Advanced Techniques for Mobile Robotics

Location-Based Activity Recognition

Wolfram Burgard, Cyrill Stachniss,

Kai Arras, Maren Bennewitz

UNI FREIBURG

Activity Recognition

Based on

- L. Liao, D. J. Patterson, D. Fox, and H. Kautz

  Learning and Inferring Transportation Routines
  Journal Artificial Intelligence, 2007

- L. Liao, D. Fox, and H. Kautz

  Extracting Places and Activities from GPS Traces
  Using Hierarchical Conditional Random Fields
  Int. Journal of Robotics Research, 2007

Motivation (1)

- Long-term monitoring of activities of daily living
- Learn typical navigation / transportation routines from user locations (GPS traces)
- Real-time tracking and predicting a user’s behavior
- Recognizing user errors
- Guidance for people with cognitive disabilities (e.g., Alzheimer's patients)

Motivation (2)

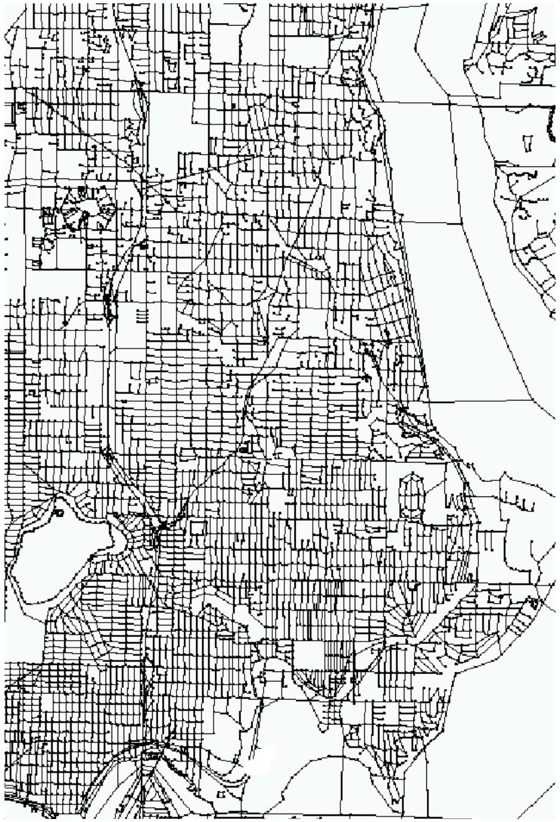
- Recognize daily activities (working, visiting friends, shopping, ...)
- Infer significant places (home, workplace, friends, stores, restaurants, ...)
- To provide location-based information services (e.g., searching nearby restaurants)
- For behavior analysis / personal guidance systems to help cognitively impaired people

Learning and Reasoning About Transportation Routines

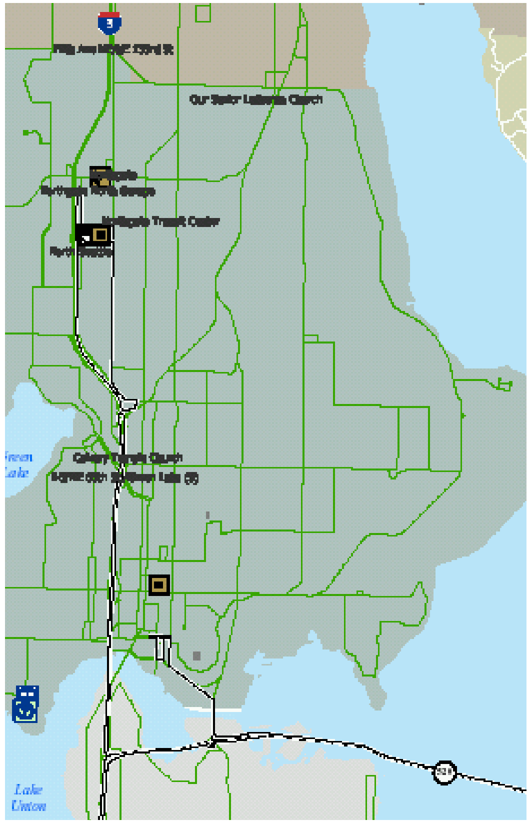
Given the data stream of a GPS device

- Track a user's location
- Infer the user's mode of transportation (foot, car, bus, ...)
- Predict the future movements (short-term and distant goals)
- Detect novel behavior / user errors

Geographic Information Systems



Street map

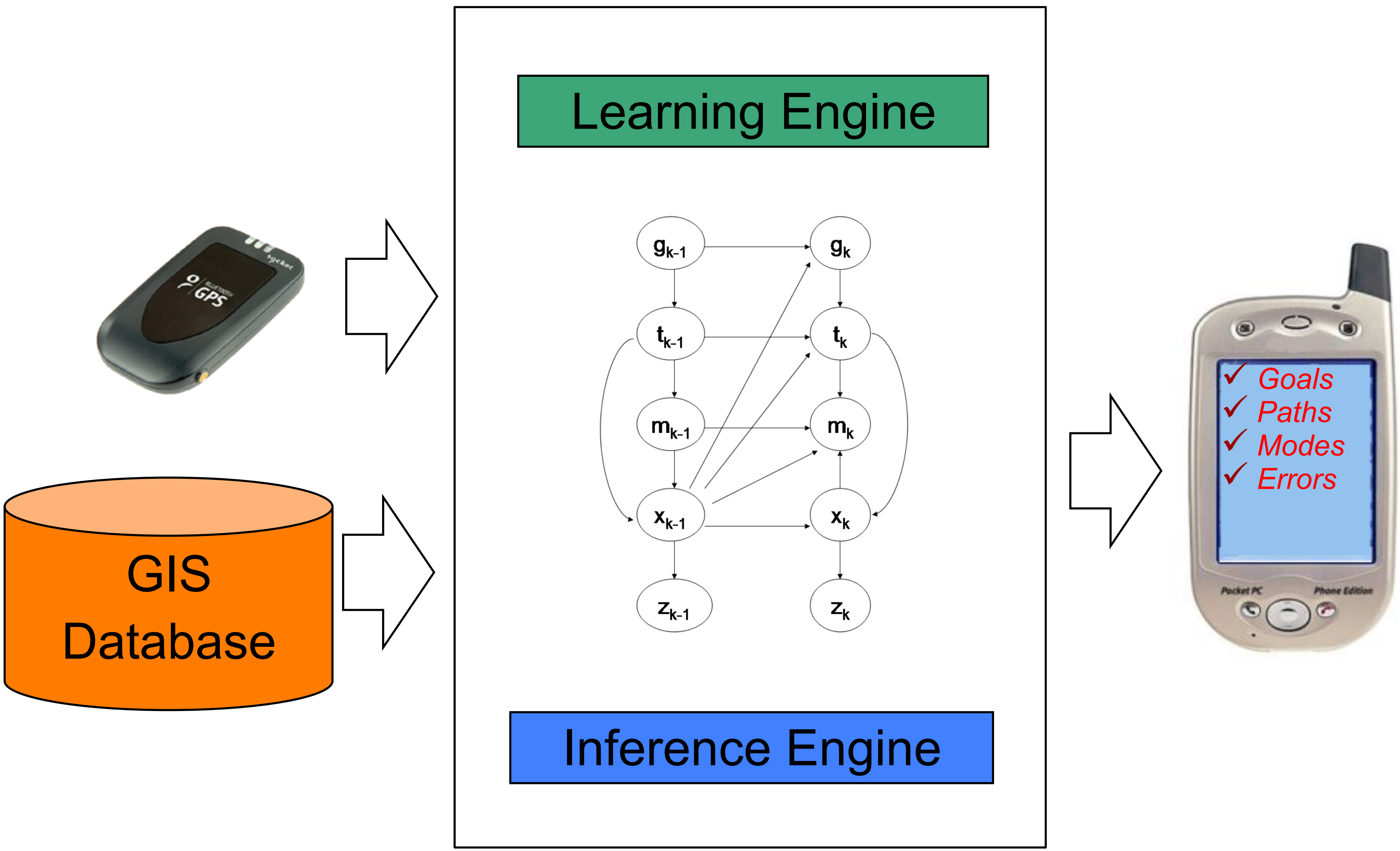


Bus routes and bus stops

GPS-Tracking is not Trivial

- GPS errors
- Dead zones near buildings, trees, ...
- Sparse measurements inside vehicles (bus)
- Multiple possible paths
- Inaccurate street map

Architecture



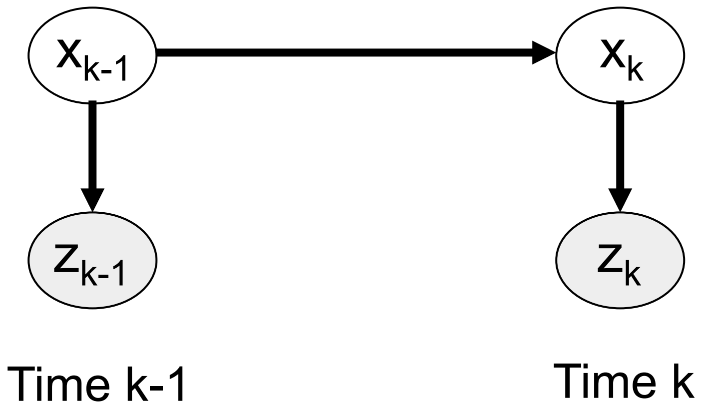
Probabilistic Inference

- Hierarchical activity model:
  3-level dynamic Bayesian network (DBN) to model temporal dependencies as well as
  - Novel behavior (top level)
  - Navigation goal (second level)
  - Transportation mode, location, and velocity (lowest level)
- Inference via Rao-Blackwellized particle filter in combination with a Kalman filter
- Parameter learning via Expectation-Maximization (EM)

Lowest Level of the DBN

- Estimation of transportation mode, location, and velocity
- Use the given street map as a directed graph
- Define a location as:
  - An edge/street with a direction (up/down)
  - Distance from start vertex of edge
- Prediction:
  - Move along the edges according to the velocity model
- Correction:
  - Update the estimate based on GPS readings

