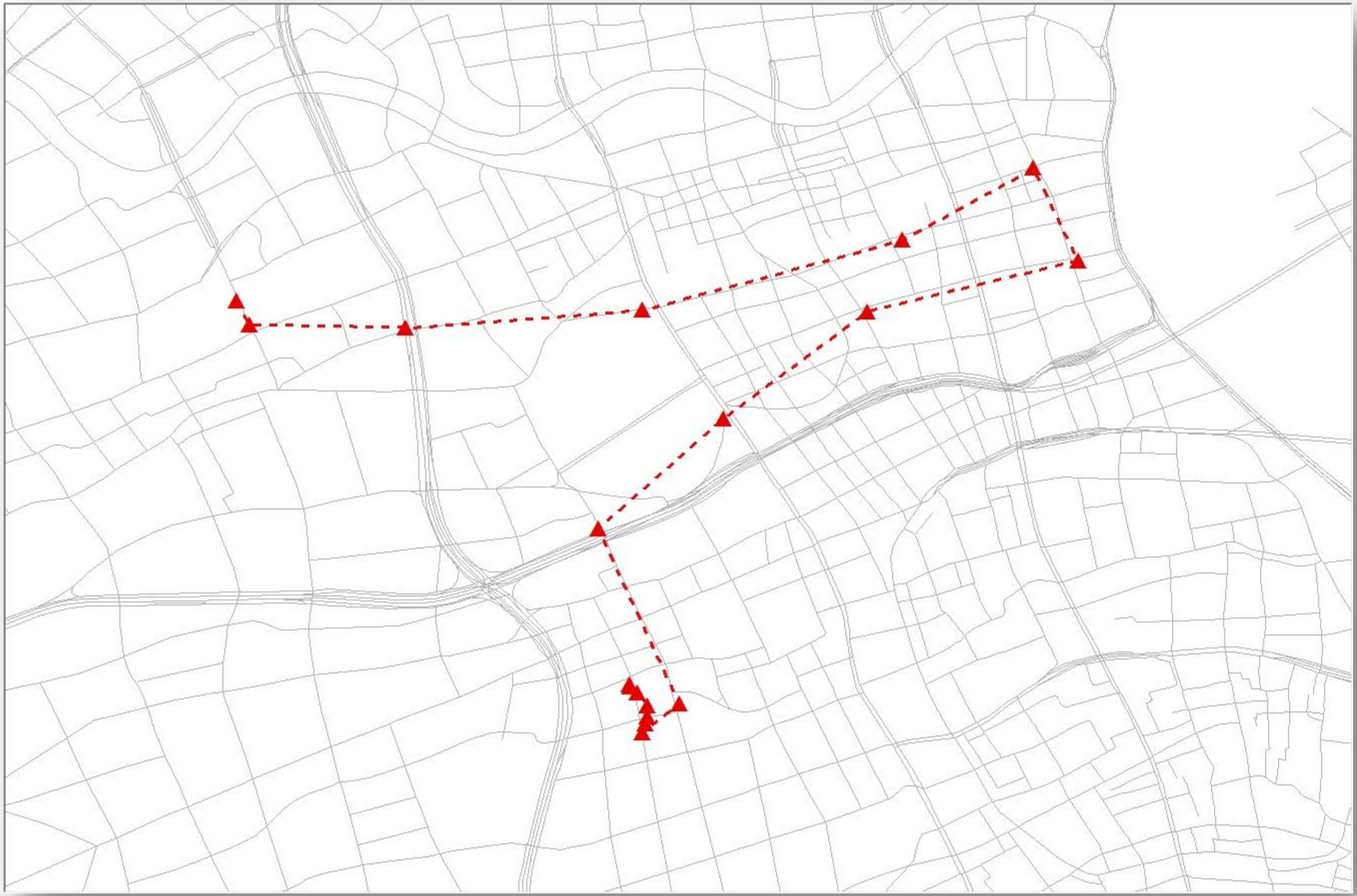
Feature Selection in CRFs for Map Matching of GPS Trajectories

Jian Yang, Liqiu Meng

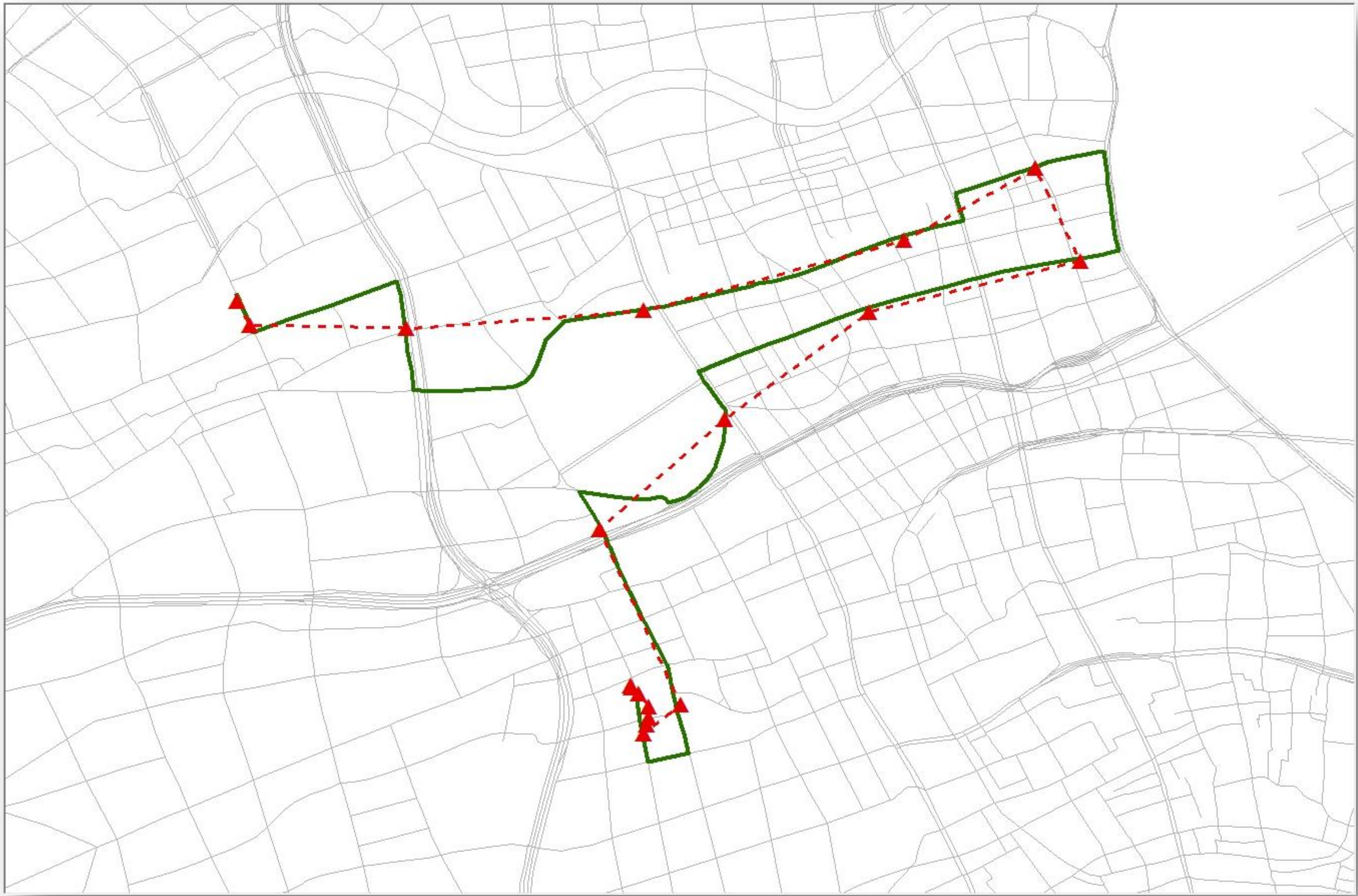
Lehrstuhl für Kartographie
Technische Universtität München
Munich, Germany


