Calling all cars: cell phone networks and the future of traffic

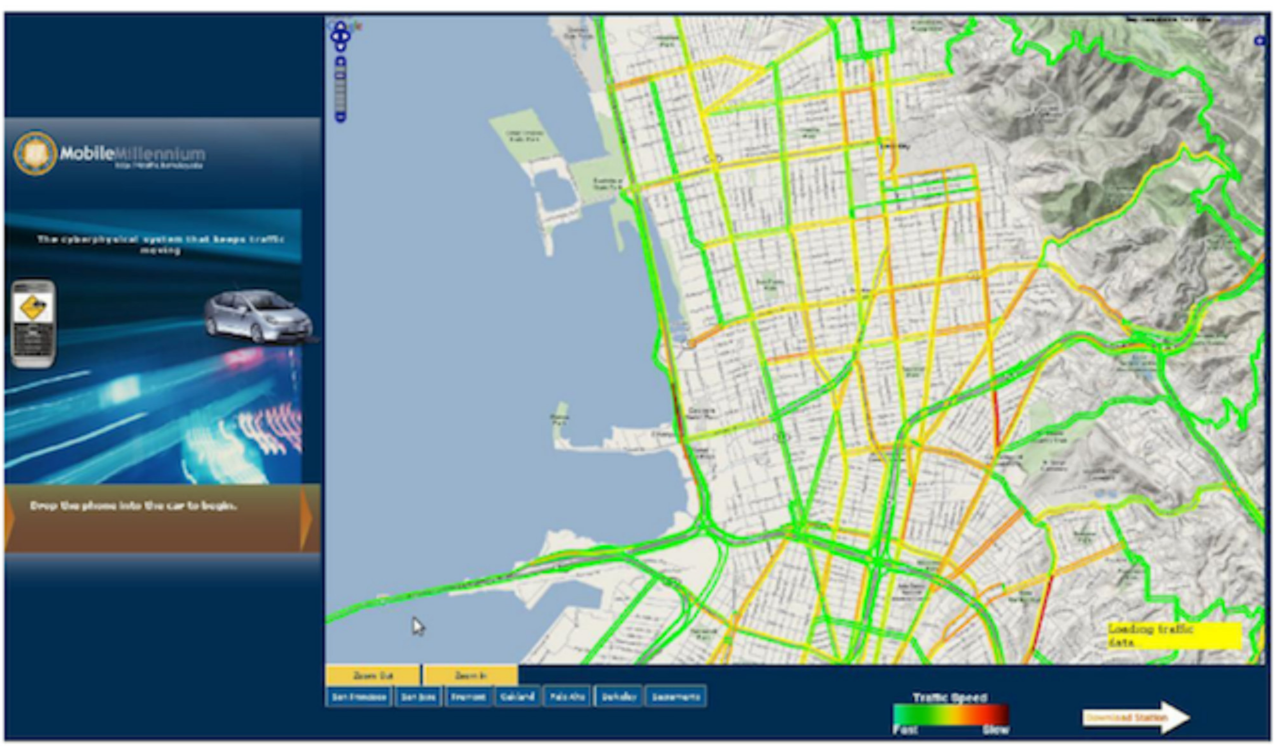
By Haomiao Huang | Last updated 6 days ago

Ask someone what they think the future of driving is, and the most likely response involves self-driving cars. And it's true that sensing and autonomy are dramatically changing the modern car, but there's another information revolution taking place outside the windows. Cheap sensors and network availability are not only making individual cars smarter, but they're also boosting the brainpower the environment cars drive in.

Networks of sensors connected by the Web are making it possible to monitor traffic, parking availability, air pollution, road quality, and more in real time and across large distances. Traffic monitoring in particular has been revolutionized by these changes. This kind of data gives drivers real-time travel time predictions, makes it possible to create smart roads where tolls and signals can adapt to changing conditions, and provides urban planners with accurate and detailed pictures of traffic usage and its effects, improving city layout and planning for the future.

