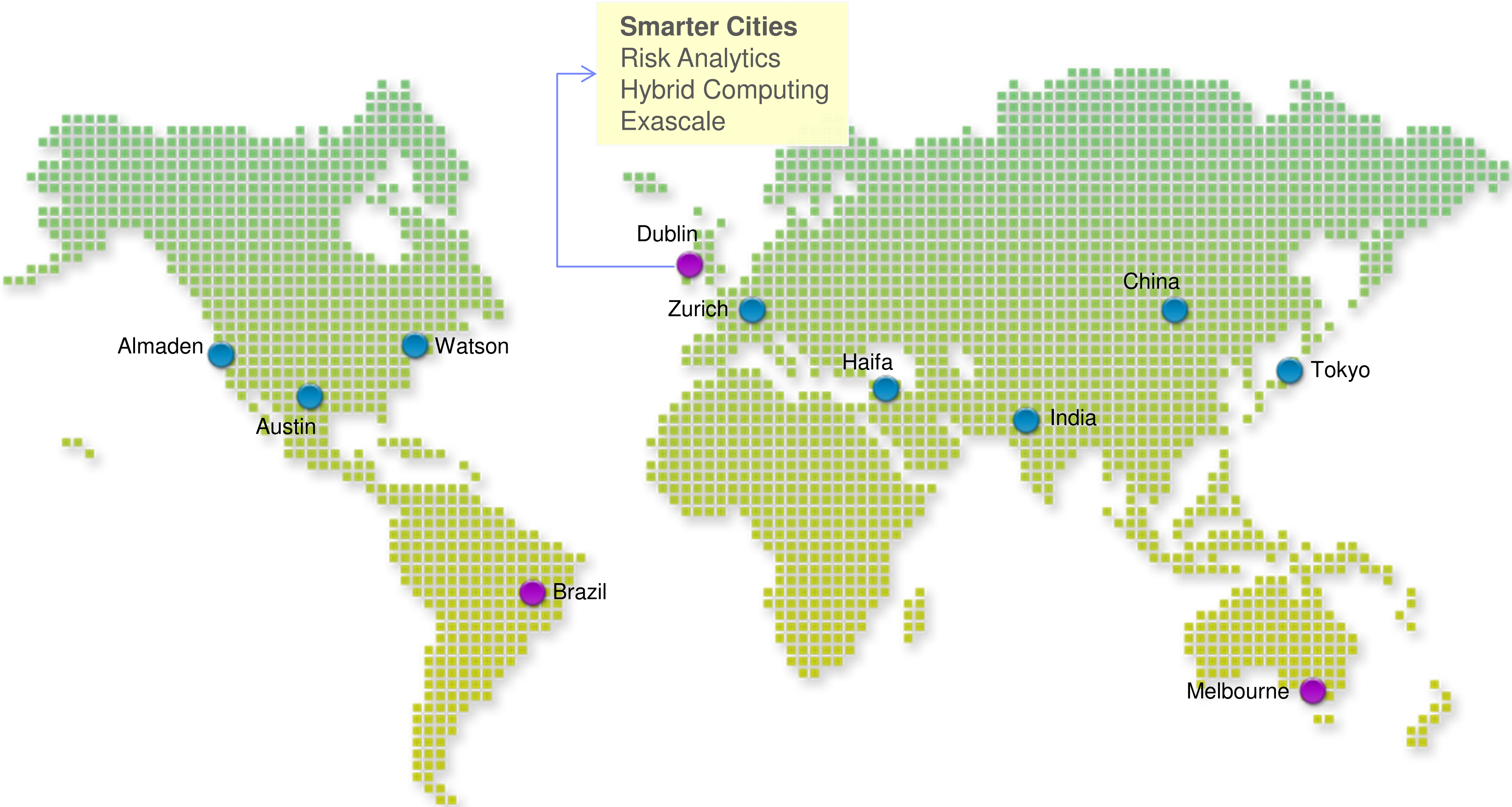
Smart Cities –
How Can Data Mining and Optimization Shape Future Cities?

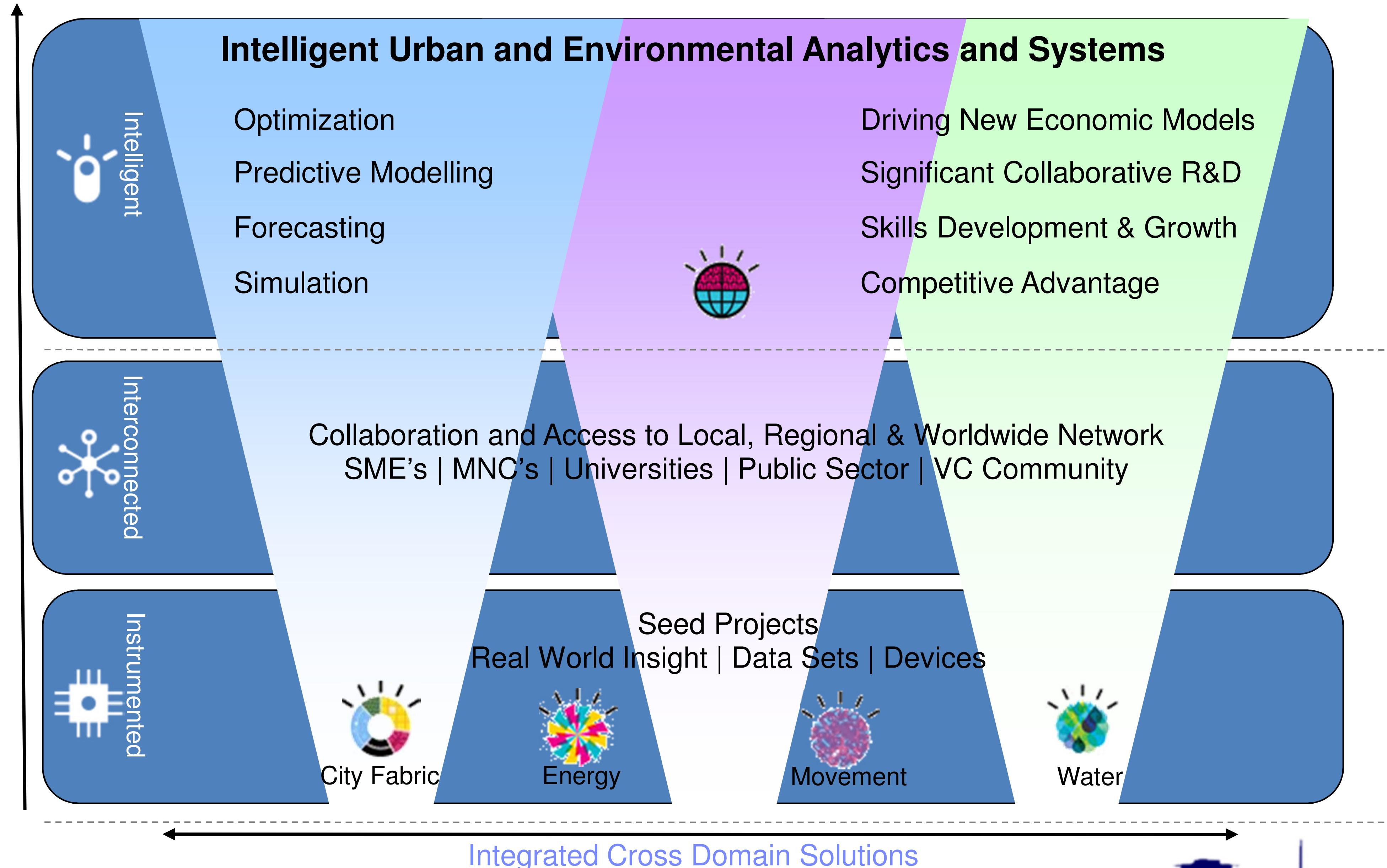
Francesco Calabrese

Advisory Research Staff Member, Analytics & Optimization
Smarter Cities Technology Centre
IBM Research and Development - Ireland

IBM Research Worldwide



Smarter Cities Technology Centre



SmarterCities Technology Centre

Many Visions of what a Smarter City might be

A "mission control" for infrastructure

A showcase for urban planning concepts

A totally "wired" city

A self-sufficient, sustainable eco-city

But we know they'll intensively leverage ICT technologies

Intelligent Transportation Systems
- Integrated Fare Management
- Road Usage Charging
- Traffic Information Management

Telecommunications
- Fixed and mobile operators
- Media Broadcasters

Energy Management
- Network Monitoring & Stability
- Smart Grid – Demand Management
- Intelligent Building Management
- Automated Meter Management

Public Safety
- Surveillance System
- Emergency Management Integration
- Micro-Weather Forecasting

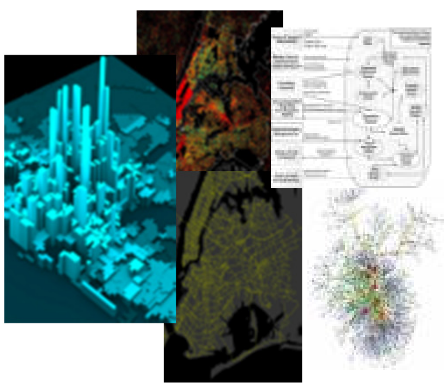
Water Management
- Water purity monitoring
- Water use optimization
- Waste water treatment optimization

Environmental Management
- City-wide Measurements
- KPI's
- CO2 Management
- Scorecards
- Reporting

How can we help cities achieve their aspirations?