Credit: Google
Driving Assistance
Credit: Google
Traffic Management
Urban Mobility Pattern
(Zheng, 2008)
Map Matching
TAXI
TOMTOM

GPS pts(red) with **120** seconds sampling rate



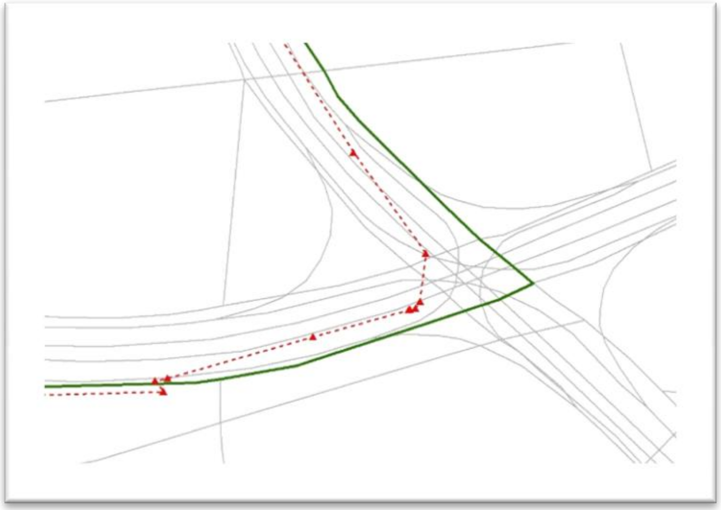
A GPS trajectory (red) with **120** seconds sampling rate.



A GPS trajectory (red) with **120** seconds sampling rate and the ground truth (green) in road network.

Two subtasks of Map Matching

1) Localize individual GPS pts



Nearest roads

2) Path between GPS pts

Shortest, Fastest

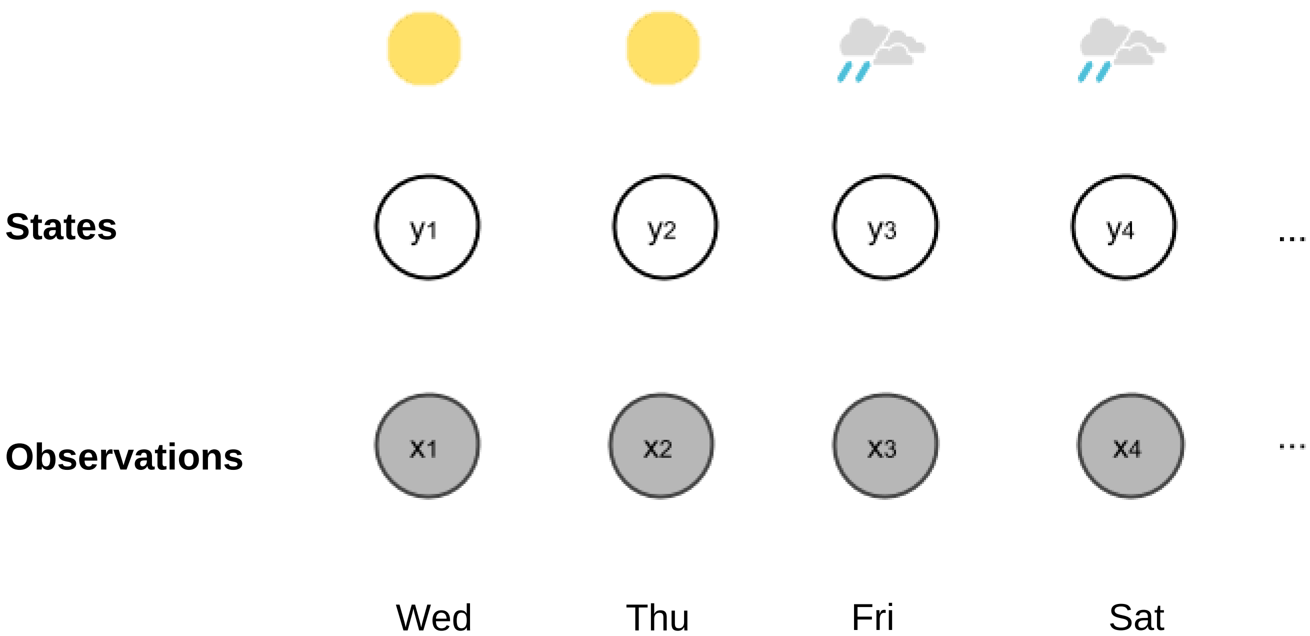
Fewest turns

Map Matching Begins…

Can we combine the modeling efforts (observation & transition)?
&
Possible to identify most relevant ones?

