
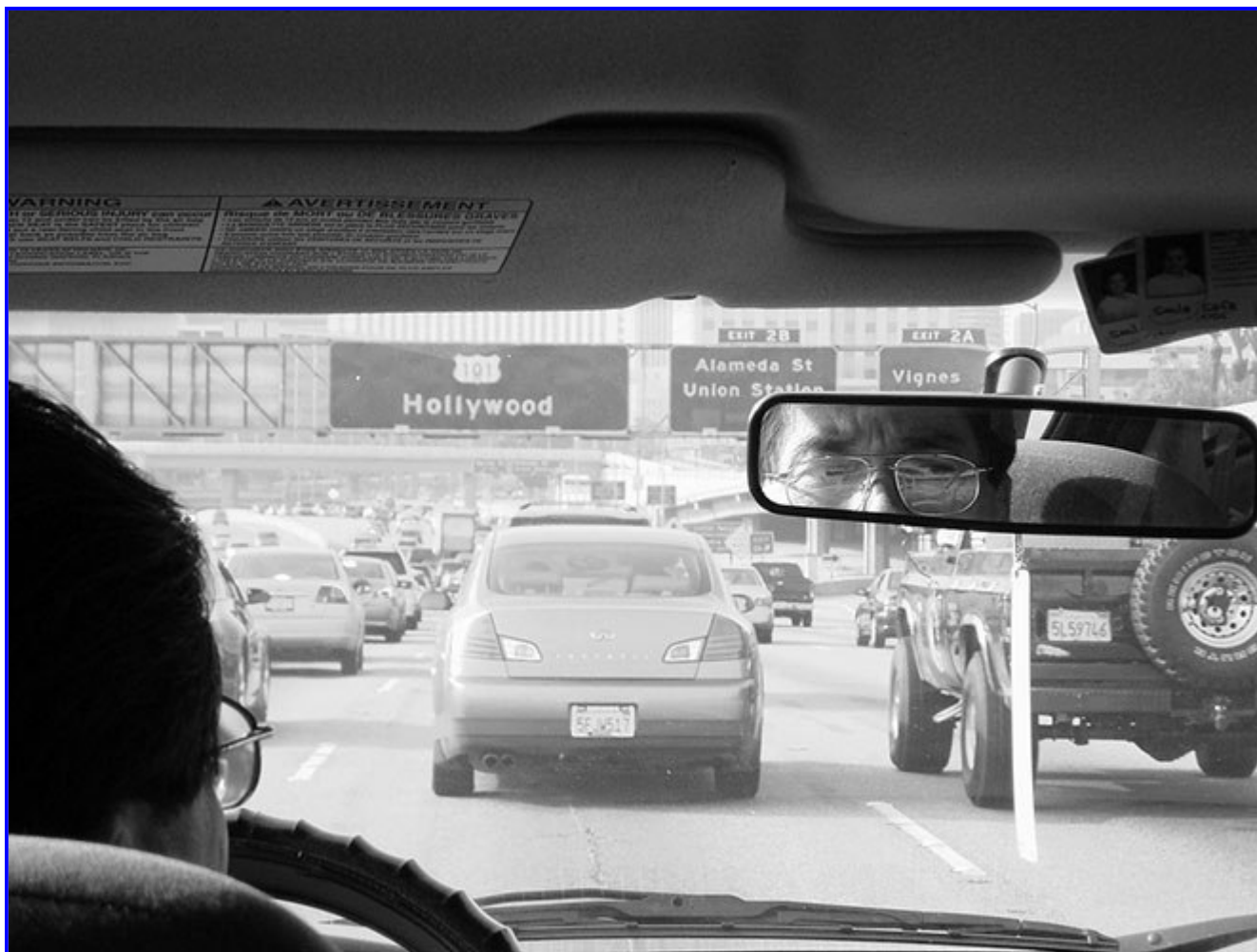


[Autopia](#)[Planes, Trains, Automobiles and the Future of Transportation](#)[Previous post](#)[Next post](#)

# Cellphone Networks and the Future of Traffic

By [Autopia Blog](#)  March 2, 2011 | 7:00 am | Categories: [Infrastructure](#)



Ask someone what they think the future of driving is and the most likely response is [autonomous cars](#). It's true sensing and autonomy are dramatically changing cars, but there's another information revolution afoot. Cheap sensors and network availability aren't just making cars smarter, they're boosting the brainpower of the environment cars drive in.

