Point-of-interest Recommendation in Location-based Social Networks

ZHAO, Shenglin
Department of Computer Science & Engineering
The Chinese University of Hong Kong
slzhao@cse.cuhk.edu.hk

Supervisors: Prof. Michael R. Lyu and Prof. Irwin King

Thesis Defense

Dec. 12, 2017

香港中文大學
The Chinese University of Hong Kong


Geo info is
everywhere!
08:08
11963
HUAWEI
Samsung SMART TV


Location-based
Services
Samsung
SMART TV

Applications

Search and Recommendation	Transportation	Healthcare	Public Safety	Game	Environment Monitoring

Location-based Services

Applications

Search and Recommendation	Transportation	Healthcare	Public Safety	Game	Environment Monitoring

Location-based Services

facebook

Facebook Places
Who. What. When. And now where.

foursquare

街旁 Jiepang.com

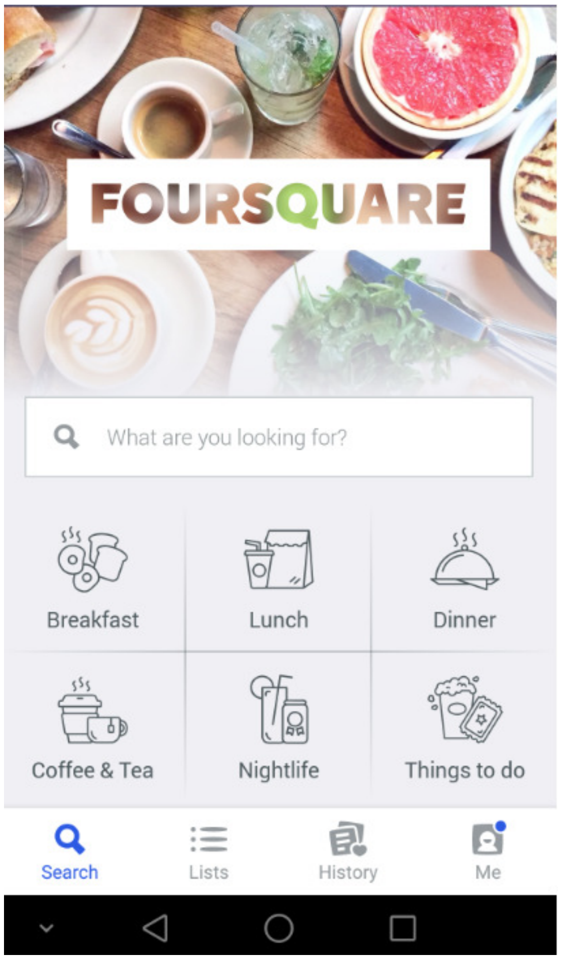
yelp

美团 meituan.com

twitter

大众点评 dianping.com

Foursquare Case



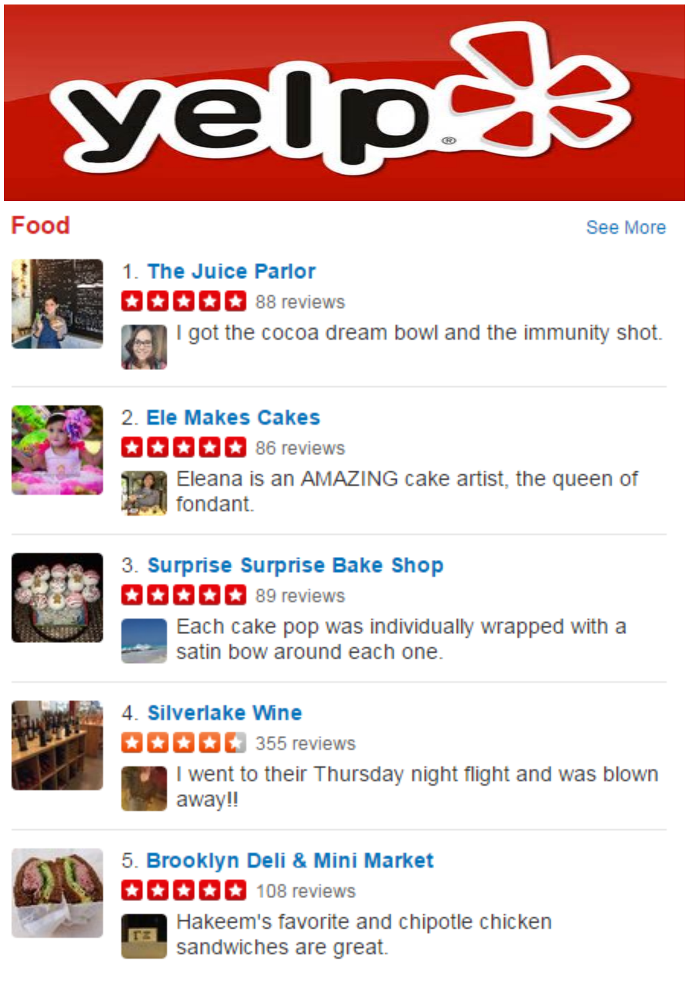
By Oct. 2017

- Over 50 million people each month
- Over 12 billion check-ins

https://foursquare.com/about

http://mashable.com/2012/12/18/apple-foursquare-maps

Yelp Case



By Oct. 2017

- Over 30 million people each month
- Over 3.7 billion dollars

https://yelp.com/about

Location-based social networks

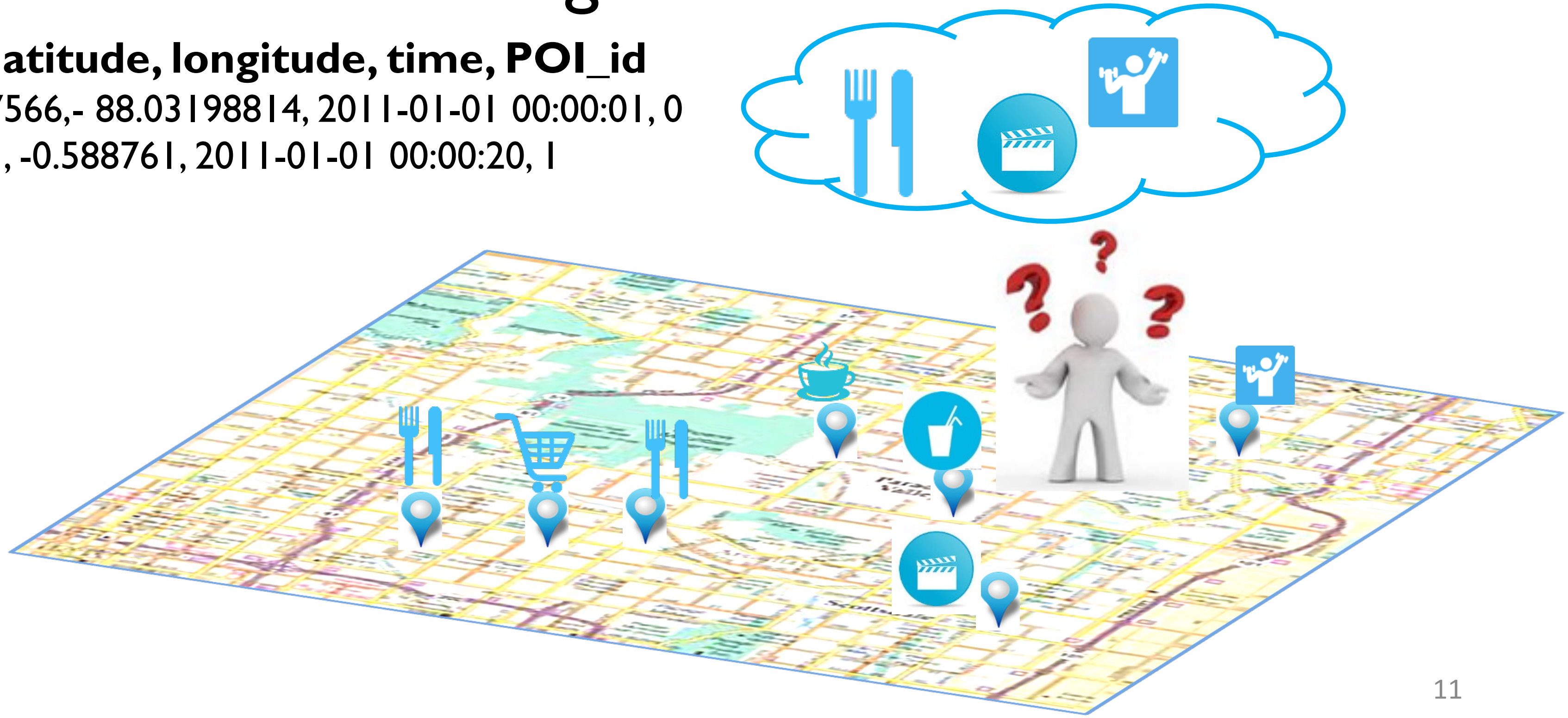
Problem Statement

- Given the user check-in data in LBSNs, point-of-interest (POI) recommendation is an application to answer where to go next

Problem Statement

- Given the user check-in data in LBSNs, point-of-interest (POI) recommendation is an application to answer where to go next

user_id, latitude, longitude, time, POI_id
0, 41.72757566,- 88.03198814, 2011-01-01 00:00:01, 0
0, 51.31791, -0.588761, 2011-01-01 00:00:20, 1
 ……
?



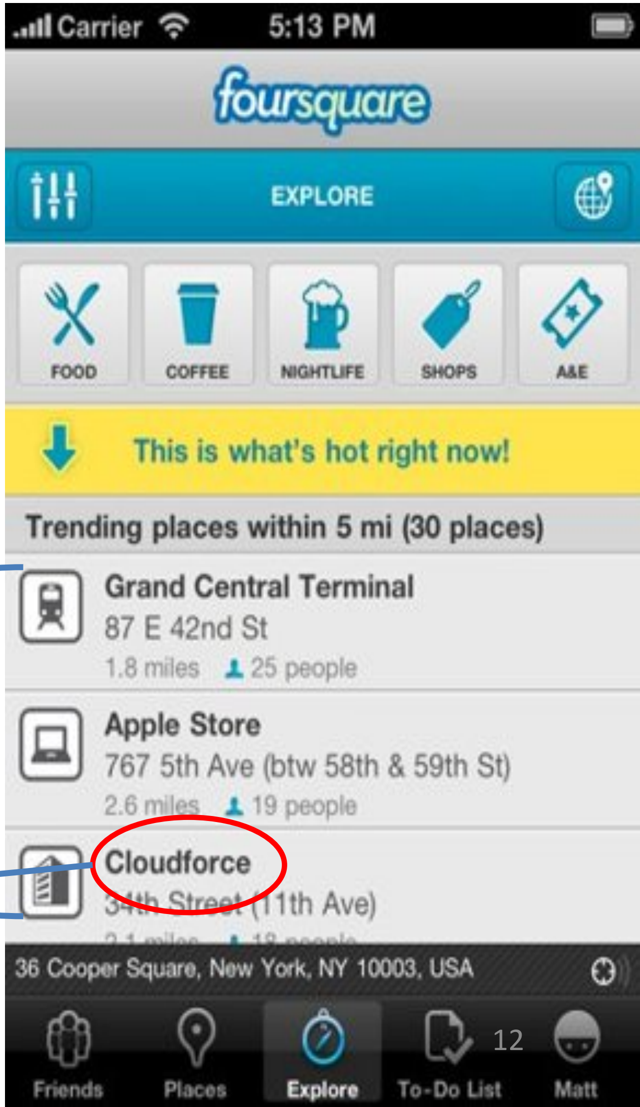
Application Example

- A POI is a specific point location that someone may find interesting and be willing to check-in

- POI recommendation suggests a personalized POI list for each user

Recommended POI list

POI



Problem Significance

- Help users explore the city and find interesting places

Problem Significance

- Help millions of businesses launch advertisements

The Tulsa Rib Company

Barbeque, Caterers

Liz P., Owner
Orange, CA

"Yelp revolutionized how we reach customers."

Watch video

Liz recommends: Yelp Ads Business analytics Video Production and Hosting

Matt Mark Service Company

Electricians

Steve H., Owner
Kansas City, MO

"I advertised with Yelp and was amazed with the results."

Watch video

Steve recommends: Yelp Ads Call to action button Business analytics

https://biz.yelp.com/support/case_studies

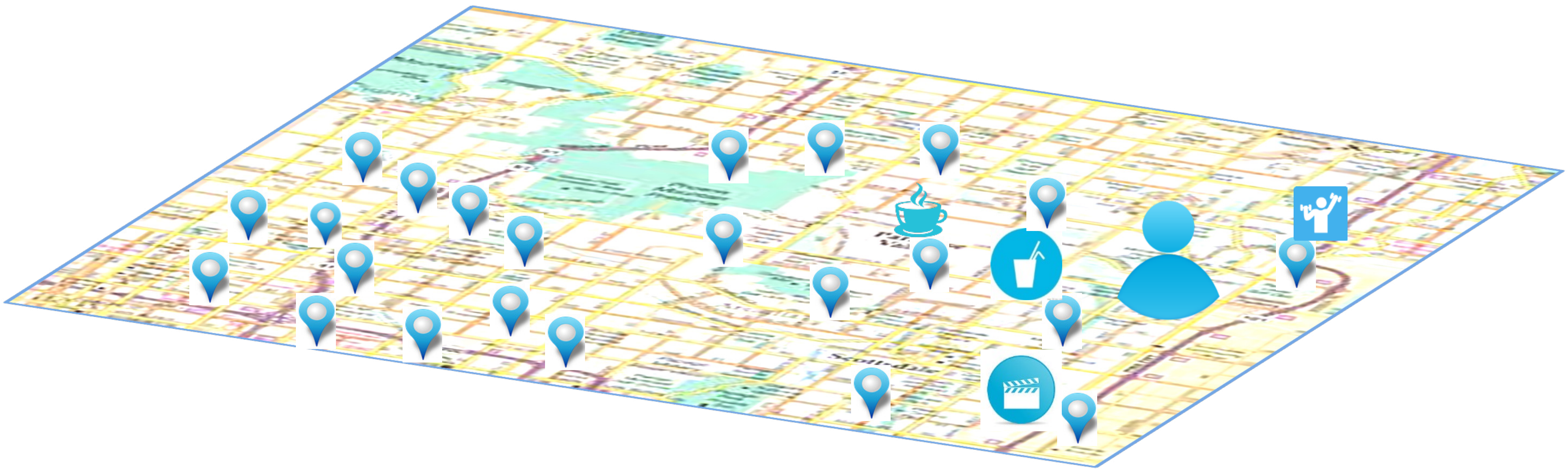
Challenges

- Physical and temporal constraints

http://subscribe.mail.10086.cn/subscribe/readAll.do?columnId=38378&itemId=3935935

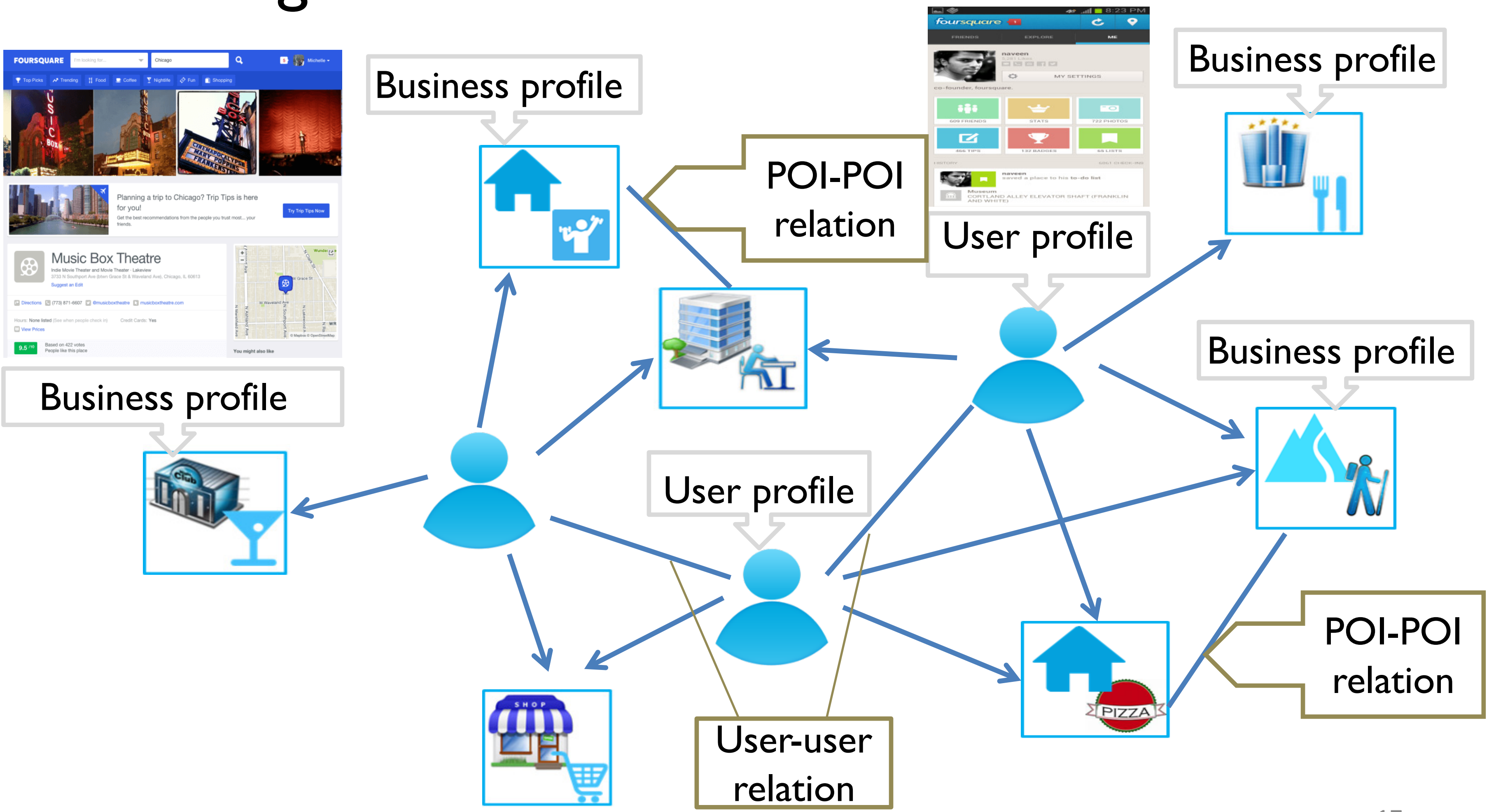
Challenges

- Sparse data



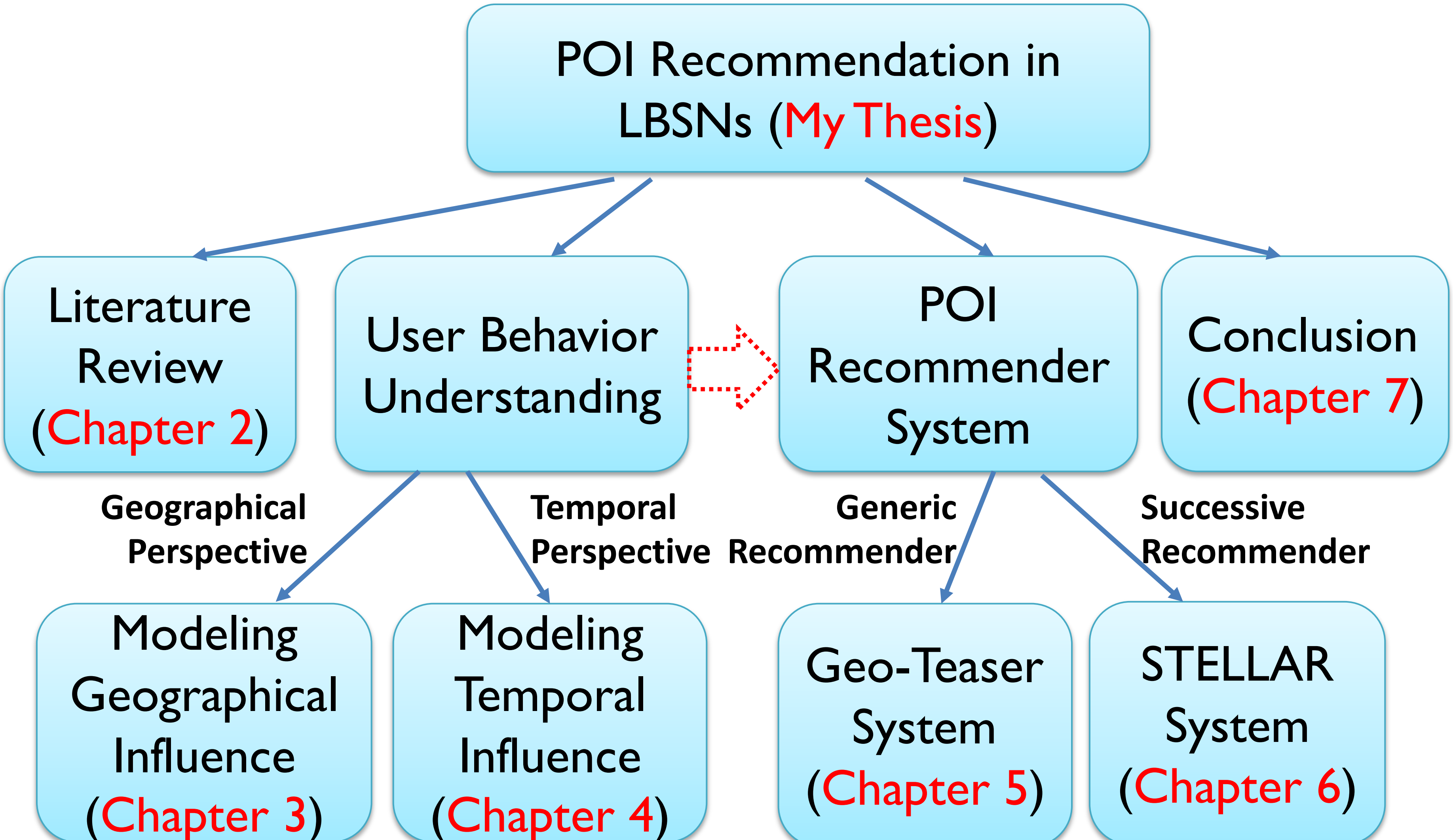
Challenges

- Heterogeneous information

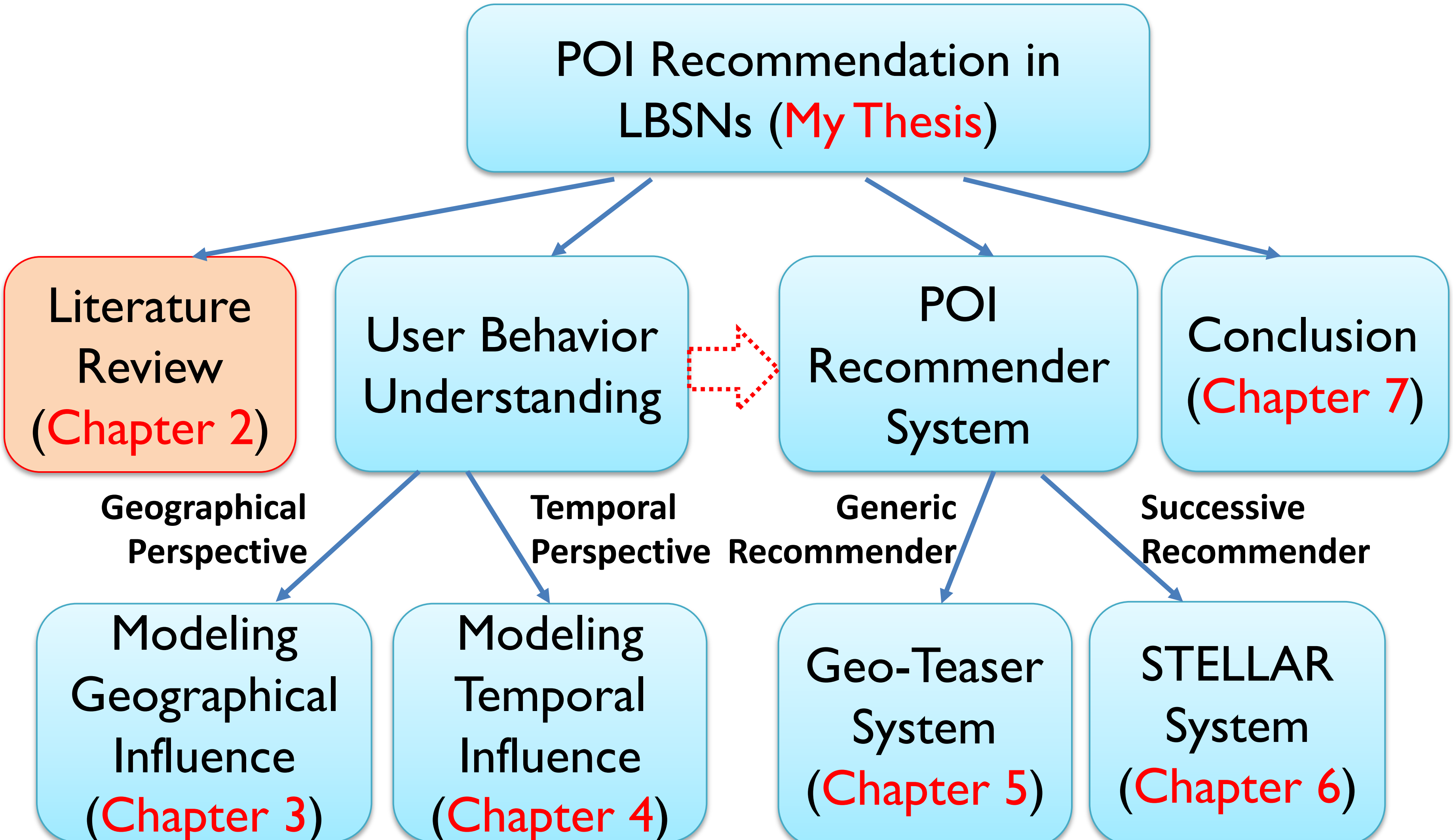


http://images.mobile-patterns.com/1363118149895-2013-03-03%2020.23.04.png
http://www.socialmediaexaminer.com/wp-content/uploads/2016/02/mp-social-media-rebrand-foursquare.png

