

AVL System performance indexes

Performance	Average	Definition
Instantaneous throughput	619Bytes/s	Number of bytes received per second by the data center from one cell phone, measured every 10 seconds. *Note: these statistics do not include measurements taken when there is a communication outage.*
AVL data availability	99.6%	The number of bytes received by the data center divided by the number of original bytes sent by the signal controller to the cell phone, measured every hour
AVL dat Latency	2 s	The time a packet takes to travel from the source (only the GPS message has its original time stamp, so the source originates from the GPS satellites) to the data center. Due to a lack of high resolution timestamp, the latency is estimated to be roughly 2s in most observations.

The average throughput data was obtained for all the clients under test. The tests were carried out at Richmond Field Station.

Figure 5-18 shows the cumulative distributions of instantaneous throughput. It shows that, the instantaneous rates (regardless of the communication outage) of the AVL modem are highly probably greater than 335B/s most of the times. While the required throughput for second by second GPS data is less than 100B/s (general length of raw GPS sentence). This rate is accomplished with a probability of over 96%, while rates higher than 335B/s over 90% of the time can be sustained over the long term when outage and other losses are taken into account.

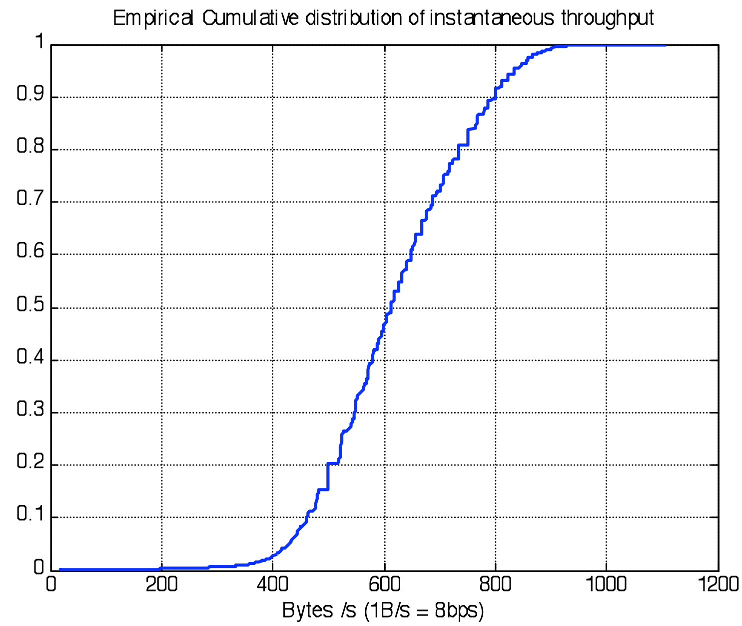


Figure 5-18: Cumulative distribution of the instantaneous throughput (Bytes/s)

The service availability is defined as the number of bytes received by the data center divided by the total number of bytes the original signal controller sent to the client (cell phone). It is always less than 1.0. From Figure 5-19, the probability of data lost due to flow control being greater than 2% is only 2%.

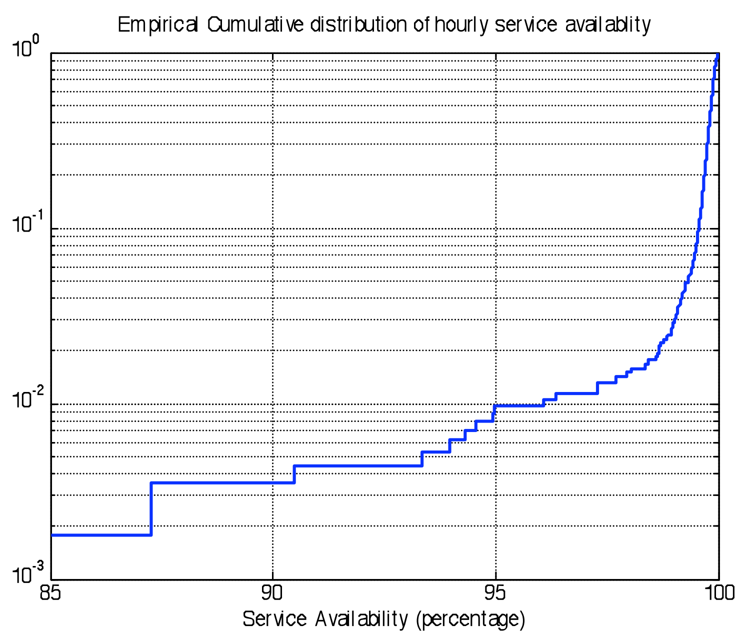


Figure 5-19: Hourly system service availability

The end-to-end latency is measured by the time difference of the GPS UTC time and the recorded time at data server.