Networks of sensors connected by the web make it possible to monitor traffic, parking availability, air pollution, road quality and more in real time across vast distances. Traffic monitoring in particular has been revolutionized. This kind of data gives drivers real-time travel time predictions, fosters creation of smart roads where tolls and signals can adapt to changing conditions and provides urban planners with accurate pictures of traffic usage and its effects, improving planning.

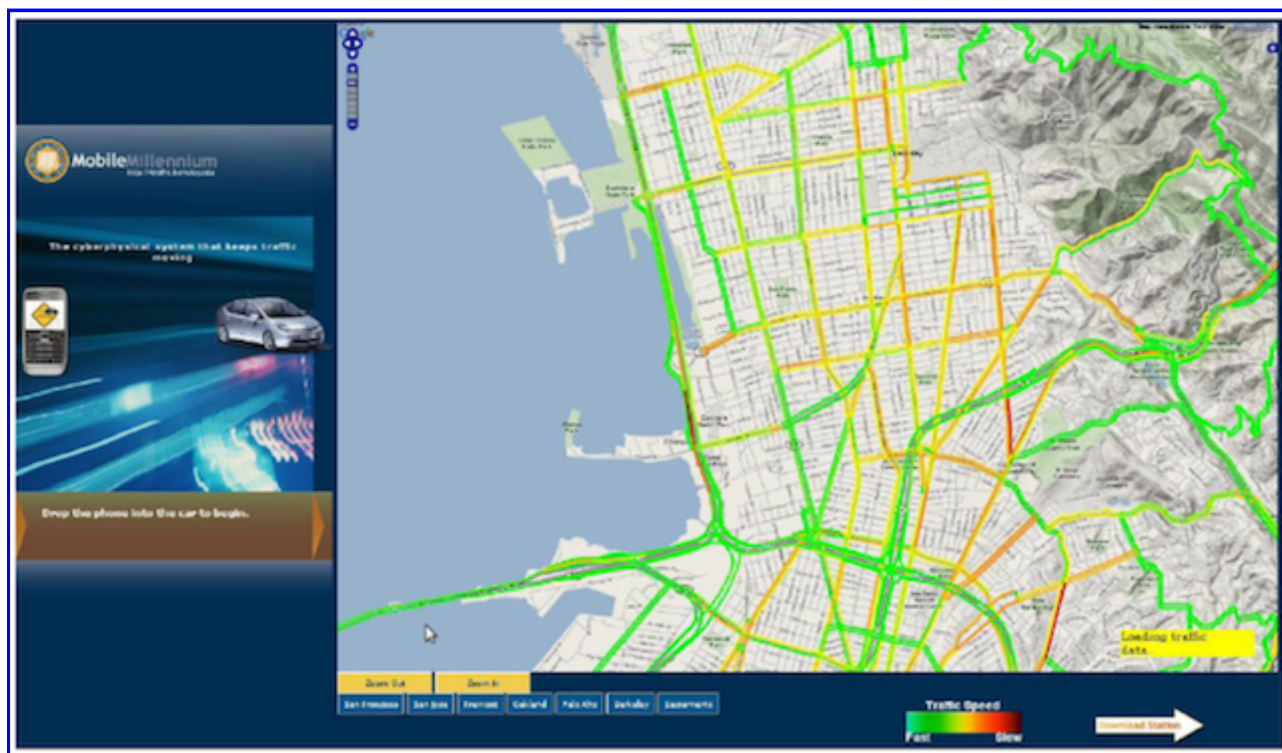


One of the most widespread and powerful sensors is the mobile phone. With their GPS and internet access, [smartphones](#) are an important source of information used to provide traffic data. Google Maps, for example, makes extensive use of data collected from users on mobile phones.

[Mobile Millennium](#) was among the first large-scale phone-based traffic monitoring projects in the United States. The project, launched by Nokia, NAVTEQ and UC Berkeley in 2007, is intended to develop and demonstrate technologies needed to allow large-scale data collection for traffic monitoring. The project combines data from a smartphone app and traditional traffic sensors to provide accurate real-time monitoring of traffic conditions in the San Francisco Bay Area.

Designing and running these sensor networks is no trivial task. Data floods in from many sources in many places, and useful data must 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, we go behind the scenes at Mobile Millennium to examine the technology behind a distributed sensor network. We look at how the system protects user privacy, examine how data from thousands of mobile phones and hundreds of static sensors are combined to measure traffic flow, and we'll look at how this technology will impact [the future of driving](#).