One of the most widespread and powerful sensors is the mobile phone. Equipped with GPS and connected to the Internet, modern smartphones are an important source of information that many companies use to provide traffic data. Google Maps, for example, makes extensive use of data collected from users on mobile phones.

Mobile Millennium was one of the first large-scale phone-based traffic monitoring projects in the US. An ongoing pilot project run by Nokia, NAVTEQ, and UC Berkeley started in 2007, its goal is to develop and demonstrate technologies needed for large-scale data collection for traffic monitoring. The project combines data from a smartphone app distributed to the public and traditional traffic sensors to provide accurate real-time monitoring of traffic conditions in the San Francisco Bay Area.



The Mobile Millennium traffic monitoring software. *Source: UC Berkeley.*

Designing and running these sensor networks is no trivial task. Data is flooding in from many different sources in many different places, and useful data has to be separated from noise. Algorithms and models are needed to fuse the incoming data into a comprehensible whole, and protecting individual privacy is also a major challenge. Yet the potential gains are huge, so there is an unceasing demand for more and better data.

In this article, Ars goes behind the scenes at Mobile Millennium to examine the technology behind a distributed sensor network. We'll look at how the system is designed to protect user privacy, examine how data from thousands of mobile phones and hundreds of static sensors are combined to measure traffic flow, and look at how this technology will impact the future of driving.

An intelligent highway

The most obvious use of traffic data is to give drivers options for reducing the effects of traffic jams and accidents, either by taking alternate routes or simply by changing their travel times. Trip-planning software already can use traffic speed information to minimize travel time or fuel usage over a trip, and future hybrids and electric vehicles might use traffic predictions to help the onboard computer optimize battery usage.

This kind of real-time data also lets civil engineers create traffic control schemes that react intelligently; for example, smart signals could eliminate the need to wait at red lights and empty intersections. Larger scale efforts might involve roads that actively change direction in response to changing traffic flows.

The data is also of more than just immediate importance. Having good data on current traffic and road usage is vital to predicting future patterns of traffic, which is important for planning purposes. Ars recently explored the issues related to congestion pricing, one of the most popular tools for alleviating congestion. Congesting pricing uses dynamic tolls that are adjusted according to road usage to try to reduce traffic during peak conditions. The success of such schemes is heavily dependent on being able to measure the effects of pricing changes on driving patterns.

Finally, accurately measuring traffic is also useful beyond the immediate realm of driving. Cars and roads have a huge impact on our societies, and traffic has many secondary effects. For example, traffic is a major source of potentially harmful noise, and the generation of noise maps of the city is one of several projects underway that piggybacks on the Mobile Millennium data and network. By correlating noise patterns to population maps, it's possible to assess the impact of noise on the city's people. Cars are also a major source of air pollution, and traffic data can be correlated and combined with measurements taken by pollution sensors to build a map of pollutants produced by cars around the city.

Going mobile

For a long time, traffic sensing was mostly reliant on static sensors. Inductive loop detectors—metal rings embedded in the road—detect the metal in cars that pass over them. Traffic cameras are another common sensor type, and the RFID tags used for electronic toll payment can be tracked to provide data.

These types of sensors are generally pretty accurate, but fixed infrastructure is expensive to deploy and operate. When these sensors break, they're also expensive to repair and replace, so they're typically placed at key places like intersections and highway on- and off-ramps. This means that when traffic conditions change, like when an accident occurs, the changes aren't detected until their effects propagate back through the traffic flow upstream to a sensor.

The need for ever more data from ever more widespread sensors has meant that going to mobile sensors is a necessity, and mobile phones are an obvious choice. It's an oft-quoted statistic that worldwide, there are more cell phones in use than toothbrushes. And an ever-growing fraction of those phones, especially in the US, are smartphones, equipped with GPS and Internet connectivity.

Calling all cars: cell phone networks and the future of traffic

By Haomiao Huang | Last updated 6 days ago

Mobile Millennium was one of the first large-scale projects to take advantage of these developments for traffic monitoring. Professor Alexandre Bayen is the principle investigator on the Mobile Millennium project.