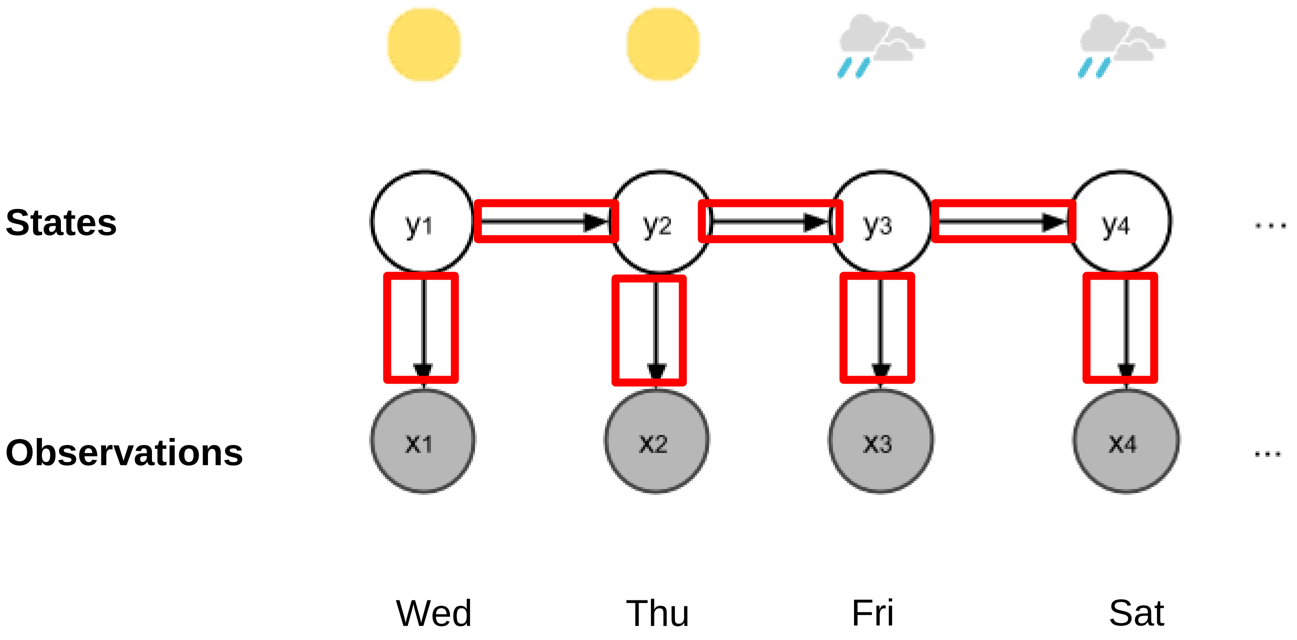
literature	summary
Hummel 06	probabilistic of the GPS sequence, HMM
Krumm 07	HMM with travel time constraint
Lou 09	Low-sampling-rate, ST-analysis (HMM alike)
Newson 09	HMM with geometric transition probability
ACM GIS CUP 12	probabilistic and HMM are the top 5 solutions
Bierlaire 13	path set generation algorithm
Chen 13	Multi-model, smart phone with Bluetooth
Hunter 13	CRFs with small feature set

What’s the weather for the next few days in Wien?

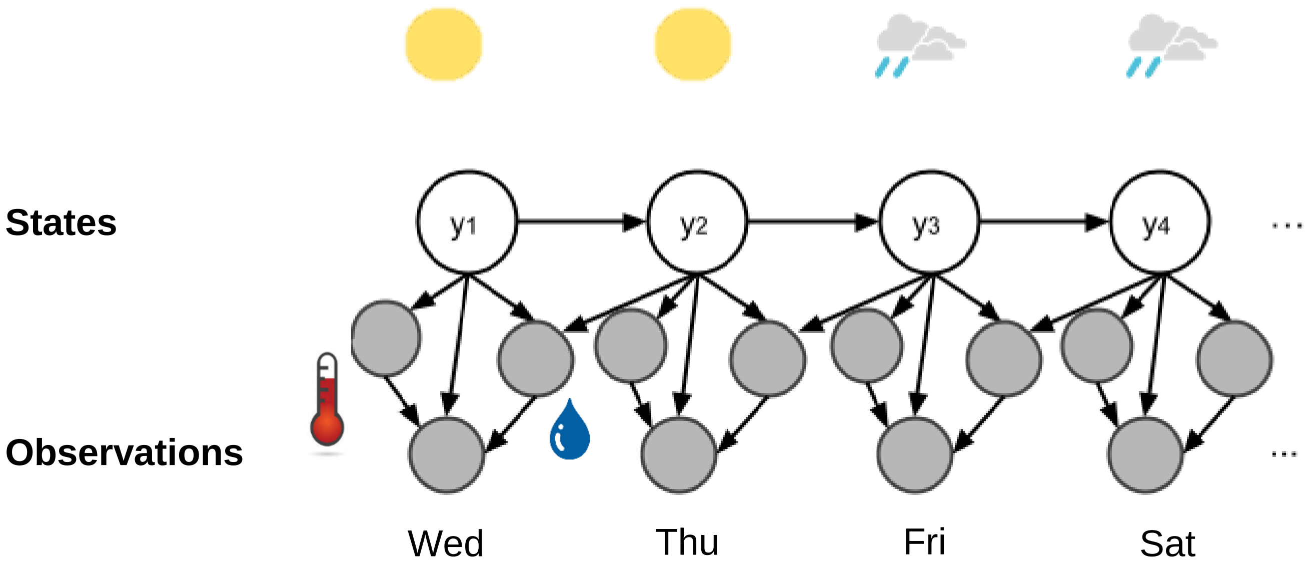


HMM – two assumptions

$$p(\mathbf{x}, \mathbf{y}) = p(x_1, \ldots, x_T, y_1, \ldots, y_T) = \prod_{t=1}^{T} p(x_t | y_t) p(y_t | y_{t-1})$$