1. **Sensor data assimilation**
   - Data diversity, heterogeneity
   - Data accuracy, sparsity
   - Data volume

1. **Modelling human demand**
   - Understand how people use the city infrastructure
   - Infer demand patterns



1. **Operations & Planning**
   - Factor in uncertainty

Outline

Sensor data assimilation

- Continuous assimilation of real-time traffic data

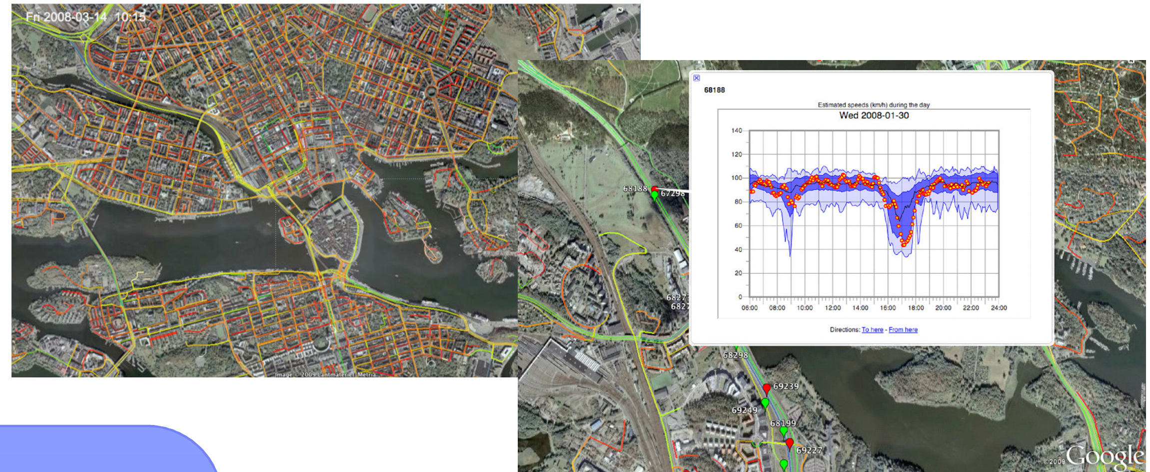
Understanding/Modeling human demand

- Characterizing urban dynamics from digital traces

Operations & Planning

- Leveraging mathematical programming for planning in an uncertain world

Our Stockholm Experience (2009)



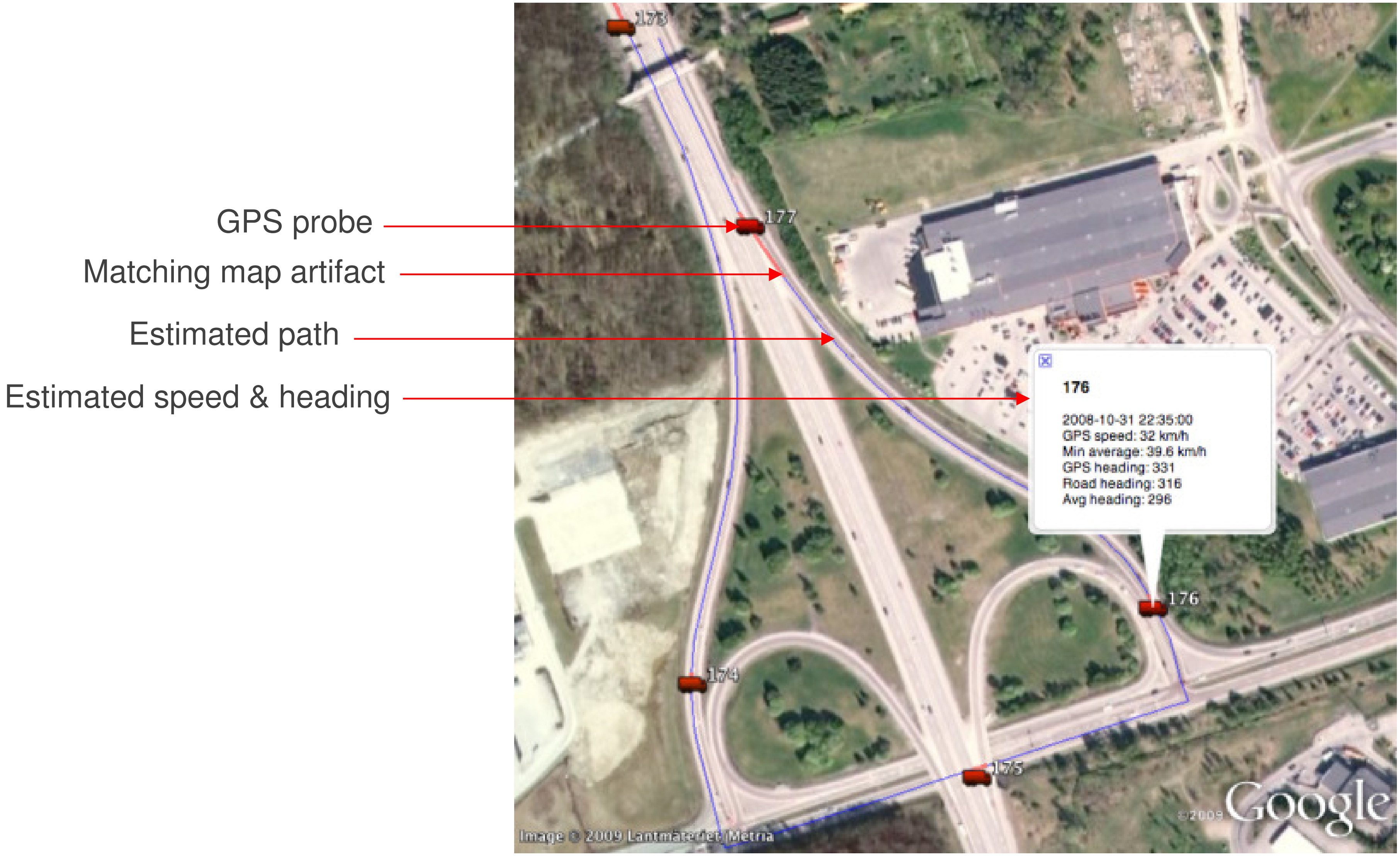
GPS devices
Induction loops
Axle counters
Traffic lights
Parking meters
Cameras
MCS
Weather stations
Microblogs

Real-time assimilation, mediation (e.g. de-noising), aggregation (e.g. key traffic metrics)

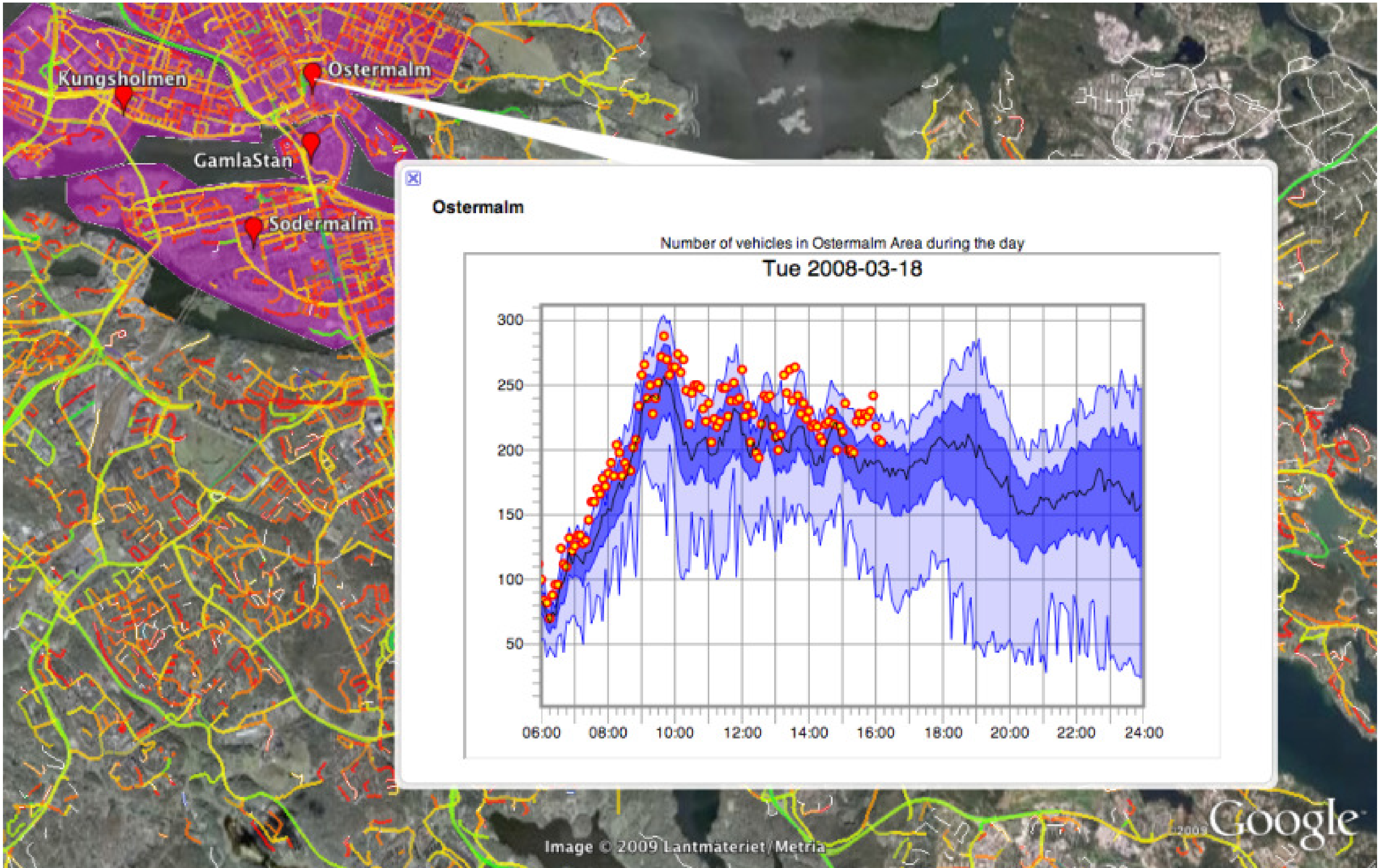
Noisy GPS Data

- To become useful, GPS data has to be related to the underlying infrastructure (e.g., road or rail network) by means of map matching algorithms, which are often computationally expensive
- In addition, GPS data is sampled at irregular possibly large time intervals, which requires advanced analytics to reconstruct with high probability GPS trajectories
- Finally, GPS data is not accurate and often needs to be cleaned to remove erroneous observations.

Real-Time Geomapping and Speed Estimation

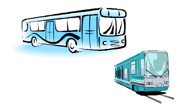


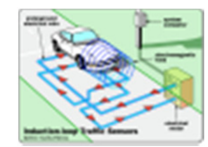
Real-Time Traffic Information

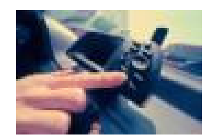


Our Dublin Experience (2011)

- Complex system & analytics challenges
  - Data diversity, heterogeneity
  - Data accuracy, sparsity
  - Data volume