Mobile Millennium traffic monitoring software. *Image: UC Berkeley*

## An Intelligent Highway

The most obvious use of traffic data is providing drivers with 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 can use traffic speed information to minimize travel time or fuel usage, and [hybrids](#) and [electric](#)

[vehicles](#) might use the data to help 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 for red lights at empty intersections. Large-scale efforts might involve roads that actively change the direction of traffic in response to changing traffic flows.

The data is of more than immediate importance. Good data on road usage is vital to predicting future traffic patterns, which is important for planning purposes. [Congesting pricing](#), for example, uses dynamic tolls adjusted according to road usage in an effort to ease traffic at peak times. The success of such schemes depends heavily on being able to measure the effects of pricing changes on driving patterns.

Accurately measuring traffic also is useful beyond the immediate realm of driving. Cars and roads have a huge impact, and traffic has many secondary effects. It is a major source of noise, for example, and creating “noise maps” of the city is one project piggybacking on the Mobile Millennium data and network. By correlating noise patterns to population maps, it’s possible to assess the impact of noise on residents. Cars also are 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 vehicular pollutants around the city.

## Going Mobile

For a long time, traffic sensing relied largely on static sensors. Inductive loop detectors — metal rings embedded in the road — detect the metal in cars passing over them. Traffic cameras are another common tool, and RFID tags used for electronic toll payment can be tracked to provide still more data.

Such tools are generally accurate, but fixed [infrastructure](#) is expensive to deploy and operate. It’s also expensive to repair and replace, so these tools typically are installed at key places like intersections and on/off-ramps. But when traffic conditions change downstream — say, during an accident — those changes aren’t detected until the impact ripples upstream to the sensor.

The need for more data from more sensors has made mobility a necessity, and mobile phones are an obvious choice. It’s often said there are more cellphones than toothbrushes in the world, and a growing number of them are smartphones with GPS and internet connectivity. Mobile Millennium was among the first large-scale projects to take advantage of this development for traffic monitoring.

“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,” said Prof. Alexandre Bayen, the principle investigator on the Mobile Millennium project. “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 GPS-enabled phones was crucial. Using cellphone signals to measure traffic flow had been attempted before, but cell tower triangulation isn’t very accurate. It 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 provides a simple way of collecting it without special access to the cell network’s infrastructure. It also provides an incentive to drivers to participate — accurate real-time traffic info can be displayed in the same app used to collect data.

Nokia, NAVTEQ, and UC Berkeley teamed up to explore these possibilities with funding from the California Department of Transportation. Nokia provided phones for initial testing and the technology to gather the data. NAVTEQ provided the mapping information needed to match collected measurements to roads. The university developed data fusion techniques to make sense of it all.

The group had to address several interrelated technical challenges. First, information collection had to be done in such a way to preserve the privacy of the users so individual cars could not be tracked using the data gathered. The server architecture had to be designed and set up to do this. Then, theory and algorithms had to be developed to make sense of the incoming data and aggregate the measurements into a unified picture of the state of traffic.

## Gathering Data, Privately

[User privacy](#) was an overriding concern from the beginning. Project leaders knew users would participate only if their information was protected, and that dictated the 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 from being 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 the service provider knows where to send the bill. This data needs to be anonymized before processing; 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, but instead of going straight to that server, it goes to a proxy server that strips the packet of any identifying information. Then the packet is passed on to a virtual trip line (VTL) server that processes it and sends it 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 although 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 trickier and required the use of virtual trip lines (VTLs), something Nokia developed 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.