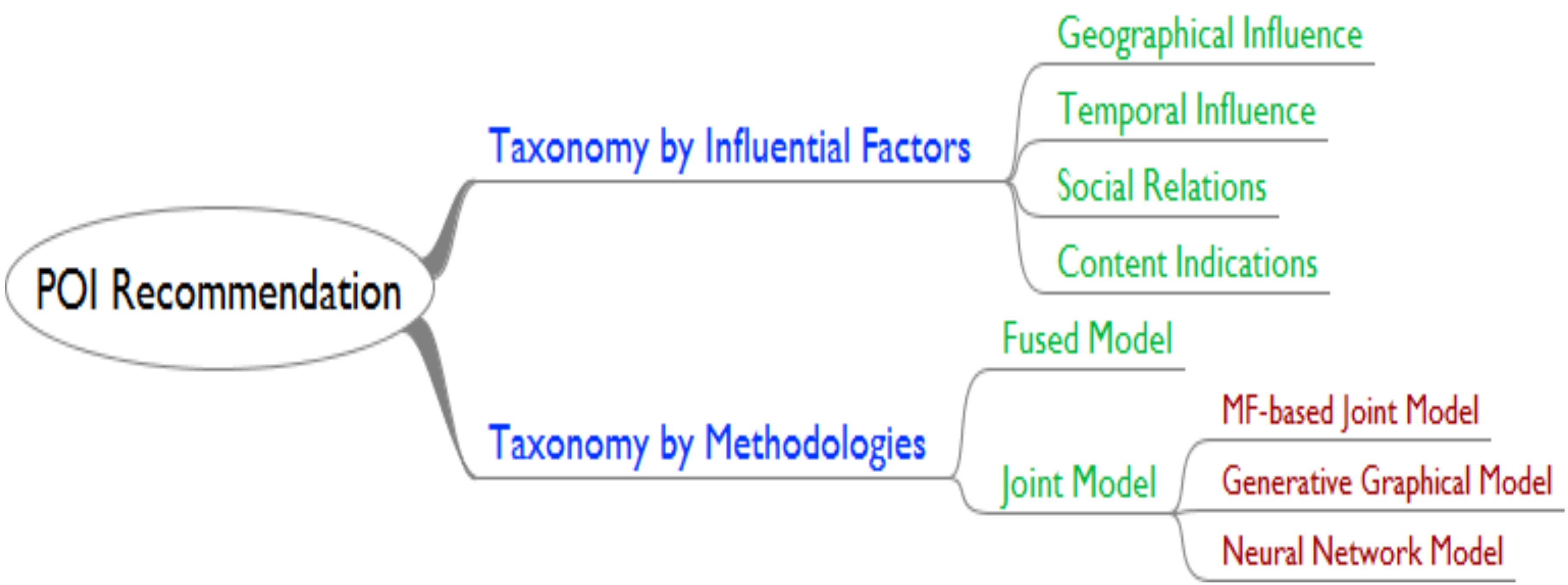
Thesis Structure



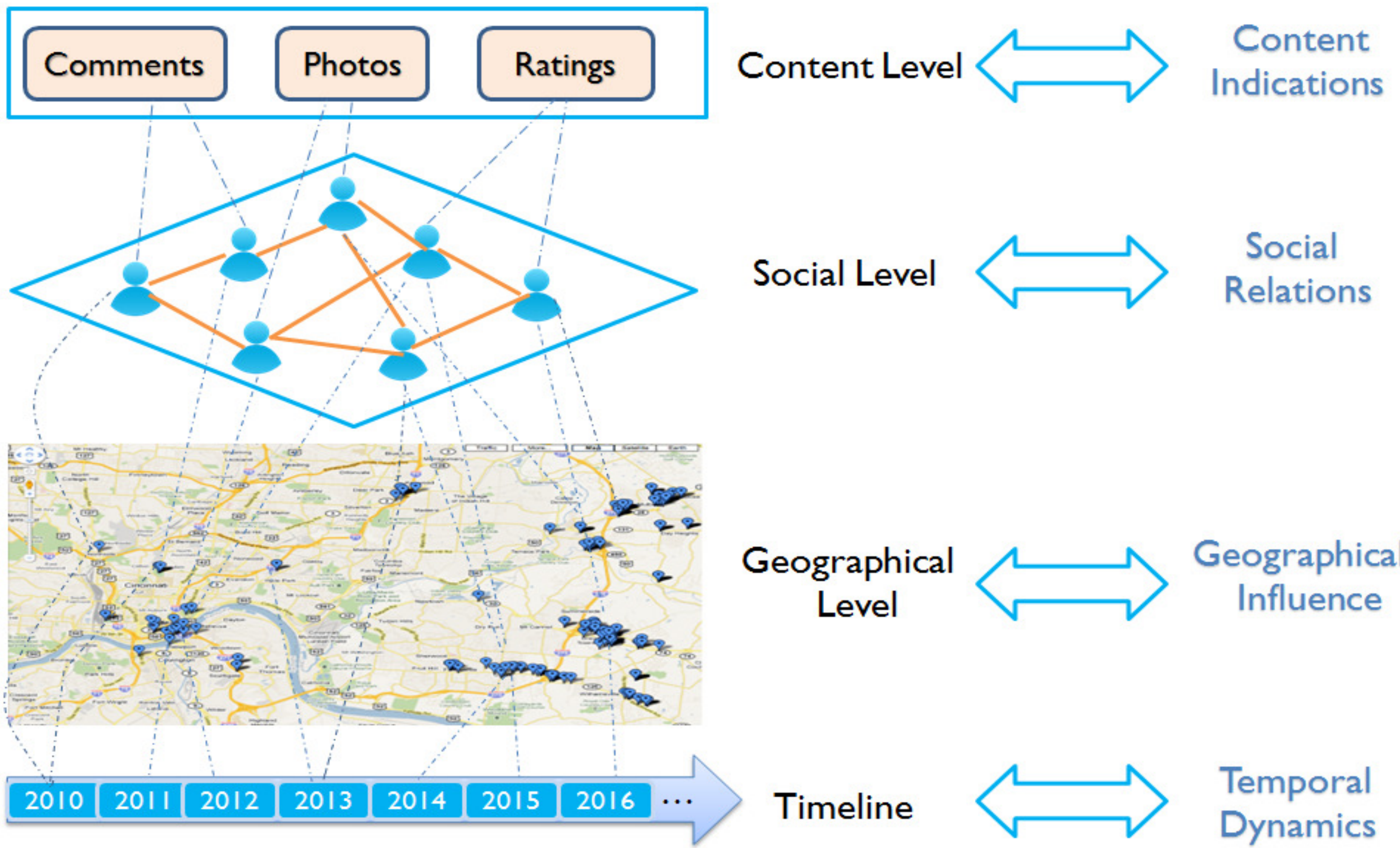
Thesis Structure



Literature Review



Influential Factors in LBSNs

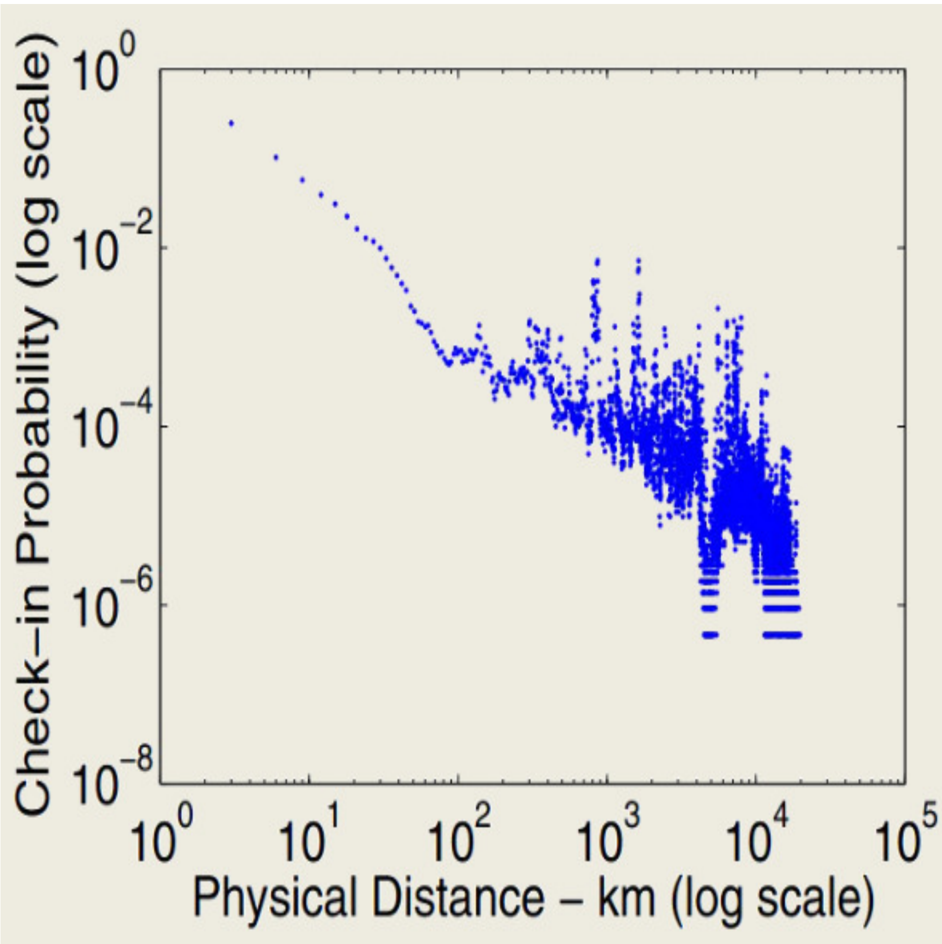


Geographical Influence

- Geographical influence depicts the physical constraints

- Models
  - Power law distribution based model
  - Gaussian distribution based model
  - Kernel density estimation model

Power Law Distribution



- Model formula

Parameters

$$y = a * x^{b}$$

Check-in probability

Distance

- Recommending

$$Pr(l_j|L_i) = \frac{Pr(l_j \cup L_i)}{Pr(L_i)} = \prod_{l_y \in L_i} Pr(d(l_j, l_y))$$

Candidate POI

Visited POIs

Longer distance, less probability

[Ye et al., 2010, Ye et al., 2011, Yuan et al., 2013]

Gaussian Distribution

- Multi-center Gaussian Model (MGM) [Cheng et al., 2012]

$$P(l|C_u) = \sum_{c_u=1}^{|C_u|} P(l \in c_u) \frac{f_{c_u}^{\alpha}}{\sum_{i \in C_u} f_i^{\alpha}} \frac{N(l|\mu_{C_u}, \Sigma_{C_u})}{\sum_{i \in C_u} N(l|\mu_i, \Sigma_i)}$$

Probability belonging to one center, inverse to distance

Normalization of center frequency

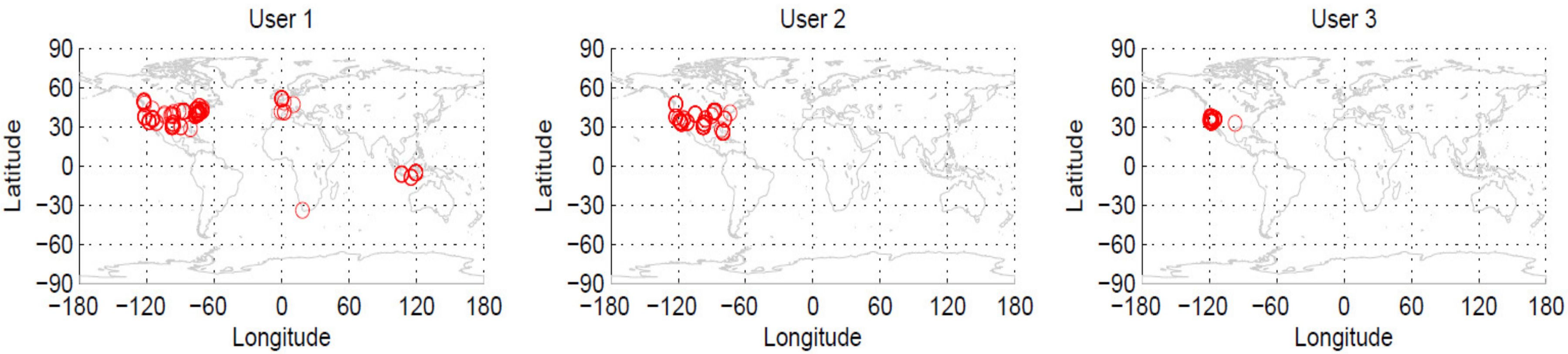
Gaussian distribution

Stay around activity centers

[Cho et al., 2011, Cheng et al., 2012, Zhao et al., 2013]

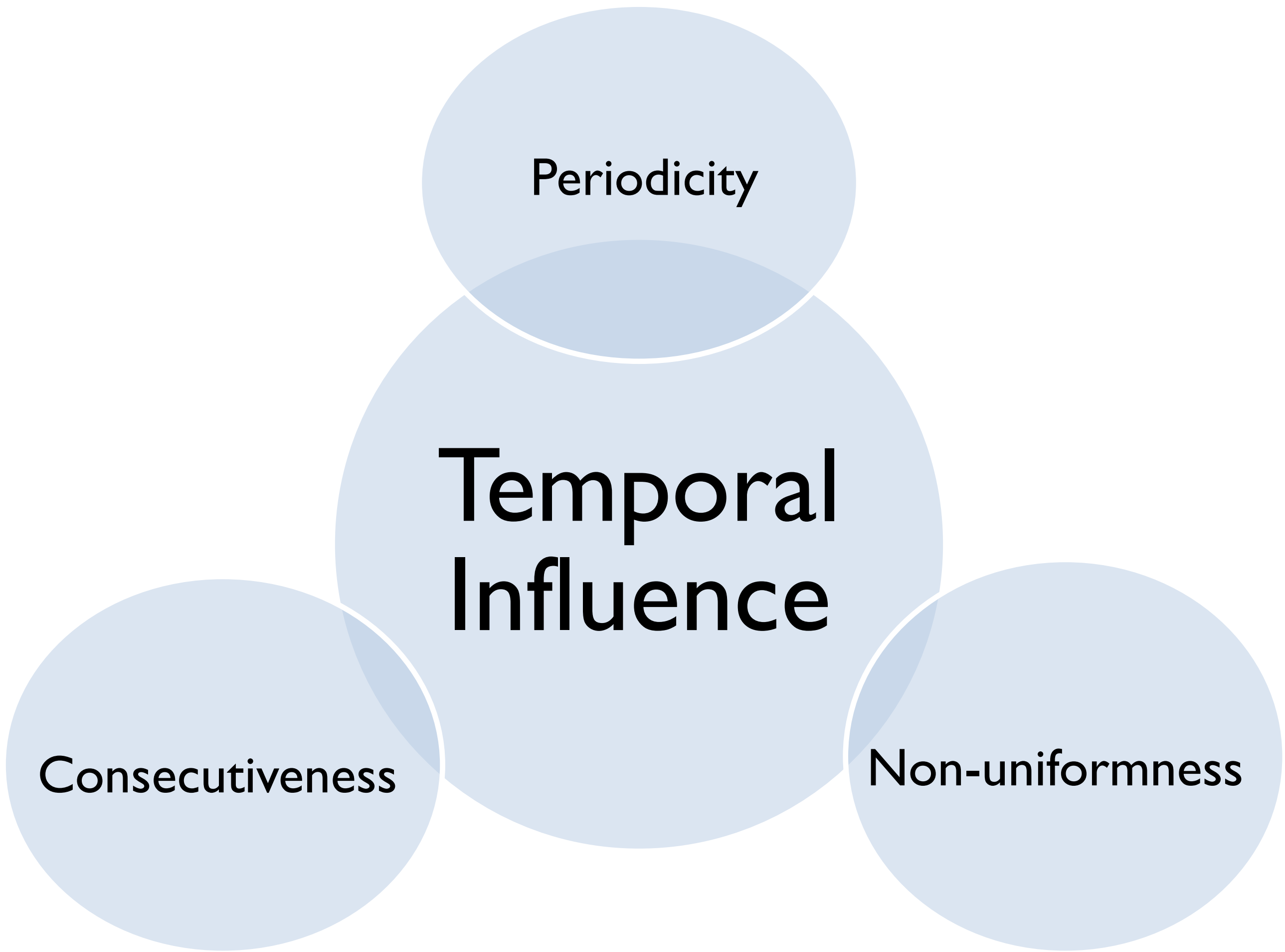
Kernel Density Estimation

- Modeling each user's geographical distribution separately [Zhang et al., 2013]



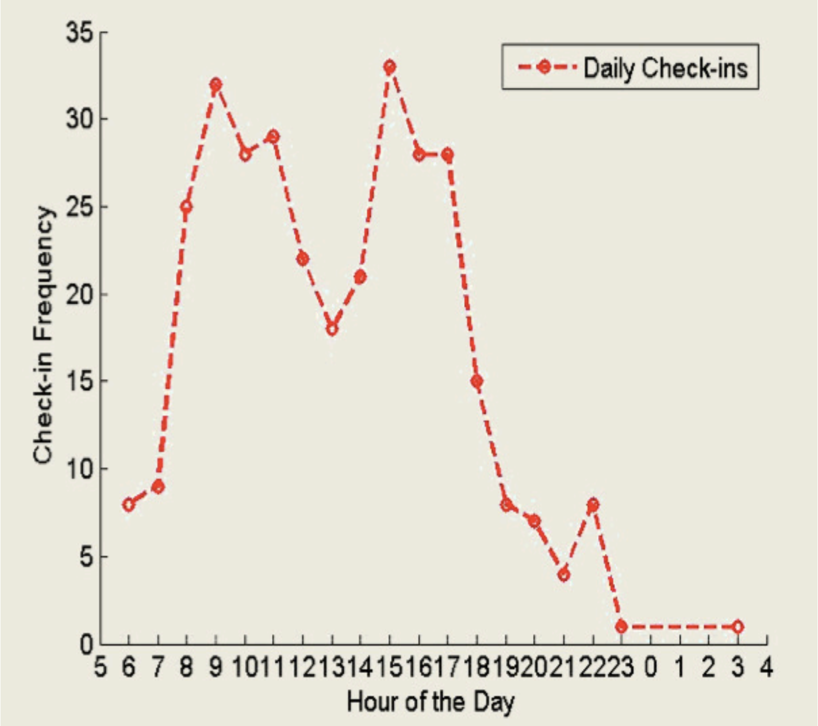
Personalization

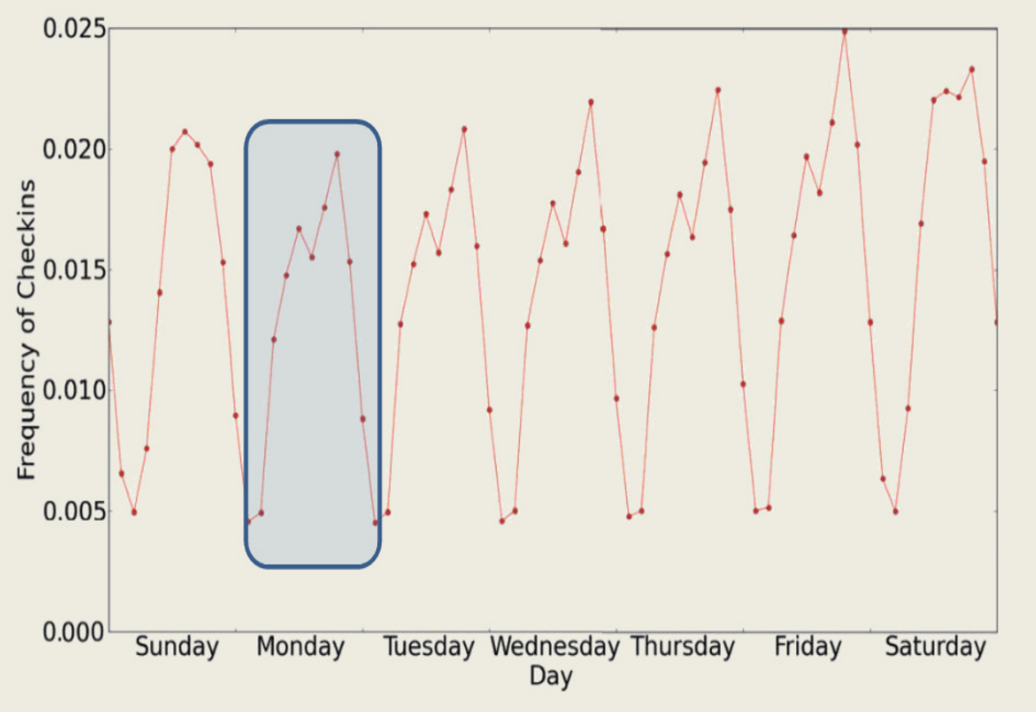
Temporal Influence



Periodic Pattern

- Periodic pattern exists in two levels: day and week



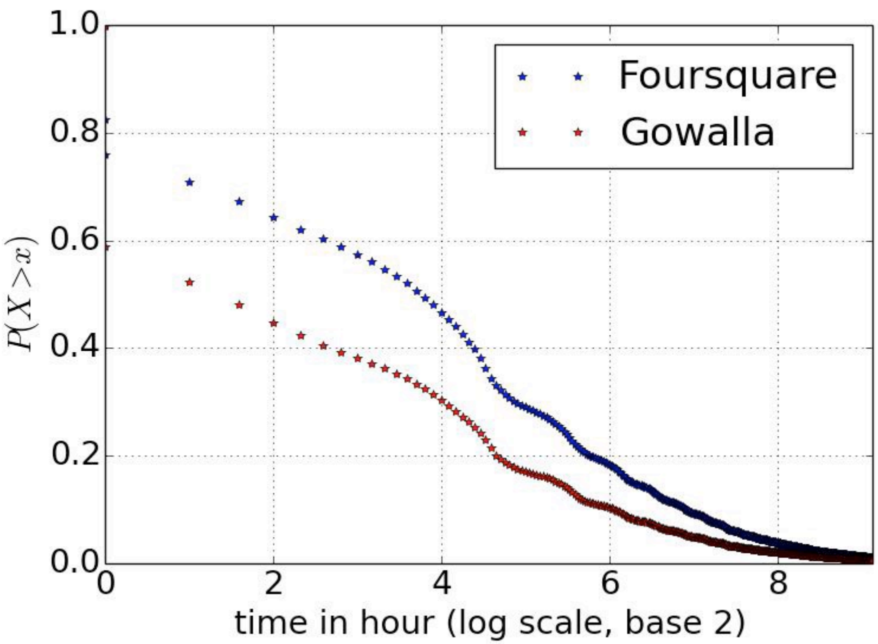


[Ye et al., 2011]