Dynamic Bayesian Network



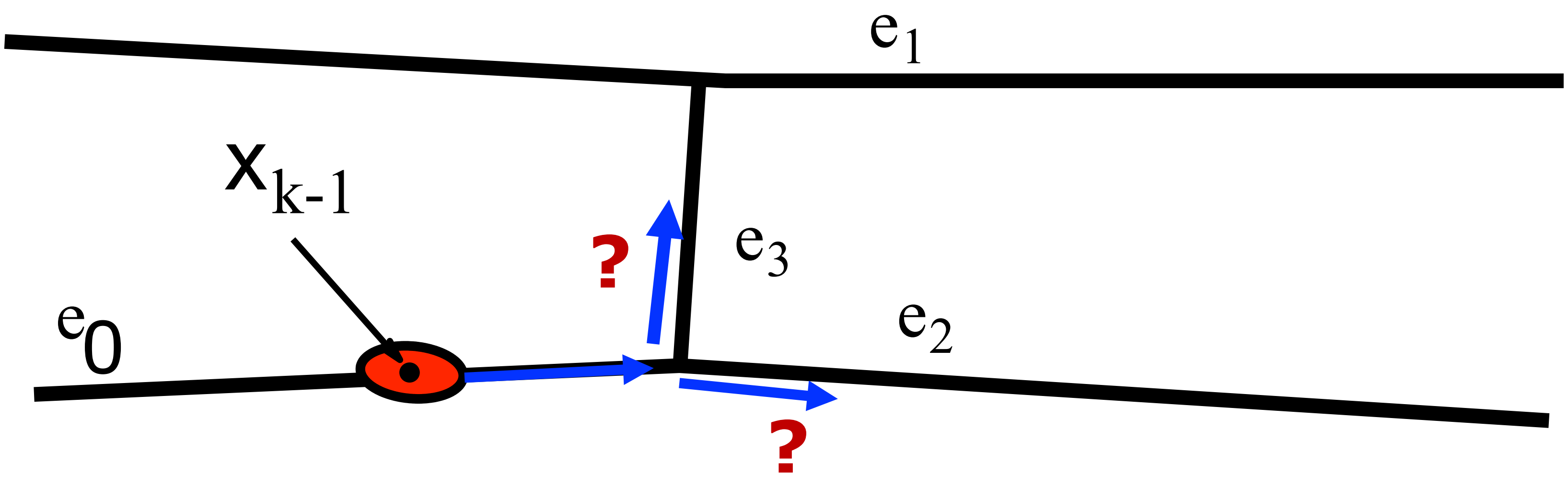
Edge, velocity, position

GPS reading

Task: Estimate the posterior over the hidden variables

slide credit: D. Fox

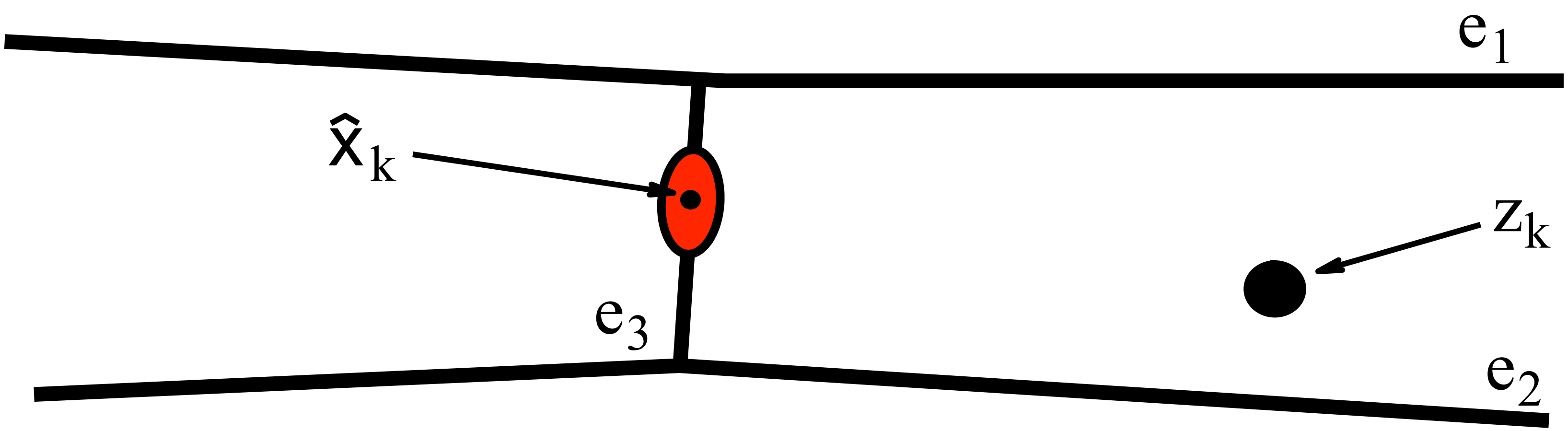
Kalman Filtering on a Graph: Prediction Step



Problem: Predicted location is multi-modal

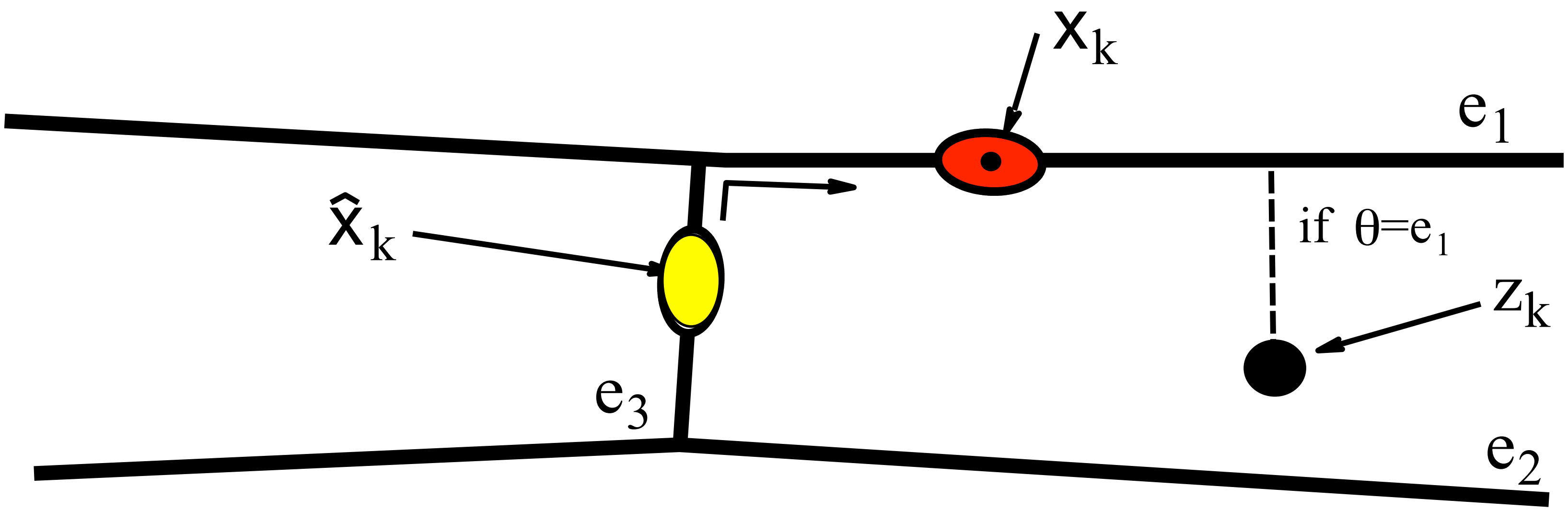
slide credit: D. Fox

Kalman Filtering on a Graph: Correction Step



Problem: GPS reading is not on the graph

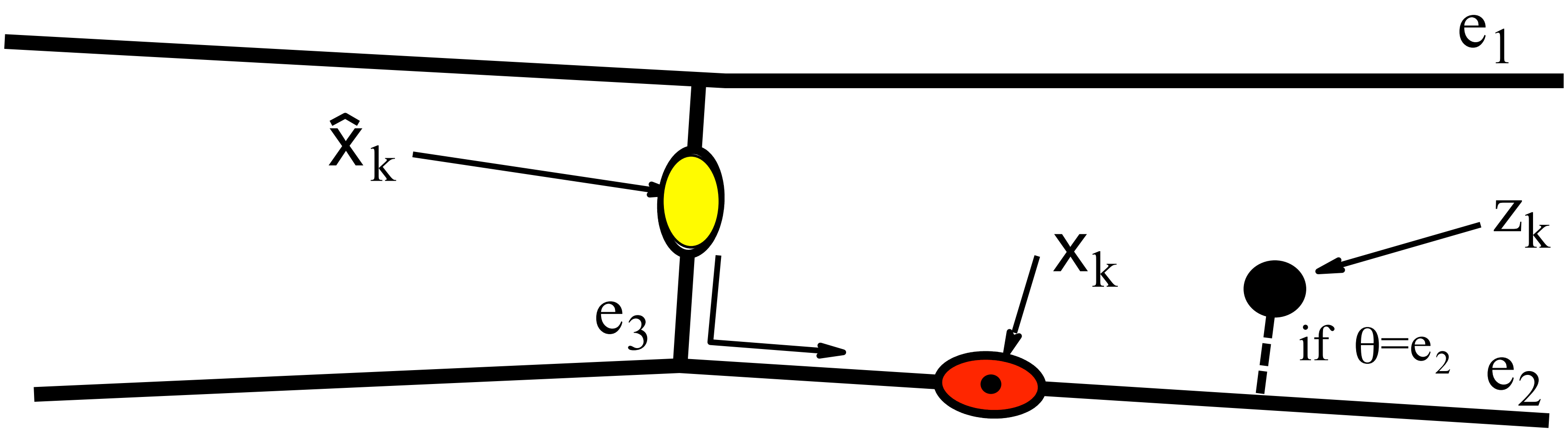
Kalman Filtering on a Graph: Correction Step



Problem: GPS reading is not on the graph

slide credit: D. Fox

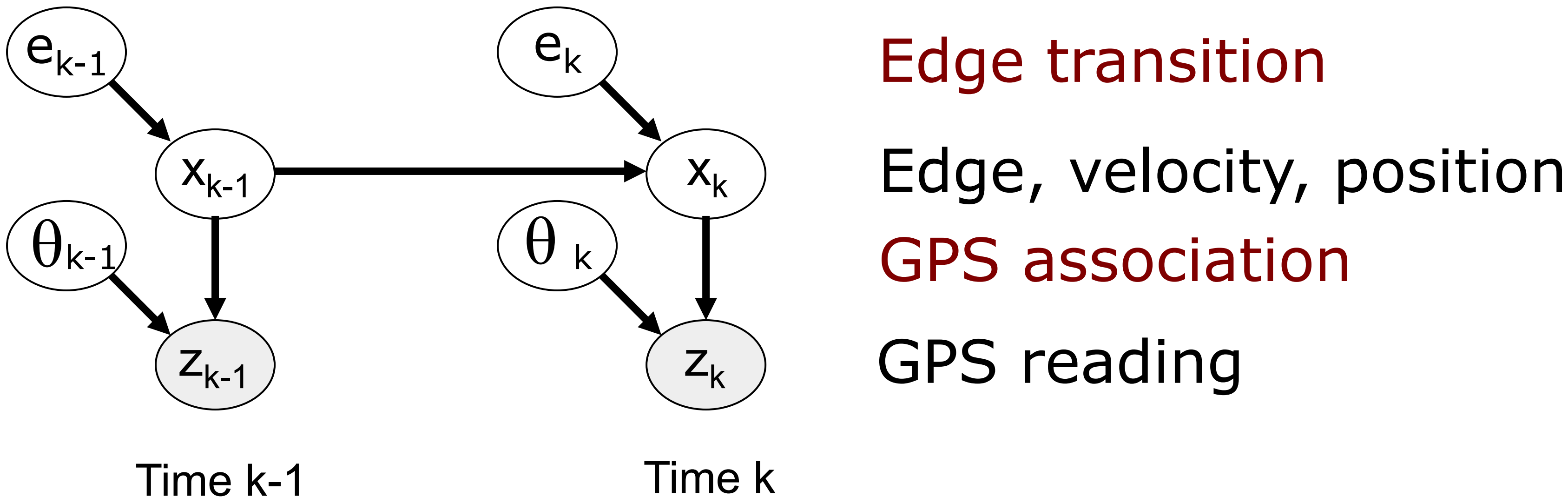
Kalman Filtering on a Graph: Correction Step



Problem: GPS reading is not on the graph

slide credit: D. Fox

Dynamic Bayesian Network



Task: Estimate the posterior over **all** hidden variables

slide credit: D. Fox

Rao-Blackwellized Particle Filtering (RBPF)

- Inference: Estimate the posterior given all past sensor measurements
- Particle filtering
  - Approximation of the posterior using samples
  - Supports multi-modal distributions
  - Supports discrete variables (e.g., transp. mode)
- Rao-Blackwellization
  - Sample some variables of the state space and solve the others analytically conditioned on sampled values

Factorization

$$p(\overset{\text{loc}}{l_k}, \overset{\text{vel}}{v_{1:k}}, \overset{\text{trans}}{e_{1:k}}, \overset{\text{assoc}}{\theta_{1:k}} \mid \overset{\text{obs}}{z_{1:k}})$$

