# Mobile Location Prediction in Spatio-Temporal Context

Huiji Gao, Jiliang Tang, Huan Liu

Data Mining and Machine Learning Lab
Arizona State University
June 18, 2012

# Mobile Location Prediction

- Applications:
  - Mobile Advertising
  - Traffic Planning
  - User oriented coupon dispersion
  - Disaster Relief

- Challenges:
  - **Over-fitting**
    Long spatial-temporal trajectories with massive n-gram spatial patterns and sparse temporal patterns, smoothing techniques are indispensable.
  - **Integration**
    Seek a good way to integrate both spatial information and temporal information.

# Problem Statement

**Given a user with a series of his historical visits in a previous time section, and a context of the latest visit location with the time of the next visit, the location prediction problem in Nokia Mobile Data Challenge can be described as finding the probability of**

$$p(v_i = l \mid t_i = t, v_{i-1} = l_k)$$

**Where**

$v_i = l$ **: the i-th visit at location l;**

$t_i = t$ **: the i-th visit happens at time t;**

**set as the** "ending time of current visit"

$v_{i-1} = l_k$ **: the (i-1)-th visit happened at location $l_k$.**

DMML

ARIZONA STATE UNIVERSITY

# Problem Statement

**Using Bayes' rule,**

$$
\begin{aligned}
& p(v_i = l \mid t_i = t, v_{i-1} = l_k) \\
& = \frac{p(v_i = l, t_i = t \mid v_{i-1} = l_k)}{p(t_i = t)} \\
& \propto p(v_i = l, t_i = t \mid v_{i-1} = l_k) \\
& = \boxed{p(t_i = t \mid v_i = l, v_{i-1} = l_k)}\, p(v_i = l \mid v_{i-1} = l_k) \\
& = \boxed{p(t_i = t \mid v_i = l)}\, p(v_i = l \mid v_{i-1} = l_k)
\end{aligned}
$$

**Consider** $p(t_i = t \mid v_i = l, v_{i-1} = l_k) = p(t_i = t \mid v_i = l)$ **under the assumption that the probability of current visit time is only relevant to the current visit location.**

DMML

Arizona State University

# Problem Statement

$$p(v_i = l \mid t_i = t, v_{i-1} = l_k)$$
$$= \boxed{p(t_i = t \mid v_i = l)}\,\boxed{p(v_i = l \mid v_{i-1} = l_k)}$$

## Temporal Constraint

The probability of the i-th visit happening at time t, observing that the i-th visit location is l.

## Spatial Prior

The probability of next visit at location l given the current visit at $l_k$

DMML

ARIZONA STATE UNIVERSITY

# Spatial Prior

❖Two properties of spatial prior $p(v_i = l \mid v_{i-1} = l_k)$

- Power Law Distribution

People tend to go to few places many times, and many places few times.

- Short Term Effect

The current visit location is more relevant to the latest visit than the older visit.

❖Correspondences between language and LBSN modeling

| Language Modeling | | LBSN Modeling | |
|---|---|---|---|
| Corpus | | Check-in collection | |
| Document | | Individual check-ins | |
| Document Structure | Paragraph | Check-in Structure | Monthly check-in sequence |
| | Sentence | | Weekly check-in sequence |
| | Phrase | | Daily check-in sequence |
| | Word | | Check-in location |

[Gao et al. Exploring Social-Historcal Ties on Location-Based Social Networks. ICWSM 2012]

# Spatial Prior

❖ Hierarchical Pitman-Yor (HPY) Language Model to generate the spatial prior $p(v_i = l \mid v_{i-1} = l_k)$ (HPY spatial prior)

➢ Generate the probability of next location based on an observation of historical visiting sequence.

➢ Consider a combination of all the n-gram patterns in the previous visits with various pattern weights.

➢ The latest visit has higher weight than the older visit.

