

Intellimotion

Keeping up with California PATH Research in Intelligent Transportation Systems

Volume 17, No. 1 2012

Page 1
Connected Corridors
and Hybrid Data Fusion

Page 3
The TOPL Project:
Tools for Active Traffic Management

Page 8
Mobile Millennium—An Overview

California PATH—Partners for Advanced Transportation Technology—is a collaboration between the California Department of Transportation (Caltrans), the University of California, other public and private academic institutions, and private industry.

PATH's mission: applying advanced technology to increase highway capacity and safety, and to reduce traffic congestion, air pollution and energy consumption.

CALIFORNIA
PATH

Connected Corridors and Hybrid Data Fusion Integrated Corridor Management (ICM) in California and Weaving Wireless Data into Traffic

Connected Corridors is a large scale Field Operational Test of next generation traffic management concepts that will maximize traffic flow on America's most complex network - California's transportation network. This pilot test is being developed under the oversight of Caltrans Division of Traffic Operations - by a large interdisciplinary team of leading transportation scientists, researchers, and engineers at California PATH. They are supported by a broad industry group of participants and traffic professionals: the Connected Corridors Consortium.

The goal of Connected Corridors is to maximize the free flow of vehicles, people and goods on a traffic corridor's roadways and to reduce travel times across the system. At its heart, this means managing all controllable components of a highway corridor to keep the number of vehicles, and the speed they are traveling, in optimal balance for the network. It means maximizing the productivity of roads without building more lanes, and it means getting travelers to collaborate closely with the traffic system and help everyone work together to lessen congestion.

This pilot ICM program uses Caltrans' real time/historical traffic data, its existing network of more than 20,000 detectors and probes, along with its institutional understanding of California's unique congestion landscape. Connected Corridors then fuses that with GPS point speed data collected in near real time to knit together a sophisticated estimate of current corridor traffic. Based on this a forecasting system models, simulates, and predicts how traffic will develop over the next minutes or hours depending on congestion, time of day, weather, incidents, events and more; then it can apply various potential highway control changes and dynamically show the predicted results of each. It is built to spit out various best options for different scenarios, giving traffic managers decision support tools to enable true corridor choices. Finally all of this is integrated into adaptive ramp and arterial and speed and message control systems that give traffic management centers (TMCs) the ability to actually affect traffic flows, meters, and signs before congestion occurs, ahead of real-time.

The Connected Corridors program integrates a number of formerly discrete tools under one umbrella and is built on top of a rich and detailed model of California traffic. It uses University of California Berkeley breakthroughs such as Tools for Operations & Planning (TOPL) for freeway control and simulation, the statewide Caltrans Performance Measurement System (PeMS) traffic monitoring and detection system, and Mobile Millennium software systems to read, map and analyze GPS point speed data in near real time.

Connected Corridors is also designed to bring commuters and travelers into the system as full partners in reducing congestion. The same traffic monitoring system that is directing traffic managers, can provide travelers with real-time alternative routings, incentives, integration with transit and parking data, incident alerts, and crowdsourcing to encourage reductions in demand for congested corridor roadways at key times.



continued on next page

Connected Corridors is an unprecedented effort to bring California drivers and their transportation system together. The pilot for the system is planned for 2014.

HYBRID DATA FUSION

Weaving Wireless Data Into Traffic

Bringing Mobile Millennium's capabilities to the real world of California's highways means not only getting data from GPS point speeds and commercial data feeds, but making the output work with one of the most capable traffic monitoring systems in the world: the existing, extensive Caltrans PeMS (Performance Measuring System). Built at UC Berkeley for the California DOT over the past decade, it offers unprecedented real time information about the highways and roads in California, allows Caltrans traffic managers to control many aspects of their roads, and is available as a data source to California drivers, researchers, and engineers.

Mobile Millennium was originally an engineering research project in the university designed to show how cell phone signals could be used to monitor current traffic conditions. Once that was proven, the goal of buying commercial GPS data, processing it with Mobile Millennium software, and fusing it to PeMS data had to be put to the test. Caltrans asked California PATH to work out the framework, the technologies, the systems, the algorithms and software routines, and the policies and procedures that the agency could use to buy and process commercial data feeds and integrate them seamlessly into PeMS.

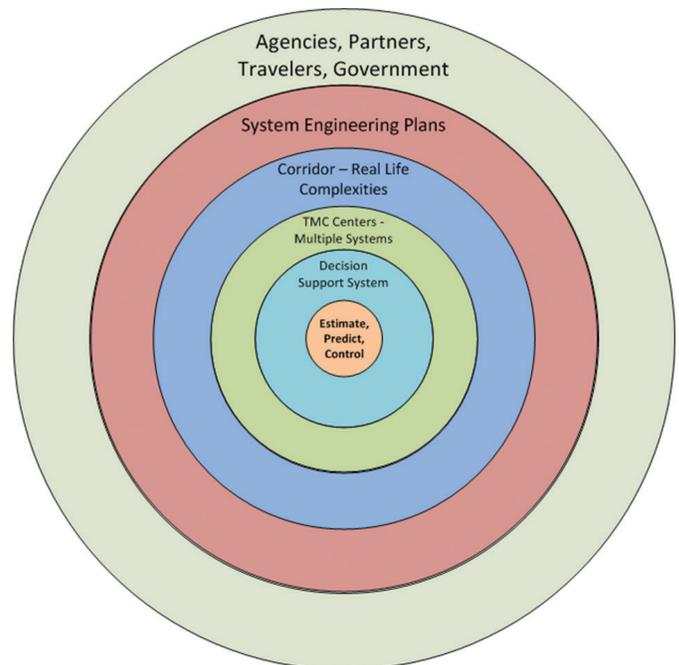
Essentially the assignment was to create a blueprint for Caltrans to put the widespread availability of commercial GPS fleet data feeds to work to enhance and extend the current monitoring system. Reading the new torrent of digital signals, making sense of them, integrating the synthesis in a meaningful and scientific manner to current maps of the roadway network, extending the already powerful capabilities of PeMS, and delivering it all is a complex challenge. Called Task Orders 1 & 2, this project is also known as Hybrid Data Fusion.

Two vendors survived the RFP process, TeleNav and Navteq, and both fleet data suppliers streamed raw, uninterpolated feeds which are in turn being processed by the university researchers. The project involves teams working on the mathematics of filtering—normalizing all the data sources so they can be compared analytically and displayed meaningfully—and “map matching” or tightly linking the traffic feed to maps. The team received raw commercial traf-

fic feeds, stored them, mapped them to corridor traffic derived from PeMS data; then they did a comparison of both commercial feeds to raw data sensed from Bluetooth (BT) readers running at the same times and places. These industrial-grade “readers” were put in place during two-week installations to triangulate the process. The BT readers record the unique identifier of an equipped and powered up smartphone as it passes under an underpass or gantry or pole; then the system records BT signals at several more locations along the corridor. By extracting speed and route information from the Bluetooth units (all personal identifiers are discarded) the traffic scientists can verify that the third-party probe data feeds being received relating to that area are accurate...and also use that kind of feedback to fine tune their algorithms and routines to get a more accurate picture of current conditions.

Another key element of the Hybrid Data Fusion project is to document the business process and the system requirements, for Caltrans to be able to buy, manipulate, integrate, and use private data sources in the future. The UC Berkeley researchers have had to work out how to funnel and refine the massive commercial GPS data feeds, how to store them, what to ask for in contracts, and how much to pay, in effect pricing the data for the public sector.

Currently the Hybrid Data Fusion project is in the midst of analyzing their results and preparing final reports, using installations from several corridors in southern California. ☼



The Connected Corridors program is providing methods, tools and leadership for each circle of the ICM implementation diagram.

Bluetooth Diaries: Fusing Data in the Field

On Wednesday, March 28th, Bluetooth reader units, brackets, batteries and mounting supplies were transported to Southern California by van. Supported by Caltrans staff from District 8, PATH staff and post-doctoral students surveyed the proposed installation sites along I-15. Thanks to the detailed and conscientious site pre-selection process, only minor refinements had to be made to the final installation locations.

On Friday, March 30th, a Caltrans maintenance crew—Al Washington and Frank Chumacero—installed the Bluetooth units along the freeway with support from Patire and Juha Distribution. We are grateful for the cooperation and support of Caltrans District 8, including Thomas Ainsworth, Peter Acosta and their respective groups.

On the home front, work was underway to improve the reliability of our data feeds. Small issues kept popping up which needed to be solved. For instance, the Bluetooth data feed was fixed to handle daylight savings time...

because the time change occurred in the middle of collecting data. The systems lead for the project also completed the setup of a backup machine to be used to record the TeleNav feed, and another machine as backup for NavTeq and PeMS feeds.