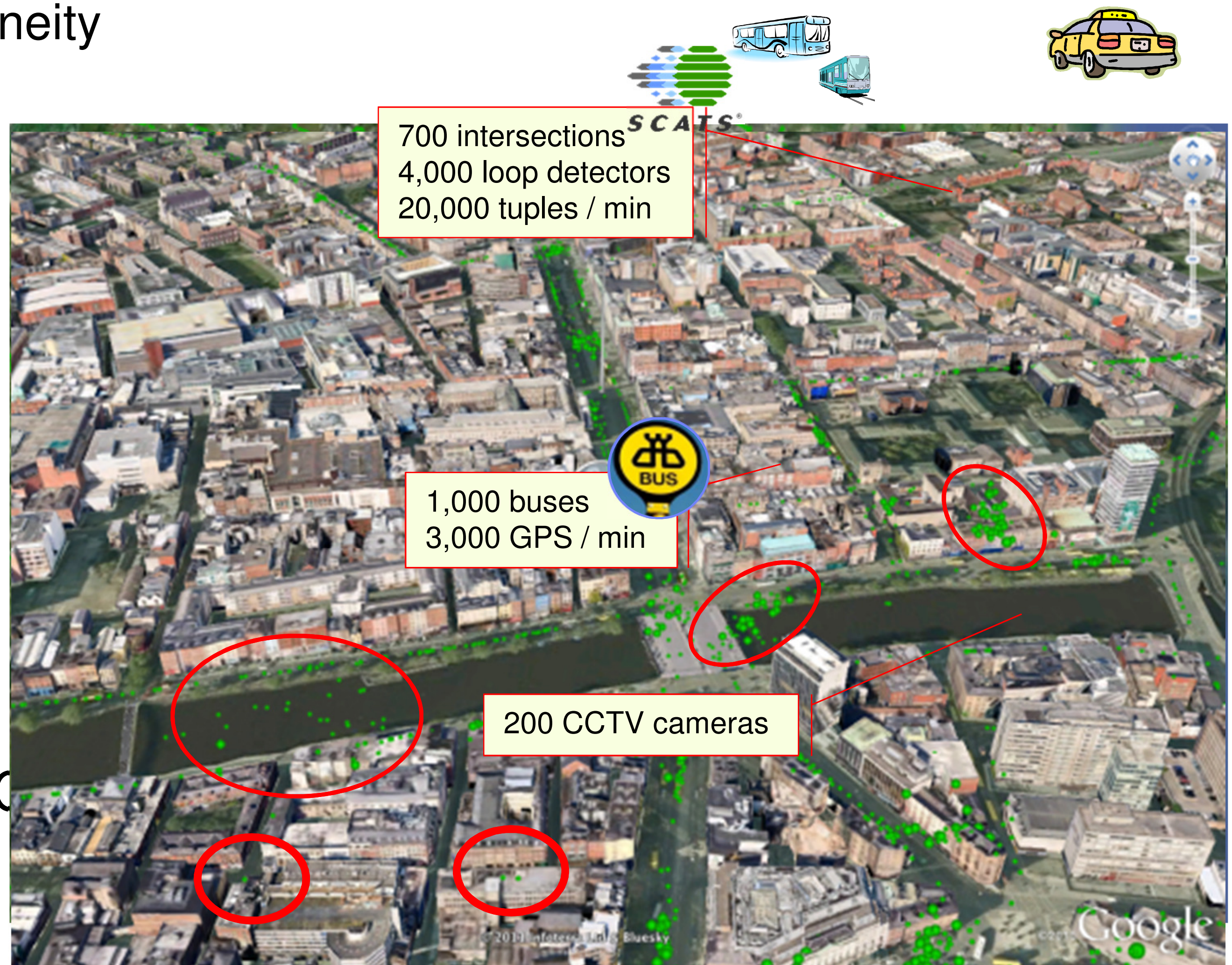




- Active relationship with DCC
- Deployed in Dublin's DoT

Our Dublin Experience (2011)

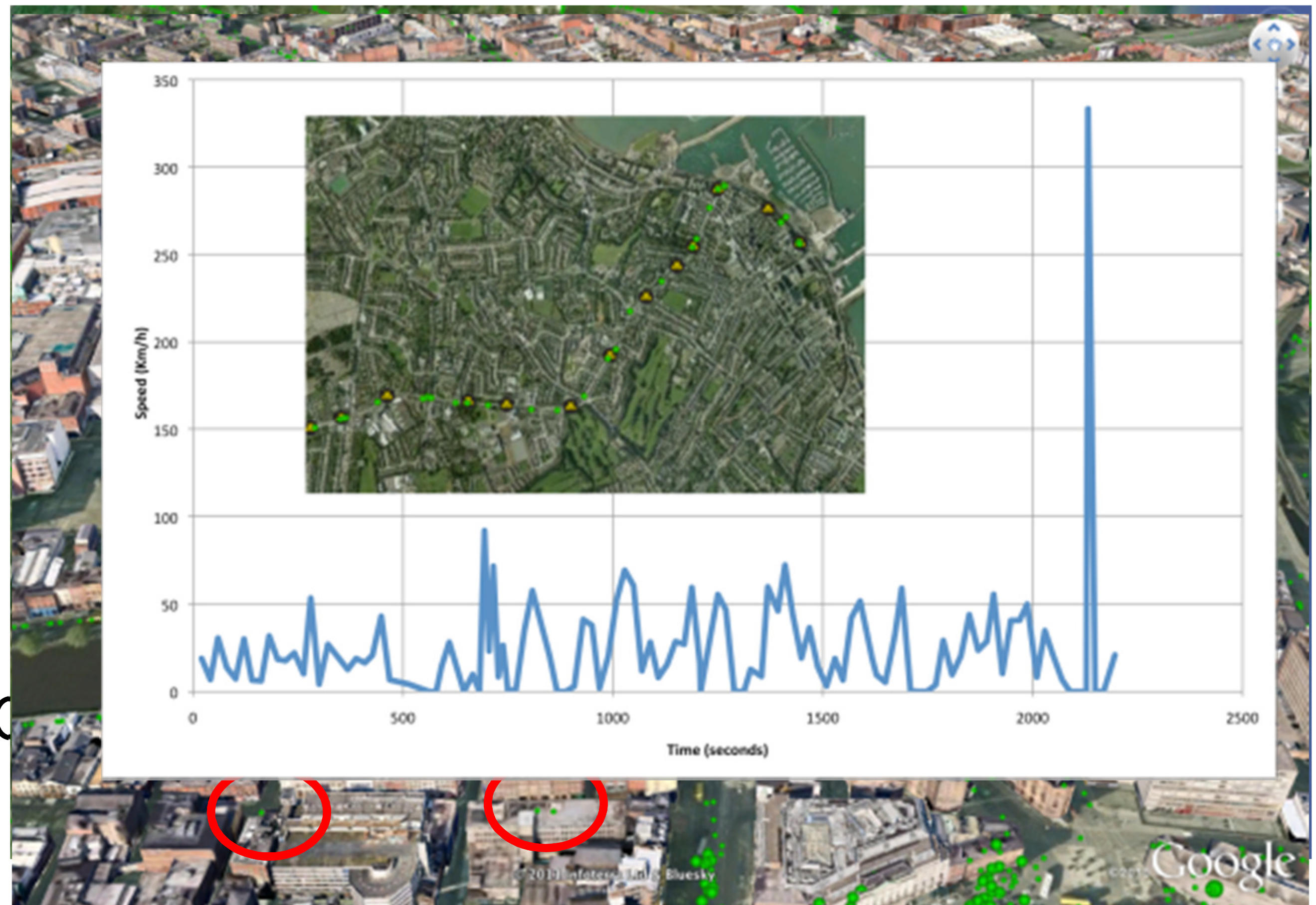
- Complex system & analytics challenges
  - Data diversity, heterogeneity
  - Data accuracy, sparsity
  - Data volume



- Active relationship with DC
- Deployed in Dublin's DoT

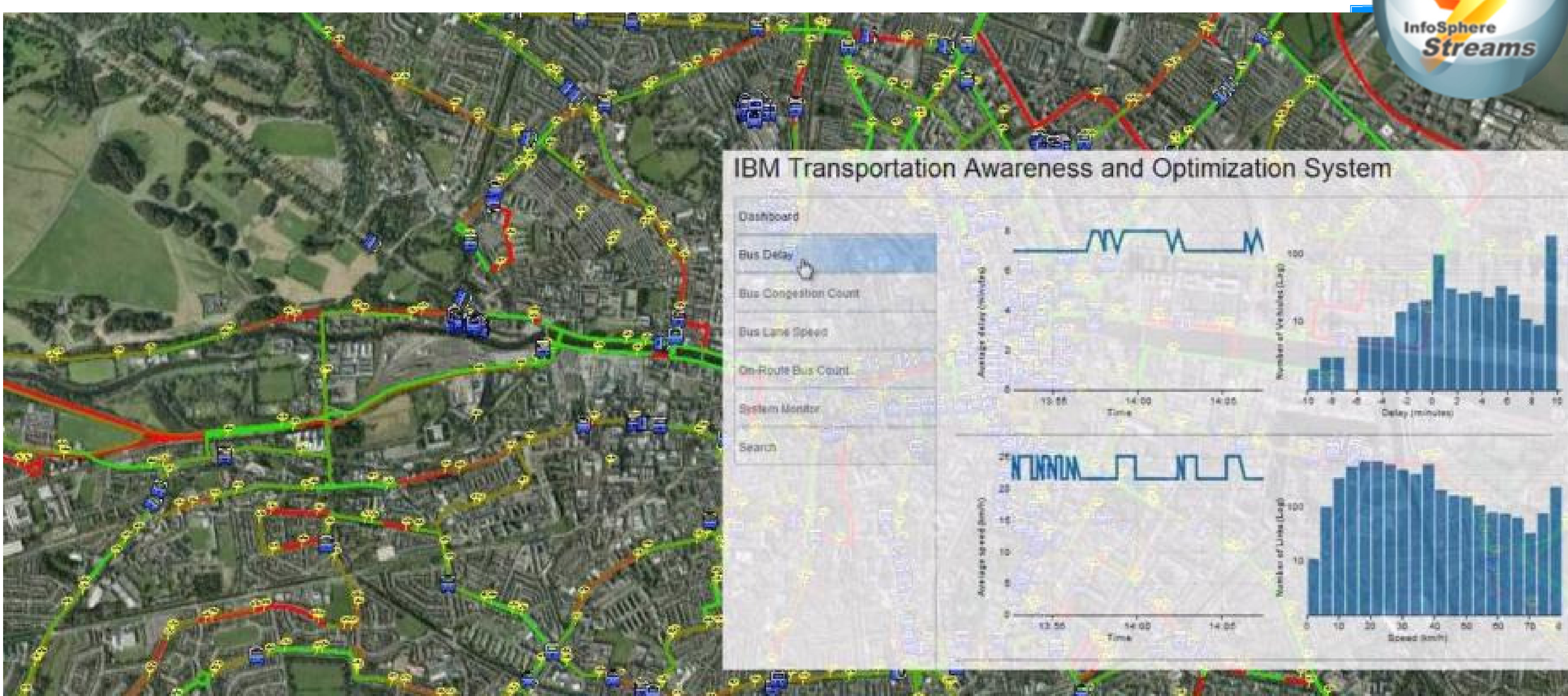
Our Dublin Experience (2011)

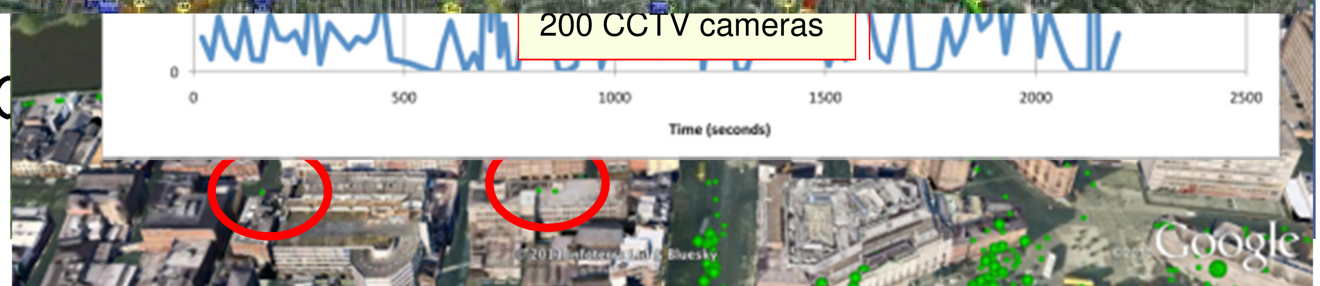
- Complex system & analytics challenges
  - Data diversity, heterogeneity
  - Data accuracy, sparsity
  - Data volume