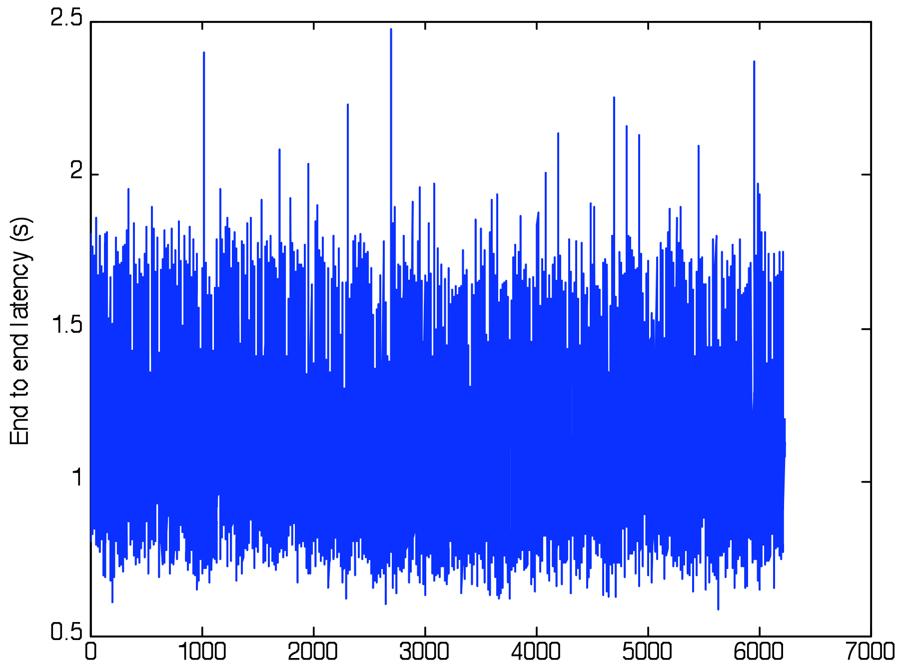


Figure 5-20: The End-to-End Latency of AVL Data

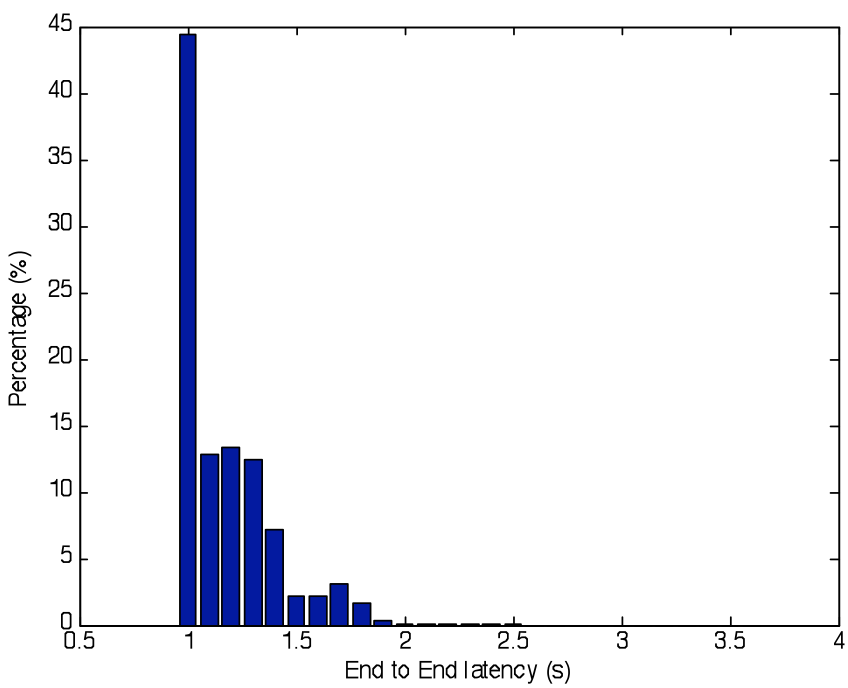


Figure 5-21: Histogram of the End-to-End AVL Data Latency

Statistics showed that (Figure 5-20 and Figure 5-21) is less than 2 seconds for over 98% of the packages. Considering the requirement for the NT transit service, the 2 second latency is within the acceptable range.

In addition to the end-to-end latency, we also need to measure the percentage of outages which is defined as the time period that the device losses network connectivity due to a wireless networking issue, or due to the GPS blockage by buildings, trees, etc. This is measured separately for the Caltrain and the VTA buses since they are running on different routes. The Caltrain locomotives go through several tunnels during the route so there are more GPS outages.

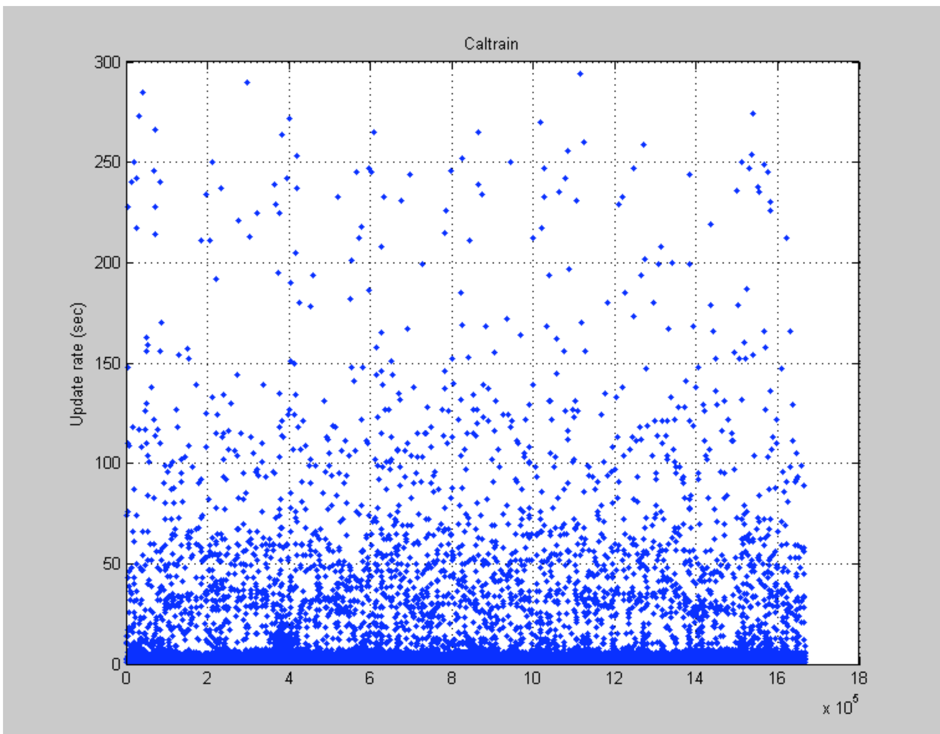


Figure 5-22: Caltrain GPS Outage Occurrences

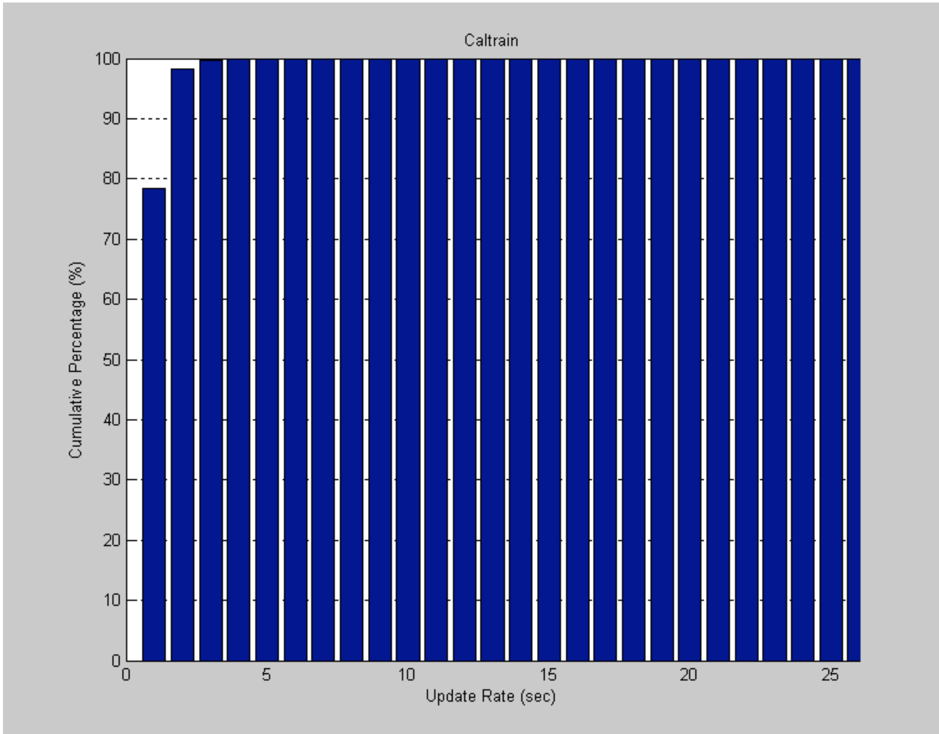


Figure 5-23: Statistics of the Caltrain GPS Outage

From Figure 5-22 and Figure 5-23 we know that with over 98 percent of the packages received on the server, they have a gap from the last sample of less than 3 seconds, or in another word, only less than 2% of the GPS packages received on the server has a gap of more than 3 seconds.

