AALBORG UNIVERSITY

Mobile Services: Background, Platform, and Applications — Using Streamspin as An Example

Hua Lu

www.cs.aau.dk/~luhua

Outline

- Background
- Streamspin
- Tracking of Moving Objects

Background

- Technology Advances
  - Positioning technologies
  - Wireless communication
  - Mobile devices
- Geo-Enabled Mobile Services
  - Traditional client/server based
  - Non-traditional

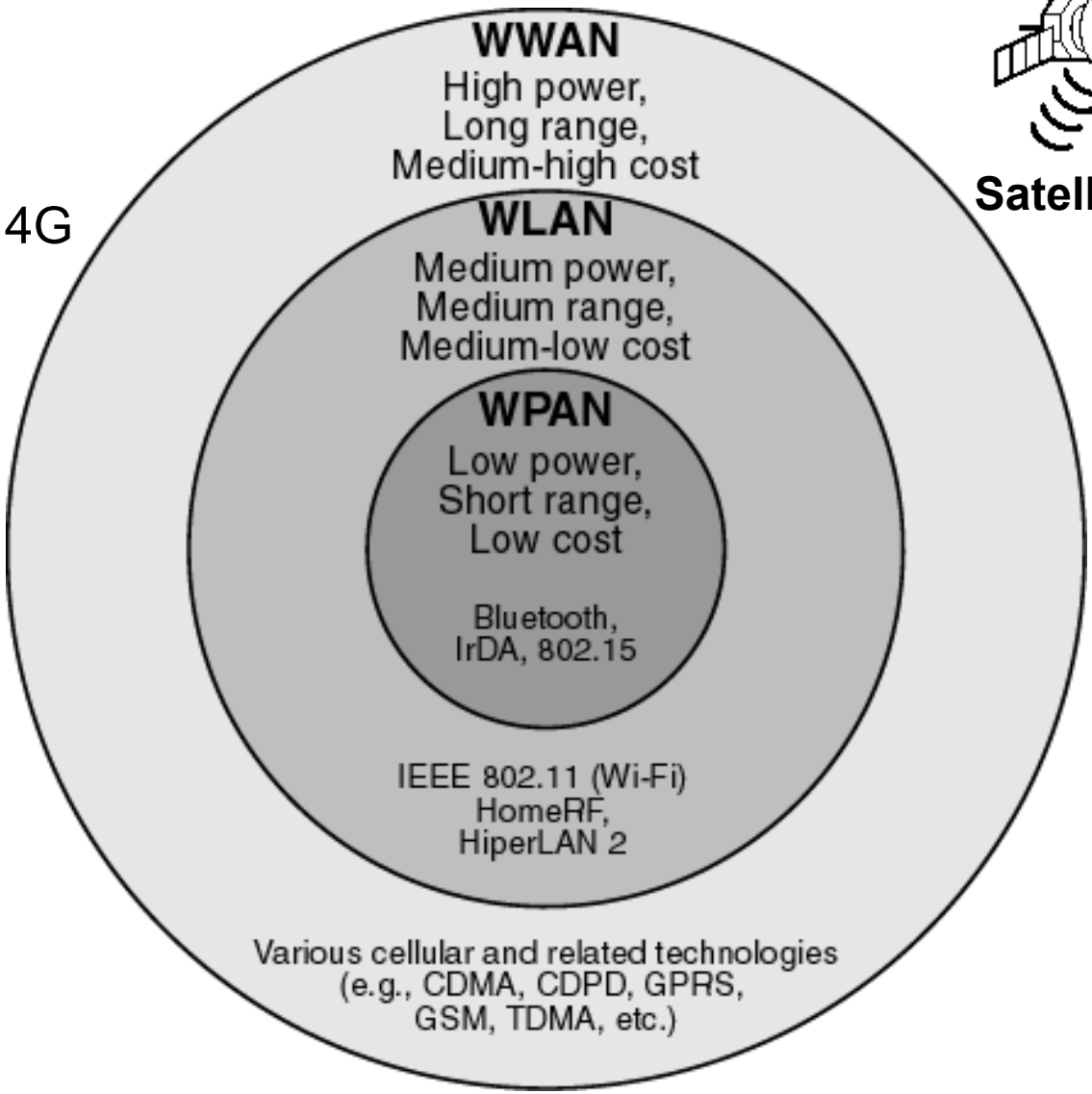
Positioning Technologies

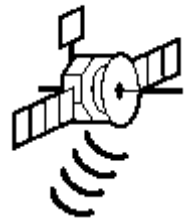
- Outdoor space
  - GPS
  - Galileo
  - WiFi
- Indoor space
  - InfraRed
  - WiFi
  - RFID

- We will see some concrete applications in the late part of this presentation

Wireless Communication

- WWAN
  - Wireless WAN
  - Cellular systems
  - 1G -> 2G -> 3G -> 4G
- WLAN
  - Wireless LAN
- WPAN
  - Wireless Personal Area Networks

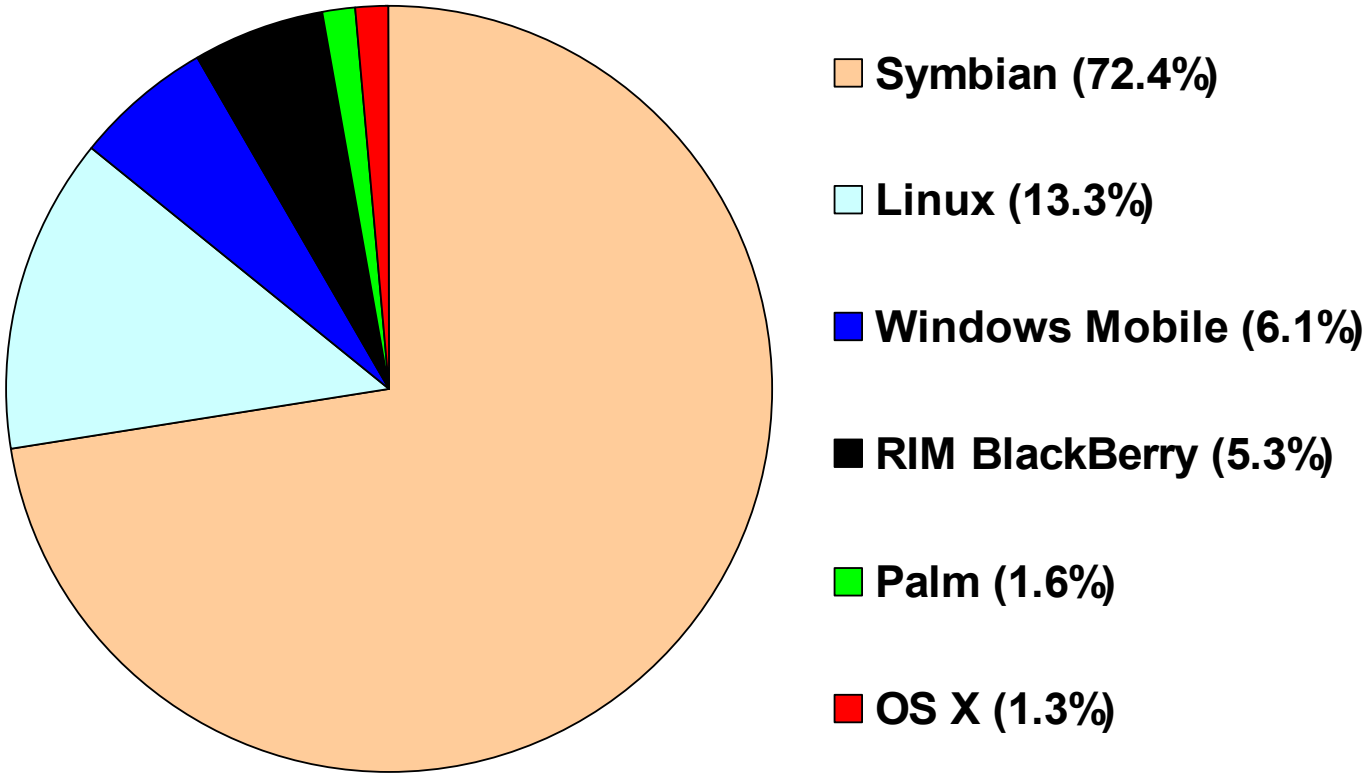




Mobile Devices

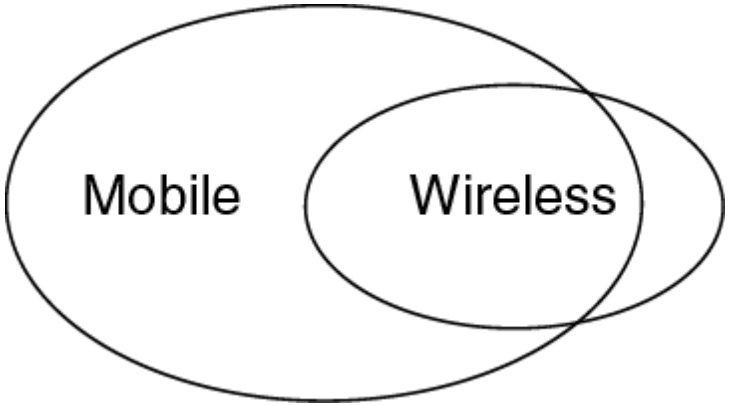
- Different kinds of handheld devices have started to evolve towards a common direction
  - Smartphones, PDAs, Laptops
- A currently cool mobile device setting
  - GPS receiver
  - Mobile phone functionality
  - Short range wireless
    - WiFi, bluetooth, InfraRed
  - Limited but decent computing capability
    - CPU, memory, storage
- Except drives from hardware advances, software also plays a crucial role
  - Android: open source mobile phone OS from Google
  - Symbian: Nokia is taking it open source too

Mobile Operating Systems



- Data from Canalys report “Worldwide smart mobile device market, Canalys Q2 2007”
- Google Android was not included

Variability of The Mobile Environment



"Mobile and Wireless Design Essentials" (Mallick 2003)

Mobility

- stationary
- nomadic (pedestrian speed)
- mobile (vehicular speed)
- roaming (mobile across networks)

Connectivity

- connected
- semi-connected (asymmetric)
- disconnected

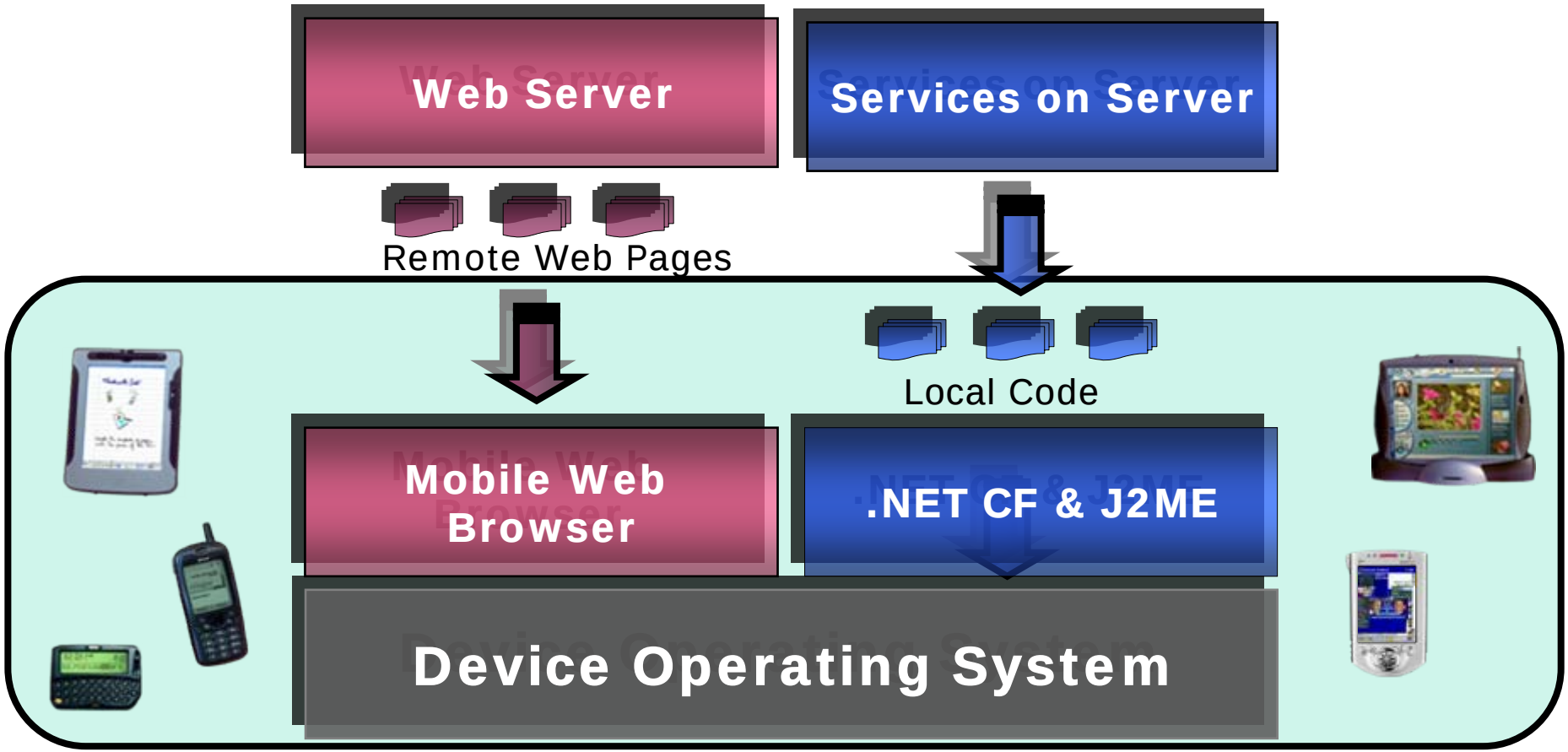
Mobile Device Capability

- form factor
- GUI
- multimedia
- real-time multimedia

Geo-Enabled Mobile Services

- Infrastructure
  - To manage all information effectively and efficiently
- Business content
  - Information services
    - Weather, stock, restaurant, … you name it
  - Deliver relevant information to right service users
- Architecture setting
  - Traditional client/server through long range wireless communications
  - Peer-to-peer via short range wireless communication
  - Hybrid model