Another challenge involved the fact that 60% of the disk space on a new database computer had been filled during the first month of the three month experiment. One of the two vendors unleashed a giant stream of data at us—their raw feed for the entire United States—and the group had to figure out how to discard most of it. The main problem had little to do with storage capacity, and more to do with the fact that we cannot write the data to the database fast enough. As a part of the analysis, we wanted to keep track of the amount of data that was thrown away. Keeping track of thrown away data was done too precisely, and this is what was filling up the disk space. So, to fix it, the team created a scheme to aggregate statistics more coarsely on the discarded data points.

The TOPL Project: Tools for Active Traffic Management

TOPL Group, UC Berkeley

The transition from infrastructure building to operations and demand management has proven challenging on many levels, including agency culture, extant skills, institutional fragmentation, and a paucity of adequate business practices and information systems. Yet progress is now being made. In the San Francisco Bay Area alone, the past couple of years have seen the implementation of congestion pricing on the Bay Bridge, the opening of the first of several regional high-occupancy tolling (HOT) lanes, and the start of the SF Park initiative, which provides real-time parking availability from sensors deployed at hundreds of parking spots and sets prices dynamically. This came in addition to the expansion of signal timing and ramp metering programs. Such advanced strategies are being promoted and funded by the Federal Highway Administration (FHWA) under the general designation of Active Transportation and Demand Management (ATDM)

ATDM is the ability to dynamically manage recurrent and non-recurrent congestion based on prevailing traffic conditions in order to maximize the effectiveness and efficiency of traffic networks. ATDM measures can be primarily categorized as supply side or demand side. Supply side measures seek to improve traffic network operation by directly changing capacity, speed and/ or signal timing plans. Demand side measures seek to improve traffic conditions by affecting the demand for travel through pricing, information, and marketing.

ATDM strategies evolve as quickly as the detection, communication, data processing, and actuation technologies they employ.

TOPL (Tools for Operations Planning) [1] is a suite of software tools for specifying such operational improvements and for quickly estimating the benefits such improvements are likely to provide. TOPL simulation software is based on the dynamic macroscopic cell transmission model (CTM) of a traffic network [2].

The ATDM workflow and the role of the TOPL technology are shown in Figure 1. This workflow consists of three loops. The purpose of the strategic loop is to process the historical traffic data, create travel demand forecasts, and perform cost-benefit analysis of potential ATDM measures, based on which the necessary infrastructure is deployed. It generally takes between 2 and 5 years to complete. Tactical loop refers to the real-time (or near real-time) traffic operations, where the traffic situation for the next day or few hours is modeled under different scenarios, and the best performing out of available ATDM strategies are selected. The third one, dealing only with the supply side measures, is the automation loop, where the traffic controllers are sophisticated enough to properly adjust to the traffic situation without interference of the operator. We envision TOPL not only as part of strategic and tactical loops of the ATDM workflow, but also as a necessary component of the automation loop, since non-recurrent traffic scenarios cannot be handled without simulation.

ATDM postulates substantial flow improvements as a result of close network monitoring and proactive application of both traffic control and pricing schemes. A related concept, integrated corridor management (ICM), seeks further gains from interconnecting different transportation subsystems: say a freeway, a parallel rail transit corridor, and adjacent arterial streets. ATDM and ICM are being demonstrated at several pilot sites around the county, and represent an evolution from existing traffic management techniques. However, proactive, concurrent management of prices, ramp meters, traffic signals and other control devices across an entire metropolitan network sets a whole new level of complexity.

As a result, roadway operators, both public and private, as well as the consulting and engineering firms that serve them, require a new generation of transportation management tools. Specifically, transportation practitioners are suddenly recognizing the need for Decision Support Systems (DSS) that can simulate network operations and guide day-to-day management as well as operational planning. While there are numerous traffic simulation software packages available on the market today, none of these is even close to running fast enough to support operations. Conversely, traffic management software that is used by operators reports current conditions and controls field devices, but it offers no comparative assessment between possible alternatives. The general concept of real-time traffic management with the proposed DSS is presented in Figure 2.

Given the specification, the road network is built with the Network Editor and stored in the Network Layout database. Traffic data needed for model calibration and data sets for simulation input are stored in the Historic Data repository. Network Demand and Network Capacities databases contain the quantities computed by the Data Processor invoking TOPL calibration and traffic flow imputation routines [3, 4]. The Network Events database stores possible events such as lane closures, incidents, change of control, traffic re-direction, etc. The Operations Toolbox is a database of controllers that operate on the link, node or network level and can be activated if necessary. All this data can be accessed, modified and used for scenario creation by the traffic operator. Network Editor is especially handy for fast road network modifications, such as creating detours in case of large incidents.

A fast and trusted simulator is the key component of the DSS. It runs simulations in off- and on-line modes. In the off-line mode performance measures are computed using statistically forecasted demand for different operational strategies, and the best performing strategy is selected as default for the next day. The off-line mode is also used for post factum daily traffic evaluations – determining if there were alternative operational strategies yielding better results. Finally, the off-line mode is used for quick assessment of Traffic Management Plans describing the traffic network operation under expected disruptions, such as planned construction.

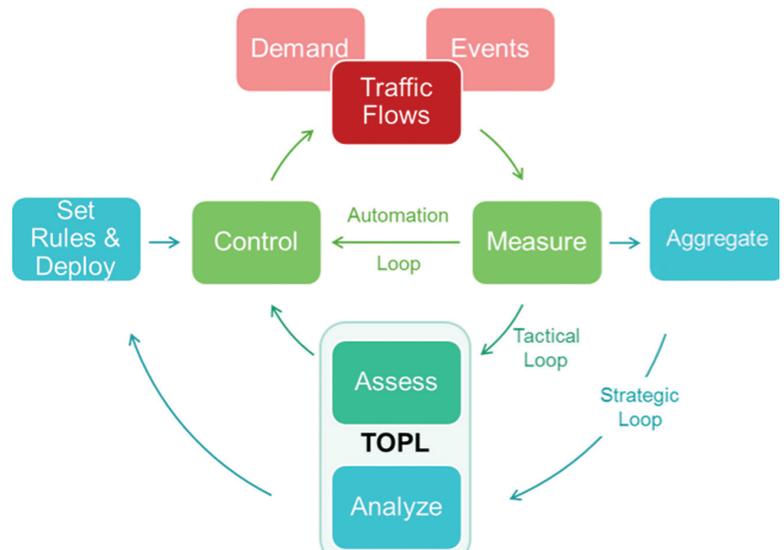


Figure 1: TOPL technology in ATDM workflow.

The on-line mode, on the other hand, provides short-term prediction operating continuously correcting the forecast with the real-time traffic data feeds. In the case of non-recurring and unexpected events, such as large accidents, the on-line prediction is computed for alternative operational strategies to select the best performing one. The other purpose of the on-line simulation mode is to filter the incoming measurements, passing them through the state estimator, before feeding them into field traffic controllers. The on-line simulation mode is the enabler of the real-time DSS.

Data to Models

One of the main ideas of TOPL is to enable quick and convenient creation of dynamic traffic models from available detector measurement data. PeMS [5] is an excellent source of California freeway data, while data for arterials (major urban streets) are in general not systematically collected and processed in the US. To obtain a traffic model ready for simulation, the following steps are required.

Step 1: Building Road Network

The Flash based Network Editor, shown in Figure 3, runs in a web browser and allows the user to create

road networks from scratch as well as to edit existing road network configurations. It loads and saves road networks from and to the user's local disk in the XML format of the TOPL simulator.

Step 2: Placing Detectors

Network Editor allows the user to create detectors by clicking with the mouse at desired locations on the Google Map, or by loading a text file in with detector locations provided, for example, by PeMS. Once the detectors are created, they must be associated with road links by dragging and dropping them onto desired links.

Step 3: Specify Detector Data Source

The user can specify URLs pointing to measurement data files for each detector individually, or for a list of selected detectors. A single PeMS data file, for instance, contains daily measurement data for all detectors in one Caltrans district – that is, several counties. In a real-time system, the URLs assigned to detectors will point to measurement data streams coming from the field.

Step 4: Calibrating Traffic Model

The automatic model calibration utility is invoked from the Network Editor. Using density and flow values from measurement data files assigned to sensors, the calibration utility rapidly estimates the key parameters of road links – capacity, free flow speed and jam density following the methodology of [3]. Missing flow data is imputed following the methodology of [4].

Step 5: Creating Scenarios

The Network Editor allows the user to create events by clicking on a mouse at given locations and specifying the time of their activation. The events are used for modeling incidents and other non-recurrent traffic situations. It is also possible to specify controllers (e.g. ramp meters) in the Network Editor or edit timing plans for the signalized intersections. Thus, the impact to certain incidents can be assessed and available response strategies evaluated.