- Active relationship with DC
- Deployed in Dublin's DoT

Our Dublin Experience (2011)





- Active relationship with DC
- Deployed in Dublin's DoT

Outline

Sensor data assimilation

- Continuous assimilation of real-time traffic data

Understanding/Modeling human demand

- Characterizing urban dynamics from digital traces

Operations & Planning

- Leveraging mathematical programming for planning in an uncertain world

Pervasive Technologies Datasets as Digital Footprints

Understand how people use the city's infrastructure

- Mobility (transportation mode)
- Consumption (energy, water, waste)
- Environmental impact (noise, pollution)

Potentials

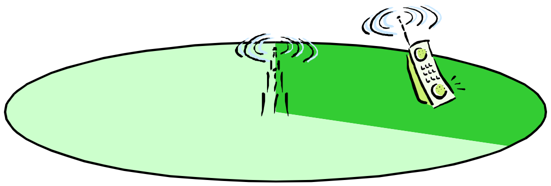
- Improve city’s services
  - Optimize planning
  - Minimizing operational costs
- Create feedback loops with citizens to reduce energy consumption and environmental impact

Understanding Urban Dynamics

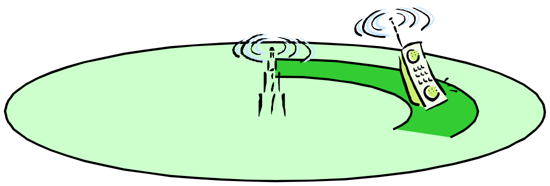
- Research goals
  - Understanding human behavior in terms of mobility demand
  - Analyzing and predicting transportation needs in short & long terms

- Outcome
  - Help citizens navigating the city
  - Design adaptive urban transportation systems
  - Support urban planning and design

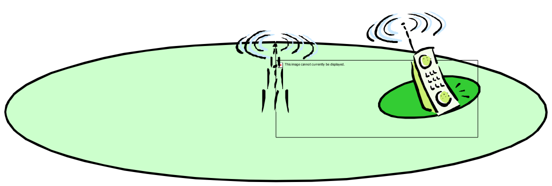
Mobile phones to detect human mobility and interactions



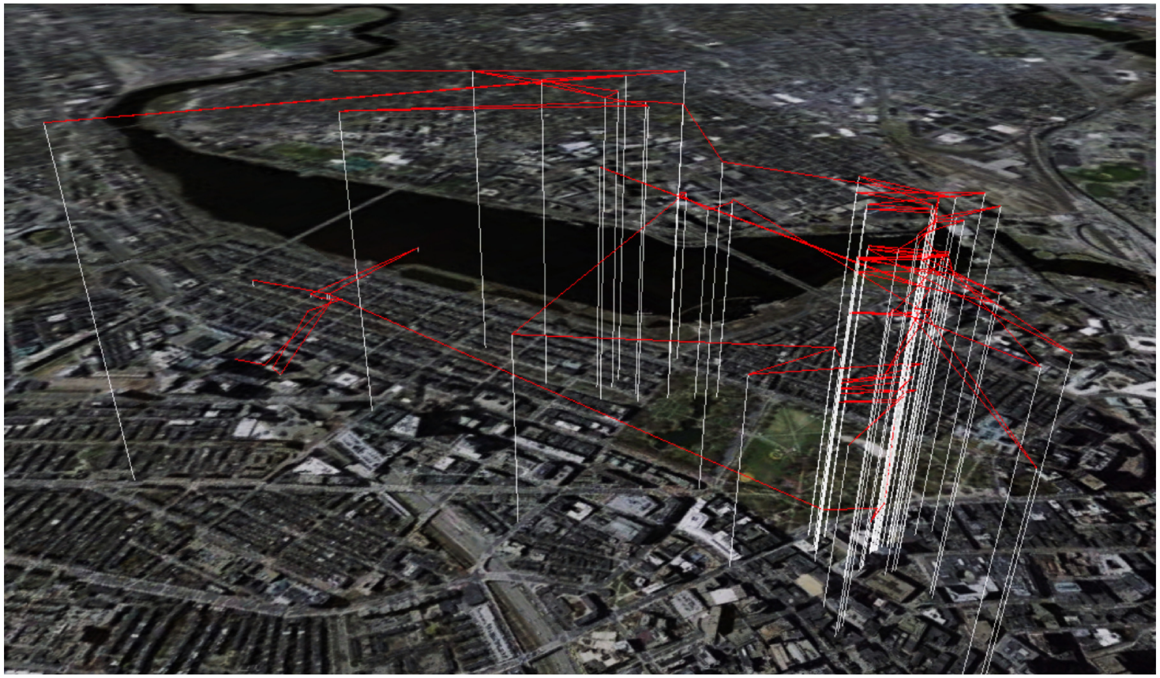
Angle of Arrival (AOA)



Timing Advance (TA)



Received Signal Strength (RSS)

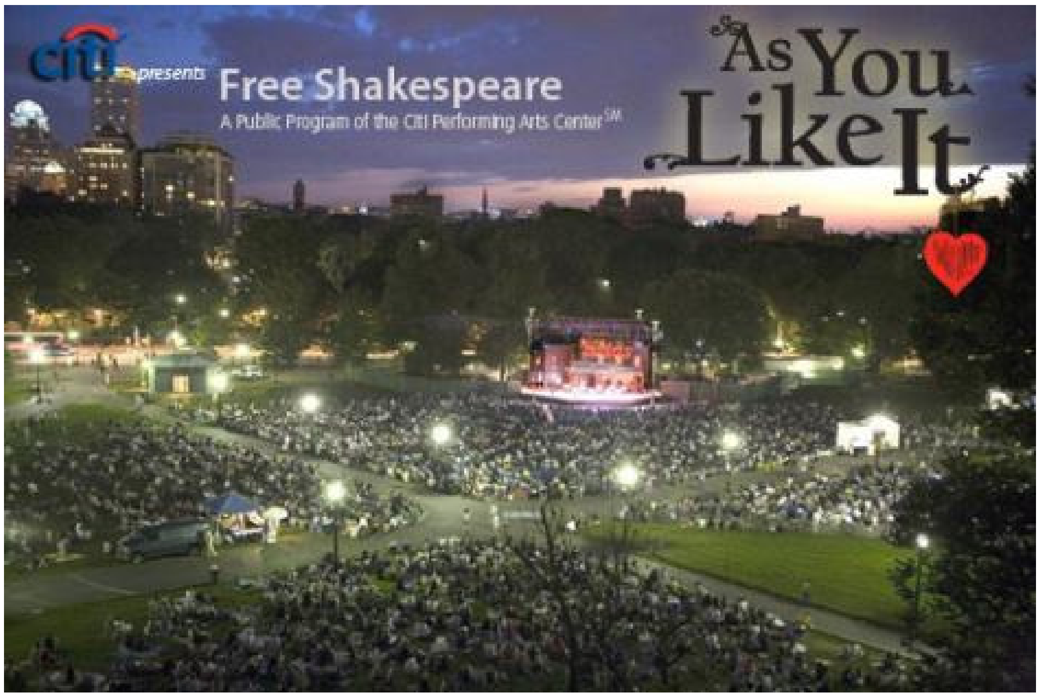


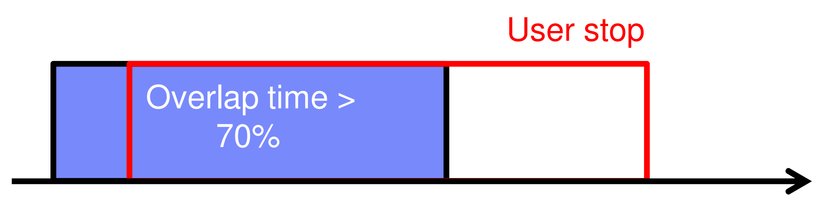
Example of extracted trajectory over 1 week

F. Calabrese, M. Colonna, P. Lovisolo, D. Parata, C. Ratti, Real-Time Urban Monitoring Using Cell Phones: a Case Study in Rome, IEEE Transactions on Intelligent Transportation Systems, 2011.

How social events impact mobility in the city

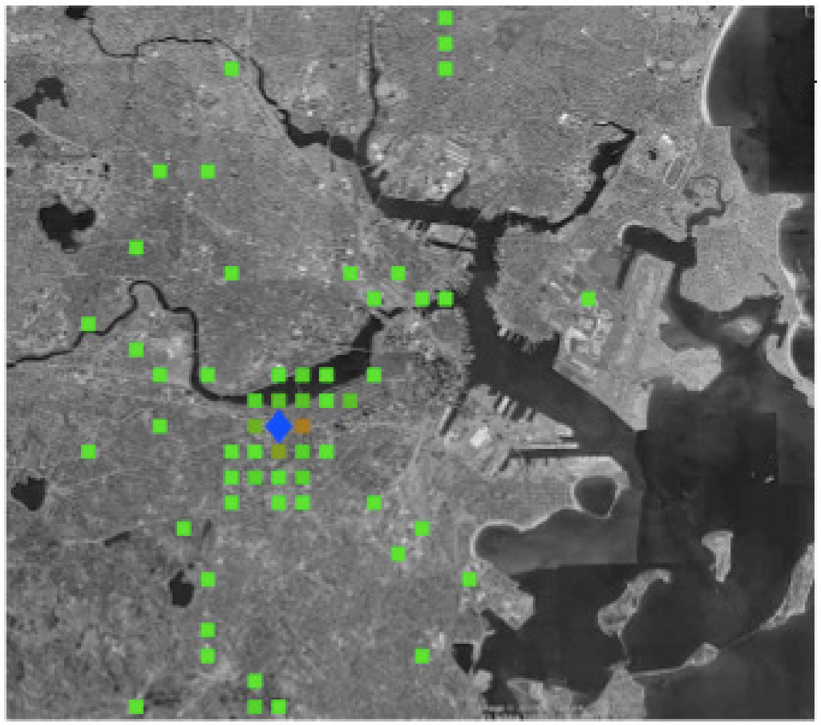
Modeling and predicting non-routine additive origin-destination flows in the city