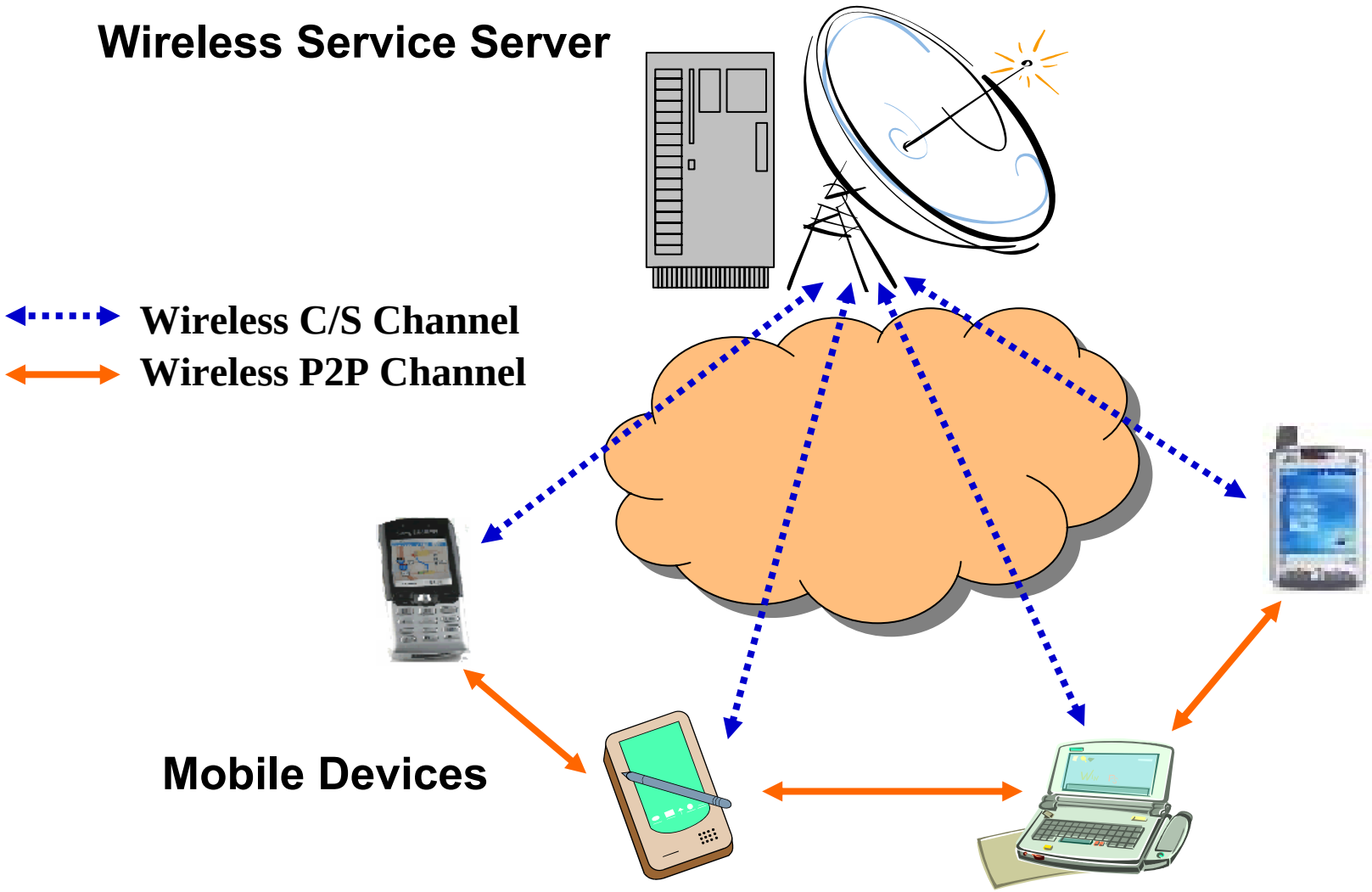
Client/Server Scenarios



Thin Client

Fat Client

Hybrid Architecture



Streamspin

- Motivation
  - Web 2.0
  - The Mobile Internet
- Mobile Services on Streamspin
  - A middleware system for service creation, sharing, and delivery.
  - A platform for activating research advances.

Web 2.0

- Crucial concepts about Web 2.0
  - Creativity
  - Information sharing
  - Collaboration among users
- Typical websites
  - Wiki
  - Blogs
  - Social-networking sites
    - Facebook, MSN space, MySpace
  - Video sharing sites
    - YouTube
- A new kind of data to manage
  - User-generated content

Web 2.0: User-Generated Content

- Primary content
  - Text – wiki's, blogs
  - Photos – e.g., Facebook, Flickr, Plazes
  - Video – e.g., YouTube
- Secondary content – feedback
  - Content about the primary content
  - Rating schemes – e.g., ratings of sellers and buyers at auctions, ratings of content
  - Tagging
    - Tag clouds, folksonomies

The Internet Is Going Mobile

- A mobile Internet infrastructure is emerging.
  - Mobile devices, e.g., mobile phones, laptops, cameras, MP3 players, navigation systems, etc.
  - Communication networks
  - Users with access
- Technologies are emerging that enable the accurate geo-positioning of all objects we care about.
  - Network-assisted GPS reduces power consumption.
  - Galileo is underway.
- We are at the unique point in history where the Internet is going mobile.
  - The mobile Internet will be “bigger” than the conventional Internet.
    - Why?

The Mobile Internet Is Important

Mobile phones are cheaper than PCs, there are three times more of them, growing at twice the speed, and they increasingly have Internet access. What is more, the World Bank estimates that more than two-thirds of the world's population lives within range of a mobile phone network. Mobile is going to be the next big Internet phenomenon. It holds the key to greater access for everyone - with all the benefits that entails.

Eric Schmidt

CEO, Google, [Financial Times, May 2006]

Service Types

- Transportation-related services
  - Road pricing generalized: payment based on where, when, and how much one drives; taxes, insurance
  - Spatial pay per use, or metered services
- “Safety”-related services
  - Warnings about slow-moving traffic or icy or slippery road conditions ahead
  - Monitoring of traffic offenders
- Games and ”-tainment” (edu-, info-, enter-) services
  - Treasure hunting (geocaching)
  - Paintball (Botfighters)
  - Catch the monster (Raygun)
  - Tell me about that!

Mobile Is Different

- The conventional Internet
  - Computers with large screens and convenient qwerty keyboards
  - In controlled environments, at home or at work
- The mobile Internet
  - Small screens, inconvenient keyboards
  - The user is out and about – yields high variation in use situations
    - In a meeting or at a café
    - On the move, e.g., on foot, using collective transport, driving a car
  - Disruptive surroundings
  - Service use is often not the primary activity
    - Assist the user in accomplishing the primary activity
  - Push services
  - Delivery of the right service at the right time is important.

Context Awareness

- Demographic user data
  - Age, gender, marital status, job, etc.
- Users may define profiles that may be (de-)activated
  - Interests and preferences, subscriptions
- A user's social network
  - Friends, colleagues
- Geo-context
  - Current location (and speed)
  - Destination and route for users on the move

- Ranges from static to dynamic; ranges from user supplied to automatic

Location Awareness

- The mobile Internet will be location-aware.
- According to Nokia, within a couple of years, GPS will be expected in mobile phones similarly to how a camera is expected today.
- It will be possible to accurately geo-position all objects we care about.
  - Galileo is underway.
  - Infrastructure-based positioning, WiFi, Bluetooth, RFID

Streamspin in a Nutshell

To create data management technology that enables sites that are for mobile services what YouTube is for video.

- Challenges
  - Enable easy mobile service creation
  - Offer advanced spatial and social context functionality
  - Enable service sharing with support for community concepts
  - Be an open, extensible, and scalable service delivery infrastructure
- The Streamspin project maintains an evolving platform that aims to serve as a test bed for exploring solutions to these challenges.
- Streamspin is a platform for activating research advances.

Streamspin.com

Service Development w/o Streamspin

- Create a server that sends content to and receive location data from mobile clients.
- Create a mobile client capable of communicating with the server.
  - Create a GPS component for the mobile client
- Handle client-server disconnects.
- Program the service, integrating with the existing infrastructure

- It is quite difficult, tedious, and error-prone

What Streamspin Offers

- With Streamspin, no mobile-client development — simply send the content using web services to the Streamspin server.
- All client communication is handled by Streamspin.
- Offers efficient and easy-to-use service-development functionality
  - Service creation
  - User subscription
  - Content push
  - Location tracking
- Client location updates are handled by the Streamspin project.

Service Creation

- Streamspin- or user-provided *templates* are available for service creation
  - Point-and-click service creation
  - Example templates: tour builder, e-mail, RSS push
- Streamspin-provided *web services* are available for creating custom services and templates
  - Content publishing
  - Service creation
  - Current location context for a user, using call back
  - Destination and route context, using call back (pending)
  - Location privacy (pending)
- Visual Studio *C# add-ins* for custom service creation
- Accessible from all web-service enabled languages

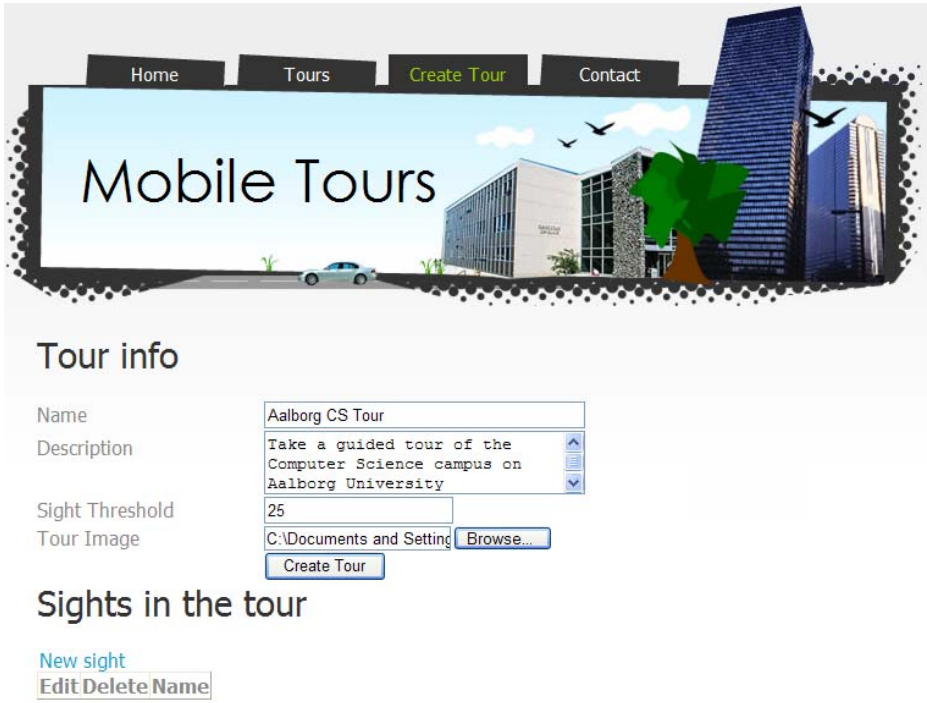
Service Sharing

- Content providers
  - Approval
  - Meta data: interest profile, location, location range, time to live, age range
- Publication of services in a service directory
- Interest hierarchy
  - Used by content providers for tagging
  - Used by service users for the specification of profiles
- Content rating
  - Recipients of content can rate the content
- Content discussions
  - Recipients of content can comment on the content and see the comments provided by other users
- Friends