Simulation and Analysis

The high speed of the TOPL traffic simulation results from the simplicity and scalability of the CTM model. There are three modes of operation of the simulation software: (1) pure simulation; (2) traffic state estimation, where the simulation results combined with the available detector measurements are used to determine the state of traffic in the parts of the road network with lack of detection; and (3) short term prediction, where based on the current state and the historical demand patterns the traffic behavior is forecasted for the next 1-2 hours. The simulation computes such standard performance measures as Vehicle Miles Traveled (VMT), Vehicle Hours Traveled (VHT), travel time, delay and productivity loss of the selected routes or road networks. Comparing these performance measures in addition to speed, flow and density contours for different scenarios helps in evaluating operational strategies in terms of the benefits they provide

Reports analyzing simulations of given scenarios can be generated and exported in PowerPoint, Excel or PDF formats, as shown in Figure 4.

I-680 South Case Study

We shall describe the concept of the TOPL DSS and its application in the concrete setting of Bay Area's I-680 South freeway between post miles 41 and 32, shown in Figure 5(a). This stretch of freeway has an HOV and three general purpose (mainline) lanes. The traffic model for this freeway segment was calibrated using the methodology of [3] with the measurement data from PeMS [5]. The missing demand data were imputed using the algorithm described in [4]. In this example, our simulation reconstructs the morning from 6.30 to 11.00 AM of November 15, 2011. An accident blocking two right lanes occurs near post mile 37, upstream of an off-ramp to Crow Canyon Road (Figure 5(a)), and it takes 10 minutes to clear this accident. According to California Highway Patrol data, for 50% of incidents

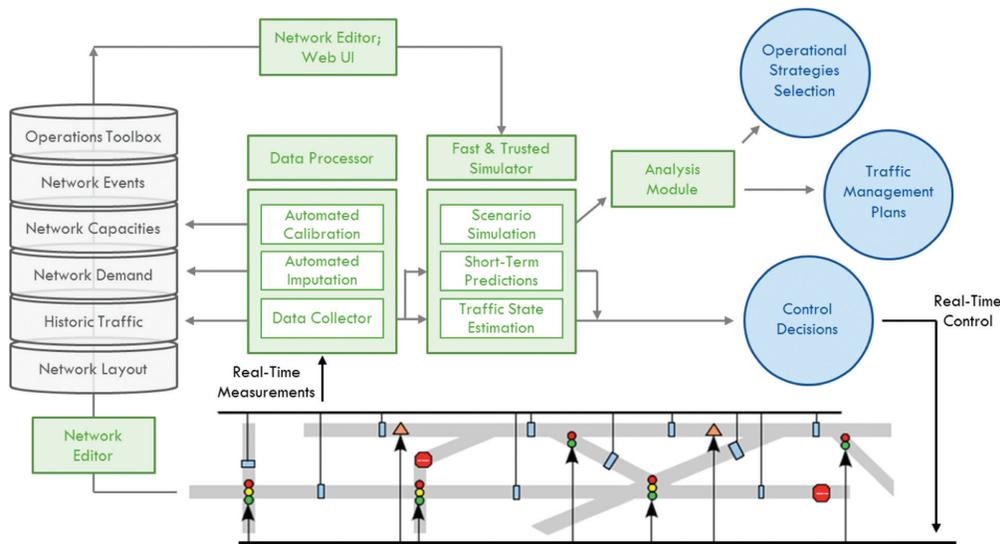


Figure 2: The general concept of the DSS for the real-time traffic management

continued on next page

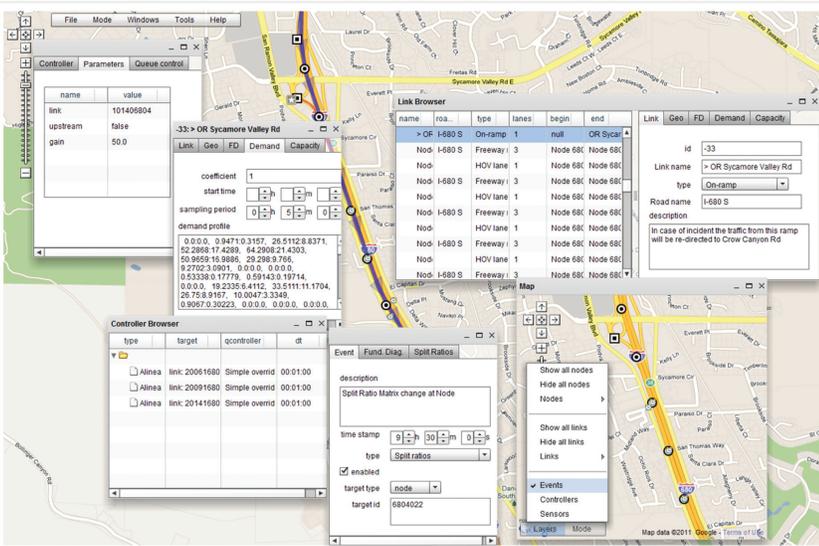


Figure 3: Network Editor – screen shot.

occurring in this particular freeway segment during the last 5 years the clearance time was longer than 10 minutes. Simulating this accident scenario under the assumption that no specific response strategies were deployed, we compute the total network delay, which includes ramp queue delay, shown in Figure 5(b), and speed contours for the mainline and HOV lanes, shown in Figure 5(c, d) – traffic moves from left to right, time increases from bottom to top. Other general performance measures retrieved from the simulation include: travel time for given route; vehicle miles traveled (VMT); vehicle hours traveled (VHT); productivity loss; vehicle queue sizes on the ramps. An additional general performance measure we plan to add is waiting time in ramp queues.

It is the task of traffic operators to come up with response strategies that would mitigate the consequenc-

es of such an accident. Here we present two possible strategies, named Response I and II.

Response I consists of three measures: (1) open the HOV lane for everyone upstream of the accident; (2) start aggressive ramp metering (RM), which in this particular case is ALINEA [6] with queue control at El Cerro Boulevard and Diablo Road on-ramps; and (3) redirect traffic from Sycamore Valley Road to Crow Canyon Road on-ramp, which is downstream of the accident, as indicated in Figure 6(a). These measures can be readily tested in the simulation and are shown to result in dramatic delay reduction – from 120 to 65 vehicle-hours per 5 minutes, as displayed in Figure 6(b). The simulation also shows more even distribution of congestion between the mainline and HOV lanes (Figure 6(c, d)), while reducing the extent of congestion. Can anything else be done to even further reduce the impact of the accident?

Response II, in addition to the three above mentioned measures, suggests divergence of traffic from I-680 South to a parallel arterial, San Ramon Valley Boulevard, through Sycamore Valley Road off-ramp via CMS, as presented in Figure 7(a).

Without considering the impact of the strategy on the arterial road, one may conclude that Response II is preferable to Response I: congestion on the freeway is further reduced by diverting vehicles (compare speed contour of Figure 7(c) with that of Figure 6(c)). The simulation, however, predicts that due to the signal at the intersection of Sycamore Valley Road and San Ramon Valley Boulevard, delays on the detour will outweigh reductions on the freeway. Furthermore, the

Specify report parameters

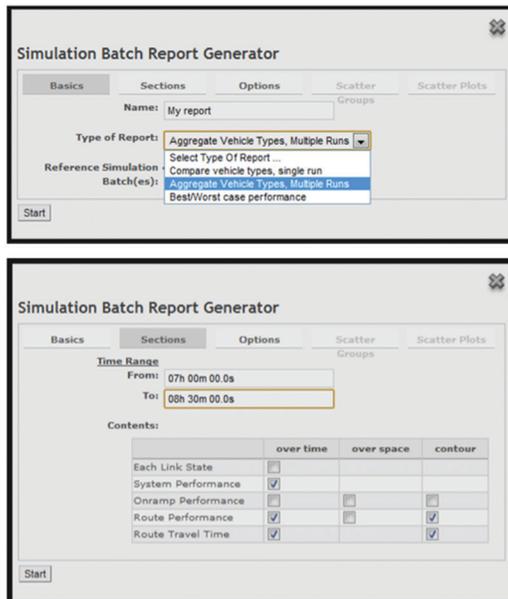
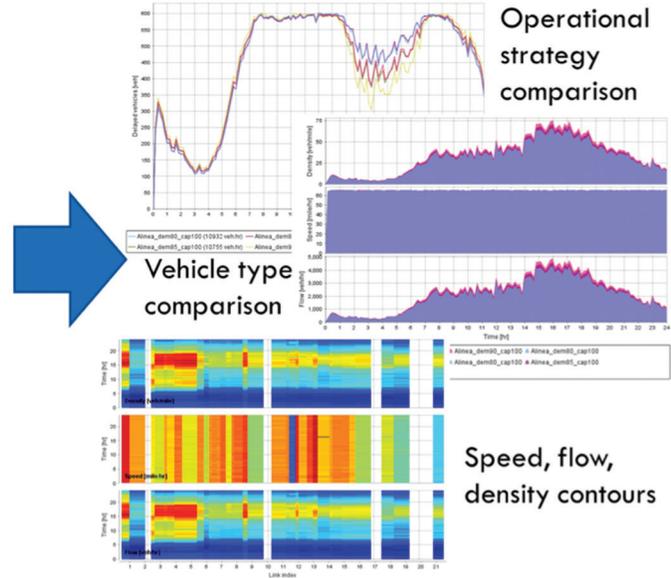


Figure 4: Simulation data analysis and reports.

