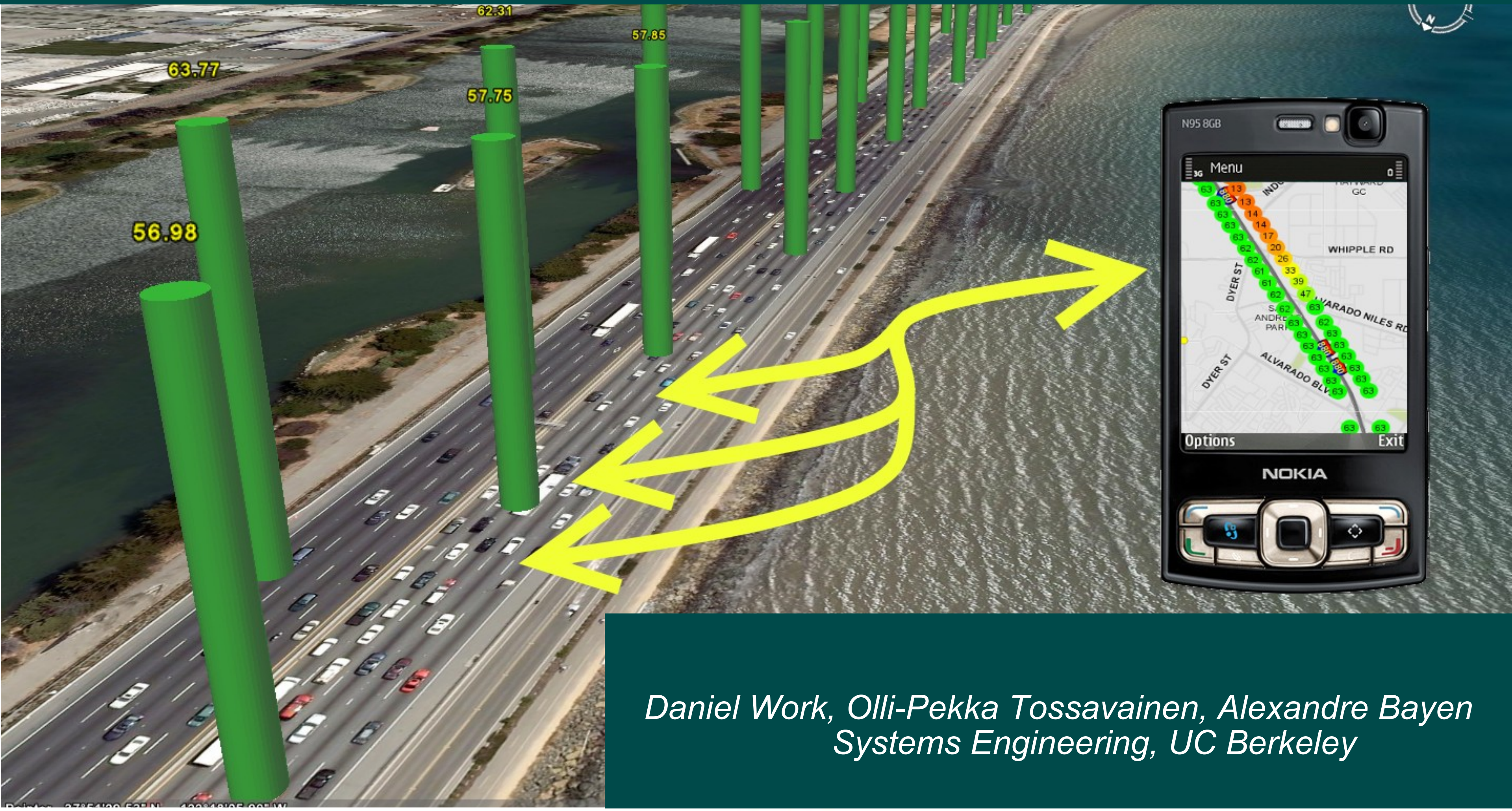
MOBILE CENTURY – Using GPS Mobile Phones as Traffic Sensors



Daniel Work, Olli-Pekka Tossavainen, Alexandre Bayen
Systems Engineering, UC Berkeley

NOKIA Connecting People CITRIS California Center for Innovative Transportation Caltrans Institute of transportation studies ITS TEKES

Outline

- Motivation
- GPS-based traffic flow monitoring
- Incorporation of Lagrangian measurements in estimation
- First order flow models
- Velocity field reconstruction
- Twin experiments (numerical simulation)
- February 8th, 2008: 100 car experiment

Outline

- Motivation
- GPS-based traffic flow monitoring
- Incorporation of Lagrangian measurements in estimation
- First order flow models
- Velocity field reconstruction
- Twin experiments (numerical simulation)
- February 8th, 2008: 100 car experiment

Societal need for [real-time] traffic monitoring

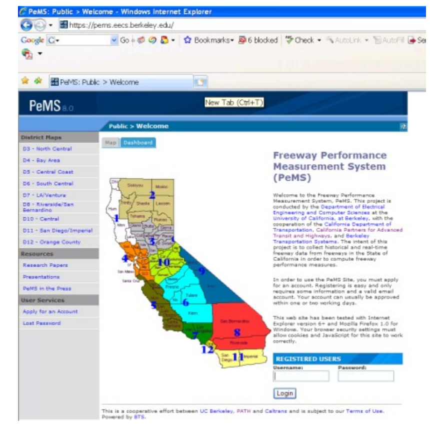
UC Berkeley

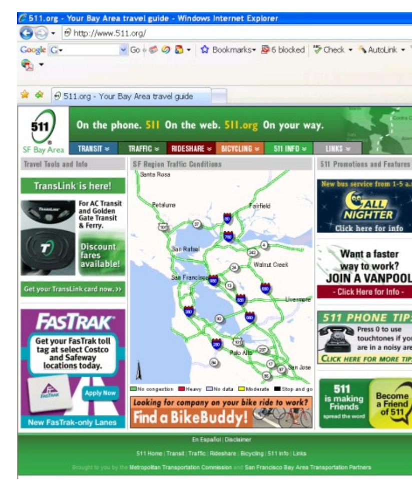
- Rough estimates of congestion impacts
  - 4.2 billion hours extra travel in the US alone
  - Accounts for 2.9 billion gallons of fuel
  - Congestion cost of 78 billion dollars

[2007 Urban Mobility Report, September 2007, Texas Transportation Institute, David Schrank & Tim Lomax]

- Need for [real-time] monitoring
  - Commuter services:
    - Call in numbers (511)
    - Websites: www.511.org, www.traffic.com
    - Changeable Message Signs (CMS)
    - Phone services
  - Planner services
    - PeMS, data archival

Chronicle / Chris Stewart

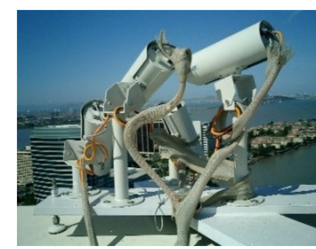




Current sensing devices in use

- Currently available sensors for monitoring traffic
  - Self inductive loops, embedded in the pavement
  - Wireless pavement sensors
  - FasTrak, EZ-pass transponders
  - Cameras
  - Radars
  - License plate readers

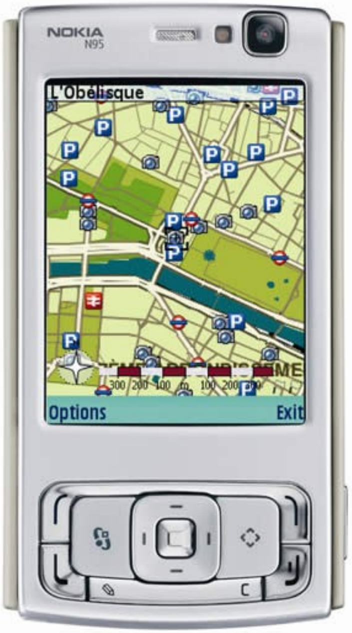
- Issues with using currently available sensors
  - Loop detectors
    - Costly and hard to install
    - Need maintenance,
    - Reliability issues
    - Inaccuracy due to inherent data aggregation
  - FasTrak and license plate readers are privacy intruding
  - Only parts of the US are equipped



[Choe, Skabardonis, Varaiya, Transportation Research Board, 2002]

Convergence of multimedia, sensing and communication

- N95 is the incarnation of the convergence of multi-media, sensing and communication platforms
  - GPS
  - mp3 and movie player
  - Multiple sensors (accelerometers, light)
  - Radio, wireless, bluetooth, various ports, infrared, etc.
  - 5 megapixel camera