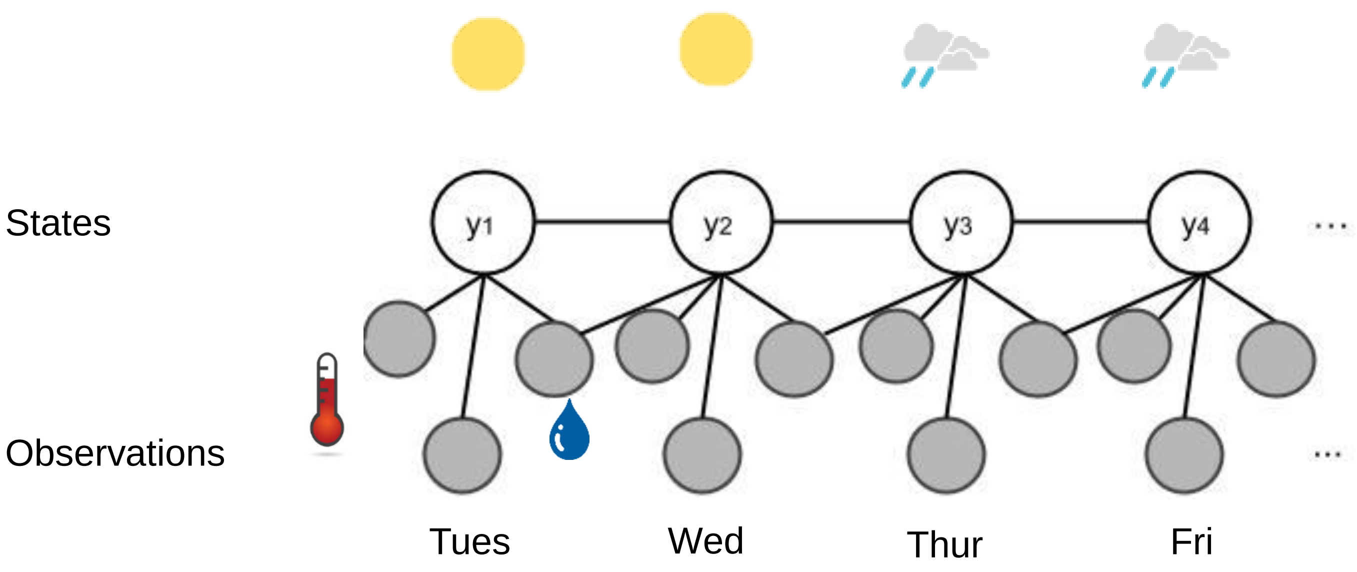


HMM(ctd.)

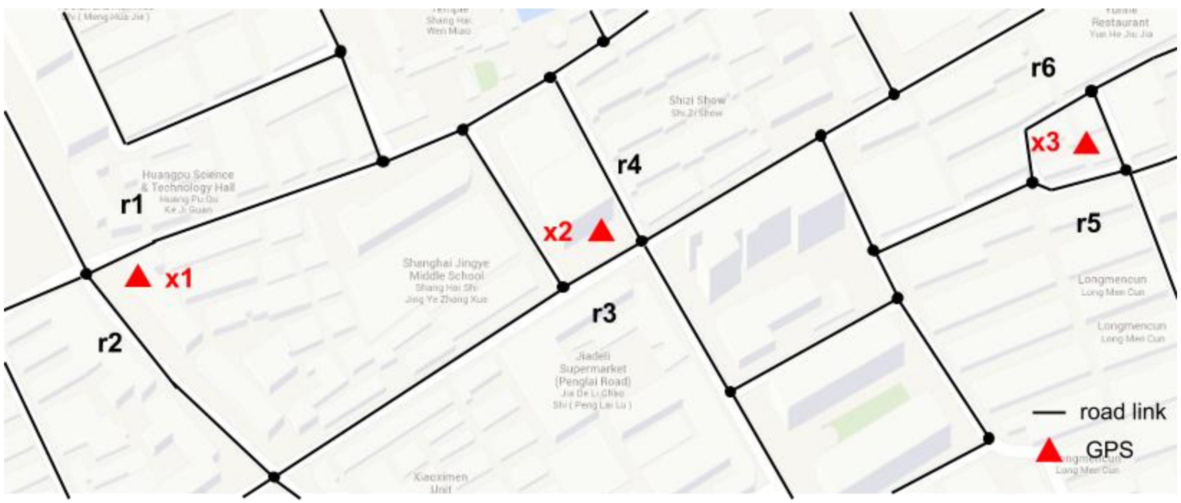


Conditional Random Fields (CRFs)

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \prod_{t=1}^{T} \exp\left(\sum_{k} \omega_k f_k(y_{t-1}, y_t, \mathbf{x})\right)$$