[Gao et al. Exploring Social-Historcal Ties on Location-Based Social Networks. ICWSM 2012]

DMML

Arizona State University

# Temporal Constraint

Temporal Constraint:

$$p(t_i = t \mid v_i = l)$$
$$= p(h_i = h, d_i = d \mid v_i = l)$$
$$= p(h_i = h \mid v_i = l)\, p(d_i = d \mid v_i = l)$$

**Hourly Constraint** **Daily Constraint**

h: Hour of the day, i.e., 10:00am, 3:00pm
d: Day of the week, i.e., Monday, Sunday

# Temporal Constraint

Compute $p(h_i = h \mid v_i = l)$ and $p(d_i = d \mid v_i = l)$

- For a visit location l that has happened at h (d) in the previous visits, it's easy to get $p(h_i = h \mid v_i = l)$ and $p(d_i = d \mid v_i = l)$

- For a visit location l that has not happened at h (d) in the previous visits (majority part in the training set),

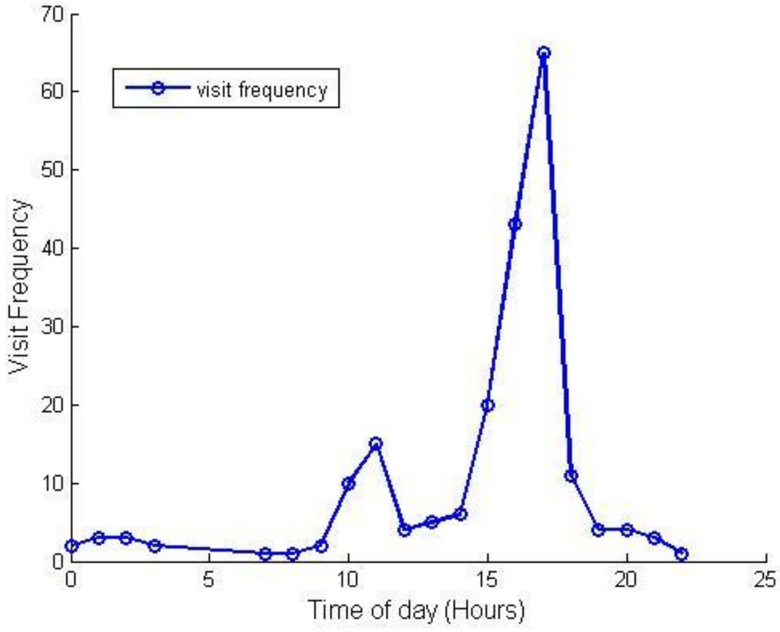
how to compute?

# Temporal Constraint

Compute $p(h_i = h \mid v_i = l)$ and $p(d_i = d \mid v_i = l)$

- Distribution of a user's visits at a specific location in 24 hours.
  (user id: 013; place id: 3 )