Reports: PPT, PDF, XLS



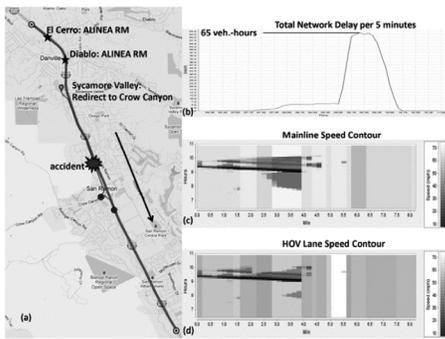


Figure 5: I-680 South - Accident near Crow Canyon Road.

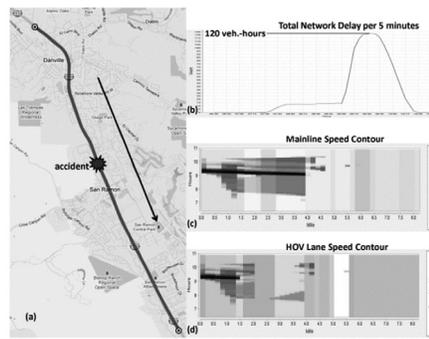


Figure 6: I-680 South accident Response I - open HOV lane, ALINEA RM, on-ramp flow redirection.

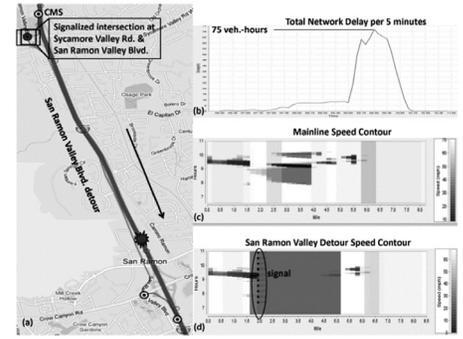


Figure 7: I-680 South accident Response II - alternate route through San Ramon Valley Boulevard.

queue from the intersection backs up into the freeway, causing additional problems (Figure 7(d)). Figure 7(b) displays the plot of the total network delay, which is larger than the total network delay under Response I. Thus, Response I is expected to be a better solution than Response II.

If the demand is not known exactly, but is expected to be within certain range (e.g. +/- 2% of some nominal time series), and/or road network parameters are determined with some uncertainty (e.g. the capacity of certain road segment(s) lies within +/- 2% of some average value), the TOPL simulation software can either run a batch of stochastic (Monte Carlo) simulations to obtain a distribution of possible outcomes, or compute the best and worst case bounds of performance measures, within which the system is expected to perform, as shown in Figure 8. ☼

References

- [1] The TOPL Project: <http://path.berkeley.edu/topl>
- [2] A.A. Kurzhanskiy and P. Varaiya. Active Traffic Management on Road Networks: A Macroscopic Approach. Philosophical Transactions of Royal Society, Part A, 368, pp. 4607-4626, 2010.
- [3] G. Dervisoglu, G. Gomes, J. Kwon, A. Muralidharan, P. Varaiya and R. Horowitz. Automatic Calibration of the Fundamental Diagram and Empirical Observations on Capacity. 88th Annual Meeting of the Transportation Research Board, Washington, D.C., USA, 2009.
- [4] A. Muralidharan and R. Horowitz. Imputation of Ramp Flow Data for Freeway Traffic Simulation. Transportation Research Record (2099):58-64, 2009.
- [5] PeMS - California Performance Measurement System. <http://pems.dot.ca.gov>
- [6] M. Papageorgiou, H. Haj-Salem, J.M. Blosseville. ALINEA: A Local Feedback Control Law for On-Ramp Metering. Transportation Research Record 1320, pp. 58-64, 1991.

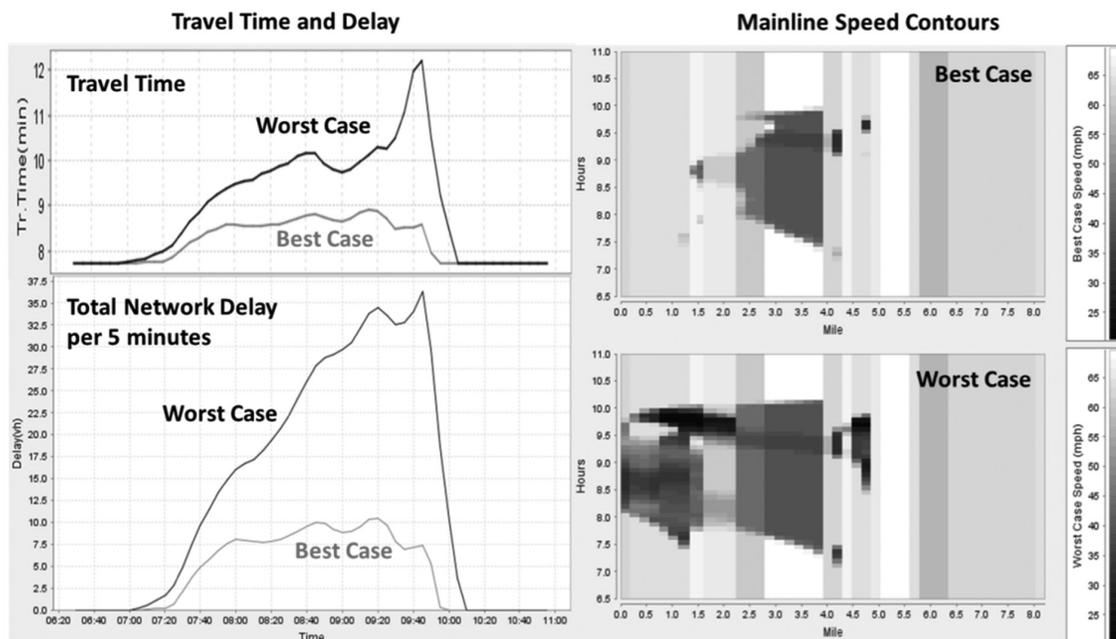


Figure 8: Traffic prediction under uncertain demand and model parameters.