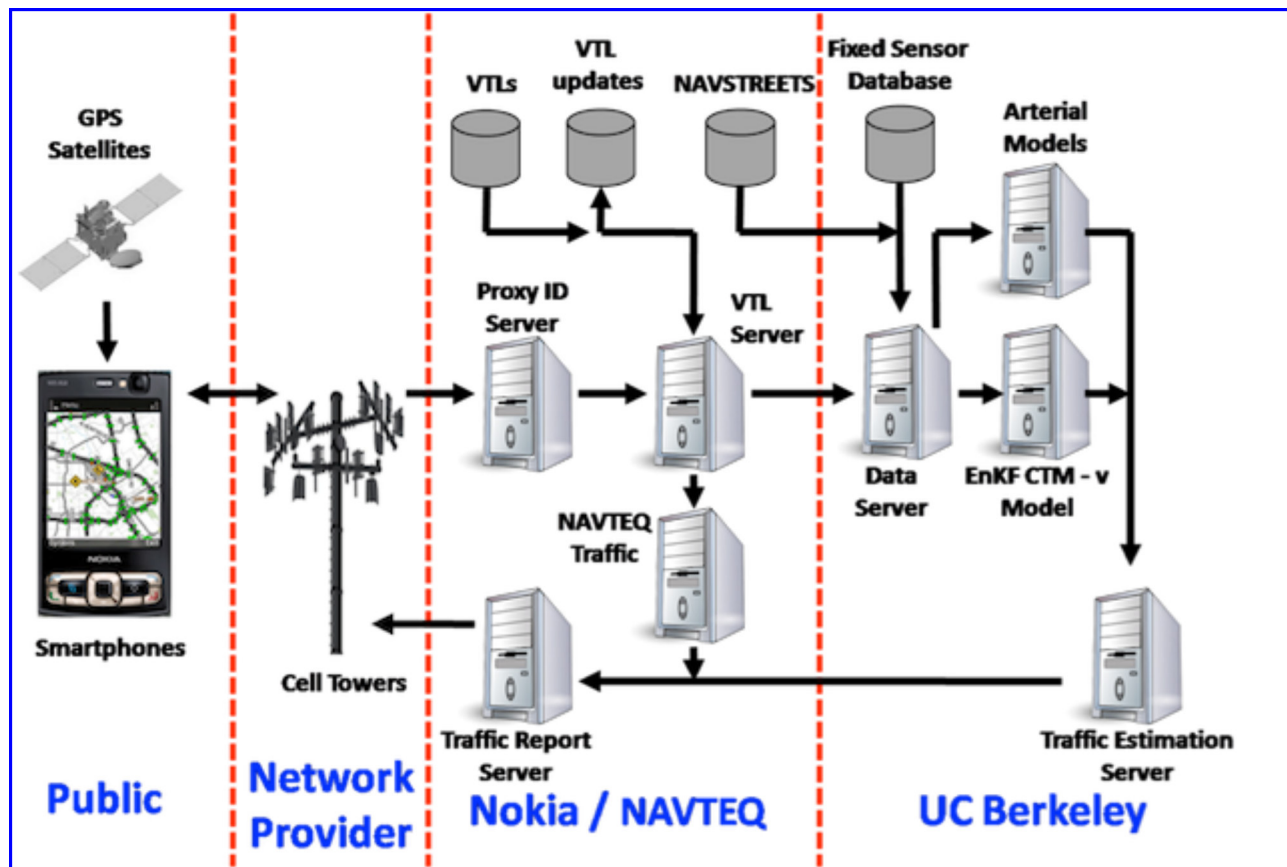
Data is only collected at virtual trip lines placed around the city, helping to maintain user privacy. *Image: UC Berkeley.*

This still leaves the possibility a sequence of measurements can be processed to build up a trajectory. Nokia created 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 determine which belong together. The algorithm uses the number of cars on the road and their speeds to determine 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 processed the incoming data, making it less likely anyone could manipulate VTL placement to make tracking a car easier.

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

The final architecture is illustrated below, showing the multilayered 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 architecture for gathering and processing data. *Image: UC Berkeley*

## Making Sense Of It All

Developing the algorithms for data fusion fell to researchers at UC Berkeley. In addition to the GPS measurements from the phones, the system incorporates GPS data from buses, taxis and other fleet vehicles. Data from static sensors in the region, such as loop detectors and RFID tag readers, also are included. The question that the data fusion algorithms try to answer is: Given all the measurements gathered from a given road, what is the best estimate of the number of cars on that road and how fast are they going?

GPS tracks in general are hard to process for traffic monitoring, and there were many challenges. One of the first was figuring out what road the measurements were coming from.

“You had to create a fully integrated geolocating system to fuse the data,” Bayen said. “You need the underlying road network on which you map measurements to.”

NAVTEQ’s mapping information was vital, but there was a lot of post-processing 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 tough. People may be walking alongside the road with their phone in their pocket, or they may park the car and forget to turn the GPS off. In urban canyons like downtown San Francisco, many of the GPS data points do not exactly match known roads because buildings obscure satellites. Measurements have to be associated with particular roads using machine-learning methods. These methods attempt to find the most likely road for a particular data point and

reject those not likely to be moving cars.

The biggest challenge, and one that remains, 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 said. "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 said. "There's no completed theory for it."