$$p(h_i = h \mid v_i = l) = N_l(h \mid \mu_h, \sigma_h^2)$$




$$p(H \mid v_i = l) = \prod_{i=1}^{N_l} N_l(h_i \mid \mu_h, \sigma_h^2)$$

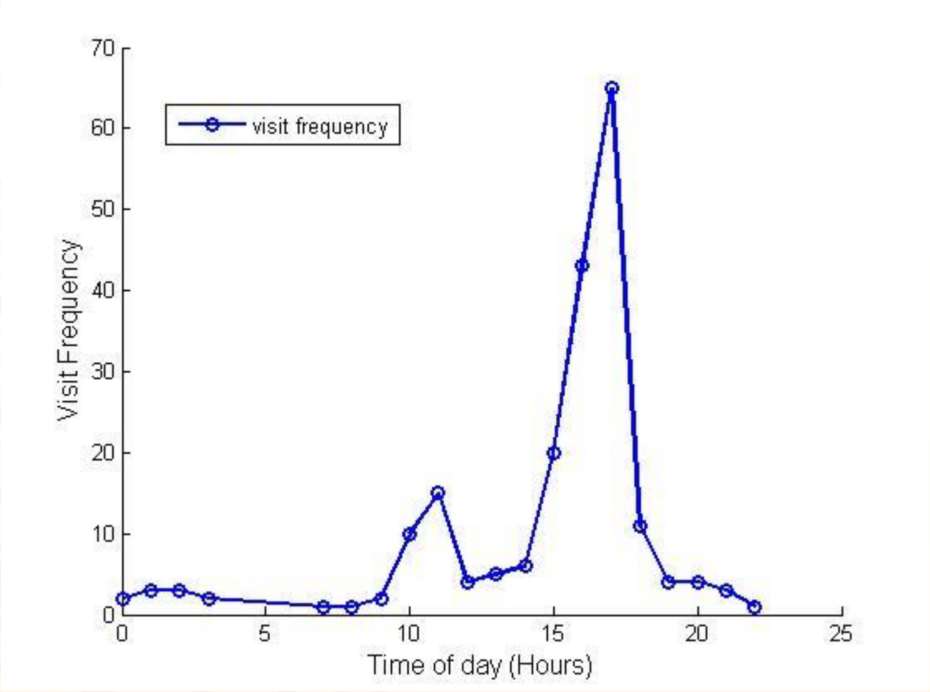
$$(h_i \in H, |H| = N_l)$$

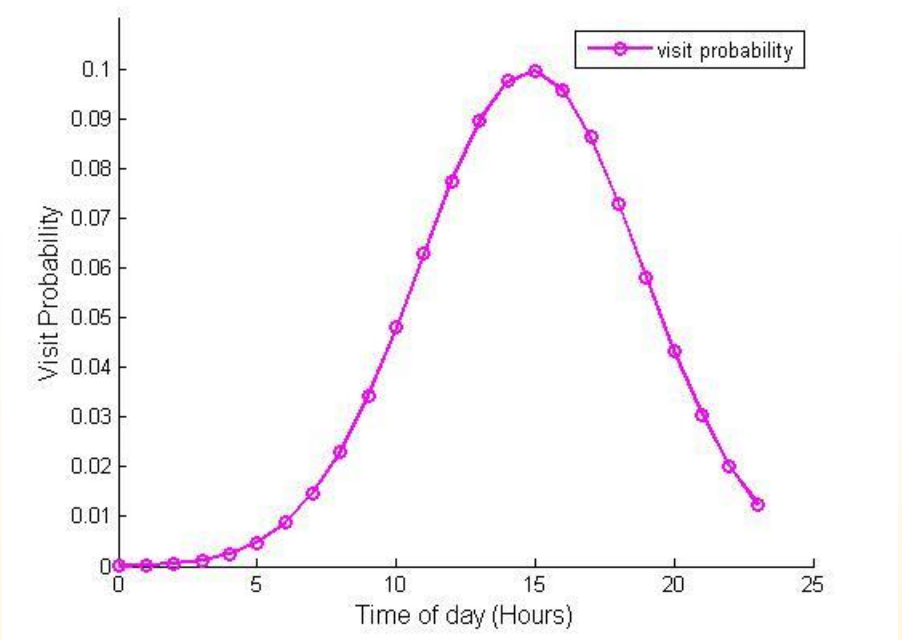
Maximizing Likelihood $\Rightarrow \begin{cases} \mu_h \\ \sigma_h^2 \end{cases}$

# Temporal Constraint

Curve Fitting:

[user id: 013; place id: 3]



# Location Prediction

Probability of visiting location l at time t with the latest visit at $l_k$

$$p(v_i = l \mid t_i = t, v_{i-1} = l_k)$$
$$= p(v_i = l \mid v_{i-1} = l_k) p(h_i = h \mid v_i = l) p(d_i = d \mid v_i = l)$$
$$= p(v_i = l \mid v_{i-1} = l_k) N_l(h \mid \mu_h, \sigma_h^2) N_l(d \mid \mu_d, \sigma_d^2)$$

HPY Prior　　Gaussian　　Gaussian

HPY Prior Hour-Day Model (HPHD)

# Experiments

Experiment Setting

❖ For Submission:

Training set: Set A

Testing set: Set C (no ground truth)

Number of users: 80

❖ For Evaluation

Divide set A into training and testing parts.

➢Testing set: Toy data provided by Nokia with 3373 unknown locations.

➢Training set: For each user, all the visits in set A that happened before the visits in testing set.