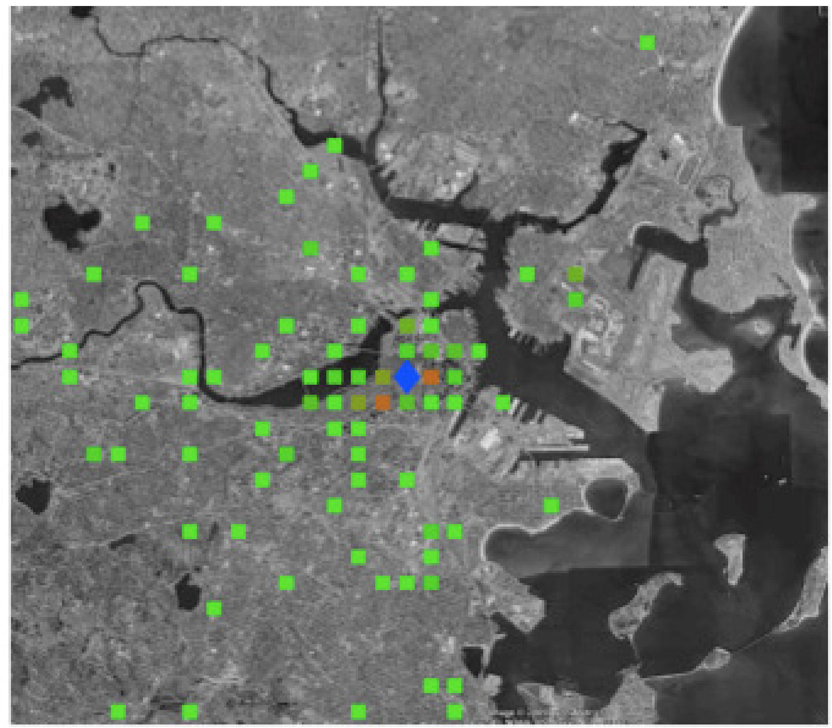


F. Calabrese, F. Pereira, G. Di Lorenzo, L. Liu, C. Ratti, “The geography of taste: analyzing cell-phone mobility and social events”, In International Conference on Pervasive Computing, 2010.

Detecting and predicting travel demand



(b) Boston Red Sox vs. Baltimore

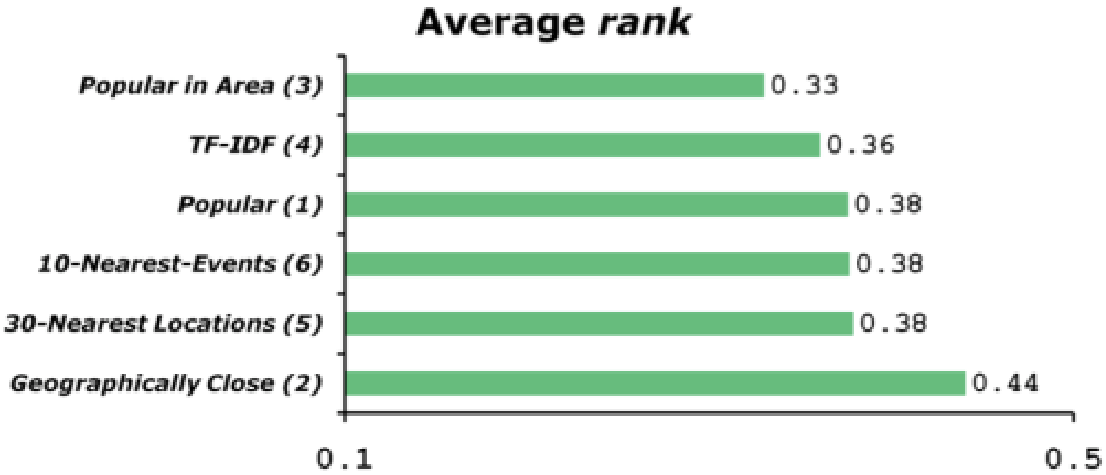


(d) Shakespeare on the Boston

Applications

- Improving event planning & management
  - Predicting the effect of an event on the urban transportation
  - Adapting public transit (schedules and routes) to accommodate additional demand

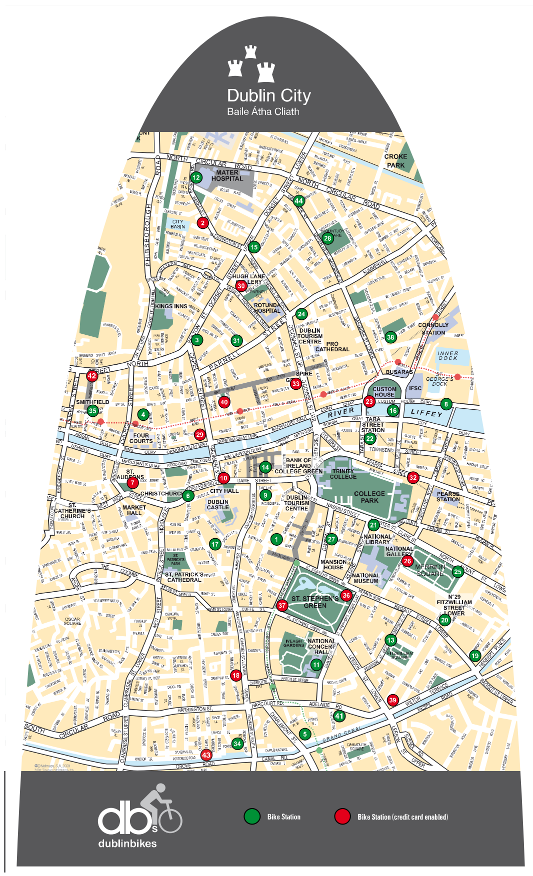
- Location based services
  - Recommending social events
  - Cold start problem



Modeling Urban Mobility: bike sharing system

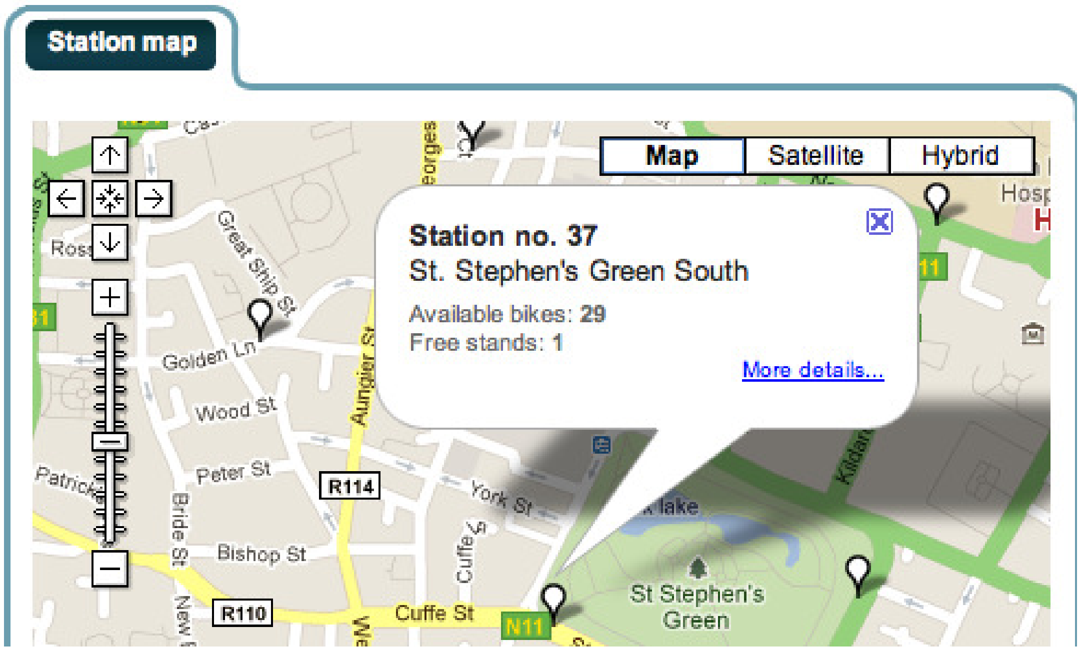
Bike sharing systems

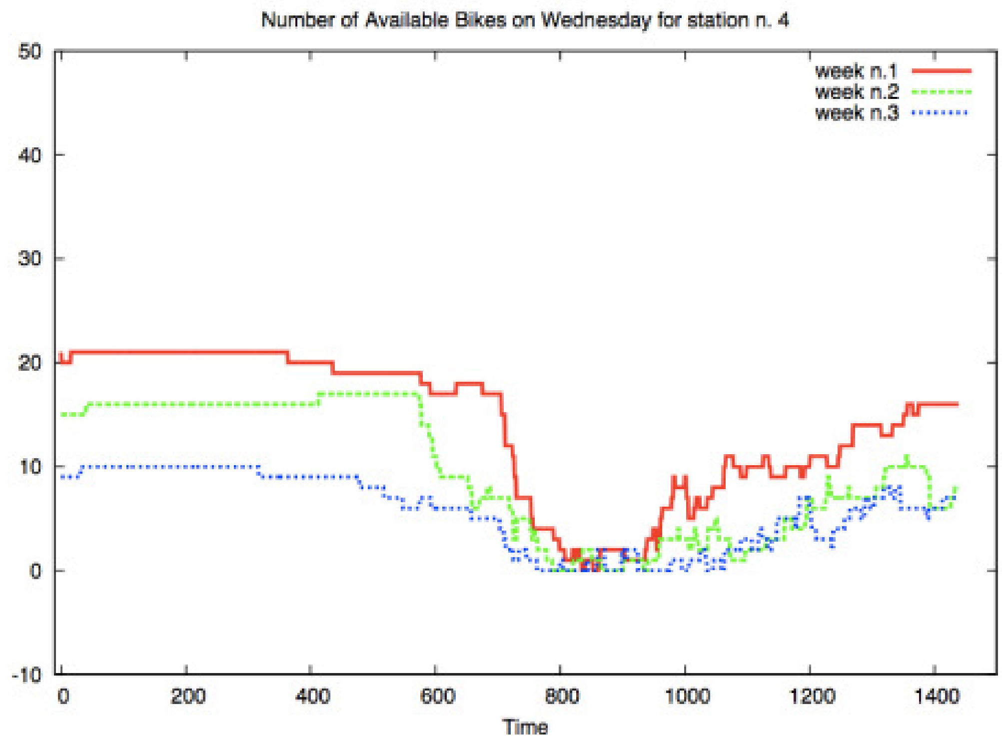
- Implemented in many cities, starting from europe
- Used by locals and tourists
- Reducing private and other public transport demand



