FCTUC - UNIVERSIDADE DE COIMBRA

UBIQUITOUS SYSTEMS

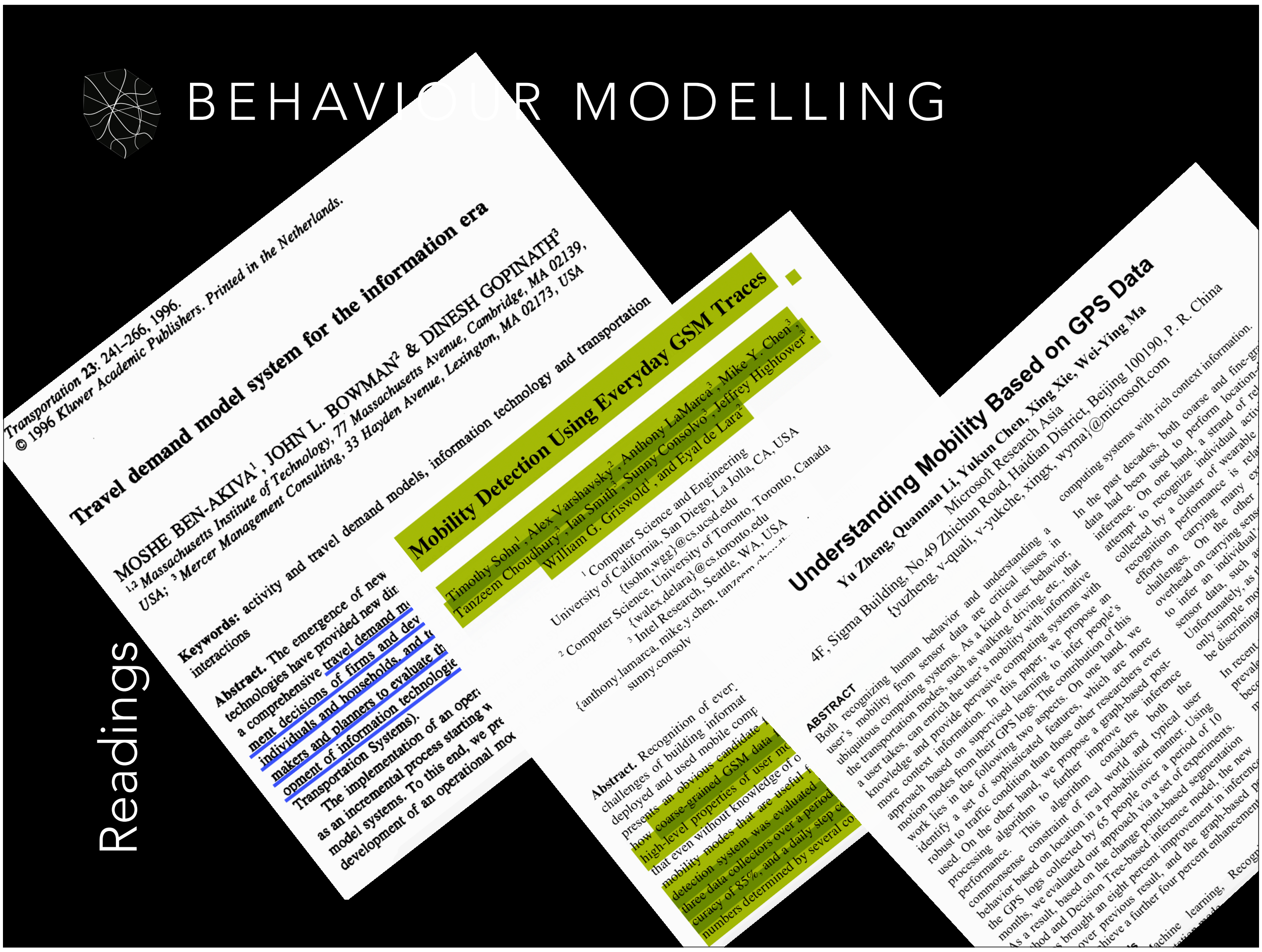
BEHAVIOUR MODELLING
(FOR URBAN COMPUTING)

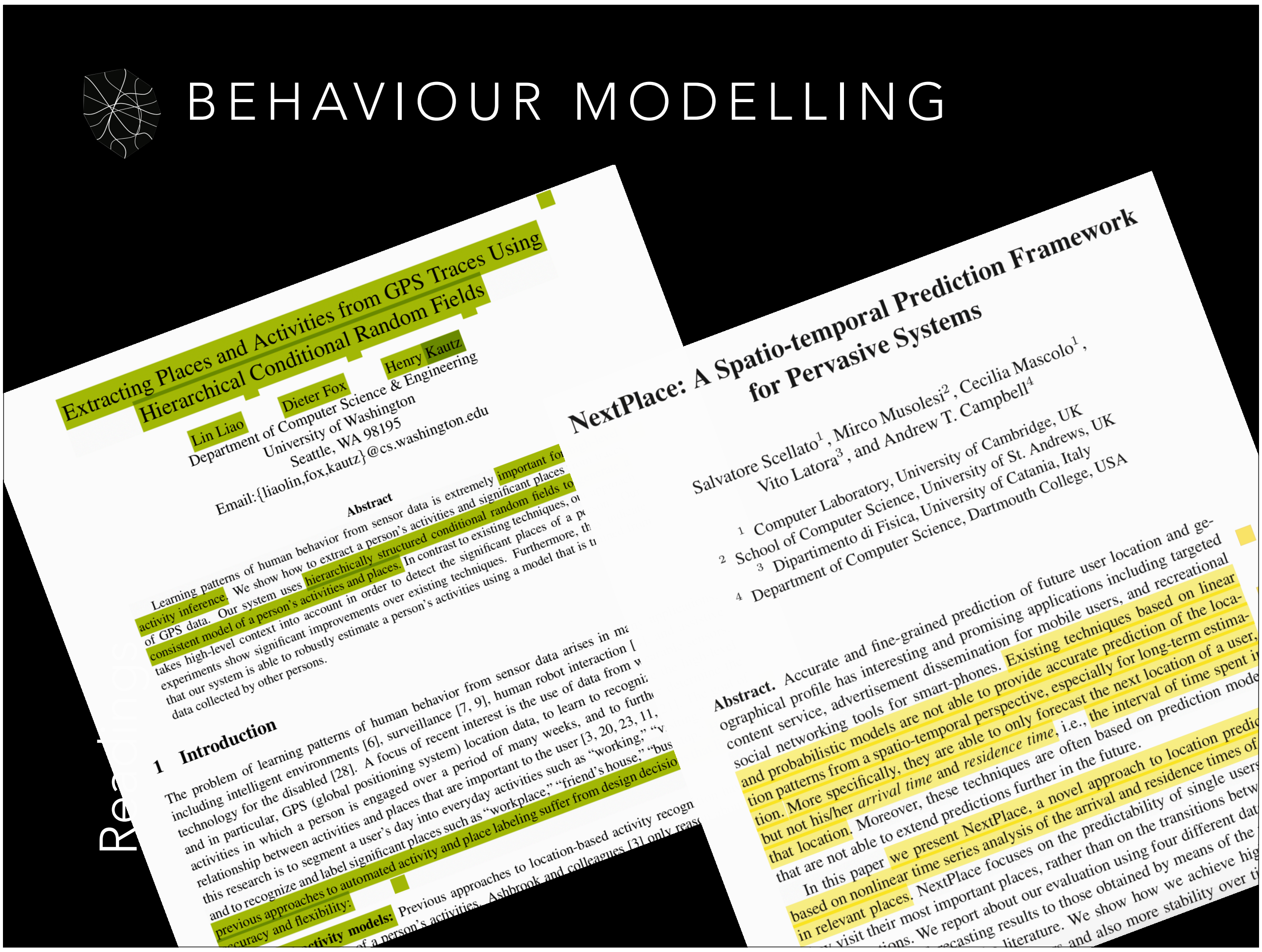
CARLOS BENTO
MEI - FCTUC

GOALS

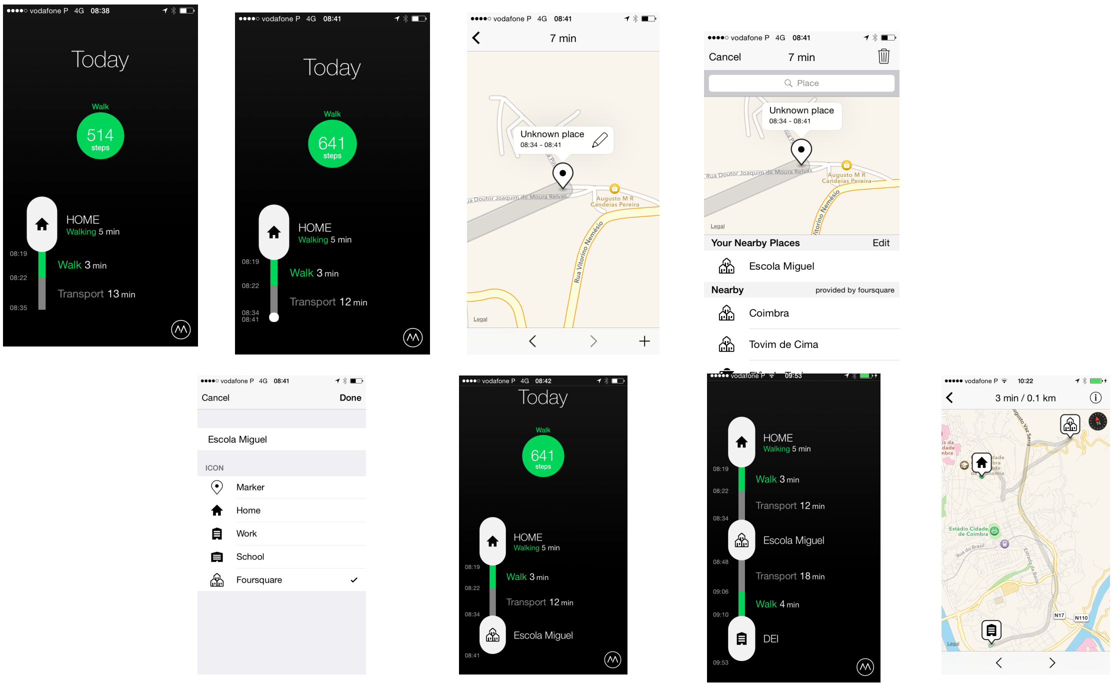
After this session the following topics should become clear to everybody:

- Why modelling user behaviour is important
- The role of mobile data on behaviour modelling
- Methods and algorithms for behaviour modelling
- Privacy issues

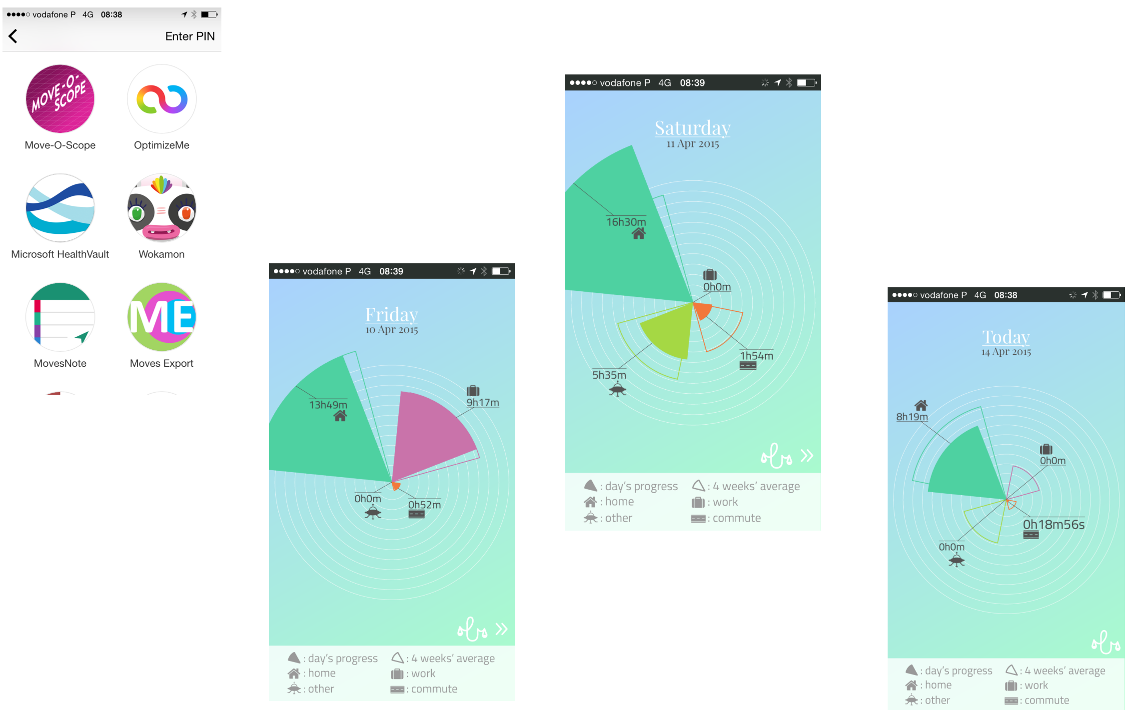

BEHAVIOUR MODELLING
Readings
Transportation 23: 241–266, 1996.
© 1996 Kluwer Academic Publishers. Printed in the Netherlands.
Travel demand model system for the information era
MOSHE BEN-AKIVA, JOHN L. BOWMAN & DINESH GOPINATH
Mobility Detection Using Everyday GSM Traces
Understanding Mobility Based on GPS Data
Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia


BEHAVIOUR MODELLING
Readings
Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields
Lin Liao Dieter Fox Henry Kautz
Department of Computer Science & Engineering
University of Washington
Seattle, WA 98195
NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems
Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, and Andrew T. Campbell