Unlike traditional static sensors, instead of measuring all cars passing a particular location, a GPS measurement gives a single measurement for a single car. This is hard to deal with. To understand why, we must look at how traffic flow is modeled.

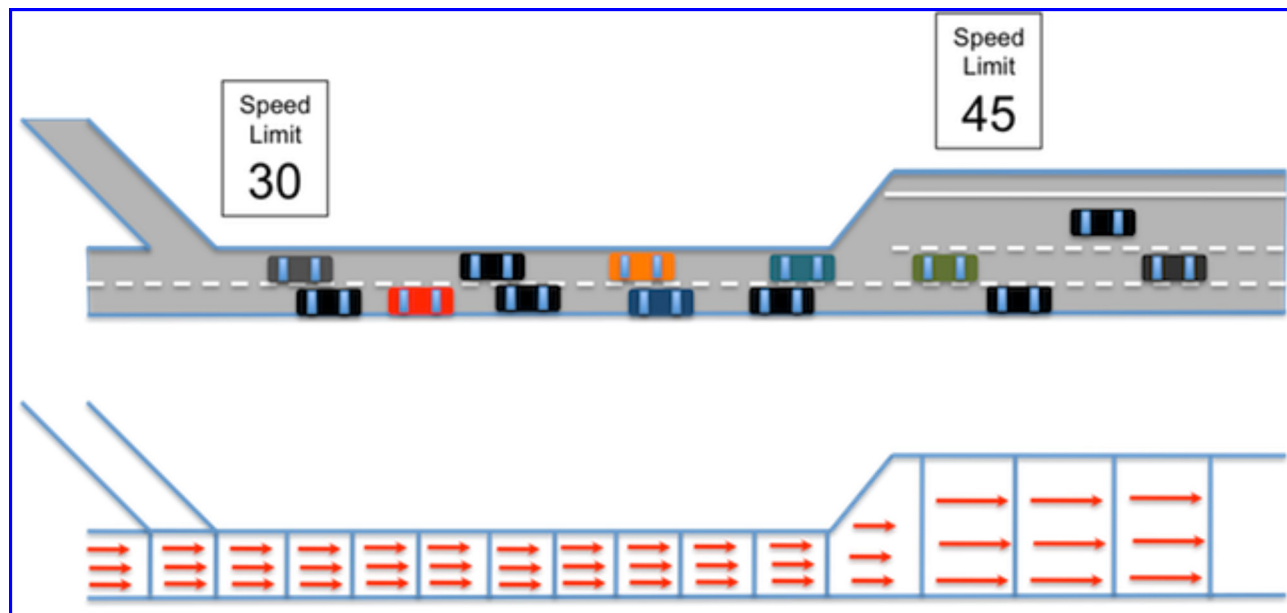
## The Flow of Traffic

The obvious thing to do when modeling cars on a roadway is track 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 get expensive quickly.

To get around this limitation, researchers often treat the movement of cars as liquid flowing through a series of tubes. Each segment of tube is a portion of road; 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 the network down into a connected sequence of small volumes, where each volume is a cell connected to others. The flow properties in each cell affect those neighboring 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. *Image: UC Berkeley*

Two more pieces of information are needed to solve the equations. First, the boundary conditions must 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 this information is 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 beginning 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 awhile and are pretty accurate within the limitations of the sensors.

This would be fine if the cars truly were a fluid, but driver actions lead to perturbations that cause slowdowns or accidents. These disruptions can't be detected until their effects ripple down 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 augmented as needed, perhaps to get more measurements on roads where the state of traffic is changing rapidly.

Although virtual sensors can be placed more densely than physical ones, their measurements are less complete. A physical sensor will count and measure the speed of every car passing it. Even complete GPS trajectories from vehicles being tracked provide 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 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 many different algorithms for a variety of different models. The details are arcane and described in [papers available 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 phones and with GPS trajectories from other sources, like buses. The resulting estimates of traffic flow are much better than those available from static sensing alone.



Field experiments validated the technology behind Mobile Millennium and captured an accident in real time.  
*Image: UC Berkeley*

## Mobile Century

The initial design of the Mobile Millennium system culminated in a proof-of-concept test called Mobile

Century on Feb. 8, 2008. One hundred 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 the Bay Area. 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 posted students with video cameras on overpasses.