Modeling Urban Mobility: Spatio-Temporal Patterns

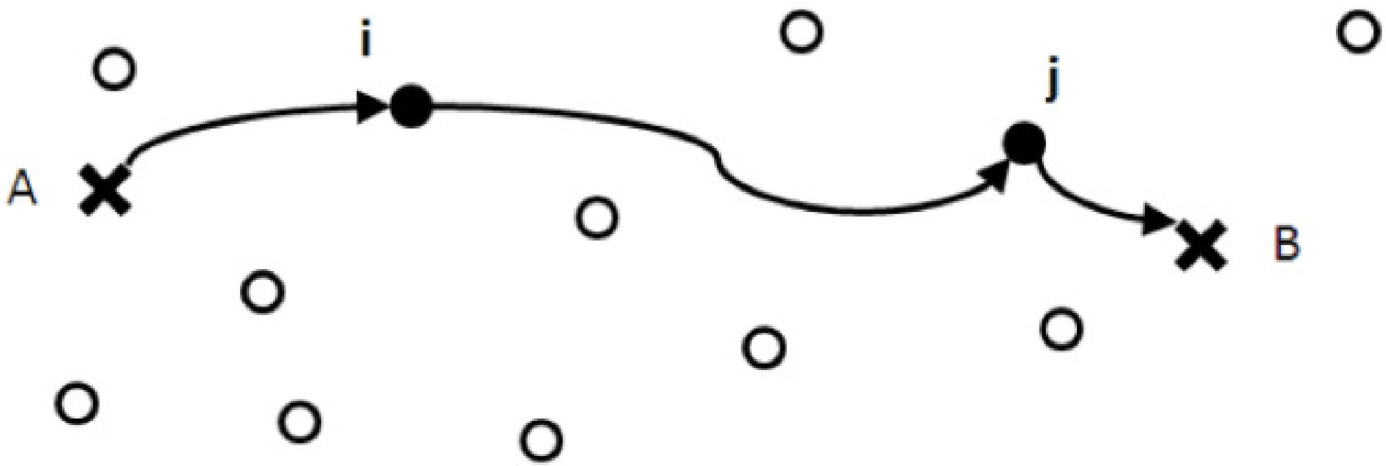
- Analyze spatio-temporal pattern of bike availability
- Infer correlation between stations (origin and destination of bike rides)
- Predict, long and short term
  - Number of available bikes
  - Number of available returning spots

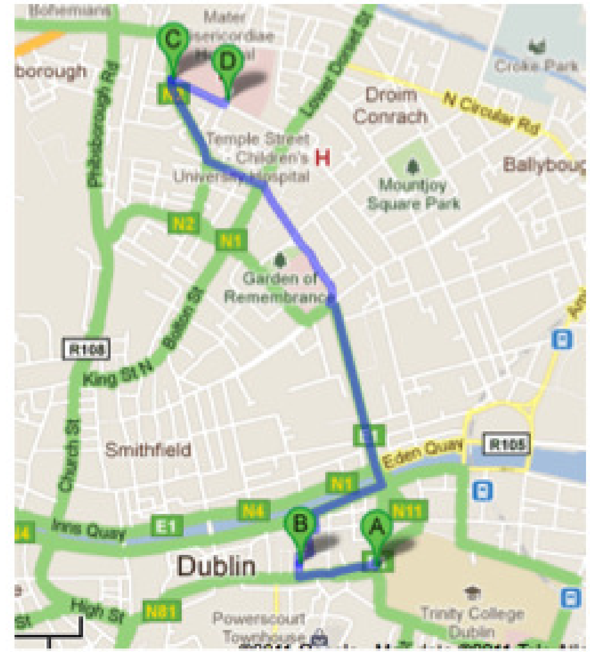




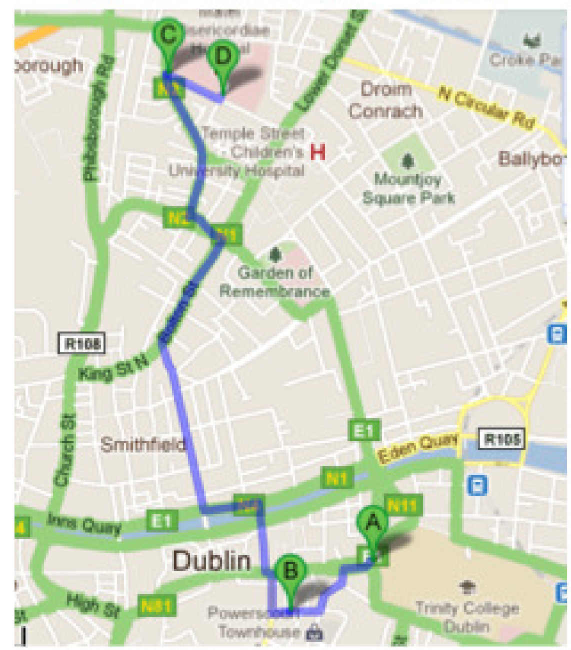
Modeling Urban Mobility: journey advisor

- Build a journey advisor application able to suggest which station to use to
  - Minimize travel time
  - Maximize probability to find and return bike





(a) the best ($t = 600$)



(c) the best ($t = 800$)