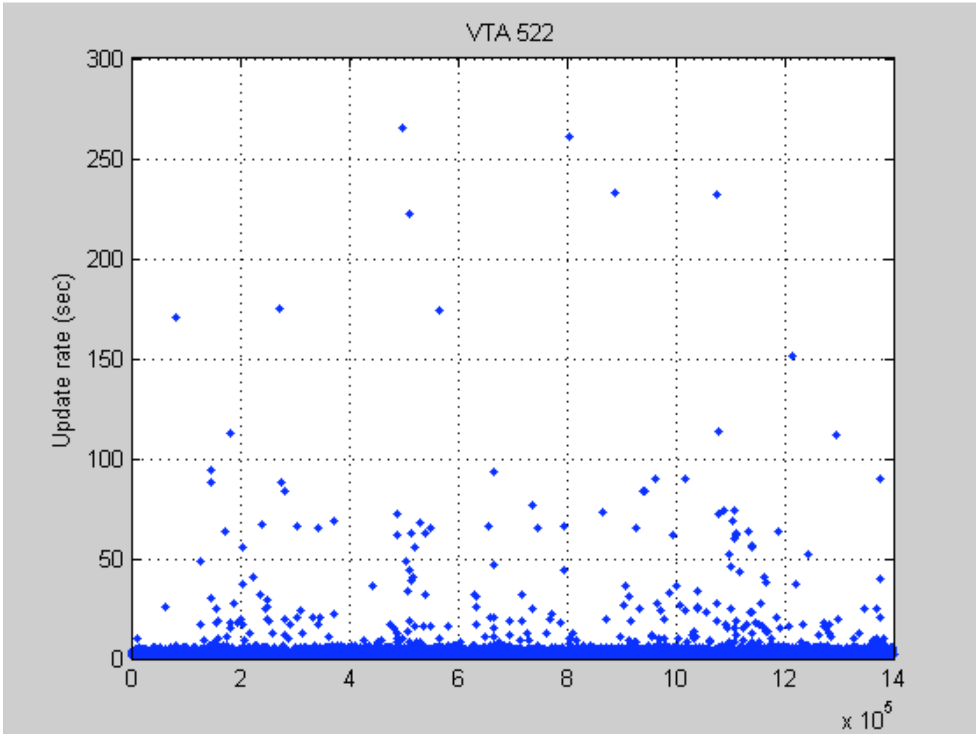


Figure 5-24: VTA 522 GPS Outage Occurrences

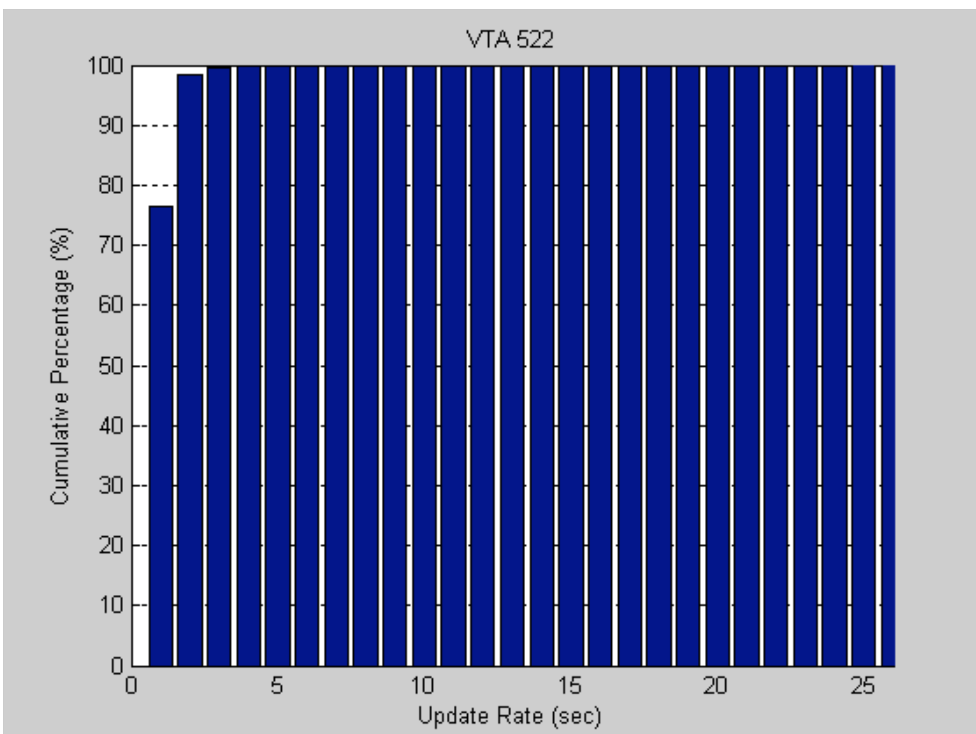


Figure 5-25: Statistics of the VTA 522 GPS Outage

For VTA the outage statistics is even better. Both the VTA and Caltrain AVL data outage probabilities meet the requirement for the NT transit services.

5.6.2.2 __Arrival Time Prediction Accuracy__

The accuracy of the arrival time prediction results need to be evaluated from two different perspectives, one is the objective evaluation which compares the prediction results to the actual bus /train arrival times (obtained from post processing the GPS data) and learn the objective accuracy of the predictions. Another one is the users' perspective, which is the statistics of the user feedback on the accuracy of the arrival time prediction results when presented as part of the en route transit information. For the second part, we will present the results based on the survey results from the users. Here we will show the objective evaluation of the results.
The MOEs of the arrival time prediction include

- Successful transit trip association rate (number of the trips that real-time predictive arrival time information is available)
    - Percentage of trips that train/bus is sending GPS data back to data server
    - Percentage of AVL data being latched with a transit trip
- Prediction error
    - Time difference of predicted arrival time and actual arrival time

For the successful association rate, the test was conducted at Oct 1st and Oct 2nd. Total number of trips made was 14 (7 train trips + 7 bus trips). And the total number of the trips that successfully showed real-time information was 12 (7 train trips + 5 bus trips) with a rate of 85.7%.
The missed trips were mainly because of the powering of the AVL devices. Bus drivers sometimes forget to turn on the headlight as they are supposed to do. The AVL devices are powered by the headlight circuitry.