Mobile Millennium – An Overview

The Mobile Millennium Research Team

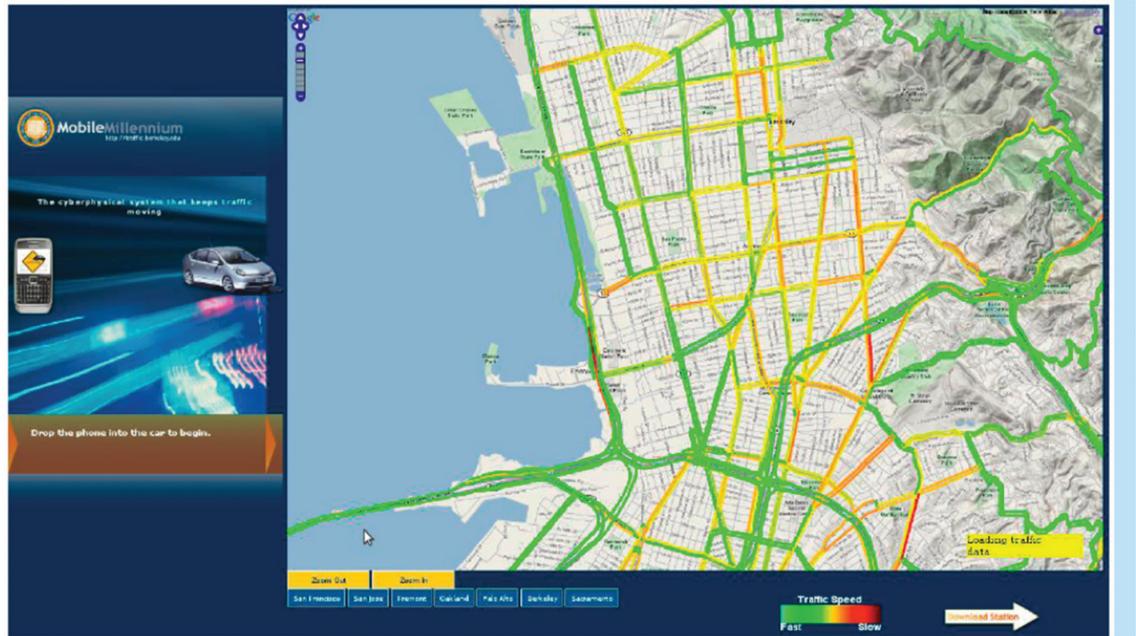


Figure 1 Mobile Millennium Smartphone Interface

The Mobile Millennium project was a groundbreaking pilot deployment that pushed forward the state of the art in using cellphone links and smartphone applications to map live traffic at scale. It explored San Francisco Bay Area traffic crowdsourcing—the Internet fueled idea of getting thousands of users sharing location and speed data effortlessly through a smartphone, then feeding it through a system of software and algorithms and maps which in turn supplied up to the minute traffic estimates of current state back to all participants. It built and operated a web application using participatory sensing to make it effortless, supported and maintained it, and explored and evaluated the technical feasibility of creating this kind of traffic information system relying on probe data from cellphones. More than 5,000 drivers in the San Francisco Bay Area participated and the traffic estimation engine ran successfully between 2008 and 2009. The crowdsourcing link between participants who used the Mobile Millennium application on their smart phones to provide location and speed, and the system’s ability to create accurate traffic estimations on an industrial scale in return, was at the heart of the project and paved the way for many of today’s commercial mapping and traffic services.

Traffic Information Systems

Traffic problems can be sorted in three main categories:

- Traffic information systems
- Operation systems (traffic control)
- Planning

Mobile Millennium focuses on the first of these three categories: traffic information systems. The fundamental question of interest was: “would it be possible to create traffic information systems based principally on probe data?” This question is not only of interest to the government, but also to industry, as demonstrated by the large amount of traffic applications that emerged from the private sector since the start of the project. Solving the problem of determining if such a traffic information system can be built relies on numerous other questions that had to be investigated as well as part of this work: In particular, how to create the proper infrastructure to collect data appropriate for building a next generation traffic information system, mapping it, and distributing it to travelers in a useful form.

Mobile Millennium was essentially a technology driven project. While it investigated some of the policy questions pertaining to the collection of probe data, in particular privacy, it did not investigate the implication of shifting from loops to crowd sourced / participatory sensing data. The goal was mainly to establish the technical feasibility of such a system, and to demonstrate that a proper recruitment strategy could be built that would lead to successful operations of a pilot through a field operational test. Solving the technical challenges leading to such a test required numerous breakthroughs in research, which are part of the achievements of this work as well as significant development of systems capabilities, which were done at UC Berkeley and Nokia.

The Mobile Millennium project is one of the most successful examples of a tripartite collaboration and joint effort between Government, Academia and Industry in the field of transportation. The U.S. DOT SafeTrip-21 (Safe and Efficient Travel through Innovation and Partnerships for the 21st Century) Initiative was established to test and evaluate integrated, intermodal, intelligent Transportation Systems (ITS) applications. As part of this program an award was made to establish the California Connected Traveler (CACT) Test Bed, which involved integrated locations in the San Francisco Bay Area.

The California Connected Traveler test bed was led by the California Department of Transportation in partnership with the University of California at Berkeley and two organizations administered by UCB's Institute of Transportation Studies: California Partners for Advanced Transportation Technology (PATH), with a mission to develop solutions to the problems of California's surface transportation systems through cutting edge research; and California Center for Innovative Transportation (CCIT), with a focus on deployment and a mission to accelerate the implementation of research results. Mobile Millennium was hosted by CCIT and a consortium of funders, the most important of which were the US Department of Transportation and the California Department of Transportation.

The Mobile Millennium project was also a partnership established with the Nokia Research Center (NRC) in Palo Alto, one of the research branches of Nokia in the US. Nokia provided hundreds of highly subsidized GPS phones to enable the deployments, and equipped its employees with thousands of these internally which represented a considerable investment of resources. In collaboration with the Berkeley team, the Nokia Research Center set the technical agenda of the deployment. The NRC staff continued to design, develop and operate client-side and server-side software that served as the engine for traffic data collection and dissemination during the Mobile Millennium field operational test.

At the time the project started, Nokia was in the process of acquiring NAVTEQ. Being the industry leader in the production and distribution of digital maps and real-time traffic information, NAVTEQ brought technical know-how and operational expertise to the program. They also brought maps that were used in the project, as well as NAVTEQ traffic patterns, a NAVTEQ product that contains historical traffic data. NAVTEQ also led the dissemination of the Mobile Millennium technology and information at the 2008 World Congress in New York City. The Mobile Millennium project team worked with NAVTEQ's subsidiary Traffic.com to fuse probe data with current traffic information and to leverage existing distribution channels.

Systems Architecture

Developing the systems architecture for an undertaking of this magnitude was the work of a vast cadre of researchers, graduate students and post-docs.

The overall system is a combination of phones, cellular networks, and back end software. As Figure 3 shows, processing was split between Nokia and UC Berkeley. A cell phone user would download a client from the Berkeley web site. This client would connect over the cellular network to servers at Nokia. The client would also download VTLs or virtual trip lines to the phone. The virtual trip lines determined at which points the cell phone client would measure velocity and location. As a phone crossed a trip line this information was sent back to the Nokia servers where it was cleaned of all user identifiable information. The anonymous information was then sent to UC Berkeley where it was cleaned, filtered, associated with an actual roadway in the Navstreets (NAVTEQ) database, fused with other data sources, and passed to algorithms that used the data to estimate the speed of traffic on Bay Area roads. These speed estimates were sent back to the Nokia servers where they were used to construct traffic maps where the color of a road corresponded to the average speed of vehicles on the road. This process occurred in real time every weekday for over a year.



Figure 2 Map of NYC arterial route used for each NYC road test and the ITS World Congress demonstration. The red dot indicates the traffic signal nearest to the Javits Convention Center; the arrows show how many lanes and the blue dots are traffic signals.

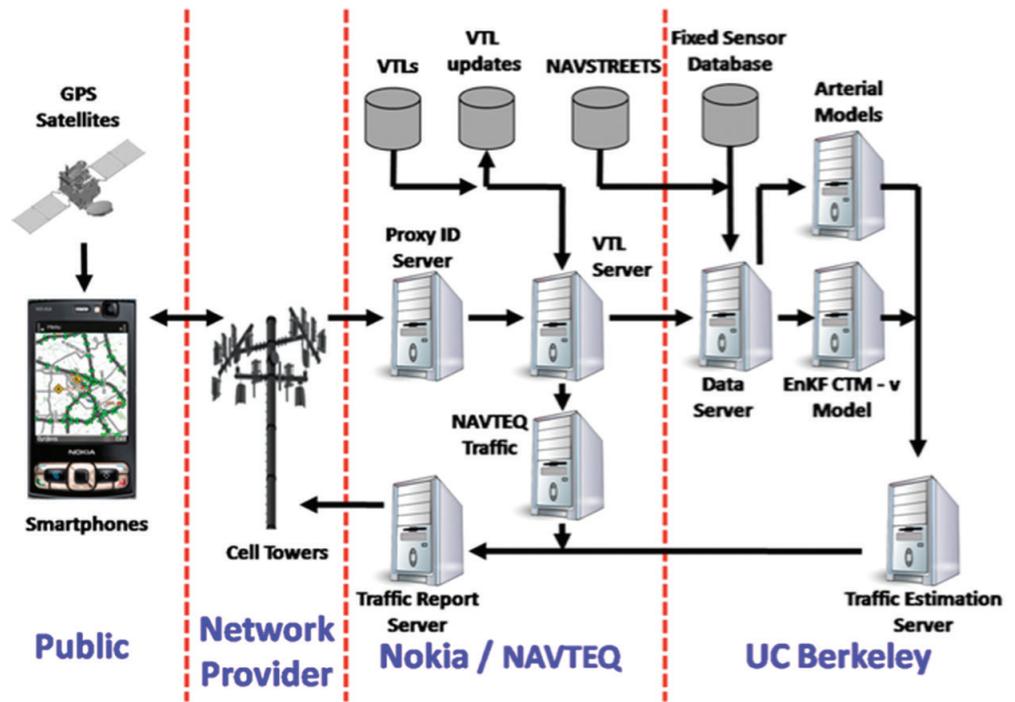


Figure 3: Mobile Millennium System Architecture (Nokia & Berkeley)

The majority of the data processing and complex algorithm execution occurred on the Berkeley servers. The Mobile Millennium system at UC Berkeley is a combination of software, hardware, procedures, and personnel all focused on providing high volume, high reliability, and real-time data processing. It was developed and used by over 50 students, interns, engineers, and scientists. It is composed of roughly half a million lines of code, over 100 database tables, multiple hardware servers, and several development environments.