Outline

Sensor data assimilation

- Continuous assimilation of real-time traffic data

Understanding/Modeling human demand

- Characterizing urban dynamics from digital traces

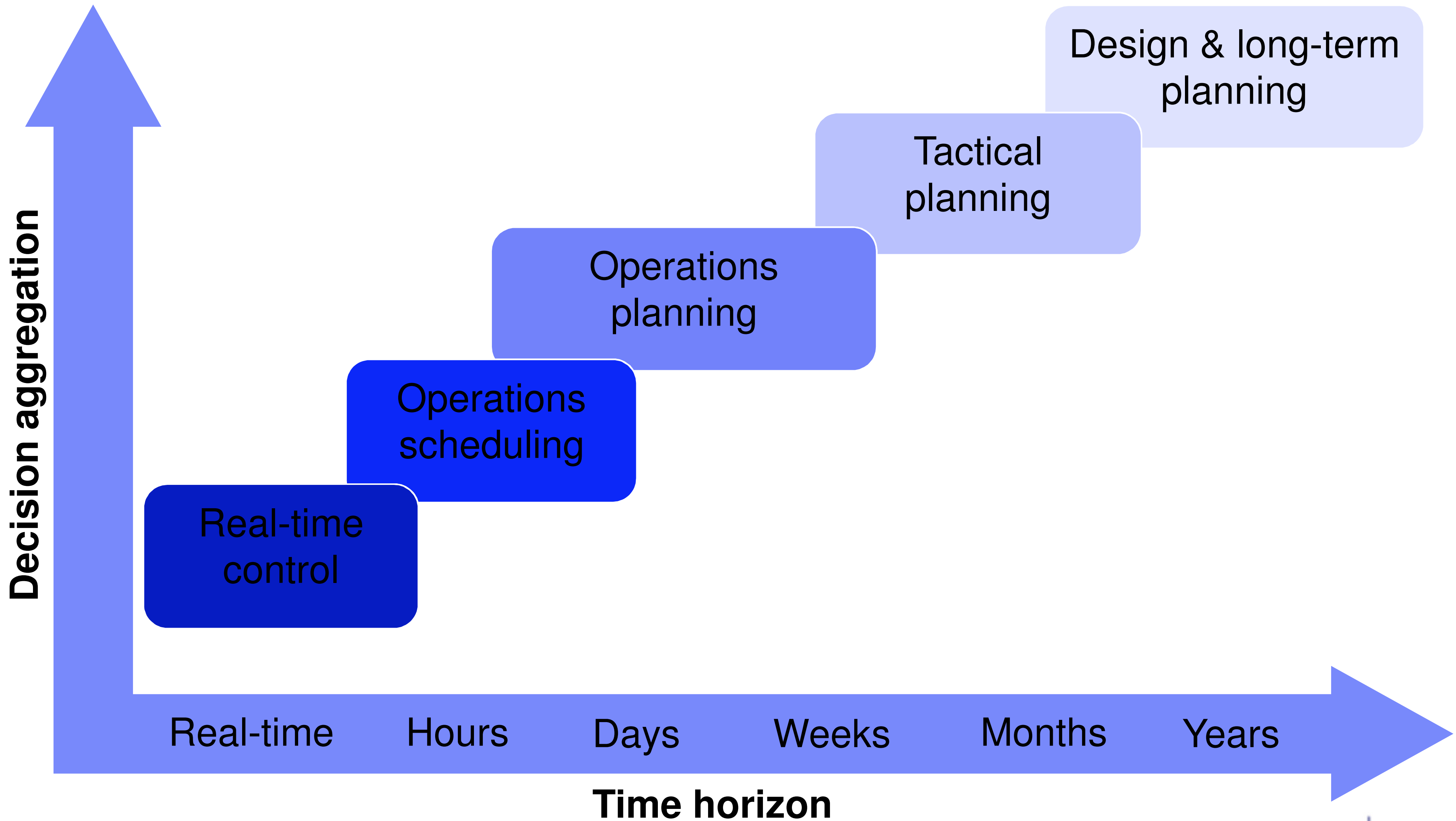
Operations & Planning

- Leveraging mathematical programming for planning in an uncertain world

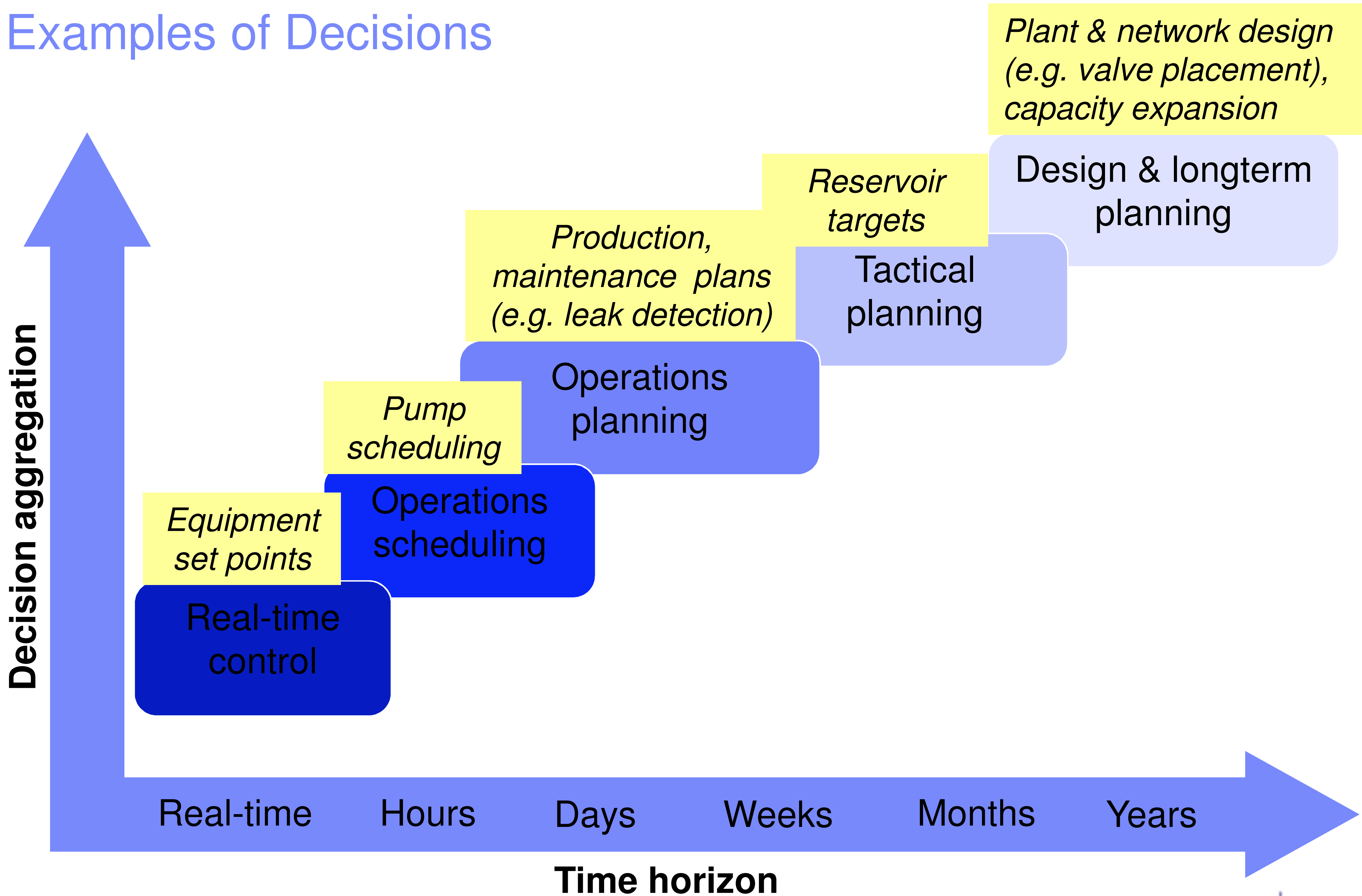
Overview

- Design and planning of urban infrastructures
  - Transportation
  - Water distribution and treatment
  - Energy
- “Standard” optimization approaches minimize costs while meeting demand
- Additional environmental objectives
  - Minimize carbon footprint
  - Meet pollution reduction targets
- Additional challenge – capturing uncertainty, such as:
  - Population growth and urban dynamics
  - Rainfall
  - Renewable energy sources
  - Energy costs