The prediction accuracy is measured by the time difference of predicted arrival time and actual arrival time. The error of schedule deviation is also calculated. The data used was from September 28th 2009 to October 2nd 2009.

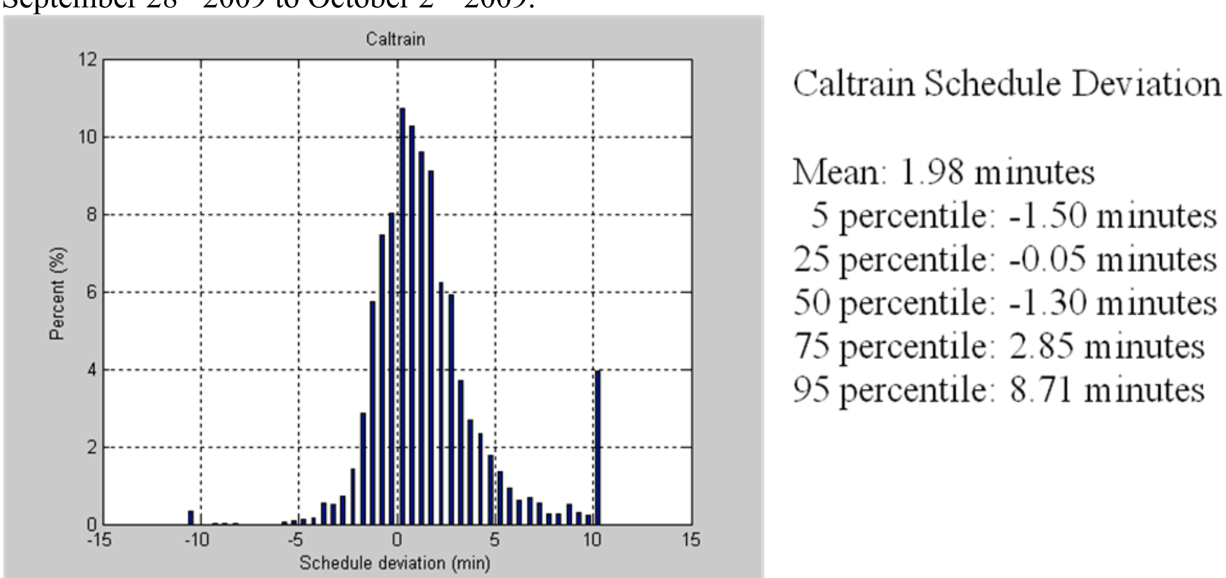
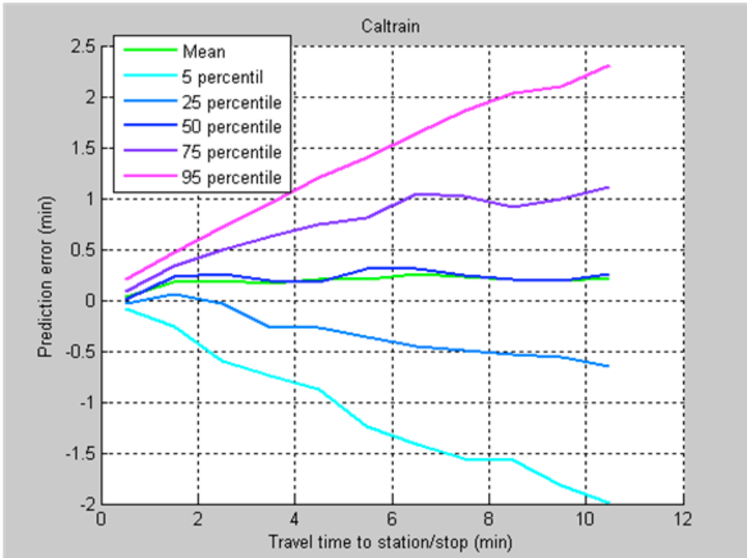


Figure 5-26: Caltran Schedule Deviation

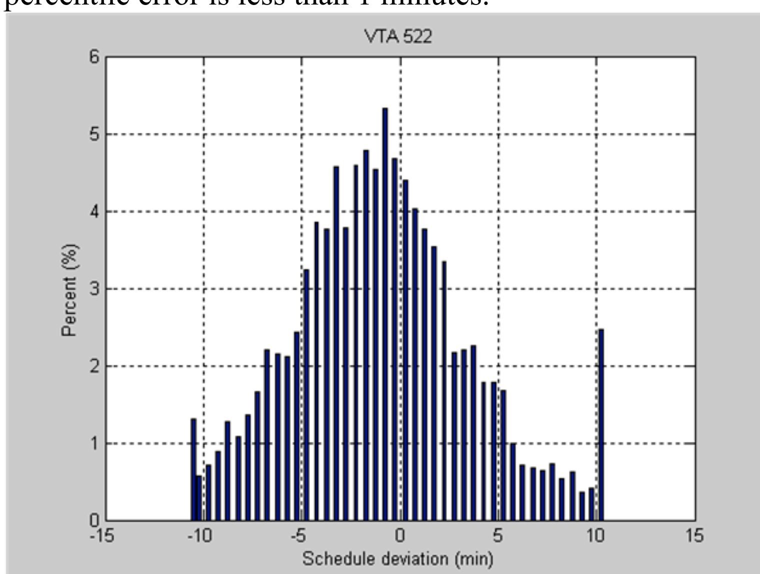


Caltrain predictive Arrival Time Deviation

Mean: 0.25 minutes
5 percentile: -1.80 minutes
25 percentile: -0.6 minutes
50 percentile: -0.25 minutes
75 percentile: 1 minutes
95 percentile: 2.2 minutes

Figure 5-27: Caltrain Arrival Time: Actual vs Predictive

The results showed that the mean deviation of the Caltrain predictive arrival time is averaged about 0.25 minutes for prediction over 10 minutes before the arrival at the stop. The 75 percentile error is less than 1 minutes.