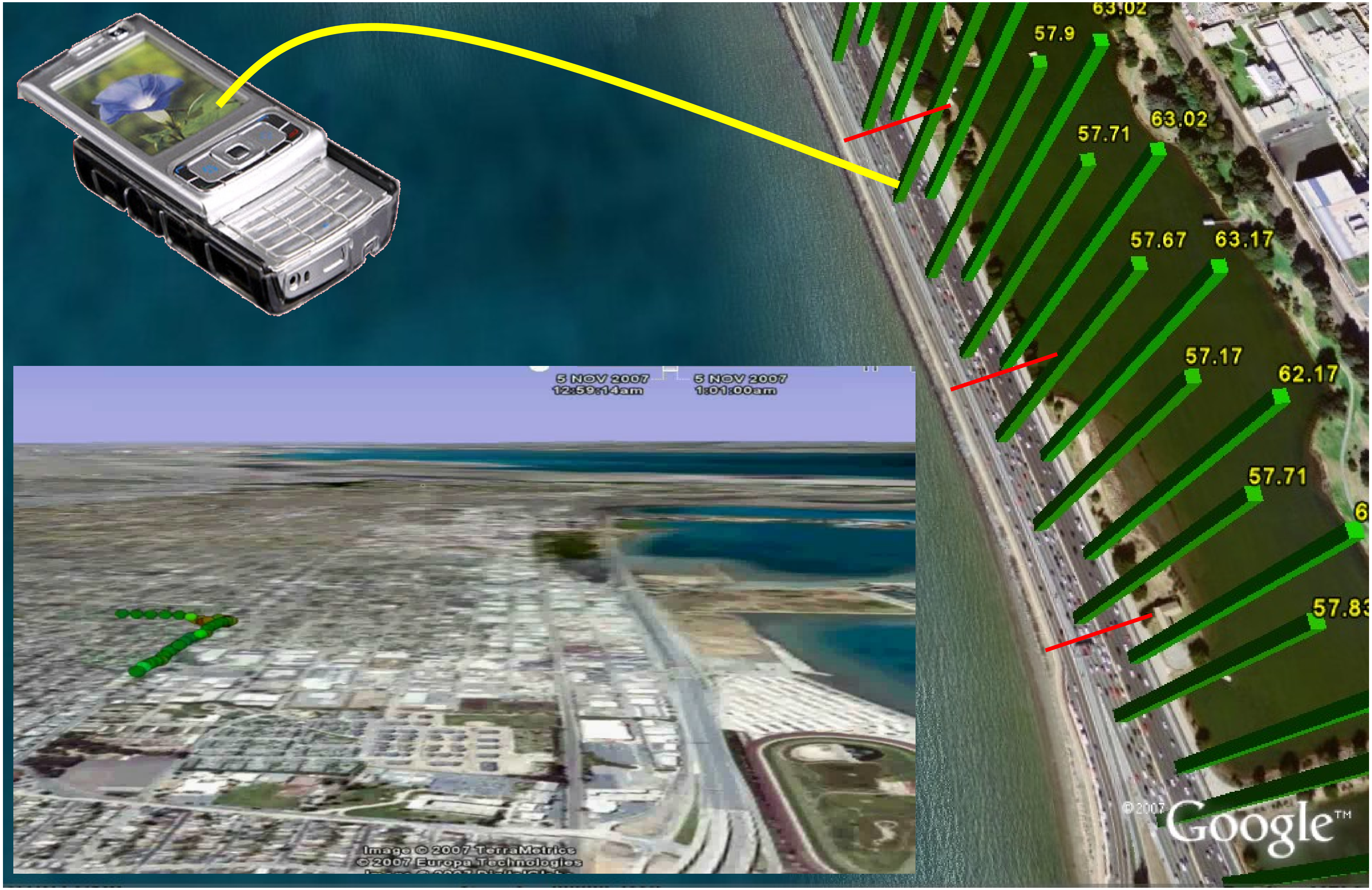
- Smart phones open the door for
  - Location based services
  - Context awareness
  - Mobility tracking

- Ubiquitous Sensing Platform (Nokia)
  - 3 billion mobile devices by 2009
  - 1. 5 million devices per day

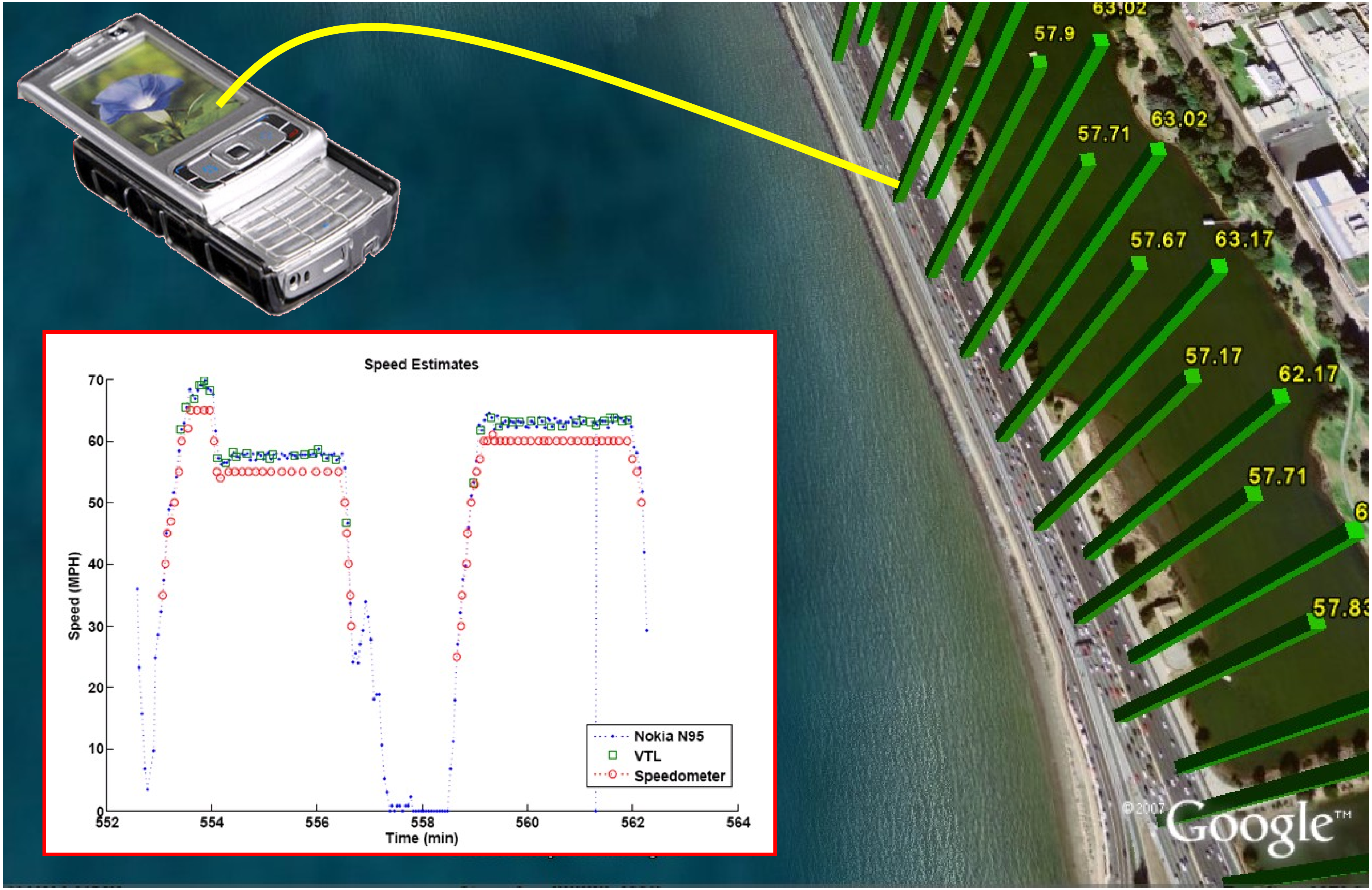
Outline

- Motivation
- GPS-based traffic flow monitoring
- Incorporation of Lagrangian measurements in estimation
- First order flow models
- Velocity field reconstruction
- Twin Experiments (numerical simulation)
- February 8th, 2008: 100 car experiment