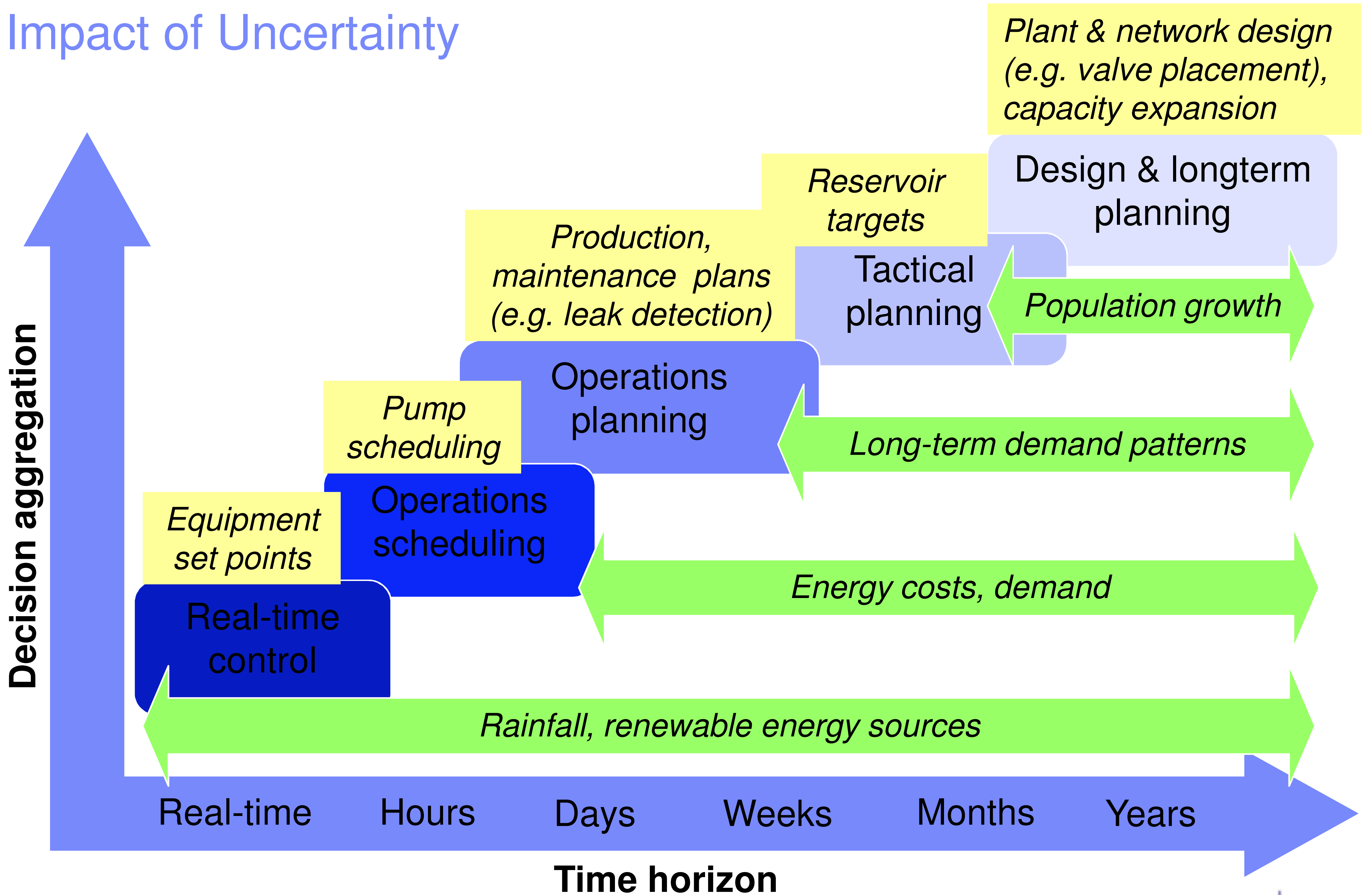
Planning Levels



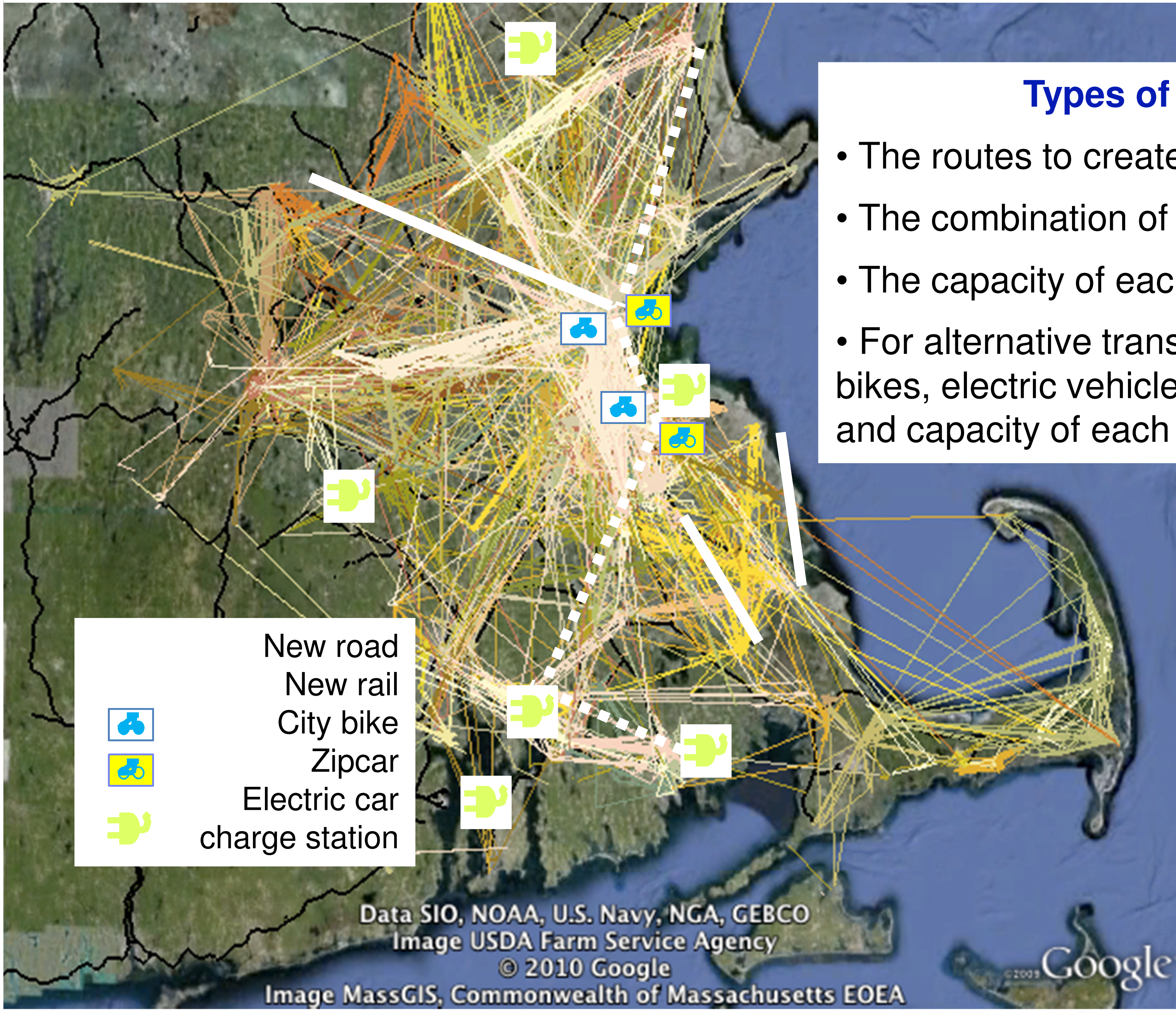
Examples of Decisions



Impact of Uncertainty



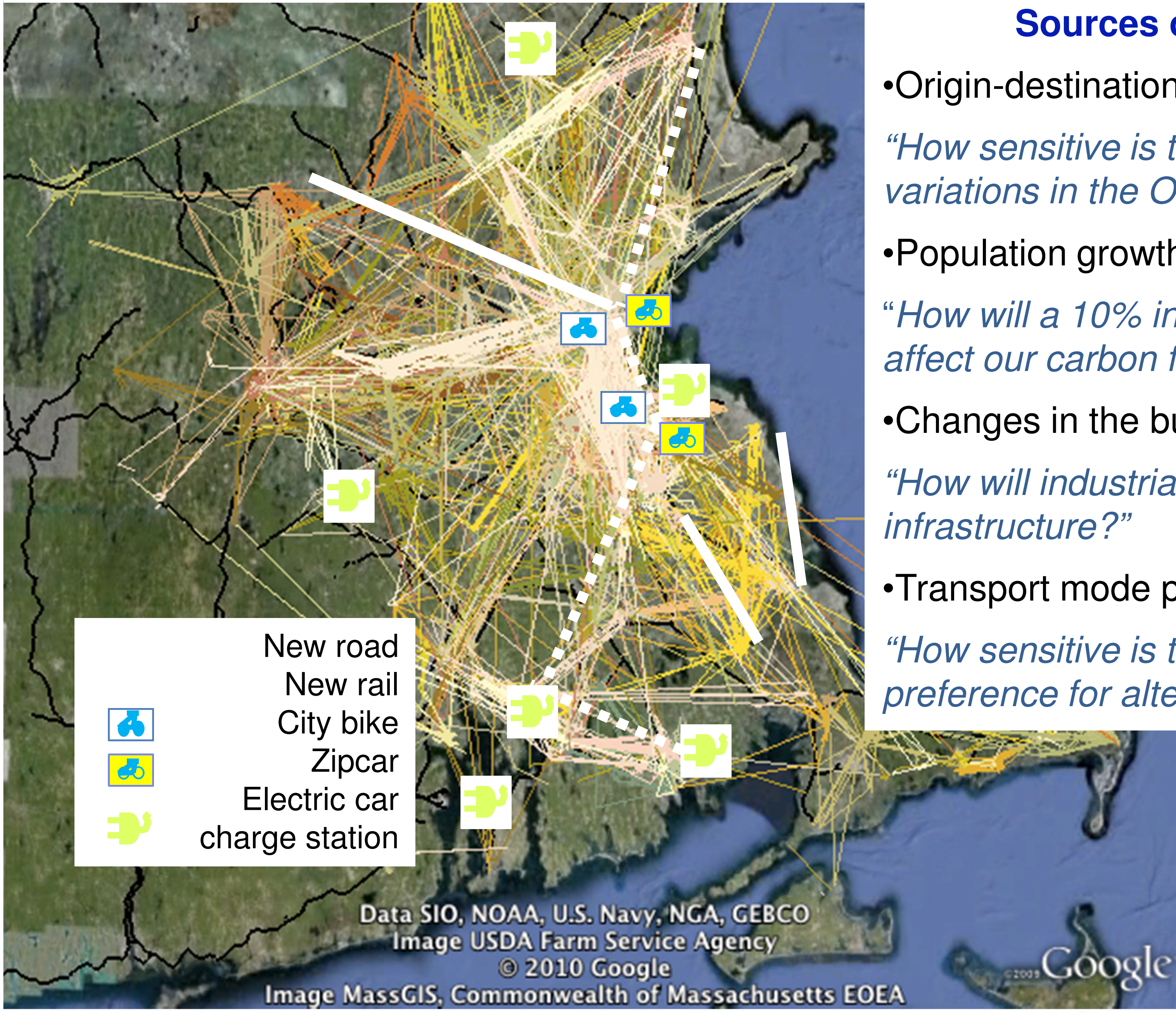
Example: Transportation infrastructure



Types of decisions:

- The routes to create or expand
- The combination of transport modes
- The capacity of each route
- For alternative transport (e.g. "zipcar", city bikes, electric vehicle stations), the location and capacity of each station

Example: Transportation infrastructure



Sources of uncertainty:

- Origin-destination matrices

"How sensitive is the investment plan to variations in the O-D matrices?"

- Population growth

"How will a 10% increase in population affect our carbon footprint?"

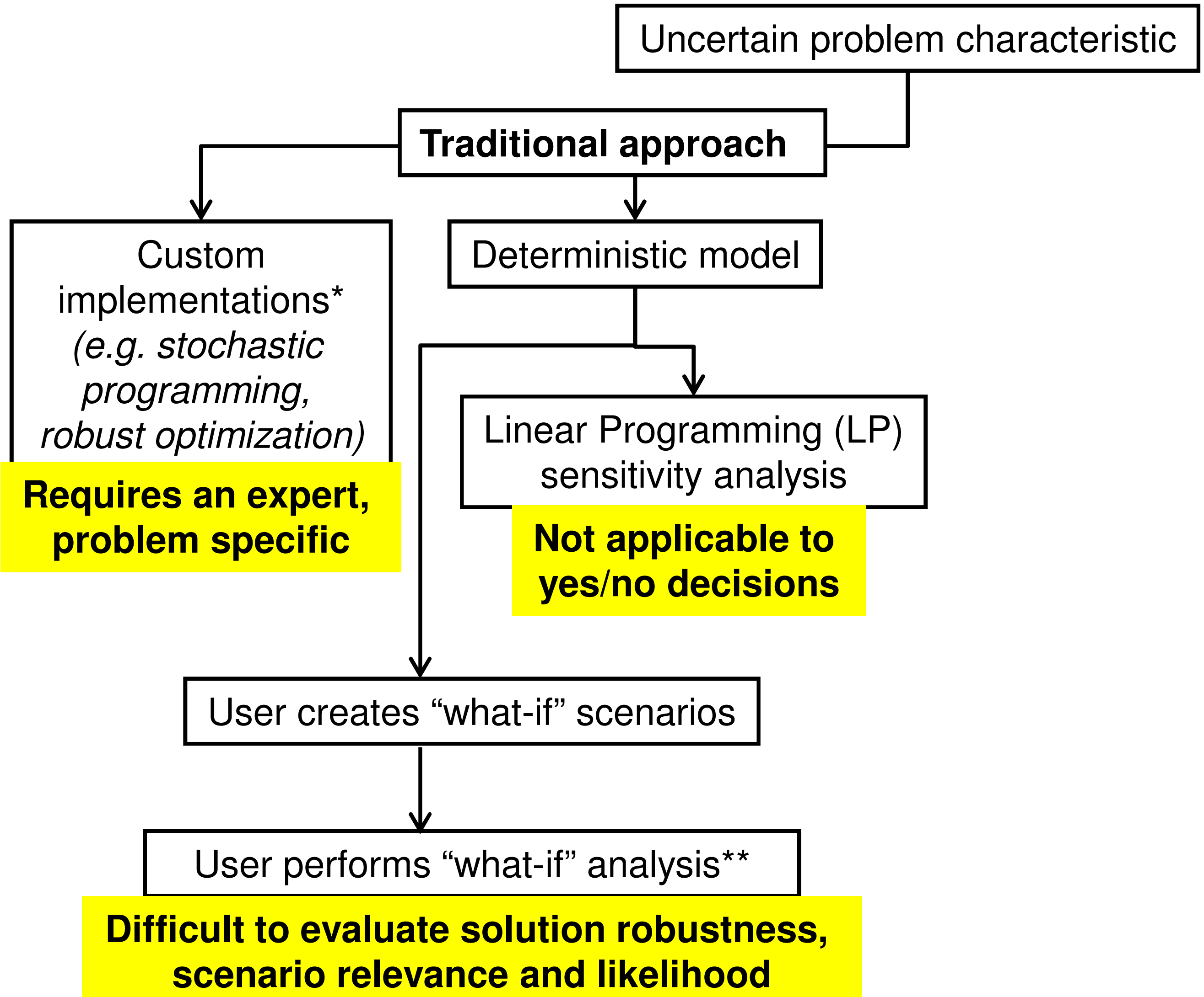
- Changes in the built environment

"How will industrial expansion affect the infrastructure?"

- Transport mode preference

"How sensitive is the plan to people's preference for alternative transport?"

Traditional vs. Proposed Approach



Traditional vs. Proposed Approach

Uncertain problem characteristic

Traditional approach

- Custom implementations* *(e.g. stochastic programming, robust optimization)*
  - **Requires an expert, problem specific**
- Deterministic model
  - Linear Programming (LP) sensitivity analysis
    - **Not applicable to yes/no decisions**
  - User creates "what-if" scenarios
    - User performs "what-if" analysis**
      - **Difficult to evaluate solution robustness, scenario relevance and likelihood**

Proposed approach

Part 1:
Scenario creation tool

Part 2:
Uncertainty & sensitivity analysis tool

Part 3:
Reports

- Solutions ranked by optimality, robustness, and feasibility
- Solutions satisfying user-specified criteria (e.g. plan with < 10% likelihood of failure)
- Reports highlight solution weaknesses

Generalized tool for non-experts: ease-of-implementation and ease-of-use

Challenges

- Capturing and generalizing user requirements
- Identifying and comparing best existing methods for
  - Scenario creation
  - Uncertainty and sensitivity analysis (e.g. stochastic programming, robust optimization, simulation, genetic algorithms)
- Researching new methods where current methods are lacking

How can we help cities achieve their aspirations?

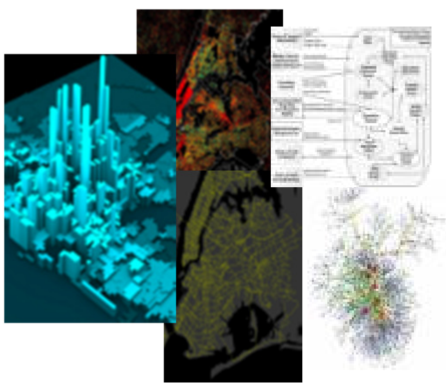
✓ **Sensor data assimilation**

From noisy data…

… to uncertain information

✓ **Modeling human demand**

Capturing uncertainty



✓ **Operations & Planning**

Factoring in uncertainty

Thanks

Francesco Calabrese
fcalabre@ie.ibm.com

Publications

- The Connected States of America. Can data help us think beyond state lines?, **Time Magazine**, 11 April 2011
- F Calabrese, D Dahlem, A Gerber, D Paul, X Chen, J Rowland, C Rath, C Ratti, The Connected States of America: Quantifying Social Radii of Influence**, International Conference on Social Computing**, 2011.
- F. Calabrese, G. Di Lorenzo, L. Liu, C. Ratti, “Estimating Origin-Destination flows using opportunistically collected mobile phone location data from one million users in Boston Metropolitan Area”, **IEEE Pervasive Computing**, 2011.
- G. Di Lorenzo, F. Calabrese, "Identifying Human Spatio-Temporal Activity Patterns from Mobile-Phone Traces”, **IEEE ITSC**, 2011
- F. Calabrese, Z. Smoreda, V. Blondel, C. Ratti, “The Interplay Between Telecommunications and Face-to-Face Interactions-An Initial Study Using Mobile Phone Data”, **PLoS ONE**, 2011.
- D. Quercia, G. Di Lorenzo, F. Calabrese, C. Ratti, “Mobile Phones and Outdoor Advertising: Measurable Advertising”, **IEEE Pervasive Computing**, 2011.
- F. Calabrese, M. Colonna, P. Lovisolo, D. Parata, C. Ratti, “Real-Time Urban Monitoring Using Cell Phones: a Case Study in Rome”, **IEEE Transactions on Intelligent Transportation Systems**, 2011.
- L. Gasparini, E. Bouillet, F. Calabrese, O. Verscheure, Brendan O’Brien, Maggie O’Donnell, "System and Analytics for Continuously Assessing Transport Systems from Sparse and Noisy Observations: Case Study in Dublin”, **IEEE ITSC**, 2011
- A. Baptista, E. Bouillet, F. Calabrese, O. Verscheure, "Towards Building an Uncertainty-aware Multi-Modal Journey Planner”, **IEEE ITSC**, 2011
- T. Tchrakian, O. Verscheure, "A Lagrangian State-Space Representation of a Macroscopic Traffic Flow Model”, **IEEE ITSC**, 2011