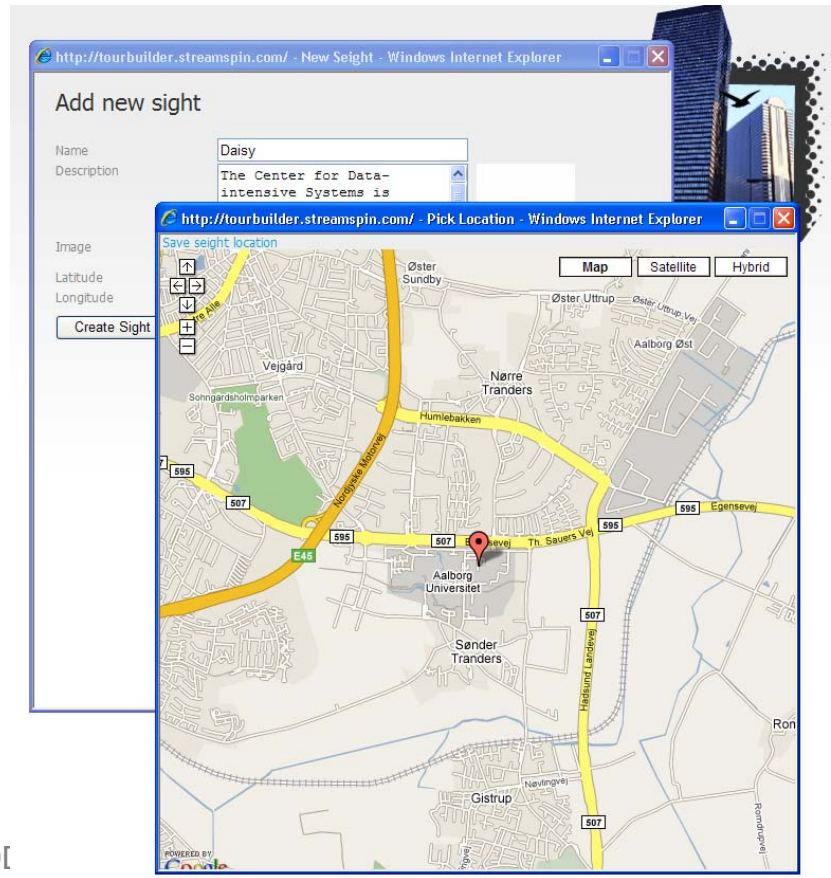
Service Delivery

- Distribution center content filtering
  - Based on the user's context and the meta data of the content
  - Publish/subscribe functionality
- Socket-based content push
  - Maintains socket-based connections to mobile clients.
  - No HTTP or web-service overhead (only TCP headers)
    - Text messaging at 2.5% of the normal cost, using current text messaging and GPRS pricing from DK
- Support for content that consists of a text header and a text (html) or URL body.
  - Text is used if the content is text.
  - Otherwise, a URL is provided that the client can then access.
- Very loosely connected architecture, supporting distributed computation

Streamspin Service Example

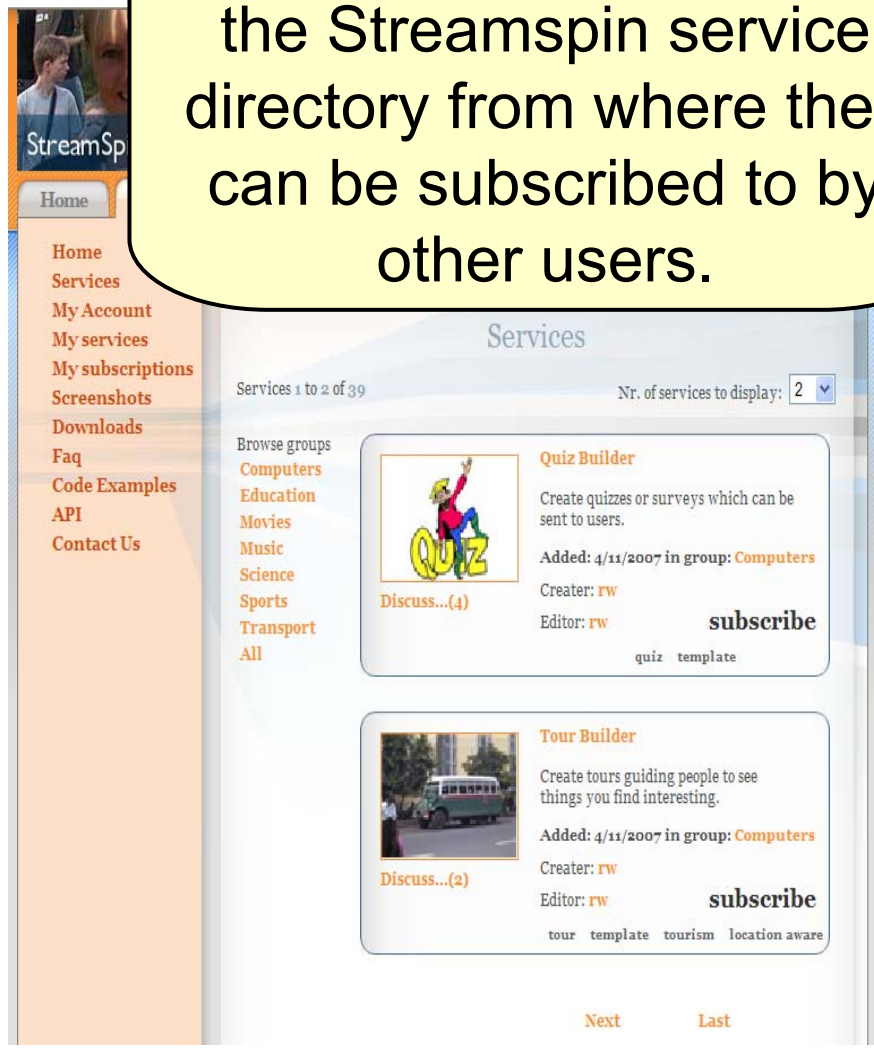


You can build and publish your own mobile service for any location-dependent applications



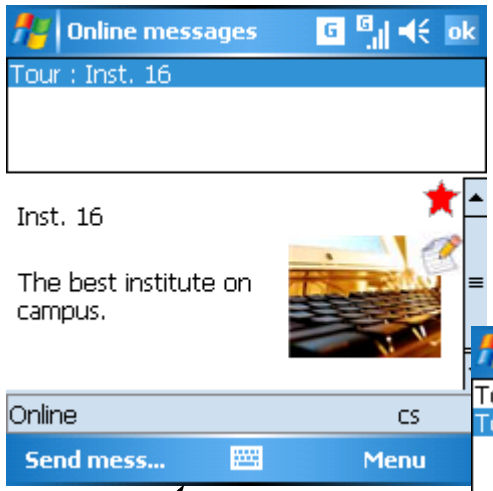
A mobile tour associates text and images with locations – the content is pushed to subscribers when they get close to the locations

The Streamspin Service Directory



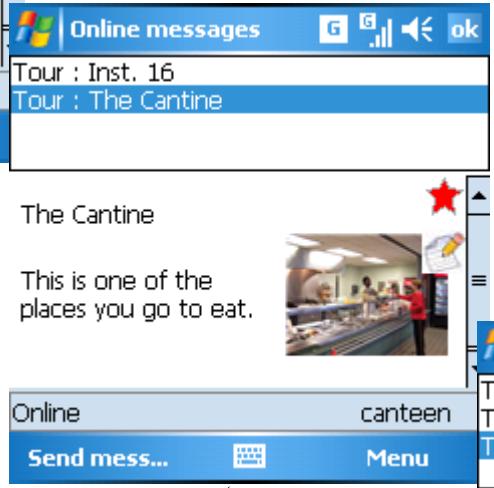
Services are published in the Streamspin service directory from where they can be subscribed to by other users.

Streamspin Service Example, Cont.

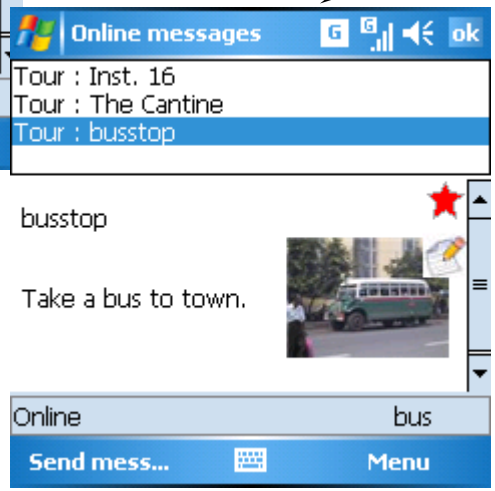


Screenshots from the phone of a user who has subscribed to the "Inst. 16" Tour

The user approaches the CS department



The user approaches the bus stop



The user approaches the canteen

Streamspin Summary

- Streamspin aims to apply Web 2.0 concepts to mobile services.
  - Easy creation and sharing of mobile services, scalable delivery of services.
- Online documentation
  - An introduction about Streamspin and how to create your own services.
    - http://www.streamspin.com/codeexamples/StreamSpindocumentation250907.pdf
  - A programmers guide to Streamspin
    - http://www.streamspin.com/codeexamples/harrysguidetostreamspin.aspx
- A detailed tutorial on Streamspin will be given later in this course

Tracking of Moving Objects

- Background and Motivation
  - Tracking basics
  - Services
- GPS-based outdoor tracking
  - System architecture
  - Communication reduction techniques
- RFID-based indoor tracking
  - Indoor space and RFID reader deployment
  - Pre-processing of raw trajectories
  - Off-line and on-line trajectory construction
  - Use of movement data for progressive trajectory construction

Trajectories and Tracking

- We assume that the extent of an object can be modeled as a point.
- The *trajectory* of a moving object is a total, continuous mapping from a time interval $[t_{start}, t_{end}]$ to point locations.
  - The current time t_{now} may belong to this interval.

- Tracking is about maintaining a *representation of the trajectory* of an object.
  - From t_{start} to the time of the last observation
  - From the time of the last observation to t_{now}.
  - From t_{now} to t_{end}.

Positioning and Sampling

- The maintenance is based on sampling.
  - A sample contains the time and a possibly inaccurate position.
- The positioning may be done in many ways.
  - GPS, Galileo.
  - WiFi.
  - Bluetooth.
  - Different RFID-based technologies.
  - Combinations.
  - Handset-based versus infrastructure-based.
- The sampling may be based on different events.
  - Time-based.
  - As needed to satisfy accuracy guarantee.
  - Presence in sensor range.
  - Arrival at timing point.

Infrastructures

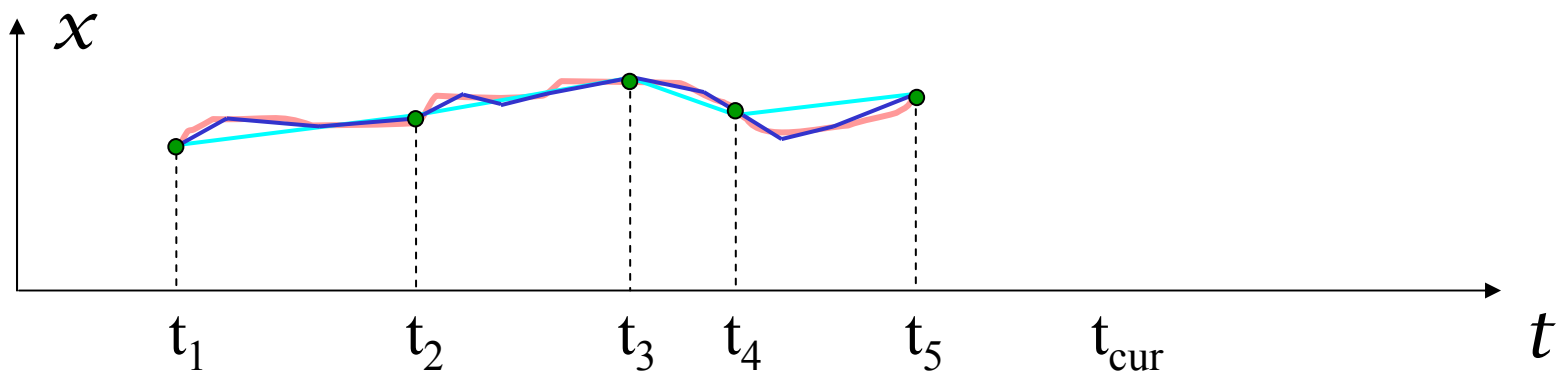
- The movement of an object occurs in an infrastructure.
- In outdoor space, the movement may be constrained.
  - By obstacles, termed constrained Euclidean space; relevant for, e.g., pedestrians and vessels at sea
  - By a spatial network; relevant for vehicles
- In indoor space, the movement of an object is constrained.
  - By the indoor space
- We will assume that a representation of the relevant infrastructure is available.
  - Maps, floor plans

Trajectory Representation

- Many representations of a trajectory are possible.
  - Depend on the infrastructure and applications.
- Examples
  - Outdoor Euclidean space: A polyline with accuracy guarantees
  - Spatial network: A polyline constructed from connected sub-segments from the network, again with accuracy guarantees
  - Indoor space: A “polyline” constructed from cells in indoor space, possibly decorated with room information, etc.
- All representations are inaccurate.
- Obtaining adequately accurate representation is often an objective.
- Reducing the cost of doing so is often an objective.
- The accuracy requirements may relate to the current position, or it may relate also to past or future positions.