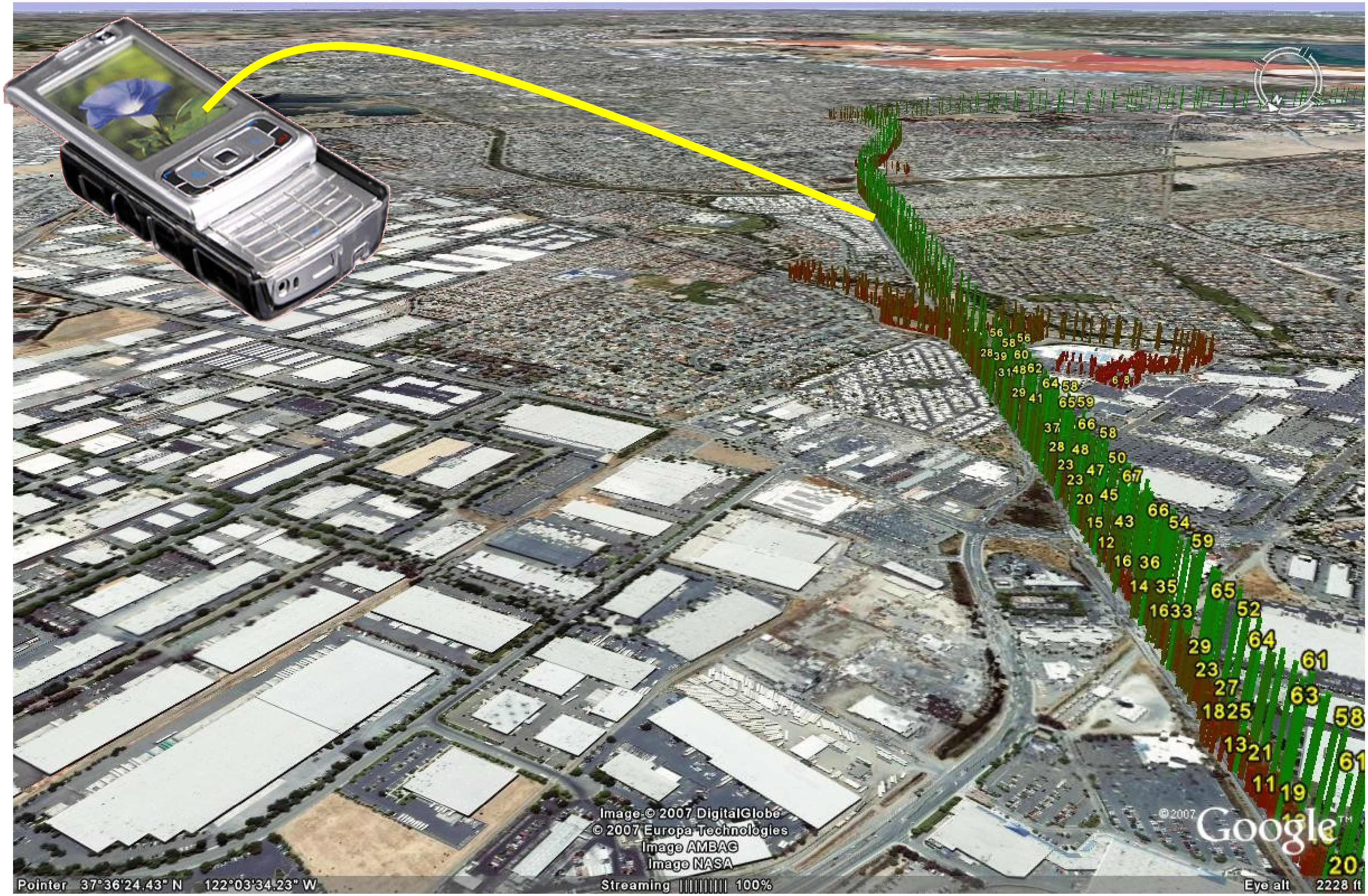
NextGen mobile highway sensor: the N95



NextGen mobile highway sensor: the N95



NextGen mobile highway sensor: the N95



Outline

- Motivation
- GPS-based traffic flow monitoring
- Incorporation of Lagrangian measurements in estimation
- First order flow models
- Velocity field reconstruction
- Twin experiments (numerical simulation)
- February 8th, 2008: 100 car experiment

Key questions which arise with Lagrangian sensing

- Estimation accuracy
  - Quantification of set of measurement samples
  - Specification of selection procedure
- Privacy intrusion
  - Data anonymization (removal of cell phone ID / user ID)
  - Cryptography (encryption of the data sent)
  - Mobility tracking (sampling selection procedure)

    KEY concept for this: VTL, virtual trip line
- Bandwidth / energy requirements
  - GPS tracking uses the phone battery
  - Large sample sizes use bandwidth
- Travel time definitions
  - Instantaneous travel time (assume conditions on the road are static)
  - Dynamic travel time (use flow model to forecast conditions)
  - A posteriori travel time (travel time of a completed trip)
  - Heterogeneous travel time (HOV / non HOV travel times)
  - Personal travel time (drivers behave differently, the phone can learn)

One possible answer to privacy: VTL

- Virtual trip lines
  - Virtual trip lines are geographical markers stored in the client (application running in the phone) which trigger a position and speed update whenever a probe vehicle passes.
  - If VTLs are placed properly, they provide a data source in a privacy preserving environment.

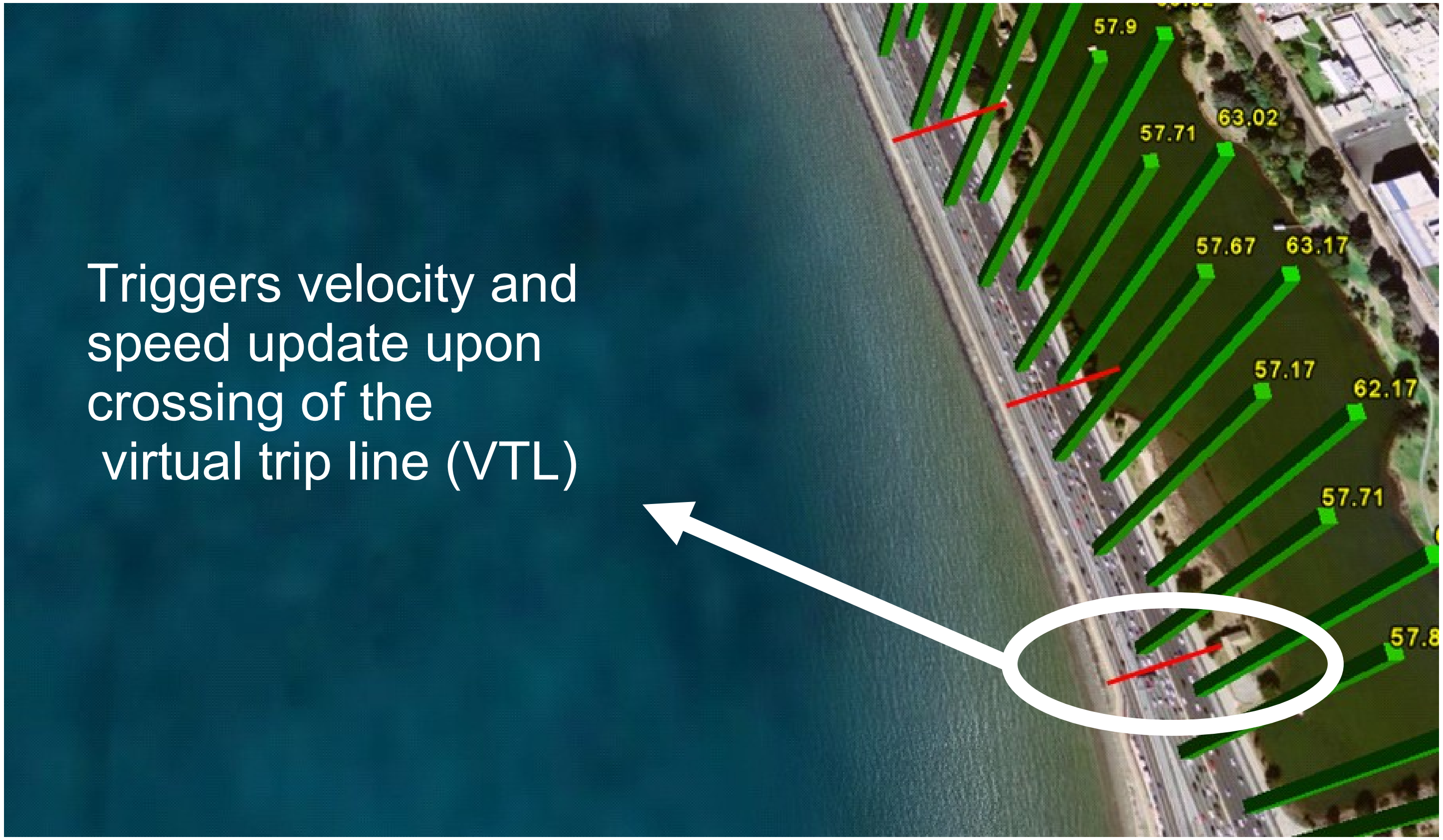
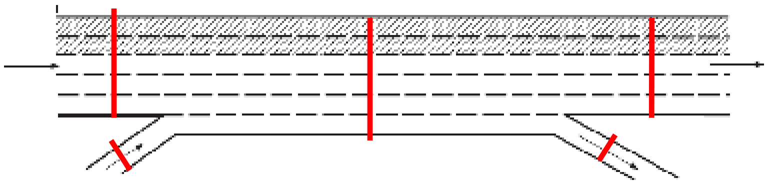
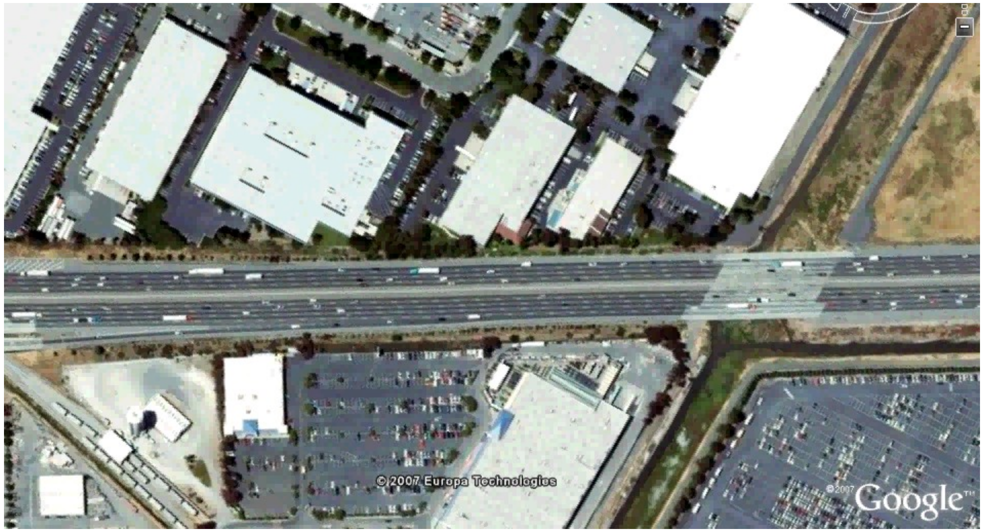


Illustration of VTLs: Pilot test 1, Nov. 2nd, 2007



- Virtual Trip Lines
  - Virtual Eulerian measurements
  - Privacy preserving
  - Provide sampled speed information
  - Deployable almost instantaneously
  - Flexible
- Can be deployed on
  - Mainline
  - Off ramps
  - On ramps
  - HOV lane