APPLICATIONS AND SYSTEMS

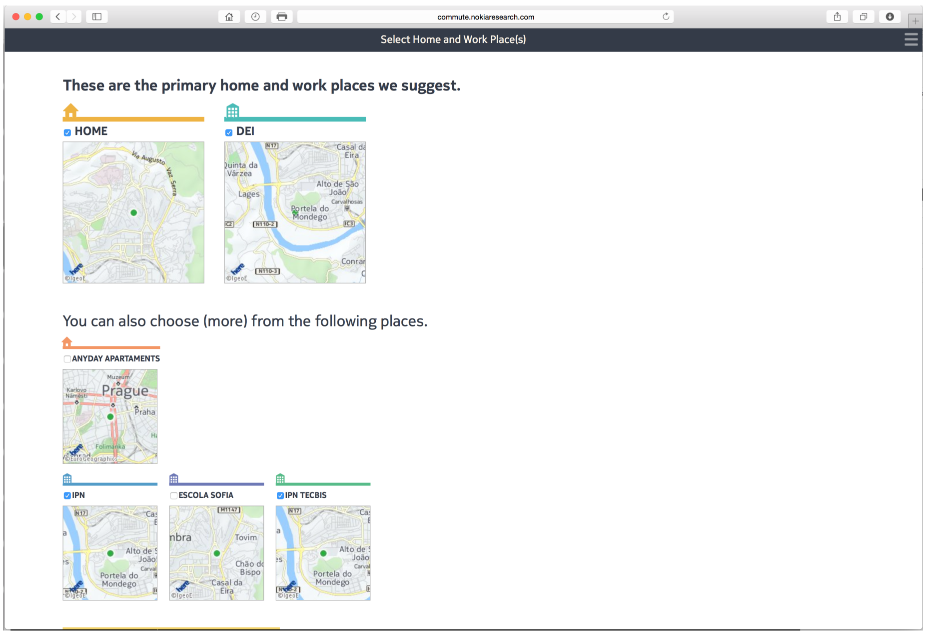


APPLICATIONS AND SYSTEMS

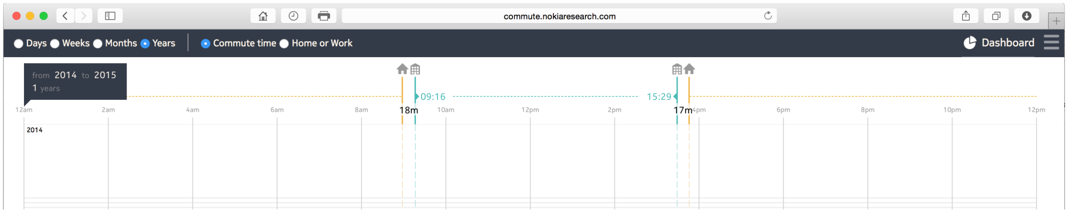


APPLICATIONS AND SYSTEMS

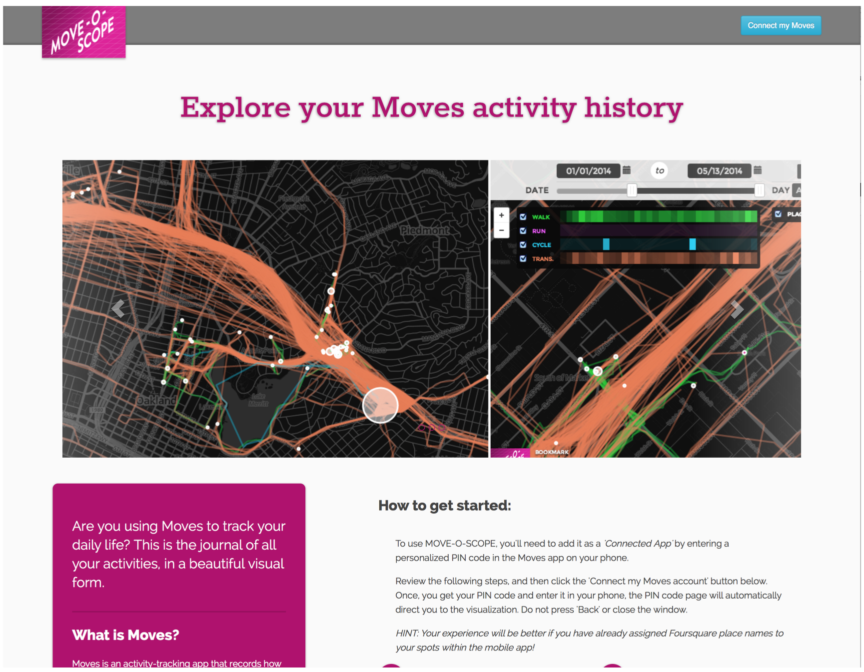
APPLICATIONS AND SYSTEMS



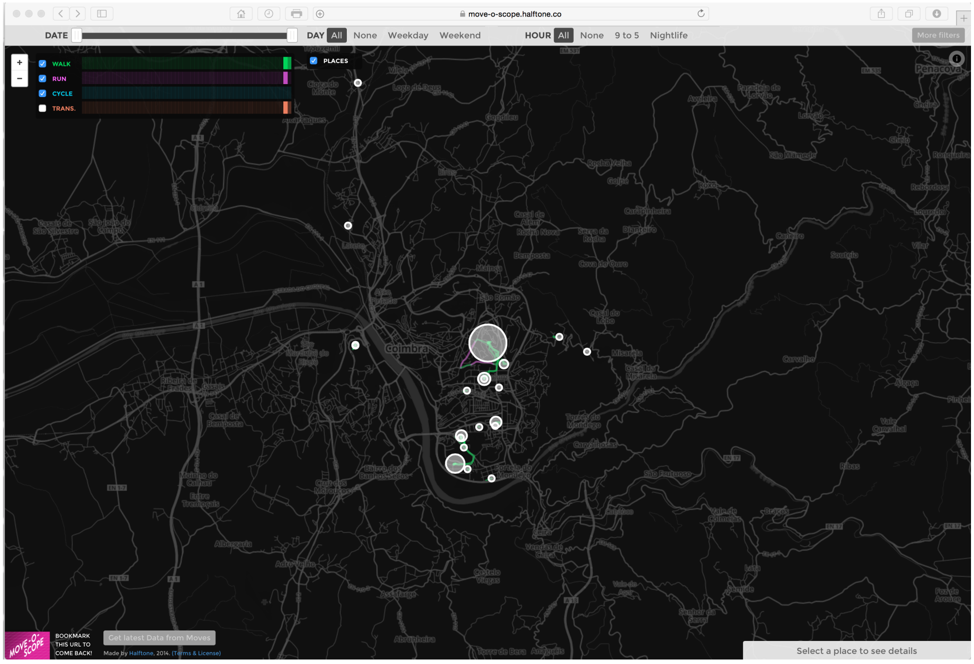
APPLICATIONS AND SYSTEMS



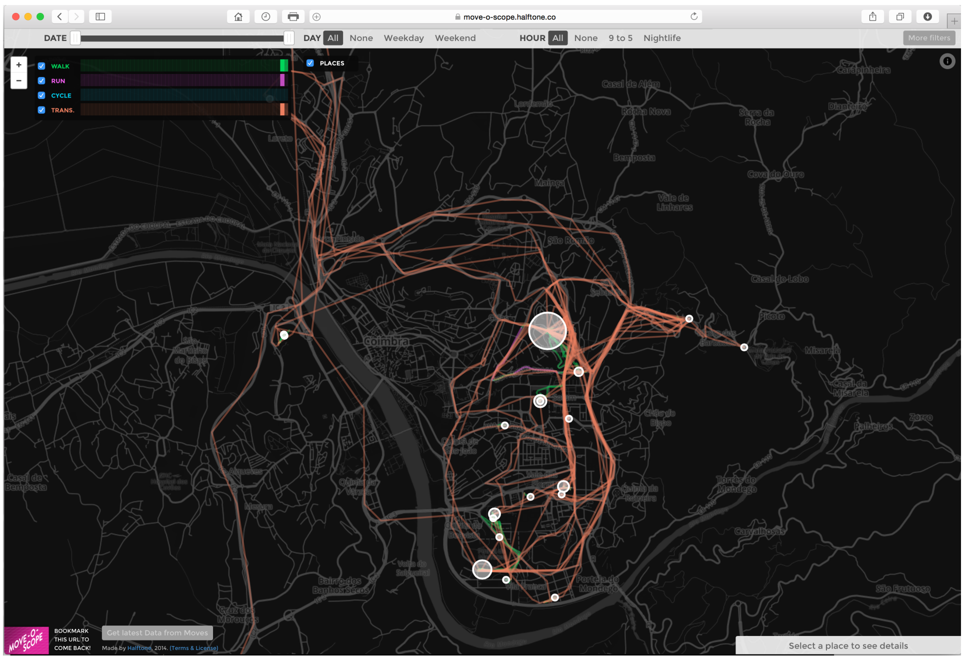
APPLICATIONS AND SYSTEMS



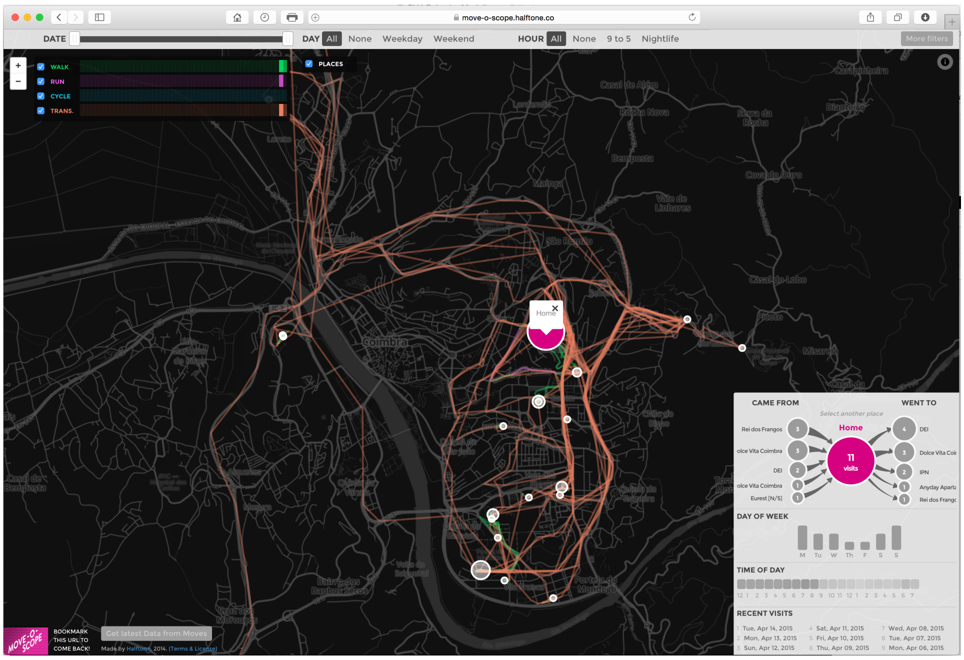
APPLICATIONS AND SYSTEMS



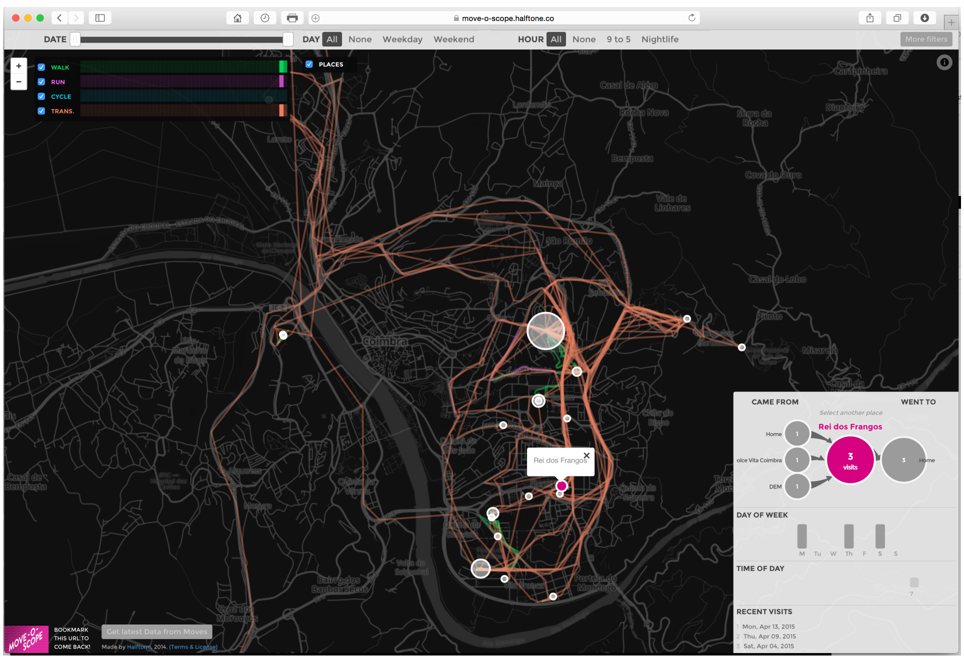
APPLICATIONS AND SYSTEMS



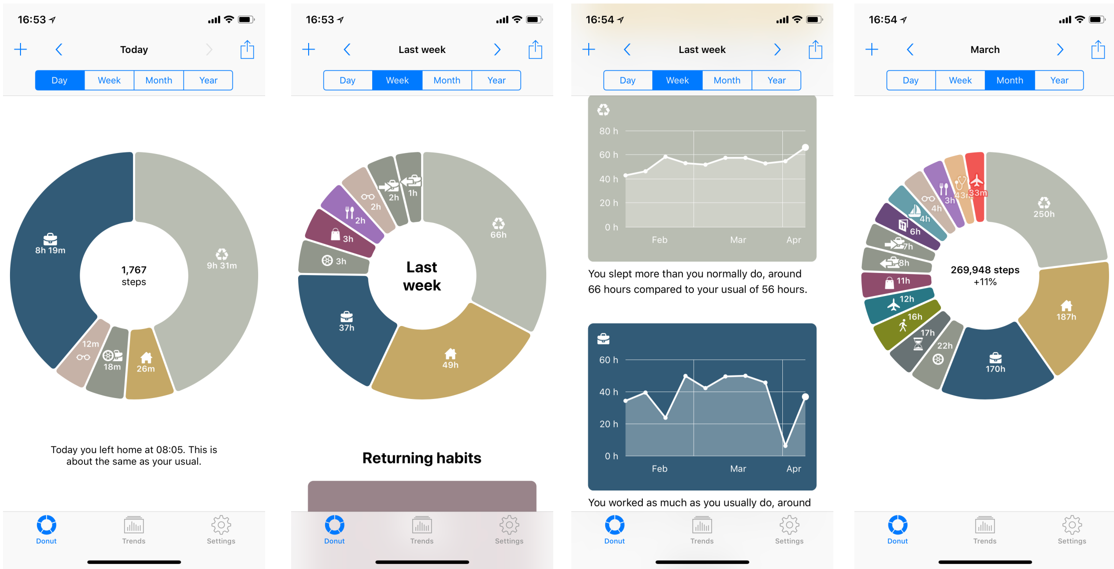
APPLICATIONS AND SYSTEMS



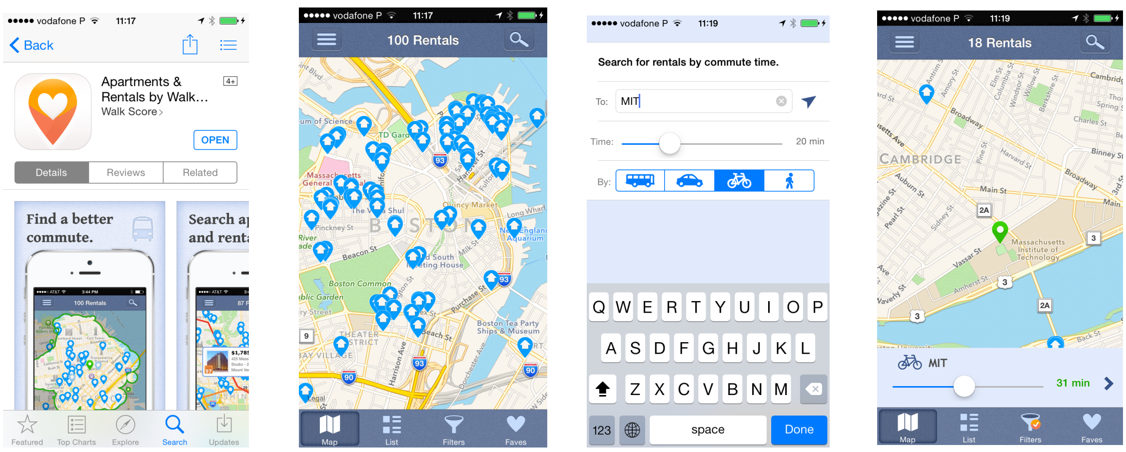
APPLICATIONS AND SYSTEMS



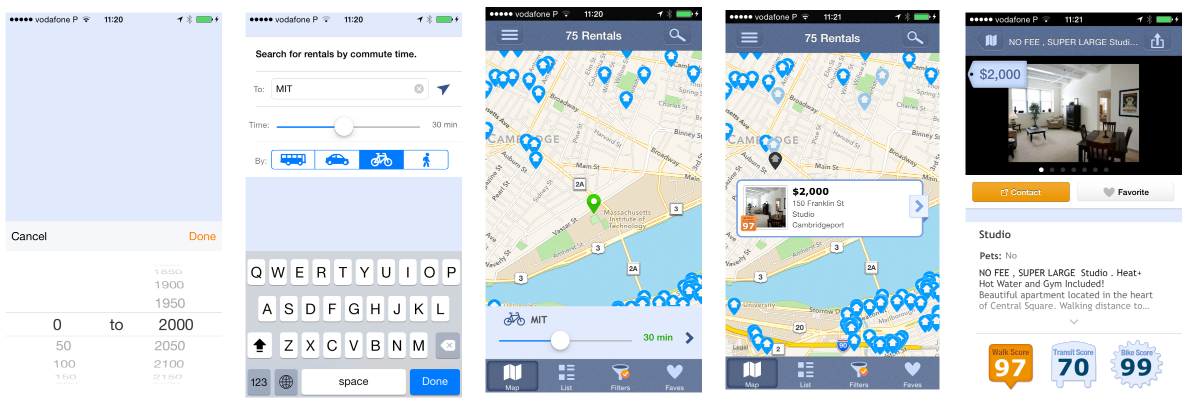
APPLICATIONS AND SYSTEMS



APPLICATIONS AND SYSTEMS



APPLICATIONS AND SYSTEMS



WHY TRACKING BEHAVIOURAL DATA?
(BEHAVIOUR MODELLING)

apps that use this kind of data

- understand my mobility pattern
  ==> IMPROVE MY MOBILITY

- understand to which point a city is cyclable,
  which is the average travel time (commuting) of citizens,
  which are the places they go in the city
  ==> IMPROVING URBAN LIFE