"This was back in 2007, and at the time we were trying to do traffic estimation using these aftermarket GPS units that you put on your dashboard," he told Ars. "Right around this time, Nokia put out some of the first phones with GPS—this was before the iPhone—and it became obvious that with [Internet] connectivity and GPS and the explosion of the cell market that this was a way more cost-effective way to get information."

The rise of phones with GPS is the crucial element. Using cell phone signals to measure traffic flow had been attempted before, but earlier experiments ran into the problem that cell tower triangulation just isn't very accurate. Triangulation also requires direct access to cell towers, which would be expensive and difficult to negotiate with service providers.

Built-in GPS provides accurate data, and the net connection gives a simple way of collecting the data that doesn't require special access to the cell network's infrastructure. As a bonus, it also provides a way to give drivers an incentive to participate. Accurate real-time traffic information can be sent right back to be displayed in the same app used to collect data.

A partnership to explore these possibilities was set up between Nokia, NAVTEQ, and UC Berkeley, with funding from the California Department of Transportation. Nokia provided phones for initial testing as well as the technology to gather the data, while NAVTEQ provided the mapping information needed to match collected measurements to roads. Berkeley would be responsible for developing data fusion techniques to make sense of the incoming measurements.

The group had to address several interrelated technical challenges associated with using cell phones as sensors this way. First, information collection had to be set up in a way that would preserve the privacy of the users, so that individual cars could not be tracked using the data gathered. The server architecture had to be designed and set up to support this. Then, theory and algorithms had to be worked out for how to make sense of the incoming data and aggregate the measurements into a unified picture of the state of traffic in the city.

Gathering data, but privately

User privacy was an overriding concern from the beginning of the project. The project's leaders knew that users would only be willing to contribute to the system if their information was protected, and this factor dictated the entire structure of the system. How the data was to be gathered would heavily influence both the hardware infrastructure and the algorithms used to process the data.

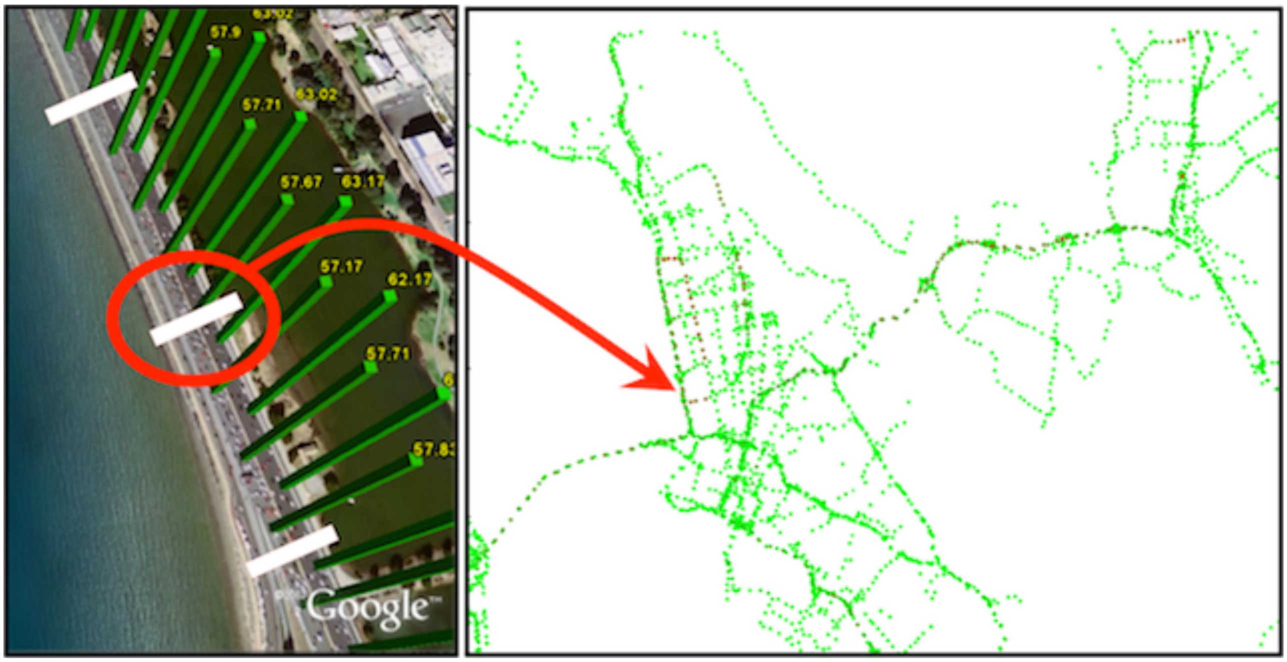
Maintaining user privacy meant meeting two main needs: preventing, as much as possible, the path of a single vehicle to be reconstructed over time, and separating the identification of the phones from the measurements.