Three Cases of Tracking

- Off-line
  - Given a set of samples, derive trajectory up until the time of the last sample.
- On-line
  - Maintain a representation of the current (and near-future) location.
- Prediction
  - Predict the future movement of an object.

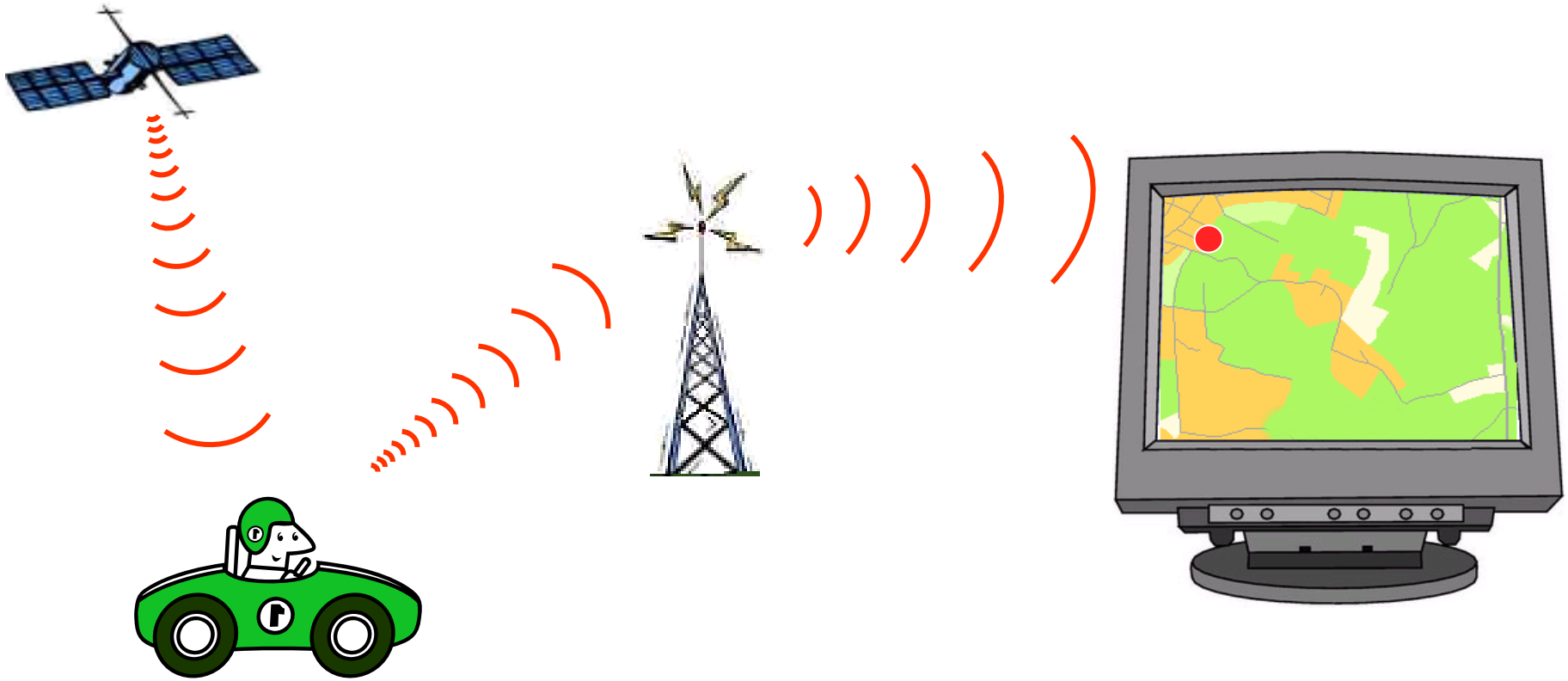


Why Tracking—Services

- Flow analyses
  - How do people use the indoor space
- Presence in pre-defined zones
  - Waiting times in lines
    - At the ticket counter
    - At security
  - How long do people spend in the tax free zone?
  - Travel times between zones
  - Heat map
- Frequent visitor analysis
- Movement patterns
- Security
- Indoor navigation
  - Assist an individual in finding the right train track
- Alerts
  - Alert an individual who is about to arrive late for a departure
- Nearby
  - Friends, service personnel
- Help!