- understand which are the main origin places of the collaborators of an enterprise
  ==> PROVIDE MOBILITY PLANS FOR THE COLLABORATORS,
  ==> PROMOTE CAR SHARING,
  ==> REDUCE CO2 FOOTPRINT,
  ==> REDUCE TRANSPORT COSTS

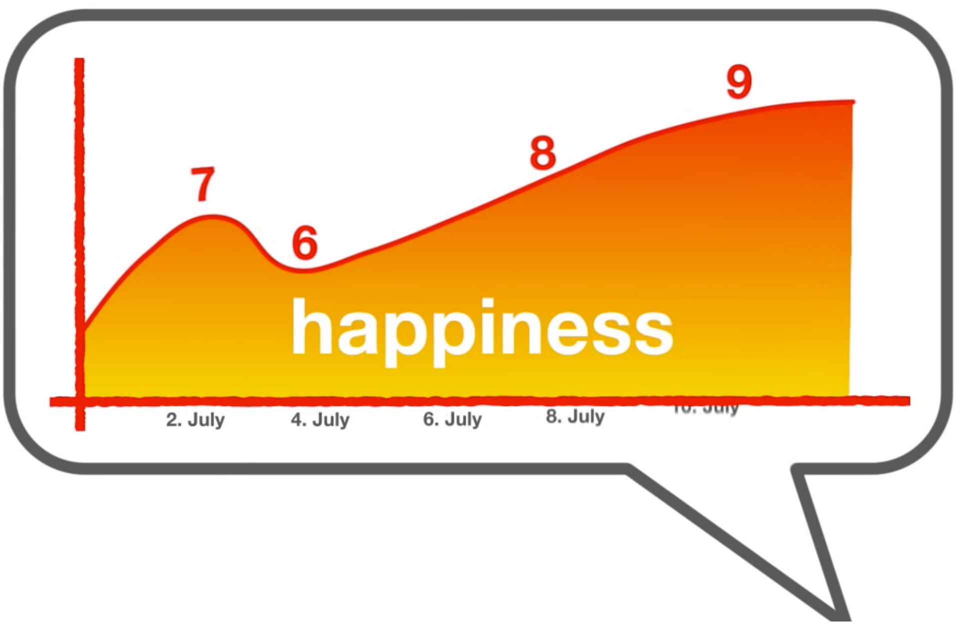
- study demand/supply in transportation
  ==> IMPROVE THE TRANSPORTS NETWORK

- understand the willingness of people to adopt soft modes of transport
  ==> IMPROVE THE CONDITIONS IN THE CITY FOR SOFT MOBILITY

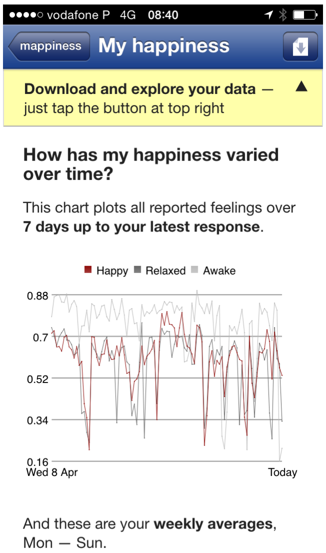
- understand where is the best place to live
  ==> SPOT THE PLACES THAT PROVID CYCLABILITY,
  EASY ACCESS TO MAINS PLACES IN THE CITY

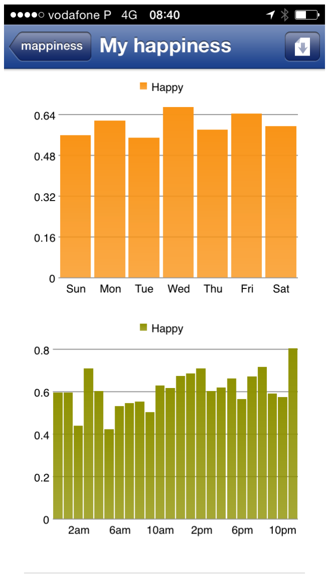
APPLICATIONS AND SYSTEMS

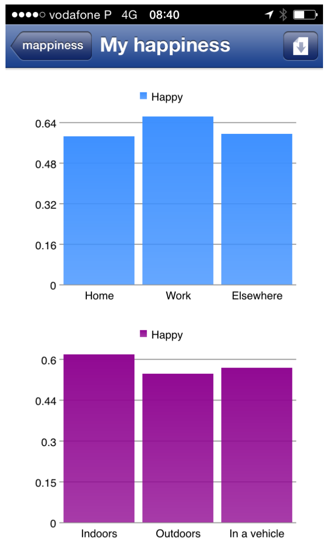
APPLICATIONS AND SYSTEMS

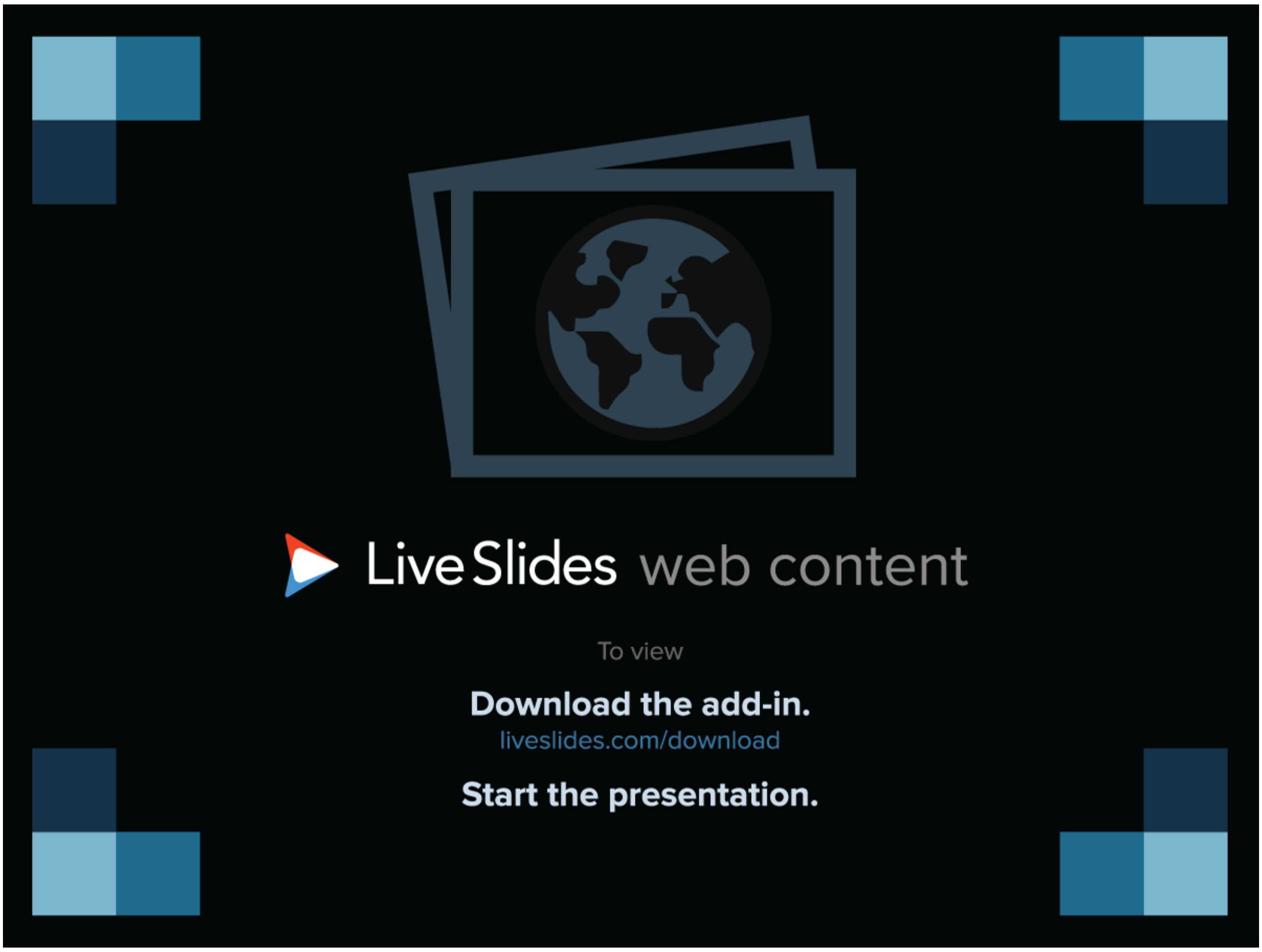


APPLICATIONS AND SYSTEMS





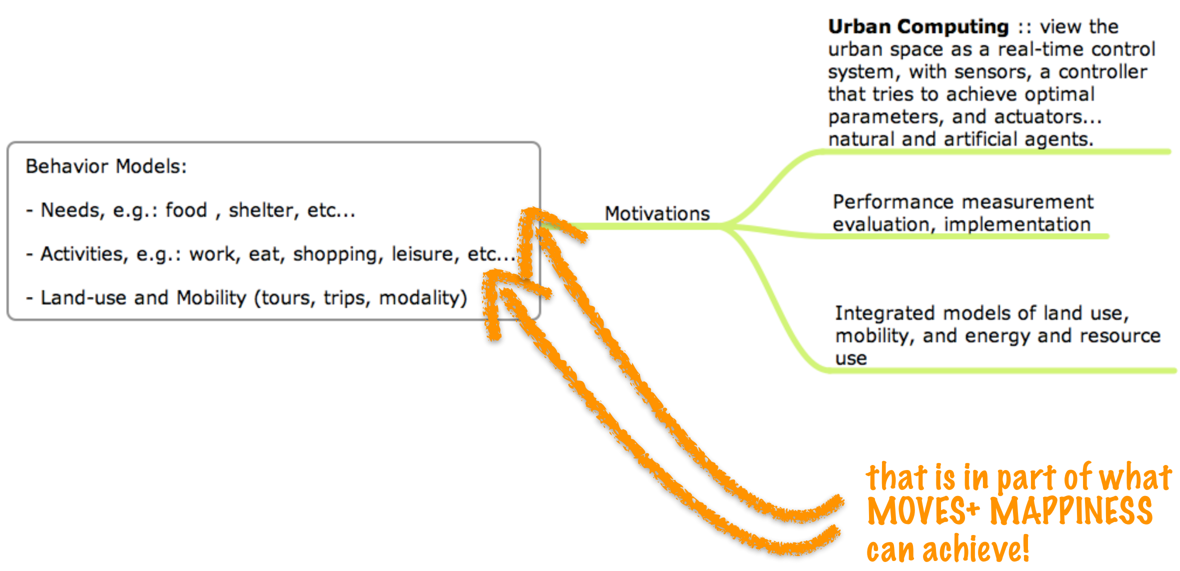




WHY TRACKING HAPPINESS?
(BEHAVIOUR MODELLING)

- many people consider that nowadays the MOST IMPORTANT DECISION on people's life is WHERE TO LIVE (and OKAY with WHOM)
  ==> SELECT BEST PLACE,
  WHERE YOU HAVE MORE CHANCES TO BE HAPPY
- understand which are the best enterprises,
  the best weather conditions,
  the best days of the week, the best hours
- (where, when, in which conditions) people are more happy)
  ==> SELECT THE BEST ENTERPRISES
  ==> ENTERPRISES PUBLICIZING BASED ON EVIDENCIES
  ==> ORGANIZE AGENDA IN THE MOST PSYCHOLOGICALLY FAVOURABLE WAY
  (e.g. I ask, if possible, not to have classes on Monday ;-))

BEHAVIOUR MODELS



BEHAVIOUR MODELS

Source: 1. Ben-Akiva, M., Bowman, J. and Gopinath, D. (1996), 'Travel Demand Model System for The Information Era', Transportation, vol. 23, pp. 241-266.
2. Bowman, J. L. and M. E. Ben-Akiva (2001) Activity-based disaggregate travel demand model system with activity schedules, Transportation Research Part A, 35(2001), pages 1-28.

Activity Schedule = chain of activities + travel

Long Term Choices
e.g.: residential/workspace locations,
auto ownership

Activity Pattern
number/purpose/sequence of independent tours

Primary Tours
timing, destination, and mode

Secondary Tours
timing, destination, and mode

that is in part of what MOVES can achieve!

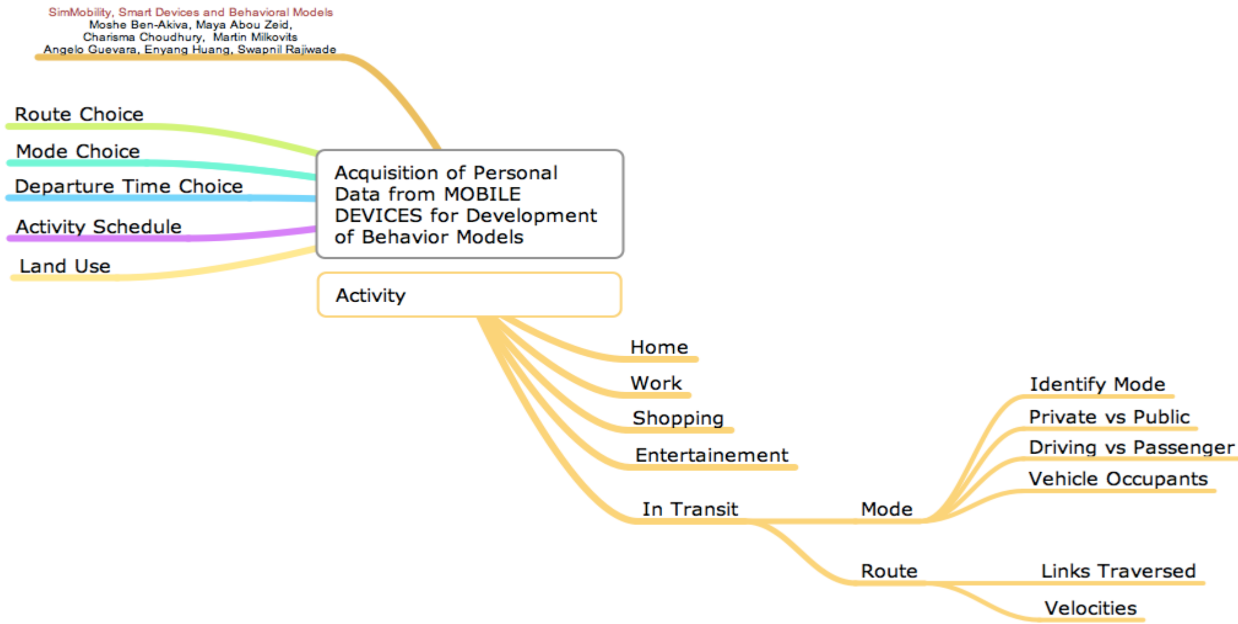
DATA ACQUISITION (BEHAVIOUR MODELS)

SimMobility, Smart Devices and Behavioral Models
Moshe Ben-Akiva, Maya Abou Zeid,
Charisma Choudhury, Martin Milkovits
Angelo Guevara, Enyang Huang, Swapnil Rajiwade