Anonymity was, in some ways, the easy part. Data sent from phones is tagged so that the service provider knows where to send the bill. This data needs to be anonymized before processing, and this requires passing it through two sets of servers.

When a phone takes a measurement, it creates a data packet containing its position, speed, and anything else that might be of interest. This packet is encrypted using the public key of the data processing server and sent off, but instead of going straight to that server, it goes to a proxy server that strips the packet of any identifying information. Only then is the packet passed on to a VTL server that processes it and passes it on to the data aggregation servers.

Reading the contents of the packet requires a decryption key. The proxy doesn't have the private key needed to perform the decryption, so while it knows the identity of the phone, it doesn't know where the data comes from. The packets that arrive at the VTL server have no identifying information. There isn't a single machine that can be compromised to provide position and speed information that can be attached to a particular phone.

Preventing paths from being reconstructed was more involved, and required the use of virtual trip lines (VTLs), something that Nokia came up with for this purpose. Instead of constantly reporting location and speed, each phone checks its current location against a downloaded database of VTL positions, and measurements are only sent when the phone crosses a VTL location. This drastically reduces the amount of data collected from any one phone, lessening the likelihood that someone could reconstruct individuals' paths from the data.



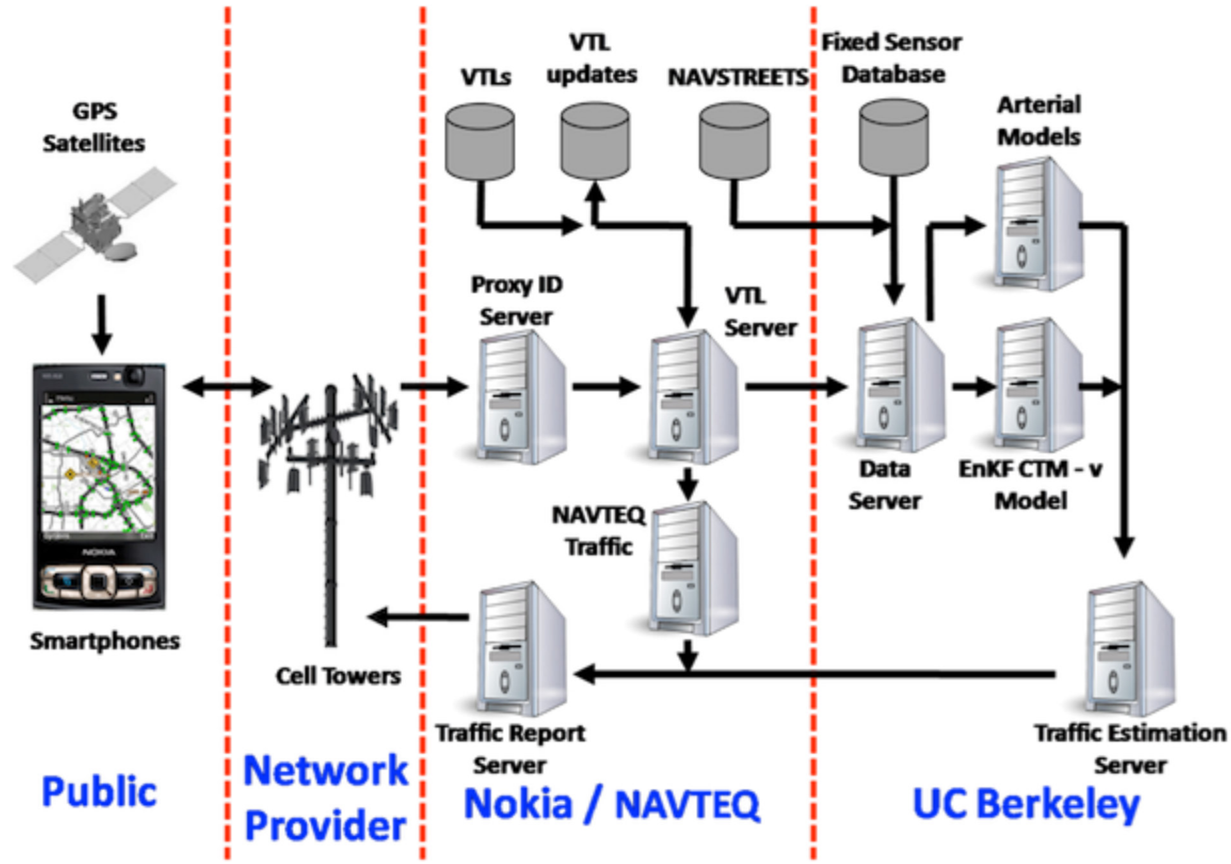
Data is only collected at virtual trip lines placed around the city, helping to maintain user privacy. *Source: UC Berkeley.*