The test ran nearly 10 hours and required more than 150 student drivers; the results were a great success. Although Mobile Century cars accounted for no more than 2 to 5 percent of cars on the road at any given time, the system very accurately measured the speed and density of traffic, and at a much higher spatial resolution than the fixed system of loop detectors. The test also provided a startling demonstration of the potential of using mobile phones to gather data quickly.

Traffic estimates calculated with the test data were displayed in real time at a control center and observed by researchers and various transportation officials. At 10:50 a.m. the team noticed its data displaying a serious slowdown in traffic, while data from Google Maps, which at the time drew data primarily from static loop detector sensors, showed things were all clear.

“We were getting nervous,” Professor Bayen said. “There were all these officials watching, and we thought maybe something had gone wrong.”

Everyone let out a sigh of relief when the Google display slowly caught up and beepers sounded as automated alerts went out to the visiting transportation officials. There had been a five-car pileup exactly where the Mobile Century system first reported the slowdown. It was clear validation of the project. The sudden slowdown 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.

## Until 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 some 5,000 users with it driving around the San Francisco Bay Area.

The concepts and technology demonstrated in Mobile Millennium are now widespread. Google’s mobile Maps app also fuses mobile GPS data with static sensors and other sources. Many companies that provide traffic monitoring data do something similar, either using phones or other dedicated mobile sources. A large number of cities use similar means of combining static and mobile sensors to measure traffic patterns.

The future of mobile sensing isn’t limited to traffic monitoring. The [CarTel project at Massachusetts Institute of Technology](#) demonstrated the use of accelerometers mounted on a fleet of a local limo company to 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 also was 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. It’s only a matter of time before our [cars are fully networked](#) and the traffic flow becomes all but self-aware. Tighter integration of

phones and data networks with cars will make still 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.

It's 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 the companies building such systems are as wary of the potential dangers as they are hopeful about the rewards.



This story was written by [Haomiao Huang](#) and [originally published by Ars Technica](#).

Main photo: [silva613/Flickr](#)

See Also:

- [The Future of Cars: P2P Mesh, 4G and the Cloud](#)
- [‘Talking’ Cars Are Coming Soon to Keep Us Safe](#)
- [High-Tech Car Allows the Blind to Drive](#)
- [Volvo Tests Almost-Autonomous ‘Road Train’](#)
- [Audi’s Robotic Car Climbs Pikes Peak](#)
- [Autonomous Autos Play Well With Others](#)

Tags: [Infrastructure](#), [Smartphone Apps](#), [Streets & Highways](#)  
[Post Comment](#) | [Permalink](#)

#### Also on Wired.com

- [Gallery: Apple's Jobs Unveils Upgraded, More Colorful iPad 2](#)
- [Military Drops Crazy 'ICBMs vs. Terrorists' Plan](#)
- [Peugeot Brings Diesel Hybrid Hotness to Racing](#)
- [Live Blog: Apple Unveils Thinner, Lighter iPad 2](#)
- [Nintendo Unveils 3DS Netflix, \*Super Mario\* at GDC](#)
- [‘Nude’ Airport Scanners: Are They Safe?](#)

#### Related Topics:

- [Cellular Phones](#),
- [Communications Products](#),
- [Consumer Electronics](#),
- [Electronics](#),
- [Global Positioning Systems](#),
- [Google Inc.](#)

[Login](#)

## Add New Comment



Type your comment here.

Real-time updating is **paused**. ([Resume](#))

Sort by popular now

## Showing 4 comments



**waltep**

Bandwidth is going to have to increase drastically. 250 GB limits are not going to be enough.

[3 hours ago](#)

[Like](#) [Reply](#)



**Narciso Torrente**

Great Article....we are kidding ourselves by thinking that traffic congestion is going to get any better. We need less cars on the road. We need alternatives, now. It is too costly to maintain from a hard to soft costing. Don't get me wrong....I love technology: cell phone, bicycles, solar energy, high speed trains, clean fuel buses, GPS, .....

[4 hours ago](#)

[Like](#) [Reply](#)



**Frédéric CHARLES**, Management des systèmes d'information et collaboratifs

Vodafone has this service operational in Europe to produce info trafic  
I worked in 2001 on that subject with SFR to follow speeds on an highway trunc (European project)  
Anyway go bears!

[5 hours ago](#)

[Like](#) [Reply](#)



**jibran**

one day there will be cars without driver and that will be the future of traffic

6 hours ago

Like Reply

[M](#) [Subscribe by email](#) [S](#) [RSS](#)

[blog comments powered by DISQUS](#)