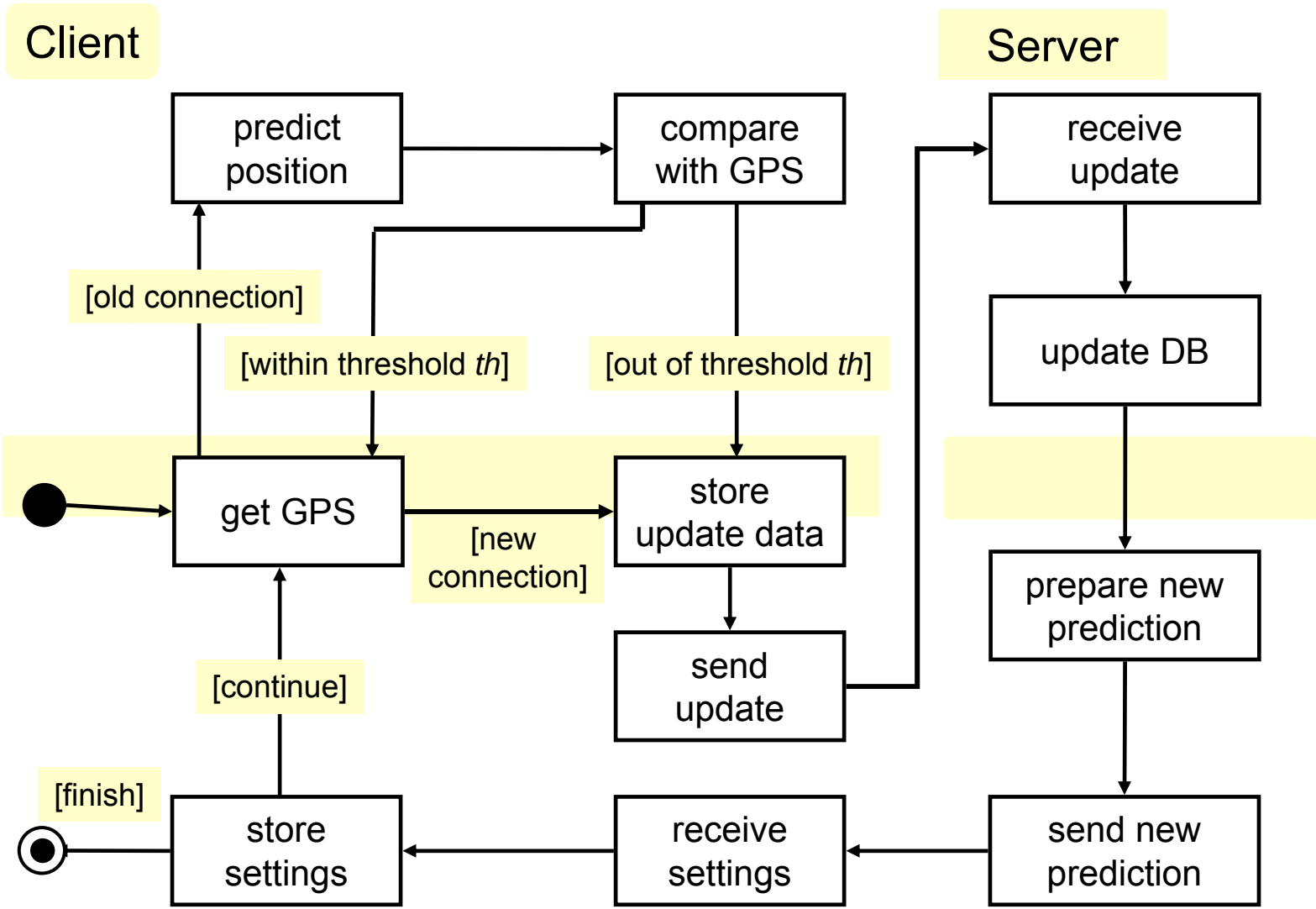
GPS-Based Outdoor Tracking

Aim: To track moving buses with accuracy guarantees



Objective: To reduce communication cost between moving client and server

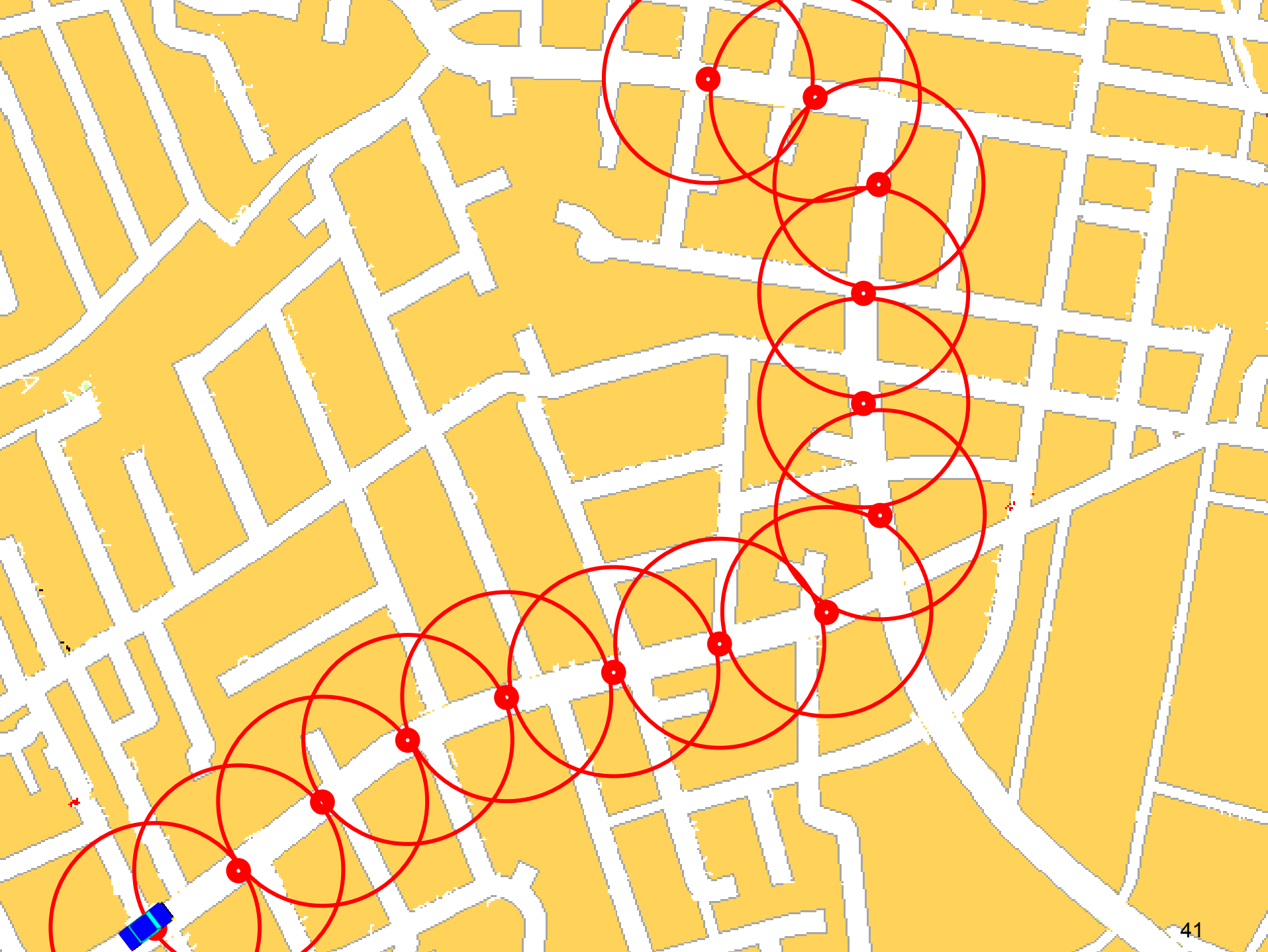
Tracking Approach

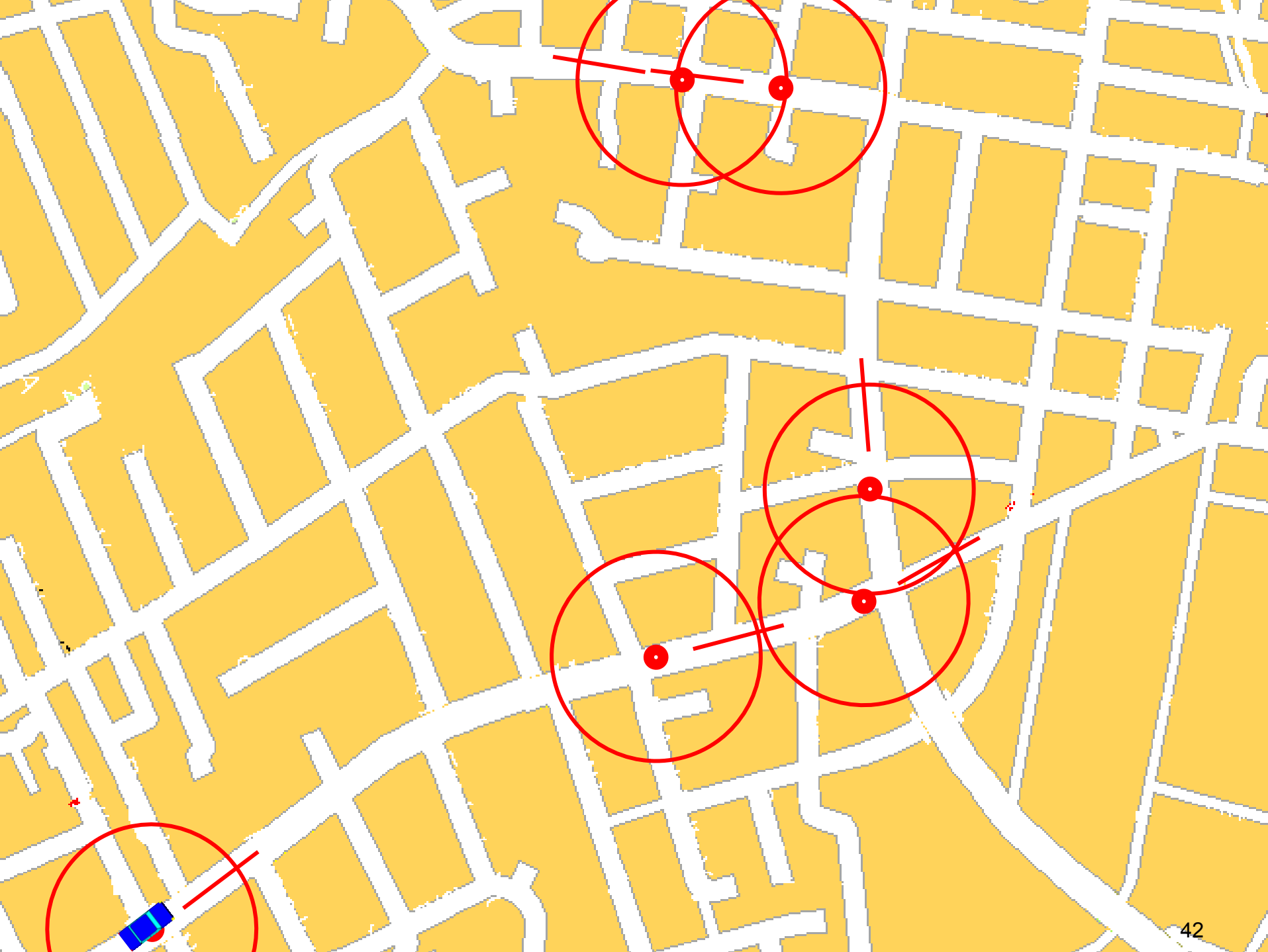


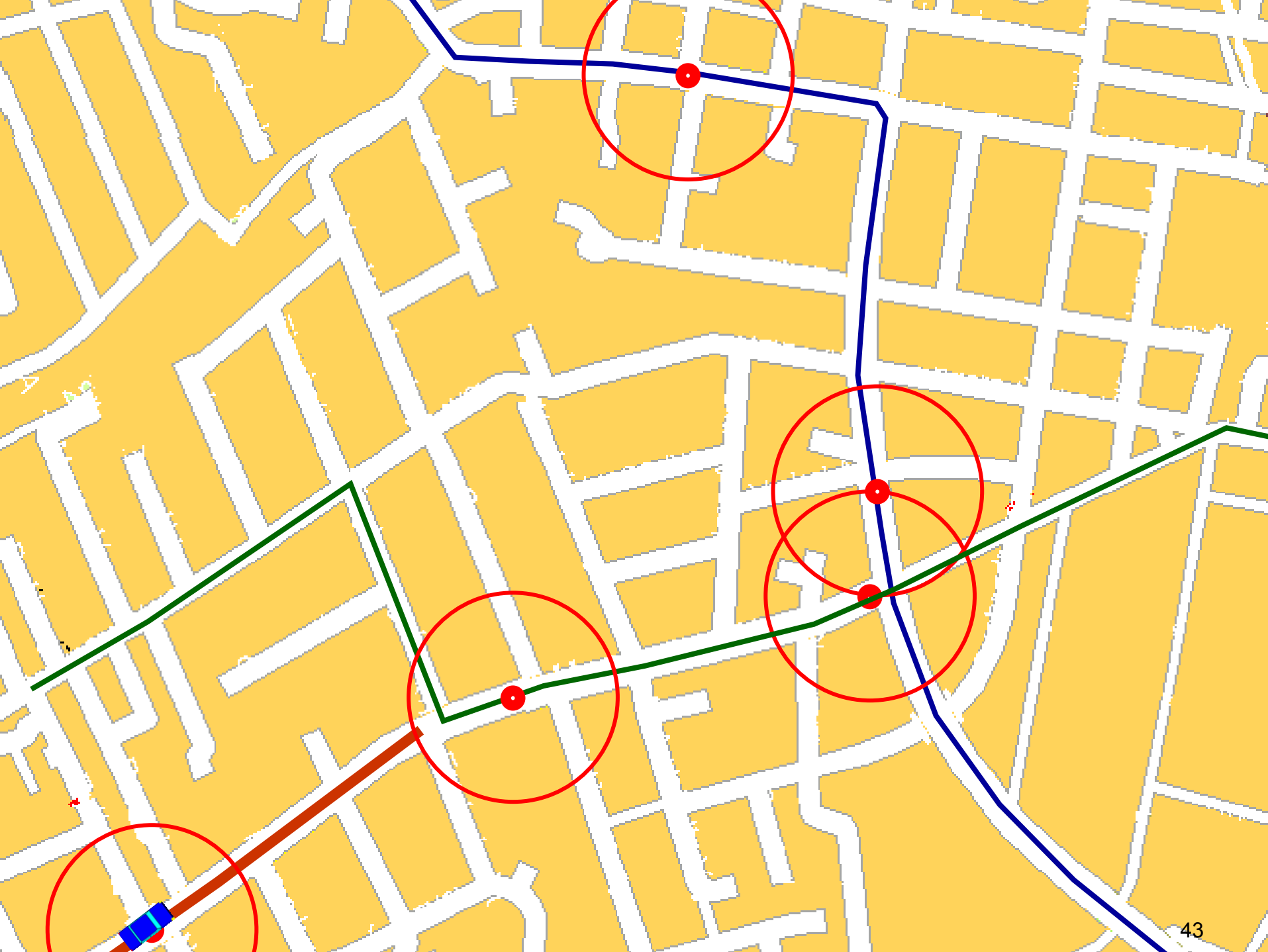
Client Side Update Policies

- Point based
  - As long as the client is within the circle centered at a known point, no updates sent to the server
    - The radius of the circle is the threshold
- Vector based
  - As long as the client's current position is still close enough to the vector, no updates sent to the server
- Segment based
  - Applicable to road network based moving client
  - As long as the client is still close enough to the identified road segment (a polyline), no update

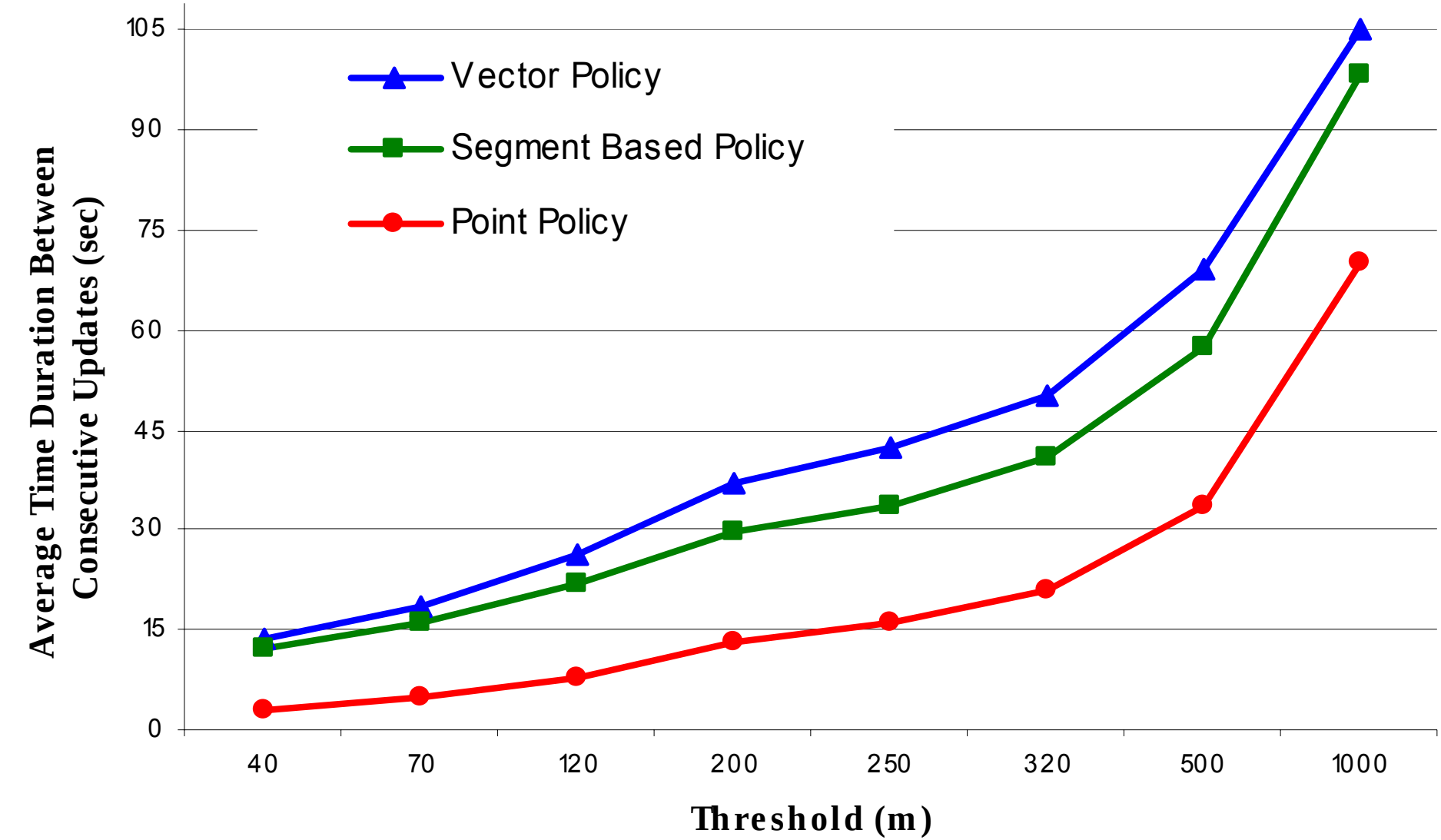
- Each policy means different tracking accuracies maintained on the server side