Factorization

$$p(\overset{\text{loc}}{l_k}, \overset{\text{vel}}{v_{1:k}}, \overset{\text{trans}}{e_{1:k}}, \overset{\text{assoc}}{\theta_{1:k}} \mid \overset{\text{obs}}{z_{1:k}})$$

$$= \quad p(l_k \mid v_{1:k}, e_{1:k}, \theta_{1:k}, z_{1:k}) p(v_{1:k}, e_{1:k}, \theta_{1:k} \mid z_{1:k})$$

- Histories over the velocity, edge transition, and edge association, represented by samples in the PF

Factorization

$$
\begin{aligned}
p(\overset{\text{loc}}{l_k}, \overset{\text{vel}}{v_{1:k}}, \overset{\text{trans}}{e_{1:k}}, \overset{\text{assoc}}{\theta_{1:k}} \mid \overset{\text{obs}}{z_{1:k}}) \\
= \quad & p(l_k \mid v_{1:k}, e_{1:k}, \theta_{1:k}, z_{1:k}) p(v_{1:k}, e_{1:k}, \theta_{1:k} \mid z_{1:k})
\end{aligned}
$$

- Histories over the velocity, edge transition, and edge association, represented by samples in the PF
- Location of the person on the graph, estimated by a KF conditioned on samples

Rao-Blackwellized Particle Filter

- Represents the posterior by a set of *n* weighted particles and applies sampling

$$S_k = \{\langle s^{(i)}, w^{(i)} \rangle, i = 1, \cdots, n\}$$

- Here: Particles include distributions over variables, not just single samples

Rao-Blackwellized Particle Filter

- Represents the posterior by a set of *n* weighted particles and applies sampling

$$S_k = \{\langle s^{(i)}, w^{(i)} \rangle, i = 1, \cdots, n\}$$

- Here: Particles include distributions over variables, not just single samples

- Each particle of the RBPF has the form

$$s^{(i)} = \langle \underbrace{e^{(i)}, v^{(i)}, \theta^{(i)}}_{\text{sampled values:}}, \underbrace{\mathcal{N}^{(i)}(\mu, \sigma^2)}_{\text{KF for the location}} \rangle$$

sampled values:
- edge transitions
- velocities
- edge associations

Sampling Step

- Sample the velocity $v^{(i)}$ from a mixture of Gaussians, which is conditioned on the transportation mode (described later on)

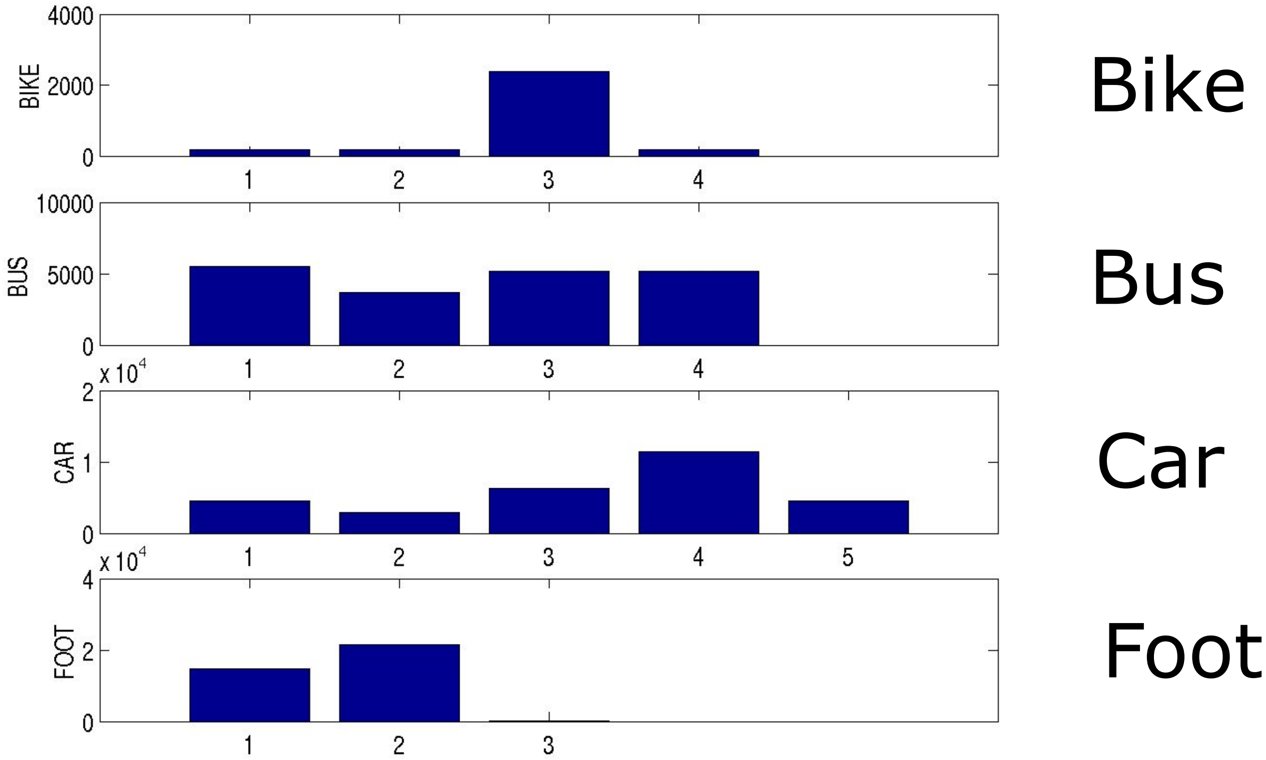


image source: H. Kautz

Sampling Step

- Sample the velocity $v^{(i)}$ from a mixture of Gaussians, which is conditioned on the transportation mode (described later on)
- Sample the edge transition $e^{(i)}$ based on the previous position of the person and a learned transition model
- Sample the edge association $\theta^{(i)}$ based on the distance between z_k and the streets in the vicinity

Kalman Filter

- Update of the position estimate based on the sampled values and the measurement
- Prediction:
  - Use sampled velocity to predict traveled distance
  - Use sampled edge transition if predicted mean transits over a vertex
- Correction:
  - Find shortest path between the prediction and the “snapped” measurement
  - Apply a 1-dimensional Kalman filtering correction step

Prediction Step

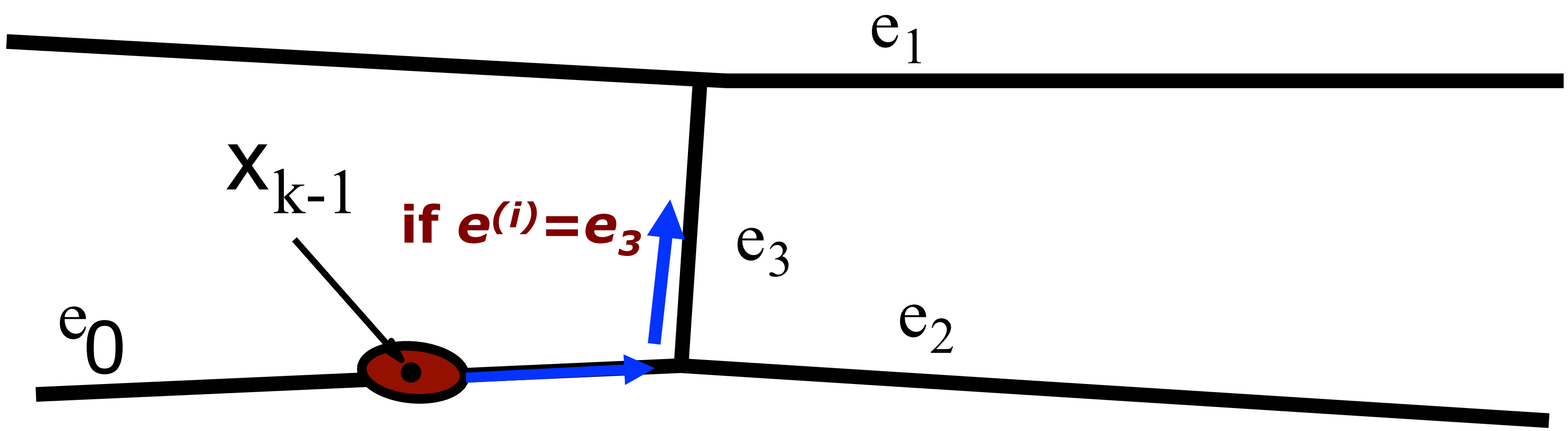
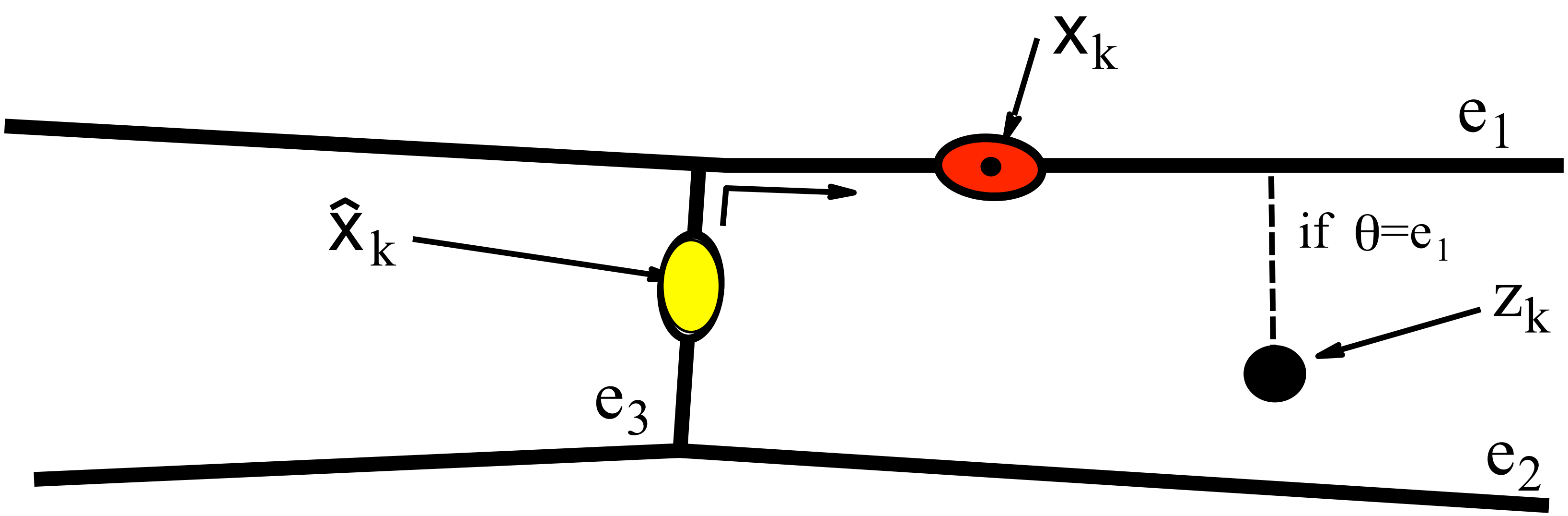