522 Schedule Deviation

Mean: -0.66 minutes
5 percentile: -8.18 minutes
25 percentile: -3.7 minutes
50 percentile: -0.88 minutes
75 percentile: 2.01 minutes
95 percentile: 7.91 minutes

Figure 5-28: VTA 522 Schedule Deviation

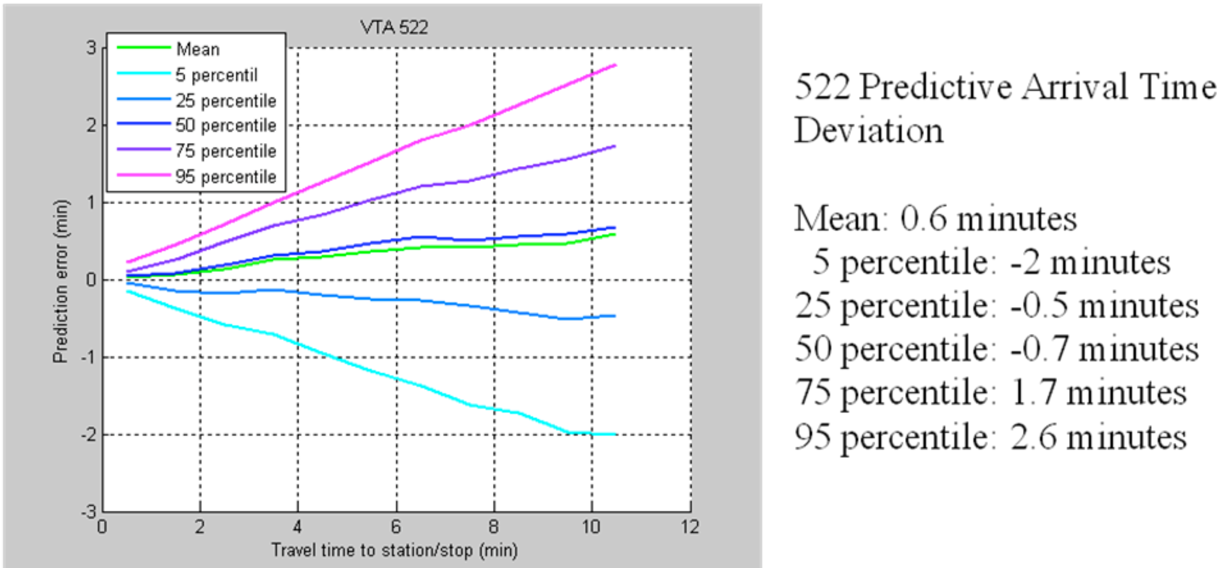


Figure 5-29: VTA 522 Arrival Time: Actual vs Predictive

The results showed that the mean deviation of the VTA bus predictive arrival time is averaged about 0.6 minutes for prediction over 10 minutes before the arrival at the stop. The 75 percentile error is less than 1.7 minutes.

5.6.3 Field Evaluation of the Bus-as-a-probe Travel Time Estimation

To verify the effectiveness of the developed method, we conducted a field test along a 3-mile segment of the three-lane arterial El Camino Real in Palo Alto, California. The test site is between Oxford Avenue and Jordan Avenue. This section is a part of VTA Rapid 522 bus line. There are only 3 bus stops in this section with 15 signalized intersections.

To measure the travel times of general traffic, we used the license plate matching method. After the installations of three video cameras at each side of the test site, we recorded license plates of all approaching vehicles from 5PM to 7PM on July 1st 2009. The total numbers of arrival vehicles at the origin and the destination are 3399 and 3461, respectively. The number of matched license plates is 497, which is 14.36% of all arrival vehicles. In average, it is about 83 sample travel times per 15-minute. Although this license plate matching method did not give detailed trajectories, it can provide enough samples to calculate a good ground truth of arterial travel time, which can help us calibrate and verify our model.

Based on the collected bus trajectories, the parameters for the bus stop model have been calibrated. The threshold velocity V^{stop} and the radius of bus stop area R are determined as 3m/s and 20m in this study. It is noted that the threshold values can be different for other sites and different GPS devices due to various reception strength and data accuracy.

By assuming traffic only delayed by traffic signal control and the resulted queues, the average intersection delay for all traffic is simply the average trip travel time minus the free flow travel time. The signal waiting time for all traffic was calculated by using the imaginary trajectory method. The queuing delay is the difference between the intersection delay and the signal waiting time. The average intersection delay for bus probes can be calculated by filtering the bus stop effects and cruise speed differences. The bus waiting time at signals was calculated based on the signal status data and the time when the bus departed the upstream intersection. The bus queuing delay is the difference between intersection delay and the signal waiting time. Figure 5-30 shows the comparisons of total intersection delay and queuing delay for all traffic and buses, respectively. According to the results, the bus probes and all traffic have similar intersection delay and also queuing delay. If we compare the model to use bus intersection delay to estimate average traffic intersection and the model to use bus queuing delay to estimate average traffic queuing delay, we found out the root mean square error (RMSE) for the queuing delay model is 34.9*sec*, which is about 9% better than 37.9*sec* for the intersection delay model.