This still leaves open the possibility that a sequence of measurements can be processed to build up a trajectory, so Nokia came up with an algorithm for placing the virtual trip lines in order to minimize the probability that two measurements from consecutive VTLs could be linked to the same vehicle.

Matching up measurements means taking a reading from one VTL and correctly associating it with another reading taken at the next VTL down the road. The more measurements there are from the next VTL that could match the first, the harder it is to say for sure which ones belong together. The algorithm uses the number of cars on the road and their speeds to figure out the best spacing to maximize the number of cars that might match going through any given VTL pair. In addition, the server that decides where to put the VTLs is separated from the one that actually processed the incoming data, to make it less likely that an attacker could manipulate the VTL placement to try to make it easier to track a car.

Finally, another layer of protection comes from randomizing measurements. Instead of transmitting when crossing every VTL, the phones perform a virtual coin flip each time to decide whether to transmit. This makes it much harder for someone looking at the data afterward to reconstruct individual trajectories.

The final architecture is illustrated below, showing the multi-layered server architecture. These precautions aren't foolproof, especially in an extreme case like a single car driving down an empty road at night, but they provide a pretty stiff layer of protection.



The final architecture for gathering and processing data. *Source: UC Berkeley.*

Calling all cars: cell phone networks and the future of traffic

By Haomiao Huang | Last updated 6 days ago

Making sense of it all

Coming up with the algorithms for data fusion was the main task of the researchers at UC Berkeley. In addition to the GPS measurements from the phones, the system incorporates GPS data from buses, taxis, and cars from other dedicated fleets. Data from all of the static sensors in the region, such as loop detectors and RFID tag readers, are also included. The question that the data fusion algorithms are trying to answer is this: given all of the measurements being gathered and a stretch of road of interest, what is the best estimate of the number of cars on that road, and how fast they're going?

GPS tracks in general are hard to process for traffic monitoring, and there were a number of problems to overcome. One of the first issues was to just figure out what road the measurements were coming from. "You had to create a fully integrated geo-localizing system to fuse the data," Bayen noted. "You need the underlying road network on which you map measurements to."

NAVTEQ's mapping information was vital, but even so there was a significant amount of post-processing that had to be done. "The maps aren't perfect, you have roads that lead to nowhere, that kind of thing," Bayen said. In fact, one of the side benefits of the Mobile Millennium data was that GPS measurements collected for traffic monitoring also improved the map data by revealing and filling in gaps.