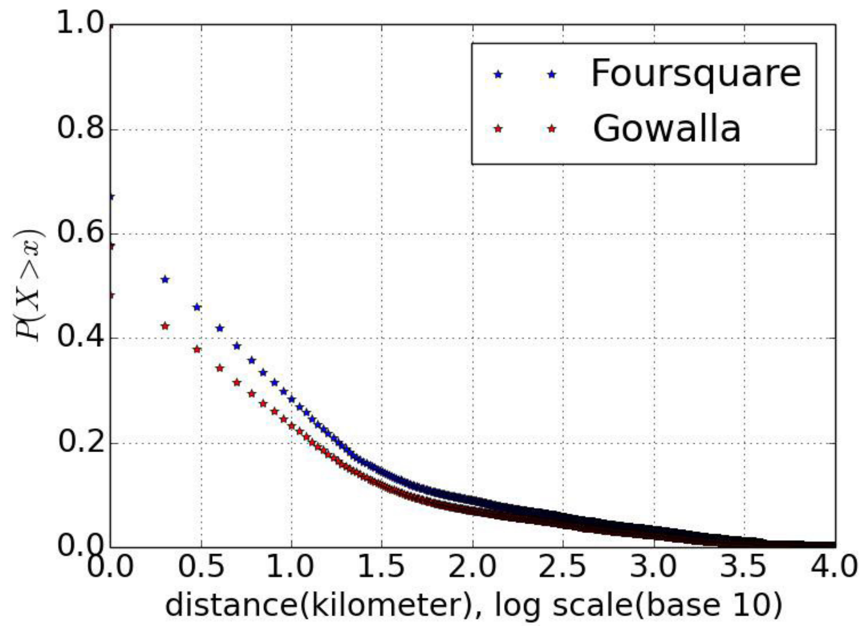
Consecutive Pattern

- Two consecutive check-ins are highly correlated

Temporal property



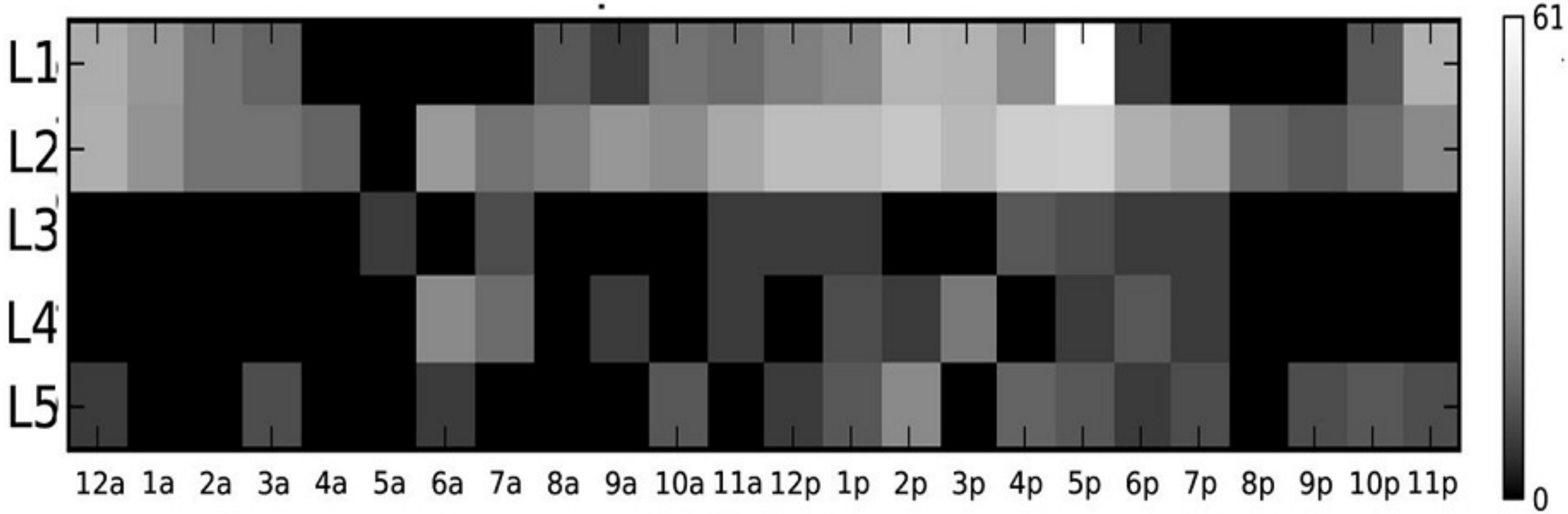
Spatial property



[Cheng et al., 2013, Zhao et al., 2016a, Zhao et al., 2016b]

Non-uniformness

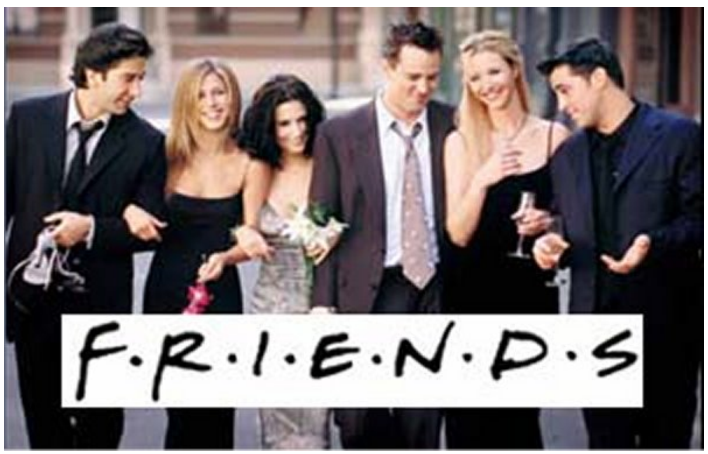
- Check-in preference changes at different time



[Gao et al., 2013, Zhao et al., 2016b]

Social Influence

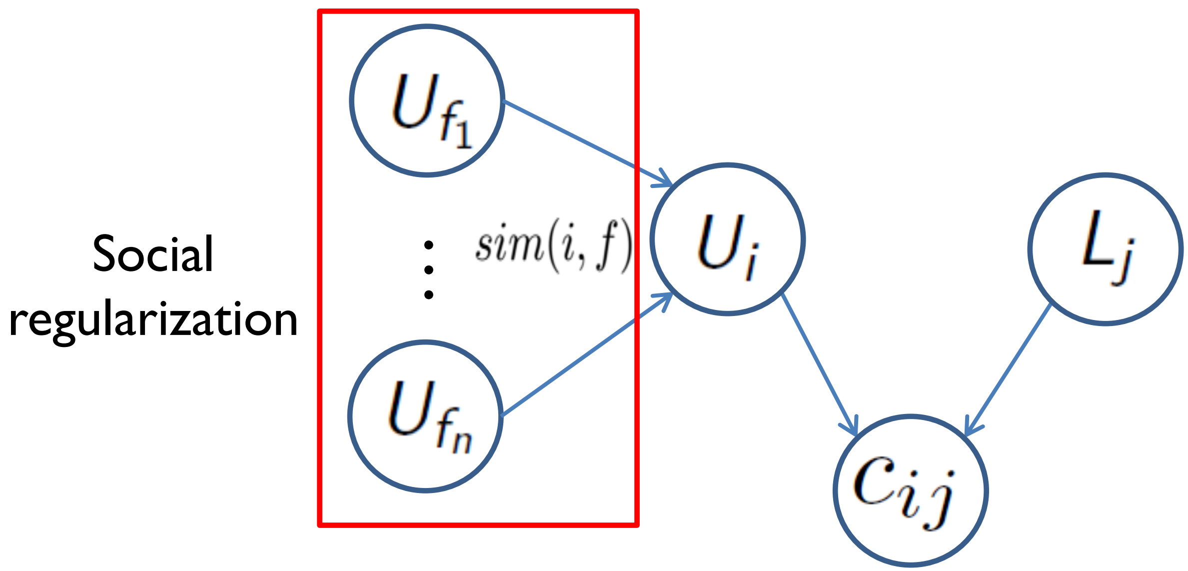
- Friends share similar check-in preferences



Have fun together

Social Influence

- Modeling social influence with regularization in matrix factorization



$$\arg\min_{U,L} \sum_{i=1}^{|\mathcal{U}|} \sum_{j=1}^{|\mathcal{L}|} I_{ij}(g(c_{ij}) - g(U_i^T L_j))^2 + \lambda_1 ||U||_F^2 + \lambda_2 ||L||_F^2 + \boxed{\beta \sum_{i=1}^{N} \sum_{u_f \in F_i} sim(i,f) ||U_i - U_f||^2}$$

[Ma et al., 2011, Cheng et al., 2012, Yang et al., 2013]

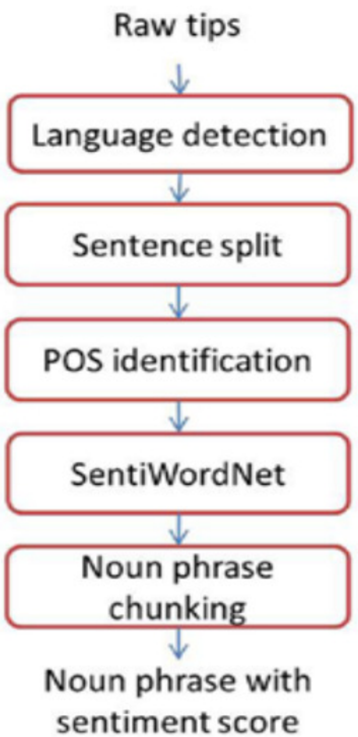
Content Indications

- Using comments to imply the user preference

[Yang et al., 2013, Gao et al., 2015]

Content Indications

- Sentiment-aware matrix factorization[Yang et al., 2013]



Sentiment Score	Preference Measure
[-1, -0.05]	1
(-0.05, -0.01]	2
(-0.01, 0.01)	3
[0.01, 0.05)	4
[0.05, 1]	5

$$\arg\min_{U,L} \sum_{(i,j)\in\Omega} (\hat{C}_{i,j} - U_i L_j^T)^2 + \alpha ||U||_F^2 + \beta ||L||_F^2$$

Combined rating value

Summarization

- Geographical Influence
  - Differentiate POI recommendation from traditional recommendations
- Temporal Influence
  - Periodicity, Consecutiveness, Non-uniformness
- Social Influence
  - Useful but limited improvements
- Content Indications
  - Useful but limited source data

Summarization

- Geographical Influence
  - Differentiate POI recommendation from traditional recommendations
- Temporal Influence
  - Periodicity, Consecutiveness, Non-uniformness
- Social Influence
  - Useful but limited improvements
- Content Indications
  - Useful but limited source data

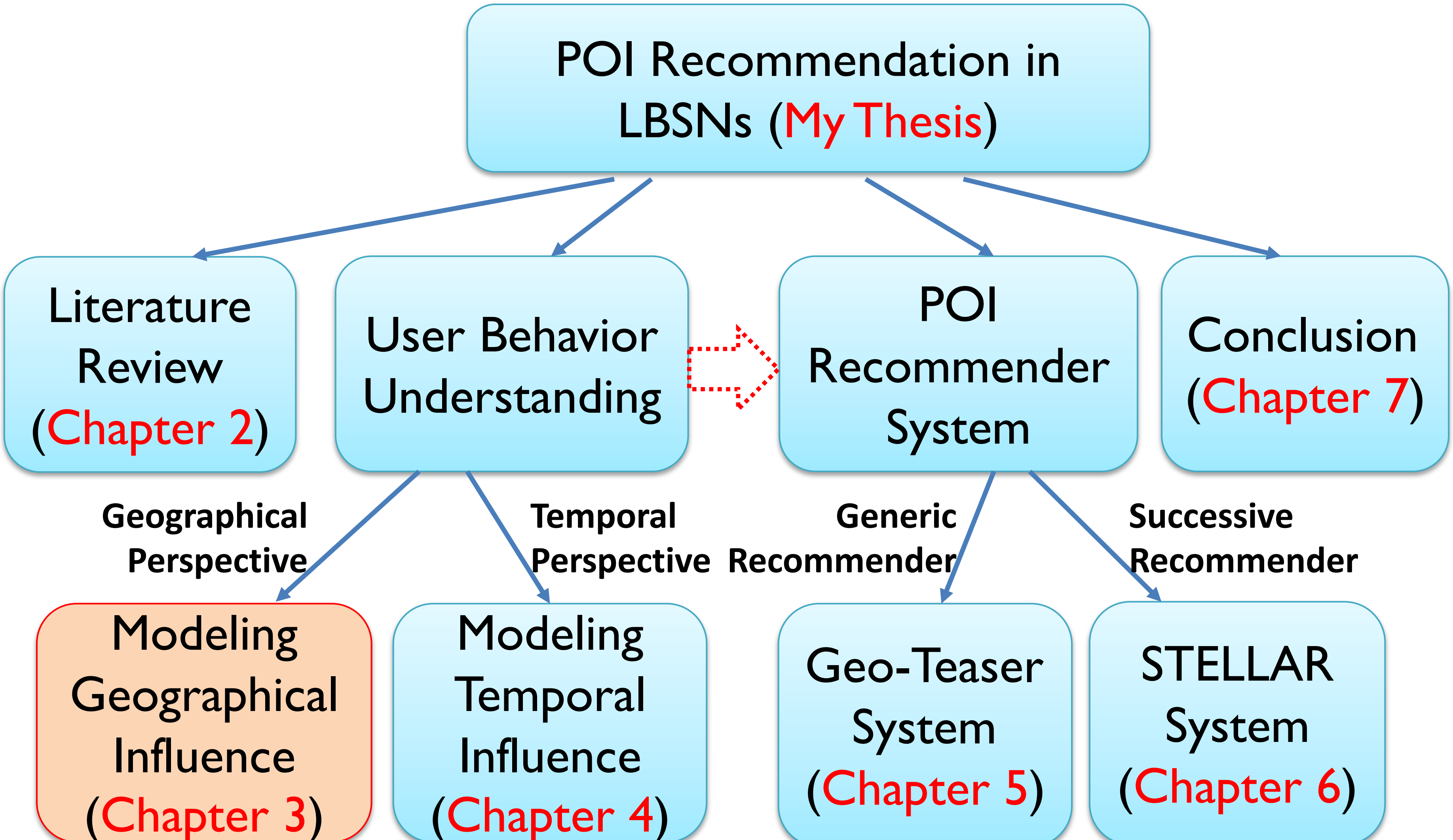
Taxonomy by Methodology

- Fused model
  - Separately recommend POIs and then combine the results
- Joint Model
  - Jointly learn the factors in a model together and recommend POIs
  - MF-based, Generative graphical model, Neural network model

Summarization

- Fused models
  - Easy to extend
  - Independent models
- Joint models
  - Better reflect the real scenario than fused models
  - Attract more attention

Thesis Structure

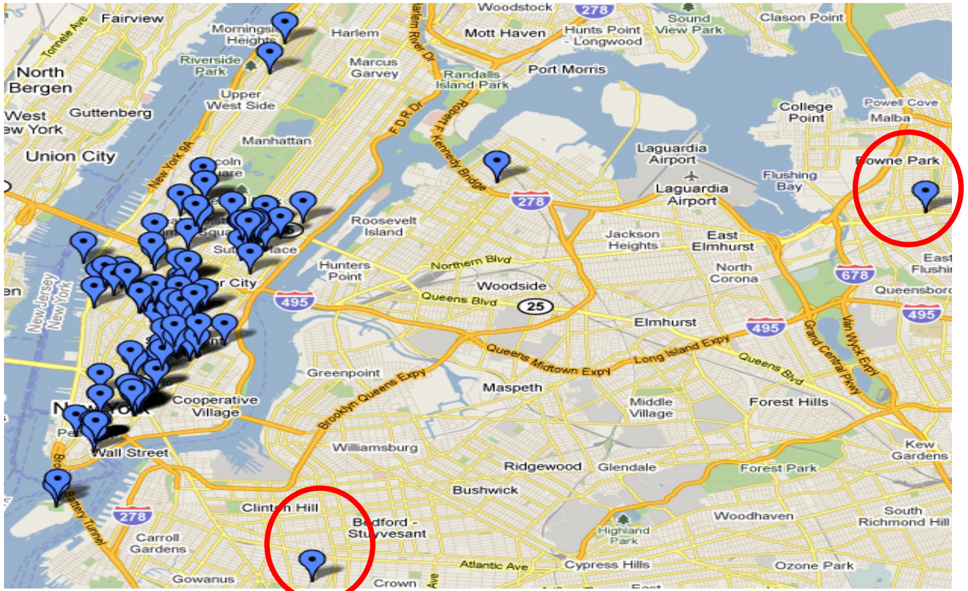


Motivations

- Geographical influence from physical constraints
- Noise data against geographical models

Clustering phenomenon

Noise



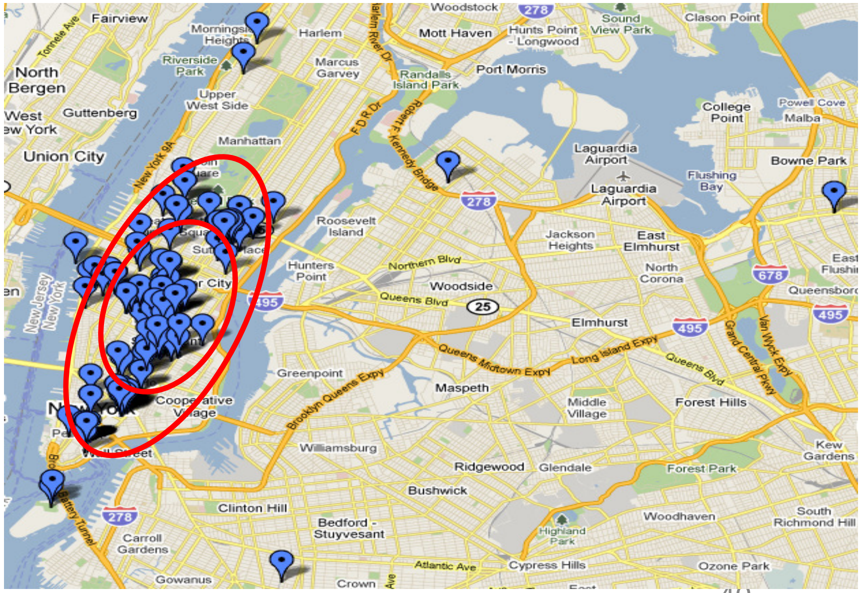
Gaussian Mixture Model (GMM)

- GMM

$$P(loc_i) \propto p(x_i) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x_i | \mu_k, \Sigma_k)$$

- Learning via maximum likelihood (ML)

$$\log p(X|\Theta)_{ML} = \sum_{i=1}^{n} \log p(x_i|\Theta)$$



Genetic Algorithm Based GMM

- GA-GMM
  - Trimmed likelihood estimate (TLE) method
  - Select a subset of the data to filter noise

$$\log p_{TLE}(X|\Theta) = \sum_{i=1}^{n} w_i \log p(x_i|\Theta)$$

$$\text{where } \forall i = 1, 2, ..., n, w_i \in \{0, 1\}$$

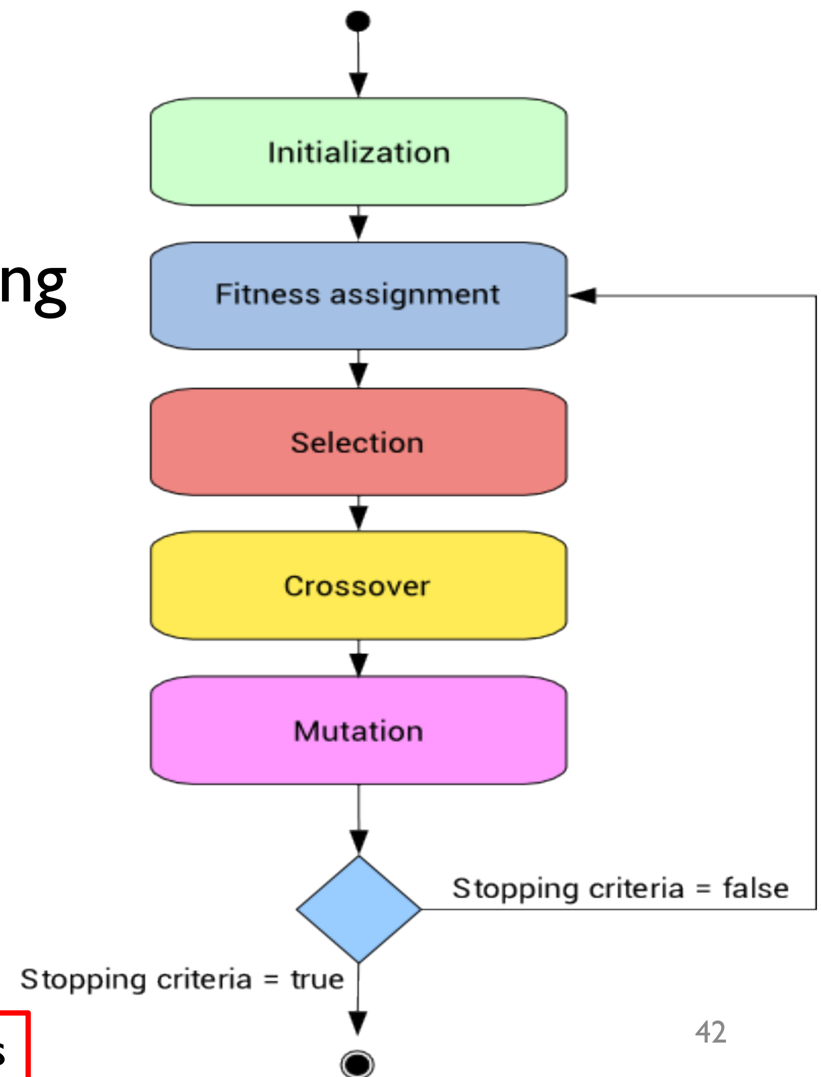
Genetic Algorithm Based GMM

- GA-GMM
  - Fitness function: TLE value
  - Coding scheme: binary string representing the data occurrence
  - Guided mutation: tend to increase TLE value
- Complexity

#Average check-ins of one user #centers

$$O(n * |P| * (mlogm + |C| * m * k) * T)$$

#Users Population size #EM cycles #iterations



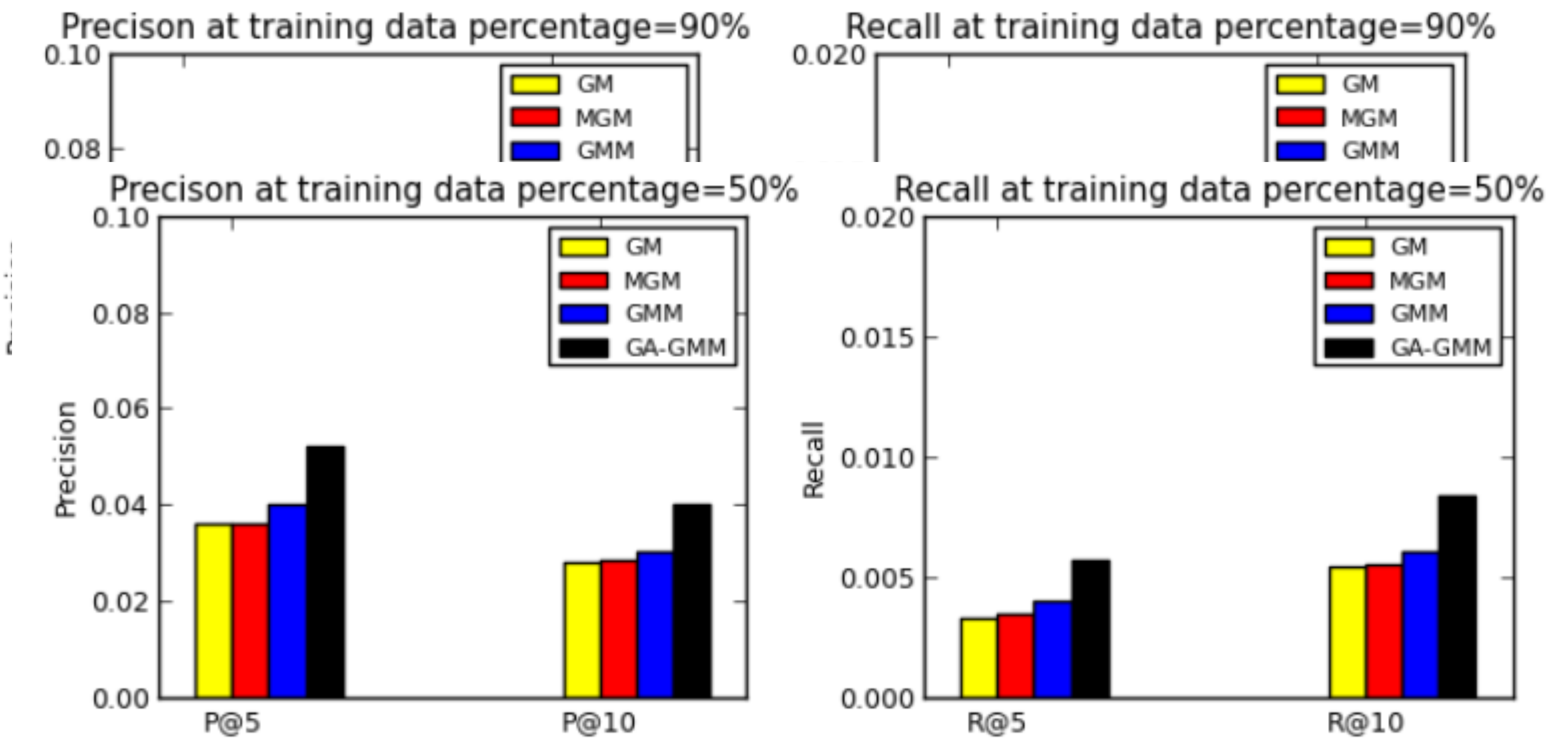
Experiment

- Data
  - Gowalla during Feb. 2009 to Sep. 2011
  - 3836 users and 183667 locations

- Baselines
  - Gaussian model (GM) [Cho et al., 2011]: modeling human movement as a stochastic process distributed around a center
  - Multi-center Gaussian model (MGM) [Cheng et al., 2012] : state-of-the-art geographical model for POI recommendation

Experimental Results

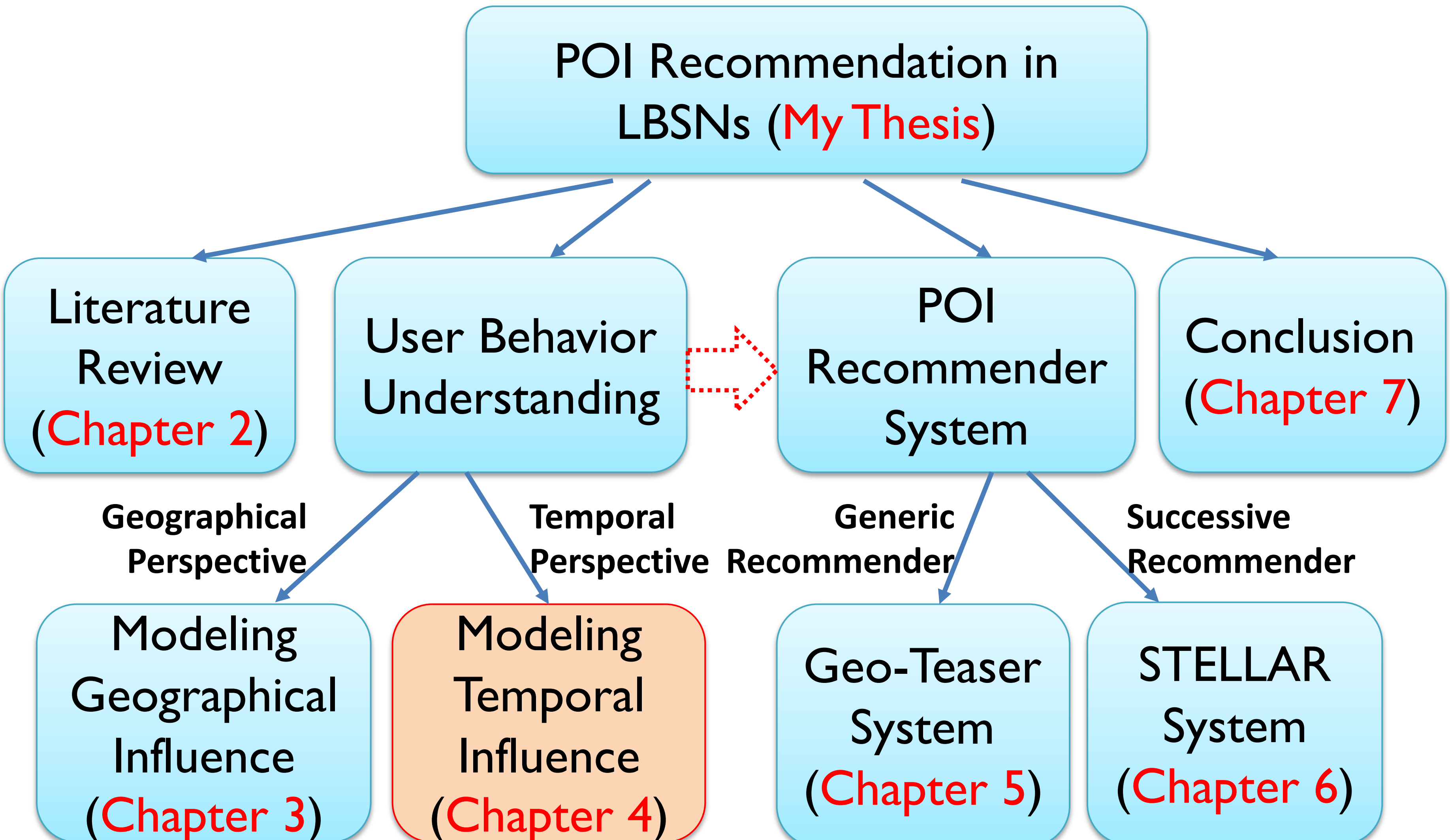
- Following [Cheng et al., 2012], two types of data separation: 90% and 50%



Chapter Contributions

- Model the geographical influence via GA-based GMM
- Model the human mobility better
- Better performance for POI recommendation than state-of-the-art methods

Thesis Structure

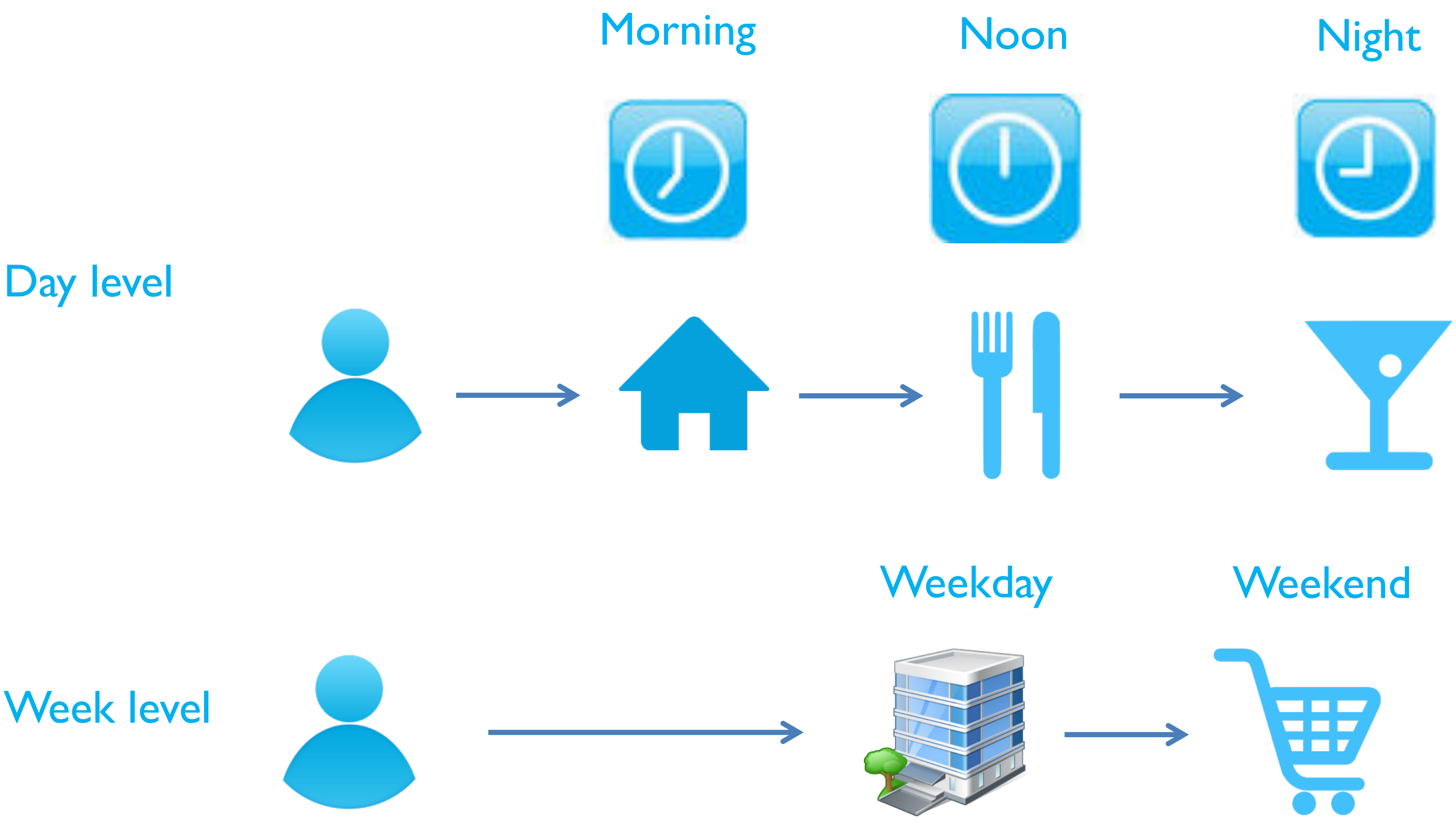


Motivations

- Temporal influence
  - Periodicity [cho et al., 2011, Yuan et al., 2013]
  - Consecutiveness [cheng et al., 2013, Gao et al., 2013]
  - Non-uniformness [Gao et al., 2013]

Motivations

- Temporal patterns happen at different scales

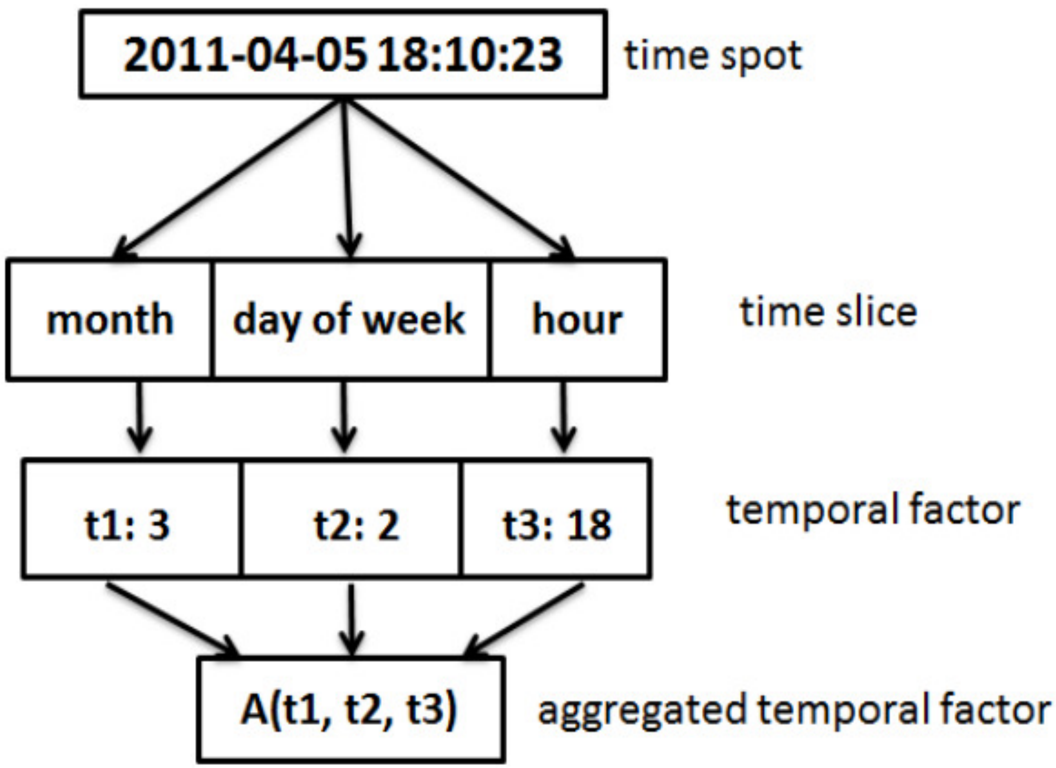


Model Description

- Aggregated Temporal Tensor Factorization (ATTF) model
  - Time labeling scheme
  - Formulate the preference score function
  - Infer the model via BPR criteria
  - Learn the model via SGD algorithm

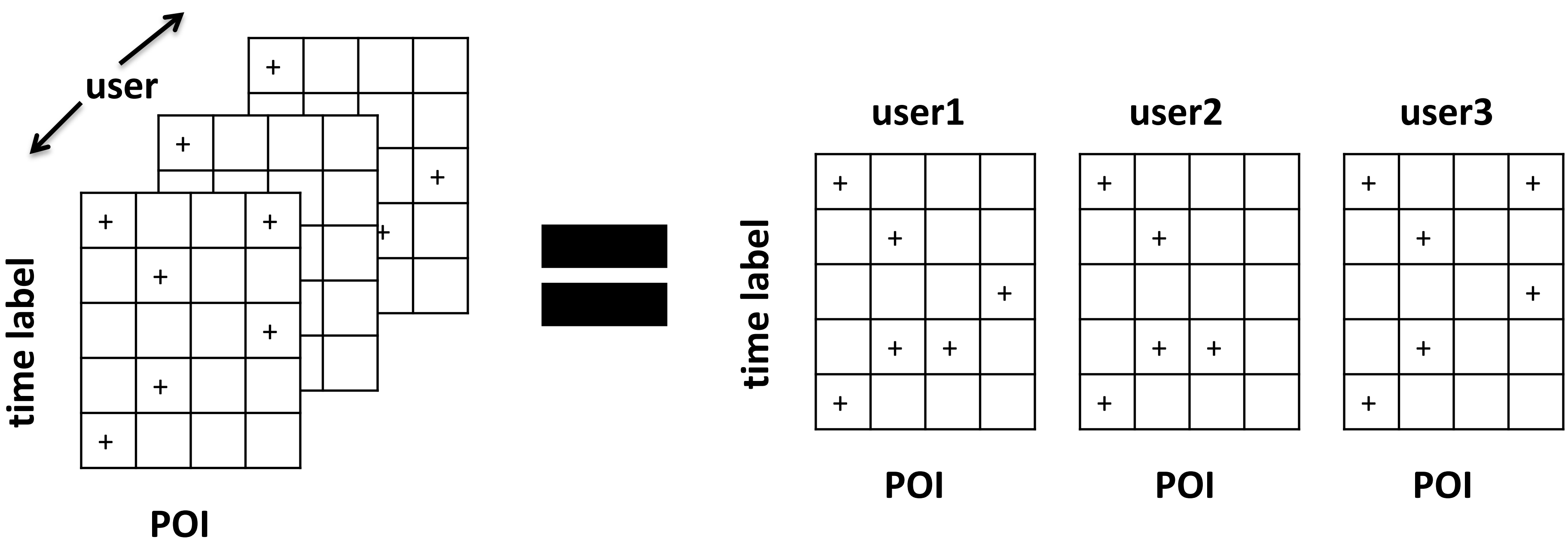
Time Labeling Scheme

- Represent time stamps with latent vectors



Tensor Construction

- Each time label is a tuple (t1, t2, t3) generated from the time labeling scheme



Model Formulation

- Score function formulation
  - Preference score of user u for a POI l given time t:

$$f(u,t,l) = \langle U_u^{(L)}, L_l^{(U)} \rangle + \langle A(T_{1,t_1}^{(L)}, T_{2,t_2}^{(L)}, T_{3,t_3}^{(L)}), L_l^{(T)} \rangle$$

- Aggregate operator definition

$$A(\cdot) = \alpha_1 \cdot T_{1,t_1}^{(L)} + \alpha_2 \cdot T_{2,t_2}^{(L)} + \alpha_3 \cdot T_{3,t_3}^{(L)}$$

Model Inference

- BPR criteria: user prefers the visited POIs than the unvisited

$$p(l_i >_{u,t} l_j) = \sigma(f(u,t,l_i) - f(u,t,l_j))$$

- Inferring the model via BPR criteria, we get the objective function

$$\arg\max_{\Theta} \prod_{(u,t,l_i,l_j)\in D_S} p(l_i >_{u,t} l_j)$$

$$D_S := \{(u,t,l_i,l_j) | l_i >_{u,t} l_j, u \in \mathcal{U}, t \in \mathcal{T}, l_i, l_j \in \mathcal{L}\}$$

Learning

- Parameter updating rules

$$U_u^{(L)} \leftarrow U_u^{(L)} + \gamma \cdot (\delta \cdot (L_{l_p}^{(U)} - L_{l_n}^{(U)}) - \lambda \cdot U_u^{(L)})$$
$$L_{l_p}^{(U)} \leftarrow L_{l_p}^{(U)} + \gamma \cdot (\delta \cdot U_u^{(L)} - \lambda \cdot L_p^{(U)})$$
$$L_{l_p}^{(T)} \leftarrow L_{l_p}^{(T)} + \gamma \cdot (\delta \cdot A(\cdot) - \lambda \cdot L_{l_p}^{(T)})$$
$$L_{l_n}^{(U)} \leftarrow L_{l_n}^{(U)} - \gamma \cdot (\delta \cdot U_u^{(L)} + \lambda \cdot L_n^{(U)})$$
$$L_{l_n}^{(T)} \leftarrow L_{l_n}^{(T)} - \gamma \cdot (\delta \cdot A(\cdot) + \lambda \cdot L_{l_n}^{(T)})$$
$$T_t^{(L)} \leftarrow T_t^{(L)} + \gamma \cdot (\delta \cdot \alpha \cdot (L_{l_p}^{(T)} - L_{l_n}^{(T)}) - \lambda \cdot T_t^{(L)})$$

- Complexity

  O(N*k*d) , where N is #training examples, d is the latent vector dimension, k is #sampled negative POIs

Experiment

- Data

Source	#users	#POIs	#check-ins
Foursquare	10,180	16,561	867,107
Gowalla	3,318	33,665	635,600

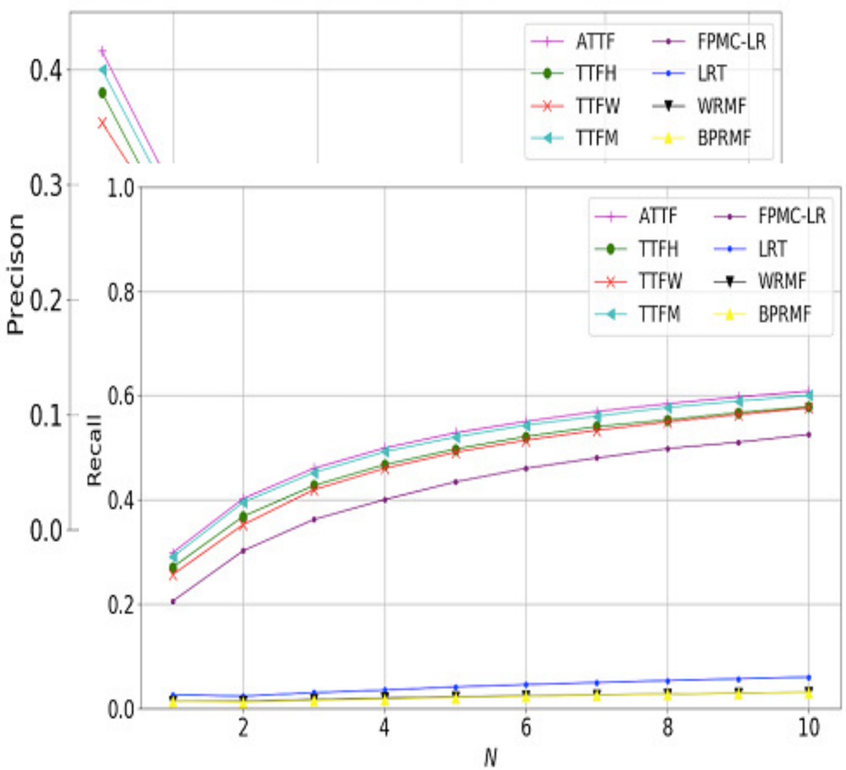
- Baselines
  - BRPMF [Rendel et al., 2009]
  - WRMF [Hu et al., 2008]

CF model for implicit feedback

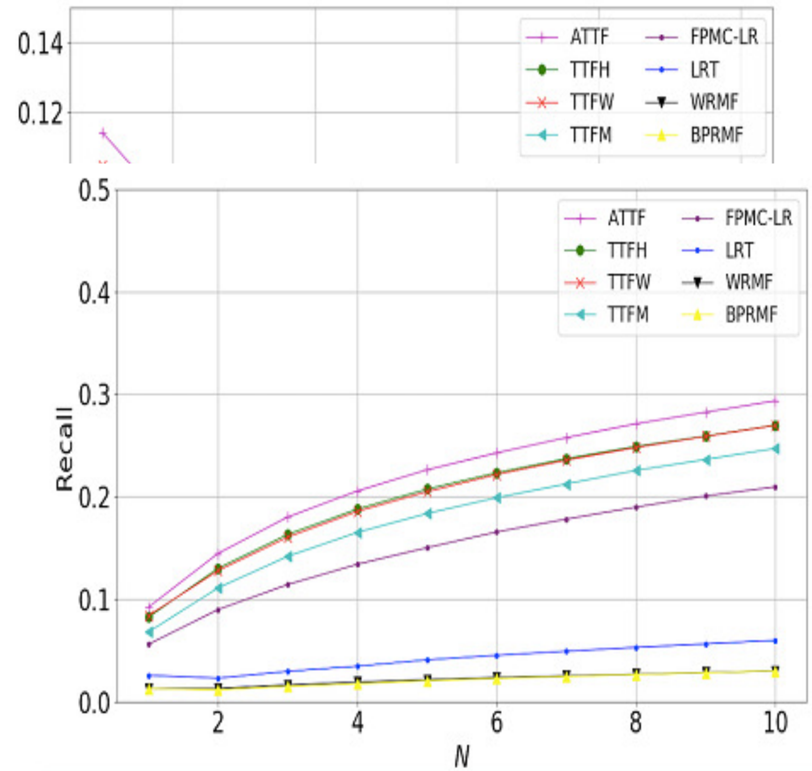
  - LRT [Gao et al., 2013]
  - FPMC-LR [Cheng et al., 2013]

POI recommendation model

Experimental Results



(a) Foursquare



(b) Gowalla

Recall on Foursquare and Gowalla

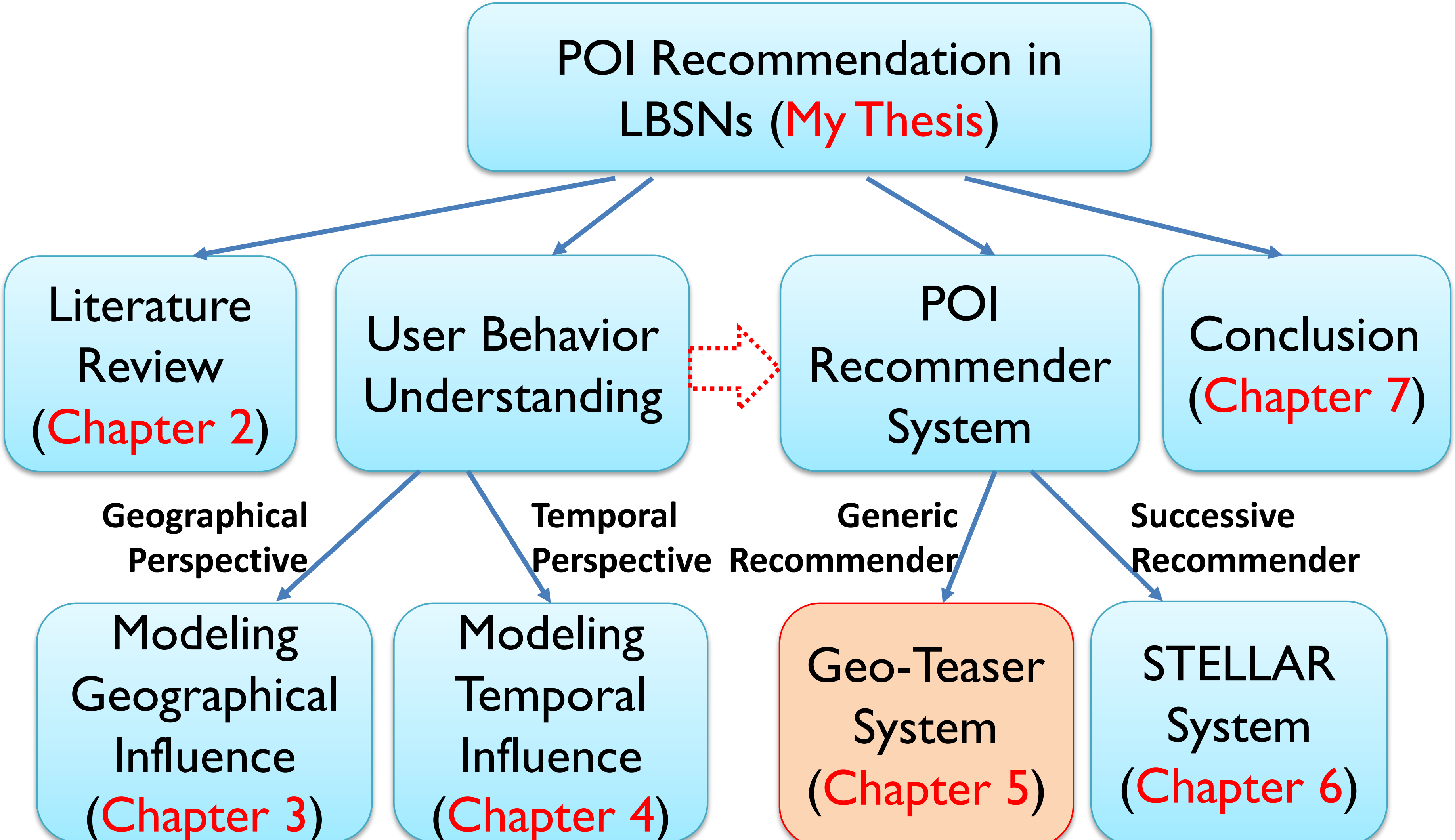
Chapter Contributions

- Propose ATTF model subsuming all the three temporal properties

- General framework to capture the temporal features at different scales

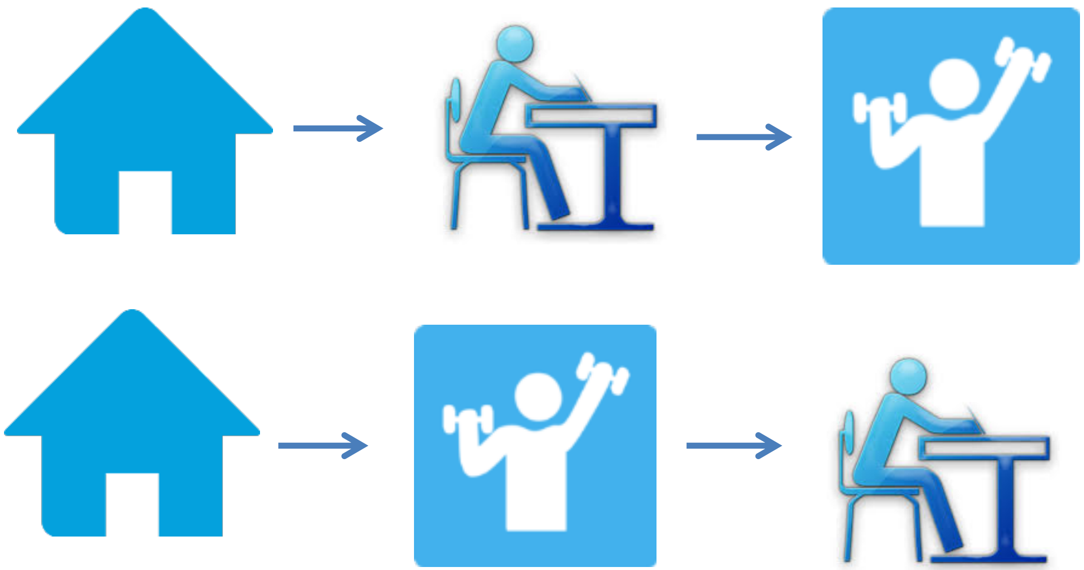
- Outperform prior temporal models

Thesis Structure



Motivations

- Observation 1

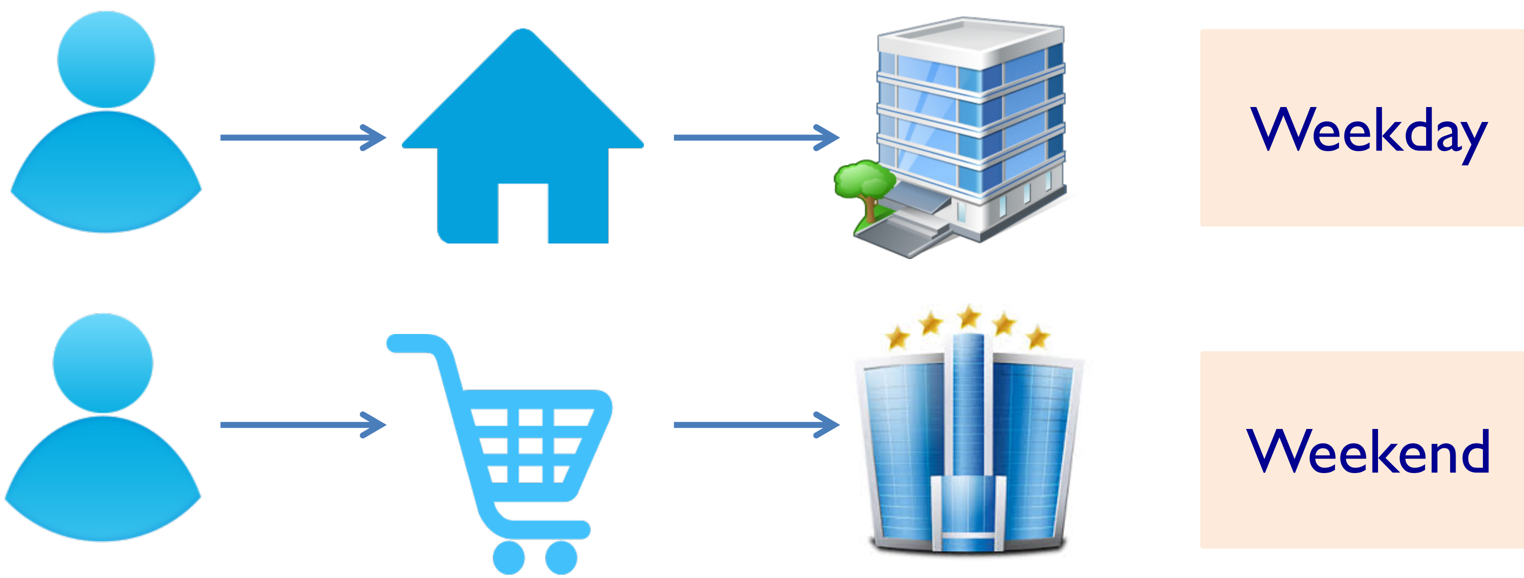


~~Markov chain~~

- Target
  - Capture contextual relations of POIs in the sequence

Motivations

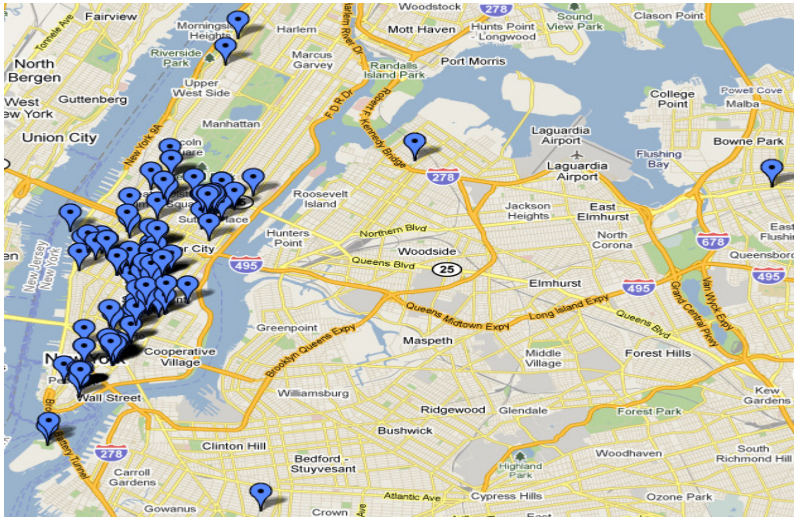
- Observation 2



- Target
  - Capture the temporal variance among sequences

Motivations

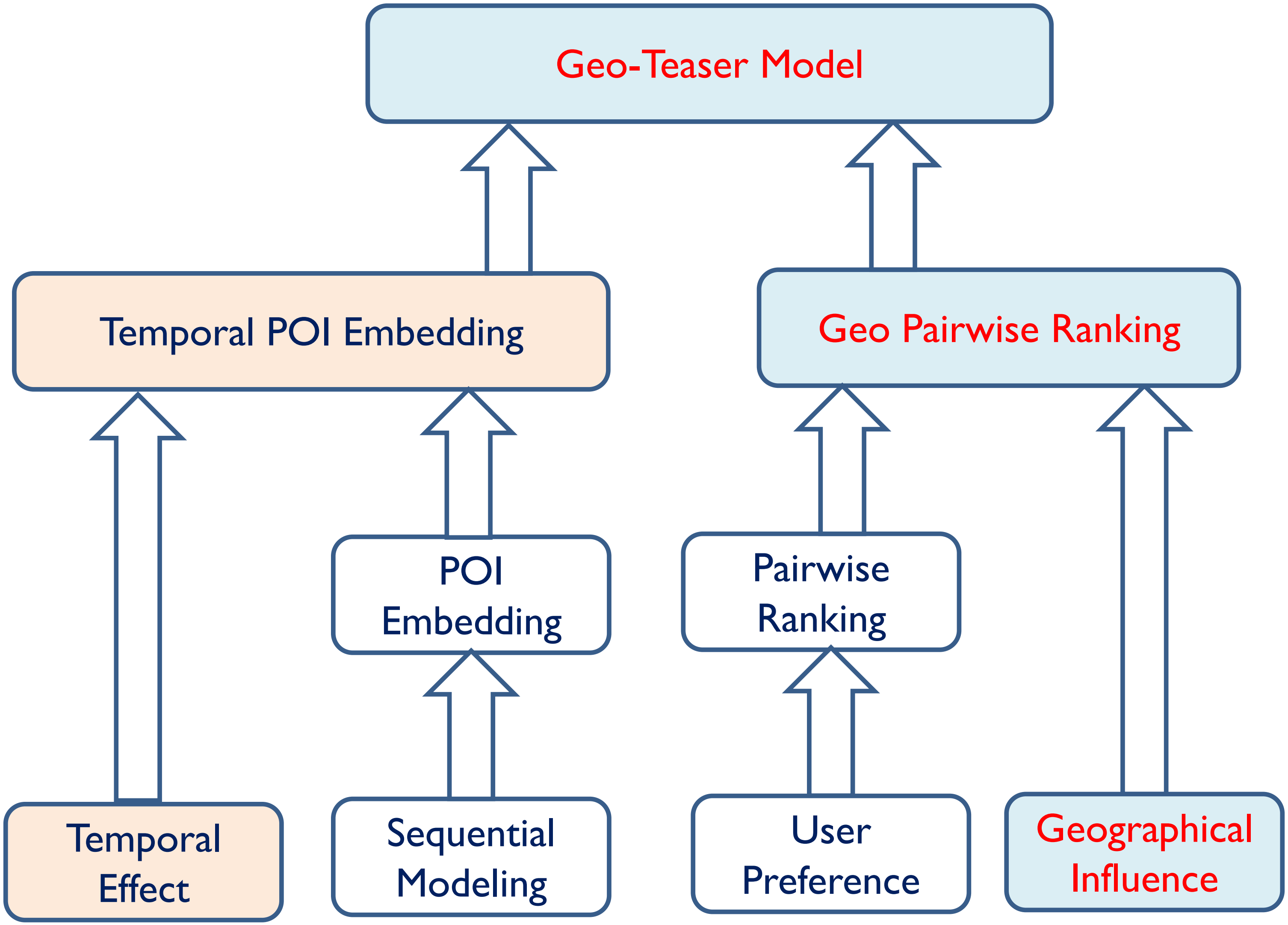
- Observation 3



Clustering phenomenon

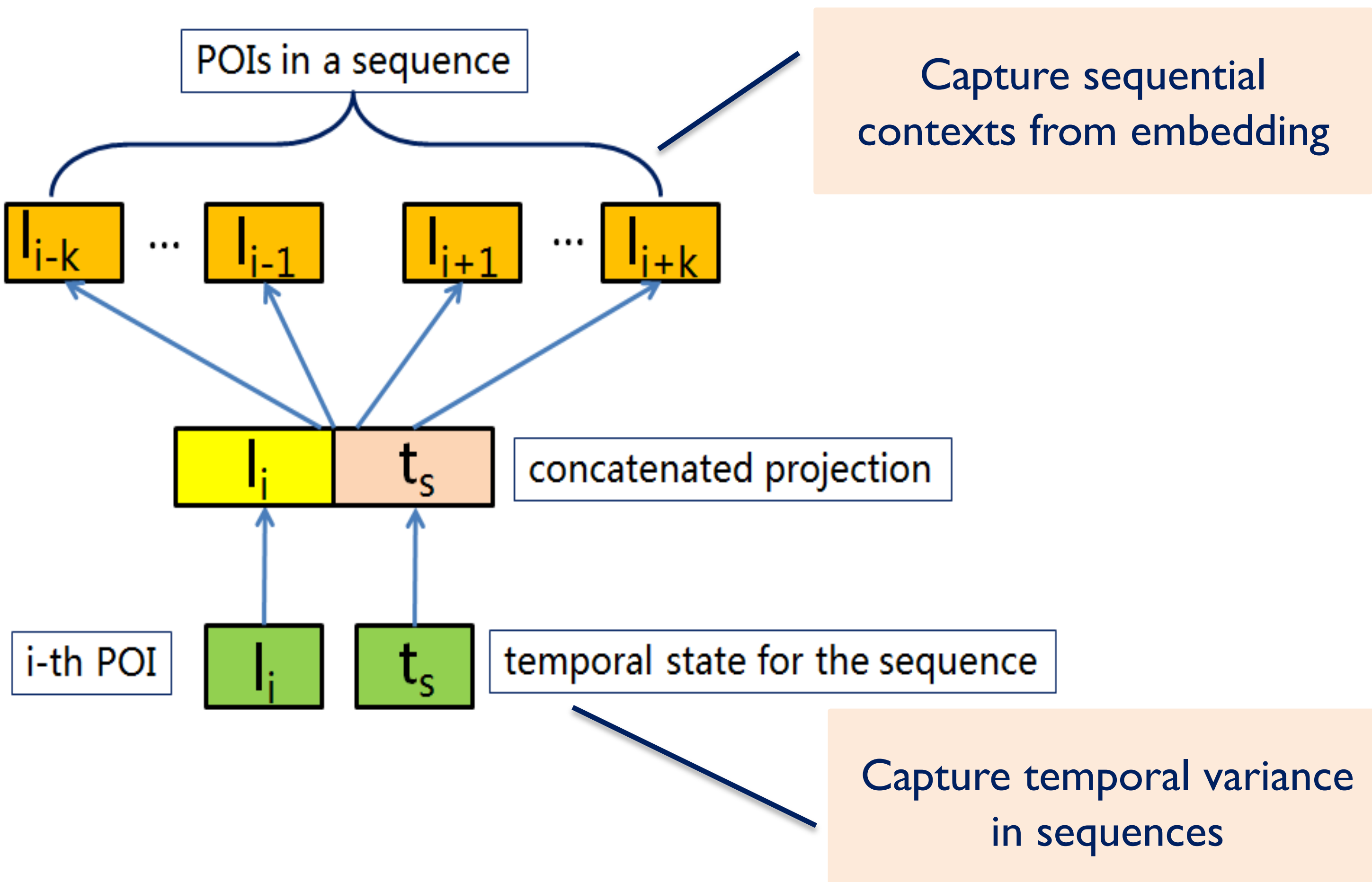
- Target
  - Capture the geographical influence

Proposed Model Framework



Temporal POI Embedding

- Temporal POI Embedding



Temporal POI Embedding

- Temporal POI Embedding

$$\mathcal{L}_{TPE} = \sum_{S_u \in S} \frac{1}{|S_u|} \sum_{l_i \in S_u} \sum_{-k \le c \le k, c \ne 0} \left(\log \Pr(l_{i+c} | l_i, t_s) \right)$$

Weekday, weekend

Softmax

Negative sampling

$$\mathcal{L}_{TPE} = \sum_{S_u \in S} \frac{1}{|S_u|} \sum_{l_i \in S_u} \sum_{-k \le c \le k, c \ne 0} \left(\log \sigma(\hat{\mathbf{l}}'_c \cdot \mathbf{l}^t_i) + \sum_h E_{k'} \log \sigma(-\hat{\mathbf{l}}'_{k'} \cdot \mathbf{l}^t_i) \right)$$

$$\boxed{\hat{\mathbf{l}}'_c = \mathbf{l}'_c \oplus \mathbf{l}'_c, \ \mathbf{l}^t_i = \mathbf{l}_i \oplus \mathbf{t}_s}$$

Geo Pairwise Ranking Model

- Geo pairwise ranking
  - Assumption: users prefer the POIs that are near the visited POIs
  - Target: discriminate the unchecked POIs
- Geo pairwise preference

Unchecked POI

$$l_i >_{u,d(l_i,l_{ne})\leq s} l_{ne} \;\vee\; l_{ne} >_{u,d(l_i,l_{nn})>s} l_{nn}$$

Checked POI

Distance

Distance threshold

Geo Pairwise Ranking Model

- Geo pairwise ranking formulation

$$\mathcal{L}_{GPR} = \sum_{S_u \in S} \sum_{(u,l_i,l_n) \in D_{S_u}} \log \sigma(\mathbf{u} \cdot (\mathbf{l}_i - \mathbf{l}_n))$$

Geo pairwise preference

$$\boxed{\begin{aligned} D_{S_u} = \{(u, l_i, l_{ne}) \vee (u, l_{ne}, l_{nn}) | l_i \in S_u, d(l_i, l_{ne}) \leq s, \\ d(l_i, l_{nn}) > s, l_{ne}, l_{nn} \in L \setminus L_u\}. \end{aligned}}$$

Geo-Teaser Model

- Objective function

$$\mathcal{O} = \arg\max_{\mathbf{U},\mathbf{L},\mathbf{T}} \sum_{S_u \in S} \sum_{l_i \in S_u} \Big(\sum_{-k \le c \le k, c \ne 0} \alpha \log \sigma(\hat{\mathbf{l}}'_c \cdot \mathbf{l}^t_i) + \sum_{h} \alpha E_{k'} \log \sigma(-\hat{\mathbf{l}}'_{k'} \cdot \mathbf{l}^t_i)$$

$$+ \sum_{D_{S_u}} \beta \log(\sigma(\mathbf{u} \cdot (\mathbf{l}_i - \mathbf{l}_n)))\Big)$$

Parameter to trade-off the sequential modeling and geo pairwise ranking

Model Learning

- SGD algorithm

$$\Theta^{t+1} = \Theta^{t} + \eta \times \frac{\partial \mathcal{O}(\Theta)}{\partial \Theta}$$

- Complexity

Window size

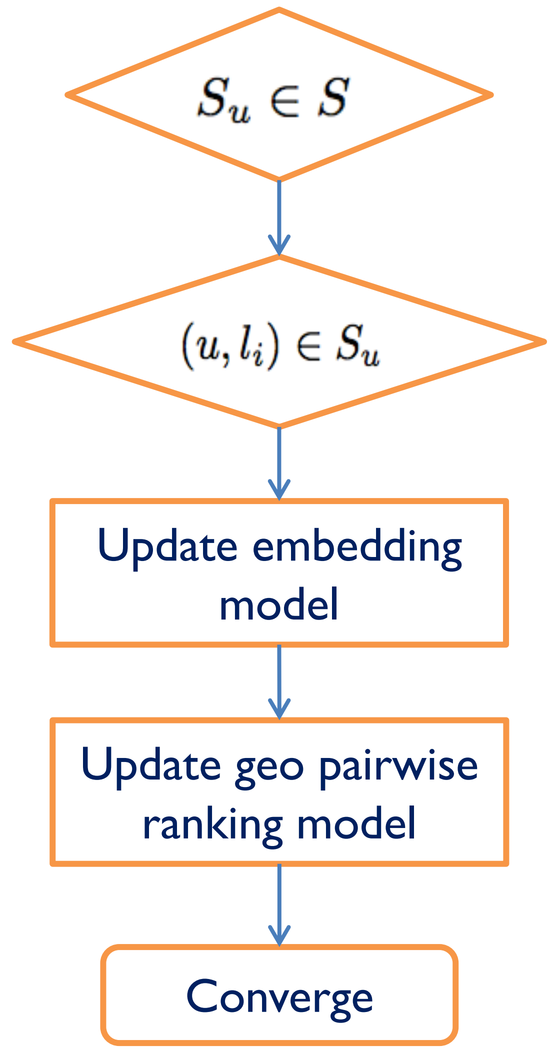
Vector dimension

All check-ins

$$O\left(\left(k \cdot h + m^2\right) \cdot d \cdot |C|\right)$$

Number of negative context POIs

Number of sampled unchecked POI



Experiment

- Data

Source	#users	#POIs	#check-ins
Foursquare	10,034	16,561	865,647
Gowalla	3,240	33,578	556,453

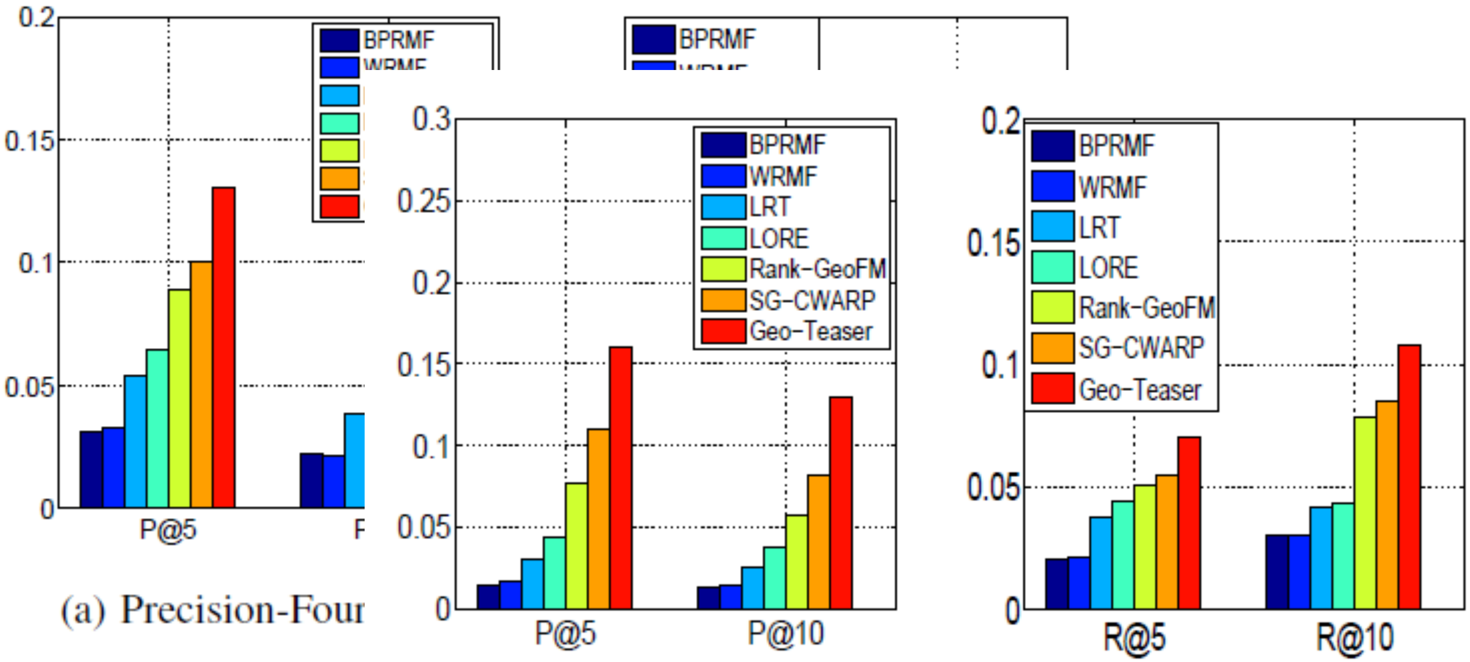
- Baselines
  - BRPMF [Rendel et al., 2009]
  - WRMF [Hu et al., 2008]

CF model for implicit feedback

  - LRT [Gao et al., 2013]
  - LORE [Zhang et al., 2014]
  - Rank-GeoFM [Li et al., 2015]
  - SG-CWARP [Liu et al., 2016]

POI recommendation model

Experimental Results



(c) Precision-Gowalla

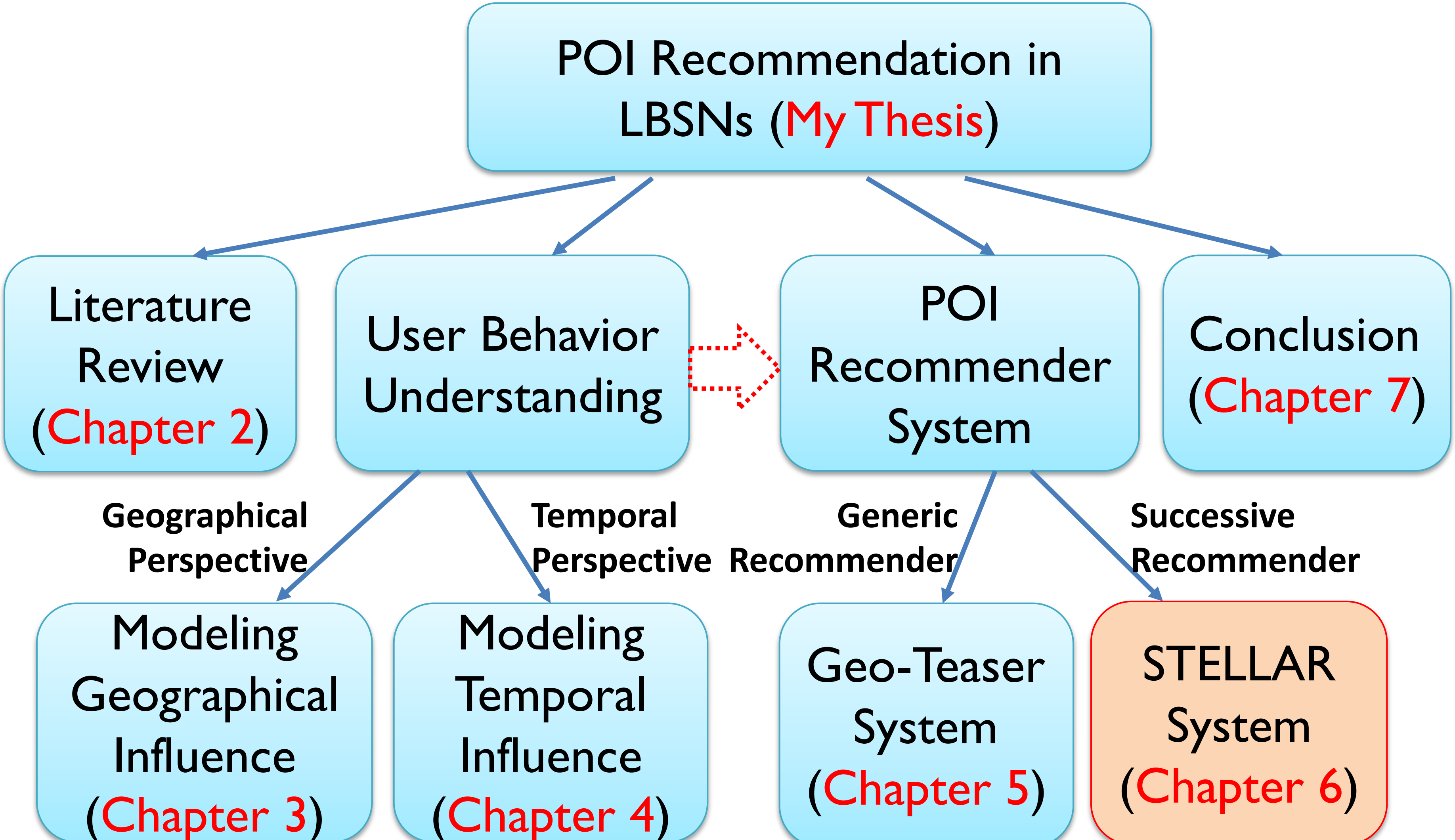
(d) Recall-Gowalla

- Geo-Teaser performs the best
- Good performance in SG-CWARP and Geo-Teaser: embedding learning works

Chapter Contributions

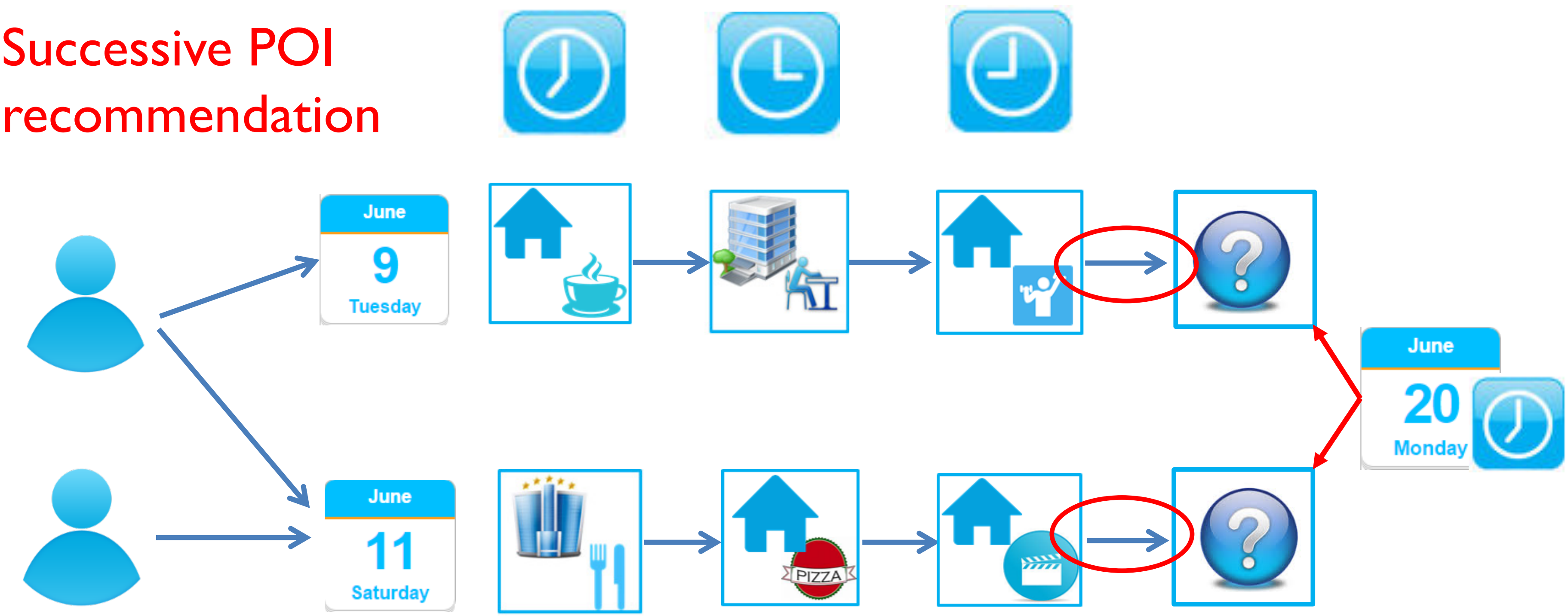
- Propose temporal POI embedding model to capture check-ins' sequential contexts
- Propose a new way to incorporate the geographical influence
- Propose the Geo-Teaser model as a unified framework incorporating sequential patterns, geographical and temporal influences

Thesis Structure



Motivations

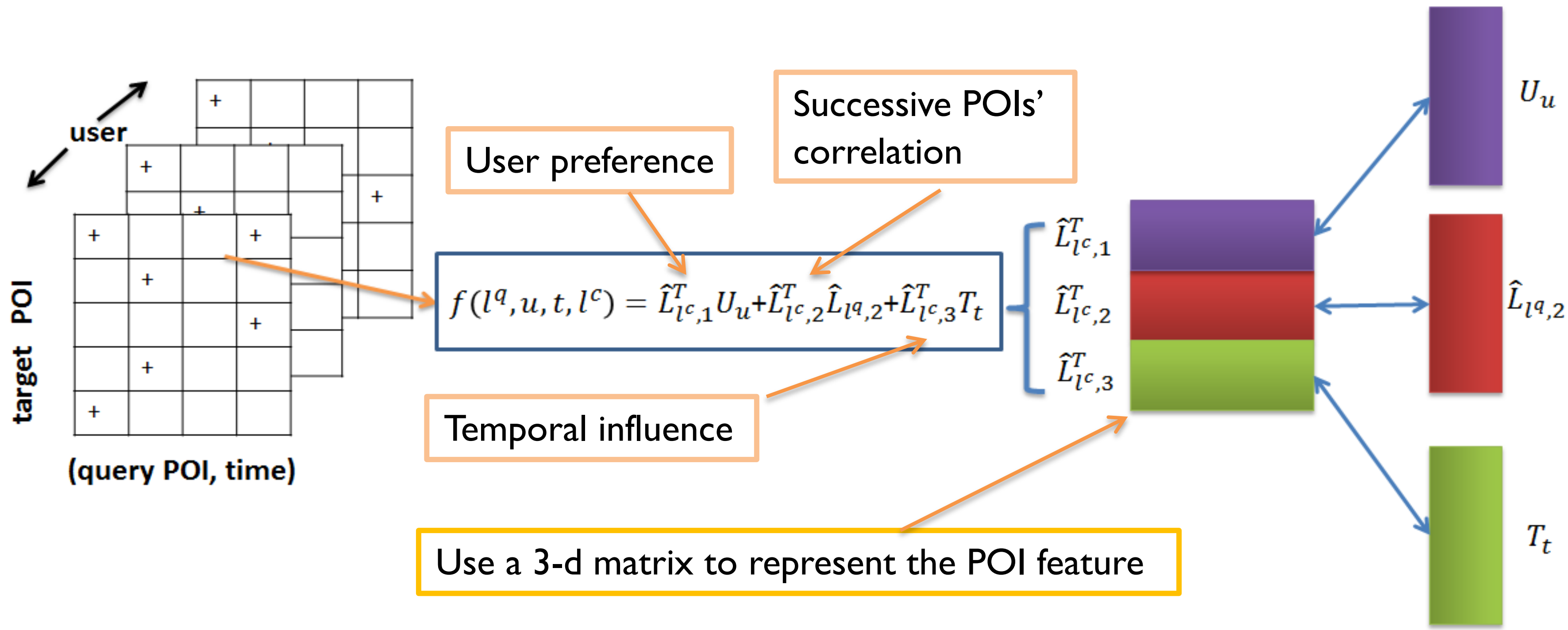
Successive POI recommendation



Successive POI recommendation is a time-subtle task

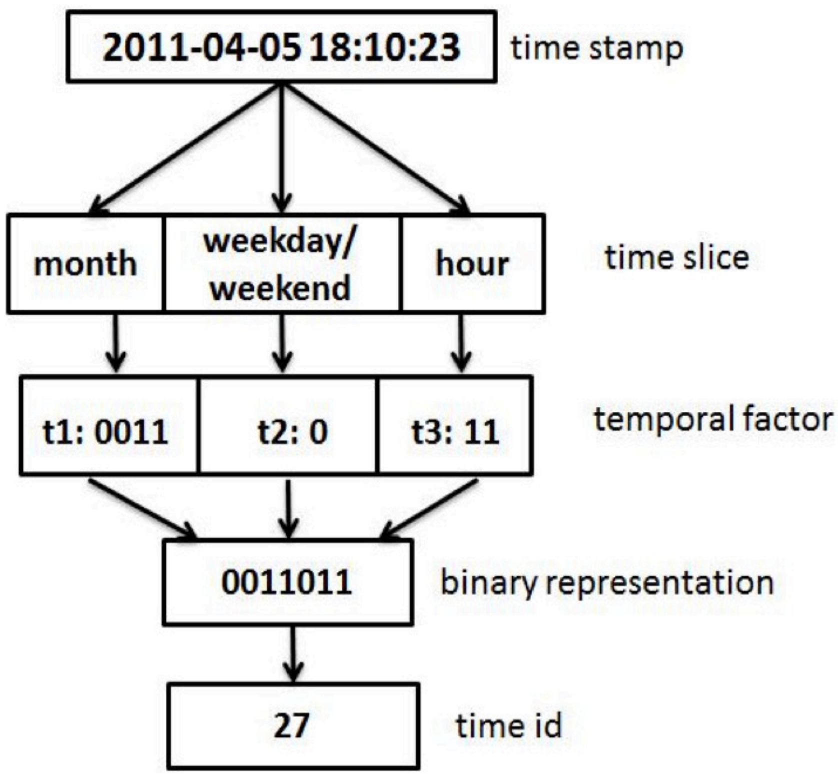
STELLAR Framework

- Spatial-Temporal Latent Ranking
  – Learn the preference score for a given user *u* to a candidate POI l^c at the time stamp *t* given his/her last check-in as a query POI l^q



Time Encoding Scheme

- Transform a time stamp to a unique id



Model Formulation

- Introduce interval-aware weight utility function

$$w = \begin{cases} 0.5 + \frac{2}{\Delta T} & \Delta T \geq s \\ 1 & otherwise \end{cases}$$

- New score function

$$f(u, l^q, t, l^c, w) = \hat{L}_{l^c,1}^T U_u + w \cdot \hat{L}_{l^c,2}^T \hat{L}_{l^q,2} + \hat{L}_{l^c,3}^T T_t$$

- Inference via BPR

$$\arg\min_{\Theta} \sum_{(u,l^q,t,l_p^c,l_n^c)\in D_S} - ln(\sigma(f(u, l^q, t, l_p^c) - f(u, l^q, t, l_n^c))) + \lambda||\Theta||_F^2$$

- Complexity

O(N*k*d) , where N is #training examples, d is the latent vector dimension, k is #sampled negative POIs

Experiment

- Data

Source	#users	#POIs	#check-ins
Foursquare	10,034	16,561	865,647
Gowalla	3,240	33,578	556,453

- Baselines
  - BRPMF [Rendel et al., 2009]
  - WRMF [Hu et al., 2008]

CF model for implicit feedback

  - LRT [Gao et al., 2013]
  - FPMC-LR [Cheng et al., 2013]

POI recommendation model

Experiment

- Model comparison

subset

		BPRMF	WRMF	LRT	FPMC−LR	TLAR	SLAR	STELLAR
Gowalla	P@5	0.025	0.031	0.033	0.048	0.053	0.050	**0.059**
	R@5	0.020	0.025	0.030	0.167	0.204	0.197	**0.226**
Foursquare	P@5	0.031	0.033	0.061	0.109	0.119	0.114	**0.129**
	R@5	0.027	0.028	0.053	0.347	0.373	0.368	**0.425**

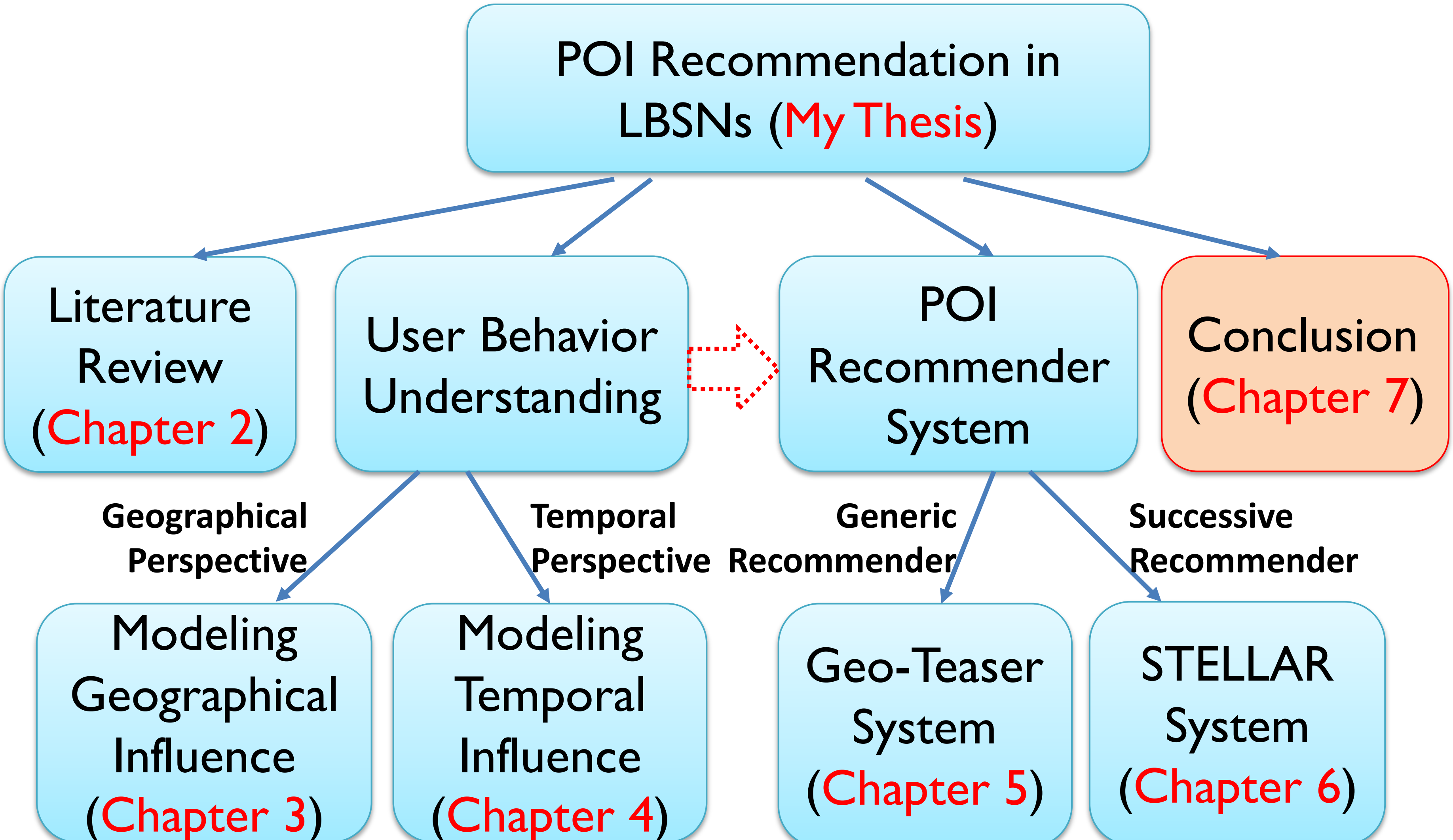
- Comparison of different time schemes

		M+W	M+D	W+D	M+W+D
Gowalla	P@5	0.051	0.053	0.054	**0.059**
	R@5	0.207	0.208	0.219	**0.226**
Foursquare	P@5	0.118	0.120	0.121	**0.129**
	R@5	0.371	0.389	0.398	**0.425**

Chapter Contributions

- Propose a time-aware successive POI recommendation method: STELLAR model
- Design a novel three-slice time indexing scheme to represent the time stamps
- Introduce a interval-aware weight utility function to differentiate successive check-ins' correlations

Thesis Structure



Conclusion

- A systematic literature review
- Understand the user behavior from geographical and temporal perspective
- Propose two systems: Geo-Teaser and STELLAR

Future Work

- Ranking Based Model
- Online Recommendation
- Deep Learning Based Recommendation

Ranking Based Model

Rating based model

estimate the value

Ranking based model

estimate preference order

$l_1 > l_k > l_{|\mathcal{L}|-1}$

	l_1	l_2	$\cdots$	l_k	$\cdots$	$l_{\lvert\mathcal{L}\rvert-1}$	$l_{\lvert\mathcal{L}\rvert}$
u_1	46	?	$\cdots$	2	$\cdots$	1	?
u_2	1	17	$\cdots$	5	$\cdots$	4	?
$\vdots$	?	1	$\vdots$	10	$\vdots$	?	4
u_k	1	?	$\cdots$	4	$\cdots$	1	?
$\vdots$	?	?	$\vdots$	?	$\vdots$	?	?
$u_{\lvert\mathcal{U}\rvert-1}$	?	?	$\cdots$	?	$\cdots$	1	?
$u_{\lvert\mathcal{U}\rvert}$	1	?	$\cdots$	?	$\cdots$	?	?

$l_2 > l_{|\mathcal{L}|}$?

Ranking Based Model

Rating based model < **Ranking based model**

estimate the value

estimate preference order

$l_1 > l_k > l_{|\mathcal{L}|-1}$

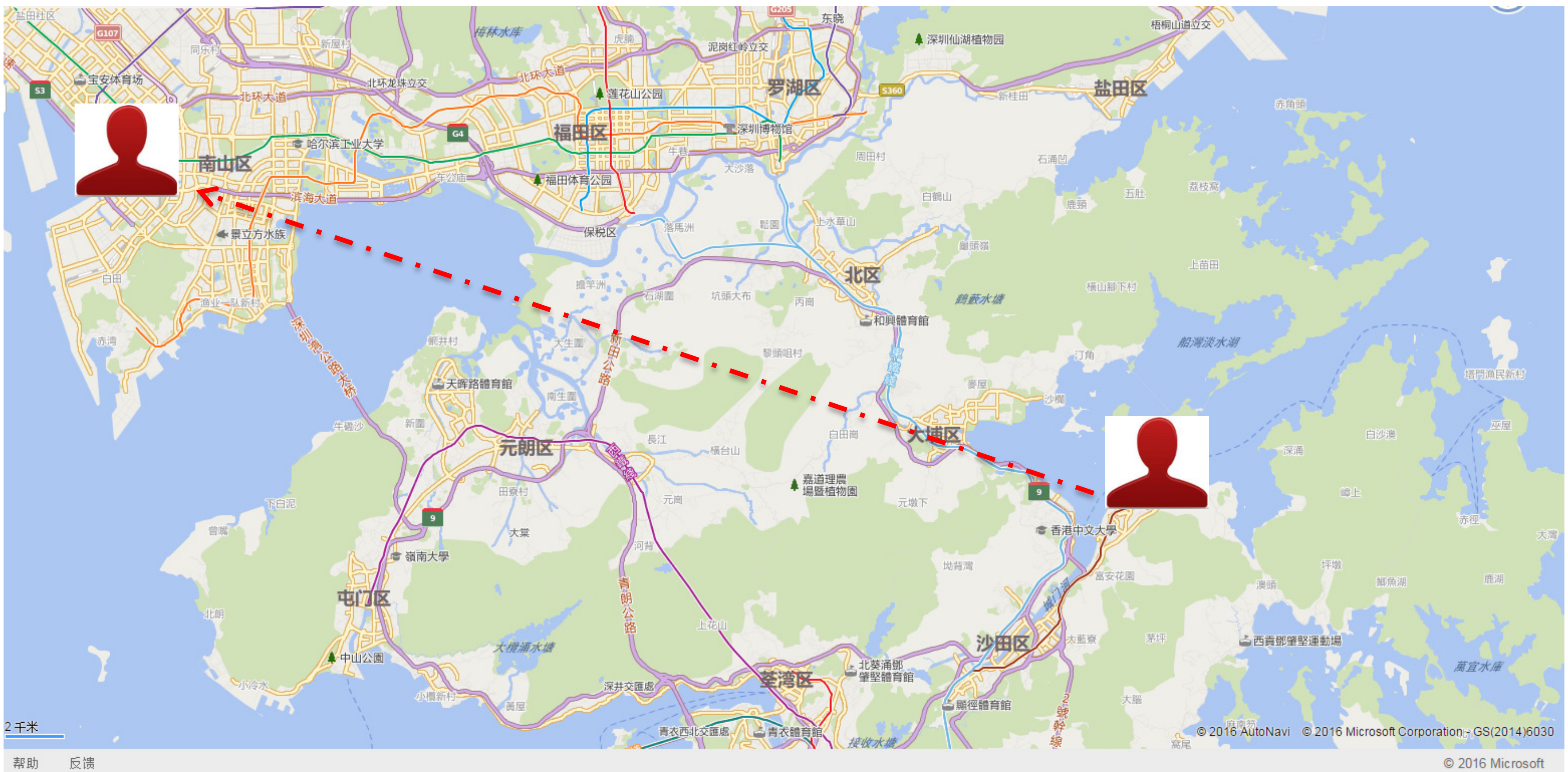
$l_2 > l_{|\mathcal{L}|}$?

	l_1	l_2	$\cdots$	l_k	$\cdots$	$l_{\mid\mathcal{L}\mid-1}$	$l_{\mid\mathcal{L}\mid}$
u_1	46	?	$\cdots$	2	$\cdots$	1	?
u_2	1	17	$\cdots$	5	$\cdots$	4	?
$\vdots$	?	1	$\vdots$	10	$\vdots$	?	4
u_k	1	?	$\cdots$	4	$\cdots$	1	?
$\vdots$	?	?	$\vdots$	?	$\vdots$	?	?
$u_{\mid\mathcal{U}\mid-1}$	?	?	$\cdots$	?	$\cdots$	1	?
$u_{\mid\mathcal{U}\mid}$	1	?	$\cdots$	?	$\cdots$	?	?

[Li et al., 2015, Zhao et al., 2016a, Zhao et al., 2016b, Liu et al., 2016b]

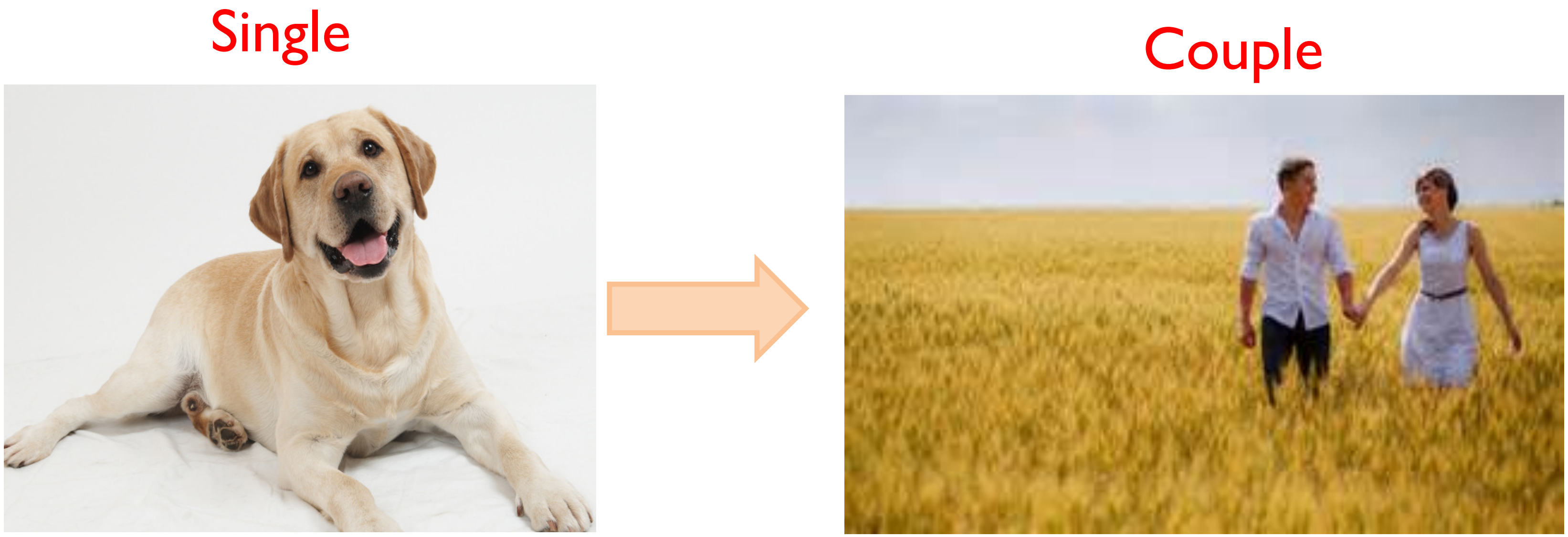
Online POI Recommendation

- Geographical characteristics change

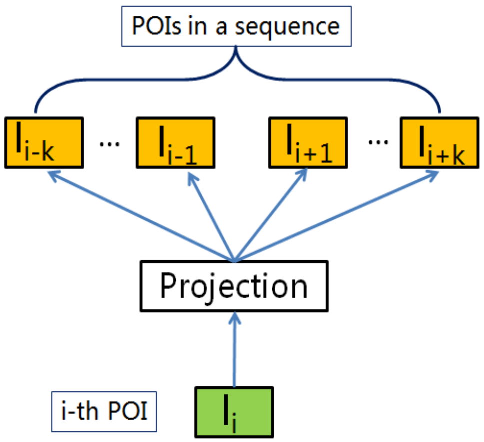


Online POI Recommendation

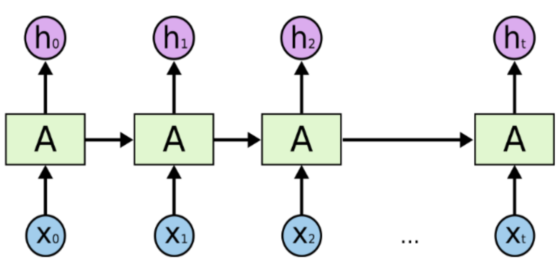
- Geographical characteristics change
- Check-in preferences change



Deep Learning Based Recommendation

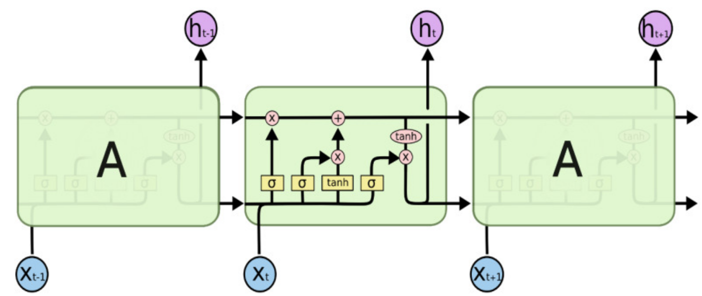


Skip-Gram



RNN

LSTM



[Liu et al., 2016a, Liu et al., 2016b]

http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Publications

Journal and Book chapter

[1] Shenglin Zhao, Michael R. Lyu, and Irwin King. "Aggregated Temporal Tensor Factorization Model for Point-of-interest Recommendation". Neural Processing Letters. (Chapter 4).

[2] Shenglin Zhao, Irwin King, and Michael R. Lyu. "A Survey on Point-of-interest Recommendation in Location-based Social Networks". ACM TWeb (Under Review). (Chapter 2).

[3] Shenglin Zhao. "Location-based Social Networks Analysis" in book "Encyclopedia of Social Network Analysis and Mining". (Accepted). (Chapter 1 and Chapter 2).

Conference

[4] Sheng Zhang, **Shenglin Zhao**, Mingxuan Yuan, Jia Zeng, Jianguo Yao, Irwin King, and Michael Lyu. "Traffic Prediction Based Power Saving in Cellular Networks: A Machine Learning Method". SIGSPATIAL 2017.

[5] Shenglin Zhao, Irwin King, and Michael R. Lyu. "Geo-Pairwise Ranking Matrix Factorization Model for Point-of-interest Recommendation". ICONIP 2017 (Best Paper Runner-up).

[6] Jiajun Cheng, **Shenglin Zhao**, Jiani Zhang, Irwin King, Xin Zhang, and Hui Wang. "Aspect-level Sentiment Classification with HEAT (*H*ierarchical *At*tention) Network". CIKM 2017.

[7] Shenglin Zhao, Michael R. Lyu, Irwin King, Jia Zeng, and Mingxuan Yuan. "Mining Business Opportunities from Location-based Social Networks". SIGIR 2017 (Short Paper).

Publications

[8] Shenglin Zhao, Tong Zhao, Irwin King, and Michael R. Lyu. "Geo-Teaser: Geo-Temporal Sequential Embedding Rank for Point-of-interest Recommendation". WWW 2017 (Cognitive Computing). (Chapter 5).

[9] Shenglin Zhao, Michael R. Lyu, and Irwin King. "Aggregated Temporal Tensor Factorization Model for Point-of-interest Recommendation". ICONIP 2016. (Chapter 4).

[10] Qi Xie, **Shenglin Zhao**, Zibin Zheng, Jieming Zhu, and Michael R. Lyu. "Asymmetric Correlation Regularized Matrix Factorization for Web Service Recommendation". ICWS 2016.

[11] Shenglin Zhao, Tong Zhao, Haiqin Yang, Michael R. Lyu, and Irwin King. "STELLAR: Spatial-Temporal Latent Ranking for Successive Point-of-Interest Recommendation". AAAI 2016 (AI and Web). (Chapter 6).

[12] Shenglin Zhao and Haiqin Yang. "Scalable Point-of-interest Recommendation via Geo-embedding Pairwise Matrix Factorization". WSDM 2015 workshop on Scalable Data Analytics. (Chapter 5).

[13] Shenglin Zhao, Irwin King, and Michael R. Lyu. "Capturing Geographical Influence in POI Recommendations". ICONIP 2013. (Chapter 3).

Selected References

- [Cheng et al., 2012] Chen Cheng, Haiqin Yang, Irwin King, and Michael R Lyu. Fused matrix factorization with geographical and social influence in location-based social networks. In Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, pages 17–23. AAAI Press, 2012.
- [Cheng et al., 2013] Chen Cheng, Haiqin Yang, Michael R Lyu, and Irwin King. Where you like to go next: successive pointof-interest recommendation. In Proceedings of the Twenty-Third international joint conference on Artificial Intelligence, pages 2605–2611. AAAI Press, 2013.
- [Cho et al., 2011] Eunjoon Cho, Seth A Myers, and Jure Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1082–1090. ACM, 2011.
- [Gao et al., 2013] Huiji Gao, Jiliang Tang, Xia Hu, and Huan Liu. Exploring temporal effects for location recommendation on location-based social networks. In *Proceedings of the 7th ACM conference on Recommender systems*, pages 93–100. ACM, 2013.
- [Gao et al., 2015] Huiji Gao, Jiliang Tang, Xia Hu, and Huan Liu. Content-aware point of interest recommendation on locationbased social networks. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 1721–1727. AAAI Press, 2015.
- [Lian et al., 2014] Defu Lian, Cong Zhao, Xing Xie, Guangzhong Sun, Enhong Chen, and Yong Rui. GeoMF: Joint geographical modeling and matrix factorization for point-of-interest recommendation. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 831–840. ACM, 2014.
- [Ma et al., 2011] Ma H, Zhou D, Liu C, et al. Recommender systems with social regularization. Proceedings of the fourth ACM international conference on Web search and data mining. ACM, 2011: 287-296.

Selected References

- [Liu et al., 2013] Bin Liu, Yanjie Fu, Zijun Yao, and Hui Xiong. Learning geographical preferences for point-of-interest recommendation. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1043–1051. ACM, 2013.
- [Yang et al., 2013] Dingqi Yang, Daqing Zhang, Zhiyong Yu, and Zhu Wang. A sentiment-enhanced personalized location recommendation system. In *Proceedings of the 24th ACM Conference on Hypertext and Social Media*, pages 119–128. ACM, 2013.
- [Ye et al., 2010] Mao Ye, Peifeng Yin, and Wang-Chien Lee. Location recommendation for location-based social networks.In Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems, pages 458–461. ACM, 2010
- [Ye et al., 2011] Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, pages 325–334. ACM, 2011.
- [Rendle et al., 2009] Rendle, Steffen, et al. "BPR: Bayesian personalized ranking from implicit feedback." Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence. AUAI Press, 2009.
- [Hu et al., 2008] Hu, Yifan, Yehuda Koren, and Chris Volinsky. "Collaborative filtering for implicit feedback datasets." 2008 Eighth IEEE International Conference on Data Mining. Ieee, 2008.

Selected References

- [Yuan et al., 2013] Quan Yuan, Gao Cong, Zongyang Ma, Aixin Sun, and Nadia Magnenat Thalmann. Time-aware point-ofinterest recommendation. In roceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, pages 363–372. ACM, 2013.
- [Zhang et al., 2013] Jia-Dong Zhang and Chi-Yin Chow. igslr: personalized geo-social location recommendation: a kernel density estimation approach. In Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pages 334–343. ACM, 2013.
- [Zhang et al., 2015] Jia-Dong Zhang and Chi-Yin Chow. Geosoca: Exploiting geographical, social and categorical correlations for point-of-interest recommendations. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 443–452. ACM, 2015.
- [Zhao et al., 2013] Shenglin Zhao, Irwin King, and Michael R Lyu. Capturing geographical influence in poi recommendations. In International Conference on Neural Information Processing, pages 530–537. Springer, 2013.
- [Zhao et al., 2016] Shenglin Zhao, Tong Zhao, Haiqin Yang, Michael R Lyu, and Irwin King. Stellar: Spatial-temporal latent ranking for successive point-of-interest recommendation. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.
- [Liu et al, 2016a] Liu X, Liu Y, Li X. Exploring the Context of Locations for Personalized Location Recommendations, IJCAI. 2016: 1188-1194.
- [Liu et al., 2016b] Liu, Qiang, et al. "Predicting the Next Location: A Recurrent Model with Spatial and Temporal Contexts." AAAI. 2016.

Thanks!

Q & A

Supplementary Slides

FAQ

- 1. Data set selection reason
- Answer: There are two kinds of data: city-based and universal. Yelp usually provide the city-based data. Our work aims to construct a global model for all users. We choose the universal data from Gowalla and Foursquare. The data are crawled in 2011. Now due to the privacy issue, a user's sequential check-ins are not allowed to attain.

- 2. Generic POI recommendation v.s. successive POI recommendation
- Answer: The difference lies in the successive POI recommendation suggests POIs given the current check-in. Generic POI recommendation does not have this constraint, which is similar to traditional movie recommendation---report suggestions given all prior records.

- 3. Relation between POI recommendation and traditional recommendation

FAQ

- Answer: We have the user-POI matrix with the following map relations: user---user, POI---movie, check-in frequency---rating. Based on this user-POI matrix, we can use CF methods for POI recommendation.

	l_1	l_2	$\cdots$	l_k	$\cdots$	$l_{\vert\mathcal{L}\vert-1}$	$l_{\vert\mathcal{L}\vert}$
u_1	46	?	$\cdots$	2	$\cdots$	1	?
u_2	1	17	$\cdots$	5	$\cdots$	4	?
$\vdots$	?	1	$\vdots$	10	$\vdots$	?	4
u_k	1	?	$\cdots$	4	$\cdots$	1	?
$\vdots$	?	?	$\vdots$	?	$\vdots$	?	?
$u_{\vert\mathcal{U}\vert-1}$	?	?	$\cdots$	?	$\cdots$	1	?
$u_{\vert\mathcal{U}\vert}$	1	?	$\cdots$	?	$\cdots$	?	?

FAQ

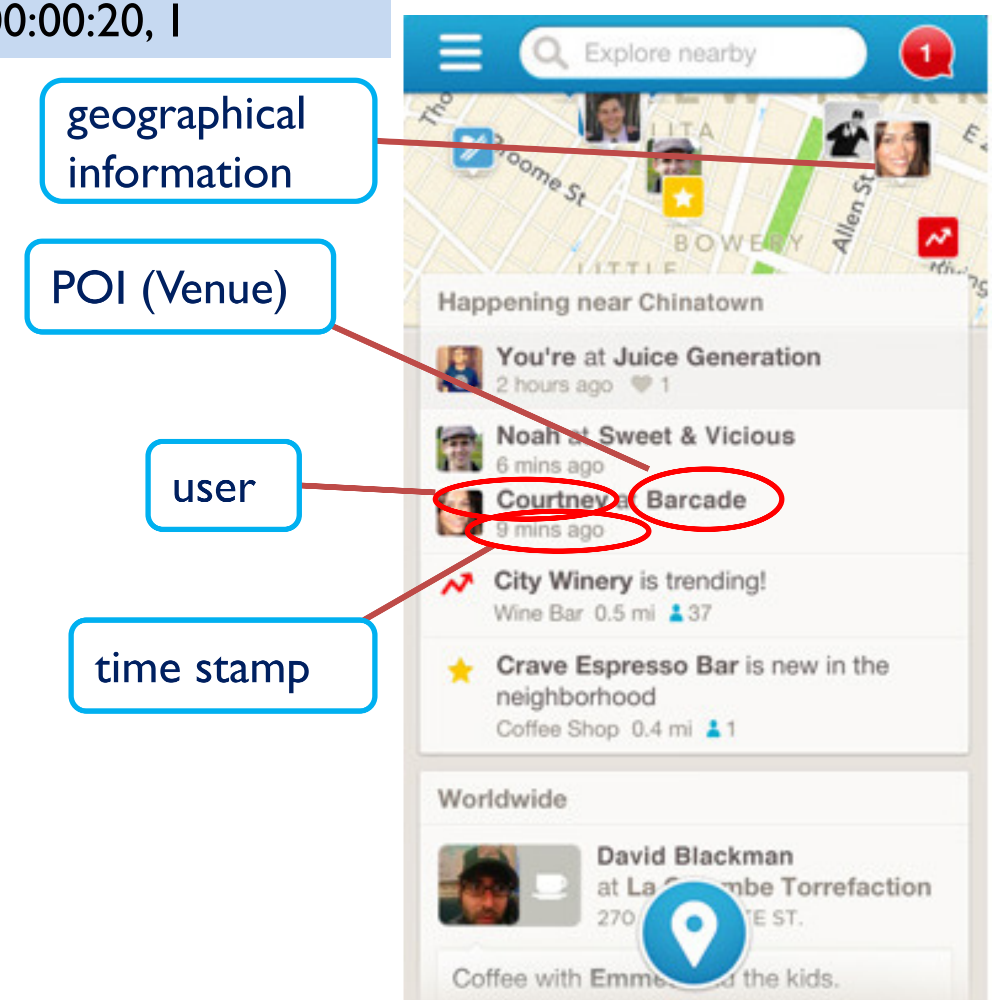
- Why the accuracy is very low?
- Answer: The big challenge is the data sparsity. Even in a city, we have hundreds of thousands of POIs, but each user only check-ins at hundred of POIs, even less than one hundred.

- How to improve the accuracy to make the application useful in industry?
- Answer: First, filter the few visited POIs. According to the long tail effect, a lot of POIs visited less than 5 users. Including these POIs does harm to the model. This method has been used in our recent paper for data preprocessing. Second, add constraints to specify the needs. In real scenarios, we have more information, such as the time and current check-in. This method has been used in our ATTF model (chapter 4) and STELLAR system (chapter 6), which really improve the accuracy, especially the recall.

- Challenge of the metrics.
- Answer: They are data sensitive.

Data Format

user_id, latitude, longitude, time, location_id
0, 41.72757566,- 88.03198814, 2011-01-01 00:00:01, 0
1, 51.31791, -0.588761, 2011-01-01 00:00:20, 1



Metrics

- Precision and recall

$$P@N = \frac{1}{|U|} \sum_{u \in U} \frac{|L_{visited} \cap L_{N,rec}|}{N}$$

$$R@N = \frac{1}{|U|} \sum_{u \in U} \frac{|L_{visited} \cap L_{N,rec}|}{|L_{visited}|}$$

Literature Review

Kernel Density Estimation

- Step 1:
  - Sample a check-in set for a user and compute the density function over the distance

Gaussian Kernel

Pair distance of visited POIs

$$f(d) = \frac{1}{|X_u|\sigma} \sum_{d' \in X_u} K(\frac{d - d'}{\sigma})$$

Distance

Smoothing parameter

Check-in set

[Zhang et al., 2013, Zhang et al., 2015]

Kernel Density Estimation

- Step 2:
  - Recommend POIs according the distance

$$p(l_j|L_u) = \frac{1}{|L_u|} \sum_{l_i \in L_u} f(d_{ij})$$

Candidate POI

Visited POI set

Distance between l_i and l_j

[Zhang et al., 2013, Zhang et al., 2015]

Fused Model

Representative model: MGM-PMF [Cheng et al., 2012]

$$P(l|C_u) = \sum_{c_u=1}^{|C_u|} P(l \in c_u) \frac{f_{c_u}^{\alpha}}{\sum_{i \in C_u} f_i^{\alpha}} \frac{N(l|\mu_{C_u}, \Sigma_{C_u})}{\sum_{i \in C_u} N(l|\mu_i, \Sigma_i)}$$

MGM

$$P(l|C_u)$$

PMF

$$P(F_{ul}) = \sigma(U_i^T L_j)$$

$$P_{ul} = P(F_{ul}) \cdot P(l|C_u)$$

$$\min_{U,L} \sum_{i=1}^{N} \sum_{j=1}^{M} I_{ij}(g(c_{ij}) - g(U_i^T L_j))^2 + \lambda_1 ||U||_F^2 + \lambda_2 ||L||_F^2$$

Joint Model

Representative model of MF-based joint model: GeoMF[Lian et al., 2014]

User latent matrix

User activity area latent matrix

$$\underset{P,Q,X}{\arg\min} ||W \odot (R - PQ^T - XY^T)||_F^2 + \gamma(||P||_F^2 + ||Q||_F^2) + \lambda||X||_1$$

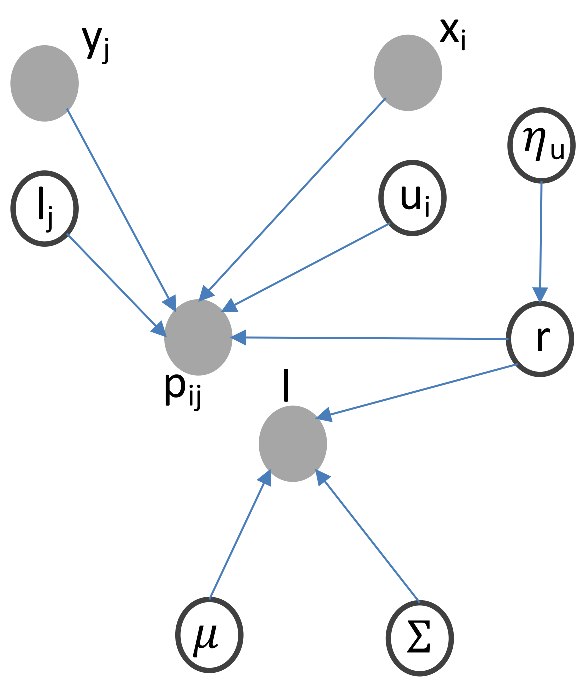
POI latent matrix

POI influence area latent matrix

Sparsity constraint for limited user activity areas

Joint Model

Representative model of generative graphical model: Geo-PFM[Liu et al., 2013]



Algorithm 1 Model generative process

1: Draw a region $r \sim \text{Multinomial}(\eta_u)$
2: Draw a location $l \sim \mathcal{N}(\mu_r, \sum_r)$
3: Draw a user preference
4: Generate user latent factor $\mathbf{u}_i \sim P(u_i; \Phi_\mathbf{u})$
5: Generate POI latent factor $\mathbf{l}_j \sim P(\mathbf{l}_j; \Phi_{\mathbf{l}_j})$
6: User-item preference $\alpha(i,j) = \mathbf{u}_i^T \mathbf{l}_i + x_i^T W y_j$
7: Generate $p_{ij} \sim P(f_{ij})$, where
$p_{ij} = (\mathbf{u}_i^T \mathbf{l}_j + x_i^T W y_j) \rho_j (d_0 + d(u_i, l_j))^{-\tau}$

Joint Model

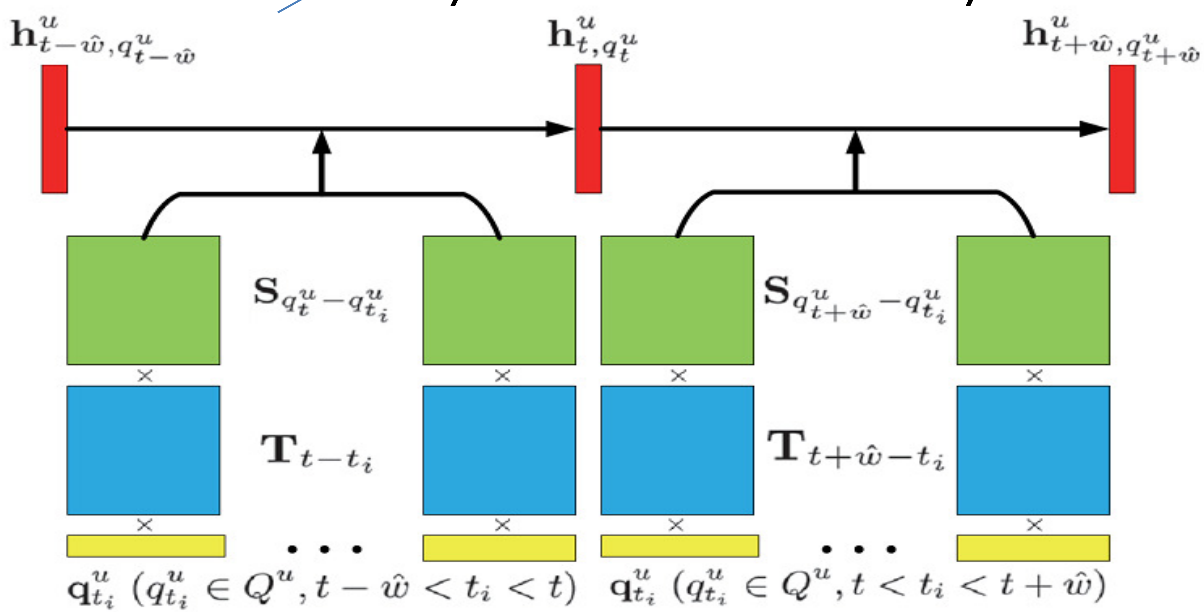
- Representative model of neural network model: ST-RNN[Liu et al., 2016]

User latent matrix

POI latent matrix

$$o_{u,t,l} = (\mathbf{h}^{u}_{t,q_l} + \mathbf{p}_u)^T \mathbf{q}_l$$

Dynamic interest modeled by RNN



GA-GMM

Term Definitions

Definition 1. *Encoding scheme. The chromosome is encoded into a binary string and each bit represents the existence of corresponding observed data. Each chromosome and its corresponding mixture model will be a possible solution to our problem.*

Definition 2. *Fitness function. The fitness score function is set as the trimmed logarithm likelihood of the corresponding GMM of a chromosome*—$\log p_{TLE}(X|\Theta)$.

