When the Internet Hits the Road

Christian S. Jensen

www.cs.aau.dk/~csj
Roadmap

Motivation
- Web 2.0
- The Mobile Internet

Vision
- "To be for mobile services what YouTube is for Videos."
- The Streamspin Project

Results
- Tracking
- Route Awareness
Web 2.0

• Web 2.0 captures the sense that there is something qualitatively different about today's web.
• Leveraging the collective intelligence of communities
• New ways of interacting

• Sharing of user-generated content
• Text
  ■ Wiki’s, e.g., Wikipedia
  ■ Blogs
• Photos
  ■ E.g., Flickr, Plazes, 23
• Video
  ■ E.g., YouTube
Web 2.0

- Community concepts abound…
- Feedback and rating schemes
  - E.g., ratings of sellers and buyers at auctions, ratings of content
- Social tagging, tag clouds, folksonomies
- Wiki’s
  - Collaborative authoring

- RSS feeds
- Active web sites, Ajax

- Fueled by Google-like business models
Flickr

• From the Flickr entry on Wikipedia

• “In addition to being a popular Web site for users to share personal photographs, the service is widely used by bloggers as a photo repository. Its popularity has been fueled by its innovative online community tools that allow photos to be tagged and browsed by folksonomic means.”


• “On December 29th, 2006 the upload limits on free accounts were increased to 100Mb a month (from 20Mb)”
YouTube

• From the YouTube entry on Wikipedia

• “The domain name "YouTube.com" was activated on February 15, 2005…”

• “According to a July 16, 2006 survey, 100 million clips are viewed daily on YouTube, with an additional 65,000 new videos uploaded per 24 hours.”

• “Currently staffed by 67 employees, the company was named TIME magazine's "Invention of the Year" for 2006. In October 2006, Google Inc. announced that it had reached a deal to acquire the company for US$1.65 billion in Google's stock.”
Roadmap

Motivation

Web 2.0

Vision

The Mobile Internet

The Streamspin Project

Results

Tracking

Route Awareness
The Internet Is Going Mobile

• We are at a unique point in history.
  ▪ The Internet is just about ready to go mobile.
  ▪ The mobile Internet has the potential for having more users than the conventional Internet.

• A mobile Internet infrastructure is emerging.
  ▪ Mobile devices, e.g., mobile phones, PDAs, laptops, cameras, MP3 players, navigation systems, etc.
  ▪ Communication networks, e.g., GPRS, EDGE, 3G, HSDPA, Wimax
  ▪ Users with access

• Technologies are becoming practical/available that enable the accurate geo-positioning of all objects we care about.
  ▪ The emerging network-assisted GPS reduces power consumption.
  ▪ Galileo is underway.
Service Types

• Traffic and traffic-management related services
  ■ Emergency vehicle dispatching
  ■ 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
  ■ Tracking of hazardous cargo
  ■ Warnings about accidents, slow-moving traffic ahead, icy or slippery road conditions
  ■ Monitoring of traffic offenders
  ■ Monitoring of tourists traveling in dangerous environments, reacting to emergencies
Service Types, cont.

- Games and "-tainment" (edu-, info-, enter-)
  - Treasure hunting (geocaching)
  - Paintball (Botfighters)
  - Catch the monster (Raygun)
  - Escape the monster
  - 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

• Context awareness will be important!

• 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
Roadmap

Motivation
- Web 2.0
- The Mobile Internet

Vision
- The Streamspin Project

Results
- Tracking
- Route Awareness
Streamspin in a Nutshell

• Vision:
  To create data management technology that enables sites that are for mobile services what Flickr is for photos and YouTube is for video.

• Challenges
  - Enable easy mobile service creation
  - Enable service sharing with support for community concepts
  - An open, extensible, and scalable service delivery infrastructure

• The streamspin project maintains an evolving platform that aims to serve as a testbed for exploring solutions to these challenges.
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)

• Visual Studio C# add-ins for custom service creation
• Accessible from all web-service enabled languages
Service Sharing

• Public content providers
  ■ Approval
  ■ Meta data: interest profile, location, location range, time to live, age range

• Publication of services in a service directory

• Interest hierarchy
  ■ For tagging of content by public providers
  ■ For specification of profiles by service users

• 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

- Filtering of public content
  - 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.
Example: Tour Services

- Users can create tours.
  - Tours are created using a point-and-click tour builder template.
  - Tours associate content (e.g., photos and text) with locations.

- Users can subscribe to tours.
  - They receive content when they get within a specified range of the locations associated with the content.

- Example tours
  - Walking and driving tours with directions
  - Tours that involve public transportation can tell their users when to get off busses in real time
Example: Gasoline Service

• Services benefit from geo-context awareness.

• No current location awareness
  ■ The user is notified when a gas station near their normal location offers gas at least 10% below the “list” price.

• Current-location awareness
  ■ The user is notified when within 3 km of cheap gas.
  ■ Close-by gas

• Location- and route-awareness
  ■ The user is notified when cheap gas is close to the route ahead.
  ■ Cheaper gas and smaller detours.
Roadmap

Motivation
- Web 2.0
- The Mobile Internet

Vision
- The Streamspin Project

Results
- Tracking
- Route Awareness
Problem Setting

Aim: To track moving objects with accuracy guarantees

Objective: To reduce cost of communication between client and server and server-side update, client-side costs
Tracking Approach

Client

- predict position
- get GPS
- store settings

[old connection]
[within threshold $th$]
[new connection]
[continue]
[finish]

Server

- receive update
- update DB
- find road
- send new prediction

[segment based tracking]
[not segment based tracking]
Point-Based Tracking
Vector-Based Tracking
Segment-Based Tracking
Two Implementations

• Centralized, Oracle-based implementation
  ■ Ideal for testing implementations of the algorithmic aspects of the techniques.
  ■ Well suited for simulation-based experiments with pre-recorded data.