Comparison of Techniques

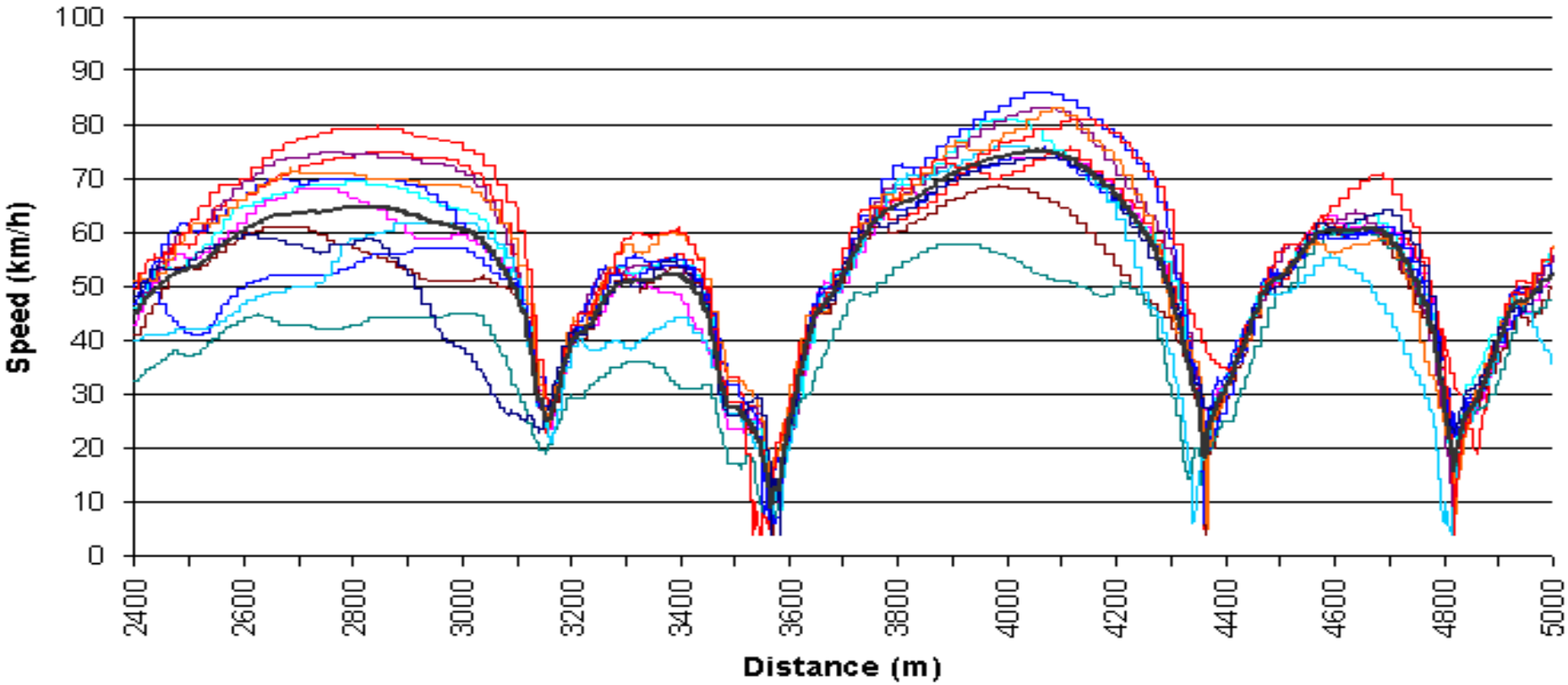


Use Other Knowledge

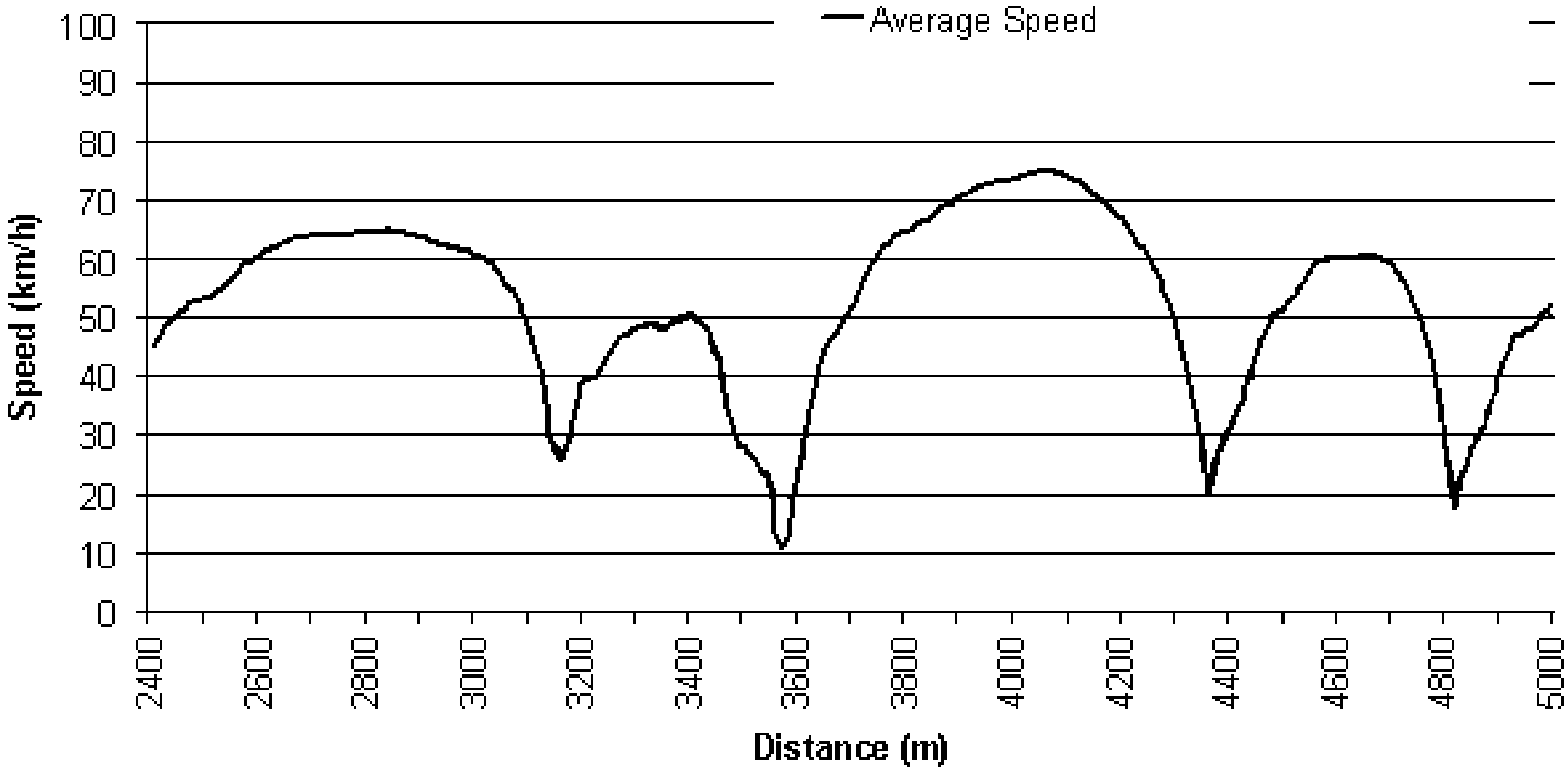
- We can make use other knowledge to further reduce the tracking cost, mainly communication cost

- Use of Routes
  - If we know the current route of a user, we can avoid segment changes altogether.
  - As routes are (long) segments, segment-based tracking works.
- Use of Acceleration Profiles
  - Repeated route traversals exhibit a clear speed pattern.
  - An acceleration profile is created for each route
    - Distance intervals with positive and negative acceleration are found using *average* speeds.
    - An average acceleration is calculated for each such interval.