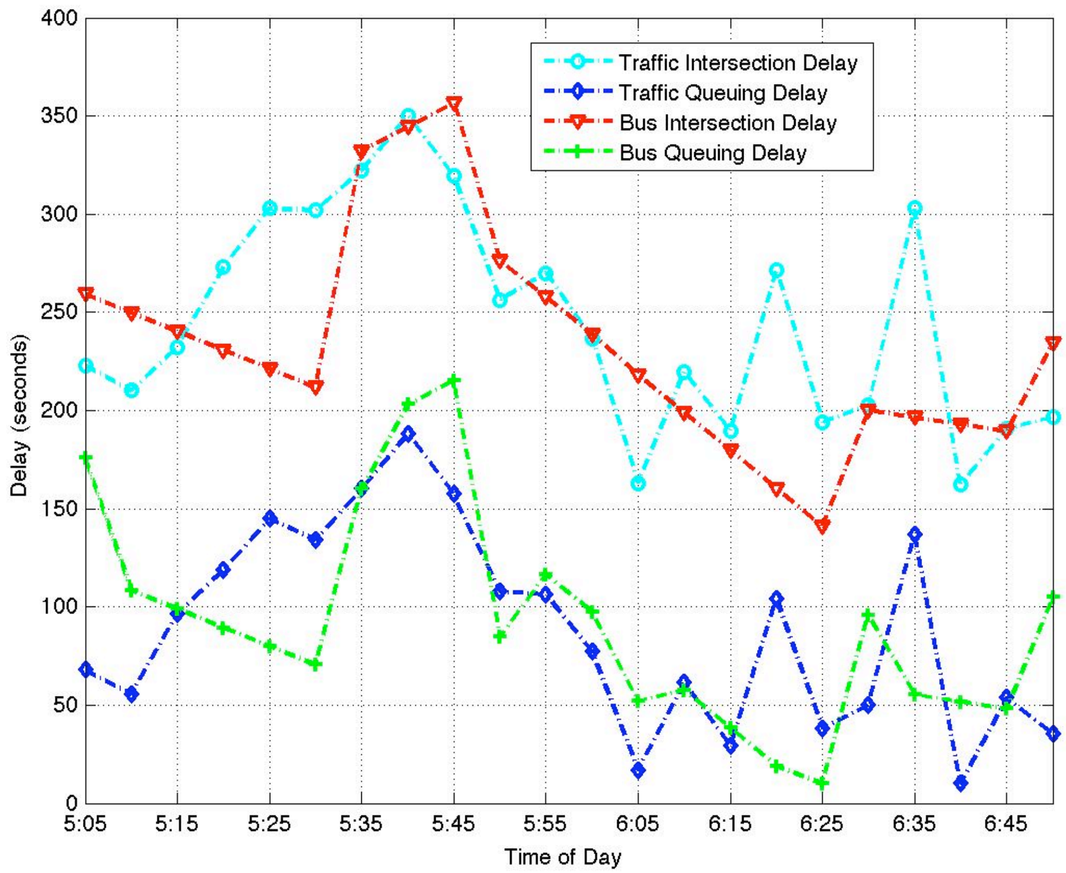


Figure 5-30: Delay comparisons between all traffic and bus probes

The arterial travel time was estimated by using the bus queuing delay plus free flow travel time and the average waiting time at signals for traffic. Figure 5-31 shows two measures of the arterial performance: travel time and level of service. For the travel time, the curve for estimation results traces the ground truth travel time well. The RMSE of the estimation is 49 seconds and the root mean square percentage error (RMSPE) is just 9%. For the level of service, the estimation model can well estimate the level of service with accuracy rate 73%. It is noted that the headway of Rapid 522 service is 15 minutes during peak hour and 30 minutes during non-peak time. The model linearly interpolated the results for every five minutes, which led to some deterioration of model RMSE.

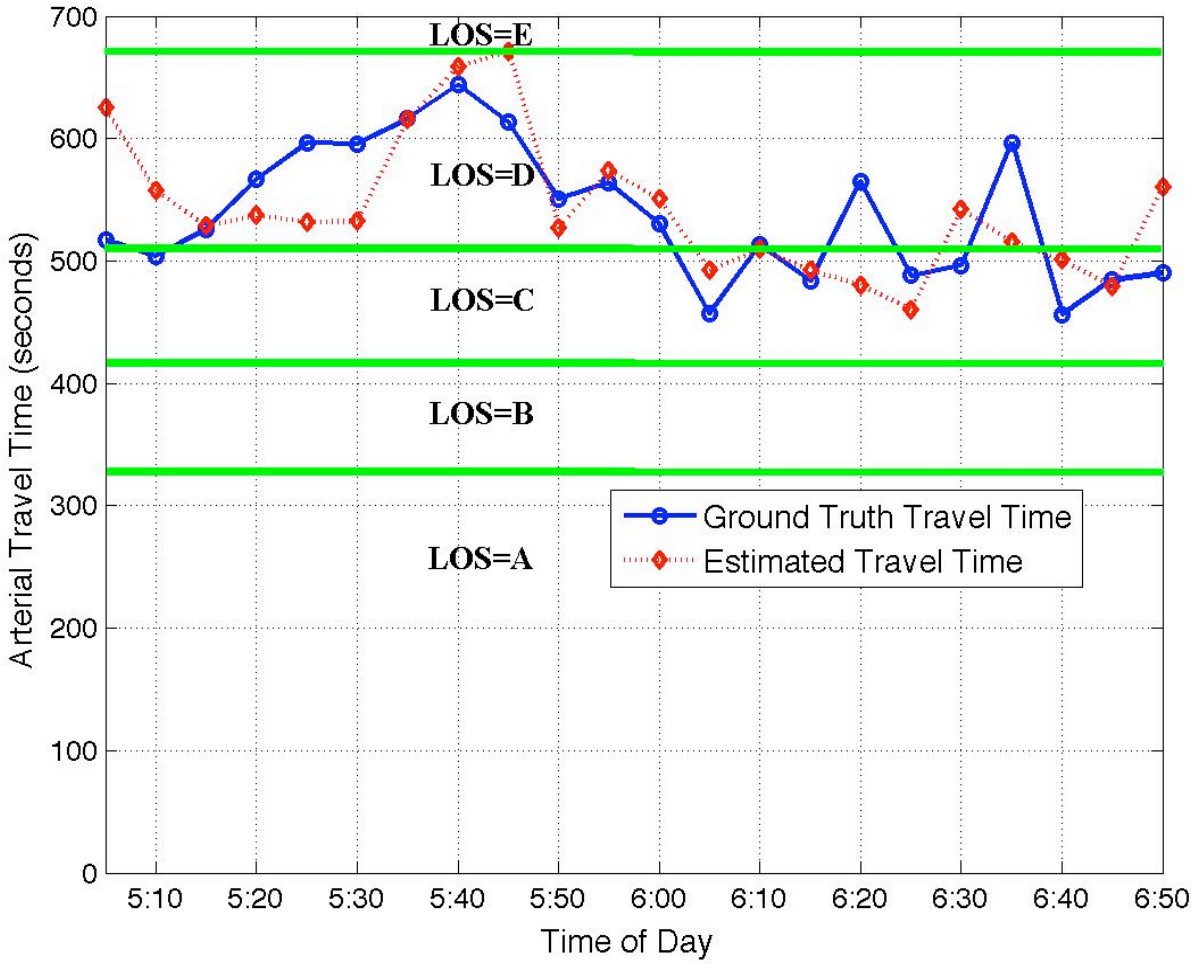


Figure 5-31: Arterial Performance Measurements

5.6.4 Field Evaluation of the NT Transit Services

Due to the fact that the NT transit services have not been publicly released and field testing has not yet started at the moment of this report, we are not able to present the results of the FOT.