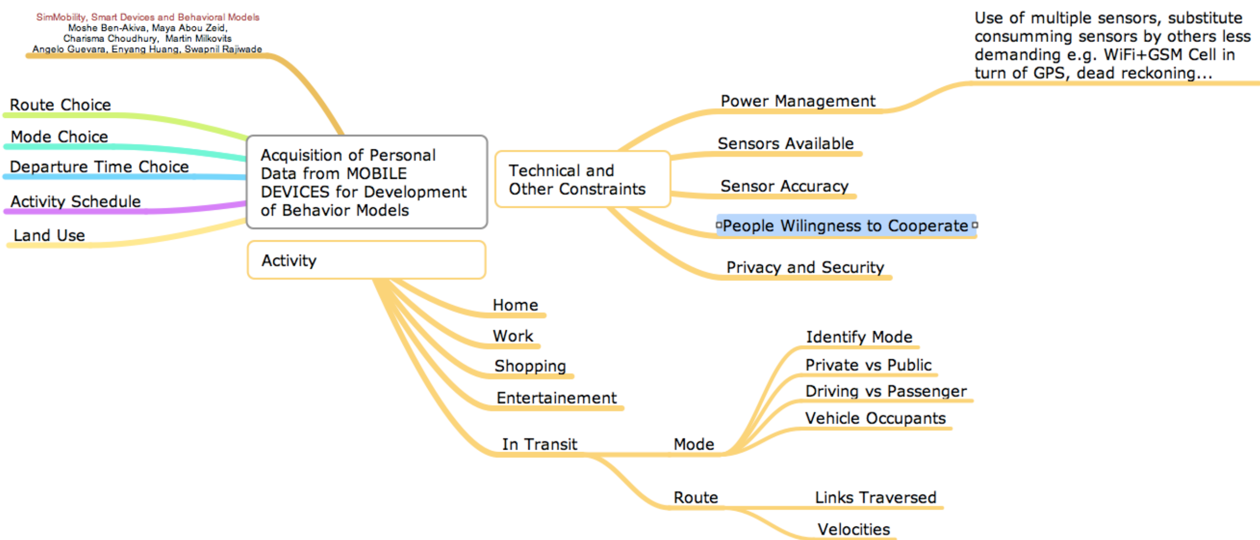
Acquisition of Personal Data from MOBILE DEVICES for Development of Behavior Models

- Route Choice
- Mode Choice
- Departure Time Choice
- Activity Schedule
- Land Use

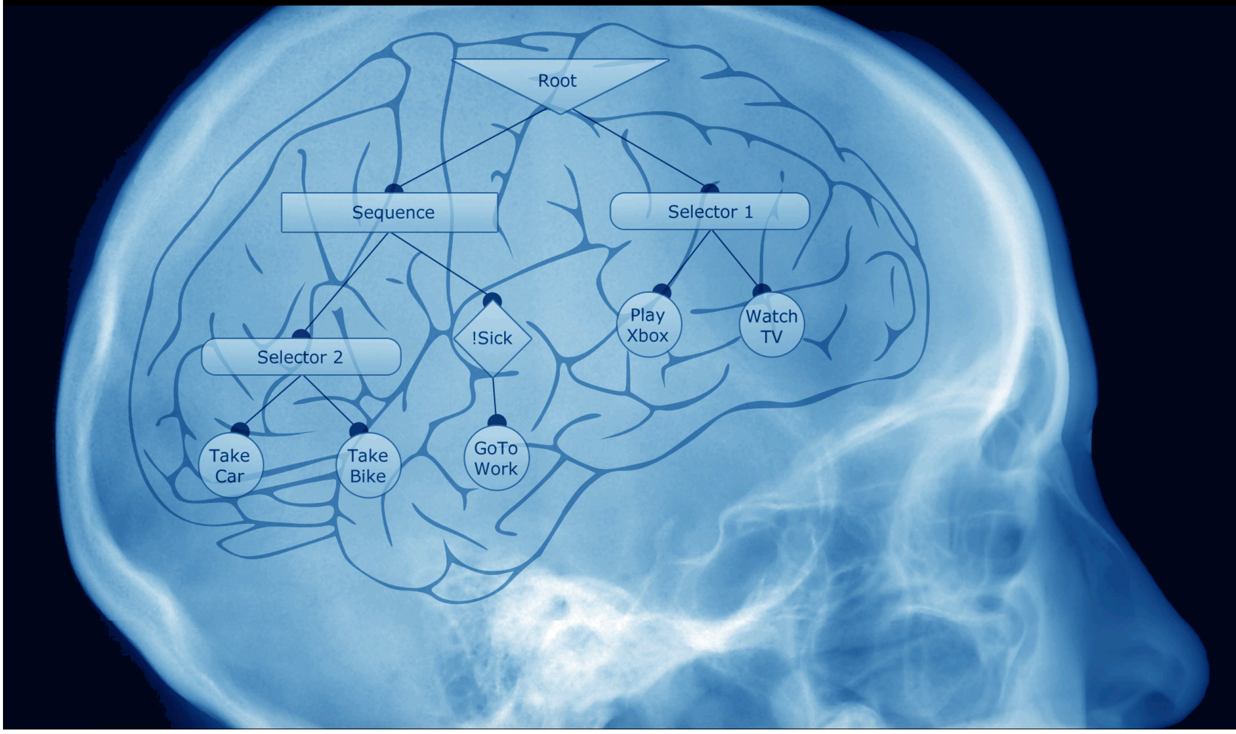
DATA ACQUISITION
(BEHAVIOUR MODELS)



DATA ACQUISITION
(BEHAVIOUR MODELS)



BEHAVIOUR MODELLING



Projects :: funf

friends and family study
Open Sensing Framework
MIT Media Laboratory

Projects :: funf

friends and family study
Open Sensing Framework
MIT Media Laboratory

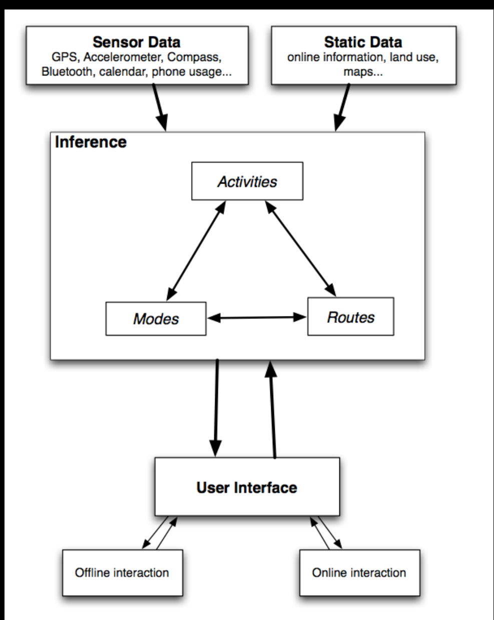
Active Probes

- Bluetooth Scan (sometimes with RSSI)
- WiFi scan (Access points in proximity)
- Location
  - GPS
  - WiFi-based location service (e.g. skyhook)
- 3-axis Accelerometer
- Compass
- Light (proximity) sensor
- Temperature
- Microphone (for ambient sound energy)

Phone DB Probes

- Call Log
- SMS Log
- Contacts
- Phone State:
  - Running processes
  - Version, Model, network IDs..
- ...
  - Web history?
  - Email?

Projects :: SAL (Smartphone Activity Logger)



senseMyCity (FEUP)

Shake UP your city News : SenseMyCity Application

Analysis of fuel consumption and stress levels

SenseMyCity is the name of the application developed for smart phones. Through the use of sensors it is possible to register the everyday life of users for further analysis. Information such as fuel consumption per journey, possibility of car sharing or levels of stress are some of the data which can be studied. Ana Aguiar, researcher and professor at Faculdade de Engenharia da Universidade do Porto, is the responsible for this application.

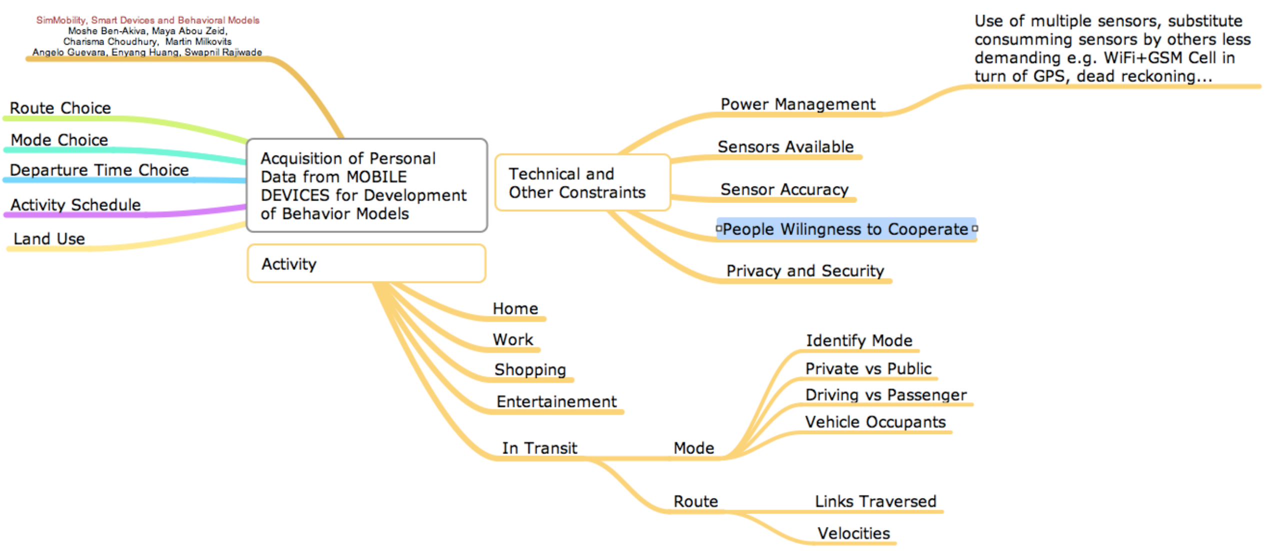
Designed to work in conjunction with other research projects and not separately, the application can be used in research of various areas such as engineering and psychology. Allows users to record, consciously and voluntarily, their daily routine.

The process is simple. The user records his routine through sensors embedded in his smart phone and then views it on a webpage created for that purpose. The collection and analysis of such data (algorithms) can lead to some conclusions such as identifying areas with slower traffic, places or situations that increase stress levels of drivers, among others. This same analysis can be used to optimize routes and consumption. The use of SenseMyCity allows, for example, the identification of people with similar mobility patterns (boosting car sharing and carpooling) and map the city suggesting bike routes with little slope and flat floor.

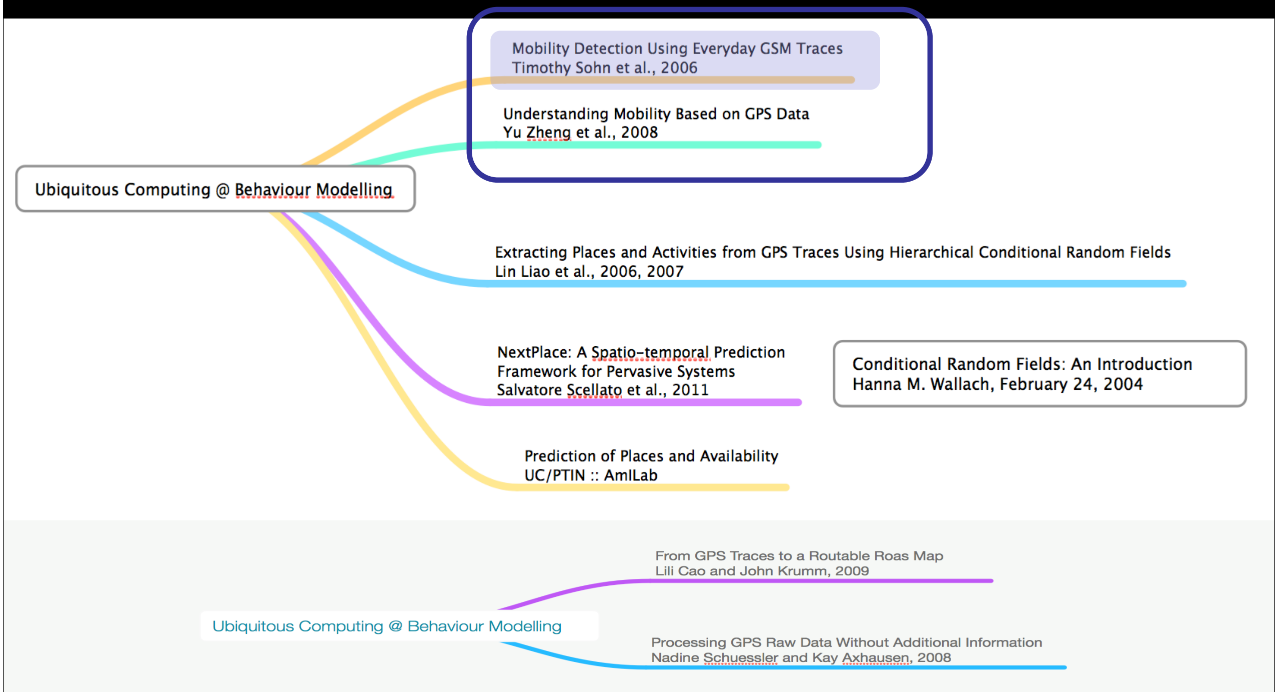
senseMyCity (FEUP)

- Motion
  - Accelerometer - Measures proper acceleration
  - Gyroscope - Measures the rotation rate
- Position
  - Magnetometer - Measures the ambient geomagnetic field
  - Proximity - Measures the proximity of an object relative to the view screen
- Location
  - Network - Provides location based on Cell-ID and Wi-Fi
  - GPS - Provides location and satellite information
- Environmental
  - Light - Measures the ambient light
  - Pressure - Measures the ambient air pressure
  - Temperature - Measures the ambient room temperature
  - Humidity - Measures the relative ambient humidity
- Transceiver
  - Wifi - Logs Access Points
  - Bluetooth - Logs Bluetooth devices information
- External
  - OBD - Logs vehicle related information
  - Zephyr - Logs heart related information
  - Vital Jacket - Logs ECG information
  - FREMU - Logs air condition information

Data Acquisition



Methods and algorithms



Methods and algorithms

Mobility Detection Using Everyday GSM Traces

Timothy Sohn[1], Alex Varshavsky[2], Anthony LaMarca[3], Mike Y. Chen[3], Tanzeem Choudhury[3], Ian Smith[3], Sunny Consolvo[3], Jeffrey Hightower[3], William G. Griswold[1], and Eyal de Lara[2]

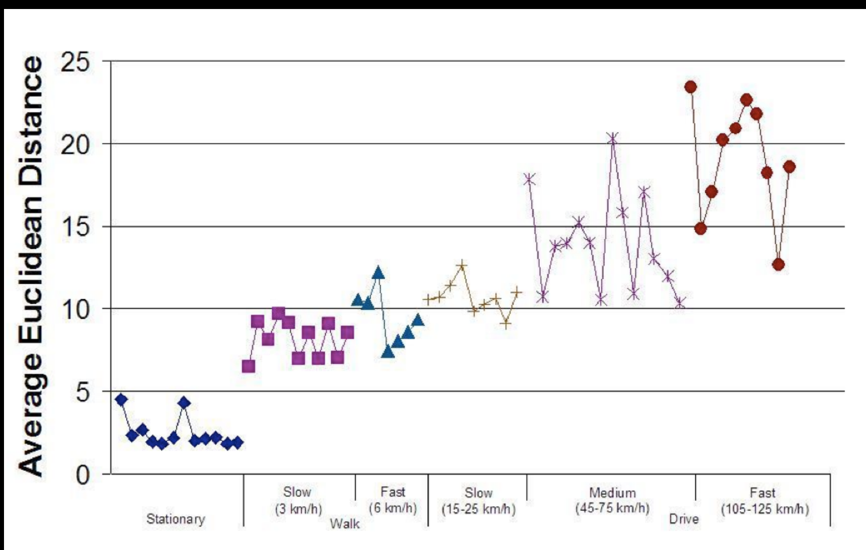
- Based on GSM fingerprints
  (assumes RSS is consistent in time but variable in space)
- Calculates the Euclidean Distance between measurements
  e.g. measurement A, 3 cells/channels $\{R_{1,A}, R_{2,A}, R_{3,A}\}$ and
  measurement B, $\{R_{1,B}, R_{2,B}, R_{3,B}\}$

  dist = $SQRT((R_{1,A} - R_{1,B})^2 + (R_{2,A} - R_{2,B})^2 + (R_{3,A} - R_{3,B})^2)$

Methods and algorithms

Mobility Detection Using Everyday GSM Traces

Timothy Sohn[1], Alex Varshavsky[2], Anthony LaMarca[3], Mike Y. Chen[3], Tanzeem Choudhury[3], Ian Smith[3], Sunny Consolvo[3], Jeffrey Hightower[3], William G. Griswold[1], and Eyal de Lara[2]



Carlos Lisboa Bento, FCTUC 9 - 45

Methods and algorithms

Mobility Detection Using Everyday GSM Traces

Timothy Sohn[1], Alex Varshavsky[2], Anthony LaMarca[3], Mike Y. Chen[3], Tanzeem Choudhury[3], Ian Smith[3], Sunny Consolvo[3], Jeffrey Hightower[3], William G. Griswold[1], and Eyal de Lara[2]

- Based on these findings they extract seven different features to use in classifying a set of GSM measurements as either
  - Stationary
  - Walking
  - Driving

Methods and algorithms

Mobility Detection Using Everyday GSM Traces

Timothy Sohn[1], Alex Varshavsky[2], Anthony LaMarca[3], Mike Y. Chen[3], Tanzeem Choudhury[3], Ian Smith[3], Sunny Consolvo[3], Jeffrey Hightower[3], William G. Griswold[1], and Eyal de Lara[2]

- The seven features:
  - Euclidean distance between two consecutive measurements
  - Spearman rank correlation coefficient between two consecutive measurements (how closely the signal strengths from common cell towers were ranked)
  - The number of common cell towers between two consecutive measurements
  - Mean Euclidean distance over a window of measurements where the values are calculated between consecutive measurements and then averaged together.
  - Variance in Euclidean distance values over a window of measurements where the values are calculated between consecutive measurements.
  - The variance in signal strengths for each tower seen within a given window.
  - Euclidean distance value between the first and the last measurement of a window.