As data flows through the system, numerous data feeds are filtered, the refined output fed to state of the art estimation and fusion algorithms and the resulting information visualized and presented for data exploration. Figure 4 is a data flow diagram for the system. The headings across the top are major processing steps, the small rectangles on either side are data sources and sinks, the circles within are individual processes, and the rectangles at the bottom are cross cutting functionalities.

The output of these models is used in a number of applications before being sent to third parties for consumption or analysis. Underneath the flow of data through the system are several components needed for quality analysis and visualization of each step of the process. With this in mind, the Mobile Millennium team built an evaluation framework and internal visualizer for comparing and analyzing data from multiple sources using several quality metrics. These allow for quick checking of the data through all steps of the process, from raw data to filtered data to model outputs.

Another way of looking at the Mobile Millennium system is depicted in Figure 5. This figure illustrates the way the components of the system interact.

In general, the database is the central point for communication between processes. This allows for a modular system where components can function independently without worrying if another component fails. The system software was designed so that one core module directly interacts with the database and requires that all requests to receive or send data be passed through that module.

Goals of the Mobile Millennium Project

After the success of Mobile Century, a precursor UC Berkeley project which demonstrated the feasibility of traffic monitoring by cell phone information in a controlled environment, it became clear that in order to push the concept to an operational level, it would be necessary to demonstrate similar capabilities of a system in an environment in which users were now representative of the general public. Of course such a setting comes with significant challenges, some of which were known at the time, some of which appeared later during the field operational test. Mobile Millennium was developed at a large scale from the start.

Another goal of the project was to create a proper intellectual property (IP) framework, to enable academia, industry, and government to work together. The challenge of this goal was to find a proper way for both research teams (UC Berkeley and Nokia) to be able to work jointly, and preserve each player's abilities to

protect their IP. The process of creating an IP agreement was a preliminary to starting the joint work, and was a major achievement of the project.

The most important technological goal of the work was to build a prototype traffic monitoring system, which would serve as the backbone for operations during the field operational test. Nokia and UC Berkeley chose an extremely challenging path, which was to have a part of the system reside at UC Berkeley, and another reside at Nokia (the two being interfaced). This was successfully achieved and demonstrated the ability of academia and industry to work together at a scale rarely achieved before. The goal of the prototype was to test the new technology (procedures and algorithms) and tailor it so it could be moved to production later (at the end of the project). Building the prototype included building a smartphone client and a backend system capable of collecting the smartphone data and capable of processing the data in real-time.

One of the important requirements at the launch of the system was to bootstrap it and to have operational traffic maps displayed on the phone from day one. This included working together with NavTeq, who contributed by giving their maps and traffic pattern product that contained historical data for the secondary network (arterial). Launching the system included numerous tests and steps, which were completed as part of the project.

Mobile Millennium was the first traffic application launched in North America by Nokia at the scale of

the nation. While it was only available to the driving public in California, the system was designed to cover the entire US and was beta tested for the entire US. In the pre-app store era, launching a traffic application of this nature was very challenging. It included doing media outreach, building a website for downloading the application, creating versions of the application for several platforms (Symbian Nokia phones and Blackberry phones), and having a specific customer support service (web-based and forum based).

One of the most difficult goals of the project was to create sustained engagement of the users. At the time, a relatively small number of apps were available for smartphones, so user behavior with respect to apps was relatively unknown. The project enabled the team to understand user behavior, i.e. the conditions under which people continue using apps. This was one of the goals of the field operational test.

Field Experiments

Numerous field experiments were conducted to test the Mobile Millennium system.

The initial Mobile Millennium Century experiments in the East Bay near the university were followed by the first true Mobile Millennium experiment in Manhattan during the ITS World Congress in 2008. Attendees were able to watch a fleet of rental cars negotiating a loop in Manhattan and see traffic mapped live. However San Francisco was the focus of the program on a consumer scale. Other experiments included:

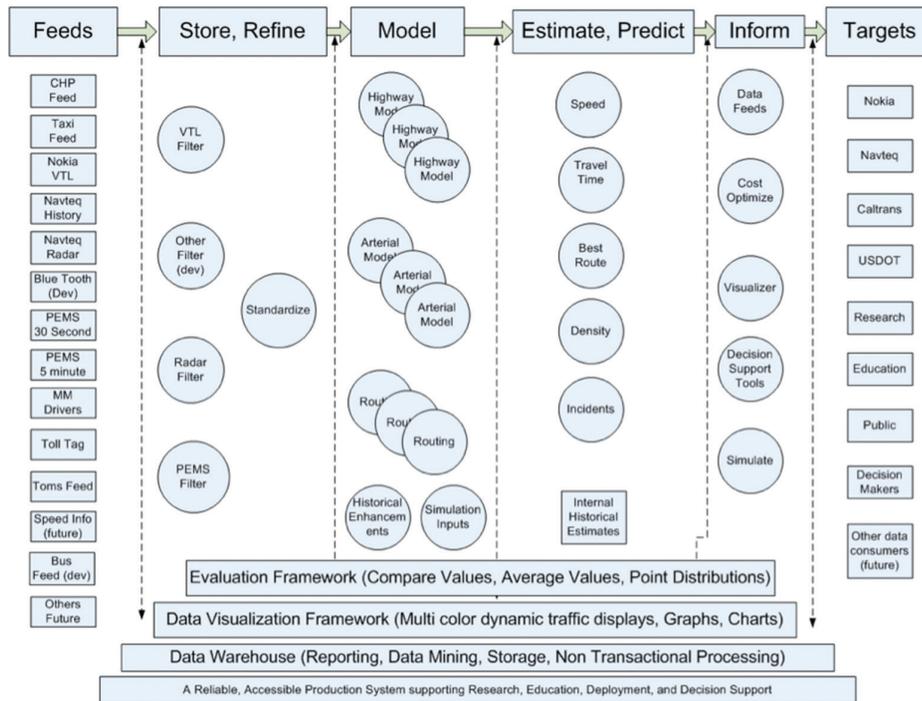


Figure 4: An overview of the Mobile Millennium system

Three experiments in June/July 2008 in Berkeley, CA.

These experiments were run on a small set of arterial roads in Berkeley as preparation for the Manhattan experiment that was run later in 2008. The primary question of interest to be answered by this set of experiments was if the GPS in the phones could provide good enough data on arterials. Additionally, the experiment was designed to specifically study the travel patterns through one of Berkeley's busiest intersections at San Pablo Avenue and University Avenue. Each experiment

Significance of the Research

The main contribution for the project was to demonstrate that with this new technology, it was possible to create an information system capable of relying principally on probe data. Because this was an open problem at the time, the secondary objective of the project was to demonstrate that with the amount of probe data available from the system (or more generally from industry), one could complement the already existing infrastructure of the Departments of Transportation to

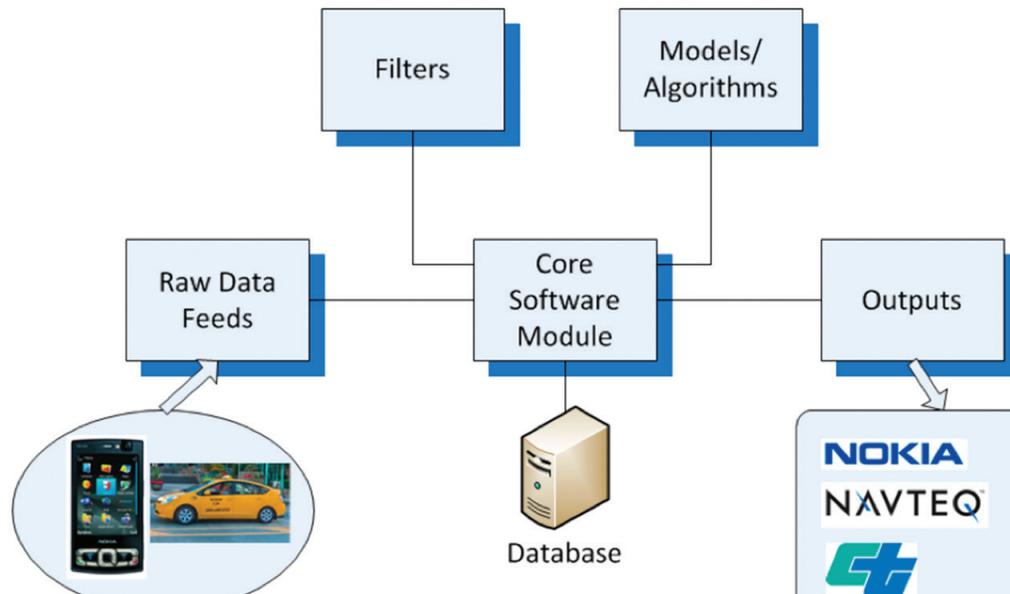


Figure 5: Mobile Millennium system database-centric architecture.

involved 20 vehicles driving several loops through Berkeley, with all 20 vehicles traversing the segment of University Avenue going west through the San Pablo Avenue intersection. Each test lasted approximately two hours.