Even with complete maps, matching measurements to a road can be a struggle. People may be walking alongside the road, or they may park and forget to turn their GPS program off. In urban canyons, like downtown San Francisco, most of the received GPS data points do not match up exactly to known roads, because the buildings obscure satellites. The measurements have to be associated with particular roads using machine-learning methods. These methods try to find the most likely road for a particular data point, and reject data points that aren't likely to be actual cars moving on the road.

The biggest challenge, and one that is ongoing, is using the measurements with mathematical models of traffic flow to estimate and predict traffic that isn't directly measured. Sensors only give a partial picture of the world at the time and place where a measurement is taken. "There's no way you can have sensors everywhere all the time," Bayen points out. "Look at Google; they have the most data of anyone, and even they don't have enough to cover the secondary network."

Models of the physical world are needed to relate those measurements to the rest of the world. The problem is that existing models aren't well equipped to integrate the kind of data mobile phones provide. "The integration of mobile data into physical models is difficult, from a scientific perspective," Bayen says. "There's no completed theory for it."

Unlike traditional static sensors, instead of measuring all of the cars that pass a particular location, a GPS measurement gives a single measurement for a single car, and this is hard to deal with. To understand why this is a problem, we need to look at how traffic flow is modeled.

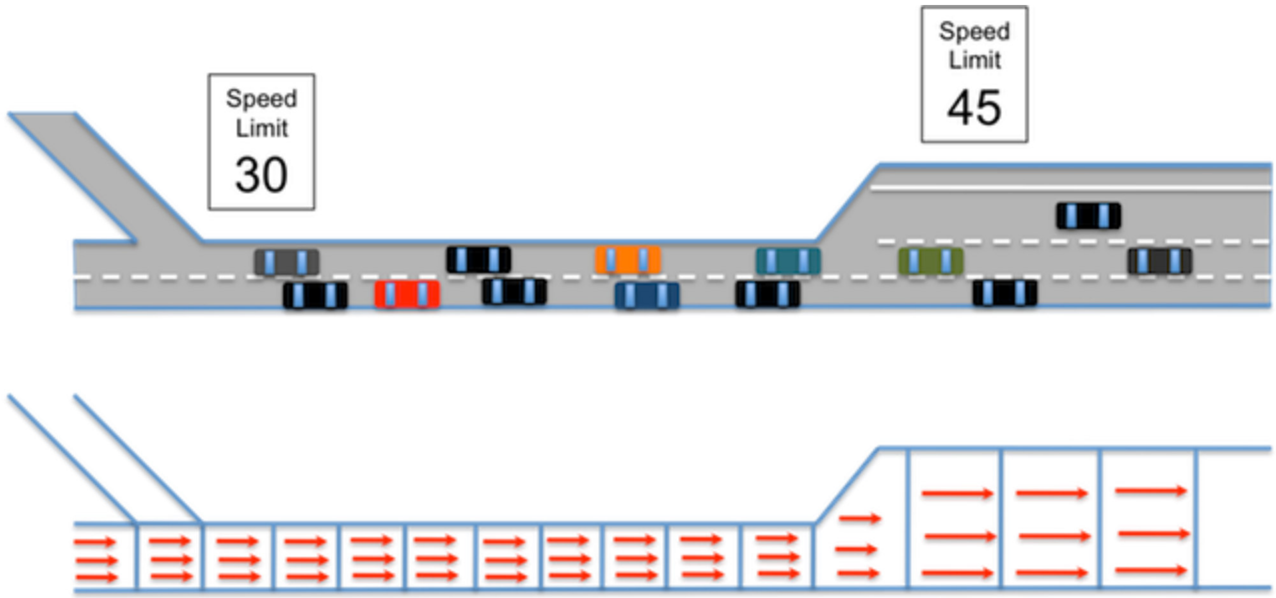
The flow of traffic

The obvious thing to do to model cars on a roadway is to track the movements of each car individually. This is important in some applications, but the computational resources needed to track thousands of cars, and the spatial relationships between them, quickly get expensive.