Outline

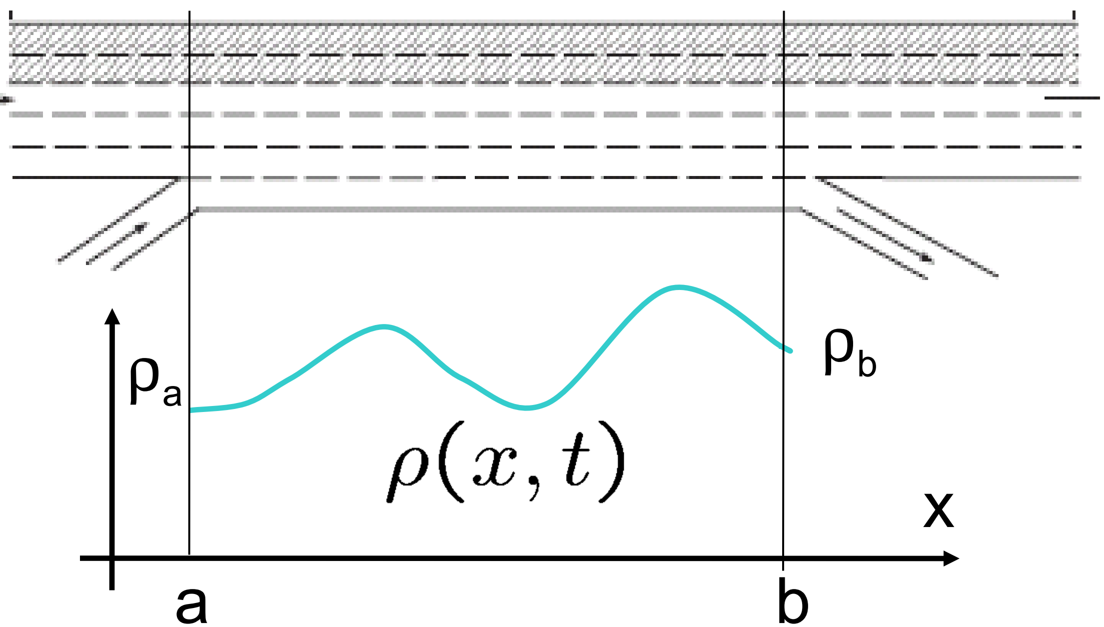
- Motivation
- GPS-based traffic flow monitoring
- Incorporation of Lagrangian measurements in estimation
- First order flow models
- Velocity field reconstruction
- Twin Experiments (numerical simulation)
- February 8th, 2008: 100 car experiment

A very brief introduction to traffic flow modeling

- Seminal hydrodynamic model: the Lighthill-Whitham-Richards partial differential equation
  - Nonlinear first order hyperbolic scalar conservation law
  - Concave flux function (empirical fundamental diagram)
  - Weak boundary conditions

    [Lighthill-Whitham, 1955; Richards, 1956; Bardos Leroux Nedelec, 1979; Bayen Strub 2006]

$\rho(x,t)$ is the vehicle density.



$$\frac{\partial \rho}{\partial t} + \frac{\partial q(\rho)}{\partial x} = 0$$

$$\rho(a,t) = \rho_a(t)$$

$$\rho(b,t) = \rho_b(t)$$

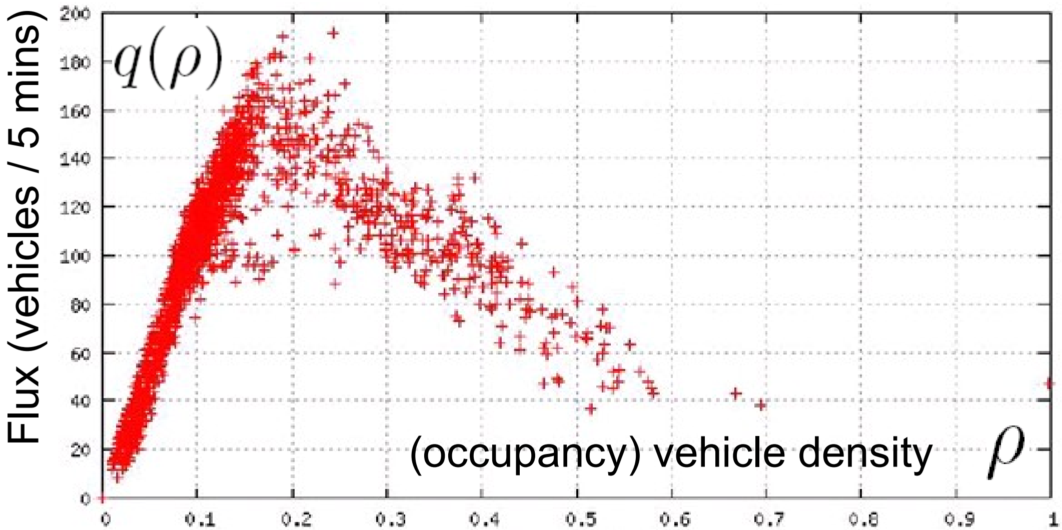
$$\rho(x,0) = \rho_0(x)$$

Lighthill-Whitham-Richards PDE

- Basic Fundamental Diagram – Greenshields.
  - Express velocity as a function of density to obtain a PDE in density only.

$$v(\rho) = v_{\max}\left(1 - \frac{\rho}{\rho_{\max}}\right)$$

$\rho(x,t)$ is the vehicle density.



$$\frac{\partial \rho}{\partial t} + \frac{\partial q(\rho)}{\partial x} = 0$$

$$\rho(a,t) = \rho_a(t)$$

$$\rho(b,t) = \rho_b(t)$$

$$\rho(x,0) = \rho_0(x)$$

Mathematical formulation of the problem

First order hyperbolic PDE (LWR PDE)

$$\frac{\partial \rho}{\partial t} + \frac{\partial q(\rho)}{\partial x} = 0$$

Initial conditions

$$\rho(x,0) = \rho_0(x)$$

Weak boundary conditions

$$\begin{cases} \rho(a,t) = \rho_a(t) \text{ or} \\ q'\left(\rho\left(a,t\right)\right) \leqslant 0 \text{ and } q'\left(\rho_a\left(t\right)\right) \leqslant 0 \text{ or} \\ q'\left(\rho(a,t)\right) \leqslant 0 \text{ and } q'\left(\rho_a(t)\right) \geqslant 0 \text{ and } q\left(\rho(a,t)\right) \leqslant q\left(\rho_a(t)\right) \end{cases}$$

$$\begin{cases} \rho(b,t) = \rho_b(t) \text{ or} \\ q'\left(\rho\left(b,t\right)\right) \geqslant 0 \text{ and } q'\left(\rho_b\left(t\right)\right) \geqslant 0 \text{ or} \\ q'\left(\rho(b,t)\right) \geqslant 0 \text{ and } q'\left(\rho_b(t)\right) \leqslant 0 \text{ and } q\left(\rho(b,t)\right) \leqslant q\left(\rho_b(t)\right) \end{cases}$$

[Lighthill-Whitham, 1955; Richards, 1956; Bardos Leroux Nedelec, 1979; Strub, Bayen 2006]

Evolution equation for velocity

- All GPS measurements are velocities, not densities. This motivates the need for an alternative model.
- If we choose the Greenshields model, the relationship:

$$v(\rho) = v_{\max}\left(1 - \frac{\rho}{\rho_{\max}}\right)$$

can be inverted, namely: $\rho(v) = \rho_{\max}\left(1 - \frac{v}{v_{\max}}\right)$

Substituting into LWR, we obtain the LWR-v PDE:

$$\frac{\partial v}{\partial t} + \frac{\partial R(v)}{\partial x} = 0$$ where $$R(v) := (v)^2 - v v_{\max}$$

[Blandin, Bayen, 2008]

Evolution equation for velocity

$$\frac{\partial v}{\partial t} + \frac{\partial R(v)}{\partial x} = 0$$

Initial conditions

$$v\left(x, 0\right) = v_0\left(x\right)$$

Weak boundary conditions

$$\begin{cases} v\left(a, t\right) = v_a\left(t\right) \text{ or} \\ R'\left(v\left(a, t\right)\right) \leq 0 \text{ and } R'\left(v_a\left(t\right)\right) \leq 0 \text{ or} \\ R'\left(v\left(a, t\right)\right) \leq 0 \text{ and } R'\left(v_a\left(t\right)\right) \geq 0 \text{ and } R\left(v\left(a, t\right)\right) \geq R\left(v_a\left(t\right)\right) \end{cases}$$

$$\begin{cases} v\left(b, t\right) = v_b\left(t\right) \text{ or} \\ R'\left(v\left(b, t\right)\right) \geq 0 \text{ and } R'\left(v_b\left(t\right)\right) \geq 0 \text{ or} \\ R'\left(v\left(b, t\right)\right) \geq 0 \text{ and } R'\left(v_b\left(t\right)\right) \leq 0 \text{ and } R\left(v\left(b, t\right)\right) \geq R\left(v_b\left(t\right)\right) \end{cases}$$

[Blandin, Bayen, 2008]

Nonlinear time invariant dynamical system

Godunov discretization scheme

$$v_i^{n+1} = v_i^n - \frac{\Delta t}{\Delta x}\left(g\left(v_i^n, v_{i+1}^n\right) - g\left(v_{i-1}^n, v_i^n\right)\right)$$