6 References:

1. Bertini, L.R. and S. Tantiyanugulchai. Transit Buses as Traffic Probes: Empirical Evaluation Using Geo-Location Data. *Transportation Research Record* No. 1870, Washington D. C. 2004. pp.35-45
2. Berkow, M., S.M. Monsere, P. Koonce, R.L. Bertini, and M. Wolfe. Prototype for Data Fusion Using Stationary and Mobile Data Sources for Improved Arterial Performance Measurement. Proceedings of the 88th *Transportation Research Board Annual Meeting*, Washington, D.C., January, 2009.
3. Bureau of Public Roads. Traffic Assignment Manual. U.S. Department of Commerce, Washington, D.C. 1964
4. Chakroborty, P. and S. Kikuchi. Using Bus Travel Time Data to Estimate Travel Times on Urban Corridors. *Transportation Research Record* No. 1870, Washington D. C. 2004. pp.18-25
5. Chen, M. and Steven I.J.C. Dynamic Freeway Travel-Time Prediction with Probe Vehicle Data: Link Based Versus Path Based. *Transportation Research Record* No. 1768, Washington D. C. 2001. pp.157-161
6. Coifman, B. Estimating travel times and vehicle trajectories on freeways using dual loop detectors. *Transportation Research Part A*, Vol. 36, No. 4, 2002, pp.351-364
7. D'Angelo, M.R., H.M. Al-Deek, and M.C. Wang. Travel-Time Prediction for Freeway Corridors. *Transportation Research Record* No. 1676, Washington D. C. 1999. pp.184-191
8. Frechette, L.A. and A.M. Khan. Bayesian regression-based urban traffic models. *Transportation Research Record* No. 1644, Washington D. C. 1998. pp.157-165
9. Hall, W.R. and N. Vyas. Buses as a Traffic Probe. Demonstration Project. *Transportation Research Record* No. 1731, Washington D. C. 2000. pp.98-103
10. *Highway Capacity Manual*. TRB, National Research Council, Washington, D.C., 2000.
11. Liu, H. X. and W. Ma. A Real-Time Performance Measurement System for Arterial Traffic Signals. Proceedings of the 87th *Transportation Research Board Annual Meeting*, Washington, D.C., January, 2008.
12. Li, M., M. Song, and G.Y. Wu, "An Online Performance Measurement Method based on Arterial Infrastructure Data", Proceedings of the 88th *Transportation Research Board Annual Meeting*, Washington, D.C., January, 2009.
13. Li, M., Yin, Y.F., Zhang, W.B., Zhou, K., and NAKAMURA, H., "Modeling and Implementation of Adaptive Transit Signal Priority on Actuated Control Systems", *submitted to Computer-Aided Civil and Infrastructure Engineering,* Malden, MA, 2009
14. 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

15. Skabardonis, A. and N. Geroliminis. Real-Time Estimation of Travel Times along Signalized Arterials. 16th ISTTT conference, Maryland, 2005.
16. Turner, S.M., T.J. Lomax and H.S. Levinson. Measuring and estimating congestion using travel time-based procedures. *Transportation Research Record* No. 1564, Washington D. C. 1996. pp.11-19
17. Uno, N., H. Tamura, Y. Iida, Y. Nagahiro, and Y. Yamawaki. A study of Evaluating Level of Service of Road Network Based on Concept of Travel Time Reliability Using Bus Probe Data. *Infrastructure Planning Review* Vol. 19, No. 2. 2002. pp. 237-244
18. Van Line, J.W.C. Reliable Real-Time Framework for Short-Term Freeway Travel Time Prediction. *Journal of transportation engineering*, Vol. 132, No. 12, 2006, pp.921-932
19. Zhang, M.H. Link-journey-speed model for arterial traffic. *Transportation Research Record* No. 1676, Washington D. C. 1999. pp.109-115
20. http://www.raosoft.com/samplesize.html
21. Lenth, R. V. (2006). Java Applets for Power and Sample Size [Computer software], retrieved Mar 02, 2009, from http://www.stat.uiowa.edu/~rlenth/Power.
22. http://www.quantitativeskills.com/sisa/calculations/sampshlp.htm

7 Appendices:

Appendix-A: Candidate FOT Sites for the Safety Application

This section contains the list (Table A1) and maps of candidate sites currently being considered for Safety Application #1.

Note: Nomenclature in the tables below:

- B: Northbound; SB: Southbound; EB: Eastbound; WB: Westbound
- PM: Post Mile as defined in California Highway Database on a certain route within a county

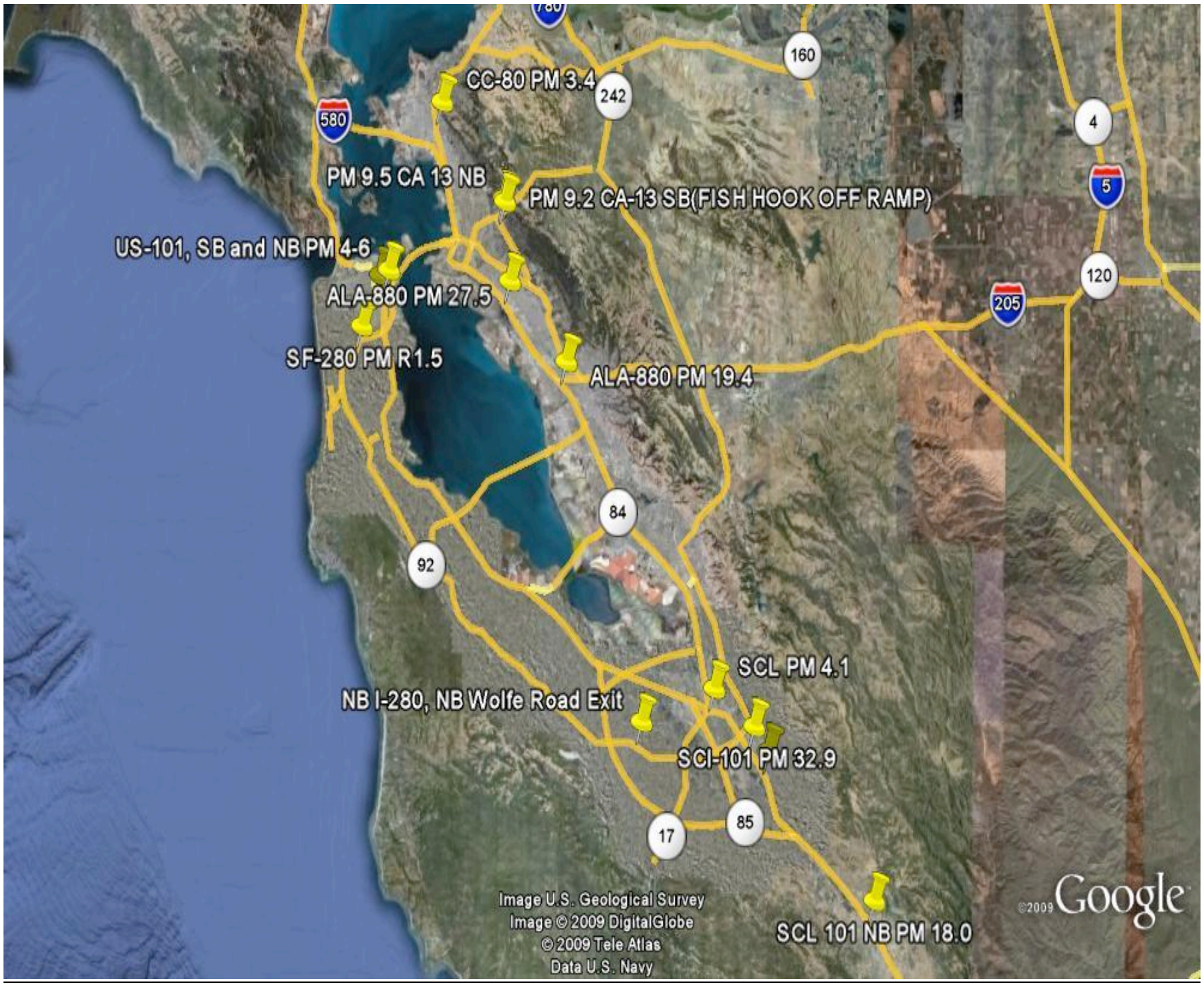


Figure 7-1: Area Wide Map of Potential Study Sites Locations

Table 7-1: Candidate Sites for Field Tests of Safety Application #1

Site No.	Site Location	Site Characteristics	Type of Situation Awareness
1	Alameda County SR-13, NB PM 9.5	• Limited line of site to sections ahead of curve • Slow queue ahead due to off-ramp backup into mainline	• Slow traffic ahead on right lane due to bottleneck • Slow traffic after curve
2	Alameda County SR-13, SB PM 9.25 Broadway Terrace Off-Ramp	• Severe fish-hoop off-ramp • Combined with a stop and traffic light at end of ramp	• Curve over-speed • Queue at off- ramp
3	Alameda County I-880, SB PM 27.5	• Combined vertical and horizontal curves • A merge of high-street	• Slow traffic after curve