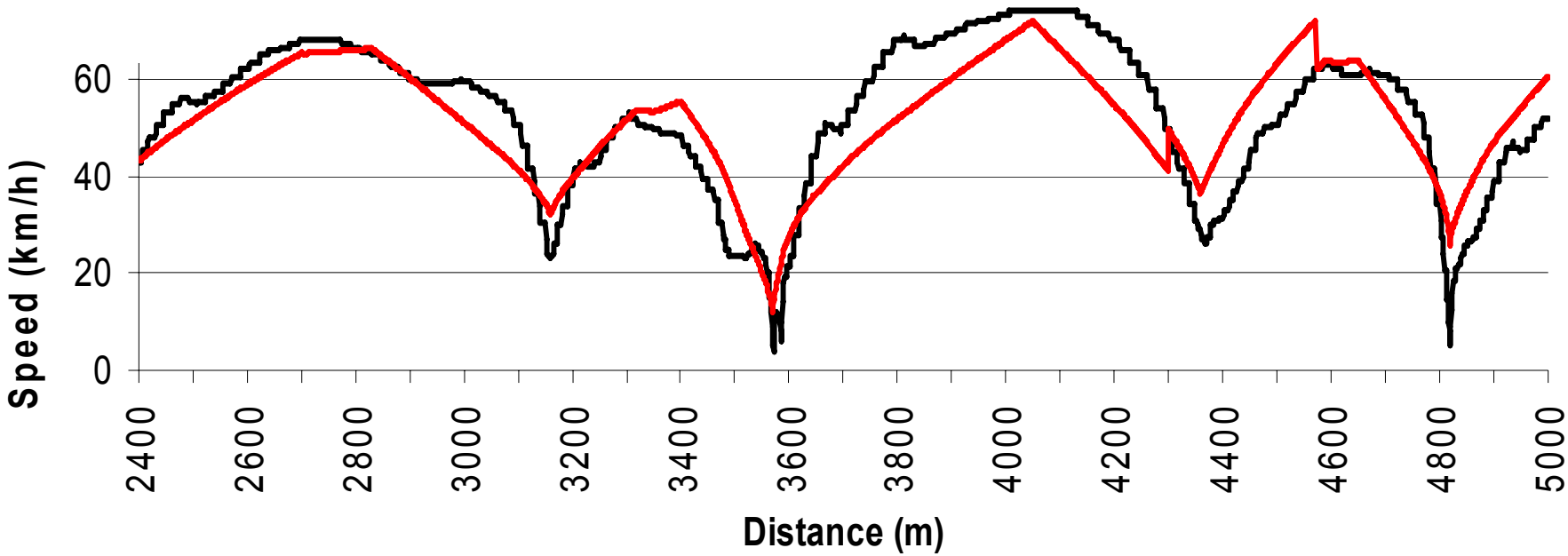
Speed Records



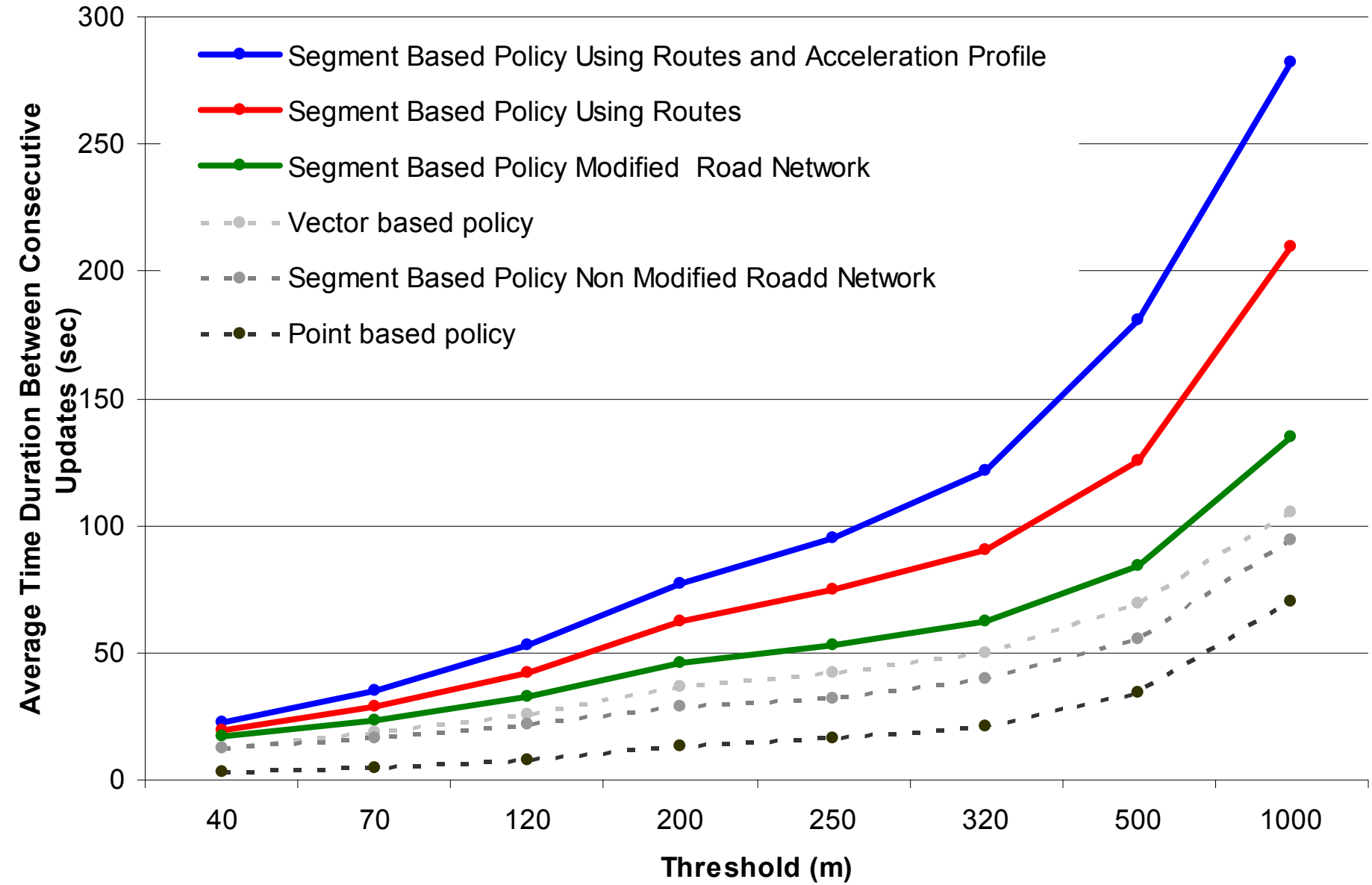
Average Speed



Simplification of Average Speed



Results – Acceleration Profiles



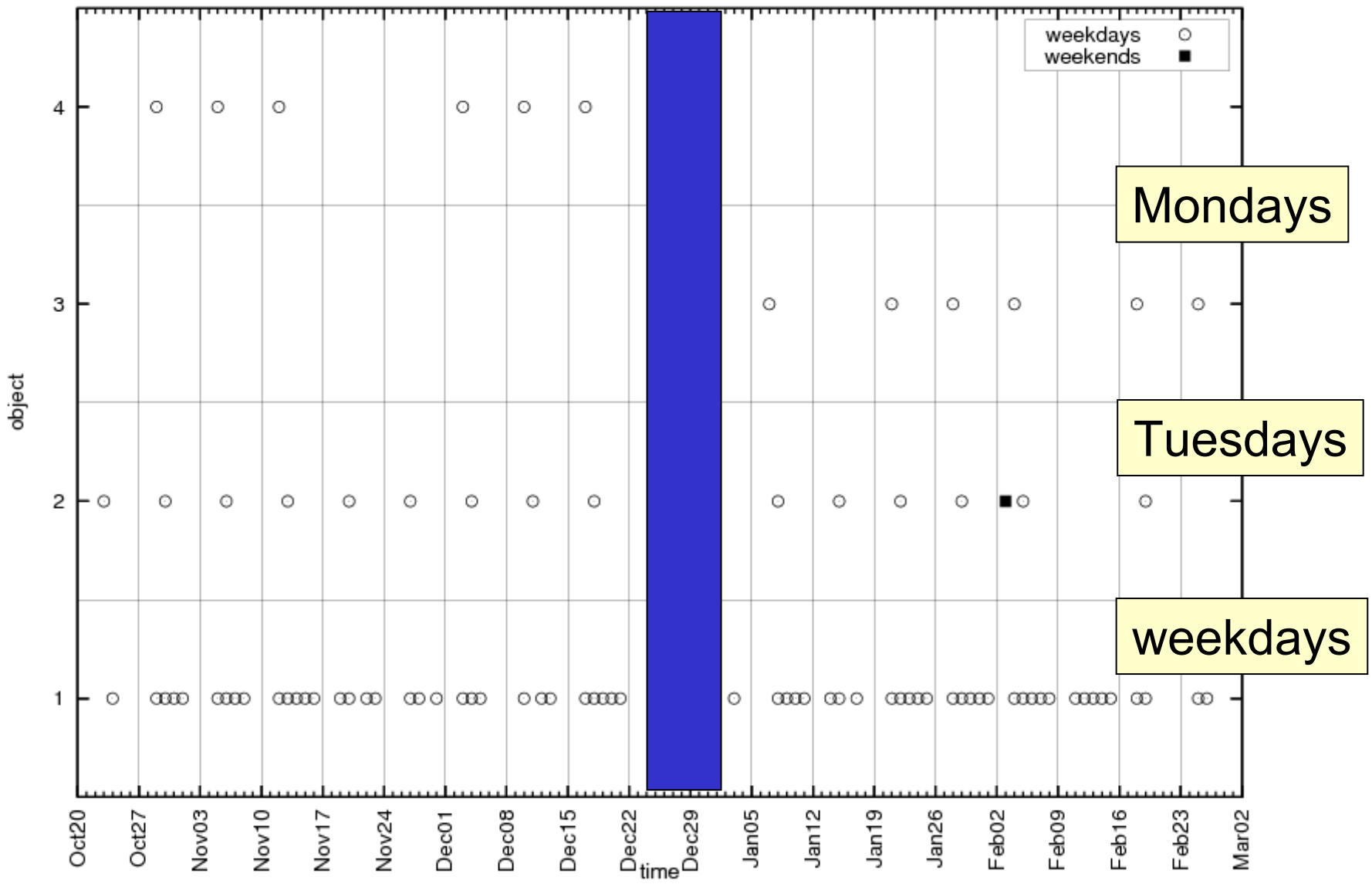
Route and Destination Awareness

- Route and destination *capture*
  - **INPUT**:
    - User IDs
    - Streams of GPS readings (position, time)
  - **OUTPUT**:
    - Routes with associated usage metadata – temporal use patterns.
- Route and destination *prediction*
  - **INPUT**:
    - User ID
    - Location
    - Time
  - **OUTPUT**:
    - Ranked list of possible routes and destinations

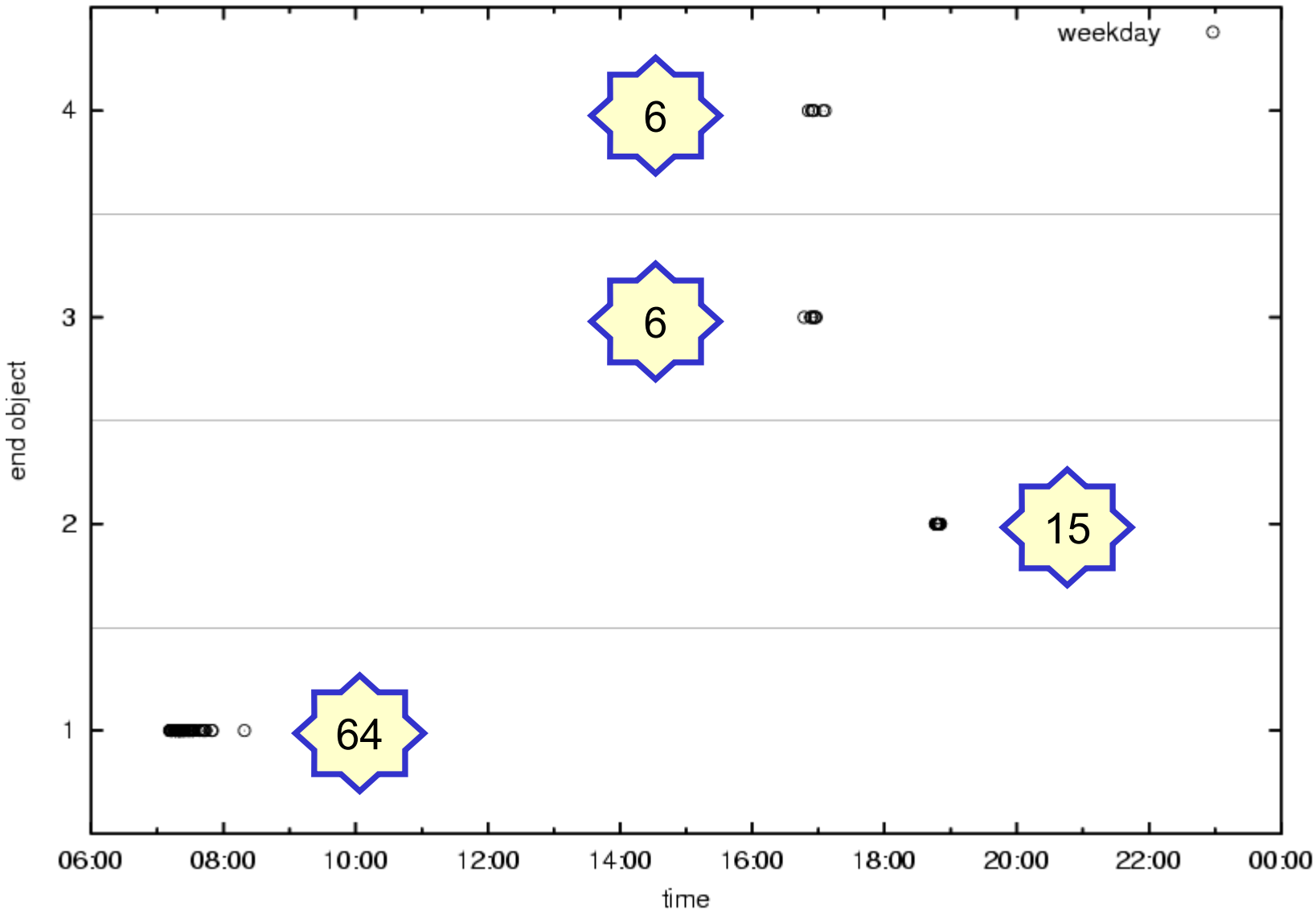
Destination Prediction

- Predict the destination of a user based on the user's past travel and the start location and time.
- Naive approach: Predicts the destination that has been used most frequently from the start location.
  - At location "Home," the route "home-work" may be the most frequently used route.
- Temporal approach: The days of the week and times of day of previous route usages are used for ranking the routes.
  - At location "Home," the route "home-work" may be used only on weekdays in the morning.

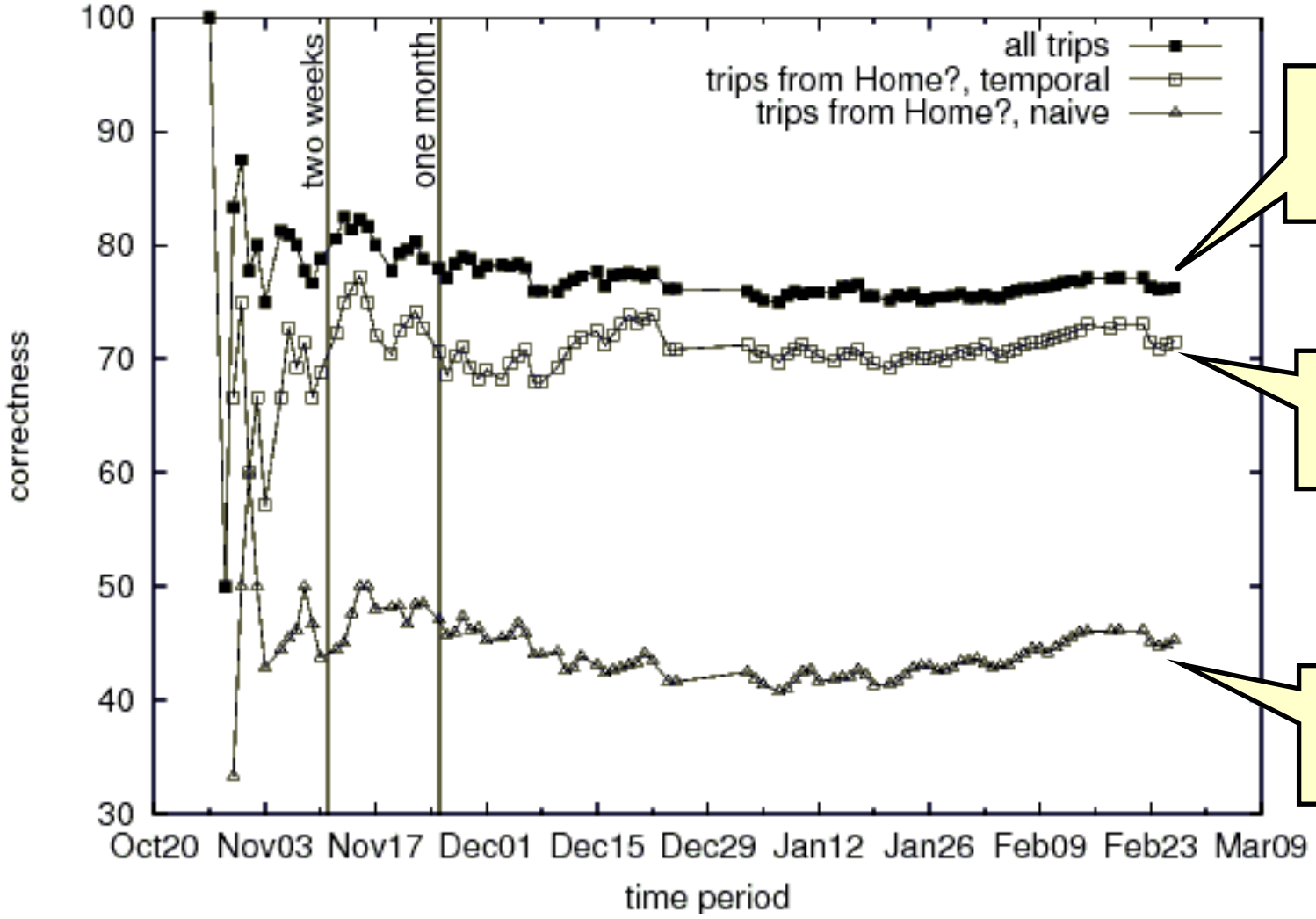
Example: Day of the Week Pattern



Example: Time of the Day Pattern



Temporal vs. Naive Approach



All user trips (temporal)

Temporal (from Home)

Naive (from Home)