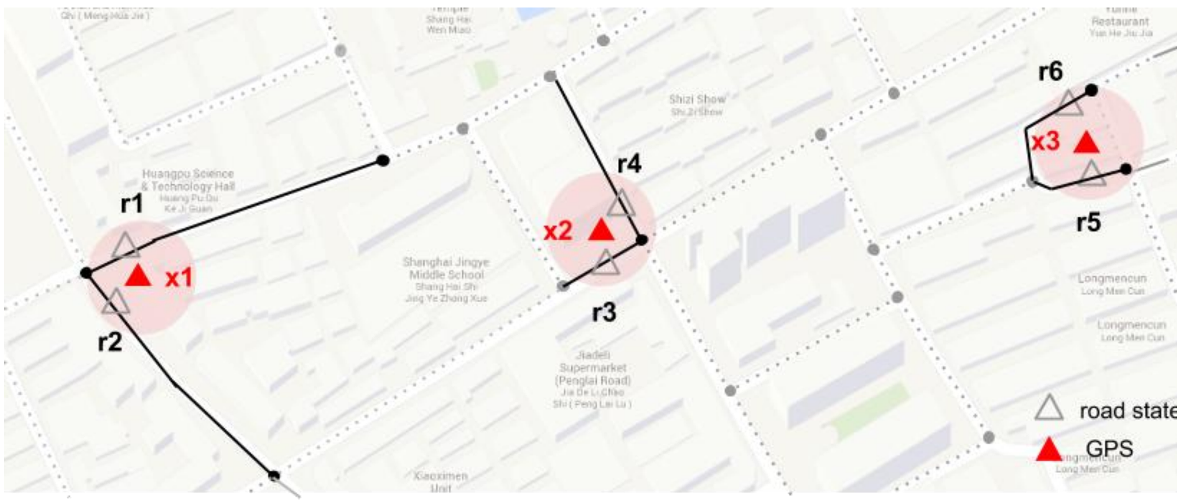


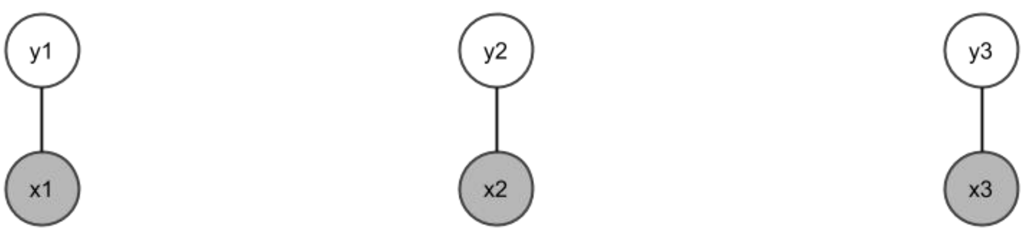
Modeling GPS trajectory using CRFs



Point nodes

$$p(y_t, x_t) = exp(\omega f)$$





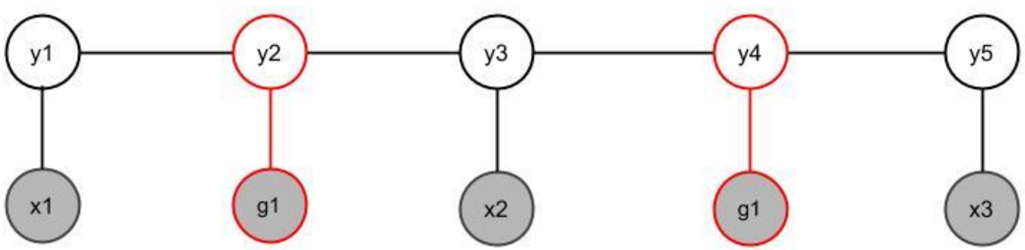
point

point

point

Path nodes

$$p(y_t, x_t){\color{red}p(y_{t-1}, y_t, y_{t+1})} = exp(\omega f){\color{red}\exp(\mu g)}$$

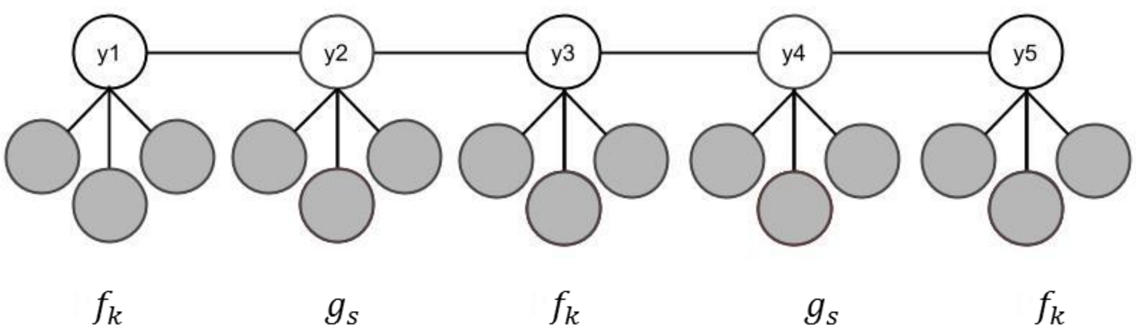


path

path

A Chain Structured CRFs for Map Matching

$$P(Y|X) = \frac{1}{Z}\prod_{t=1}^{N} exp(\sum_{k} \omega_k f_k(y_{2t-1}, x_t) + \sum_{s} \mu_s g_s(y_{2t}, y_{2t-1}, y_{2t+1}, X))$$



$f_k =$ *err_dist,*
sqr(err_dist),
bearing_err,
cos(bearing_err),
abs(cos(bearing_err)),
accu_filter(bearing_err),
…

$g_s =$ *Leng_difference,*
max_avg_speed,
min_avg_travel_time,
#left_turn,
#right_turn,
highest_road_class,
lowest_road_class,
change_road_class,
#sharp_turns,
#sharp_turn_left,
#sharp_turn_right
…

Map Matching as Inference

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \prod_{t=1}^{N} \exp(\sum_k \omega_k f_k(y_{2t-1}, x_t) + \sum_s \mu_s g_s(y_{2t}, y_{2t-1}, y_{2t+1}, \mathbf{X}))$$

Denoted as

$$p(\mathbf{y}|\mathbf{x}, \theta), \qquad \theta = (\omega_1, \ldots, \mu_1, \ldots)$$

Map matching can be cast to solve:

$$\arg\max_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}, \theta)$$

With a chain structure, it can be efficiently solve using dynamic programming, e.g. **Viterbi**

Parameter estimation and feature selection

$$p(\mathbf{y}|\mathbf{x}, \theta), \qquad \theta = (\omega_1, \ldots, \mu_1, \ldots)$$

θ Can be estimated by maximizing the ***log-likelihood*** given a set of training examples

$$\arg\max_{\theta} \log p(\mathbf{y}|\mathbf{x}, \theta)$$

A common model would use ***L2 regularization*** to prevent overfitting

$$\arg\max_{\theta} \log p(\mathbf{y}|\mathbf{x}, \theta) - \lambda_2 \sum |\theta|^2$$