Godunov flux

$$g\left(v_1, v_2\right) = \begin{cases} R\left(v_2\right) & \text{if } v_1 \leq v_2 \leq v_c \\ R\left(v_c\right) & \text{if } v_1 \leq v_c \leq v_2 \\ R\left(v_1\right) & \text{if } v_c \leq v_1 \leq v_2 \\ \max\left(R\left(v_1\right), R\left(v_2\right)\right) & \text{if } v_1 \geq v_2 \end{cases}$$

Discretization v_i^n

$$\Delta t = \frac{T}{N} \qquad \Delta x = \frac{b-a}{M}$$

Nonlinear discrete dynamical system

$$v^{n+1} = \mathcal{M}[v^n]$$

[Blandin, Bayen, 2008]

Outline

- GPS-based traffic flow monitoring
- Incorporation of Lagrangian measurements in estimation
- First order flow models
- Velocity field reconstruction
- Twin experiments (numerical simulation)
- February 8th, 2008: 100 car experiment

Observation model

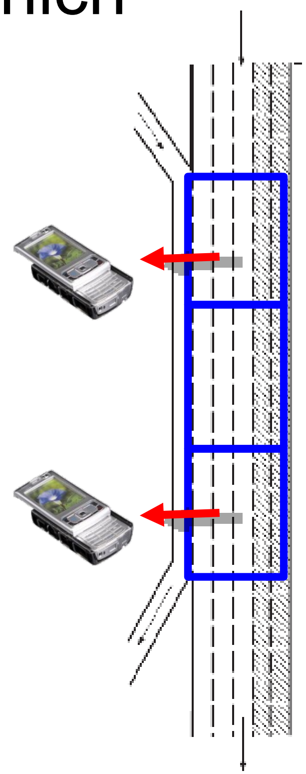
- We model the observations as follows:

$$\boxed{y^n = \mathbf{H}^n v^n + \gamma^n}$$

y^n is the vector of observed state (ie GPS measurements).

$\mathbf{H}^n \in \{0,1\}^{p^n \times M}$ is the observation matrix which encodes the discrete cells in which measurements are observed.

$$\mathbf{H}^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



$\gamma^n \sim (0, \mathbf{R}^n)$ is the measurement noise

- From GPS.

- From average velocity assumption.

Linear state estimation

- We form the linear state estimator as follows:

$$\boxed{\hat{v}^{n+} = \hat{v}^{n-} + \mathbf{G}^n(y^n - \mathbf{H}^{n+}\hat{v}^{n-})}$$

$\hat{v}^{n-}$ estimate before measurements are known

$\hat{v}^{n+}$ estimate after measurements are known

$(y^n - \mathbf{H}^{n+}\hat{v}^{n-})$ is called the *innovation*

G is the *Kalman Gain*, to be determined.

Choice of Kalman gain

- Optimality Criterion: minimize the sum of the variances of the estimation errors:

Min: $$J^n = E\left[\left(v_1^n - \hat{v}_1^{n+}\right)^2\right] + \cdots + E\left[\left(v_M^n - \hat{v}_M^{n+}\right)^2\right]$$

$$= E\left[\mathrm{Tr}\ \left(\left(v^n - \hat{v}^{n+}\right)\left(v^n - \hat{v}^{n+}\right)^T\right)\right]$$

$$= \mathrm{Tr}\ \mathbf{P}^{n+}$$

where $\mathbf{P}^{n+}$ is the estimation error covariance after the measurements are known.

Solve for the optimal Kalman gain (1/2)

- Define $\epsilon_{v,n+} = v^n - \hat{v}^{n+}$, so that: $J^n = E\left[\mathrm{Tr}\ \left(\epsilon_{v,n+}\epsilon_{v,n+}^T\right)\right]$

but $\epsilon_{v,n+} = \ v^n - \hat{v}^{n-} - \mathbf{G}^n\left(y^n - \mathbf{H}^{n+}\hat{v}^{n-}\right)$

(from linear estimator)

and $\epsilon_{v,n+} = v^n - \hat{v}^{n-} - \mathbf{G}^n\left(\mathbf{H}^{n+}v^n + \gamma^n - \mathbf{H}^{n+}\hat{v}^{n-}\right)$

(from observation model)

so $\epsilon_{v,n+} = \left(I - \mathbf{G}^n\mathbf{H}^{n+}\right)\epsilon_{v,n-} - \mathbf{G}^n\gamma^n$

(collect terms)

Substitute back into J^n :

$$J^n = \mathrm{Tr}\ \left(\left(I - \mathbf{G}^n\mathbf{H}^{n+}\right)\mathbf{P}^{n-}\left(I - \mathbf{G}^n\mathbf{H}^{n+}\right)^T + \mathbf{G}^n\mathbf{R}\left(\mathbf{G}^n\right)^T\right)$$

Solve for the optimal Kalman gain (2/2)

- From previous slide:

$$J^n = \mathrm{Tr}\left(\left(I - \mathbf{G}^n\mathbf{H}^{n+}\right)\mathbf{P}^{n-}\left(I - \mathbf{G}^n\mathbf{H}^{n+}\right)^T + \mathbf{G}^n\mathbf{R}\left(\mathbf{G}^n\right)^T\right)$$

Now solve the unconstrained minimization problem:

$$\frac{\partial J^n}{\partial \mathbf{G}^n} = 0$$

Finally, solve for $\mathbf{G}^n$

$$\boxed{\mathbf{G}^n = \mathbf{P}^{n-}\left(\mathbf{H}^{n+}\right)^T\left(\mathbf{H}^{n+}\mathbf{P}^{n-}\left(\mathbf{H}^{n+}\right)^T + \mathbf{R}^n\right)^{-1}}$$

Linear state estimation - summary

- For an initial estimate of the state at time n: $\hat{v}^{n-}$

With covariance: $\mathbf{P}^{n-}$

And observations : y^n and corresponding map: $\mathbf{H}^{n+}$

The optimal linear estimator (minimizes of the sum of the variances) is computed by:

$$\hat{v}^{n+} = \hat{v}^{n-} + \mathbf{G}^n(y^n - \mathbf{H}^{n+}\hat{v}^{n-})$$

With:

$$\mathbf{G}^n = \mathbf{P}^{n-}\left(\mathbf{H}^{n+}\right)^T\left(\mathbf{H}^{n+}\mathbf{P}^{n-}\left(\mathbf{H}^{n+}\right)^T + \mathbf{R}^n\right)^{-1}$$

Recursive state update

- We still need a state estimate at time n: $\hat{v}^{n-}$ given the best estimate at (n-1), and its covariance $\mathbf{P}^{n-}$.

We use the CTM-v state equation:

$$v^{n-} = \mathcal{M}[v^{n-1}] + \eta^n$$

where $\eta^n \sim (0, \mathbf{Q}^n)$ is the zero mean state noise with covariance $\mathbf{Q}^n$.

For a linear system, $\mathbf{P}^{n-}$ can be computed explicitly from the Lyapunov equation. The CTM-v is nonlinear.

Ensemble approximation

- Instead, we compute $\mathbf{P}^{n-}$ from a statistical ensemble of size K.

Let $\xi_k^{(n-1)+}$ be the *k*th realization of a random variable with mean $\bar{v}^{(n-1)+}$ and covariance $\mathbf{P}^{(n-1)+}$.

Update each realization according to the CTM-v:

$$\xi_k^{n-} = \mathcal{M}\left[\xi_k^{(n-1)+}\right] + \eta^n$$

Compute the mean and covariance as follows:

$$\bar{v}^{n-} = \frac{1}{K}\sum_{1}^{K}\xi_k^{n-} \quad \text{and} \quad \mathbf{P}^{n-} = \frac{1}{K-1}\sum_{k=1}^{K}\left(\bar{v}^{n-} - \xi_k^{n-}\right)\left(\bar{v}^{n-} - \xi_k^{n-}\right)^T$$

[Evenson, Journal of Geophysical Research, 1994]

The ensemble Kalman filter (initialize, forecast, update)

- We now have the tools necessary to build the *ensemble Kalman filter* (EnKF).

To initialize the algorithm, draw *K* realizations ξ_k^{0-} of the initial state v^0 with covariance $\mathbf{P}^{0-}$.

In practice, we assume some correlation in $\mathbf{P}^{0-}$ so that the initial states are smooth.

With no measurements, $\xi_k^{0+} := \xi_k^{0-}$

The ensemble Kalman filter (initialize, forecast, update)

- Starting with n=1, compute the ensemble mean and covariance at each time step:

$$\xi_k^{n-} = \mathcal{M}[\xi_k^{(n-1)+}] + \eta^n$$

$$\bar{v}^{n-} = \frac{1}{K}\sum_{k=1}^{K}\xi_k^{n-}$$

$$\mathbf{P}^{n-} = \frac{1}{K-1}\sum_{k=1}^{K}\left(\bar{v}^{n-} - \xi_k^{n-}\right)\left(\bar{v}^{n-} - \xi_k^{n-}\right)^T$$

The ensemble Kalman filter (initialize, forecast, update)

- Obtain measurements, compute the Kalman gain, update ensemble estimates.

$$\mathbf{G}^n = \mathbf{P}^{n-} \left(\mathbf{H}^{n+}\right)^T \left(\mathbf{H}^{n+}\mathbf{P}^{n-} \left(\mathbf{H}^{n+}\right)^T + R^n\right)^{-1}$$

$$\xi_k^{n+} = \xi_k^{n-} + \mathbf{G}^n (y^n - \mathbf{H}^{n+} \xi_k^{n-})$$

- Optimal estimate with innovations:

$$\bar{v}^{n+} = \frac{1}{K} \sum_{1}^{K} \xi_k^{n+}$$

Outline

- Motivation
- GPS-based traffic flow monitoring
- Incorporation of Lagrangian measurements in estimation
- First order flow models
- Velocity field reconstruction
- Twin experiments (numerical simulation)
- February 8th, 2008: 100 car experiment

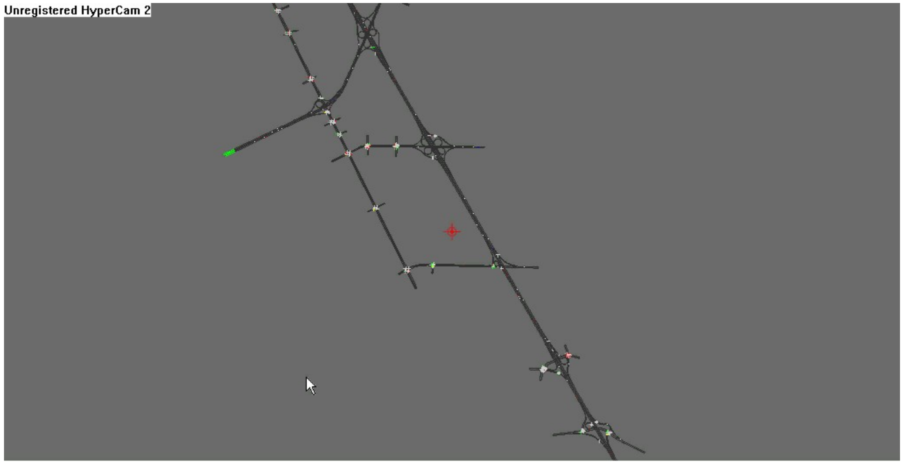
The twin experiments

- To test the algorithm, we set up twin experiments

  First model - Paramics

  – Highly calibrated traffic microsimulation.

  – "ground truth"



  Second model –CTM-v

  – Reduced order model

  – Running EnKF algorithm to estimate the state of the ground truth from limited observations

Velocity contour estimation

First model - Paramics

– Highly calibrated traffic microsimulation.

– “ground truth”

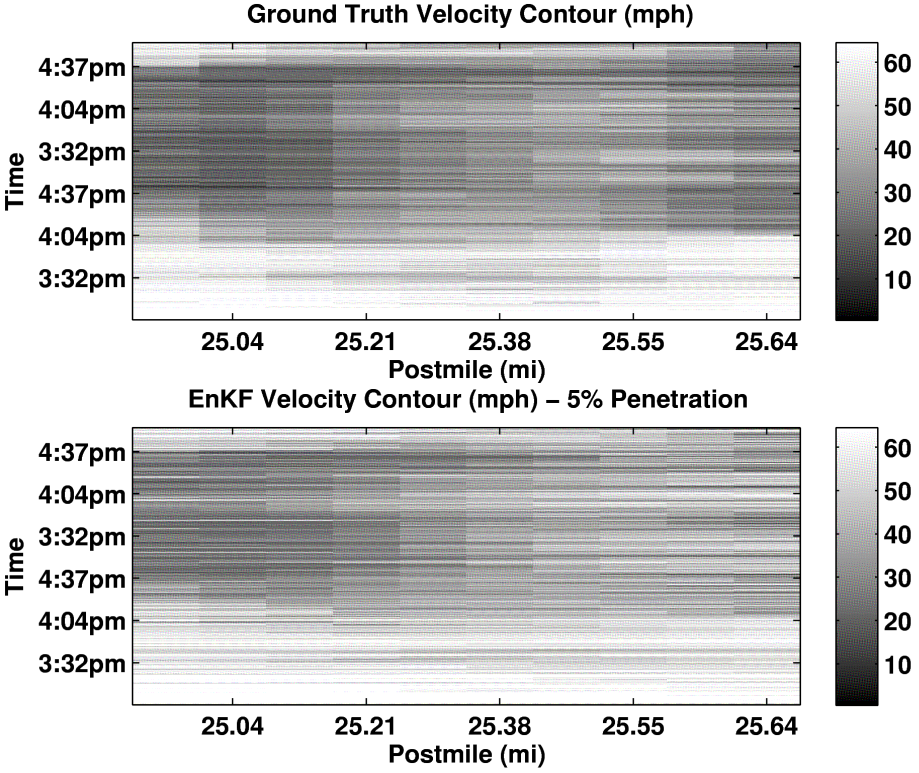
Second model –CTM-v

– Reduced order model

– Running EnKF algorithm

– 5% penetration

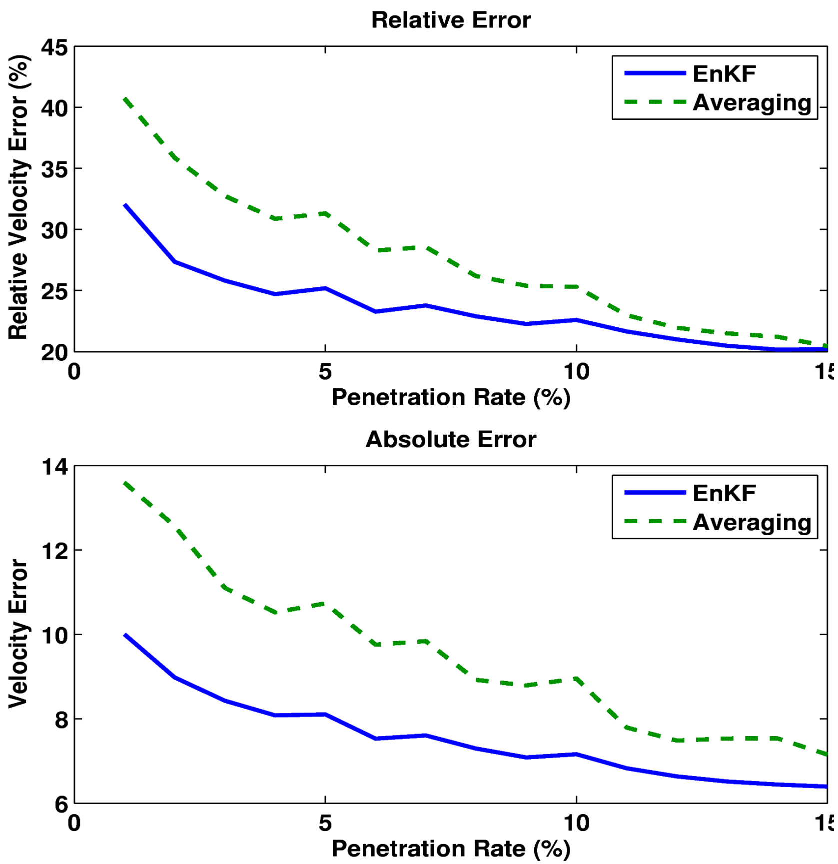
– 10 VTLs



Value of the flow model

- The EnKF algorithm outperforms trajectory averaging with less data.
- The averaging model (assuming full trajectory of equipped vehicles)

$$v_i^n = \begin{cases} v_i^{n-1} & \text{if } v_{\text{obs}} = \emptyset \\ (v_{\text{obs}})_i^n & \text{otherwise.} \end{cases}$$



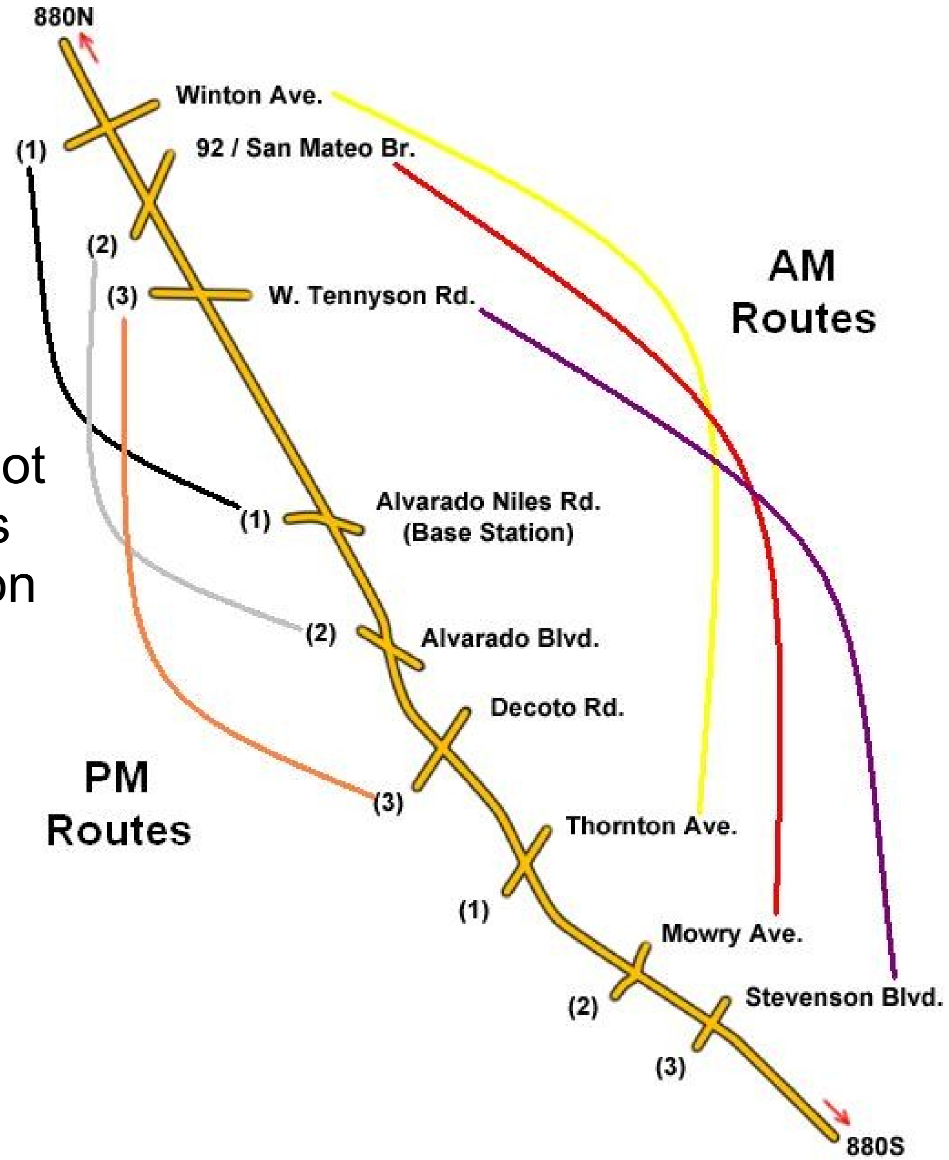
Outline

- GPS-based traffic flow monitoring
- Incorporation of Lagrangian measurements in estimation
- First order flow models
- Velocity field reconstruction
- Twin experiments (numerical simulation)
- February 8th, 2008: 100 car experiment

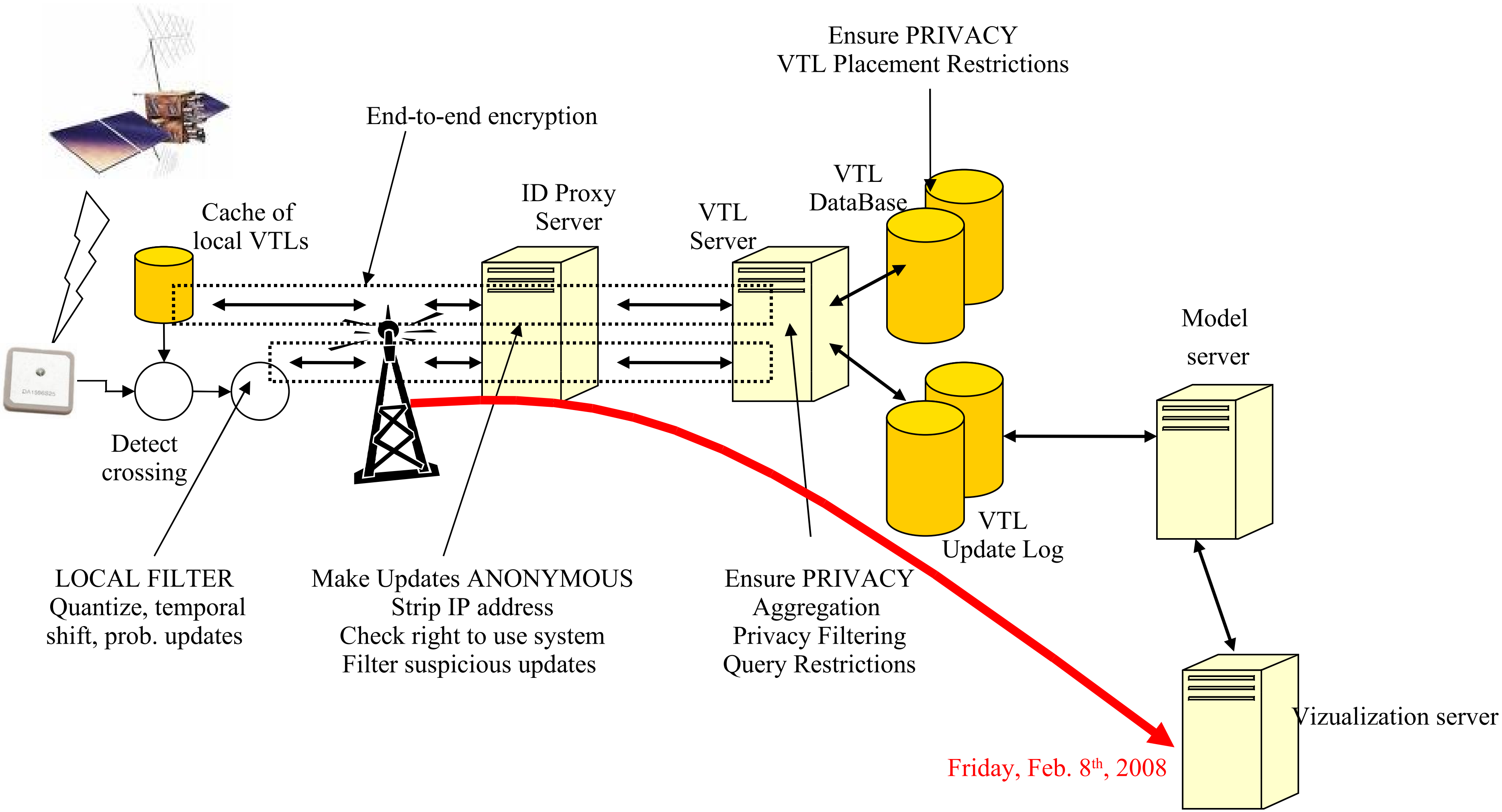
Mobile Century experiment

Cycle time expected to be close to 20 minutes.

Three different loops are used to not overload any ramp (each vehicle is expected to cross the same location 3 times per hour).

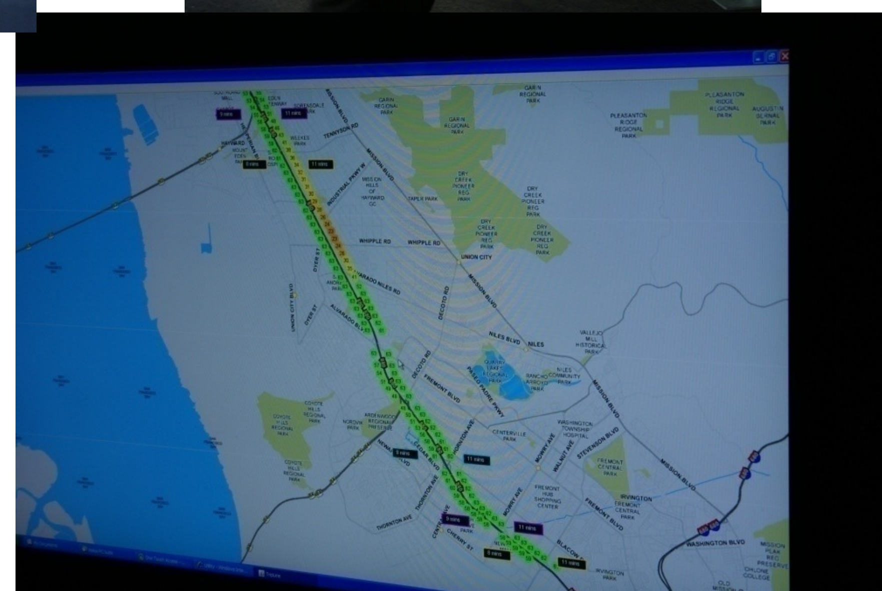


Mobile Century experiment



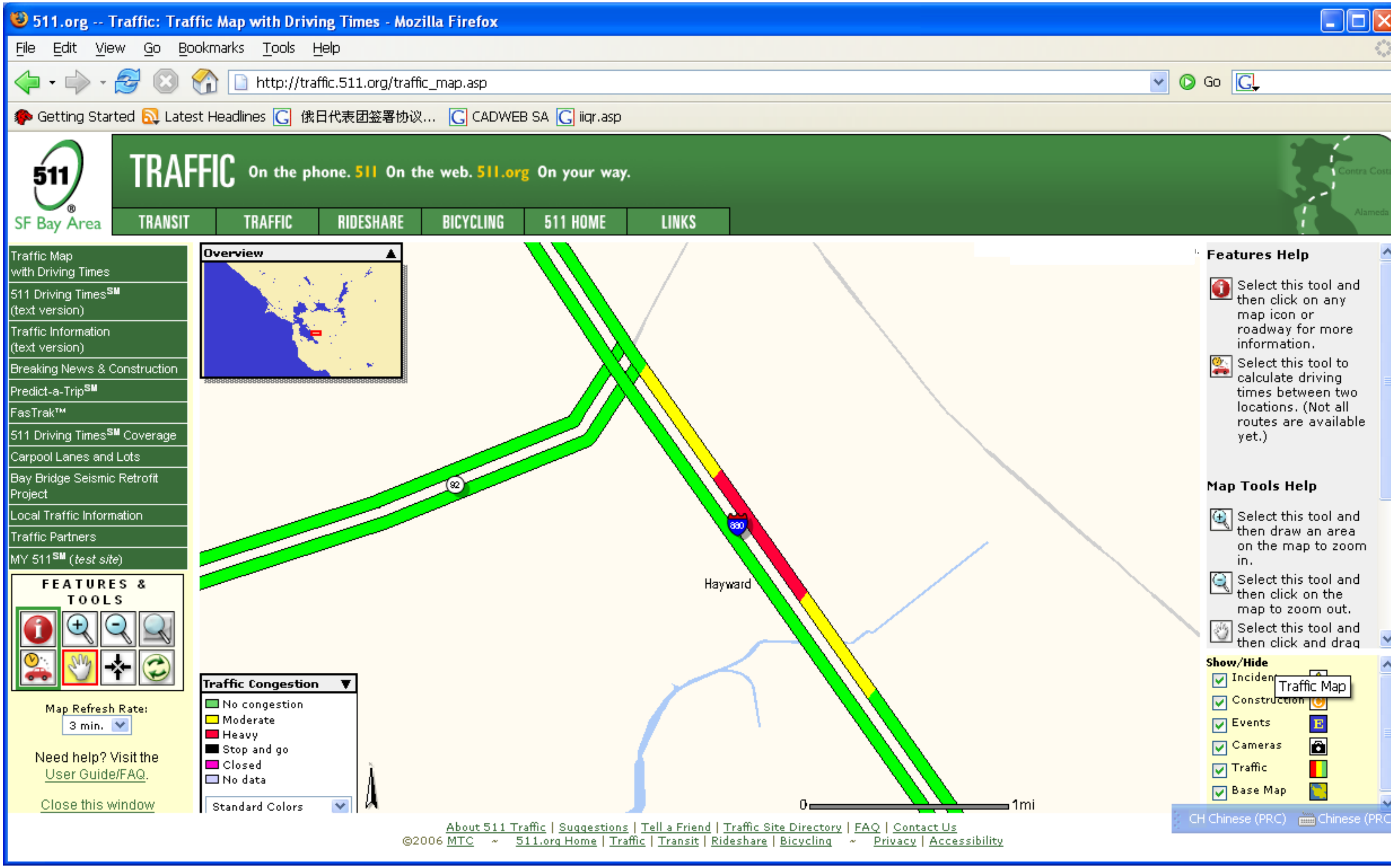
Union landing parking lot, operation is running