• Real implementation
  ■ Involves a central server, mobile terminals, GPS receivers
  ■ More complex than the centralized implementation
  ■ Enables more detailed cost modeling, e.g., of data transmission cost and server and client side loads
  ■ Offers insight into the specifics, e.g., network delays
  ■ Offers the ultimate proof of concept
Denmark on a winter’s day

Kenneth driving his BMW

Smart Phone w. built-in GPS

Camcorder

Stardas holding it all
See the video at:
http://daisy.aau.dk/projects/trax.php
Data for Experiments

- **GPS Data** – the INFATI data is used for evaluation
  - GPS receivers and computers installed in cars
  - GPS coordinates are registered every second for ~6 weeks
  - The data used has ~100,000 records per car and ~458,000 in total

- **Digital Road Network**
  - Each segment corresponds to the road in-between two crossroads
  - The geometry of a segment is represented as a polyline
Comparison of Techniques

Average Time Duration Between Consecutive Updates (sec)

Threshold (m)

- Vector Policy
- Segment Based Policy
- Point Policy
The graph shows the average time duration between consecutive updates for different network policies as a function of threshold (m).

- **Theoretically Optimal Policy**: Assumes constant speed.
- **DSC Modified Road Network**: Solid green line.
- **TSC Modified Road Network**: Dashed black line.
- **SSC Modified Road Network**: Solid orange line.
- **Vector Based Policy**: Dashed blue line.
- **Unmodified Road Network**: Solid cyan line.

The graph indicates that the theoretically optimal policy has the shortest average time duration at all thresholds, followed by the DSC modified network, TSC modified network, SSC modified network, vector based policy, and finally the unmodified network which has the longest average time duration.

*Note: The graph assumes constant speed for all policies.*
Use of Routes

• Users follow routes to reach their destinations.
• If we know the current route of a user, we can avoid segment changes altogether.
• As routes are (long) segments, segment-based tracking works.
• Routes may be obtained via a navigation system or a route acquisitioning and provisioning component (next!).
Results – Use of Routes

Average Time Duration Between Consecutive Updates (sec)

- **Segment Based Policy Using Routes**
- **Theoretically "Optimal" Policy**

Threshold (m)
The figure displays part of a user’s route from home to work. Distances are indicated for some points. Used in examples next.
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.
Average Speed
Use of Acceleration Profiles

Example tracking of one car using a 70 m threshold.
Results – Acceleration Profiles

- Segment Based Policy Using Routes and Acceleration Profile
- Segment Based Policy Using Routes
- Segment Based Policy Modified Road Network
- Vector based policy
- Segment Based Policy Non Modified Road Network
- Point based policy
Roadmap

Motivation

Web 2.0

The Mobile Internet

Vision

The Streamspin Project

Results

Tracking

Route Awareness
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
Road Network Data Model

- Base points describe roads.
- Polylines of base points approximate road geometry.
- Real distance values are given for base points.
- Connections describe crossroads.
Route Data Model

- Destination areas are circular regions.
- A subpolyline is a part of a polyline.
- A route element is a "directed" subpolyline.
- A route is a sequence of route elements that make up an uninterrupted polyline.
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

- Tuesdays
- Mondays
- Weekdays
Example: Time of the Day Pattern
Empirical Study

• Real data
  ■ Top10DK map
  ■ GPS logs from the AKTA road pricing project
    ◆ 182 users
    ◆ All usage periods were longer than 80 days
    ◆ Each user made more than 40 trips

• Destination identification
  ■ The first/last positions of the trips were combined into circular destinations.
  ■ The number of end destinations from one start varied from 1 to 86.
Temporal vs. Naive Approach

- All user trips (temporal)
- Temporal (from Home)
- Naive (from Home)
Summary

• Streamspin aims to apply Web 2.0 concepts to mobile services.
  ■ Easy creation and sharing of mobile services, scalable delivery of services.

• Services are expected to be context aware.

• Two geo-context services are being integrated into the streamspin testbed system.
  ■ Efficient continuous tracking of moving objects with accuracy guarantees.
    ✦ The use of real data was essential in guiding the design process.
  ■ Capture and subsequent prediction of a user’s routes and destinations.
    ✦ The use of real data was essential in obtaining realistic solutions.
Summary, cont.

• These two services are expected to be important
  ■ Geo-context (position, destination, route) is important for filtering in push services.
  ■ Can be provided without user interaction.

• Aside: Implications for query processing and indexing
  ■ Data is inaccurate, but accuracy guarantees are possible.
  ■ Updates are distributed non-uniformly across objects and time.
  ■ Spatial networks are important.
Acknowledgments

• The Streamspin project
  ■ Aalborg University (*Kenneth H. Pedersen, Kristian Torp, Rico Wind*)

• The M-Track and TAU-SeN projects, funded by ETRI, Korea.
  ■ Pusan National University (Ki-Joune Li, Si-Wan Kim, Kyoung-Sook Kim, Deuk-Chun Han)
  ■ Chungbuk National University (Keun Ho Ryu, Eung-Jae Lee)
  ■ Aalborg University (Dalia Tiesyte, Nerius Tradisauskas, Simonas Saltenis, Linas Bukauskas, *Alminas Civilis, Stardas Pakalnis*)

• The ContextIT project
  ■ Aalborg University (*Agne Brilingaite*)
Readings


• http://daisy.aau.dk
• http://streamspin.com
• http://www.cs.aau.dk/DBTR/
• http://www.cs.aau.dk/TimeCenter/