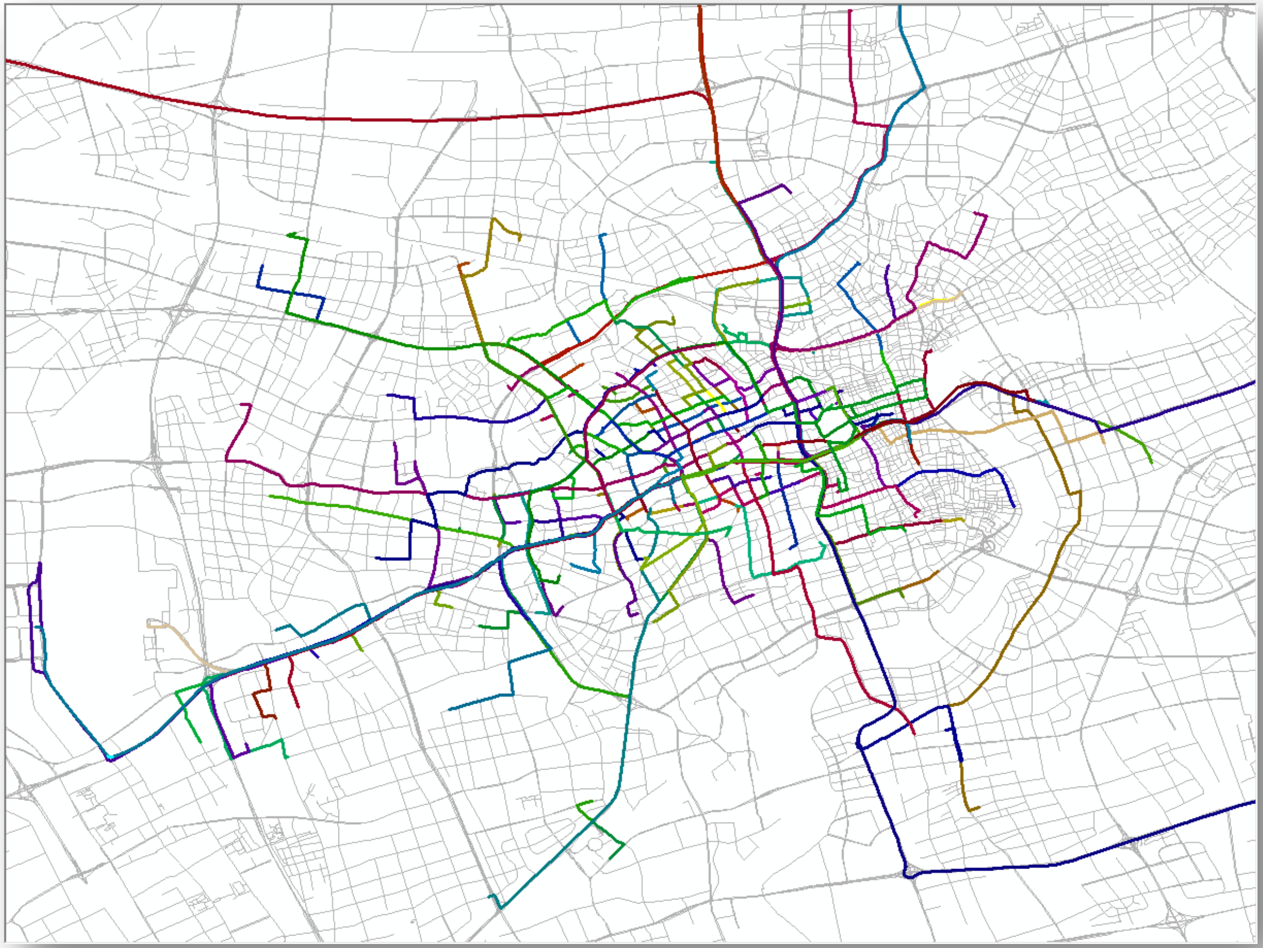
Since the cost function is convex, it can be solved by unconstrained optimization method e.g., BFGS

We use ***L1 regularization***

$$\arg\max_{\theta} \log p(\mathbf{y}|\mathbf{x}, \theta) - \lambda_1 \sum |\theta|$$

Which is non-differentiable at 0s, optimization is more difficult, but it allows sparse parameters. For efficiency concern, *Projected Scaled Sub-Gradient (PSSG) is used*