To get around this limitation, traffic researchers often treat the movement of cars as liquid flowing through, well, a series of tubes. Each segment of tube is a portion of road, and instead of having to track many individual cars, the number and speed of cars on that road is represented by the density and velocity of the liquid. By using a specialized set of equations similar to those that govern the flow of air or water, the properties of traffic flowing along a road can be modeled and computed.

The equations that govern fluid flow come from conservation relations. The basic idea is straightforward: given a volume of space and some fluid flowing through it, the amount of fluid in that space at a given time is whatever was in there to begin with, plus the amount that goes in, and minus the amount that comes out.

To get a fine-grained picture of fluids flowing through our road network, we break down the network into a connected sequence of small volumes, where each volume is a cell that's connected to others. The flow properties in each cell affect the ones that neighbor it. And matching the outflow of each cell with the inflow of the next one down the line produces a system of equations that relate the flow properties over time in each cell to its neighbors.



Instead of counting individual cars, traffic is modeled as flow in a series of cells.

Two more pieces of information are needed to solve the equations for what we want. First, the boundary conditions have to be specified—that is, the values coming into the cells on the outside edges. In the case of traffic networks, that's usually the cars coming into and going out of the road area of interest.

The second requirement is to provide initial conditions: how much fluid starts out in each cell, and how fast it's going. Once these pieces of information are provided, we can solve the equations in sequence and over time by integrating all of flow coming in and going out. The solutions give the fluid density and velocity at any given point in the network over time. Solving for fluid flow like this is known as computational fluid dynamics, and the same basic concept is used in many applications, for example, computing the flow of air over an airplane's wing or water around a ship's hull.

The fluid dynamics model of traffic flow works well with fixed sensors. Put sets of sensors at the start and end of a stretch of a road, and these give the boundary conditions for that bit of road. Cameras and satellites can provide initial conditions, and the flow density and velocity along that road can be calculated. These methods have been around for a while, and have proven to be pretty accurate within the limitations of the sensors.

This would be fine if the cars truly were a fluid flowing deterministically down a tube, but driver actions lead to little perturbations that cause slow-downs or even accidents. These disruptions can't be detected until their effects propagate down the line to a sensor, usually in the form of a traffic jam. Finer-grained spatial detail requires finer-grained placement of sensors, which is where the smartphones come in.

Using GPS measurements to augment sensors like traffic cams and loop detectors makes the entire system much more versatile. Unlike fixed sensors, the virtual trip lines can be moved and added to as needed, perhaps to get more measurements on roads where the state of traffic is changing rapidly.

However, though virtual sensors can be placed much more densely than real ones, their measurements are less complete. A physical sensor placed to cover a road will count and measure the speed of every car that passes it. Even complete GPS trajectories from vehicles being tracked without privacy considerations give only data for a single car, which must then be related to the cars around it. Virtual trip lines only generate measurements from cars with phones running the Mobile Millennium software, and even then only in accordance to the privacy-protecting randomization scheme. This makes the data fusion problem like trying to calculate the flow of a river given the properties of a few drops of water.

This means that the mobile phone measurements can't simply be fed into the system as additional boundary conditions. To use the data from the phones, the researchers and graduate students on the project had to develop new methods of solving the flow equations.

The team ultimately developed a number of different algorithms suitable for a variety of different models. The details of these algorithms are a little arcane, and are described in a number of papers that can be found on the Mobile Millennium website. Basically, the new methods allowed GPS measurements to be incorporated as special internal conditions for the flow to satisfy. Density and velocity aren't computed directly from boundary and initial conditions. Instead, the flow is calculated as the result of an optimization that finds the flow values that best match the measured data.

With these algorithms in place, the models can synthesize data from point sources. Measurements from loop detectors and cameras can be combined with GPS data from the phones, and also longer GPS trajectories from other sources, like buses. The resulting estimates of traffic flow are much better than those available from static sensing alone.