image source: D. Fox

Correction Step



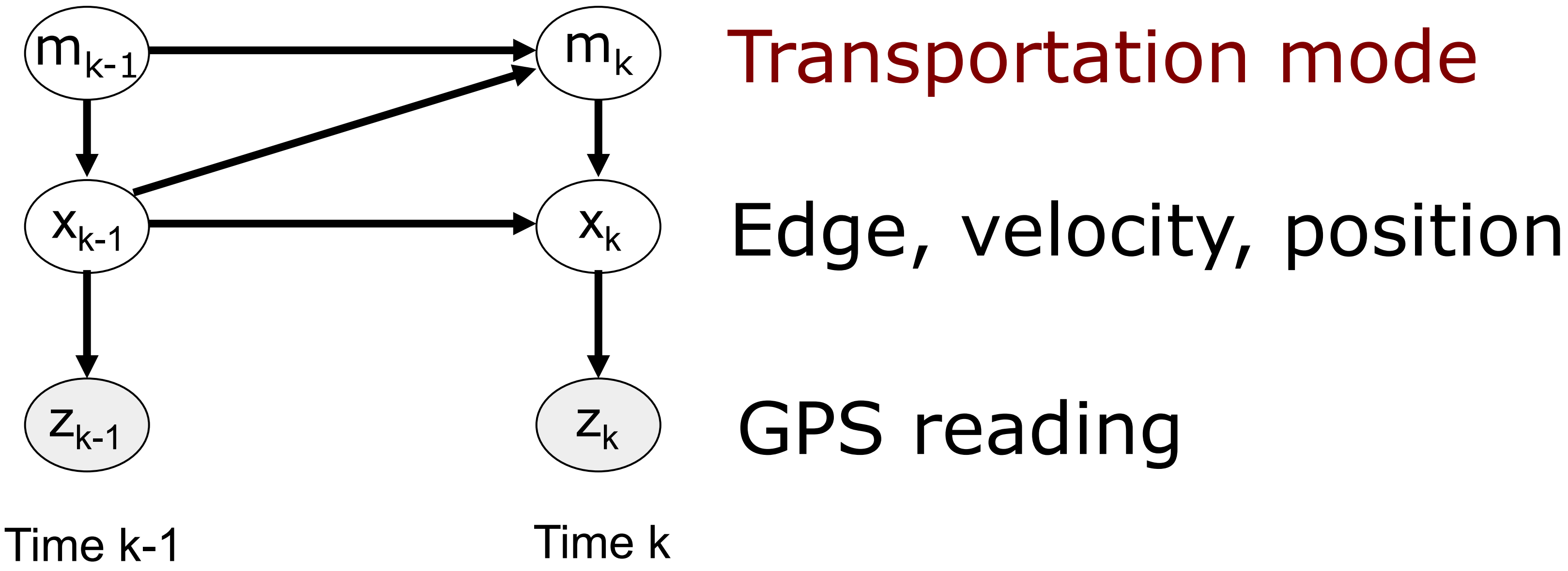
Depending on the edge association, the correction step moves the estimate up or downwards

image source: D. Fox

Mode of Transportation / Prior Knowledge

- Transportation modes have different velocity models
- Buses run on bus routes (corresponding to edge transitions)
- Get on/off the bus near bus stops
- Switch to car near car location

Dynamic Bayesian Network



$$s^{(i)} = \langle m^{(i)}, e^{(i)}, v^{(i)}, \theta^{(i)}, \mathcal{N}^{(i)}(\mu, \sigma^2) \rangle$$

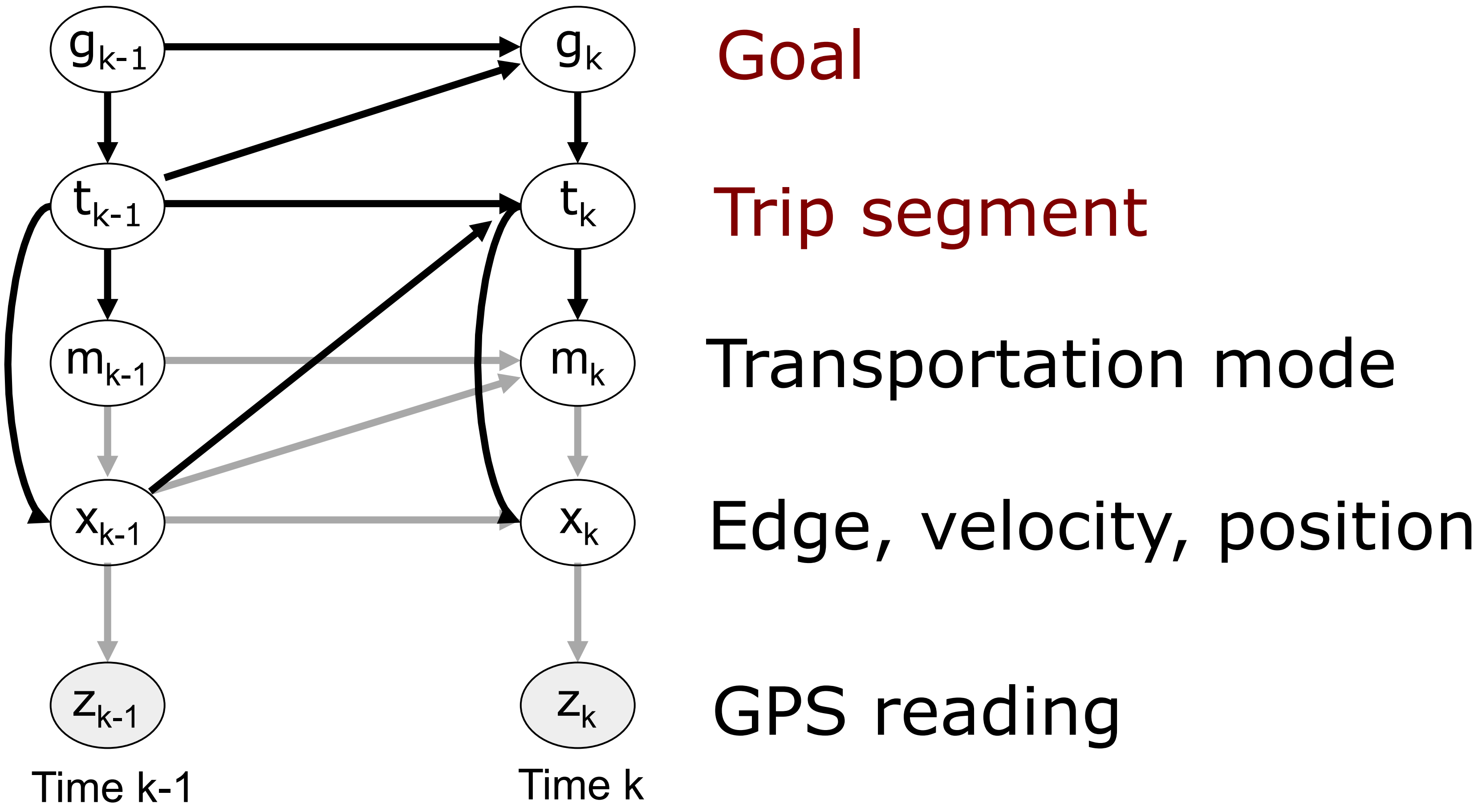
slide credit: D. Fox

Transportation Routines

- Goal (destination):
  - Workplace (could also be friends, restaurant, ...)
- Trip segments: <start, end, transportation>
  - Home to Bus stop A on Foot
  - Bus stop A to Bus stop B on Bus
  - Bus stop B to workplace on Foot

slide credit: D. Fox

Hierarchical Model



$$s^{(i)} = \langle \langle g, t \rangle^{(i)}, m^{(i)}, e^{(i)}, v^{(i)}, \theta^{(i)}, \mathcal{N}^{(i)}(\mu, \sigma^2) \rangle$$

slide credit: D. Fox

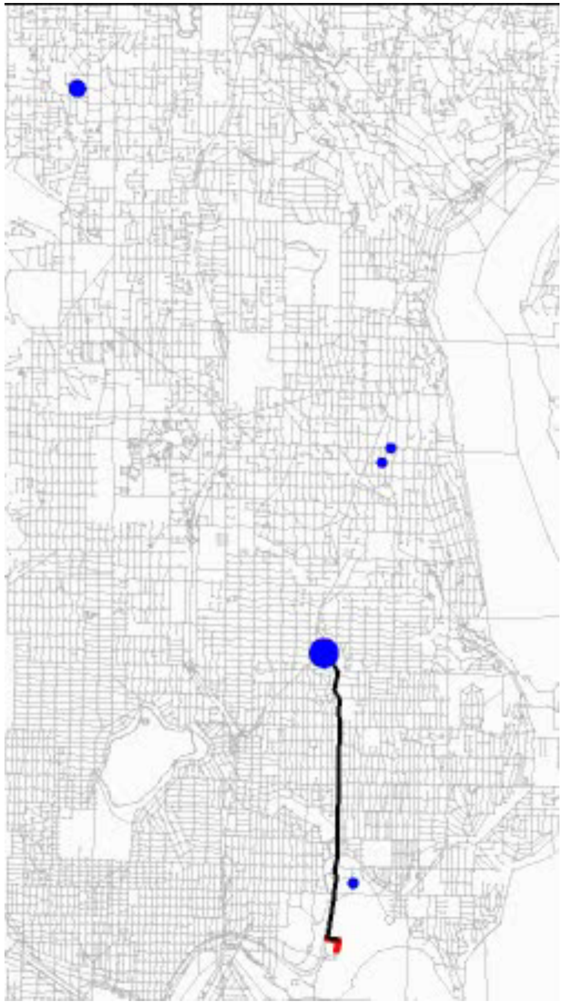
Remarks

- Note the hierarchical structure
- RBPF first samples the goal and trip segment
- Low-level model (w/o goal and trip segment) samples the edge transition solely based on the location and the transp. mode
- Hierarchical model takes the current trip segment into account
- Edge transition probabilities depend on trip segments, which leads to improved predictive capabilities

Learning the DBN Parameters

- Learn variable domains
  - Goals: Locations where the user stays for long time
  - Transition points: Locations with high transportation mode switching probability
  - Trip segments: Connect transition points and goals
- Learn transition matrices for goals, trip segments, and edges via EM
- Unlabeled data: 30 days of one user, logged at 2 second intervals