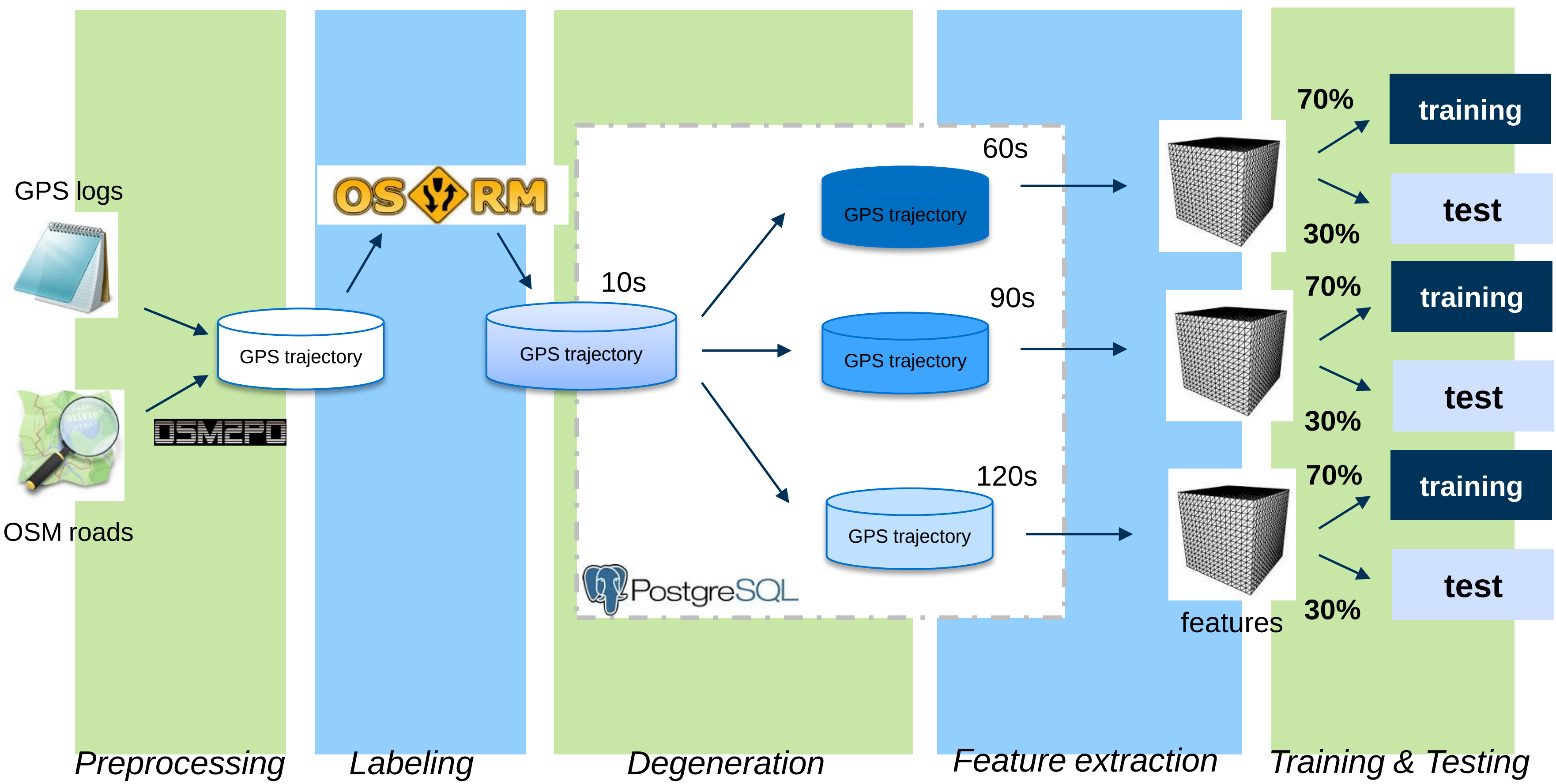
Experiment setting



GPS data from 70 taxis in road network during a day, Shanghai, China

- *124 taxis trajectories*
- *1 day*
- *14.000 GPS pts*
- *10s interval*

Experiment workflow

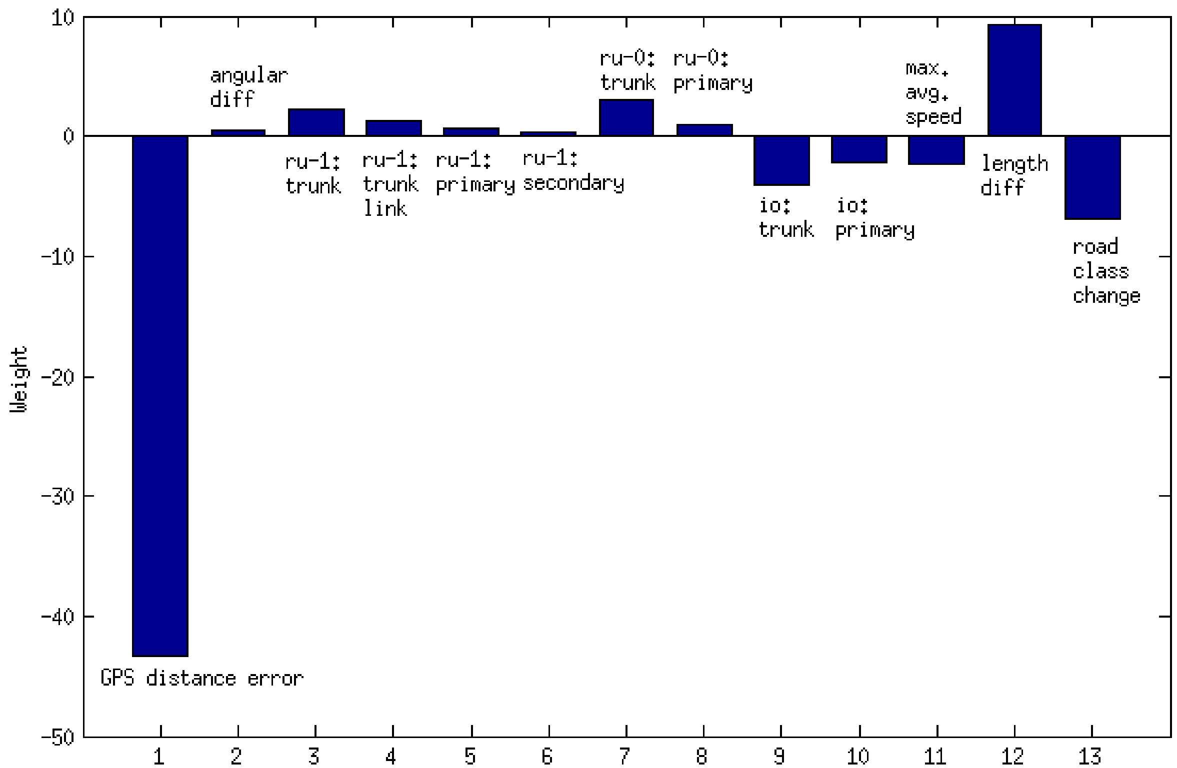


Common model (L2) vs. model with Feature Selection (L1)

Intervals	Regularizer	Feature#	Pt. err. rate	Path. err. rate
60	L2	44	.228	.299
	L1	18	.153	.194
90	L2	43	.235	.304
	L1	20	.146	.197
120	L2	43	.255	.339
	L1	17	.166	.234

- Feature selection yields
  50% feature reduction and 10% performance improve
- Surprisingly, more features do NOT outperform the *baseline*