Calling all cars: cell phone networks and the future of traffic

By Haomiao Huang | Last updated 6 days ago

Mobile century

The initial design of the Mobile Millennium system culminated in a proof-of-concept test called Mobile Century, on February 8, 2008. 100 cars, each equipped with a Nokia smartphone running the GPS tracking software, were mixed in with traffic along a 10-mile stretch of Interstate 880 in Northern California. To get ground-truth data to compare against, the project team recorded data from fixed inductive loop detectors along the same stretch of road and also posted students with video cameras on overpasses.

The Mobile Century field experiment validated the technology behind Mobile Millennium and captured an accident in real time. *Source: UC Berkeley.*

The test ran for nearly 10 hours and required over 150 student drivers, and the results were a great success. Despite the fact that the Mobile Century cars accounted for no more than 2-5% of the cars on the highway at any time, the system was able to very accurately measure the speed and density of cars driving along the highway, and at a much higher spatial resolution than the fixed system of loop detectors. The day also provided a startling demonstration of the potential of the mobile phones in getting data quickly.

The traffic estimates calculated using the test data were being displayed in real-time at a control center specially set up for the occasion, and were being observed by members of the research team as well as various local and state transportation officials involved with the project.

At 10:50am, the team noticed that the route displayed a sudden growing red blotch in the middle of the test stretch of highway. The display, projected prominently on a wall, also showed the traffic data from the same stretch of road on Google Maps, which then drew its data primarily from the static loop detector sensors. Where the Mobile Century data showed traffic coming to a near standstill, the Google display remained completely clear.

"We were getting nervous," Professor Bayen recounts. "There were all these officials watching, and we thought maybe something had gone wrong." After some tense minutes of waiting, the researchers breathed a sigh of relief as the Google displayed slowly changed, first to yellow, then to red. Beepers began going off around the room as automated alerts were sent out to the visiting transportation officials. There had been a 5-car pile-up, exactly where the first slow-down had been reported by the Mobile Century system.

The accident was a clear and stunning vindication of the project. The sudden slow-down due to the accident had been detected and reported in less than a minute, well before its effects could propagate back through the chain of cars to a static detector upstream. The phone-based measurements had dramatically out-performed the fixed sensor network.

Till all are one

After the proof-of-concept demonstration, Mobile Millennium went live in November of 2008 as an operational test and has been running ever since. Though the software is no longer available for download, there are still around 5000 users with the Mobile Millennium software driving around the San Francisco Bay Area.

The concepts and technology demonstrated in Mobile Millennium are now widespread. Google's mobile Maps app does the same thing: fusing mobile GPS data with static sensors and other sources of information. Many companies that provide traffic monitoring data do something similar, either using phones or other dedicated mobile sources. A large number of cities all over the world are using methods similar those described here to combine static and mobile sensors to measure traffic patterns. Whether it's taxis and buses (or other sources) with GPS, cities and companies are snapping up data left and right.

The future of mobile sensing isn't limited to traffic monitoring either. The CarTel project at MIT demonstrated the use of accelerometers mounted on a fleet of a local limo company to automatically detect and map potholes. A machine-learning algorithm was taught to recognize the distinctive bump associated with driving over a pothole. Each time a pothole was detected it could be instantly reported and mapped.

Although this particular experiment used a custom sensor unit with accelerometers, it's not difficult to imagine that a similar system could be designed to take advantage of the accelerometers built into smartphones. The pothole detection was also based on detecting extremes in the measured roughness of the road. With a larger base of reporting sensors, it would be possible to build a constantly updated map of road conditions everywhere in a city. Data from this could be used to warn drivers of unsafe conditions or inform maintenance planning.

In the coming years, mobile sensing is going to transform the driving experience. As Ars noted, it's only a matter of time before our cars are fully networked and the traffic flow becomes all but self-aware. The tighter integration of phones and data networks with cars will make even more data available. The CarTel project has suggested that shared engine sensor information, for example, will allow owners to see if their car is deviating from the norm, possibly indicating a maintenance problem. Yet it's also obvious that as these technologies proliferate, privacy is going to be even more of a concern, and the data collection systems that are built will need robust privacy protections. One can only hope that the companies building such systems are as wary of the potential dangers as they are hopeful for the rewards.