# Experiments

❖ Baseline Methods (Spatial Family)

1. Most Frequent Visit Model (MFV)
   Consider the most frequent visited location
2. Order-1 Markov Model (OMM)
   Consider the most frequent two-gram pattern with the latest visit as context.
3. Fallback Markov Model
   A combination of MFV and OMM
4. HPY Prior Model (HP)
   Consider the HPY prior only to predict the next location.

# Experiments

❖ Baseline Methods (Temporal Family)

5. Most Frequent Hourly Model (MFH)
   Predict the next visit at time h as the most frequent visit location at h in previous visits.
6. Most Frequent Daily Model (MFD)
   Predict the next visit at time d as the most frequent visit location at d in previous visits.
7. Most Frequent Hour-Day Model (MFHD)
   A combination of MFH and MFD (by mutiplying)

# Experiments

❖Baseline Methods (Spatio-Temporal family)

8. HPY Prior Hourly Model (HPH)
   Consider the HPY prior and hourly information.

9. HPY Prior Daily Model (HPD)
   Consider the HPY Prior and daily information.

❖ Proposed Method (Spatio-Temporal family)
HPY Prior Hour-Day Model (HPHD)
Consider the HPY prior, hourly information, and daily information.

# Experiments

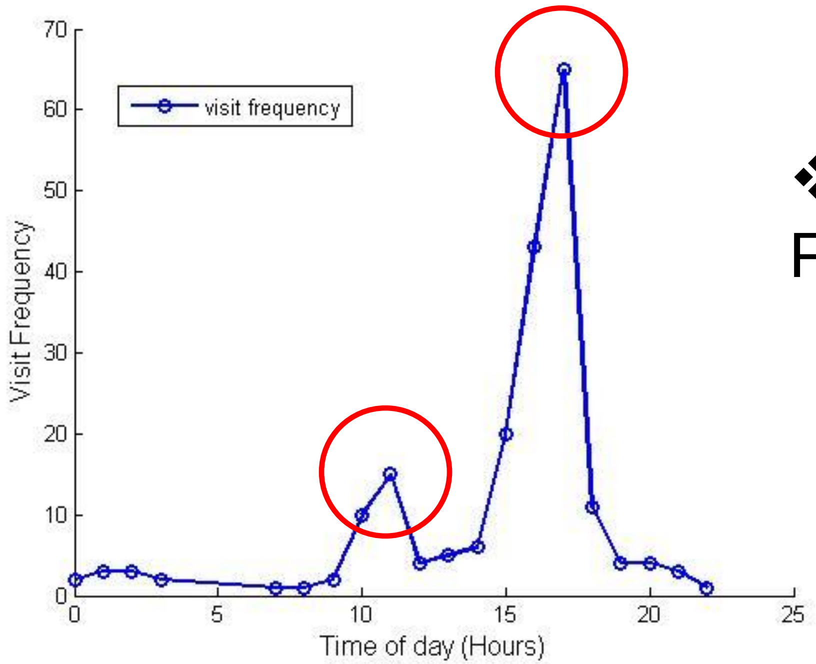
- Results

Table 1: Location Prediction Results

| | Models | Correct No. | Accuracy |
|---|---|---|---|
| Spatial-based | MFV | 1148 | 0.3402 |
| | OMM | 1466 | 0.4345 |
| | FMM | 1583 | 0.4692 |
| | HP | 1610 | 0.4772 |
| Temporal-based | MFH | 1462 | 0.4333 |
| | MFD | 1156 | 0.3426 |
| | MFHD | 1538 | 0.4558 |
| Spatio-temporal | HPH | 1680 | 0.4979 |
| | HPD | 1583 | 0.4692 |
| | **HPHD** | **1705** | **0.5053** |

# Conclusions and Future Work

❖Gaussian Distribution with Two Peaks
- ➢ An alternative version of HPHD. (AHPHD)
- ➢ Not stable, sensitive to peak detection.




❖ Five Submissions

FHD, HP, HPH, HPHD, AHPHD

# Acknowledgements

This work is supported, in part, by ONR (N000141010091).

All the data used in this work is provided by Nokia Mobile Data Challenge.

# Questions?