Prediction of Goal and Path



Predicted goal

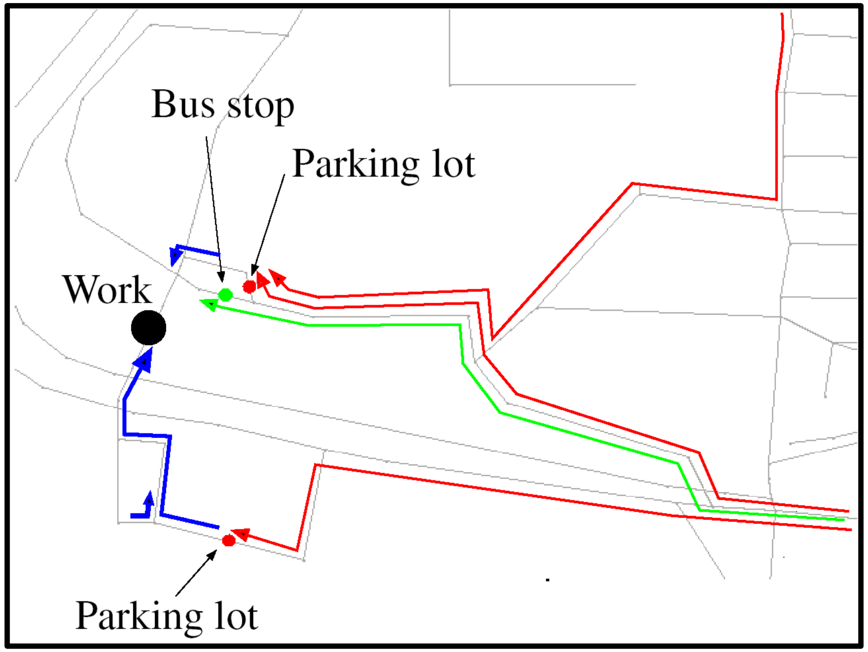
Predicted path

Correct goal and route predicted
100 blocks away

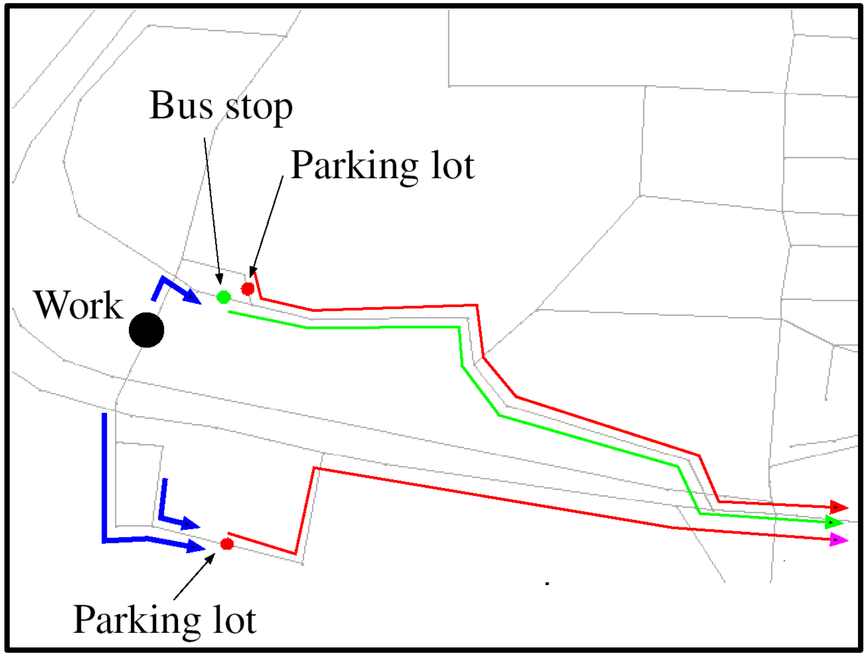
animation: D. Fox

Learned Transition Probabilities

Going to the workplace



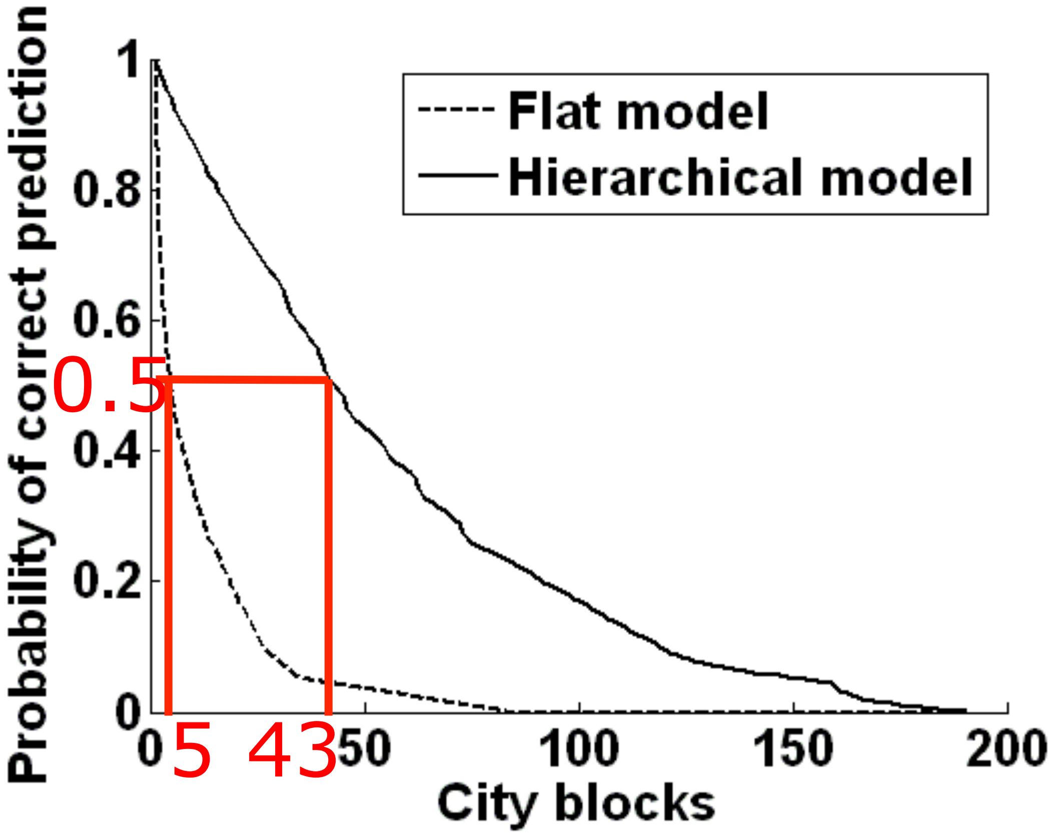
Going home



High probability transitions: **bus** **car** **foot**

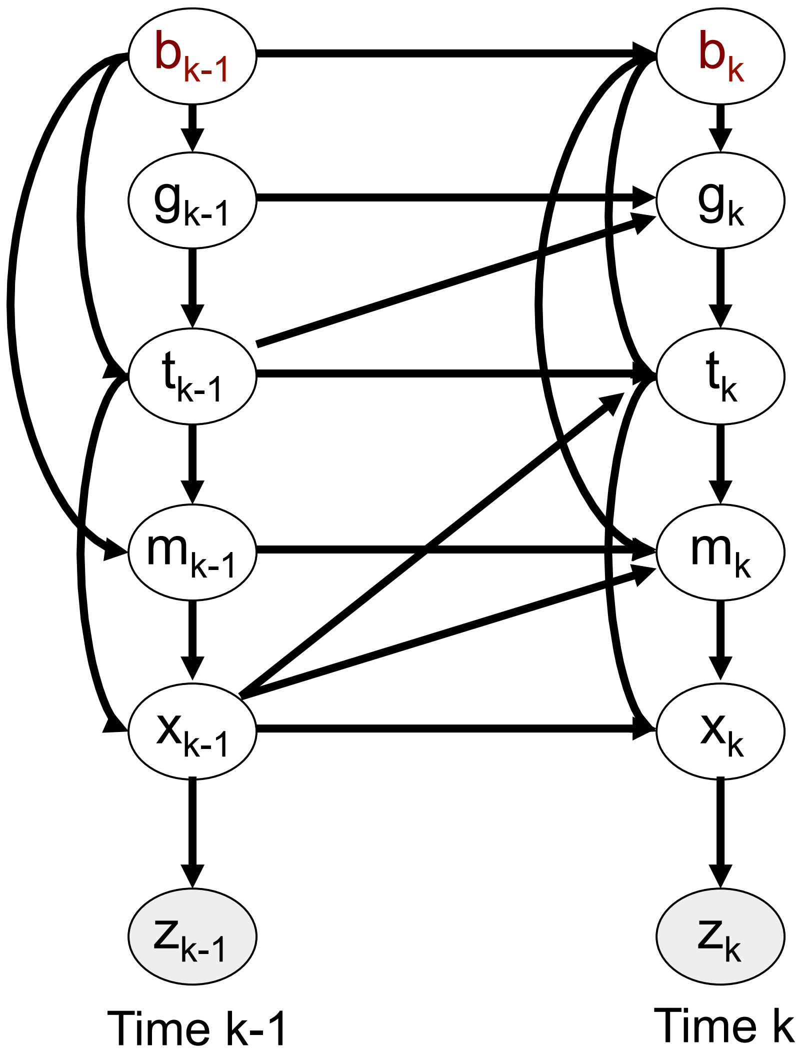
slide credit: D. Fox

Prediction Capabilities



slide credit: D. Fox

Detecting Deviations



Behavior mode
normal / unknown

Goal

Trip segment

Transportation mode

Edge, velocity, position

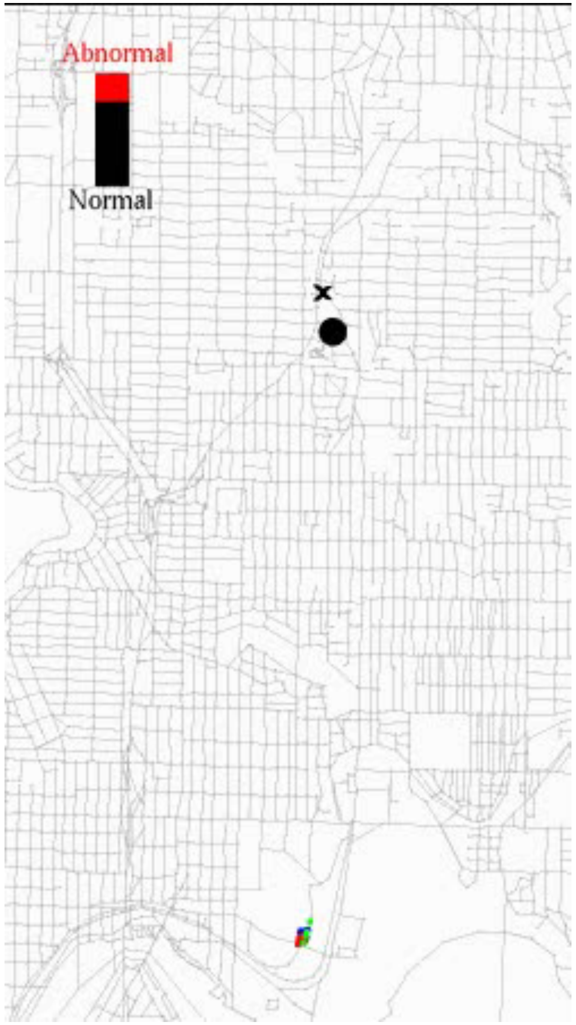
GPS reading

slide credit: D. Fox

Detecting Novel Behavior

- RBPF: Sample novelty variable
- Depending on the sampled value use
  - Hierarchical model as trained for the user
  - Untrained, flat model (no user-specific preferences for motion directions or transportation modes)