Three experiments April 27-29, 2010 in the East Bay, CA.

Along with the following set of experiments, this set was designed to be the official test of the production version of the Mobile Millennium arterial model. These experiments each had 20 drivers (although data from some of the vehicles was never captured). The goal of this set of experiments was to estimate travel times along a 2.3-mile stretch of San Pablo Avenue that went through Berkeley, Albany, and El Cerrito, CA during a three-hour time period.

Three experiments June 29-July 1, 2010 in San Francisco, CA.

These experiments were the other half of the official Mobile Millennium evaluation. The goal of this set of experiments was to estimate travel times on a 1.1-mile stretch of 3 parallel roads (Van Ness Avenue, Franklin Street, and Gough Street) over a three-hour time period. Twenty vehicles were used in each experiment, split into two loops.

provide improved travel information services, which would otherwise be of poor quality or non-existing by using static infrastructure only. The project thus focused on establishing such evidence, by building a prototype system, and tackling two problems: generating traffic information from probes only; and generating traffic information from probes and loops (and other sources of data as well), to demonstrate that it was realistic for the technology to be used and operated by practitioners and in particular by the US DOT and the California DOT.

One of the other important objectives of this research and work was to perform technology exploration. One of the missions of the University is to assist the California DOT (and to a certain extent US DOT) in understanding how technology can impact the field of transportation in significant ways. This project, having roots traceable back to 2007 was started at a time when the mobile Internet, and smartphone technology was at its infancy.

At the time, it was unclear how this technology could benefit the California DOT and the US DOT. In fact,



Figure 6: Ph.D. students helping early users to download the app on their phone on the launch of Mobile Millennium

several failed attempts to use cellular phone data to provide travel information had raised doubts in the practitioner community that cell phones were ever going to be a viable source of information for traveler information. Part of this work for the California DOT and the US DOT was specifically to demonstrate that the team could prototype a system giving evidence of such claims. Thus, the work can also be viewed as a form of active technology scouting in which the scouting activity consists of constructing a prototype system to make such an assessment.

One of the questions that arose immediately when working on the Mobile Millennium project was the problem of data fusion. At the infancy of the mobile Internet, one question needed to be answered: with a potentially new source of ubiquitous data (mobile phone data), how would one operate data fusion at a global scale in an efficient manner and how would one manage to use this data efficiently in order to modernize traffic information systems. This question is far from being answered today. However Mobile Millennium provided a first step in this direction, which was considered to be a significant breakthrough by the community.

This project started at a time when traffic information systems had just undergone a first mutation, i.e. moving from operational center owned and operated by the government and available at best by phone, to being web based. The 511 website was an example of such a system, soon followed by the emergence of numerous

web based traffic information systems. In 2008, traffic information systems underwent a second mutation: they started to become part of the suite of mobile applications available through connected networks. Their importance is critical as evidenced by Apple's recent troubles with its new mapping application. They also started to integrate numerous additional and new sources of data, the most predominant one being probe data.

These sets of information raised questions when designing next generation traffic information systems: how to integrate new sources of data; by which process does one fuse heterogeneous sources of data; how to diversify the broadcast mechanisms for traffic information (phone, data based, sound based, etc.); how to handle compatibility with recent driver distraction laws etc. The work performed in the Mobile Millennium project provided the government with numerous answers to these questions, which now move the California DOT and the US DOT one level further: with sufficient understanding of these new technologies to be able to design such a system inside the California DOT and the US DOT.

Pioneering Traffic Solutions

Mobile Millennium enabled the research community, industry partners, and government to answer several questions of interest to transportation engineering involving the difficulties of public agencies to instrument infrastructure systems at a global scale. While California had been very fortunate to have significant instrumentation (in particular the loop detector systems deployed on the major freeways of California), it was

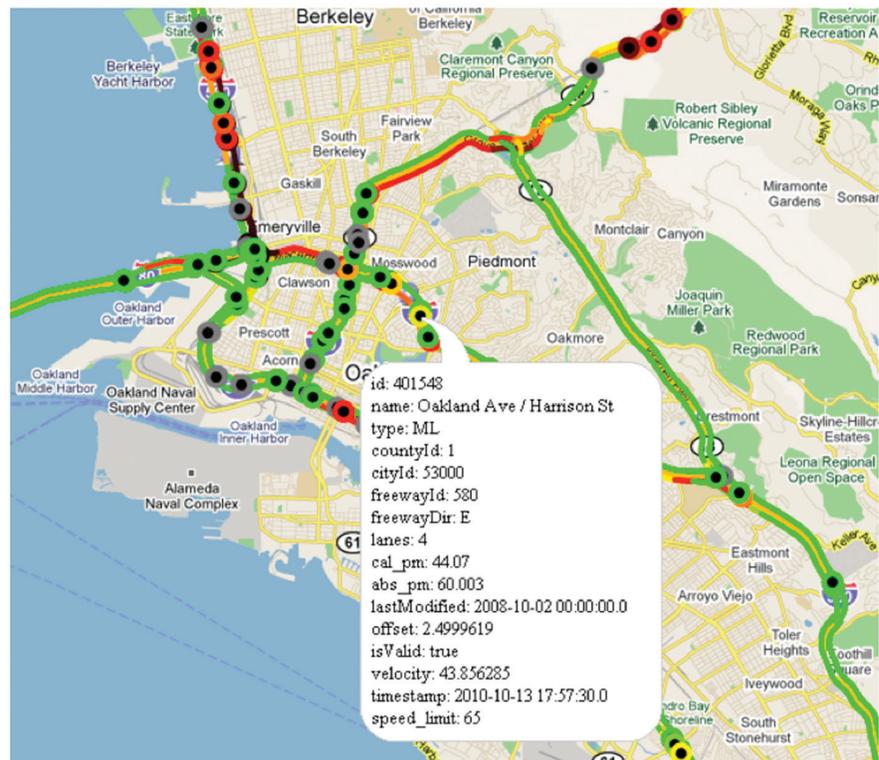


Figure 7: Mobile Millennium internal visualizer showing model outputs and locations of PeMS loop detectors (hollow circles).

clear from the start of the project (and it is still clear in the current financial situation of California) that the secondary network (arterial roads) would not be instrumented in the near future. Furthermore, the aging loop detector system suffered significant performance issues, and the natural question arose of knowing what could be done using probe data for traffic. Could smartphones provide fuller accurate traffic data at low cost?

This project took off at the very early stages of the smartphone era, at a time when the iPhone app store was just starting. As of the time this project concluded (in 2011), the iPhone app store contained more than 200,000 apps. When the Mobile Millennium project was started, the supremacy of the iPhone had not been established yet, and the Android platform did not exist. Another contribution from Mobile Millennium was the pioneering of the era of crowd-sourcing and participatory sensing. Mobile Millennium was one of the first systems in history to tackle crowd-sourcing in the context of traffic at large scale, and its success provided the foundations for many innovative traffic programs today.

Connected Corridors and Beyond...

In the specific context of traffic, it is now clear that at least for the coming years, no single entity will have enough data to sustain reliable traffic information at a global scale. The Mobile Millennium team realized this very early on, and at the time of the Mobile Millennium system integrated dozens of feeds including industry feeds (NAVTEQ, Telenav), public feeds (CHP, cab spotting, PeMS) and our own feeds (VTL / Nokia). The era of

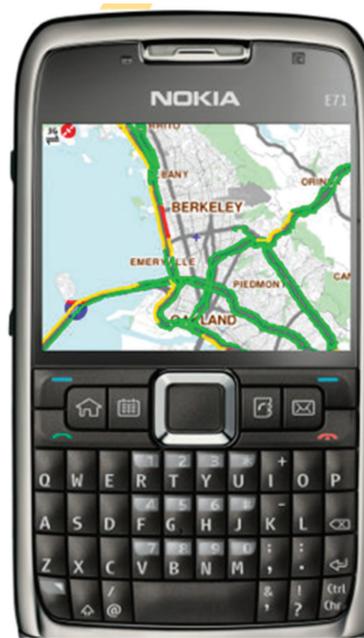


Figure 8: The Mobile Millennium phone client running on a Nokia E71. The phone client displays a color map of current traffic conditions around the driver's location while simultaneously providing VTL data to a central server.

data feed fusion has just begun, and numerous tasks still need to be completed before this can converge to practical solutions applicable to DOTs, industry and academia.