Learned patterns



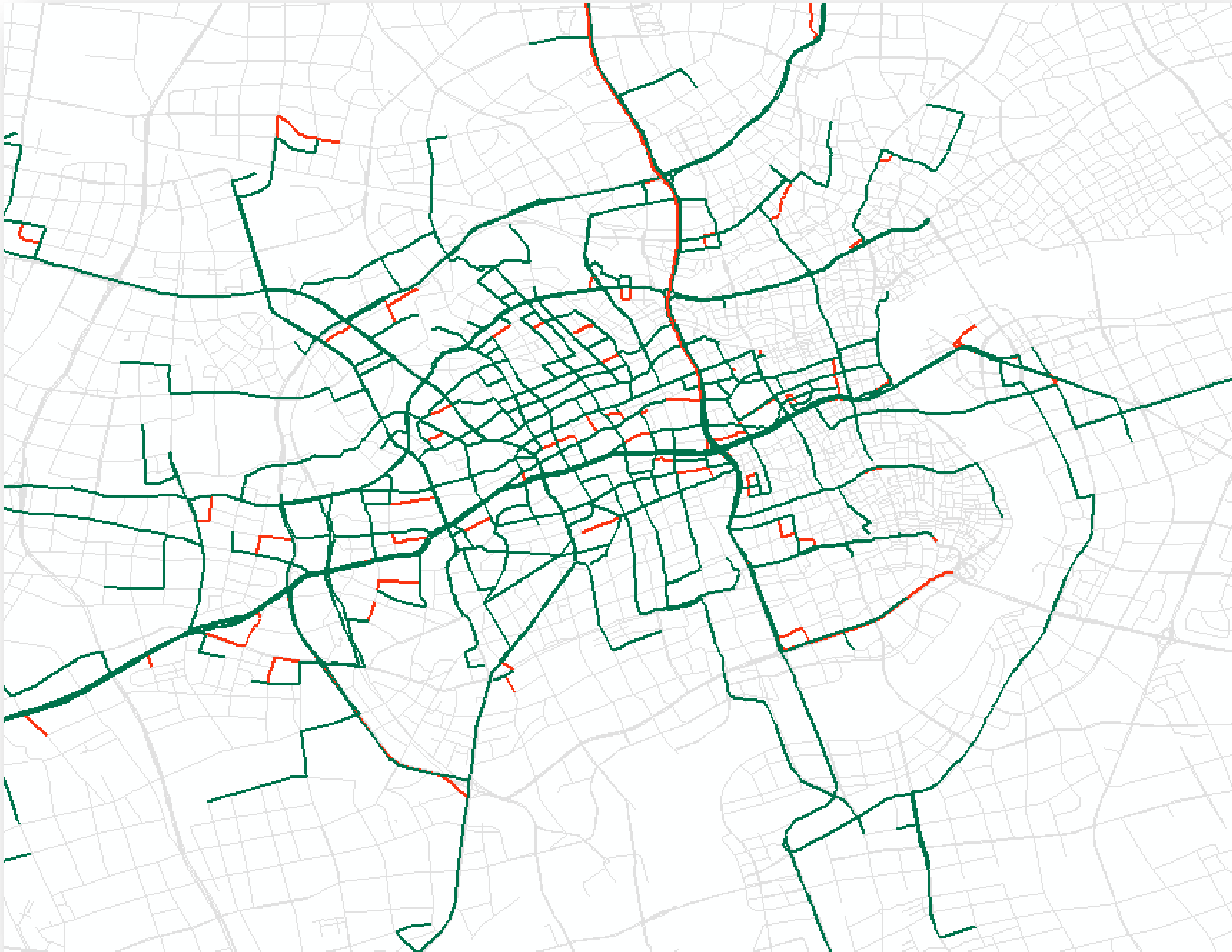
Most relevant features:

Distance error

Length difference

Road class change

Mapping the results



***Green**: Ground truth*
***Red**: recovered Route*

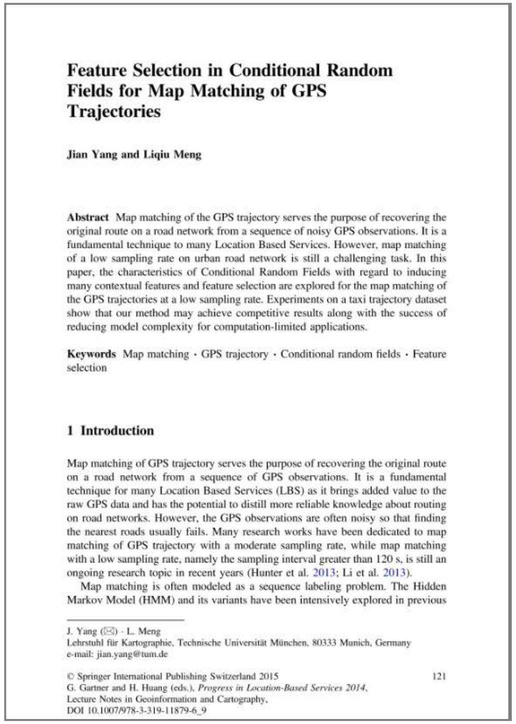
Among all errors:

***Missing label*: 18.3%**
***Parallel roads*: 13.7%**
***U-turn* 13.0%**
End points 10.0%
Position outlier 9.9%

Future work

- Analysis of the impact of using Open source Road Data
- Scale issues in movement analysis

Thanks for *your* attention!



Feature Selection in Conditional Random Fields for Map Matching of GPS Trajectories

Jian Yang and Liqiu Meng

Abstract Map matching of the GPS trajectory serves the purpose of recovering the original route on a road network from a sequence of noisy GPS observations. It is a fundamental technique to many Location Based Services. However, map matching of a low sampling rate on urban road network is still a challenging task. In this paper, the characteristics of Conditional Random Fields with regard to inducing many contextual features and feature selection are explored for the map matching of the GPS trajectories at a low sampling rate. Experiments on a taxi trajectory dataset show that our method may achieve competitive results along with the success of reducing model complexity for computation-limited applications.

Keywords Map matching · GPS trajectory · Conditional random fields · Feature selection

1 Introduction

Map matching of GPS trajectory serves the purpose of recovering the original route on a road network from a sequence of GPS observations. It is a fundamental technique for many Location Based Services (LBS) as it brings added value to the raw GPS data and has the potential to distill more reliable knowledge about routing on road networks. However, the GPS observations are often noisy so that finding the nearest roads usually fails. Many research works have been dedicated to map matching of GPS trajectory with a moderate sampling rate, while map matching with a low sampling rate, namely the sampling interval greater than 120 s, is still an ongoing research topic in recent years (Hunter et al. 2013; Li et al. 2013).

Map matching is often modeled as a sequence labeling problem. The Hidden Markov Model (HMM) and its variants have been intensively explored in previous

J. Yang (✉) · L. Meng
Lehrstuhl für Kartographie, Technische Universität München, 80333 Munich, Germany
e-mail: jian.yang@tum.de

G. Gartner and H. Huang (eds.), *Progress in Location-Based Services 2014*,
Lecture Notes in Geoinformation and Cartography,
DOI 10.1007/978-3-319-11879-6_9

Jian Yang, Liqiu Meng

jian.yang@tum.de

Lehrstuhl für Kartographie, TU München

Germany