Methods and algorithms

Mobility Detection Using Everyday GSM Traces

Timothy Sohn[1], Alex Varshavsky[2], Anthony LaMarca[3], Mike Y. Chen[3], Tanzeem Choudhury[3], Ian Smith[3], Sunny Consolvo[3], Jeffrey Hightower[3], William G. Griswold[1], and Eyal de Lara[2]

Precision

Ground Truth	Predicted Movement: Stationary	Walking	Driving
Stationary	95.4%	12.6%	6.9%
Walking	2.5%	70.2%	8.8%
Driving	2.1%	17.2%	84.3%

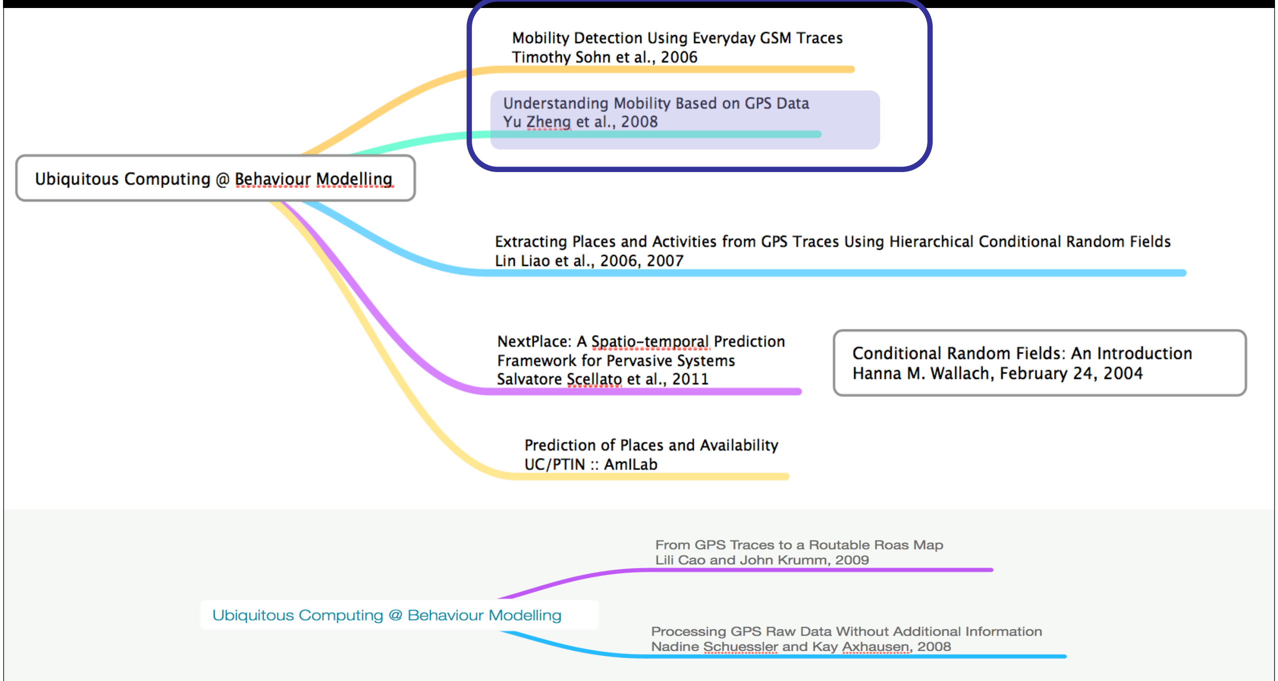
Recall

Ground Truth	Predicted Movement: Stationary	Walking	Driving
Stationary	92.5%	4.5%	3.0%
Walking	7.7%	80.0%	12.2%
Driving	4.5%	13.8%	81.7%

Precision and recall confusion matrices

Overall accuracy: 85% should:

Methods and algorithms



Methods and algorithms

Understanding Mobility Based on GPS Data
Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

- Based on GPS traces
- Logs from 65 people over 10 months
- Comprises a learning/classification phase improved by a pos-processing step
  - Change point-based segmentation + decision tree based inference model >> 8% accuracy improvement over previous results
  - Pos-processing (spatial knowledge extraction with transition probability-based enhancement) >> additional 4% accuracy improvement

Methods and algorithms

Understanding Mobility Based on GPS Data
Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

Four steps:

- Segmentation
- Feature extraction
- Inference
- Post-processing

stop here,
reading assignment

each student explain succinctly each step (3-4 slides) + plus one that gives the big picture (I'll organize slides together)

(send me the slides in PDF till Monday night)

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

Concepts:

- GPS trajectory = sequence of GPS points (latitude, longitude, timestamp)
- Segment = the division of a trajectory into trips
- Change point = a place were people change the transportation mode

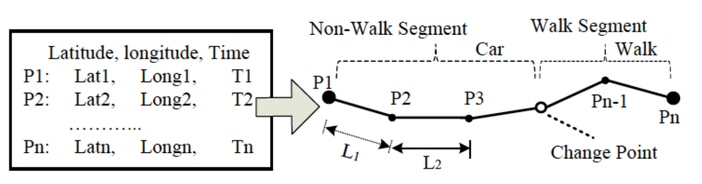


Figure 1. GPS log, segment and change point

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 1 :: Segmentation

Assumption: walking should be a transition between different transportation modes

1A :: distinguishes walk points from non walk points by *veloc. and accel.*

1B :: if the length of a segment composed by consecutive *walk points* or *non-walk points* is less than a threshold merge this into its backward segment

1C :: length of a segment exceeds a certain threshold
=> certain segment
if the number of uncertain segments exceeds a certain threshold
=> they will be merged into a non-walk segment

1D :: the start point and end point of each WALK SEGMENT are potential CHANGE POINTS to partition a trip

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 1 :: Segmentation

Assumption: walking should be a transition between different transportation modes

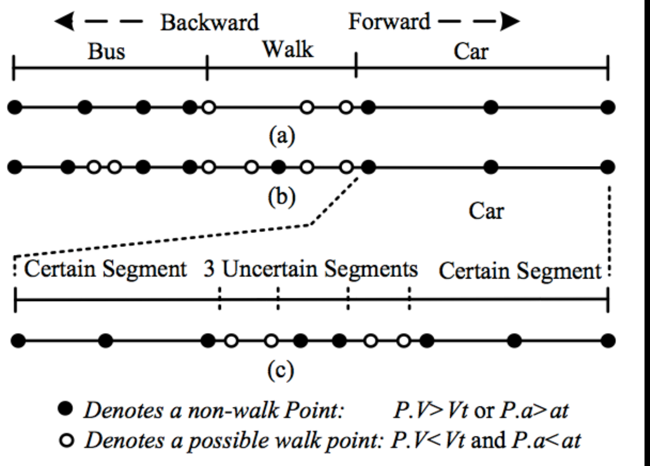


Figure 4. An example of detecting change points

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 2 :: Feature Extraction :: Calculating features from GPS points

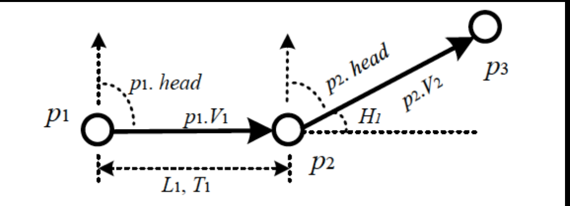


Figure 2. Feature calculation based on GPS logs

velocity of p_1

$$p_1.V_1 = L_1 / T_1$$

heading change

$$H_1 = |p_1.head - p_2.head|$$

also: • acceleration
• expectation of velocity

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 3 :: Inference :: Features for discrimination of motion mode

Features	Significance
Dist	Distance of a segment
MaxV*i*	The *i*th maximum velocity of a segment
MaxA*i*	The *i*th maximum acceleration of a segment
AV	Average velocity of a segment
EV	Expectation of velocity of GPS points in a segment
DV	Variance of velocity of GPS points in a segment
HCR	Heading Change Rate
SR	Stop Rate
VCR	Velocity Change Rate

Table 1. The features we explored in the experiment

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 3 :: Inference :: Three robust features for discrimination of motion mode

HEADING CHANGE RATE (HCR)

walking or cycling from driving a car or taking a bus

HCR = |Pc| / Distance

collection of GPS points at which a user changes his/her heading direction exceeding a certain threshold (Hc), and |Pc| represents the number of elements in Pc

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 3 :: Inference :: Three robust features for discrimination of motion mode

STOP RATE (SR)

number of GPS points with velocity below a certain threshold within a unit distance.

$$SR = |P_s| / Distance$$

$$P_s = \{p_i \mid p_i \in P, p_i.V < V_s\}$$

in general

$$SR(Walk) > SR(Bus) > SR(Drive)$$

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 3 :: Inference :: Three robust features for discrimination of motion mode

VELOCITY CHANGE RATE (VCR)

first we calculate the VRate at each point using

$$p_1.VRate = |V_2 - V_1| / V_1$$

then we have

$$VCR = |P_v| / Distance$$

points with a VRate (velocity change rate) higher than a threshold

$$P_v = \{p_i \mid p_i \in P, p_i.VRate > V_r\}$$

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 3 :: Inference :: Three robust features for discrimination of motion mode

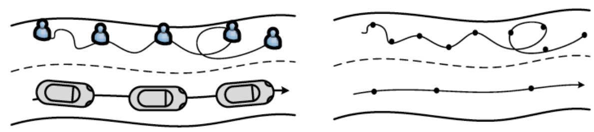


Figure 5 Heading change rate of different modes

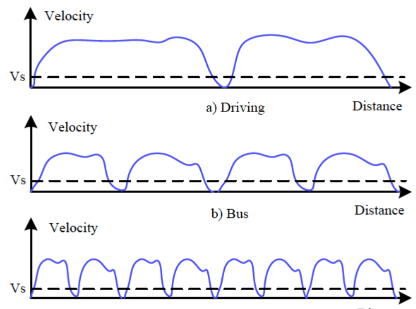
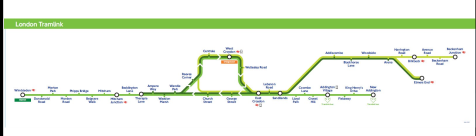


Figure 6 Velocity change rate and stop rate



Stop Rate

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 3 :: Inference

- Support vector machines (SVM)
- Decision tree (outperformed the other methods)
- Bayesian net
- Conditional Random Field (CRF)

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 3 :: Inference :: Three robust features for discrimination of motion mode

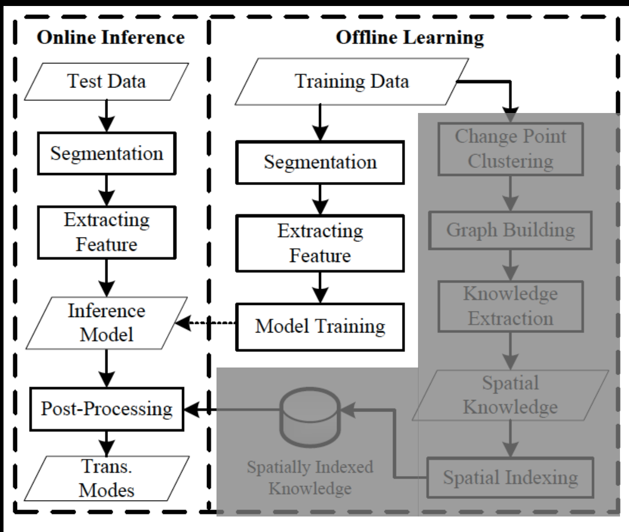


Figure 3 Architecture of our approach

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 4 :: Post-processing

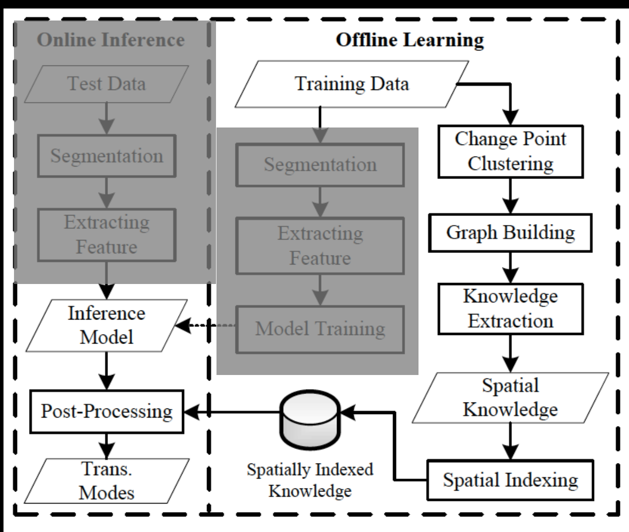


Figure 3 Architecture of our approach

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 4 :: Post-processing

4A :: clustering of change points and start/end points (using density-based clustering algorithms

4B :: building graph

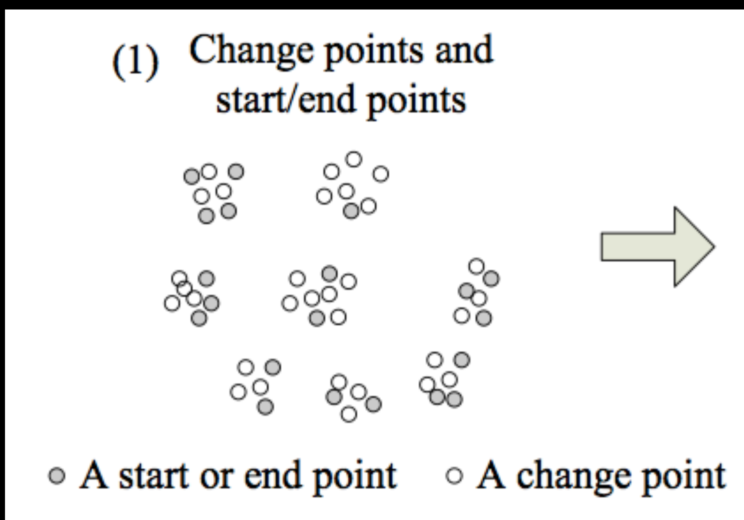
4C :: spatial indexing

4D :: probability calculation

Methods and algorithms

Understanding Mobility Based on GPS Data
Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

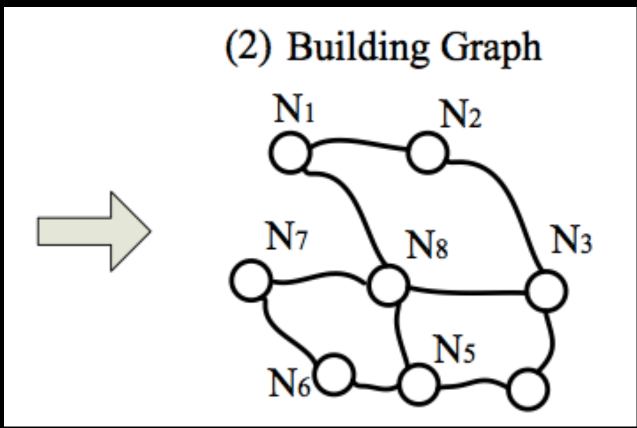
STEP 4 :: Post-processing :: clustering



Methods and algorithms

Understanding Mobility Based on GPS Data
Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 4 :: Post-processing :: Building graph

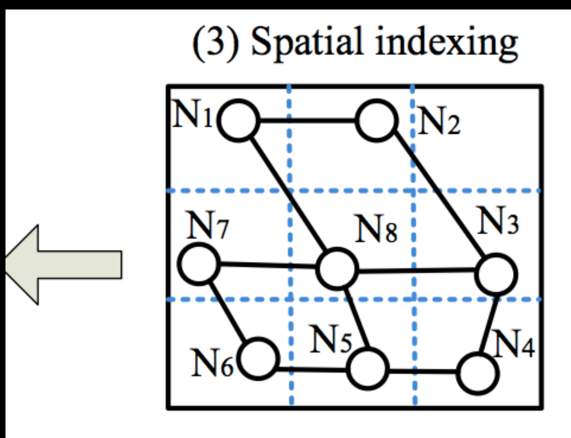


Simplification :: all trajectories passing two graph nodes are regarded as similar.

Methods and algorithms

Understanding Mobility Based on GPS Data
Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 4 :: Post-processing :: Spatial indexing

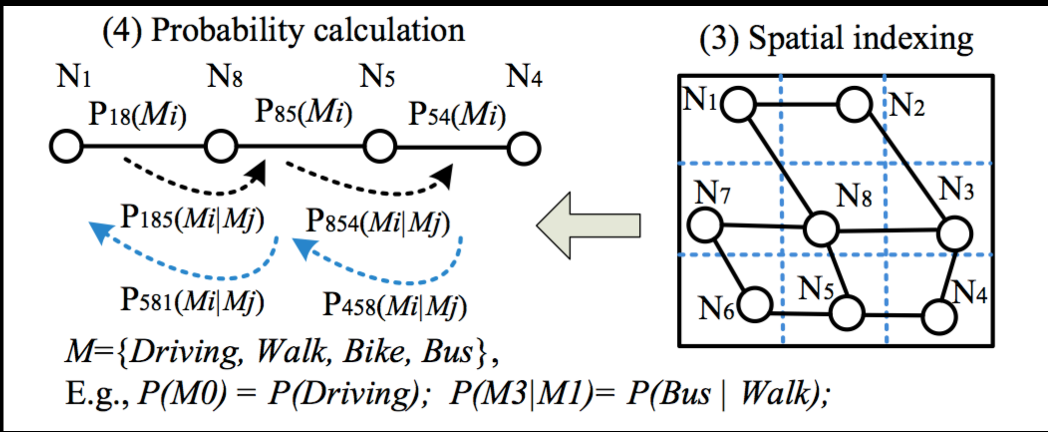


Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 4 :: Post-processing :: Probability calculation



Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 4 :: Post-processing :: Probability calculation

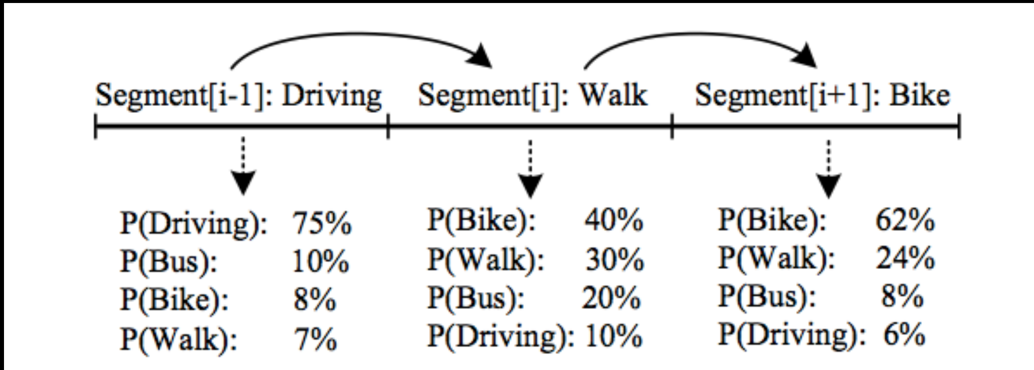


Figure 9. Perform normal post-processing on a GPS trajectory

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 4 :: Post-processing :: Probability calculation

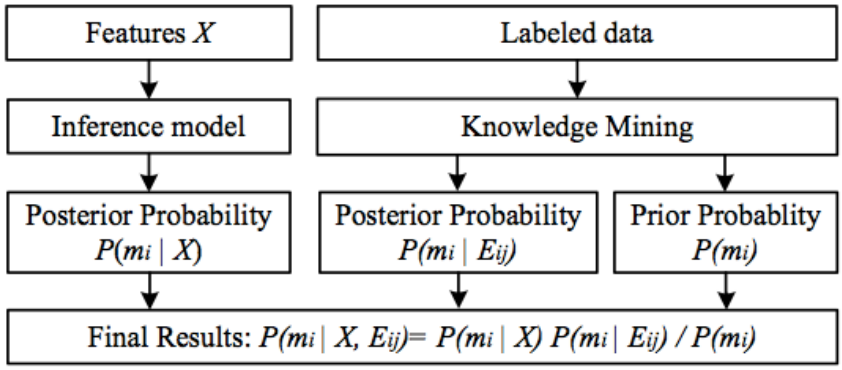


Figure 10. Prior probability-based boosting method

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 4 :: Post-processing :: Probability calculation

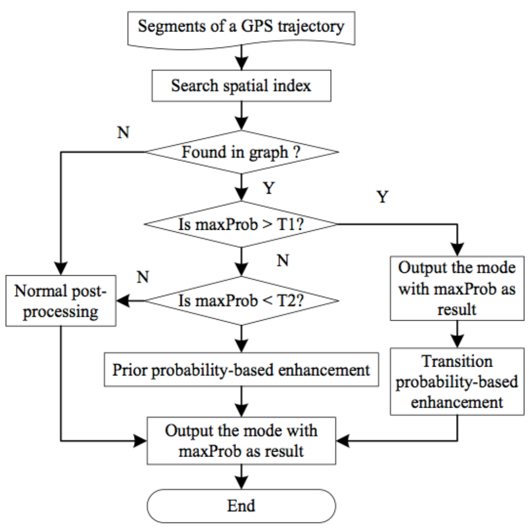


Figure 8 Flowchart of the post-processing

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

STEP 4 :: Post-processing :: Probability calculation

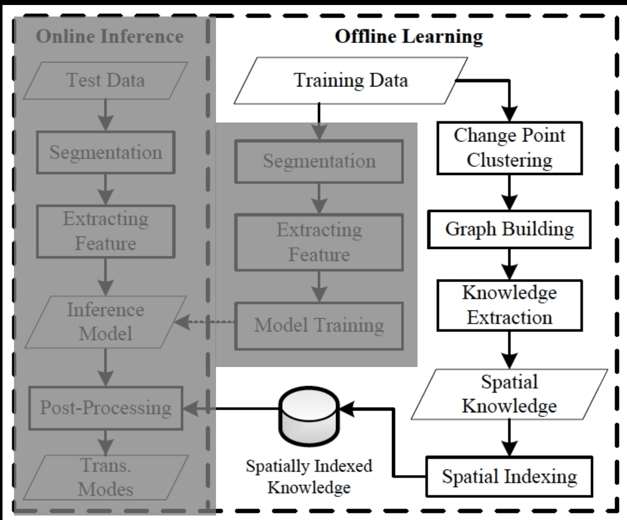


Figure 3 Architecture of our approach

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

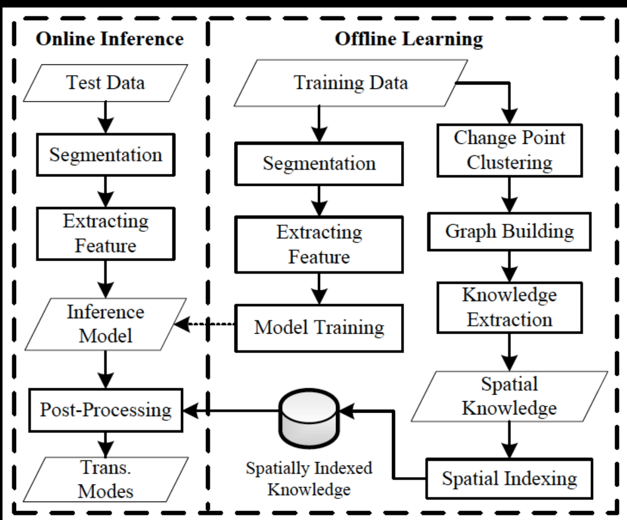


Figure 3 Architecture of our approach

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

	A_D	CP/P	CP/R
Enhanced Features (*EF*)	0.728	0.491	0.817
EF + normal post-processing	0.741	0.508	0.818
EF + graph-based post-processing	**0.762**	**0.516**	**0.818**

Table 4 Comparison between normal post-processing and graph-based post-processing

Methods and algorithms

Understanding Mobility Based on GPS Data

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, Wei-Ying Ma
Microsoft Research Asia
4F, Sigma Building, No.49 Zhichun Road, Haidian District, Beijing 100190, P. R. China
{yuzheng, v-quali, v-yukche, xingx, wyma}@microsoft.com

Ground truth	Predicted Results (KM)					
	Walk	Driving	Bus	Bike		
walk	1026.4	122.1	386.5	357.3	0.543	Recall
Driving	42.6	2477.3	458.5	235.1	0.771	
Bus	34.8	164.7	1752.4	46.2	0.877	
Bike	49.3	113.5	31.9	1234.3	0.864	
	0.891	0.861	0.666	0.659	**0.762**	
	Precision					

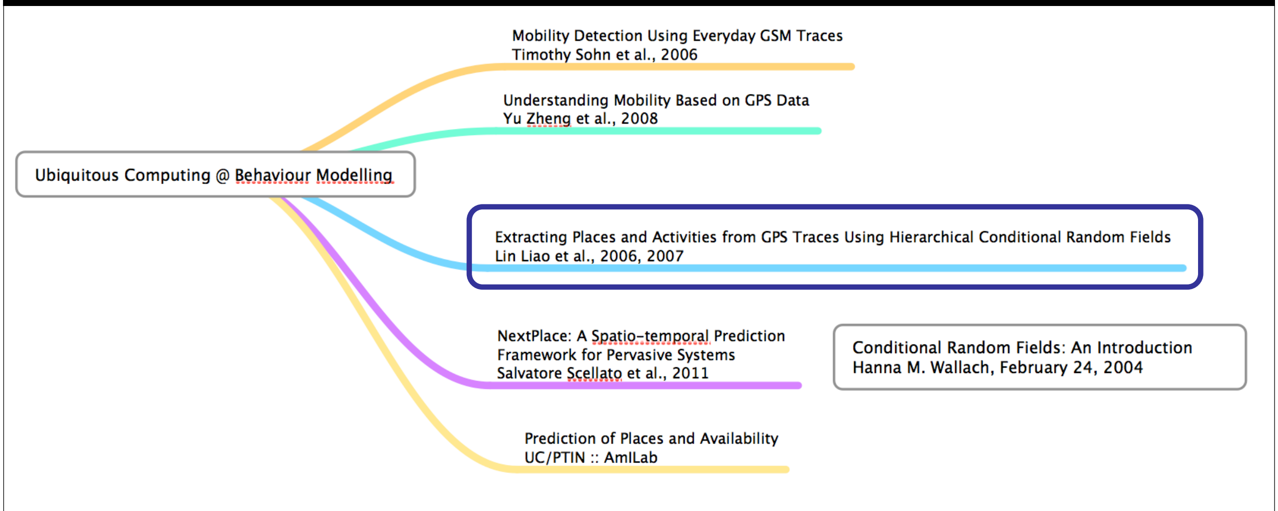
Table 5. Fusion matrix of final inference results with graph-based post-processing

BEHAVIOUR MODELLING

What is expected students learn on this topic:

- Why modelling user behaviour is important
- The role of mobile and opportunistic data on behaviour modelling
- Methods and algorithms for behaviour modelling
  - Mode choice modelling
  - Prediction of next place
- Privacy issues

Methods and algorithms



Methods and algorithms

Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields

Lin Liao Dieter Fox Henry Kautz
Department of Computer Science & Engineering
University of Washington
Seattle, WA 98195
Email:{liaolin,fox,kautz}@cs.washington.edu

Hierarchically structured conditional random fields for generation of consistent models of:

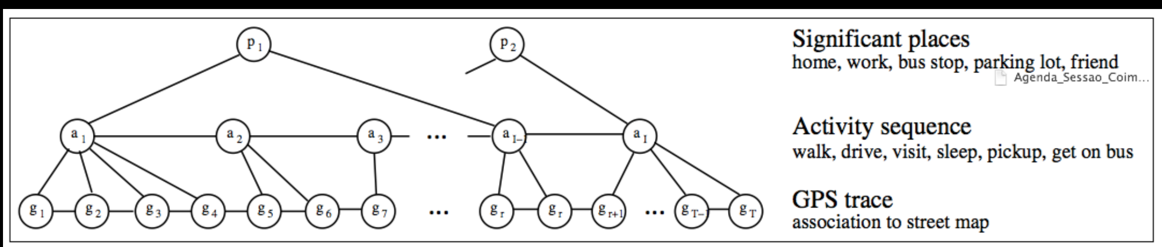
- activity
- places

Methods and algorithms

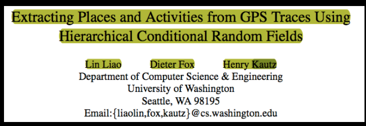
Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields

Lin Liao Dieter Fox Henry Kautz
Department of Computer Science & Engineering
University of Washington
Seattle, WA 98195
Email:{liaolin,fox,kautz}@cs.washington.edu

Hierarchy for location-based activity recognition

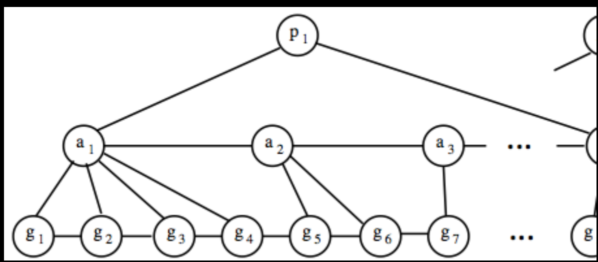


Methods and algorithms


Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields

Lin Liao Dieter Fox Henry Kautz
Department of Computer Science & Engineering
University of Washington
Seattle, WA 98195
Email:{liaolin,fox,kautz}@cs.washington.edu