The Mobile Millennium project spawned a further set of projects, known as Data Fusion or Task Orders 1 & 2, which researched how to contract for commercial data feeds, what feeds to buy, how to integrate existing public data derived from Caltrans' Performance Measurement System (PeMS) with purchased traffic data, rationalize and filter it, and use it for accurate traffic state estimations. This project is nearing completion and will be the first rigorous evaluation of the capabilities of the probe and cellphone traffic monitoring systems developed during Mobile Millennium.

Simultaneously Mobile Millennium has joined with PATH's Tools for Operations and Planning (TOPL) team to develop a comprehensive and collaborative framework for managing traffic, transit, vehicles, and people in the most highly congested traffic corridors in the state. It is called Connected Corridors, and while centered at Berkeley for the systems, software and big data analytics, it is supported by the California DOT as well as a consortium of partners from agencies and vendors including Nokia, Renault, and Google. Built using probe data integrated with loop detectors, and designed to provide predictive suggestions about control system changes (like ramp metering or arterial traffic light timings) to traffic operators as congestion patterns are developing, rather than after the fact, it is an ambitious program. Essentially the program is exploring what it will take to create a professional and effective modern traffic management system for California's travelers in the next decades. It wouldn't be possible to do it without the breakthrough work of Mobile Millennium. ☼

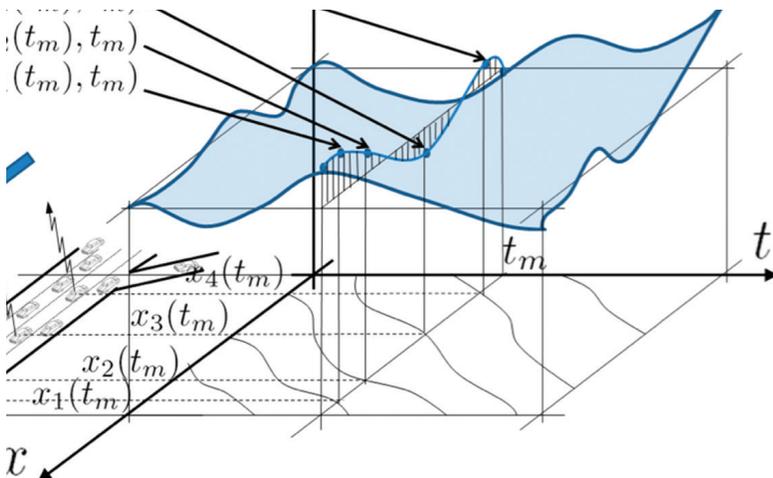


Figure 9: Illustration of the distributed velocity field $v(x, t)$ to be reconstructed from GPS samples. Four samples $v_i(x_i(t), t)$ are shown at $t = t_m$, from vehicles i transmitting their data (indicated by up-arrows above the vehicles).

Selected Mobile Millennium Research Achievements

A significant portion of the work of Mobile Millennium included developing new models to model traffic (both highways and arterials), as well as algorithms to integrate streaming data into the models (static and mobile data). This work generated a large number of papers and reports, both in academic and traffic engineering disciplines, but also in the consumer media where the possibilities of interactive traffic data like this captivated and fascinated the public.

Mathematical research contributions

Proof of existence and uniqueness of a Barron-Jensen Frankowska solution to a Hamilton-Jacobi PDE for a Cauchy problem. Extension of this proof for the case of internal boundary conditions. Derivation of an analytical solution to the problem in the case of piecewise affine conditions. This equation serves as a model for Lagrangian descriptions of traffic flow (i.e. from a mobile particle perspective).

C. Claudel and A. Bayen. Lax-Hopf based incorporation of internal boundary conditions into Hamilton-Jacobi equation. Part I: theory. *IEEE Transactions on Automatic Control*, 55(5):1142–1157, 2010. doi:10.1109/TAC.2010.2041976.

C. Claudel and A. Bayen. Lax-Hopf based incorporation of internal boundary conditions into Hamilton-Jacobi equation. Part II: Computational methods. *IEEE Transactions on Automatic Control*, 55(5):1158–1174, 2010. doi:10.1109/TAC.2010.2045439.

Derivation of a new PDE model for highway velocity evolution, called Lighthill-Whitham-Richards- v (LWR- v) model. Proof of equivalence of the new model with the classical LWR PDE, in the case of parabolic flux functions.

Derivation of a nonlinear discrete time dynamical system model for velocity that is compatible with the entropy solution of the LWR PDE, called Cell Transmission Model-Velocity (CTM- v). Extension to general networks using linear programming.

D. Work, S. Blandin, O. Tossavainen, B. Piccoli, and A. Bayen. A traffic model for velocity data assimilation. *Applied Research Mathematics eXpress (ARMX)*, 1:1–35, April 2010.

D. Work, O.-P. Tossavainen, S. Blandin, A. Bayen, T. Iwuchukwu, and K. Tracton. An ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile devices. In *47th IEEE Conference on Decision and Controls*, 2008 Cancun, Mexico, pages 5062–5068, 2008.

D. Work, O.P. Tossavainen, Q. Jacobson, and A. Bayen. Lagrangian Sensing: Distributed traffic estimation with mobile devices. *American Control Conference*, 2009, AC '09 pages 1536-1543.

D. B. Work, S. Blandin, O.-P. Tossavainen, B. Piccoli, and A. M. Bayen. A traffic model for velocity data assimilation. pages 2010(1):1–35, 2010. *Applied Mathematics Research eXpress*.

Second order traffic models based on the Colombo model, and development of the corresponding numerical schemes (Godunov scheme). This is a generalization of first order models, which enables us to characterize the variability of traffic in the congested regime.

S. Blandin, D. Work, P. Goatin, B. Piccoli, and A. Bayen. A general phase transition model for vehicular traffic. *SIAM J. Appl. Math.*, 71(1), 107–127. 2009.

Derivation of convex optimization based frameworks for data assimilation using the Hamilton-Jacobi equation, to integrate mobile data into robust estimation problems. Applications to problems of filtering (sensor placement problems), and computation of travel time (robust travel time estimation problems).

Design of a controller for exponential stability of switched linear hyperbolic PDEs.

C. G. Claudel and A. M. Bayen. Solutions to switched Hamilton-Jacobi equations and conservation laws using hybrid components. In M. Egerstedt and B. Mishra, editors,

Hybrid Systems: Computation and Control, number 4981 in *Lecture Notes in Computer Science*, pages 101–115. Springer Verlag, Saint Louis, MO, April 2008.

C. G. Claudel and A. M. Bayen. Convex formulations of data assimilation problems for a class of Hamilton-Jacobi equations. *SIAM J. Control Optim.*, 49(2), 383–402. 2011.

Derivation of new statistical models of traffic in arterial networks, based on hydrodynamic theory. Integration of signal timing into the models.

R. Herring, A. Hofleitner, P. Abbeel, and A. Bayen. Estimating arterial traffic conditions using sparse probe data. In *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems*, Madeira, Portugal, September 2010.

R. Herring, A. Hofleitner, S. Amin, T. Abou Nasr, A. Abdel Khalek, P. Abbeel, and A. Bayen. Using mobile phones to forecast arterial traffic through statistical learning. In *Proceedings of the 89th Annual Meeting of the Transportation Research Board Washington D.C.*, 2010.

Derivation of new measurement models for Lagrangian data into arterial traffic models, using hydrodynamic flow theory.

A. Hofleitner, R. Herring, and A. Bayen. Arterial travel time forecast with streaming data: a hybrid flow model - machine learning approach. *Transportation Research Part B: Methodological Volume 46, Issue 9*, November 2012, Pages 1097–1122.

A. Hofleitner, R. Herring, and A. Bayen. A hydrodynamic theory based statistical model of arterial traffic. *UCB-ITS-CWP-2011-2* (January 2011).

Intellimotion

Keeping up with California PATH Research in Intelligent Transportation Systems

Design and Layout

Bill Stone

For more information or to offer comments about this newsletter, please write, call, fax or e-mail:

California PATH

2105 Bancroft Ave Suite 300
Berkeley, CA 94720

Tel: 510-642-4522 FAX: 510-642-0910

e-mail: path@path.berkeley.edu

<http://www.path.berkeley.edu>



Partners for Advanced Transportation technology

Directors

Thomas West and Roberto Horowitz

Primary funding provided by:



Member



Intelligent Transportation Society of California

©2012 UC Regents. All rights reserved. Unless permission is granted, this material shall not be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise.

ISSN-1061-4311