Detecting User Errors



● Predicted goal

x Predicted bus stop

Missing the bus stop

animation: D. Fox

Application: Cognitive Aid

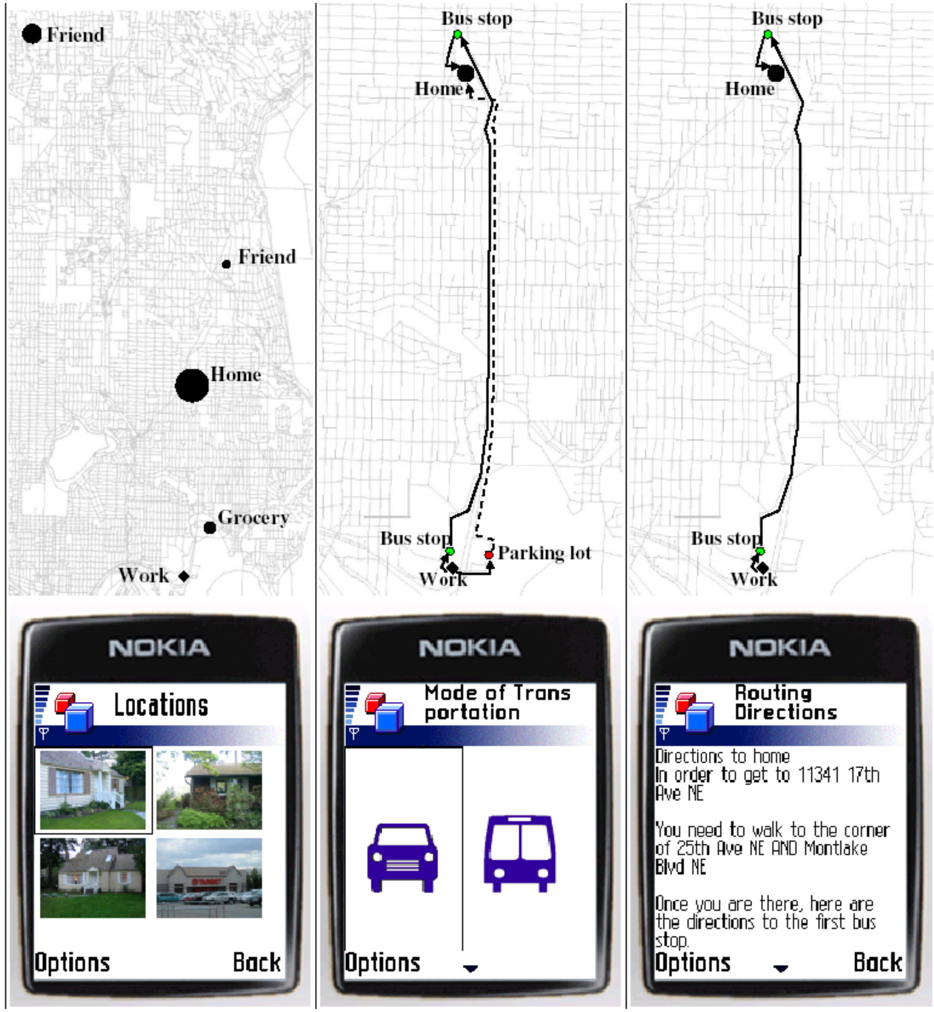


image source: D. Fox

Application: Cognitive Aid

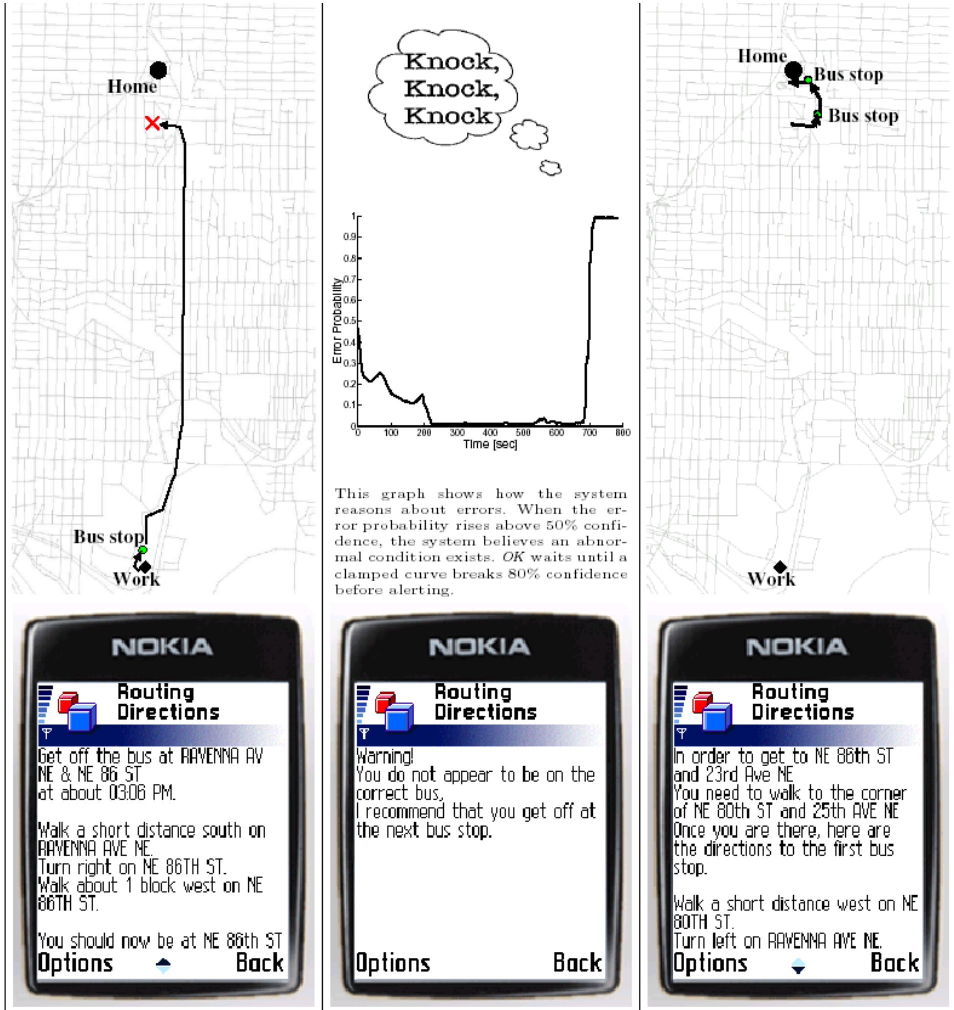


image source: D. Fox

Inferring Significant Places and Activities

- So far
  - No distinction between different types of goals
  - Fixed thresholds for the duration to extract goals and transition mode transfer locations
- However, both can have a significant influence on the inference quality
- Idea: Simultaneous identification and labeling of significant locations and estimation of activity

Give Semantic Meaning to Places

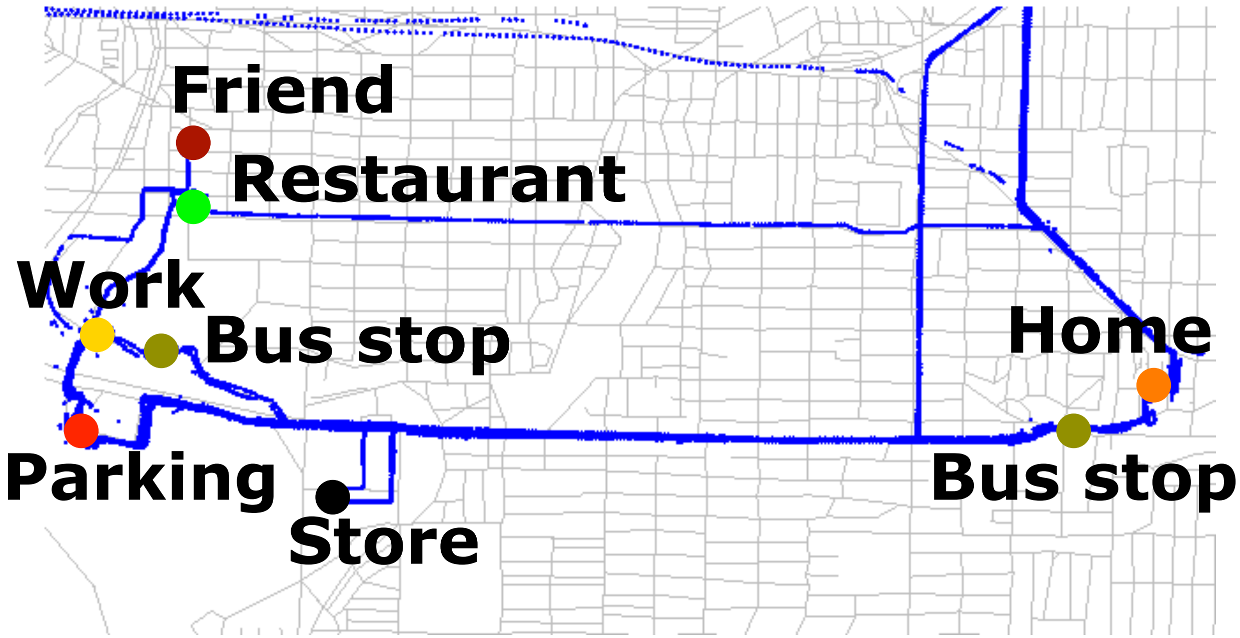
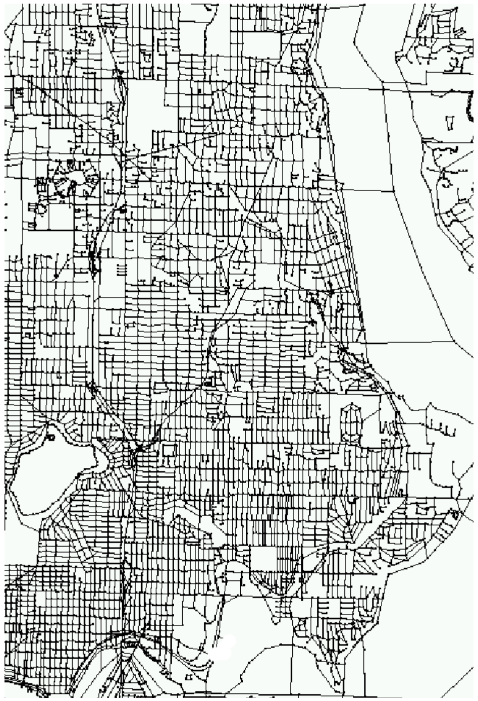
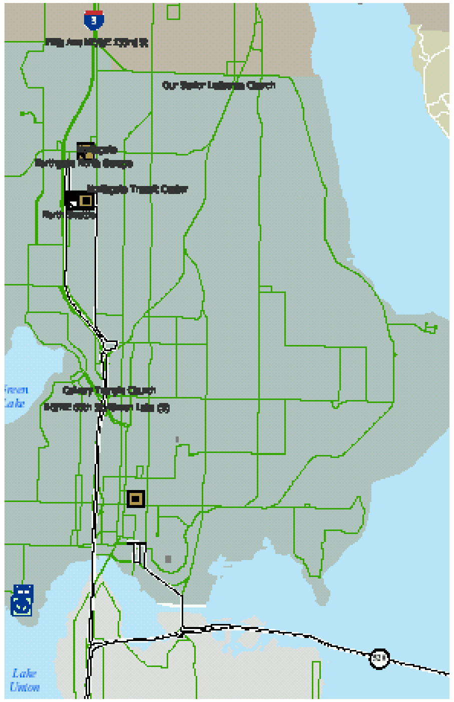


image source: D. Fox

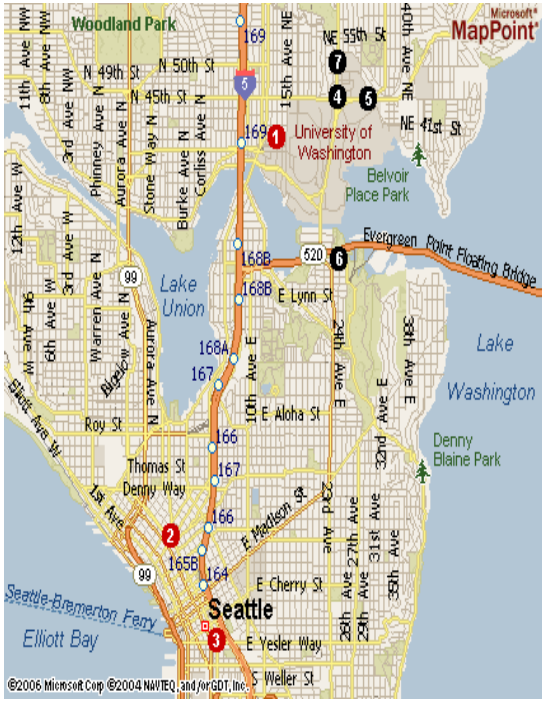
Geographic Information Systems



Street map



Bus routes / bus stops



Restaurants / Stores

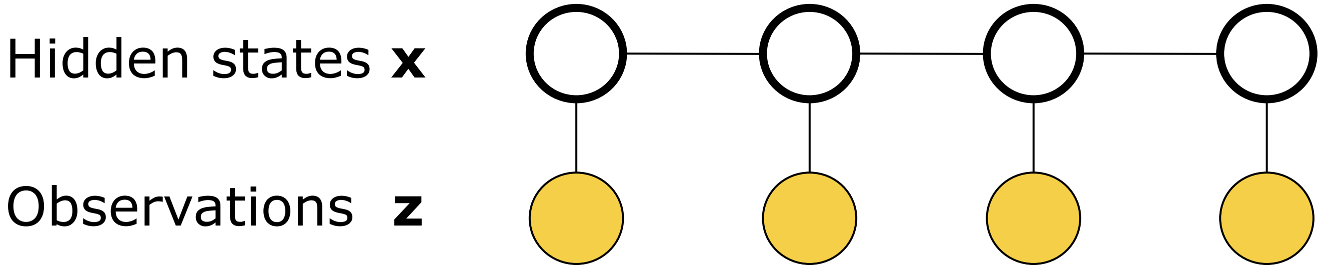
Activity Inference

- For each location (10m patch) infer the person’s activity (e.g., bus, foot, work, visit)
- Use information such as
  - Temporal pattern: duration, time of day, etc.
  - Geographic features: restaurant / store / bus stop nearby
  - Activities of neighbor cells
- Additionally consider number of occurances of labels (e.g., home, workplace; summation constraints)