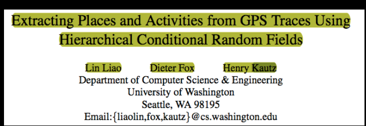
Hierarchical conditional random fields



- Hidden states $<y_1, \ldots, y_T>$
- Observations $<x_1, \ldots, x_T>$

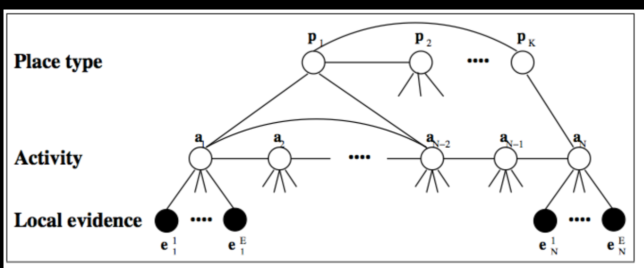
P(y|x)

Methods and algorithms


Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields

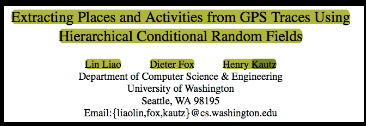
Lin Liao Dieter Fox Henry Kautz
Department of Computer Science & Engineering
University of Washington
Seattle, WA 98195
Email:{liaolin,fox,kautz}@cs.washington.edu

Hierarchical conditional random fields for labeling activities and places



- Places :: generated based on the activities inferred at the activity level… each place connected to the activities within a certain range
- Activity nodes :: range over activities
- Local evidences :: time of day, duration, motion velocity…

Methods and algorithms


Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields

Lin Liao Dieter Fox Henry Kautz
Department of Computer Science & Engineering
University of Washington
Seattle, WA 98195
Email:{liaolin,fox,kautz}@cs.washington.edu

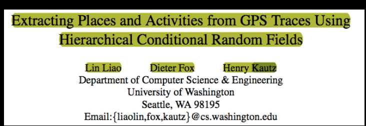
Algorithm for jointly detecting significant places and inferring activities and types of places

Input: GPS trace $\langle g_1, g_2, \ldots, g_T \rangle$
$i := 0$
// Generate activity segments and local evidence by grouping consecutive GPS readings
$(\langle a_1, \ldots, a_N \rangle, \langle e_1^1, \ldots, e_1^E, \ldots, e_N^1, \ldots, e_N^E \rangle) :=$ spatial_segmentation$(\langle g_1, g_2, \ldots, g_T \rangle)$
// Generate CRF containing activity and local evidence nodes (lower two levels in Figure 3)
$\mathrm{CRF}_0 :=$ instantiate_crf$(\langle\ \rangle, \langle a_1, \ldots, a_N \rangle, \langle e_1^1, \ldots, e_1^E, \ldots, e_N^1, \ldots, e_N^E \rangle)$
// Determine MAP sequence of activities
$\mathbf{a}_0^* :=$ MAP_inference(CRF_0)
do
    $\mathrm{i} := \mathrm{i} + 1$
    // Generate places by clustering significant activities
    $\langle p_1, \ldots, p_K \rangle_i :=$ generate_places$(\mathbf{a}_{i-1}^*)$
    // Generate complete CRF with instantiated places
    $\mathrm{CRF}_i :=$ instantiate_crf$(\langle p_1, \ldots, p_K \rangle_i, \langle a_1, \ldots, a_N \rangle, \langle e_1^1, \ldots, e_1^E, \ldots, e_N^1, \ldots, e_N^E \rangle)$
    // Perform MAP inference in complete CRF
    $\langle \mathbf{a}_i^*, \mathbf{p}_i^* \rangle :=$ MAP_inference(CRF_i)
until $\mathbf{a}_i^* = \mathbf{a}_{i-1}^*$
return $\langle \mathbf{a}_i^*, \mathbf{p}_i^* \rangle$

Agenda_Sessao

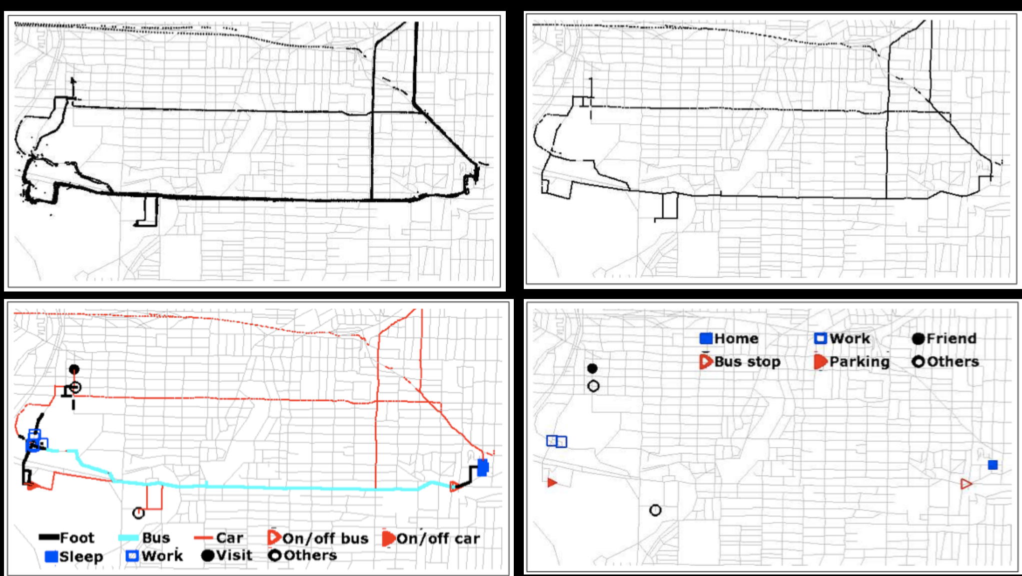
MAP = maximum a posteriori probability

Methods and algorithms


Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields

Lin Liao Dieter Fox Henry Kautz
Department of Computer Science & Engineering
University of Washington
Seattle, WA 98195
Email:{liaolin,fox,kautz}@cs.washington.edu

Experimental results



Methods and algorithms

Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields

Lin Liao Dieter Fox Henry Kautz
Department of Computer Science & Engineering
University of Washington
Seattle, WA 98195
Email:{liaolin,fox,kautz}@cs.washington.edu

Experimental results

Truth	Inferred labels							FN
	Work	Sleep	Leisure	Visit	Pickup	On/off car	Other	
Work	12 / 11	0	0 / 1	0	0	0	1	0
Sleep	0	21	1	2	0	0	0	0
Leisure	2	0	20 / 17	1 / 4	0	0	3	0
Visiting	0	0	0 / 2	7 / 5	0	0	2	0
Pickup	0	0	0	0	1	0	0	2
On/Off car	0	0	0	0	1	13 / 12	0	2 / 3
Other	0	0	0	0	0	0	37	1
FP	0	0	0	0	2	2	3	-

Methods and algorithms

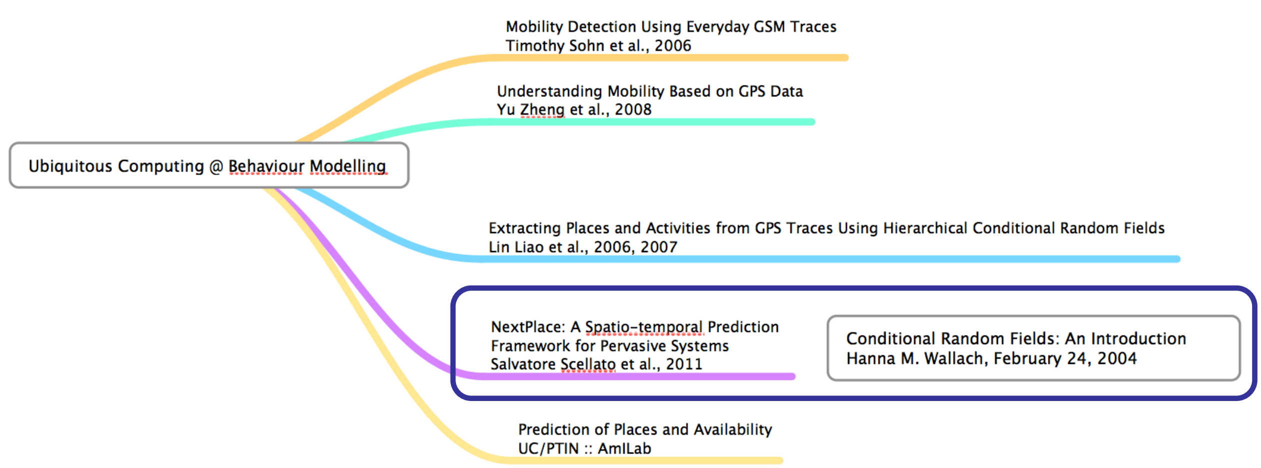
Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields

Lin Liao Dieter Fox Henry Kautz
Department of Computer Science & Engineering
University of Washington
Seattle, WA 98195
Email:{liaolin,fox,kautz}@cs.washington.edu

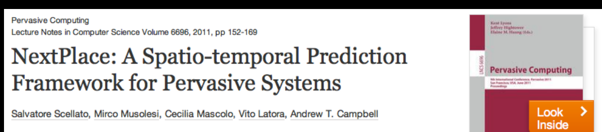
Experimental results

Truth	Inferred labels					FN
	Work	Home	Friend	Parking	Other	
Work	5	0	0	0	0	0
Home	0	4	0	0	0	0
Friend	0	0	3	0	2	0
Parking	0	0	0	8	0	2
Other	0	0	0	0	28	1
FP	0	0	1	1	2	-

Methods and algorithms



Methods and algorithms


Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Predicting User Behaviour in NextPlace

WHERE, **WHEN** and **HOW LONG** the next visit will occur

Methods and algorithms

Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Previous works based on linear or probabilistic models on predicting of next location

This work based on nonlinear time series analysis of the arrival times of users in relevant places

Focus on the estimation of the duration of a visit to a location and the intervals between two subsequent visits

Methods and algorithms

Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

A visit of a user is defined by a tuple (u, loc, t, d)

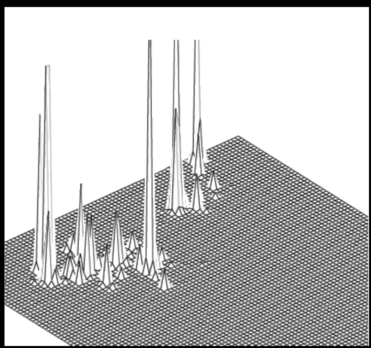
u ≈ user
loc ≈ location
t ≈ time of arrival
d ≈ residence time in location loc

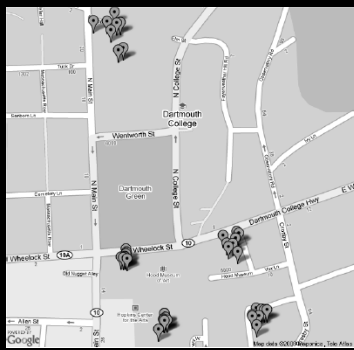
Methods and algorithms


Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169
NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems
Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Extracting significant places from GPS data

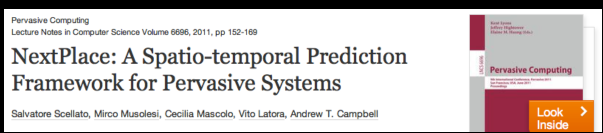
Assumption: permanence at a place is directly proportional to the importance that is attributed to the place


Frequency map. Higher peaks reveal places where user spent most of their time



Significant places with a threshold equal to 15% of the maximum residence time observed

Methods and algorithms


Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169
NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems
Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

For each place they try to predict:

1) When the next visit will take place
2) How long it will be

For each user they keep track
1) For each place
2) The historic of visits $((t_1, d_1), \ldots (t_n, d_n))$

And create two time series
1) $C = (c_1, \ldots, c_n)$ start time series
2) $D = (d_1, \ldots, d_n)$ duration time series

Methods and algorithms

Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Predicting User Behaviour Algorithm

1) Create time series $C = (c_1, \ldots, c_n)$ start time $D = (d_1, \ldots, d_n)$ duration
2) Search in C sequences of m consecutive values (c_{i-m+1}, c_i)
3) Estimate the next value of time series C by averaging all values c_{i+1} that follow each found sequence
4) At the same time in respective D time series select the corresponding sequence $(d_{i-m+1}, \ldots, d_i)$
5) Estimate the next value of time series D by averaging all the values d_{i+1} that follow these sequences

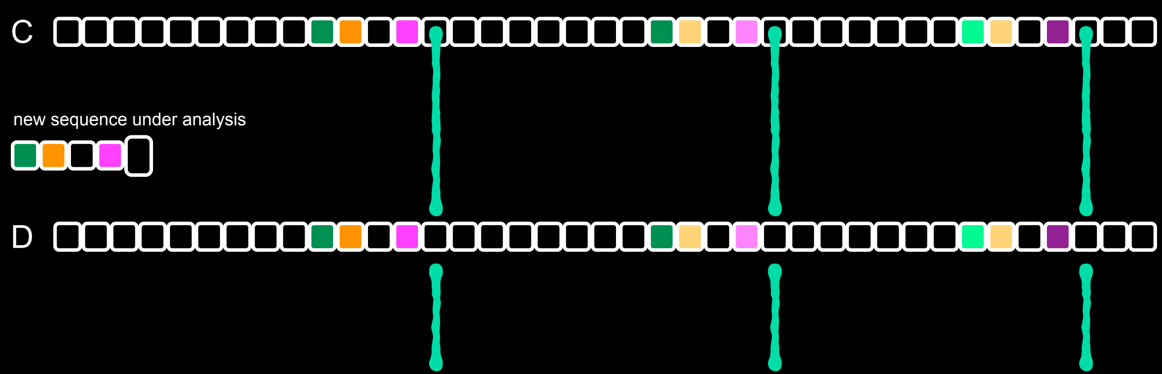
Methods and algorithms

Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

on a daily basis (we can group into week and weekend days)



calculates average time and duration of the visit to the next place

Methods and algorithms


Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

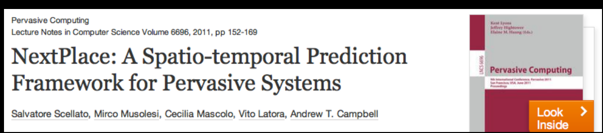
NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Concepts and assumptions

- non-linear time series analysis
- use of delay embedding
  $(s_0, s_1, \ldots, s_N)$
  embedded in a m-dimensional space with a delay v
  $B_n=(s_{n-(m-1)v}, s_{n-(m-2)v}, \ldots s_{n-v}, s_n)$
- in the current approach v=1

Methods and algorithms


Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Validation (Datasets)

- **Cabspotting :: GPS traces of taxi cabs in San Francisco :: 500 taxis, along 30 days :: used by dispatchers to efficiently reach customers**
- **CenceMe GPS :: a system for recreational personal sensing :: 20 NokiaN95 phones carried by posgrads and staff at Dartmouth College**
- **Dartmouth Wifi :: data from the SNMP logs of the wifi Dartmouth College**
- **Ile Sans Fils :: non-profit organization operates a network of free WiFi :: 45000 users with 140 hotspots :: mostly at cafes, restaurants, bars, libraries, but also outdoor around parks and commercial streets**

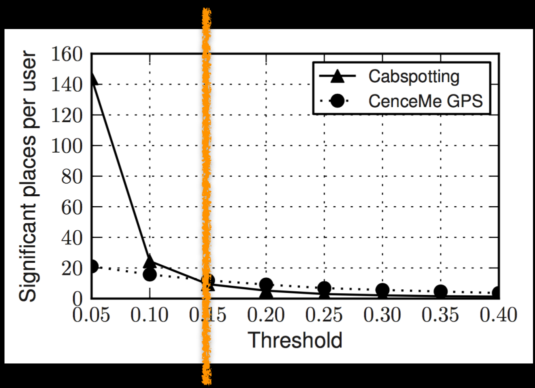
Methods and algorithms

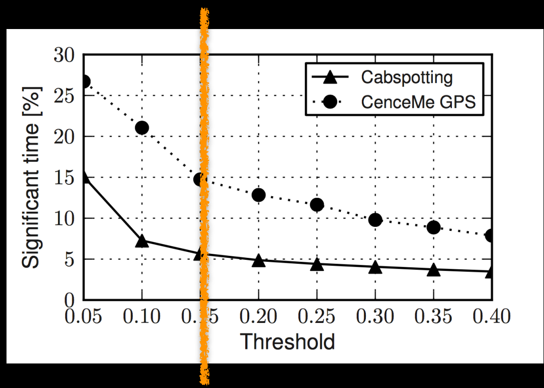

Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

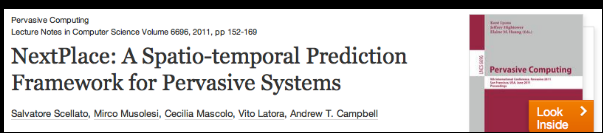
Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Validation (significant places)





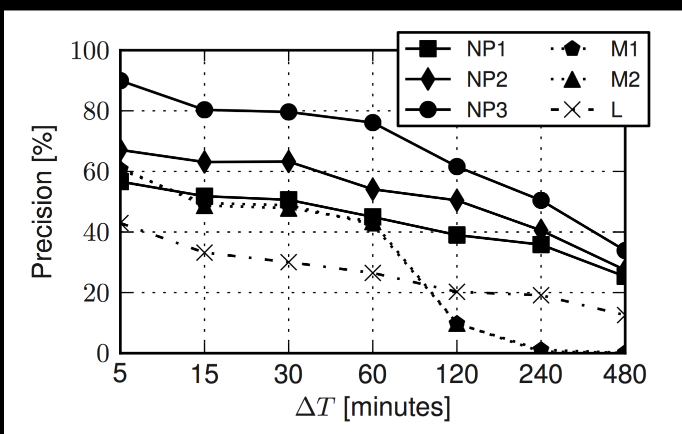
Methods and algorithms


Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Validation (m=1, 2, 3 and Markov 1st and 2nd order)



(a) Cabspotting

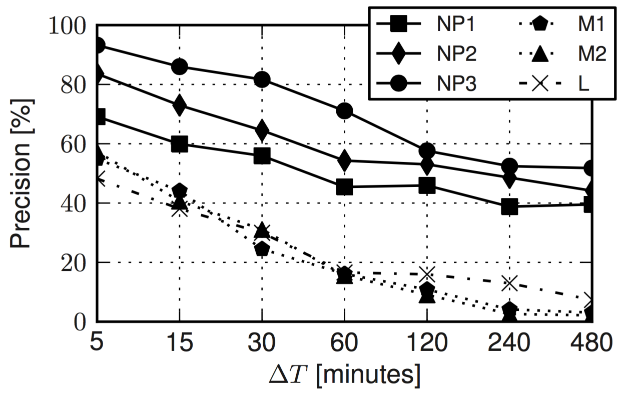
Methods and algorithms

Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

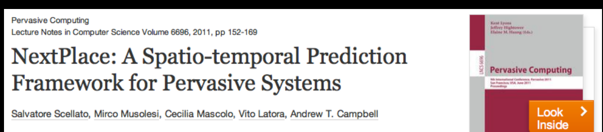
Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Validation (m=1, 2, 3 and Markov 1st and 2nd order)



(b) CenceMe GPS

Methods and algorithms

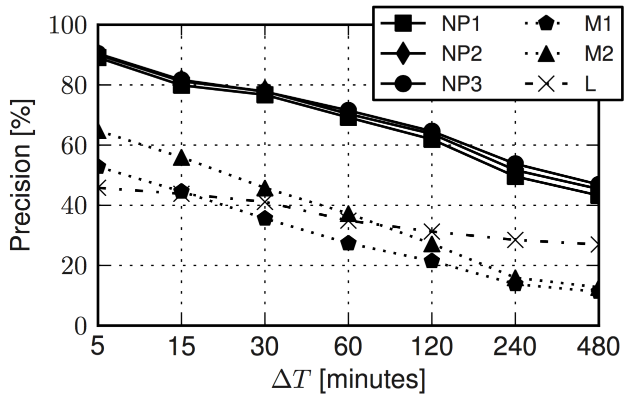


Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

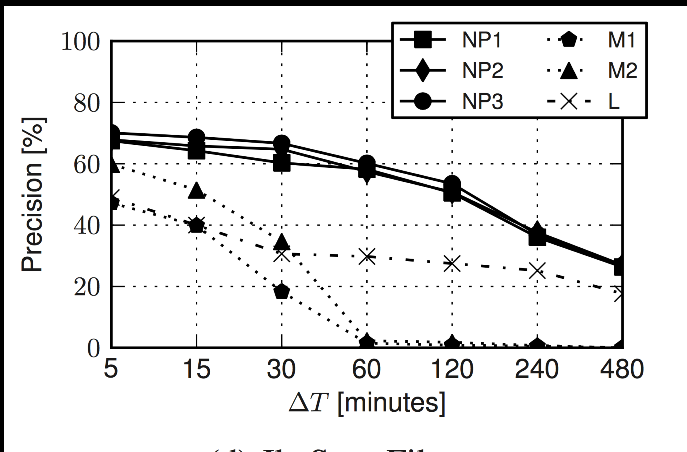
Validation (m=1, 2, 3 and Markov 1st and 2nd order)



(c) Dartmouth WiFi

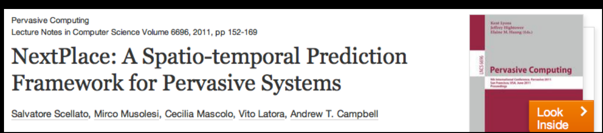
Methods and algorithms

Validation (m=1, 2, 3 and Markov 1st and 2nd order)

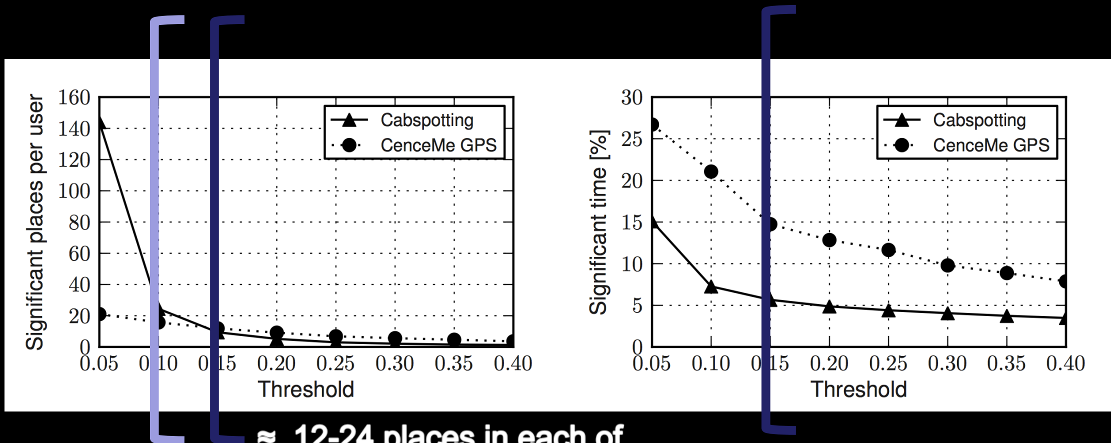


(d) Ile Sans Fils

Methods and algorithms



Validation (average number of significant places and percentage time spent)



≈ 12-24 places in each of these datasets

Carlos Lisboa Bento, FCTUC 9 - 45

Methods and algorithms

Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

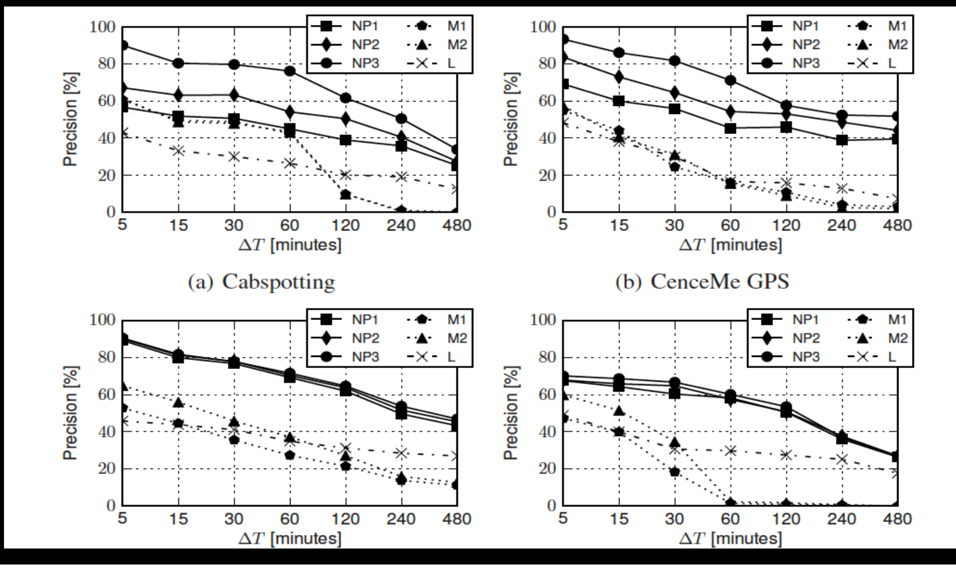
NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Pervasive Computing

Look Inside

Prediction precision as a function of time interval ΔT for the different datasets



Methods and algorithms

Pervasive Computing
Lecture Notes in Computer Science Volume 6696, 2011, pp 152-169

NextPlace: A Spatio-temporal Prediction Framework for Pervasive Systems

Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, Vito Latora, Andrew T. Campbell

Pervasive Computing

Look Inside

An approach inspired on Scellato et al paper…

use (1) your places collect from MOVES
or (2) perform POIs detection from GPS traces or at least use the DB provided at USys site with annotated traces

For each place they try to predict:

1) When the next visit will take place
2) How long it will be

For each user they keep track
1) For each place
2) The historic of visits $((t_1, d_1), \ldots (t_n, d_n))$

And create two time series
1) $C = (c_1, \ldots, c_n)$ start time series
2) $D = (d_1, \ldots, d_n)$ duration time series

Methods and algorithms

An approach inspired on Scellato et al paper…

Predicting User Behaviour Algorithm

1) Create time series C = (c_1, …, c_n) start time D = (d_1, …, d_n) duration
2) Search in C sequences of m consecutive values (c_{i-m+1}, c_i)
3) Estimate the next value of time series C by averaging all values c_{i+1} that follow each found sequence
4) At the same time in respective D time series select the corresponding sequence (d_{i-m+1}, …, d_i)
5) Estimate the next value of time series D by averaging all the values d_{i+1} that follow these sequences

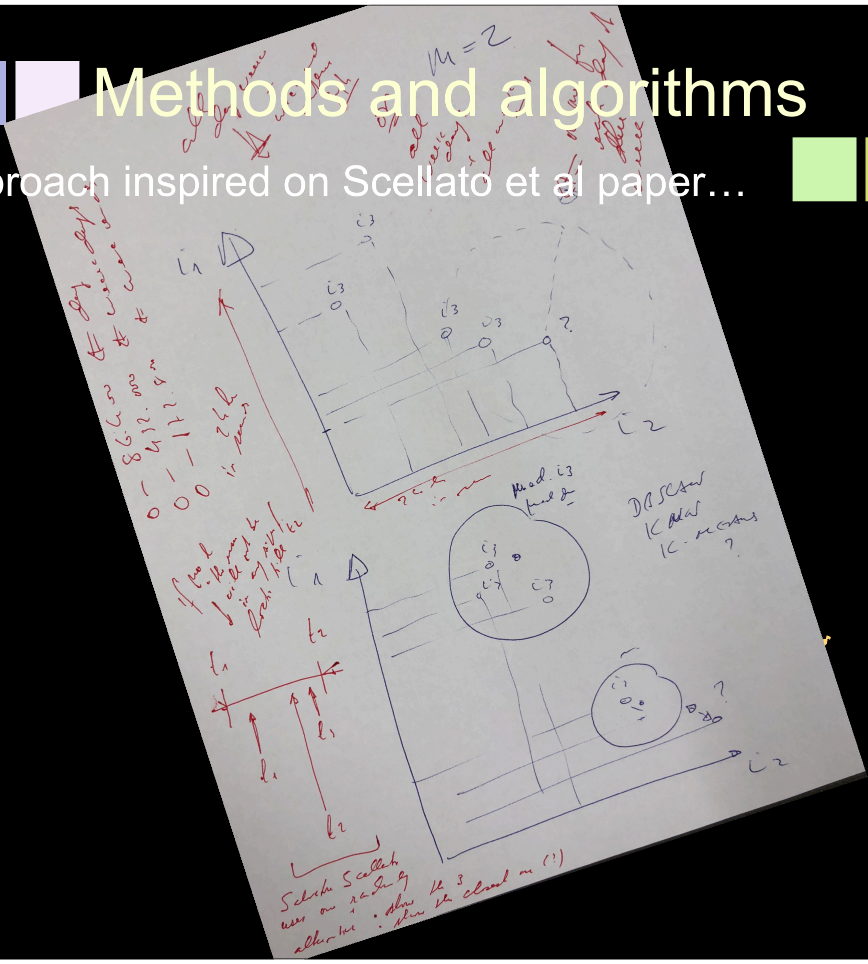
(1) try m=2, 3, 4 (show the results)
(2) organize C in m-dimension spaces, and
    (1) calculate the minimum distance between the new instance and the examples

    OR

    (2) use a spatial clustering algorithm for the points from C in the m-dimensional space, calculate the centroid of the clusters and calculate the minimum distance of the instance to the centroids; use the average time and duration for m+1 as the result for your algorithm

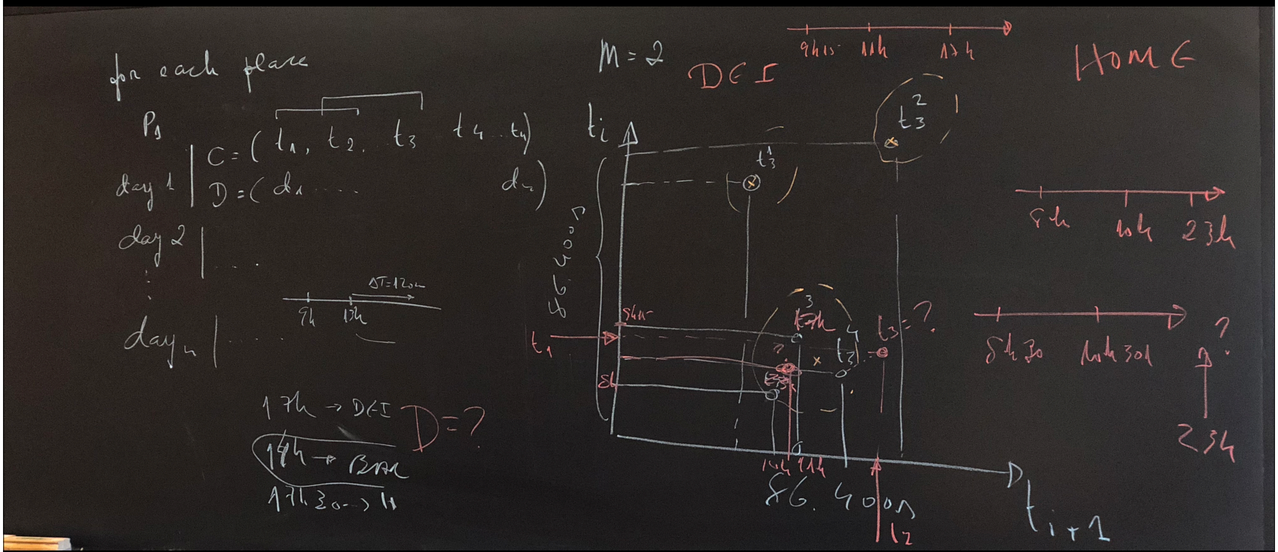
Methods and algorithms

An approach inspired on Scellato et al paper…



Methods and algorithms

An approach inspired on Scellato et al paper…



FCTUC - UNIVERSIDADE DE COIMBRA

UBIQUITOUS SYSTEMS

BEHAVIOUR MODELLING
(FOR URBAN COMPUTING)

CARLOS BENTO
MEI - FCTUC