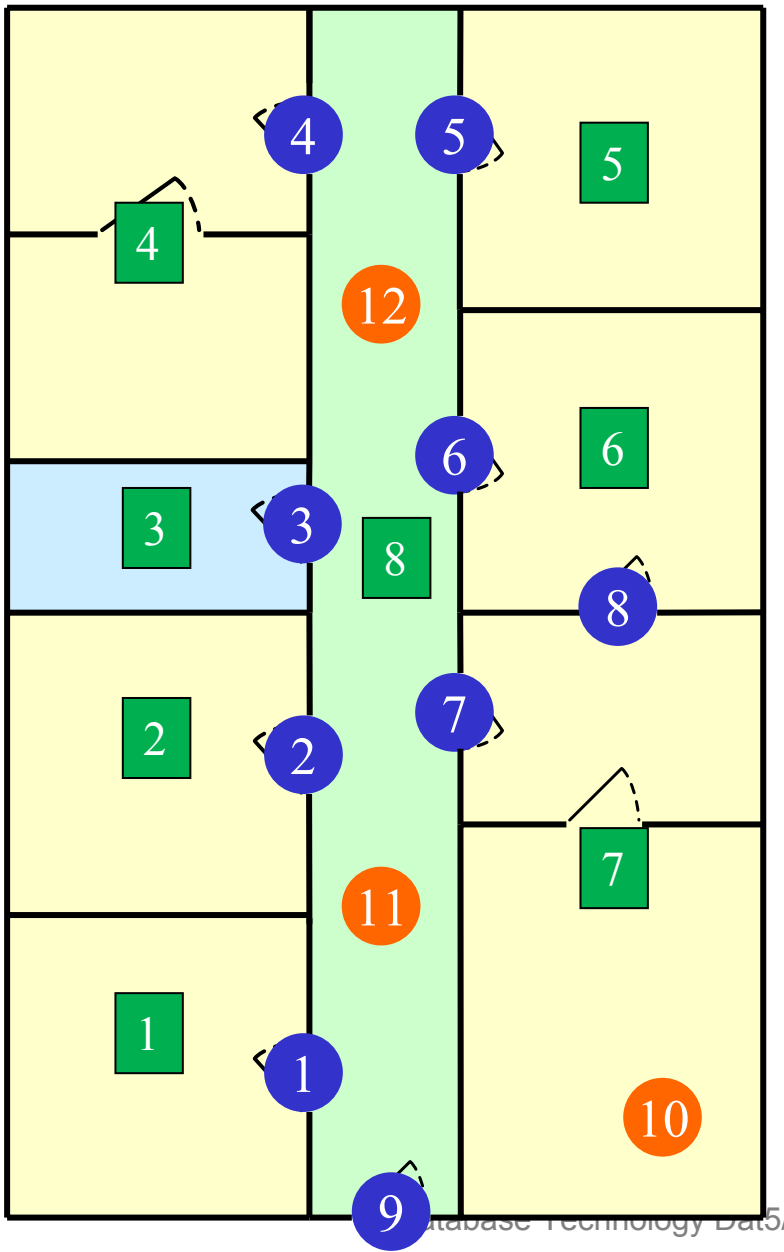
RFID Based Indoor Tracking

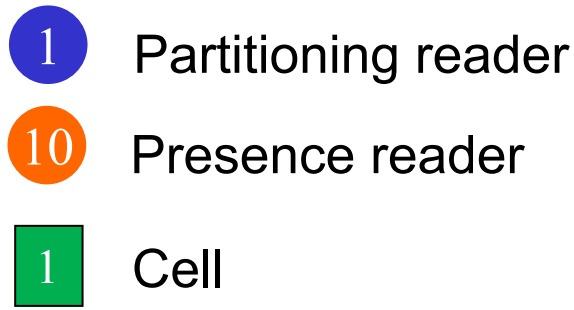
- An ongoing research project in Daisy
- Some facts
  - RFID readers are deployed within an indoor space
  - Moving objects (usually people) are attached with RFID tags.
  - A tag can be detected by RFID reader(s) when it enters the range
  - RFID readers are quite expensive, so that deployment should be carried out carefully
  - But RFID tags are cheap
- Two interesting problems
  - Given a deployment, how to track with a desired the accuracy?
  - If an accuracy is wanted, how to create/improve the deployment?

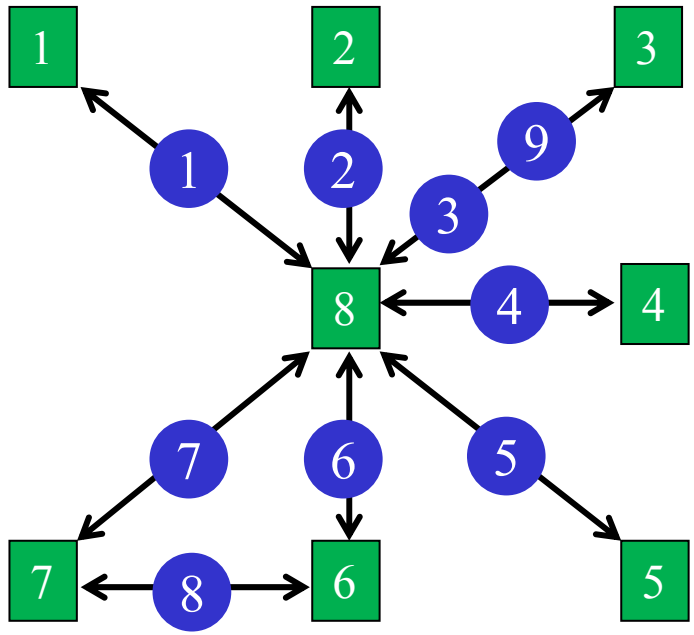
Reader Deployment

- A map of the indoor space is assumed.
- We assume that tag readers are deployed according to application requirements.
  - Partitioning readers: These partition the indoor space into cells
  - Presence readers: These do not contribute to the partitioning.
- A directed graph captures the movements possible between cells, along with the readers that sense the movements.
  - Complete: captures all cells and the movements possible.
  - Minimal: has as few vertices and edges as possible.

Example Reader Deployment

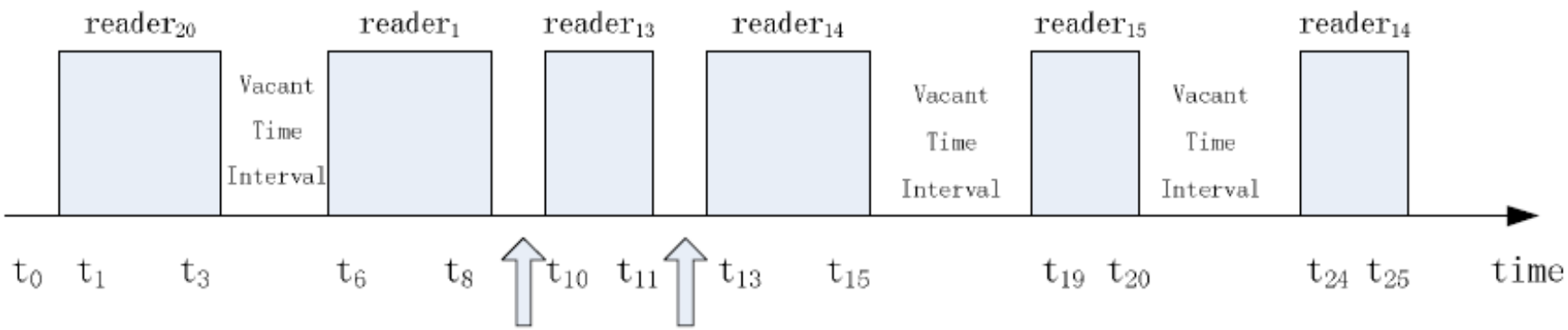






Trajectory Pre-Processing

- It is assumed that the tag readers emit readings of the tags within range at regular intervals.
- The resulting “raw” trajectories are pre-processed in two steps.
  - Step 1: Generation of arrival and departure events for on-line tracking.
  - Step 2: Generation of presence intervals for off-line tracking.



Off-line Trajectory Construction

- Try to fill the gaps between tag readings, by figuring out where the object can be in-between

- What we can make use of?
  - Topology of the indoor space, partly captured in the directed graph model
    - E.g., starting from room 1, there is no pathway to reach room 30 in the same building
  - Speed constraint
    - The maximum speed of an object within a building: v_m
    - Distance from room 1 to room 10: d
    - An object is seen in room 1 at time t_1
    - Then, it cannot be in room 10, if the following inequality holds:
      - $d > v_m * (t_1 - t_2)$

On-line Trajectory Construction

- Two overall cases occur depending on what was observed at the last sampling time t_i
  - Case 1: The object is currently in the range of r_i.
  - Case 2: The object has left the range of r_i at time t_i.
- For either case, the topology information and speed constraint are still useful.
- We also can attempt to use past movement information of the object or other objects, to better predict the location of the object.

Use of Past Movement Information

- Idea: An object is likely to move in the future like it moved in the past.
- We capture past movement by each object as association rules of the form

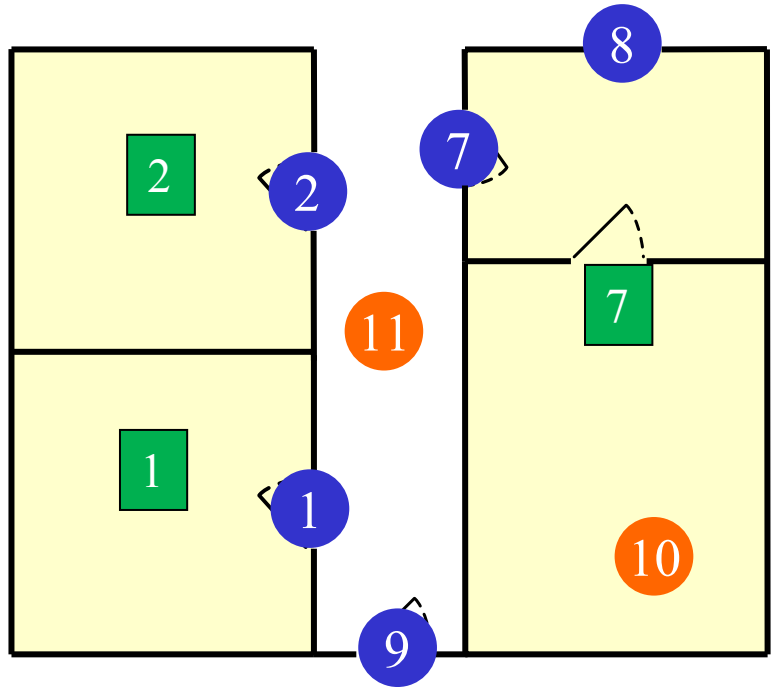
$$r_i \text{ then } r_j \ldots \text{ then } r_k \rightarrow r_i \text{ then } r_j \ldots \text{ then } r_k$$

- We may introduce delays between the readers.
- We may use cells instead of readers.
- Next, we match the movement of an object against the rules, thus forming a prediction of the future movement of the object.
  - We use rules for the object if they exist.
  - Otherwise, we use rules for all objects.

Past Movement Example

Assume there are 10,000 items and these rules exist:

Count	Rules	Supp.	Conf.
700	$r_1, r_7 \rightarrow r_8$	7%	70%
150	$r_1, r_7 \rightarrow r_{10}$	1.5%	15%
100	$r_1, r_7 \rightarrow r_2$	1%	10%
50	$r_1, r_7 \rightarrow r_9$	0.5%	5%



Having observed an object at r_1 and then r_7, we may predict that it will move to r_8.

Challenges

- Automated reader deployment.
- Tracking accuracy versus cost.
- Trajectory representation, e.g., conservative versus progressive.
- Near and long-term prediction.
- Indexing of indoor trajectories.
- Seamlessly tracking regardless of location.
  - Different positioning technologies for indoor and outdoor.
  - Overlap of positioning technologies.

Tracking Summary

- Tracking and trajectories
- Outdoor tracking
  - Using GPS
  - Accuracy guarantees
  - Spatial networks
- Indoor tracking
  - An ongoing research project

Readings

- Christian Hage, Christian S. Jensen, Torben Bach, Laurynas Speicys, Igor Timko: Integrated Data Management Real World. VLDB 2003: 1019-1030
- Lionel M. Ni, Yunhao Liu, Yiu Cho Lau, Abhishek P. Patil: LANDMARC: Indoor Location Sensing Using Active RFID. PerCom 2003: 407-415
- Christian S. Jensen: Database Aspects of Location-Based Services. Location-Based Services 2004: 115-148
- Alminas Civilis, Christian S. Jensen, Stardas Pakalnis: Techniques for Efficient Road-Network-Based Tracking of Moving Objects. IEEE Trans. Knowl. Data Eng. 17(5): 698-712 (2005)
- Christian S. Jensen: Geo-Enabled, Mobile Services-A Tale of Routes, Detours, and Dead Ends. DASFAA 2006: 6-19
- Zhiyong Huang, Christian S. Jensen, Hua Lu, Beng Chin Ooi: Collaborative Spatial Data Sharing Among Mobile Lightweight Devices. SSTD 2007: 366-384
- Christian S. Jensen: When the Internet Hits the Road. BTW 2007: 2-16

Acknowledgement

- Part of the slides and contents are created or shared by my colleagues
  - Christian Jensen
  - Bent Thomsen
  - Rico Wind
  - Bin Yang