Conditional Random Fields (CRF)

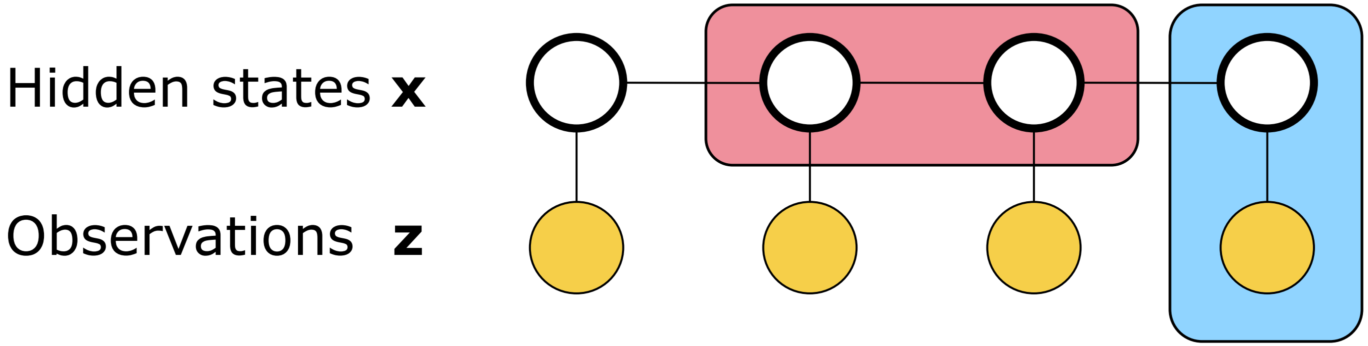
- CRF are undirected graphical models
- Developed for labeling data sequences
- Do not assume independence between the observations
- Relationships between labels of states are considered and the labeling is done simultaneously
- CRF model the distribution p(**x** | **z**)
- Hidden states **x** = activities
- Observations **z** = features

Conditional Random Fields



slide adapted from: D. Fox

Conditional Random Fields



Clique potentials Φ_c measure the "compatibility" among the variables in a clique *c*

Local potentials link states to observations

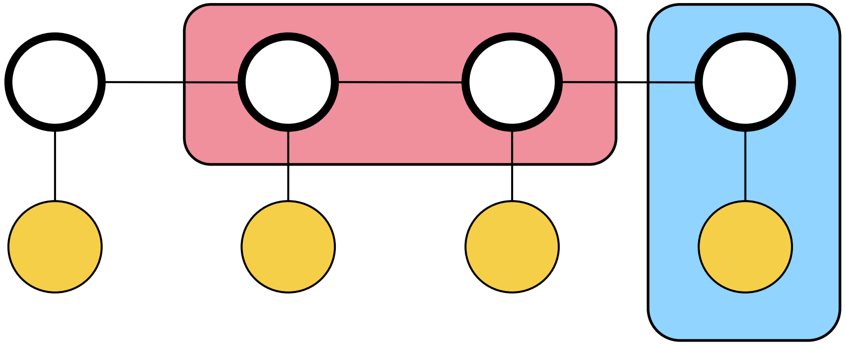
Neighborhood potentials link states to neighboring states

slide adapted from: D. Fox

Conditional Random Fields

Hidden states **x**

Observations **z**



$$p(\mathbf{x} \mid \mathbf{z}) = \frac{1}{Z(\mathbf{z})} \prod_{c \in C} \Phi_c(\mathbf{x}_c, \mathbf{z}_c) = \frac{1}{\boxed{Z(\mathbf{z})}} \exp\left\{ \sum_{c \in C} \boxed{\mathbf{w}_c^{\mathrm{T}}} \boxed{\mathrm{f}_c(\mathbf{x}_c, \mathbf{z}_c)} \right\}$$

Normalizing partition function

Weights

Feature functions

Local potentials link states to observations

Neighborhood potentials link states to neighboring states

slide adapted from: D. Fox

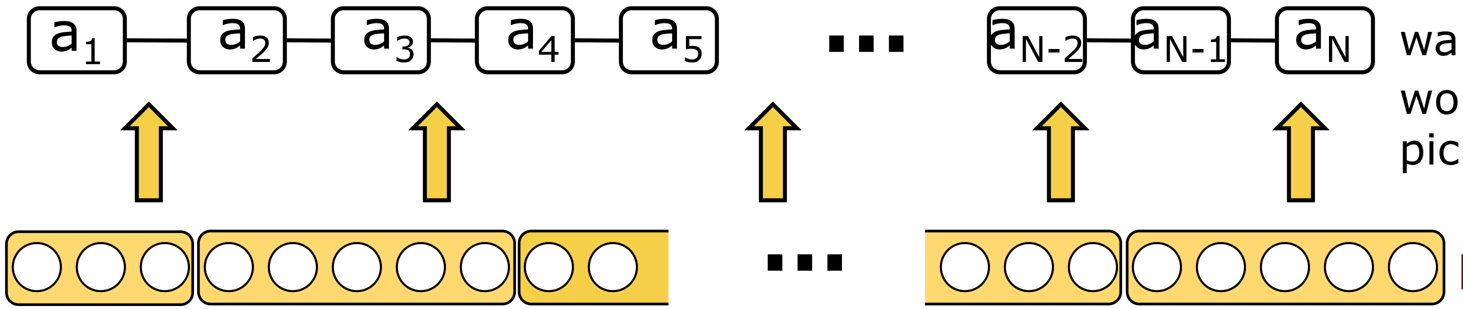
Feature Functions

- Typically designed by the user
- Extract a vector of features from variable values
- Weights represent importance of different features for correctly inferring the hidden states
- Weights are learned from labeled training data
- Approximation of the conditional distribution parameterized via the weights $p(\mathbf{x} \mid \mathbf{z}, \mathbf{w})$

Features for Place Labeling

- Temporal information: time of day / week, duration (binary indicator function)
- Average velocity (binary indicator)
- Geographic information: bus stop / restaurant / shop nearby (binary indicator)
- Transition relation: Adjacent activities (e.g., driving the car after taking the bus rather unlikely)
- Spatial context: Relation between place and activity (count + binary indicator for each combination of place, activity, frequency)
- Summation constraints: Number of places labeled home / workplace (count features)

Hierarchical CRF Model

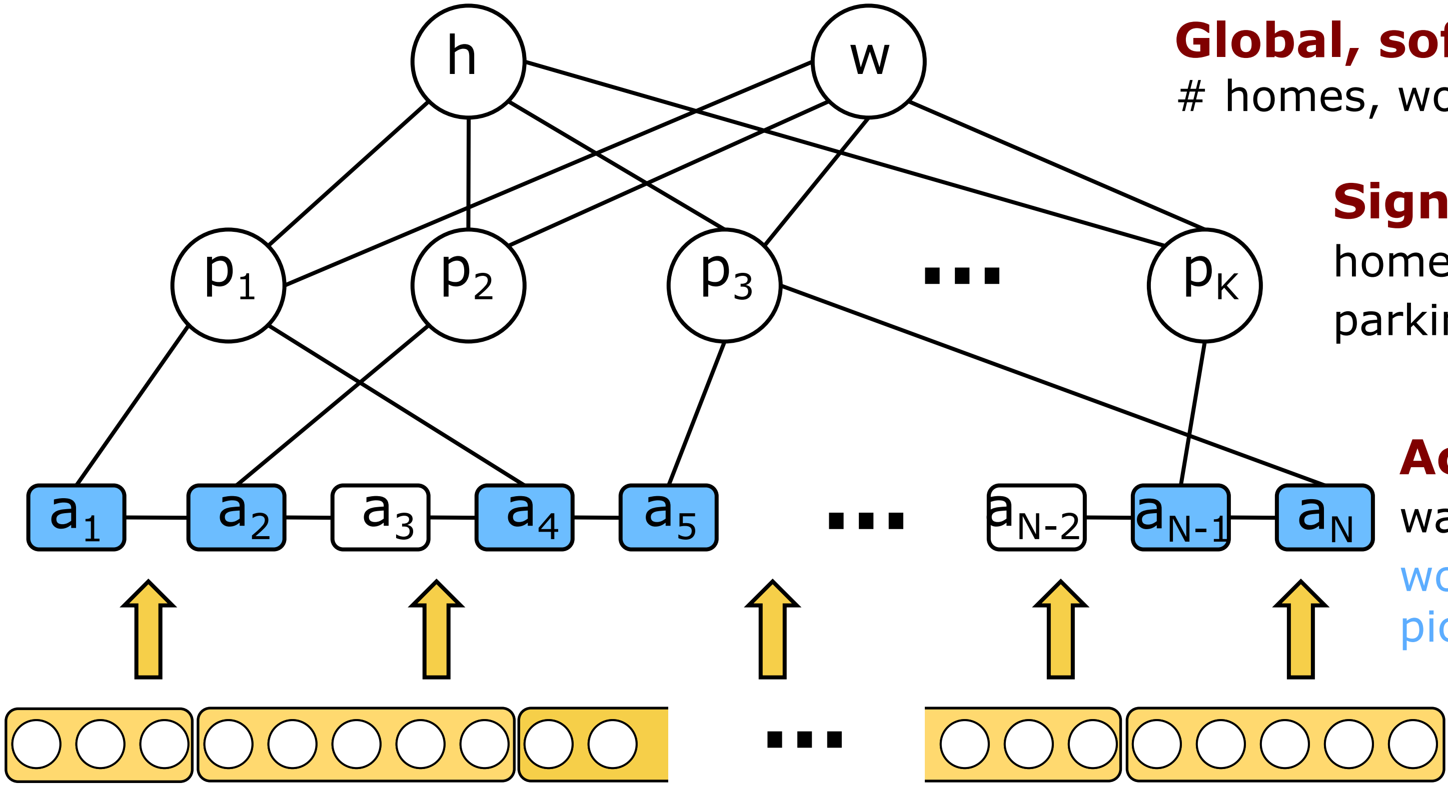


Activity sequence
walk, drive, ride bus,
work, visit, sleep,
pickup, get on/off bus

Local evidence
time, duration,
velocity, geographic
information

slide adapted from: D. Fox

Hierarchical CRF Model



Global, soft constraints

homes, workplaces

Significant places

home, work, bus stop,
parking lot, friend's home

Activity sequence

walk, drive, ride bus,
work, visit, sleep,
pickup, get on/off bus

Local evidence

time, duration,
velocity, geographic
information

slide adapted from: D. Fox

Experimental Results

- GPS data from 4 different persons / 7 days
- 40,000 GPS measurements / 10,000 activity segments
- Manually labeled activities and places
- Leave-one-out cross validation
- Maximum pseudo-likelihood for learning (1 minute to converge)
- Inference via loopy belief propagation (activities and places from 1 week within 1 minute)