		entry ramp and mainline prior to curve • Congestion in rush hours	
4	Alameda County I-880, NB PM 19.4 Connector to SR-238	• Off-tramp bottleneck with backup into right lane of mainline traffic	• Slow traffic ahead on right lane due to bottleneck
5	San Francisco County I-280, NB PM 1.5 Geneva Off-Ramp	• Off-tramp bottleneck with backup into right lane of mainline traffic	• Slow traffic ahead on right lane due to bottleneck
6	SR101 NB, PM 4.2, Mission & Dubose off Ramp	• Traffic backing up in the off ramp. • Congestion	• Traffic Weaving • Slow traffic due to congestion
7	San Francisco County US-101, SB and NB PM 4-6 (Hospital Curve)	• Combined vertical and horizontal curves • Congestion bottleneck • Traffic Weaving with on- and off-ramp traffic	• Slow queue ahead • Traffic Weaving
8	Santa Clara County I-880, NB PM 4.2 Connector to US-101 NB	• Frequent collisions with K-rail barrier on left side	• Off-ramp curve
9	Santa Clara County US-101, SB PM 32.9 Segment between Tully Road and Story Road	• Considerable traffic weaving between two exits in congestion periods	• Traffic Weaving
10	Santa Clara County US-101, NB PM 18.0 Near Cochrane Interchange	• Near 60% of collisions on left lane • Near Transition of 3-Lane into 4-lane segment with HOV lane on left • On-ramp and off-ramp nearby	• Lane Transition Traffic Weaving
11	Santa Clara County US-101, SB Tully Road EB Exit	• Speeding a major factor (more than 75%) • Rear-end Collision dominate (near 90%) • More than 85% ramp collisions at ramp exit and cross street	• Slow queue ahead • Ramp over-speeding

12	Santa Clara County I-280, NB Wolfe Road Exit	• Limited line of sight to ramp queue from mainline • Signal controlled cross street at ramp exit • More than 90% ramp collisions at ramp exit and cross street	• Slow queue ahead • Ramp over-speeding
13	Contra Costa County I-80, EB PM 3.4 Segment between San Pablo Ave and Solano Ave Exits	• Combined vertical and horizontal curves • Congestion bottleneck	• Slow traffic ahead due to bottleneck • Traffic Weaving

Appendix-B: Candidate FOT Sites for the Transit Application

This section contains the candidate sites currently being considered for Transit Application.

Table 7-2: Candidate Sites for Field Tests of Transit Application

Site No.	Site Location	Site Characteristics	Type of Tests
1	VTA buses	• Express buses and BRT buses	• Bus as probes • En-route traveler information
2	SamTrans Buses	• Regular bus service	• En-route traveler information
3.	Caltrain	• High speed rail	• En-route traveler information
4.	El Camino Real	• Major Arterial	• Buses as probes
5.	US-101	• Freeway	• Buses as probes

Appendix-C: GPS Data Mining System for Safety-related Benefit Analysis

Database Tables:

links

```
create table links(linkid bigint, fromid tinytext, toid tinytext,
shapeid tinytext, linklength float, dirtravel char default 'F'
not null, functclass int default 0 not null, speed_cat int,
primary key(linkid)); a

+------------+------------+-----+
| Field      | Type       | Key |
```

```
+------------+-----------+-----+
| linkid     | bigint(20) | PRI |
| fromid     | tinytext  |     |
| toid      | tinytext  |     |
| shapeid   | tinytext  |     |
| linklength | tinytext  |     |
| dirtravel | char(1)   |     |
| functclass | int(11)   |     |
| speed_cat | int(11)   |     |
+------------+-----------+-----+
```

nodes

create table nodesslc(nodeid int, longitude int, latitude int, functclass tinytext, primary key(nodeid));

```
+------------+----------+-----+
| Field      | Type     | Key |
+------------+----------+-----+
| nodeid     | int(11)  | PRI |
| longitude  | int(11)  |     |
| latitude   | int(11)  |     |
| functclass | tinytext |     |
+------------+----------+-----+
```

rawprobedata

create table rawprobedataslc(ID bigint not null auto_increment, TimeStamp timestamp, DeviceID integer, Latitude double, Longitude double, Speed float , Bearing float, LinkID bigint, Flag boolean default 0, primary key(ID), Unique index (TimeStamp, DeviceID));

```
+-----------+-----------+-----+
| Field     | Type      | Key |
+-----------+-----------+-----+
| ID        | bigint(20) | PRI |
| TimeStamp | timestamp | MUL |
| DeviceID  | int(11)   |     |
| Latitude  | double    |     |
| Longitude | double    |     |
| Speed     | float     |     |
| Bearing   | float     |     |
| LinkID    | bigint(20) |     |
| Flag      | tinyint(1) |     |
+-----------+-----------+-----+
```

Note: the ID field is auto_increment.

shapes

create table shapes(shapeid int, sequence int, longitude double, latitude double, numseg int, primary key(shapeid, sequence));

```
+-----------+---------+------+-----+
| Field     | Type    | Null | Key |
```

```
+-----------+---------+------+-----+
| shapeid   | int(11) | NO   | PRI |
| sequence  | int(11) | NO   | PRI |
| longitude | double  | YES  |     |
| latitude  | double  | YES  |     |
| numseg    | int(11) | YES  |     |
+-----------+---------+------+-----+
```

Note: shapes is an ordered table. The shapeid should maintain the same ordering as shapeid in the links table. The sequence ordering is also important.

traveltimes

create table traveltimes(linkid char(32), traveltime integer, timeofday timestamp);

```
+------------+-----------+
| Field      | Type      |
+------------+-----------+
| linkid     | char(32)  |
| traveltime | int(11)   |
| timeofday  | timestamp |
+------------+-----------+
```

Path Algorithm Diagram: