Tutorial: Traffic Management and AI

Biplav Srivastava, Anand Ranganathan
IBM Research

Toronto, Canada
What to Expect: Tutorial Objectives

- The aim of the tutorial is to make early and experienced researchers aware of the traffic management area, provide
  - an insightful overview of the current efforts using AI techniques
    - in Research and
    - in practice (real world pilots), and
  - whet interest for newer efforts on important open issues.

- From the call for tutorial: “a second type of tutorial provides a broad overview for an AI area that potentially crosses boundaries with an interesting application area”.

- **Disclaimer:** we are only providing a sample of the traffic management space intended to match audience profile in the available time.
Outline

1. Traffic Management Problem
2. Instrumentation
   1. Sensing traffic
   2. Traffic state estimation
   3. Optimizing and combining sensor data
3. Interconnection
   1. Middleware
   2. Traffic standards
4. Intelligence
   1. Path planning
      1. Simple Illustration
      2. Path Planning for Individual Vehicles
   2. End-user analytics
      1. Bus arrival prediction and journey planning, with state-of-art instrumentation
      2. Multi-modal journey planning, without sensors
5. Supporting topics
   1. Traffic Simulators
   2. Practical considerations for real-world pilots
Acknowledgements

All our collaborators, and especially those in:

- City agencies around the world
  - Bengaluru, India
  - Boston, USA
  - Dublin, Ireland
  - Ho Chi Minh City, Vietnam
  - New Delhi, India
  - New York, USA
  - Stockholm, Sweden

- Academia

- IBM: Many including - Raj Gupta, Ullas Nambiar, Srikanth Tamilselvam, L V Subramaniam, Chai Wah Wu, Anand Paul, Milind Naphade, Jurij Paraszczak, Wei Sun, Laura Wynter, Olivier Verscheure, Eric Bouillet, Francesco Calabrese, Tsuyoshi Ide, Xuan Liu, Arun Hampapur, Nithya Rajamani, Vivek Tyagi, Raguram Krishnapuram, Shivkumar Kalyanraman, Manish Gupta, Niterndra Rajput, Krishna Kummamuru, Raymond Rudy, Brent Miller, Jane Xu, Steven Wysmuller, Alberto Giacomel, Vinod A Bijlani, Pankaj D Luna, Tran Viet Huan, Wei Xiong Shang, Chen WC Wang, Bob Schloss, Rosario Usceda-Sosa, Anton Riabov, Magda Mourad

For discussions, ideas and contributions. Apologies to anyone unintentionally missed.
Traffic Management and AI

Section: Traffic Problem

Speaker: Biplav Srivastava
Outline

1. Traffic Management Problem

2. Instrumentation
   1. Sensing traffic
   2. Traffic state estimation
   3. Optimizing and combining sensor data

3. Interconnection
   1. Middleware
   2. Traffic standards

4. Intelligence
   1. Path planning
      1. Simple Illustration
      2. Path Planning for Individual Vehicles
   2. End-user analytics
      1. Bus arrival prediction and journey planning, *with state-of-art instrumentation*
      2. Multi-modal journey planning, *without sensors*

5. Supporting topics
   1. Traffic Simulators
   2. Practical considerations for real-world pilots
## We All See Traffic Daily. An Illustration from Across the Globe

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>New York City, USA</th>
<th>New Delhi, India</th>
<th>Beijing, China</th>
<th>Moscow, Russia</th>
<th>Ho Chi Minh City, Vietnam</th>
<th>Sao Paolo, Brazil</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 How is traffic predominantly managed</td>
<td>Automated control, manual control</td>
<td>Manual control</td>
<td>Automated control, manual control</td>
<td>Automated, manual control</td>
<td>Manual control</td>
<td>Automated, manual control, Rotation system (# plate based)</td>
</tr>
<tr>
<td>2 How is data collected</td>
<td>Inductive loops, cops, video, GPS</td>
<td>Traffic surveys, cops</td>
<td>Video, GPS, cops</td>
<td>GPS, some video, cops</td>
<td>Traffic surveys, cops</td>
<td>Video, GPS, cops</td>
</tr>
<tr>
<td>3 How can citizens manage their resources</td>
<td>GPS devices, alerts on radio, web, road signs (variable)</td>
<td>Alerts on radio</td>
<td>alerts on radio, road signs (variable), mobile alerts</td>
<td>GPS, radio, road signs, mobile alerts</td>
<td>Alerts on radio</td>
<td>GPS devices, alerts on radio, web</td>
</tr>
<tr>
<td>4 Traffic heterogeneity by vehicle types</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>5 Driving habit maturity (Low: &lt;10 yrs; Medium: 10-20; High: &gt; 20)</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>6 Traffic movement</td>
<td>Lane driving</td>
<td>Chaotic</td>
<td>Lane driving</td>
<td>Lane driving</td>
<td>Chaotic</td>
<td>Lane Driving</td>
</tr>
</tbody>
</table>

Source: Google map for New York City and New Delhi; Search done on Aug 20, 2010
Is it Just Supply-Demand Mismatch?

- **Supply**
  - Roads and their capacity
  - Personnel available
  - Capital and operational budget

- **Demands**
  - Travel needs of citizens
  - Travel needs of organizations: businesses, governments

- **Solution to mis-match?**
  - Keep building supply (roads, bridges, …)
  - Keep reducing demand (restrict citizens, businesses, …)
  - **May work** for some of the cities, for some of the time,
    - but not for all of the cities and all of the time
Levers in Cities’ Hands for Traffic Management

- Physical infrastructure
  - New flyovers, roads
  - Expanding existing capacity
  - New metros, ...

- Policies
  - Relocating businesses
  - Incentivising public transport

- IT enabled technologies
  - Intelligent Traffic/ Transportation Systems
  - Social networking
Which Mode of Transport is Best for a Particular Distance – Door-to-Door Trip Times

(a) 3 km trip

(b) 6 km trip

(c) 12 km trip

(d) 24 km trip

Car (car + motorcycle) takes less time regardless of distance unless there is congestion on road.

Data used:
- International data on things like access time to public transportation was taken.
- Studies were done for Delhi Metro.
- Minimum assumptions were made.
- E.g., walking time for metro access (5 mins) is widely acceptable.
- See speaker notes too.

Why Focus on Mere Congestion Leads to Anomalies

- The Pigou-Knight-Downs paradox
  - States that adding extra road capacity to a road does not reduce travel time.
  - This occurs because traffic may simply shift to the upgraded road from the other, making the upgraded road more congested.

- The Downs-Thomson paradox
  - States that the equilibrium speed of car traffic on the road network is determined by the average door-to-door speed of equivalent journeys by public transport.
  - This occurs when the shift from public transport causes a disinvestment in the mode such that the operator either reduces frequency of service or raises fares to cover costs.

- The Braess' paradox
  - States that adding extra capacity to a network, when the moving entities selfishly choose their route, can in some cases reduce overall performance and increase the total commuting time.

Details:
Chengri Ding and Shunfeng Song, Paradoxes of Traffic Flow and Congestion Pricing.
Pigou-Knight-Downs Paradox

- Bridge is the direct route between A-B, while highway is the circuitous route.
- The average travel time on the (congested) bridge is a linear function of the flow-to-capacity ratio and the average travel time on the uncongested highway is a constant.
- $a$, $b$, and $d$ are positive parameters with $d > a$; At equilibrium, travel times are same on both routes.
- Paradox exists for any bridge capacity less than $C_1$, because expanding the bridge will only shift travelers to the bridge but not reduce travel time.
- Using $a=10$, $b=10$, $d=15$, and $F=1,000$, Arnott and Small (1994) showed that the average travel time remains at 15 even if the bridge capacity continues to increase until it reaches 2,000, which is twice as large as the total traffic volume.
- How to address paradoxes
  - Car commuters ignore how much delay they cause on other travelers but only pay attention to how long it takes them to commute.
  - Delay cost to others must be quantified and accounted for as social cost.
  - If social cost is introduced as tolls, the paradox goes away.

Details:
Chengri Ding and Shunfeng Song, Paradoxes of Traffic Flow and Congestion Pricing,
Key Insights for a New Traffic Problem Look

- **Two category of stake-holders putting their resources**
  - Public resources
    - Example: $1M per month for a city with 5M population
    - Example: $100M available for roads and bridges for the next 3 years.
  - Private resources (often ignored)
    - Example: Average time taken to travel 1 KM
    - Example: Average $ needed to travel 1 KM

- **Impact of time-frame**
  - In short-term, new physical supply cannot be created due to long construction time, time to hire personnel, etc
  - Public resources have a short-term and a long-term cycle, while private resources for transportation need is time-frame insensitive

- **Approach: optimize both public as well as private resources**
Problem Statement for Traffic Management from Engineering Perspective

- **Problem (Short Term)**
  
  *Match traffic demand to supply with optimal usage of available public resources and concomitant optimization of citizens’ private resources for travel needs*

  - Example: For a given day,
    - minimize over-time payments to traffic personnel, while
    - minimizing average commute time per km.
  - Alternative: Service objectives can also be stated like average commute time per km be below 10 mins/km.

- **Implications**: Conflicting goals: optimality of public resources (e.g., road, traffic cops) v/s optimality of individual’s resources (e.g., travel time)

---

Problem Statement for Traffic Management from Engineering Perspective

- **Problem (Short Term)**
  
  *Match traffic demand to supply with optimal usage of available public resources and concomitant optimization of citizens’ private resources for travel needs*

  - Example: For a given day, minimize over-time payments to traffic personnel while minimizing average commute time per km. Service objectives can also be stated like average commute time per km be below 10 mins/ km.
  - **Implications**: Conflicting goals: optimality of public resources (e.g., road, traffic cops) v/s optimality of individual’s resources (e.g., travel time)

- **Problem (Long Term)**

  *For a known future period, match traffic demand to supply with optimal usage of available and planned public resources and concomitant optimization of citizens’ private resources for travel needs.*

  - Example: For the next 3 years, minimize city’s expenses (annual operational budget and capital investment in the period) while minimizing average travel time per km and average fare per km. Service objectives can also be stated towards citizens like average commute time be below 10 mins/ km and average fare be below $1/ km.

  - **Implications**: Information is used to plan city’s regions (including businesses, communities) and roads, thereby influencing traffic patterns

Starting Point from Problem Statement for Traffic Management

- **Problem (Short Term)**
  
  *Match traffic demand to supply with optimal usage of available public resources and concomitant optimization of citizens’ private resources for travel needs*

  - Example: For a given day, minimize over-time payments to traffic personnel while minimizing average commute time per km. Service objectives can also be stated like average commute time per km be below 10 mins/ km.
  
  - **Implications:** Conflicting goals: optimality of public resources (e.g., road, traffic cops) v/s optimality of individual’s resources (e.g., travel time)

- **Problem (Long Term)**
  
  *For a known future period, match traffic demand to supply with optimal usage of available and planned public resources and concomitant optimization of citizens’ private resources for travel needs.*

  - Example: For the next 3 years, minimize city’s expenses (annual operational budget and capital investment in the period) while minimizing average travel time per km and average fare per km. Service objectives can also be stated towards citizens like average commute time be below 10 mins/ km and average fare be below $1/ km.

  - **Implications:** Information is used to plan city’s regions (including businesses, communities) and roads, thereby influencing traffic patterns

- **Starting point for any solution requires traffic to be measured accurately at sustained, continuous basis, at a rate that helps effective traffic control**

**Details:** Biplav Srivastava. A new look at the traffic management problem and where to start. In 18th ITS Congress, Orlando, USA, Oct 16-20, 2011.
### Consulting Solution Approach

**A Summary of where Global Cities are within the Transportation Maturity Model**

<table>
<thead>
<tr>
<th>Level 1 Silo</th>
<th>Level 2 Centralized</th>
<th>Level 3 Partially Integrated</th>
<th>Level 4 Multimodal Integrated</th>
<th>Level 5 Multimodal Optimized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning</td>
<td>Functional Area Planning (single mode)</td>
<td>Project-based Planning (single mode)</td>
<td>Integrated system-wide planning (multi-mode)</td>
<td>Integrated regional multimodal planning</td>
</tr>
<tr>
<td>Performance Measurement</td>
<td>Minimal</td>
<td>Defined metrics by mode</td>
<td>Limited across silos</td>
<td>Shared multimodal system-wide metrics</td>
</tr>
<tr>
<td>Customer Management</td>
<td>Minimal capability, no customer accounts</td>
<td>Customer accounts managed separately for each system/mode</td>
<td>Multi-channel access interaction per mode</td>
<td>Integrated multimodal incentives to optimize multimodal use</td>
</tr>
<tr>
<td>Data Collection</td>
<td>Limited or Manual Input</td>
<td>Near real-time for major routes</td>
<td>System-wide real-time data collection across all modes</td>
<td>System-wide real-time data collection across all modes</td>
</tr>
<tr>
<td>Data Integration</td>
<td>Limited</td>
<td>Networked</td>
<td></td>
<td>Extended integration</td>
</tr>
<tr>
<td>Analytics</td>
<td>Ad-hoc analysis</td>
<td>Periodic, Systematic analysis</td>
<td>Detailed analysis in real-time</td>
<td>Multi-modal analysis in real-time</td>
</tr>
<tr>
<td>Payment Methods</td>
<td>Manual Cash Collection</td>
<td>Automatic Cash Machines</td>
<td>Electronic Payments</td>
<td>Multimodal, multimedia (fare cards, cell phones, etc)</td>
</tr>
<tr>
<td>Incident Management</td>
<td>Manual detection, response and recovery</td>
<td>Manual detection, coordinated response manual recovery</td>
<td>Automated pre-planned multimodal recovery plans</td>
<td>Dynamic multimodal recovery plans based on real-time data</td>
</tr>
<tr>
<td>Demand Management</td>
<td>Individual static measures</td>
<td>Individual measures with long term variability</td>
<td>Dynamic pricing</td>
<td>Multimodal dynamic pricing</td>
</tr>
<tr>
<td>Traveler Information</td>
<td>Static Information</td>
<td>Static trip planning limited real-time information</td>
<td>Location-based, on-journey multimodal information</td>
<td>Location-based, multimodal proactive re-routing</td>
</tr>
</tbody>
</table>

**Multimodal Network Management Maturity Model version 1.1**

© Copyright IBM Corporation 2007

An Intelligent Transportation Integration & Analytics Framework

Collection
- Mapping Services (Creation and Maintenance of maps etc)
- In-Vehicle Detectors (Probes etc)
- Infrastructure Detectors (Loops, CCTV, Tag & Beacon etc)
- Third Party Information Feeds (e.g. Floating Cellular Data)
- Mass Transit Operational Management (timetables, etc)

Integration & Analysis
- Data Fusion (Creation of Single View)
- Data Analytics (Forecasting, Journey Time Determination)
- Data Warehouse (Reporting, Archiving)
- Geographic Information System
- Vehicle Identification
- CRM (Call Centre, Interactive Voice)
- Pricing Models
- Security

Command & Control
- Traffic Control Center Dashboard
- Incident Management (Process Execution)
- Route Guidance & Trip Planning
- ITS Asset Management, Maintenance & Optimization

Dissemination
- Roadside Infrastructure (Variable Message Signs, Traffic Signals etc)
- Traveller Information Portal (Info Display, Route Planning)
- Traveller Information Gateway (B2B info feeds etc)
- Self-service Traveler Portal (view bill, pay charge, etc)
- External Systems (Vehicle Licensing, Electronic Payments, Bill Printing, etc)

Integration
- Transport Management Subsystems (Traffic Signals, Bridges, Parking, Mgt, etc)

Systems Management

© 2010 IBM Corporation
References


- A new look at the traffic management problem and where to start, by Biplav Srivastava, In 18th ITS Congress, Orlando, USA, Oct 16-20, 2011.


- Chengri Ding and Shunfeng Song, Paradoxes of Traffic Flow and Congestion Pricing,
Traffic Management and AI

Section: Traffic Sensing

Speaker: Biplav Srivastava

Toronto, Canada
Outline

1. Traffic Management Problem

2. Instrumentation
   1. Sensing traffic
   2. Traffic state estimation
   3. Optimizing and combining sensor data

3. Interconnection
   1. Middleware
   2. Traffic standards

4. Intelligence
   1. Path planning
      1. For individual vehicles
      2. For public transportation
   2. End-user analytics
      1. Bus arrival prediction and journey planning, *with state-of-art instrumentation*
      2. Multi-modal journey planning, *without sensors*

5. Supporting topics
   1. Traffic Simulators
   2. Practical considerations for real-world pilots
A Feel For Traffic Sensing

**Situation:** Sensor readings from different types, various locations; for some links, no sensor present

**Problem:** What is the overall view of traffic?

<table>
<thead>
<tr>
<th>No.</th>
<th>Data type</th>
<th>Format</th>
<th>Ability for using in sensing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Collected by manual method</td>
<td>Document</td>
<td>Can be used right after they are integrated, cleaning and conversion</td>
</tr>
<tr>
<td>2</td>
<td>Mobile data</td>
<td>CDR</td>
<td>Data will be converted to transport data in standard format during project implementation</td>
</tr>
<tr>
<td>3</td>
<td>Data from mobile which have GPS device</td>
<td>Follow format of data traffic</td>
<td>Data will be collected directly during project deployment and switch to the format needed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Label</th>
<th>Sensor Type</th>
<th>Data Format</th>
<th>Cost</th>
<th>Accuracy</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Manual, GPS</td>
<td>Document</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>2</td>
<td>Video</td>
<td>Image</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>3</td>
<td>Call Data Record (mobile)</td>
<td>Binary</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
</tbody>
</table>

© 2010 IBM Corporation
Sensed Traffic Metrics

- Traffic flow – number of vehicles* that pass a certain point during a specified time interval (vehicles/hour)
- Speed – rate of motion in distance/time (mph)
- Density – number of vehicles occupying a given length of highway or lane (vehicles per mile per lane)
- Spacing – the distance between successive vehicles in a traffic stream as they pass some common reference point on the vehicles
- Time headway – the time between successive vehicles in a traffic stream as they pass some common reference point on the vehicles

*What is a vehicle? There are 48 types in India! So measurement can be non-trivial!
Traffic Flow and Time Headway

Traffic Flow given by:

\[ q = \frac{n}{t} \]

q = traffic flow in vehicles per unit time

n = number of vehicles passing some designated roadway point during time interval t

t = duration of time interval
Time Mean Speed

- Measure of speed at a certain point in space

Arithmetic mean of vehicles speeds is given by:

\[
\bar{u}_t = \frac{\sum_{i=1}^{n} u_i}{n}
\]

\(u_t\)=time-mean speed in unit distance per unit time
\(u_i\)=spot speed of the ith vehicle
\(n\)=number of measured vehicle spot speeds
Space Mean Speed

- Time necessary for a vehicle to travel some known length of roadway

\[ \bar{u}_s = \frac{l}{\bar{t}} \]

\( u_s \) = space-mean speed in unit distance per unit time

\( l \) = length of roadway used for travel time measurements of vehicles

\( t \) = vehicle travel time

\[ \bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i \]

\( t_i \) = time necessary for vehicle \( i \) to travel a roadway section of length \( l \)
Evolution of Traffic Flow Sensor Technology

Goal: Get traffic information (speed, volume) for city region on sustained, continuous basis, at a rate that helps effective traffic control at low cost

Latest Buzz:
- Social Media
- Mobile phones

Source: IBM China Research Lab
Conventional Sensing Approach

- Take whatever sensors are in place and in immediate control of the concerned department, ignoring other sensing data

Sample References

Technology Complementarity

Inductive-loop detector

- Analysis easiness
- Rich traffic data
- Network-wide information
- Resource consolidation
- Low operational cost
- Low capital cost
- Trustiness (legal)

Video image processor

- Analysis easiness
- Rich traffic data
- Network-wide information
- Resource consolidation
- Low operational cost
- Low capital cost
- Trustiness (legal)

Floating car data

- Analysis easiness
- Rich traffic data
- Network-wide information
- Resource consolidation
- Low operational cost
- Low capital cost
- Trustiness (legal)

Mobile traffic probe

- Analysis easiness
- Rich traffic data
- Network-wide information
- Resource consolidation
- Low operational cost
- Low capital cost
- Trustiness (legal)

Source: IBM China Research Lab
Illustration of Some Sensor Types
Induction Loop System

Detector

Loop Extension Cable

Loop

Device to be Triggered by Detector

Magnetic Field

Freq

Time
Method to calculate the average speed of the links:

\[ V(t) = g(t) \times \frac{c(t)}{o(t) \times T} \]

Reliability of method depends on the estimated value for \( g \)

\[ g(t) = g_{traffic}(t) + g_{detector} \]
Video-based Vehicle Counting

Traffic counting video analytics demo

Left: 3
Right: 17

http://www.youtube.com/watch?v=KWobYJzQI_k
Available Mobility Data

1. Passive collection:
   - standby phones
   - Location Update (LU) (Location Update)
   - Location Area Dimension: 3~10km

2. Active collection:
   - on-call phones
   - Continuous comm. behaviors (SMS/MMS/Call) and HO (Handover)
   - Location Area Dimension: 100m~2km
   - AOA, TDOA & TOA
   - E-OTD & A-FLT
   - Positioning Techniques
     - Accuracy: 50~200m

Network Based

Terminal Based

3. GPS Enabled Phones:
   - Accuracy: 5~30m

4. Other Software Installed:
   - Accuracy: 50m~2km

Need to upload the Data
Power Consumption Issue

Source: IBM China Research Lab

Road-side Acoustics based Sensing

- Honking
- Road sounds
  - Engine
  - Road
  - Air
  - ...

© 2010 IBM Corporation
Honk-Ok-Please [IIT Bombay]

Summary

- Can detect speed with high accuracy (relative error < 20% in most cases)
- Can detect congestion v/s free-flow
- Results heavily dependent on positioning of sensors (recorders),
Traffic State: Audio spectrograms of various traffic density states, free flow, medium and jammed, and their corresponding cues

Road-side cumulative acoustics composed of various noise signals for eg: tire noise, engine noise, engine idling noise, occasional honks, air turbulence noise;

Traffic density state determines combination of noise signals. Therefore use the spectral modulation acoustic features along with the Gaussian mixture Hidden Markov models to statistically characterize these traffic density states.

Initial experimentation have established the feasibility and accuracy of using this approach.

<table>
<thead>
<tr>
<th>Traffic Class</th>
<th>T=5s</th>
<th>T=10s</th>
<th>T=15s</th>
<th>T=20s</th>
<th>T=25s</th>
<th>T=30s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jammed</td>
<td>95.0</td>
<td>97.7</td>
<td>98.3</td>
<td>99.5</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Medium-Flow</td>
<td>96.1</td>
<td>98.5</td>
<td>99.0</td>
<td>99.4</td>
<td>99.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Free-Flow</td>
<td>93.5</td>
<td>97.8</td>
<td>99.0</td>
<td>99.6</td>
<td>99.8</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Classification accuracy using MFCC features frames covering T sec of audio, SVM classifier
Sensor Subset Selection and Optimization
Problem Considered

- City authorities need to decide what sensors to use to get traffic data for traffic of their region
  - Slew of techniques available varying in accuracy, coverage and cost to install and maintain
  - Diversity in how they can be set up
  - How sensing methods may complement each other

- City can make an initial decision but it will need to be re-visited over time
  - Traffic patterns changes
  - Sensing technologies changes

**Problem**: find the subset of sensors from available types that give the best cost-benefit for a given traffic pattern
Effect of Increasing Number of Sensors on RMSE

[Traffic Pattern 1 on a Grid] Effect of increasing number of sensors of different type.

a) Shows that RMSE decreases with increase in percentage of sensors from 10 to 100%.

b) Shows increase in cost is highest for manual sensors whereas it is lowest for CDR.
Subset Selection Approach

- Model sensor types based on cost, accuracy and coverage
- Create a sample space of sensor combination choices
- Use a traffic simulator (MATSIM)
  - To measure the sensing error distribution entailed in each sensor combination
  - To ensure physical characteristics of the city are taken into account
- Choose Pareto sensor combinations (non-dominated); Call this “Optimal Candidate Set (OCS)”
- Optional filtering steps
  - Remove combinations above a give cost threshold
  - Remove combinations above an error threshold


- For a given set of ‘k’ optimal combinations to be returned
  - Select a preference function
  - Use OCS selection using ICP approximation (Carlyle et al)
- Return ‘k’ optimal sensor combinations

© 2010 IBM Corporation
Examples of City Preferences in Sensor Selection

- **(Case A):** A city which already has some sensors in place would want to add only those new sensor types which complement its investments
  - Maximize incremental accuracy improvement
- **(Case B):** A city building the ITS infrastructure for the first time would want the highest accuracy possible within its budget
  - Highest accuracy
- **(Case C):** A city may have low budgets and would want to have high coverage even at the loss of some accuracy (which it may compensate by putting more people on the field)
  - Minimize cost
- **(Case D):** A city may want to minimize operational costs, even trading it off with higher capital expenditure, possibly because they come from a different budget
  - Minimize operational cost
Key Results from Sensing

- Low-cost, noisy, CDR-based sensing complements existing sensors in a city due to easy coverage it provides
  - Highly beneficial when other, more accurate, sensor types are few
  - Benefit diminishes as the number of more accurate sensors increases
  - Important result for traffic management of emerging countries

- Sensor type combinations can be suggested for a city’s traffic patterns using a simulator balancing accuracy and cost of sensing
  - Returns non-dominated results
  - Return few results approximating the search space
Extracting Real City Characteristics for Simulation to Choose Sensors
Combining Multiple Sensor Data

An example from Boston, USA completed in June 2012
Traffic Context in Boston: Background

Consumers

- Urban planning users
  - Needs counts done annually (max) or more frequently
  - Currently uses 10 year old data

- Environment users
  - Needs counts every 15 mins (min) or longer time

- Transportation users
  - Collects and uses counts every 15 mins

- Others
  - Residents: Consume for their personal choices
  - Businesses: identify new opportunities

Data Sources for Traffic Count

- Data in hand
  - Inductive loop detectors (Transportation asset)
  - Manual count by consultants during urban development (Urban planning asset)
  - Tube count (by state authorities)

- Future
  - Video camera (multiple organizations)
  - GPS (city vehicles)
  - Directly by citizens (e.g., Citizen Connect)
  - Social media

All consumers demand quality over quantity and frequency, since no sensor is perfect.
Model Consolidation

• Objectives
  • Make it simple to consume (traffic count) data*; easily provide content that is needed commonly by applications
  • Enable entitled users who want to see details to access them
  • Ease the process to import, process and manage data for IT staff

• Approach
  1. Identify requirements from relevant consumers, negotiate
  2. (Domain knowledge): know what information typical consumers (traffic and other domains) want
  3. Look at sample data
  4. Select attributes commonly of interest
  5. Have additional attributes for data quality, provenance and cost, as applicable
  6. Validate selected attributes with consumers and implementers
  7. Reconcile with applicable standards (NIEM, CAP, TMDD, Datex2) and enhance description
  8. Sign-off: Freeze for development and create documentation

* Set expectation for data quality and error checking
Common Model

Standards Aligned, Uniform format, Uniform Error Semantics

A Snapshot of Common Model and Mapping to Data Sources
References

Traffic Management and AI

Section: Interconnection

Speaker: Anand Ranganathan
Outline

1. Traffic Management Problem
2. Instrumentation
   1. Sensing traffic
   2. Traffic state estimation
   3. Optimizing and combining sensor data

3. Interconnection
   1. Middleware
   2. Traffic standards

4. Intelligence
   1. Path planning
      1. Simple Illustration
      2. Path Planning for Individual Vehicles
   2. End-user analytics
      1. Bus arrival prediction and journey planning, *with state-of-art instrumentation*
      2. Multi-modal journey planning, *without sensors*

5. Supporting topics
   1. Traffic Simulators
   2. Practical considerations for real-world pilots
Challenges in Middleware

- Scalability
  - processing large volumes of real-time and static data

- Ability to incorporate various kinds of specialized as well as reusable analytics
  - Classifiers, forecast/prediction algorithms, simulators, spatio-temporal processing, image/video processing, etc.

- Extensibility
  - being able to add sources and data and new kinds of analyses on the data rapidly
  - support different kinds of one-time and continuous queries from diverse end-users
    - commuters, highway patrols, public service vehicles like re-engines and ambulances, departments of transportation, urban planners, commercial vehicle operators, etc.
Addressing these challenges on the IBM InfoSphere Streams platform

- Stream Processing
  - New data processing paradigm where large volume of real-time (sensor) data is processed on the fly, without storing in a database
  - Continuous queries (and processing)
    - Queries are long running, while data keeps moving

- Key Features of InfoSphere Streams
  - Parallel and high performance stream processing software platform
  - Analysis of structured and unstructured information
Data is being produced continuously in very large volumes in different domains.

- **Natural Systems**
  - Seismic monitoring
  - Wildfire management
  - Water management

- **Transportation**
  - Intelligent traffic management

- **Manufacturing**
  - Process control for microchip fabrication

- **Stock market**
  - Impact of weather on securities prices
  - Analyze market data at ultra-low latencies

- **Radio Astronomy**
  - Detection of transient events

- **Fraud prevention**
  - Detecting multi-party fraud
  - Real-time fraud prevention

- **Law Enforcement**
  - Real-time multimodal surveillance

- **Health & Life Sciences**
  - Neonatal ICU monitoring
  - Epidemic early warning system
  - Remote healthcare monitoring

© 2010 IBM Corporation
How Streams Works

→ continuous ingestion → continuous analysis

achieve scale by
partitioning applications into components
How Streams Works

→ continuous ingestion
→ continuous analysis

infrastructure provides services for
scheduling analytics across h/w nodes
establishing streaming connectivity

Transform
Annotate

Filter

Classify

Correlate

where appropriate,
elements can be “fused” together
for lower communication latencies

hieve scale
by partitioning applications into components
by distributing across stream-connected hardware nodes

© 2010 IBM Corporation
Application Programming

- Consumable
- Reusable set of operators
- Connectors to external static or streaming data sources and sinks

MARIO: Automated Application Composition

SPL: Stream processing dataflow scripting language

Platform Optimized Compilation

© 2010 IBM Corporation
SPL Language Highlights

- Intermediate language describing how a set of individual stream processing operators are connected to one another in a graph.

- Each individual operator has zero or more input and output streams:
  - A Built-In Operator provided by the SPL language
    - Filter, Aggregator, Join, Split, Barrier,
  - A User Defined Operator
    - Written in C++ or Java
    - Can be written in a schema independent manner to enhance reusability
  - A source or sink operator to read/write from/to external data sources

- Procedural constructs
  - To create distributed and parallel flow graphs

- Settings to influence placement of operators, fusion into processes, execution model, etc.
A simple example

```plaintext
[Application]
SourceSink trace

[Typedefs]
typedef id_t Integer
typedef timestamp_t Long

[Program]
# a source stream is generated by a Source operator – in this case tuples come from an input file
stream SenSource ( id : id_t, location : Double, light : Float, temperature : Float, timestamp : timestamp_t )
   := Source( ) [ "file:///SenSource.dat" ] {}

# this intermediate stream is produced by an Aggregate operator, using the SenSource stream as input
stream SenAggregator ( id : id_t, location : Double, light : Float, temperature : Float, timestamp : timestamp_t )
   := Aggregate( SenSource <count(100),count(1)> ) [ id . location ]
      { Any(id), Any(location), Max(light), Min(temperature), Avg(timestamp) }

# this intermediate stream is produced by a functor operator
stream SenFunctor ( id : Integer, location : Double, message : String )
   := Functor( SenAggregator ) [ log(temperature,2.0)>6.0 ]
      { id, location, “Node ”+toString(id)+ “ at location ”+toString(location) }

# result management is done by a sink operator – in this case produced tuples are sent to a socket
Nil := Sink( SenFunctor ) [ “cudp://192.168.0.144:5500/” ] {}
```
Infosphere Streams Runtime Services

Optimizing scheduler assigns operators to processing nodes, and continually manages resource allocation.

Runs on commodity hardware – from single node to blade centers to high performance multi-rack clusters.
Infosphere Streams Runtime Services

*Optimizing scheduler assigns operators to processing nodes, and continually manages resource allocation*

*Adapts to changes in resources, workload, data rates*

*Runs on commodity hardware – from single node to blade centers to high performance multi-rack clusters*

*Capable of exploiting specialized hardware*
Real-Time GPS Data Processing – Map Matching

- Inputs:
  - \{GPS-id, Latitude, Longitude, Heading\} tuples
  - Collection of poly-lines (the map)

- Output:
  - \{GPS-id, Shape-id, distance\} tuple, representing the line-segment that is the best match for the GPS reading
  - Filter based on heading and distance

- Our technique: Grid-Based Decomposition
Scaling Map-Matching in InfoSphere Streams - how can it be done?

Given those operators, building and extending the application is just a matter of wiring those operators.

© 2010 IBM Corporation
### Results

<table>
<thead>
<tr>
<th>No. of Sources</th>
<th>Map Size (# links)</th>
<th>CPU %</th>
<th>Throughput (#GPS points mapped)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200M</td>
<td>10</td>
<td>310K</td>
</tr>
<tr>
<td>2</td>
<td>200M</td>
<td>16</td>
<td>589K</td>
</tr>
<tr>
<td>3</td>
<td>200M</td>
<td>20</td>
<td>777K</td>
</tr>
<tr>
<td>4</td>
<td>200M</td>
<td>27</td>
<td>1003K</td>
</tr>
<tr>
<td>1</td>
<td>500M</td>
<td>14</td>
<td>301K</td>
</tr>
<tr>
<td>2</td>
<td>500M</td>
<td>24</td>
<td>582K</td>
</tr>
<tr>
<td>4</td>
<td>500M</td>
<td>37</td>
<td>964K</td>
</tr>
<tr>
<td>1</td>
<td>1000M</td>
<td>18</td>
<td>294K</td>
</tr>
<tr>
<td>2</td>
<td>1000M</td>
<td>33</td>
<td>596K</td>
</tr>
<tr>
<td>4</td>
<td>1000M</td>
<td>53</td>
<td>941K</td>
</tr>
</tbody>
</table>
Illustration: Streams for State Estimation
Recall: Elements of Traffic State

- Traffic flow – number of vehicles that pass a certain point during a specified time interval (vehicles/hour)
- Speed – rate of motion in distance/time (mph)
- Density – number of vehicles occupying a given length of highway or lane (vehicles per mile per lane, vpmpl)

- Spacing – the distance between successive vehicles in a traffic stream as they pass some common reference point on the vehicles
- Time headway – the time between successive vehicles in a traffic stream as they pass some common reference point on the vehicles
Using GPS Data … Prototype for Stockholm

- Historic GPS data traces from Stockholm for the year 2008
  - 1500 taxis, 40 trucks
  - 170 million GPS points for the year
  - Each taxi produces a reading every 60 seconds
  - Trucks – every 30 seconds
  - Peak rate of about 1000 readings/min
  - Often replayed at a higher rate – over 100,000 readings per sec

- Road Network
  - 628,095 links, spread over a 80km x 80km area
Overall Application Description

Real-time GPS data processing
- Cleaning
- Map-Matching
- Per-Vehicle Speed Estimation

Real Time Transformation Logic → Real Time Geo Mapping → Real Time Speed & Heading Estimation → Real Time Aggregates & Statistics

Stochastic Link Travel Time Calculation → Interactive visualization

Splitter
- Shortest Path 1
- Shortest Path n

Web Server
- Google Earth

Offline statistical analysis

User-driven computations
- Travel times, Shortest Paths

Aggregated Statistics
- Per-link / per-region

Warehouse

© 2010 IBM Corporation
Real Time Geo Mapping & Speed Estimation

- GPS probe
- Matching map artifact
- Estimated path
- Estimated speed & heading

Estimated path and speed information:
- 2008-10-31 22:35:02
- GPS speed: 32 km/h
- Min average: 36.6 km/h
- GPS heading: 331
- Road heading: 316
- Avg heading: 298
Traffic Statistics Generation

- Aggregation and Inter-Quartiles Generation

```
stream roadAttributeStream(
    timestamp : Long,
    shapeid   : Long,
    timeIndex : Long,
    weekday   : Integer,
    monthId   : Integer,
    avgSpeed  : Float,
    minSpeed  : Float,
    maxSpeed  : Float,
)

:= Aggregate(sortedGeoTrackStream
    <attrib(abs(timeIndex, 01), perGroup)>
    [ timeIndex . shapeid ]

{,
    timestamp := Any(timestamp),
    shapeid   := Any(shapeid),
    timeIndex := Any(timeIndex),
    weekday   := Any(weekday),
    monthId   := Any(monthId),
    avgSpeed  := Avg(averageSpeed),
    minSpeed  := Min(minSpeed),
    maxSpeed  := Max(maxSpeed),
    estRoute  := Any(estroute)
}

-> partition["GpsAggregation"]
```

Output Stream Schema

Aggregation operation, specifying window boundaries and attributes to group by

Logic for computing fields in output schema

Specifies which execution container to run operation in. SPADE allows fusing multiple operators into a single execution process

© 2010 IBM Corporation
User Specific Computations

- Time Dependent Shortest Paths
- Stochastic Travel Times
- Stochastic Shortest Paths
ITS Application Flow-graph (125k GPS/second)
Showing Placement of Operators on 4 nodes
Scaling of throughput as we add more nodes

Average throughput (# GPS Points processed per second)

- Average throughput (at sources) : # tuples per second
Features of InfoSphere Streams / SPADE that helps development

- Component Based Programming Model
  - Applications developed as a graph of reusable operators
  - Operators can be built-in or user defined

- Modular Application Development
  - Applications can import and export streams
  - Allows composing entire applications

- Declarative Application Configuration within SPADE
  - Degree of parallelization
  - Fusion of operators
  - Placement of operators onto nodes

- Rich toolkits of available Operators
  - Relational, Geo-Spatial, Stream Mining, …
References


- Eric Bouillet, Anand Ranganathan, Scalable, Real-time Map-Matching using IBM’s System S, In 11th International Conference on Mobile Data Management (MDM 2010), Kansas City, MO, May 23-26, 2010
Traffic Management and AI

Section: Useful Standards (or Commonly Used)

Speaker: Biplav Srivastava
Outline

1. Traffic Management Problem
2. Instrumentation
   1. Sensing traffic
   2. Traffic state estimation
   3. Optimizing and combining sensor data
3. **Interconnection**
   1. Middleware
   2. **Traffic standards**
4. Intelligence
   1. Path planning
      1. Simple Illustration
      2. Path Planning for Individual Vehicles
   2. End-user analytics
      1. Bus arrival prediction and journey planning, *with state-of-art instrumentation*
      2. Multi-modal journey planning, *without sensors*
5. Supporting topics
   1. Traffic Simulators
   2. Practical considerations for real-world pilots
Outline

- Motivation: Why we are talking standards
- Overview of immediately relevant standards
- Recommendations on standards
Motivation: Assembling the Smarter City … today’s cities

The data/information, services, structures and activities in cities are described in different ways by the agencies which manage them, by the denizens which use them and by companies which provide them.

This cacophony of views and understanding drive many of the inefficiencies, create uneconomic costs in cities and limit the growth of city quality of life.

Research efforts ongoing to provide structures for the data, and the city infrastructure and services which reduce this confusion, significantly reduce costs and improve life in the city. Example: **SCRIBE** (IBM).
Motivation: Traffic

- All parts of a city affect each other – none can work in isolation
  - Sample: Traffic incidents may affect public safety, water management, environment monitoring.
  - Sample: External events may affect traffic, like a broken water pipe.

- Issue: how do participants share common information about the domain unambiguously?
  - E.g., congestion by traffic personnel, fire dept, electrical utility, mayor office, citizens, civil contractors, IT companies implementing Intelligent Transportation Systems
  - How are concepts related to the data actually being collected. E.g., congestion to (motorized) traffic count.
Motivation: Reducing every agency’s software to understanding at most 2 representations, not n representations.
<table>
<thead>
<tr>
<th>Standard</th>
<th>Examples of supported concepts</th>
<th>Current deployments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Traveler Information Systems (ATIS), SAE 2354</td>
<td>Messages defined for traveler information, trip guidance, parking, Mayday (emergency information)</td>
<td>International standard (SAE) with traction in North America Gaining momentum in the US; Nebraska, Washington and Minnesota planning support. See Washington State Department of Transportation Advanced Traveler Information Systems Business Plan for details.</td>
</tr>
<tr>
<td>DATEX II</td>
<td>Traffic elements, operator actions, impacts, non-road event information, elaborated data, and measured data</td>
<td>European standard Version I deployed in several European countries. DATEX II deployment appears to be gaining momentum in several European countries. See DATEX Deployments for details.</td>
</tr>
<tr>
<td>IEEE 1512</td>
<td>Common incident management message sets for use by emergency management centers, traffic management, public safety, hazardous materials, and entities external to centers</td>
<td>International standard Early deployments include Washington D.C., NYC, and Milwaukee. See 1512 Public Safety Early Deployment Projects for details.</td>
</tr>
<tr>
<td>National Transportation Communication for ITS Protocol (NTCIP) 1200 Series</td>
<td>Currently thirteen data dictionaries defined, including object definitions for: dynamic message signs, CCTV, camera control, ramp meter control, transportation sensor systems</td>
<td>U.S. specific standard According to NTCIP web site, several cities and states in the U.S. have NTCIP projects underway. See NTCIP Deployment: Projects for details.</td>
</tr>
<tr>
<td>Service Interface for Real Time Information (SIRI)</td>
<td>Information exchange of real-time information about public transportation services and vehicles</td>
<td>European specific standard Based on best practises of various national and proprietary standards from across Europe.</td>
</tr>
<tr>
<td>Traffic Management Data Dictionary (TMDD)</td>
<td>OwnerCenter, ExternalCenter, Device, DateTime, Dynamic Message Sign, Event, Generic, Organization, and RoadNetwork</td>
<td>International standard (ITE and AASHTO) with traction in North America In early stages of deployment in the US. TMDD is backed by US DoT and multiple vendors (Transcom, Siemens). Multiple state DoTs are planning to move to TMDD in their next rev (including Florida and Utah). See &quot;A Report to the ITS Standards Community ITS Standards Testing Program, For TMDD and Related Standards as Deployed by the Utah Department of Transportation&quot; for TMDD testing details from the Utah DoT.</td>
</tr>
</tbody>
</table>
Relevant Standards

- Information exchanged
  - Transportation and external agencies: Datex 2
  - Between agencies: NIEM

- Format of information given: CAP

- Public Transportation: GTFS

- In addendum: Road Conditions
  - Road Surface Conditions
  - Road types
  - Types of restriction
Information Exchanged

Information exchanged with DATEX II systems is composed of different basic elements:

• Road and traffic related events (called “Traffic elements”)
• Operator actions
• Advice
  – Different types of advice: speed, lane usage, winter driving …
• Impacts
• Non road event information
  – It concerns information about events which are not directly on the road: transit information, service disruption, car parks.
• Elaborated data (derived/computed data, e.g. travel times, traffic status)
• Measured data (direct measurement data from equipments or outstations, e.g. traffic and weather measurements)
Road Events (Elements)

Road and traffic related events (called “Traffic elements”)
These are all events which are not initiated by the traffic operator and force him to undertake (re)actions. They are classified in 6 main categories:
• Abnormal traffic (long queues, stop and go, …)
• Accidents - are events in which one or more vehicles lose control and do not recover. They include collisions between vehicle(s) or other road user(s), between vehicle(s) and any obstacle(s), or they result from a vehicle running off the road. Details can be given:
  – cause type,
  – accident type,
  – overview of people involved (number, injury status, involvement role, type),
  – overview of vehicles involved per type (number, status, type, usage),
  – details of involved vehicles (type + individual vehicle characteristics)
• Obstructions :
  – animal presence,
  – vehicles presence,
  – obstructions due to environment (avalanches, flooding, fallen trees, rock falls, …),
  – obstructions due to equipments (fallen power cables, …)
• Activities
• Incidents on infrastructure equipments (Variable message sign out of order, tunnel ventilation not working, emergency telephone not working, …)
• Specific conditions: road conditions related to weather (ice, snow, …) or not (oil, …), environment conditions (precipitation, wind, …), …
Operator actions

Operator actions are classified in 4 main categories:

- Network management: road closure, alternate traffic, contra flow …
  - traffic control: rerouting, temporary limits
- Roadworks: resurfacing, salting, grass cutting …
- Roadside assistance: vehicle repair, helicopter rescue, food delivery …
- Sign settings: variable message signs, matrix signs.
Impact

- Delays and traffic status (Impact) have 5 possible values:
  - free flow,
  - heavy,
  - congested,
  - impossible,
  - Unknown
Elaborated Data - These sets of data are normally derived on a periodic basis by the Traffic Centre from input data for specified locations

- Traffic values (normally published as measured data, but can be derived on a periodic basis and published as elaborated data):
  - flow,
  - speed,
  - headway,
  - concentration and
  - individual vehicle measurements.

- Weather values (normally published as measured data, but can be derived on a periodic basis and published as elaborated data):
  - precipitation,
  - wind,
  - temperature,
  - pollution,
  - Road surface condition and
  - visibility.
Measured Data - These data sets are normally derived from direct inputs from outstations or equipments at specific measurement sites (e.g. loop detection sites or weather stations) which are received on a regular (normally frequent) basis

- **traffic values:**
  - flow, speed, headway, concentration and individual vehicle measurements.

- **weather values:**
  - precipitation, wind, temperature, pollution, road surface condition and visibility.

- **travel times:**
  - elaborated time, free flow time, normally expected time

- **traffic status:**
  - free flow, heavy, congested, impossible, Unknown
National Information Exchange Model: Summary

- "NIEM provides a common vocabulary for consistent, repeatable exchanges of information between agencies and domains. The model is represented in a number of forms, including a data dictionary and a reference schema, and includes the body of concepts and rules that underlie its structure, maintain its consistency, and govern its use.

- NIEM:
  - Is a foundation for information exchange.
  - Offers a common vocabulary so that when two or more people talk to each other they can exchange information based on common words that they both understand.
  - Is community-driven.
  - Provides a data model, governance, methodologies, training, technical assistance, and an active community to assist users in adopting a standards-based approach to exchanging information.
  - Provides technical tools to support development, discovery, dissemination, and reuse of information exchanges.
  - Provides a forum for accelerating collaboration as well as identifying common approaches and challenges to exchanging information.

- NIEM is Not:
  - The actual exchange of information. NIEM describes the data that is in motion in the exchange.
  - A database or system. NIEM does not store information.
  - Intrusive to existing systems. NIEM allows organizations to move information quickly and effectively without rebuilding systems.
  - Software.
  - A technology stack. NIEM is technology agnostic and addresses the data layer, which means you can use NIEM irrespective of the particular technologies used within an organization.
  - Just for law enforcement and justice. Fourteen domains participate in NIEM with additional domains forming.
  - Strictly for federal government. NIEM is used in all 50 states, as well as local and tribal governments and private industry.
  - Limited by national borders. NIEM is used internationally.

Source: NIEM Website
Common Alerting Protocol (CAP)

- The **Common Alerting Protocol** (CAP) is an **XML-based** data format for exchanging **public warnings** and emergencies between **alerting technologies**.
  - Alerts from the **United States Geological Survey**, the **Department of Homeland Security**, **NOAA** and the **California Office of Emergency Services** can all be received in the same format, by the same application.
  - The CAP data structure is backward-compatible with existing alert formats including the **Specific Area Message Encoding** (SAME) used in **Weatheradio** and the broadcast **Emergency Alert System** as well as new technology such as the **Commercial Mobile Alert System** (CMAS)

```
<?xml version = "1.0" encoding = "UTF-8"?>
<alert xmlns = "urn:oasis:names:tc:emergency:cap:1.2">
<identifier>43b080713727</identifier>
<sender>hsas@dhs.gov</sender>
<sent>2003-04-02T14:39:01-05:00</sent>
<status>Actual</status>
<messageType>Alert</messageType>
<scope>Public</scope>
<info>
  <category>Security</category>
  <event>Homeland Security Advisory System Update</event>
  <urgency>Immediate</urgency>
  <severity>Severe</severity>
  <certainty>Likely</certainty>
  <senderName>U.S. Government, Department of Homeland Security</senderName>
  <headline>Homeland Security Sets Code ORANGE</headline>
  <description>The Department of Homeland Security has elevated the Homeland Security Advisory System threat level to ORANGE / High in response to intelligence which may indicate a heightened threat of terrorism.</description>
  <instruction>A High Condition is declared when there is a high risk of terrorist attacks. In addition to the Protective Measures taken in the previous Threat Conditions, Federal departments and agencies should consider agency-specific Protective Measures in accordance with their existing plans.</instruction>
  <web>http://www.dhs.gov/dhspublic/display?theme=29</web>
  <parameter>
    <valueName>HSAS</valueName>
    <value>ORANGE</value>
  </parameter>
  <resource>
    <resourceDesc>Image file (GIF)</resourceDesc>
    <resourceDesc>http://www.dhs.gov/dhspublic/getAdvisoryImage</resourceDesc>
    <area>U.S. nationwide and interests worldwide</area>
  </resource>
</info>
</alert>
```

Public Transportation: General Transit Feed Specification (GTFS)

- Proposed by Google
  - Started in 2006
  - 2009: moves from Google-specific to general
  - 2011: supports real-time update

- Purpose: public transportation schedules and associated geographic information
- Format is useful even if the data is not shared with Google

Source: https://developers.google.com/transit/gtfs/examples/gtfs-feed

Source: Making Public Transportation Schedule Information Consumable for Improved Decision Making, Raj Gupta, Biplav Srivastava, Srikanth Tamilselvan, In 13th International IEEE Annual Conference on Intelligent Transportation Systems (ITSC 2010), Anchorage, USA
# General Transit Feed Specification (GTFS)

<table>
<thead>
<tr>
<th>Filename</th>
<th>Required</th>
<th>Defines</th>
</tr>
</thead>
<tbody>
<tr>
<td>agency.txt</td>
<td>Required</td>
<td>One or more transit agencies that provide the data in this feed.</td>
</tr>
<tr>
<td>stops.txt</td>
<td>Required</td>
<td>Individual locations where vehicles pick up or drop off passengers.</td>
</tr>
<tr>
<td>routes.txt</td>
<td>Required</td>
<td>Transit routes. A route is a group of trips that are displayed to riders as a single service.</td>
</tr>
<tr>
<td>trips.txt</td>
<td>Required</td>
<td>Trips for each route. A trip is a sequence of two or more stops that occurs at specific time.</td>
</tr>
<tr>
<td>stop_times.txt</td>
<td>Required</td>
<td>Times that a vehicle arrives at and departs from individual stops for each trip.</td>
</tr>
<tr>
<td>calendar.txt</td>
<td>Required</td>
<td>Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available.</td>
</tr>
<tr>
<td>calendar_dates.txt</td>
<td>Optional</td>
<td>Exceptions for the service IDs defined in the calendar.txt file. If calendar_dates.txt includes ALL dates of service, this file may be specified instead of calendar.txt.</td>
</tr>
<tr>
<td>fare_attributes.txt</td>
<td>Optional</td>
<td>Fare information for a transit organization's routes.</td>
</tr>
<tr>
<td>fare_rules.txt</td>
<td>Optional</td>
<td>Rules for applying fare information for a transit organization's routes.</td>
</tr>
<tr>
<td>shapes.txt</td>
<td>Optional</td>
<td>Rules for drawing lines on a map to represent a transit organization's routes.</td>
</tr>
<tr>
<td>frequencies.txt</td>
<td>Optional</td>
<td>Headway (time between trips) for routes with variable frequency of service.</td>
</tr>
<tr>
<td>transfers.txt</td>
<td>Optional</td>
<td>Rules for making connections at transfer points between routes.</td>
</tr>
<tr>
<td>feed_info.txt</td>
<td>Optional</td>
<td>Additional information about the feed itself, including publisher, version, and expiration information.</td>
</tr>
</tbody>
</table>

Source: [https://developers.google.com/transit/gtfs/](https://developers.google.com/transit/gtfs/)
Recommendation on Standards

- **DATEX2**
  - Use terms from Datex 2
    - *Impact, Measured Data, Elaborated Data* are directly applicable
    - Evaluate others
  - Use traffic state from *Impact*
    - Can be rule based using flow and/or speed.

- **NIEM**
  - Align time, data, organizations, address, …, based on NIEM, if in North America
  - Consider adopting NIEM in other regions, depending on relevance
    - Address needs localization

- **CAP**
  - Support CAP format as actual data exchange

- **GTFS**
  - Adopt as much as relevant
References

- Smarter city data model standards landscape, Part 2: Transportation
- CAP: http://docs.oasis-open.org/emergency/cap/v1.2/CAP-v1.2-os.html
- Datex 2: www.datex2.eu/
- NIEM: www.niem.gov
- GTFS: https://developers.google.com/transit/gtfs/
- Scribe: http://researcher.watson.ibm.com/researcher/view_project.php?id=2505
Addendum: Road Conditions
Road Surface Conditions

- **Surface Condition**
  - **Boggy Conditions**
    An unsealed road surface that is wet and soft, which may lead to a vehicle becoming bogged. Can also apply to loose dry conditions.
  - **Bulldust**
    A very fine material with a flour like texture often covering surface irregularities or deep holes.
- **Changing Surface Conditions**
  The condition of the road surface may change along the length of the road, and following road maintenance activities. Wet weather may also cause a localised change in road condition.
- **Corrugations**
  An unsealed road surface having ripples or undulations along the road.
- **Floodway**
  A localised section of road designed to accommodate temporary overtopping allowing water to flow over the surface of the road. Refer also “Stream Crossing”.
- **Loose Surface**
  A layer of unbound or coarse material on the pavement surface. Can be very fine material (bull dust or sand), or large material in coarse gravel pavements.
- **Potholes**
  Bowl-shaped depressions in the road surface, resulting from the loss of pavement material.
- **Rough**
  The consequence of irregularities, uneven in nature, in the longitudinal profile of a road with respect to the intended profile. Interpretation of conditions will not allow vehicles to use the road at customary speeds.
- **Rutting**
  A longitudinal deformation of a pavement surface formed by the wheels of vehicles. May be on a sealed or unsealed road.
- **Slippery Surface**
  Surface becomes slippery in wet conditions or from fuel or material spills.
- **Stream Crossing**
  Natural watercourses that cross the road alignment. On sealed roads the crossing would generally be a bridge, floodway or culvert. On unsealed roads there may be no structure. Water levels at floodways and stream crossings are likely to change rapidly and may present an extreme hazard. The ability to cross these floodways and streams will depend on depth of water, velocity of flow, driver experience and vehicle type. ALWAYS assess conditions prior to crossing or travelling along the affected section. See also “Water over Road”.
- **Washouts**
  Steep, irregularly sided grooves in the road surface caused by erosion of the road surface by water.

http://www.nlris.nt.gov.au/roadreport/include.jsp?pageName=term_files/index
Road Types

- Road Types
  - **Formed**
    An unsealed road that has been constructed to above the natural surface using local materials. Generally of a higher standard than unformed.
  - **Gravelled**
    An unsealed road that has been formed and strengthened with a good quality gravel material. Generally of a higher standard than formed.
  - **Sealed**
    A road that is constructed with a bitumen surface.
  - **Unformed**
    An unsealed road that is generally a flat track following the natural terrain. Often occurs as a rough track with two wheel paths, and close vegetation.
  - **Unsealed**
    A road without a bituminous surface that could be gravelled, formed or unformed.

http://www.ntlis.nt.gov.au/roadreport/include.jsp?pageName=term_files/index
Types of Restrictions

- **Detour**
  An alternative route available for traffic during a temporary closure of a road, or part of that road. Detours may be sealed or unsealed.

- **High Clearance**
  A light vehicle with clearance sufficient to travel over rough and uneven road surfaces. The clearance on these vehicles generally exceeds common passenger vehicles.

- **High Clearance 4WD**
  A light vehicle with two axles that is built for on and off road use and clearance sufficient to travel over rough and uneven road surfaces. The clearance on these vehicles generally exceeds common passenger vehicles.

- **Impassable**
  Access along the section of road is affected by flooding or other obstructions. Road conditions are likely to change rapidly and may present an extreme hazard. Persons should not attempt to use/access the road.

- **Lane Closure**
  Temporary closure of one or more traffic lanes. Traffic control devices are in place -to enable traffic to use unaffected lanes.

- **Road Closed**
  Temporary closure of the road where passage of motor vehicles is not permitted. Barriers are in place and penalties shall be applied to drivers ignoring the road closure.

- **Weight and Vehicle Type Restriction**
  A restriction placed on heavy vehicles to control axle group loads and/or vehicle size to protect the road during adverse conditions.

- **With Caution**
  The condition of part of the road may have deteriorated so that due care must be exercised by the travelling public.

- **4WD**
  A light vehicle with two axles that is built for on and off road use. All four wheels can be engaged to improve traction in adverse conditions.

http://www.ntlis.nt.gov.au/roadreport/include.jsp?pageName=term_files/index
Traffic Management and AI

Section: Intelligence

Speakers: Anand Ranganathan, Biplav Srivastava
Outline

1. Traffic Management Problem
2. Instrumentation
   1. Sensing traffic
   2. Traffic state estimation
   3. Optimizing and combining sensor data
3. Interconnection
   1. Middleware
   2. Traffic standards
4. Intelligence
   1. Path planning
      1. Simple Illustration
      2. Path Planning for Individual Vehicles
   2. End-user analytics
      1. Bus arrival prediction and journey planning, with state-of-art instrumentation
      2. Multi-modal journey planning, without sensors
5. Supporting topics
   1. Traffic Simulators
   2. Practical considerations for real-world pilots
Simple Illustration
Description

- City’s road network is a 5x5 grid
- Travelers in the city:
  - A, B, C, and D live as shown on the map
  - Their workplace is W [2,2]

Objective:

1. We want to help all the travelers plan their trips to work in the morning, and back home daily using shortest time.

2. If the office can be moved, where should it be moved to reduce the commute time for all employees?

Commute constraints

- Speed of all travelers, by default, is 1 hop per 10 minutes.
- Speed is reduced by half if sun on the face while commuting. Hence, in the morning, travelers going east are slowed by half, and in the evening, travelers going west are slowed by half. Some examples:
  - Time taken from [0,0] to [0,1] is 20 minutes in the morning, and 10 minutes for the rest of the day
  - Time taken from [0,1] to [0,0] is 20 minutes in the evening, and 10 minutes for the rest of the day
  - Time taken from [0,0] to [1,0] remains 10 minutes throughout the day
- If two persons travel on the same road, their speeds reduce by half.
- The two slowing factors work independently. An example is that two vehicles going from [0,0] to [0,1] in the morning will take 40 minutes

[Coordinated Planning] Total time for all: 200 minutes

D (Deb)
[4,0]

D (Car2) time- 60

C (Car2) time- 40

C (Chumki)
[4,4]

A (Alice)
[0,0]

A (Car0) time- 60

B (Car1) time- 40

B (Bob)
[0,4]

W (Workplace)
[2,2]
[Non Coordinated Planning] Total time for all: 260 minutes

D (Deb)

[4,0]

D (Car2) time- 80

C (Chumki)

[4,4]

C (Car2) time- 50

[2,2]

W (Workplace)

A (Car0) time- 80

B (Car1) time- 50

[0,4]

A (Alice)

B (Bob)

North

East
Path Planning for Individual Vehicles
Traffic - Dependent Route Planning

- Definition: Find a path from a source to a destination on a road network with minimum travel time delay and that considers current and future traffic conditions on all road links.

- Needs dynamic time-dependent shortest path computations
  - Cost of each link is different at different intervals of time, and
  - These time-dependent costs get updated as new traffic information is obtained.

- Challenges related to performance
  - Need high throughput and low latency
    - 100s of users, need response in less than a second
  - Scale well as we increase size of road network
    - 10,000s or 100,000s nodes and links
  - Handle real-time updates to road network conditions
Dynamic Time-Dependent Shortest Path

- Previous works have focused either on dynamism or on time-dependency, but not both

- A* adapted for time-dependent shortest path networks (Chabini and Lan)
  - Assumes First In – First Out rule on each link
    - A traveler who travels on a link will reach the end of the link before anyone who departs later from the beginning of the link.

- Our approach: Modify A* as follows:
  - As paths are extended with links during the execution of A*, time is advanced and therefore future delay values of links are used as link costs.
  - To ensure the FIFO property in a time-dependent network,
    - If the time interval changes while traveling on a link, new delay value used for the rest of the link.
    - A link can be traveled during many time intervals
  - Update link costs with new traffic information
  - Use heuristic function as \( \text{Euclidean Distance} / \text{max speed limit} \)
Implementing Shortest Path as a Stream Processing Application

- Shortest Path Algorithm implemented as a C++ user defined Operator

- One time input to Operator :
  - Road Network

- Continuous Updates to Operator:
  - Stream with Delays on different links in road network
  - Stream with Queries (containing Origin, Destination, Departure Time)

- Result : Stream containing Links in Shortest Path and Expected Travel Time

- Can Parallelize effectively
  - By Destination Point
  - Allows scaling and caching of heuristic function
Experiments on Shortest Path

- Updates to Traffic Conditions
  - Generated continuously as GPS data is streamed in
  - Divided day into 5 minute intervals
  - 61422 updates on 11177 links in 1 day
Performance Experiments

- Measured delay and throughput on 10, 20 and 50 machines when there is a burst of 10,000 shortest path queries
- Scales almost linearly as number of machines increases

![Graphs showing average delay and throughput vs. number of nodes.](image-url)
Evaluating the benefits of doing dynamic time-dependent shortest paths – quantifying the time savings

- How much do we gain from using traffic data in the shortest path calculation?
  - Compared to path obtained using static speed limits
- Ran shortest path algorithm for 50 different departure times for 500 origin-destination pairs
- Upto 62% decrease in travel time
Number of changes in Shortest Path

- Shortest Path for a certain origin-destination pair changed upto 37 times (out of the the 50 consecutive runs)
  - 30 changes for more than half of the queries
Degree of change of shortest path

- Degree of change = \((1 - \text{number of common links} / \text{total number of links in two paths}) \times 100\)
- Paths can change by up to 98% between consecutive runs
Visualizing the paths for origin-destination pair with max number of changes
End User Analytics
Bus Arrival Prediction and Travel Planner

Business Challenge

- Meanwhile, public transportation represents a major cost for governments and is one of the most visible services available to citizens.
- However, cities mostly lack the tools, information and visibility to understand and optimize how public transportation is performing and meeting citizen’s needs.
IBM’s Public Transportation Awareness tool

- Continuously analyses data from Automatic Vehicle Location (AVL) systems
- Collects, processes, and visualizes location data of all public transportation vehicles
- Analytic algorithms are optimized for high performance, on-demand access
- Automatically generated transportation routes and stop locations
Bus Delay Overview

IBM Transportation Awareness and Optimization System

Bus Delay Distribution

Bus Delay over Time

IBM Research

© 2010 IBM Corporation
Bus Delay Query

http://www.youtube.com/watch?v=VBv5XRGQ7nA
Bus Stop, Bus Route Queries

IBM Transportation Awareness and Optimization System

Enter a search query here:

- Show all bus vehicles
- Show all bus routes
- Show all bus stops

- Stop 1462 Gallymore Road Junction Gallymore Park
  Position: (53.330451, -6.323801)
  Routes seeing this stop:
  - 123 (var. 001) from Griffith Avenue to Klinamanagh Road
  - 123 (var. 003) from Cathal Brugha Street to Klinamanagh Road

- Route 123 (var. 001) from Griffith Avenue to Klinamanagh Road
  Position: (53.342875, -6.276754)

- Stop 24 Home Farm Road Junction Walsh Road

© 2010 IBM Corporation
Bus Arrival Prediction
Dublin Experience
Dublin Experience – Deployed Dublin Bus RTPI

150 bus routes / 5000 bus stops
600 buses contemporary on routes during the day
20 stops per bus
20*600 = 12,000 predictors per min
Use Case and Performance

- Deployed in a production environment at the Dublin City Council since Jan 2011
- Run on VMWare Linux RHEL 5.4, 4 cores Intel Xeon© 2.27GHz, 6GB RAM
- Has been running on the average 2-3 months between updates
- Typically 1 – 3 user sessions running simultaneously between DCC and IBM
Technical Challenges

1. Data diversity, heterogeneity
2. Data accuracy, sparsity
3. Data volume

- 700 intersections
- 4,000 loop detectors
- 20,000 tuples / min

- 1,000 buses
- 3,000 GPS / min

- 200 CCTV cameras

© 2010 IBM Corporation
PTA Design – Streams Flow Graph

(*) Returns StopMonitoring info for all stops
Data

- GPS readings once every 20 seconds, with:
  - Time of capture, GPS location, Bus line ID, Bus vehicle ID
    - Delay information => Calculated by Dublin bus with respect to timetables
    - Delay information embedded on the GPS readings and available for processing
  - Approx 19 million GPS readings per month
  - Also, GPS coordinates of Bus Stops
Pre-Processing

- Detect arrival/departure buses
  - Radius of 20m around the bus stop and track a bus from a bus line.
- Modeling of delays at every bus stop
- Recovering timetables
- Modeling of delay correlation between delay at departure bus stop and arrival bus stop
- Calculate Travel Times
Modeling delays at bus stops

- Delay: bus arrival/departure earlier/later than the expected on the timetables
- Detect when the bus is at a bus stop (within a 20m radius)
- Calculate Histograms
  - Per hour of the day,
  - Weekdays, Weekends
- Use of Kernel Density Estimators
Modeling delays at bus stops

Delays along bus stops for bus line 46A at 7am: Buses tend to be ahead of schedule while traveling

Delays along bus stops for bus line 46A at 8am: Buses tend to be behind of schedule while traveling
Recovering timetables

Histogram of all buses detected within the 20m radius of a bus stop, for a certain bus line.

Histogram of data corrected with the delay information. Patterns = Timetables
End to end trip times

- Estimate distributions of travel times from A to B on the network

\[
Trip_{A \rightarrow B}(t) = Pr(Trip_{A \rightarrow B} = t)
\]

- As a function of:

\[
Trip = f(\text{Timetables, Delays, Evolution of Delays, Travel Times per leg})
\]

- Buses early/late/missing connections => how the trip changes and probabilities
Scenario and Simulation

- Line 123 from stop 1499 to stop 271
- Walk 200 meters from stop 271 to stop 6059
- Line 46A from stop 6059 to stop 2017

- Monte Carlo Simulation for End-to-End travel times
- Arrival of Passengers, Bus, Walking time, Travel time per leg
- Walk speed modeled as normal distribution with mean 4km/h and variance 1km/h
Results

- Expected value differs by 10 minutes of the deterministic trip time
- Variance is also large
- Travel time distribution mean and variance is function of the time of the day
Multi-modal journey planning, without sensors

Based on joint work by Raj Gupta, Srikanth Tamilselvam, Biplav Srivastava

Problem

- **Input:**
  - A person wants to travel from place A to B

- **Invariant Inputs:**
  - Mode and mode categories for commuting
    - The city has taxis, buses, metros, autos, rickshaws
    - Buses and metros have published routes, frequency and stops. No other instrumentation.
    - Autos and rickshaws can be available at stands, or opportunistically, on the road
    - Taxis can be ordered over the phone
  - Preferences over commuting modes and categories
    - The person can have a vehicle (e.g., car),
    - Can walk short distances
    - Has preferences over commuting modes based on cost, time, safety, physical effort.

- **Output**
  - Suggest to the person which mode or combination of modes to select

- **Observation:** there is no coordination assumed from the transport operators

© 2010 IBM Corporation
Multi-Mode Commuting Recommender in Delhi And Bangalore, India

Get public transit directions

City: Delhi
From: [0046] aims (bus,metro)
To: [2094] rajiv chowk (metro)
Start: 5 PM  End: ANY
Mode: ANY
No. of choices: ANY
No. of hops: ANY

13 solutions found. Please click on the links to explore

Solution 1[# Hops = 0; Modes used = metro ]
Source: aims
Destination: rajiv chowk
Route Id: LINE-2
Starting Time: 17:00
Reaching Time: 17:22
Frequency: 005
Mode: metro

Solution 2[# Hops = 1; Modes used = bus , metro ]

Solution 3[# Hops = 1; Modes used = bus , metro ]

Highlights

• Published data of multiple authorities used; repeatable process
• Multiple modes searched
• Preference over modes, time, hops and number of choices supported; more extensions, like fare possible

Get public transit directions

City: Bangalore
From: [0235] bial Bgl (bus)
To: [1039] jn. of nice rd Bgl (bus)
Start: ANY  End: ANY
Mode: ANY
No. of choices: ANY
No. of hops: ANY

2 solutions found. Please click on the links to explore

Solution 1[# Hops = 1; Modes used = bus , bus ]
Source: bial Bgl
Destination: majestic (kbs) Bgl
Route Id: BIAS-1
Mode: bus

Solution 2[# Hops = 1; Modes used = bus , bus ]
Source: majestic (kbs) Bgl
Destination: jn. of nice rd Bgl
Route Id: G-7
Mode: bus
Factors which affect/ can improve accuracy
- Quality of schedule published by public transportation operators (bus and metro for us)
  • Names, spelling and conventions in stops by different agencies
  • We correct and can do more - If we correct too much, we remove the traceability to original published schedules
- Lack of co-relationship across stop names and location
  • Affects what the user sees when they select
  • We can include geo-spatial analysis when we offer choice of locations
- Increase inter-operability across agencies
  • Make traffic data into linked open data format
- Integrate with geo-spatial analysis of tools
References

References


- A. Ozal, A. Ranganathan, N. Tatbul, "Real-time Route Planning with Stream Processing Systems: A Case Study for the City of Lucerne", ACM SIGSPATIAL International Workshop on GeoStreaming (IWGS’11), Chicago, IL, USA, November 2011


Tutorial: Traffic Management and AI

Section: Supporting Topics

Biplav Srivastava, Anand Ranganathan
IBM Research

Toronto, Canada
Outline

1. Traffic Management Problem

2. Instrumentation
   1. Sensing traffic
   2. Traffic state estimation
   3. Optimizing and combining sensor data

3. Interconnection
   1. Middleware
   2. Traffic standards

4. Intelligence
   1. Path planning
      1. Simple Illustration
      2. Path Planning for Individual Vehicles
   2. End-user analytics
      1. Bus arrival prediction and journey planning, *with state-of-art instrumentation*
      2. Multi-modal journey planning, *without sensors*

5. **Supporting topics**
   1. Traffic Simulators
   2. Practical considerations for real-world pilots
Why Use Traffic Simulators

- Use them to prepare for real-world experience
- Overcome
  - Data issues
  - Try what-ifs
  - Cost reasons
- No simulator is a replacement for actual deployment
## Traffic Simulators

<table>
<thead>
<tr>
<th>availability</th>
<th>licensing</th>
<th>programing language</th>
<th>code extension</th>
<th>development started</th>
<th>current status</th>
<th>macro</th>
<th>micro traffic flow</th>
<th>location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamit</td>
<td>Academia</td>
<td></td>
<td></td>
<td>2004</td>
<td>last release 25-7-2011</td>
<td></td>
<td>both</td>
<td>mit.edu/its/dynamit.html</td>
</tr>
<tr>
<td>Freesim</td>
<td>Open Source</td>
<td>GNU</td>
<td>java</td>
<td>Possible</td>
<td>2004</td>
<td></td>
<td>both</td>
<td><a href="http://www.freewaysimulator.com/index.html">http://www.freewaysimulator.com/index.html</a></td>
</tr>
<tr>
<td>Sumo</td>
<td>Open Source</td>
<td>Sourceforge</td>
<td>C++</td>
<td>Possible</td>
<td>2002</td>
<td></td>
<td>micro</td>
<td><a href="http://sumo.sourceforge.net/">http://sumo.sourceforge.net/</a></td>
</tr>
<tr>
<td>Microsimulation of Road Traffic Flow</td>
<td>Academia</td>
<td>java applet</td>
<td>Possible</td>
<td>Last release June 2011</td>
<td>micro</td>
<td></td>
<td></td>
<td><a href="http://www.traffic-simulation.de/">http://www.traffic-simulation.de/</a></td>
</tr>
<tr>
<td>MITSIMlab</td>
<td>Open Source</td>
<td>Free to use</td>
<td>Possible</td>
<td>1999</td>
<td>No support since 2005</td>
<td>micro</td>
<td></td>
<td>mit.edu/its/mitsimlab.html</td>
</tr>
<tr>
<td>Matsim</td>
<td>Open source</td>
<td>GNU</td>
<td>Java</td>
<td>Possible</td>
<td>2003</td>
<td></td>
<td>macro</td>
<td><a href="http://www.matsim.org/">http://www.matsim.org/</a></td>
</tr>
</tbody>
</table>

© 2010 IBM Corporation
Agent-Based Transport Simulations

Key Features of MATSim
- **Fast Dynamic and Agent-Based Traffic Simulation**
  Simulate whole days within minutes
  It’s not time based.
- **Supports Large Scenarios**
  MATSim can simulate millions of agents or huge, detailed networks
- **Versatile Analyses and Simulation Output**
  E.g. compare simulated data to real-world counting stations
- **Modular Approach**
  Easily extended with your own algorithms
- **Sophisticated Interactive Visualizer**
  See what each agent is doing during the simulation
- **Open Source**
  You get the Java Source Code, which runs on all major operating systems

Easy integration with Open Street Maps
Matsim: Input and Output

Network

Plan

Network_Change

Events
Practical Considerations
High-levelSuggestions

- Follow and understand the data
  - Problem characteristic depends on it
  - Relevance of algorithms depends on it
  - Result quality depends on it

- Prototype early
  - Optionally, in a simulator, and
  - Definitely, in real-world
  - No solution works unless millions are using it

- Think long term
  - Traffic systems last for decades, not 2-3 years (typical life of a computer)
  - Support inter-operability in IT as well as in users
  - Quantify benefits and costs using standard metrics
  - Always think of people issues
    - Good to have analogies with computer networks, but remember that packets can be dropped while travelling, but not people
    - If people adopt, an imperfect solution will still be successful
Experience in Dublin, Ireland

- Close collaboration with the city
  - Jointly funded research center

- Access to a number of live data feeds
  - GPS from buses, loop, traffic signal cycles, CCTV

- City was willing to try out research prototypes

- Goal was to improve different aspects of transportation
Experience in Stockholm, Sweden

- Close collaboration with KTH and with Trafik Stockholm
  - Research funding

- Access to taxi GPS data and to “MCS” data (overhead microwave sensors)
  - Through a 3rd party collector
  - Privacy restrictions limited completeness of data

- Ongoing research collaboration

- Key goal is to improve travel time estimation to various points in the city
Experience in New York, USA

- Research Collaboration with CUNY and NYU

- Access to traffic data (only start and stop information)
- Privacy restrictions limited identification information
- Very limited in terms of ability to extract traffic information
- However, can still extract human mobility patterns
Experience in Boston, USA

- Smarter Cities Challenge (SCC), an IBM philanthropic initiative to bring IBM's experts to cities around a topic and get a recommendation on what the city needs to do.
  - SCC in Boston was on traffic in June 2012 for 3 weeks.
  - The project has been unique for two reasons: (a) academia has been involved and (b) the team does a running prototype along with a recommendation.

- Boston collects a significant amount of data from many sources that could be quite useful to researchers, developers, transportation engineers, urban planners, and above all, citizens.
  - Siloed, in multiple formats and not fully exploited.
  - Developed a prototype that created a standards-based common model, loaded > 1M data from 3 sensor types already in city, and showed the potential of insights from them; also recommendations for more sensor types and analytics

Team at work – Source: Boston Globe article

Press on the IBM SCC Boston team work:
http://www.boston.com/business/technology/articles/2012/06/29/ibm_gives_advice__on_how_to_fix_boston_traffic__first_get_an_app/

2. Popular Science, 2 July 2012

3. Others: National Public Radio (USA), and a range of local TV stations on the work.
More Attempts

- Around the World
  - Level of instrumentation can vary drastically, or even be absent
  - Ideas needed to incentivize people to reduce demand
  - Funding for large-scale supply (infrastructure) increase is reducing

- Beyond city bodies
  - Collaboration with mobile companies
  - Collaboration with bus companies
  - ...

- And by individual organizations. Any can take a lead to reduce traffic demand
  - Car-pooling. [Making Car Pooling Work – Myths and Where to Start, Biplav Srivastava, in 19th World Congress on Intelligent Transportation Systems, Vienna, Austria, 2012]