Example: Raw GPS Data

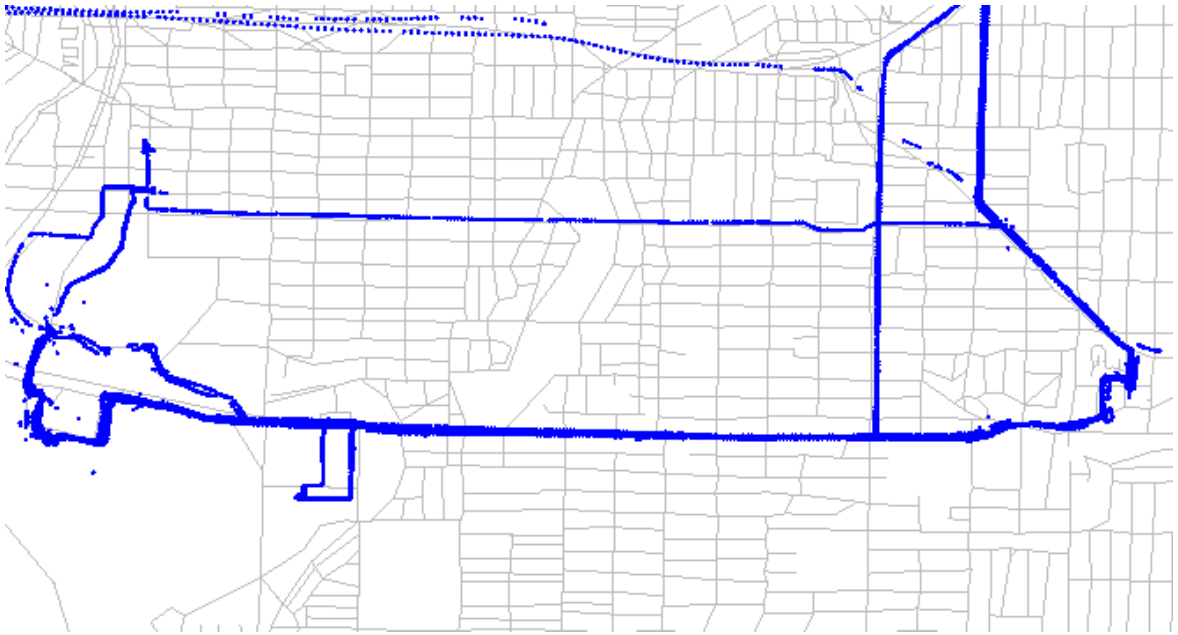


image from: D. Fox

Activities for Each Patch

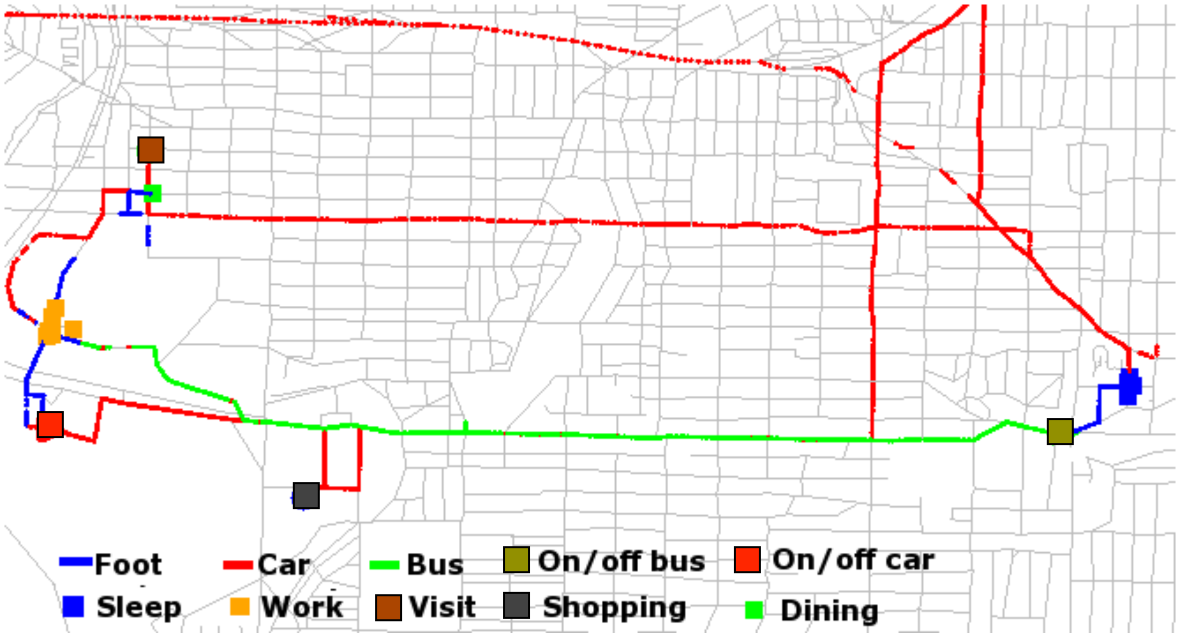


image from: D. Fox

Places by Clustering Significant Activities

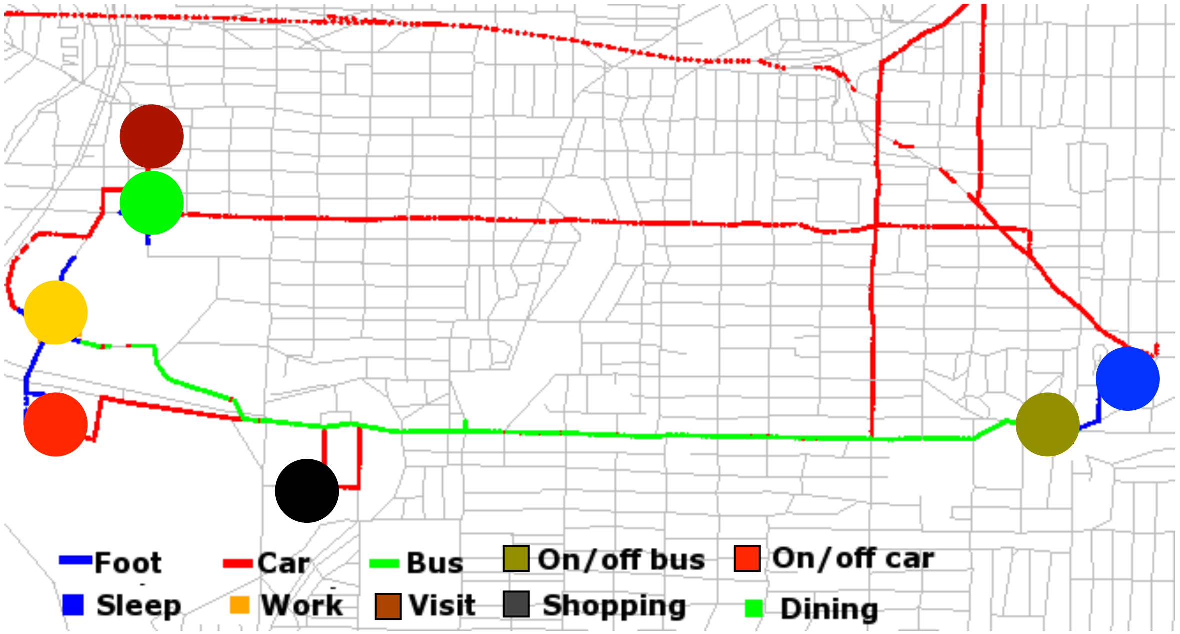
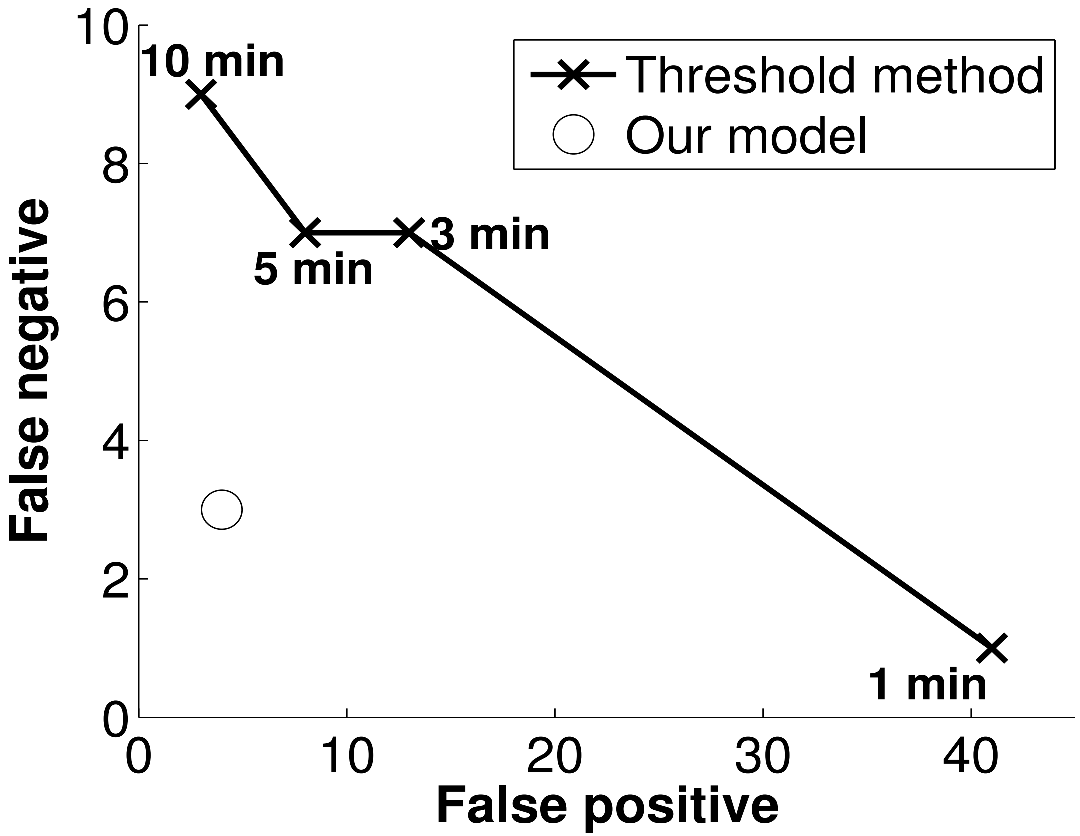


image from: D. Fox

Improved Place Finding



- New model clearly outperforms the threshold method

Summary of a Day

Time	Activity and transportation
8:15am - 8:34am	Drive from home 1 to parking lot 2, walk to workplace 1;
8:34am - 5:44pm	Work at workplace 1;
5:44pm - 6:54pm	Walk from workplace 1 to parking lot 2, drive to friend's place 3;
6:54pm - 6:56pm	Pick up/drop off at friend 3's place;
6:56pm - 7:15pm	Drive from friend 3's place to other place 5;
9:01pm - 9:20pm	Drive from other place 5 to friend 3's place;
9:20pm - 9:21pm	Pick up/drop off at friend 3's place;
9:21pm - 9:50pm	Drive from friend 3's place to home 1;
9:50pm - 8:22am	Sleep at home 1.

- Most likely sequence of activities and places

Summary

- Location-based activity recognition is possible
- Graph-based representations are well suited to compactly represent and learn typical behavior
- Hierarchical graphical models (DBN, CRF) powerful tools for bridging the gap between continuous sensor data, low-level activities, and abstract states
- Conditional Random Fields can handle high-dimensional / dependent feature vectors

Further Reading

- L. Liao, D. Fox, H. Kautz

  Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields
  Int. Journal of Robotics Research, 2007

- L. Liao, D. J. Patterson, D. Fox, H. Kautz

  Learning and Inferring Transportation Routines
  Journal Artificial Intelligence, 2007