UC Berkeley



Mobile Century validation (1/2)

Mobile Century validation (2/2)



Summary

- Convergence of multi-media, sensing and communication platforms
  - Velocity sensing becomes possible everywhere
  - In 18 months, most standard phones will have GPS
  - Mobility tracking becomes possible

- What is the future of such a system?
  - Mobility tracking, location based services:
    - Commuters
    - Multi-modal transport
    - Direct services from the phones to the phones
  - Data sources for operators, planners, data fusion with existing sources of data

- What are the important open problems?
  - Tradeoffs between privacy and safety
  - Flow based modeling of highway, arterials, roads, city traffic
  - Personalized travel time computation

Team: UC Berkeley

LAGRANGIAN SENSOR SYSTEMS LABORATORY

- Lagrangian Sensor Systems Laboratory

  Christian Claudel
  Juan Carlos Herrera
  Coralie Claudel
  Andrew Tinka
  Patrick Saint-Pierre
  Xavier Litrico
  Saurabh Amin
  Sebastien Blandin
  Jean-Severin Deckers

Team: UC Berkeley CCIT

- California Center for Innovative Transportation (CCIT)

  Tom West
  JD Margulici
  Ali Mortazavi
  Marika Benko
  Steve Andrews
  Ryan Herring
  Osama Elhamshary
  Tia Dodson
  Chris Flens-Batina
  Manju Kumar
  Jed Arnold
  Jeff Ban
  Xiaohong Pan
  Erica Sherlock-Thomas
  Arthur Wiedmer
  James Lew

Team: Nokia, Rutgers, UC Berkeley ITS

- Nokia Team
  - Quinn Jacobson
  - Toch Iwuchukwu
  - Dave Sutter
  - Ken Tracton
  - John Loughney
  - John Paul Shen
- Rutgers Team
  - Marco Gruteser
  - Baik Hoh
- Institute of Transportation Studies
  - Steve Campbell
  - Margaret Chang
  - Jillene Bohr
  - John Li
  - Meriel Ennik
  - Derek Johnson

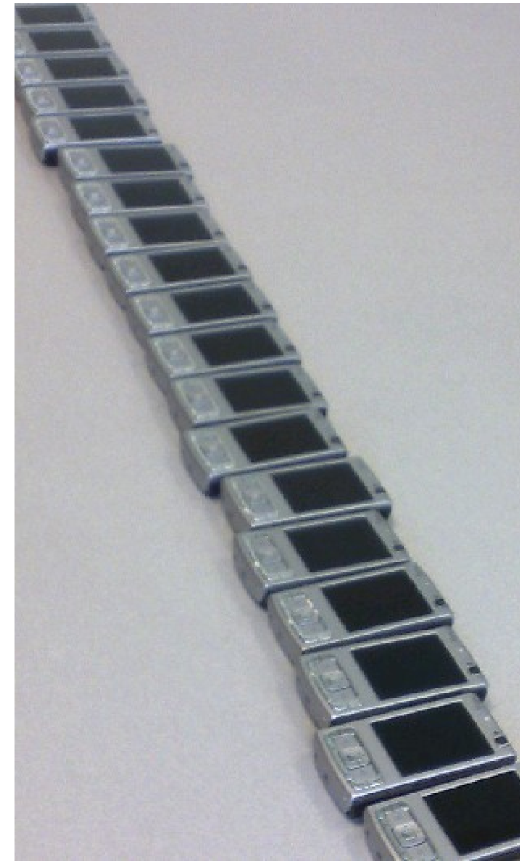
Team: Officers

- Officers

Anthony Patire
Emma Strong
Jason Wexler
Jean Parks
Jennifer Chang
Kristen Ray
Negin Aryaee
Timmy Siauw
Anurag Sridharan
Arthur Wiedmer
Madeline Ziser
Matthew Vaggione
Qingfang Wu
Tarek Rabbani
Josh Pilachowski

Kristen Parrish
Megan Smirti
Carl Misra
Christina Sedighi
Jessica Ariani
Elizabeth Kincaid
Sandy Do
Nick Semon
Trucy Phan
Timothy Racine
Alan Wang
Charlotte Wong
Irene Kwan
Karl David Cruz
Swe Shin Maung
Tyler Moser
Alexis Clinet
Julie Percelay

CCIT team

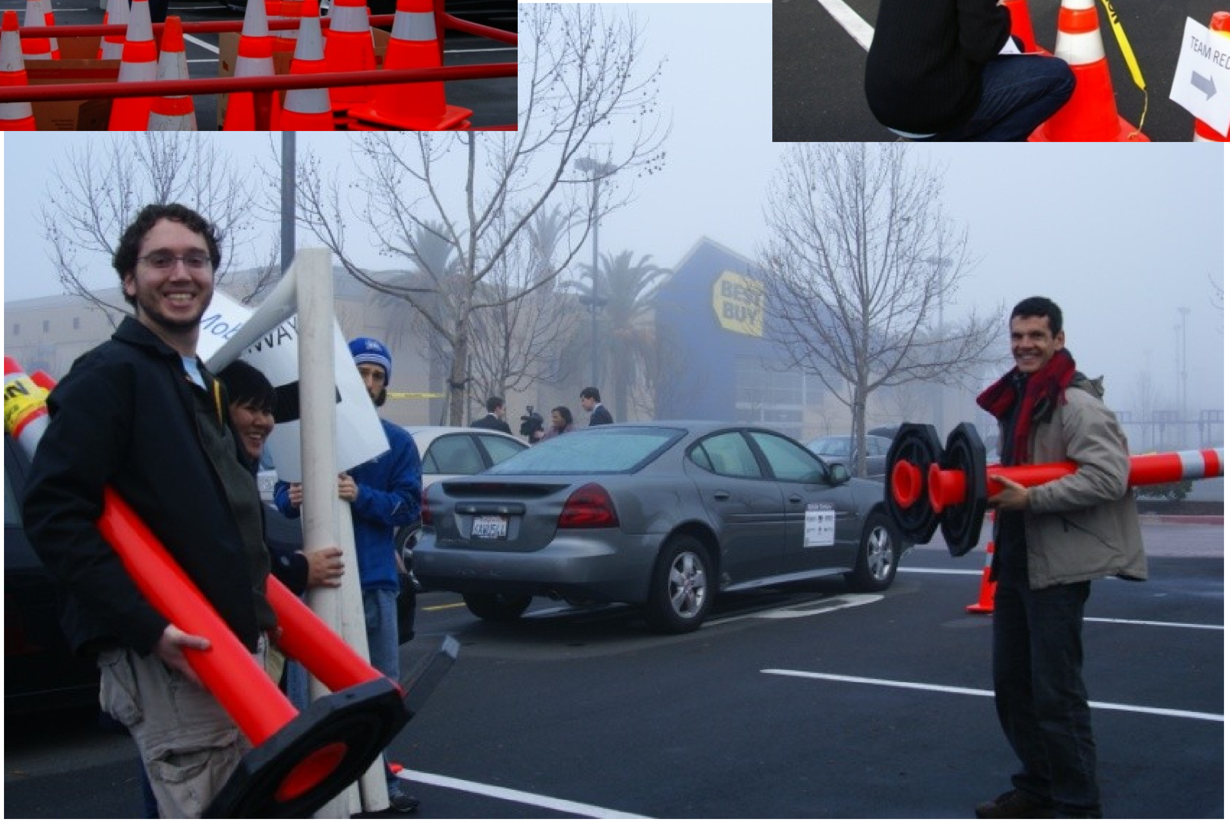


Caltrans team: Thursday 4:00pm

Union landing parking lot, Friday 7:00am

UC Berkeley

Union landing parking lot, Friday 7:15am



Union landing parking lot, operation is running

Union landing parking lot, operation is running

Caltrans, Nokia, Berkeley, CCIT