Definition 3. *Guided Mutation. Guided Mutation ensures the chromosome in a population to mutate toward maximizing fitness score. It means we choose chromosome that has higher value fitting trained GMM.*

GA-GMM Alg.

Algorithm 1 Genetic-based Expectation Maximization Algorithm

t=0;
Initialize $P_0(t)$;
for $t = 1 : G$ **do**
  $P_1(t) \leftarrow$ perform several cycles of EM on $P_0(t)$;
  $P_2(t) \leftarrow$ Guided Mutation in $P_1(t)$;
  $fScore_2 \leftarrow$ evaluate $P_2(t)$;
  $P_0(t)^{'} \leftarrow$ selection and crossover to generate offspring from $P_2(t)$;
  $P_1(t)^{'} \leftarrow$ perform several cycles of EM on $P_0(t)^{'}$;
  $P_2(t)^{'} \leftarrow$ Guided Mutation in $P_1(t)^{'}$;
  $fScore_2^{'} \leftarrow$ evaluate $P_2(t)^{'}$;
  $P_3(t) \leftarrow$ selection from $[P_2(t), P_2(t)^{'}]$;
  $iBest \leftarrow$ best individual from $P_3(t)$;
  if $iBest$ satisfies convergence condition **then** break;
  $P_0(t+1) \leftarrow P_3(t)$;
  $t = t + 1$;
Perform EM on $iBest$ until convergence;

Complexity

Parameter Settings

The number of components in GMM and GA-GMM is set as 2. We set the radius of region in MGM as 1 kilometers and the ratio of one centre to whole record is set as 0.1. For GA-GMM, we set population size $|P| = 6$, EM cycles $|C| = 4$, and discard rate $\epsilon = 1/\sqrt{n}$.

Computation Complexity

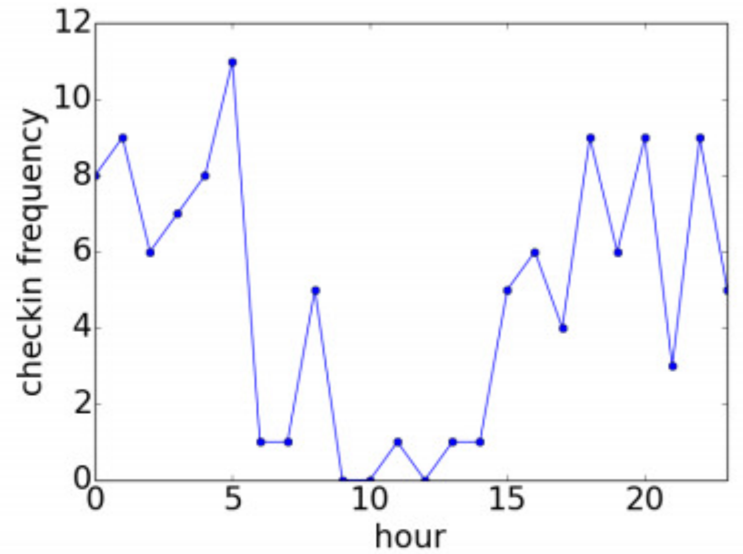
- GM: $O\big(n * m * T\big)$
- MGM: $O\big(n * m^2\big)$
- GMM: $O\big(n * m * k * T\big)$
- GA-GMM: $O\big(n * |P| * (mlogm + |C| * m * k) * T\big)$ [a]

---

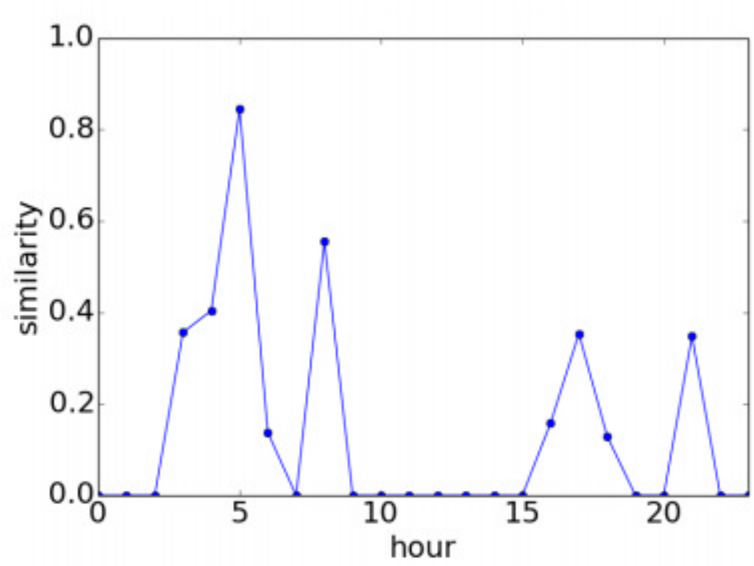
[a] n means the size of users; m means the average check-ins of one user; T means the iterations; k means the size of centers; $|P|$ is population size, $|C|$ is EM cycles.

ATTF

Data Analysis



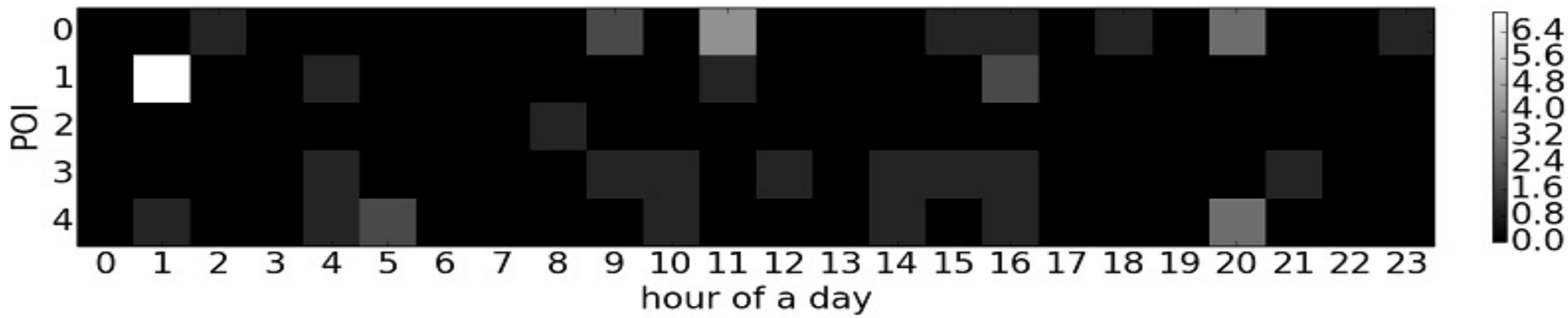
(a) A user's check-in pattern in a day



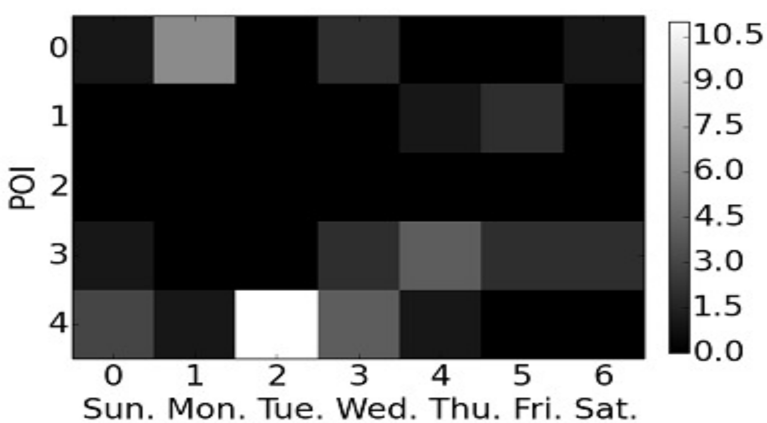
(b) Consecutive hour pair similarity

Figure 4.2: Sparsity demonstration

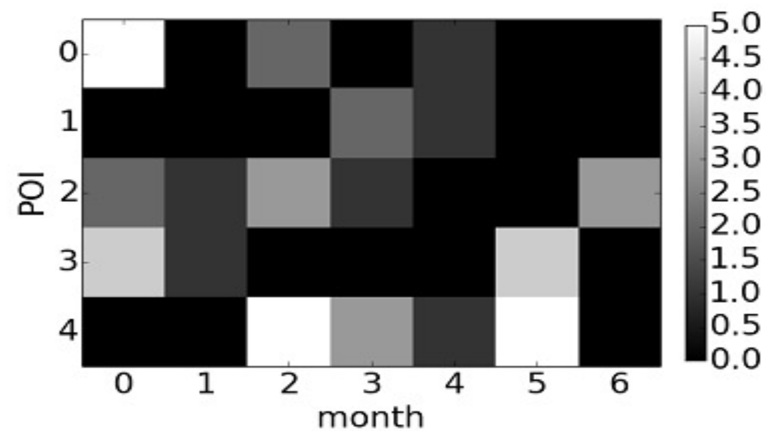
Data Analysis



(a) Non-uniformness in hour of one day



(b) Non-uniformness in day of week



(c) Non-uniformness in month

Figure 4.3: Demonstration of non-uniformness at different time scales

Model Interpretation

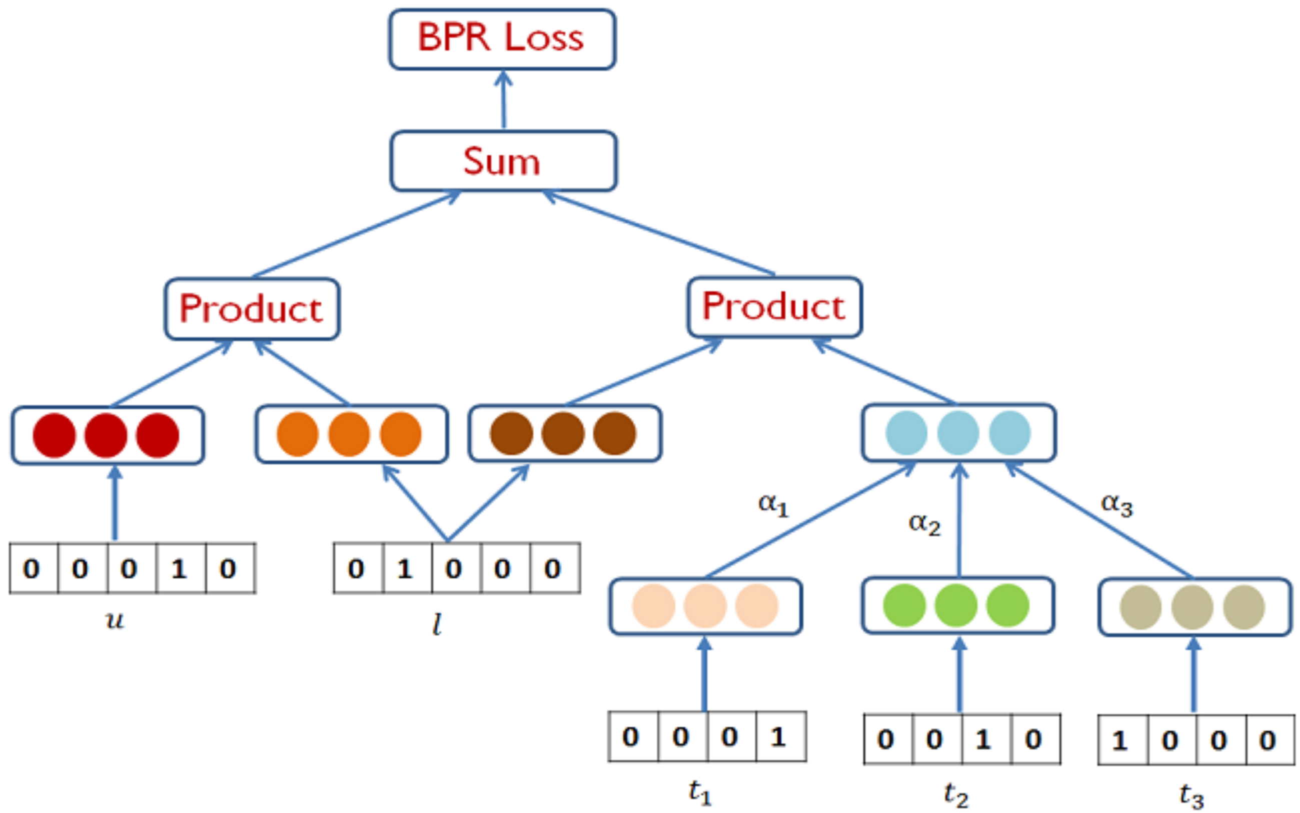


Figure 4.5: Embedding neural network for ATTF model

Summary

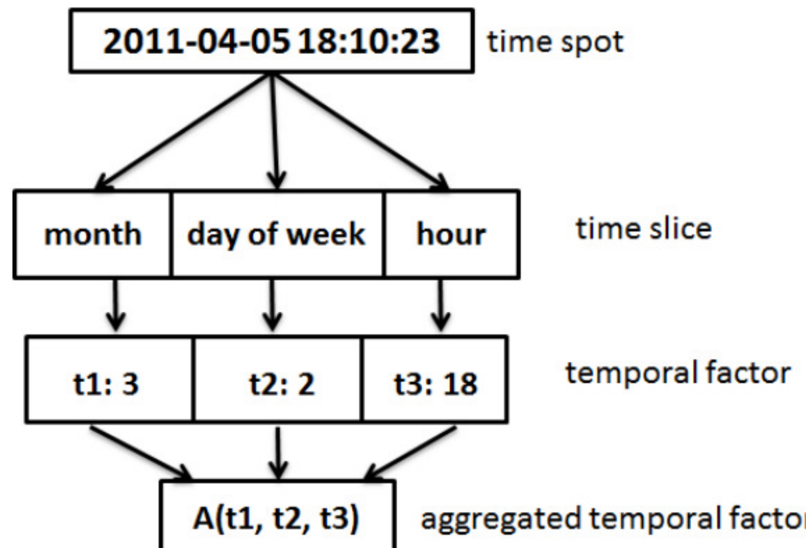
Aggregated Temporal Tensor Factorization

Time spot $\rightarrow$ Latent vector representation

Periodicity: same vector for the same time type

Non-uniformness: different vectors for different time types

Consecutiveness: learned from vectors via the check-ins



Preference score function

$$f(u, t, l) = \langle U_u^{(L)}, L_l^{(U)} \rangle + \langle A(T_{1,t_1}^{(L)}, T_{2,t_2}^{(L)}, T_{3,t_3}^{(L)}), L_l^{(T)} \rangle$$

$$A(\cdot) = \alpha_1 \cdot T_{1,t_1}^{(L)} + \alpha_2 \cdot T_{2,t_2}^{(L)} + \alpha_3 \cdot T_{3,t_3}^{(L)}$$

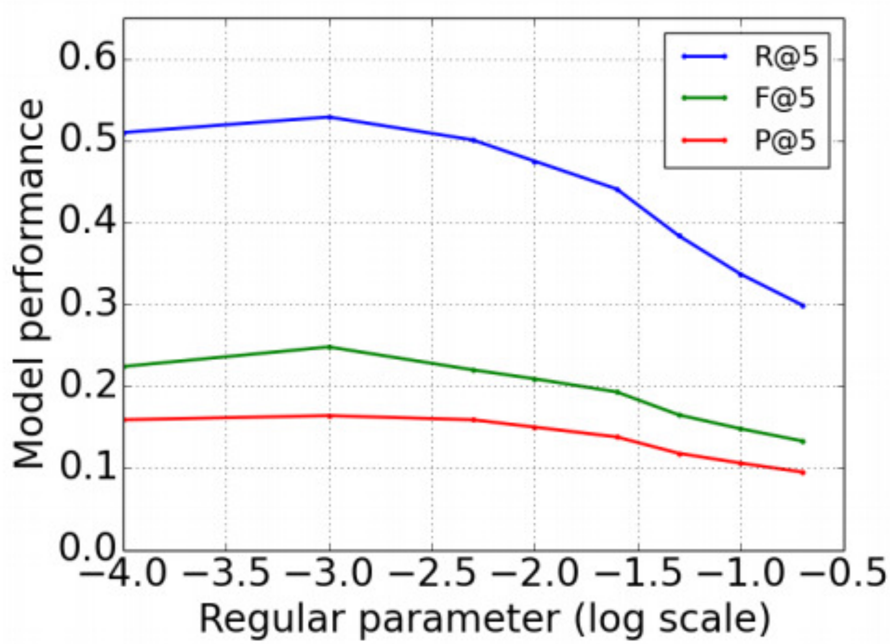
Learning ATTF

- ATTF model learning algorithm

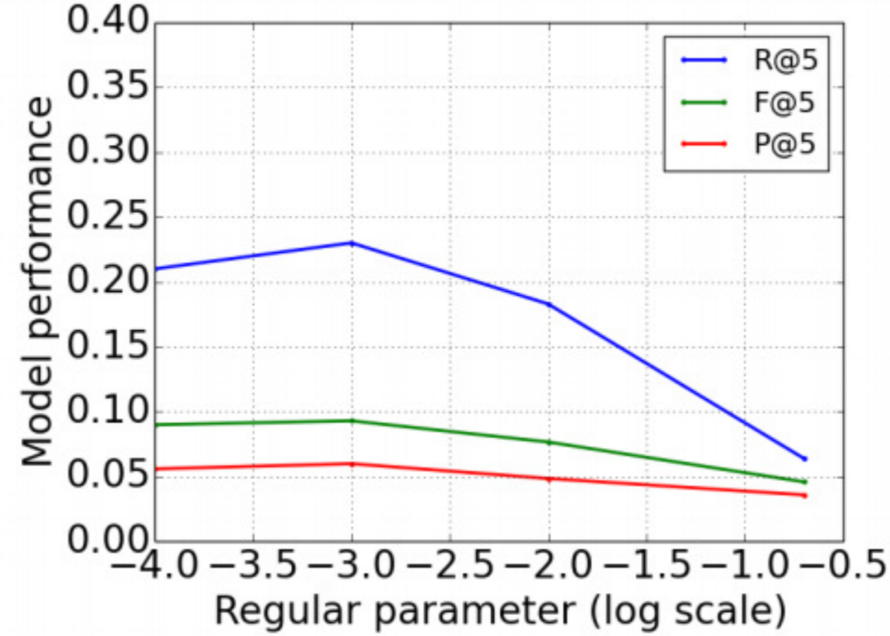
Input: Training tuples $\{(u_i, t_i, l_i)\}_{i=1,\ldots,N}$
Output: $U^{(L)}, T_1^{(L)}, T_2^{(L)}, T_3^{(L)}, L^{(U)}, L^{(T)}$

```
Initialize U^(L), T_1^(L), T_2^(L), T_3^(L), L^(U), L^(T)
repeat
    Draw (u, t, l_p) uniformly from training tuples
    For s = 1, 2, ..., k, where k is #sampled negative POIs
        Draw (u, t, l_p, l_n) uniformly
        y_{u,t,l_p,l_n} ← y_{u,t,l_p} − y_{u,t,l_n}
        δ ← 1 − σ(y_{u,t,l_p,l_n})
        Update parameters according to Eq. (4)
until convergence
return U^(L), T_1^(L), T_2^(L), T_3^(L), L^(U), L^(T)
```

Parameter Effect



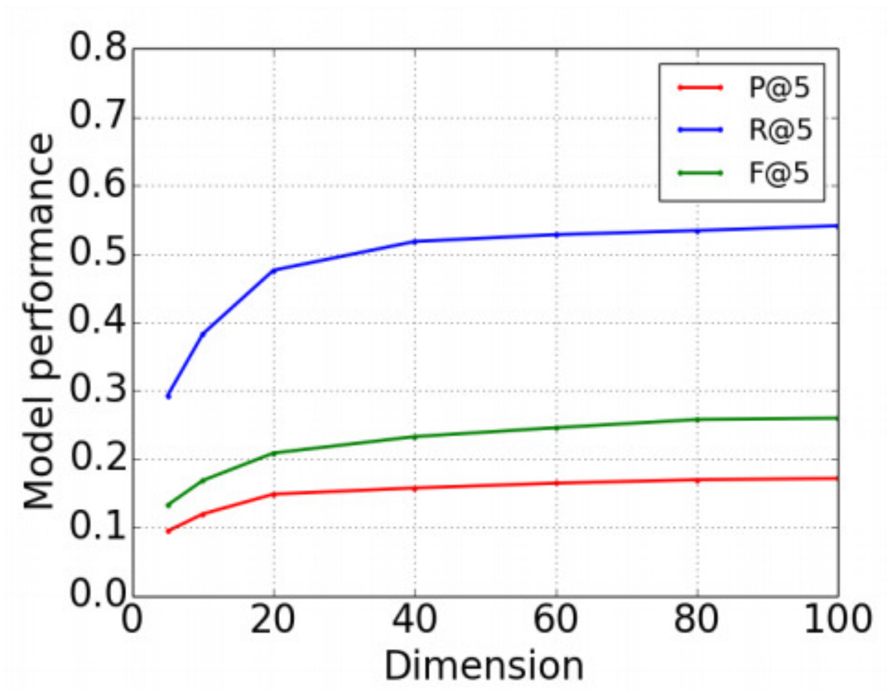
(a) Foursquare



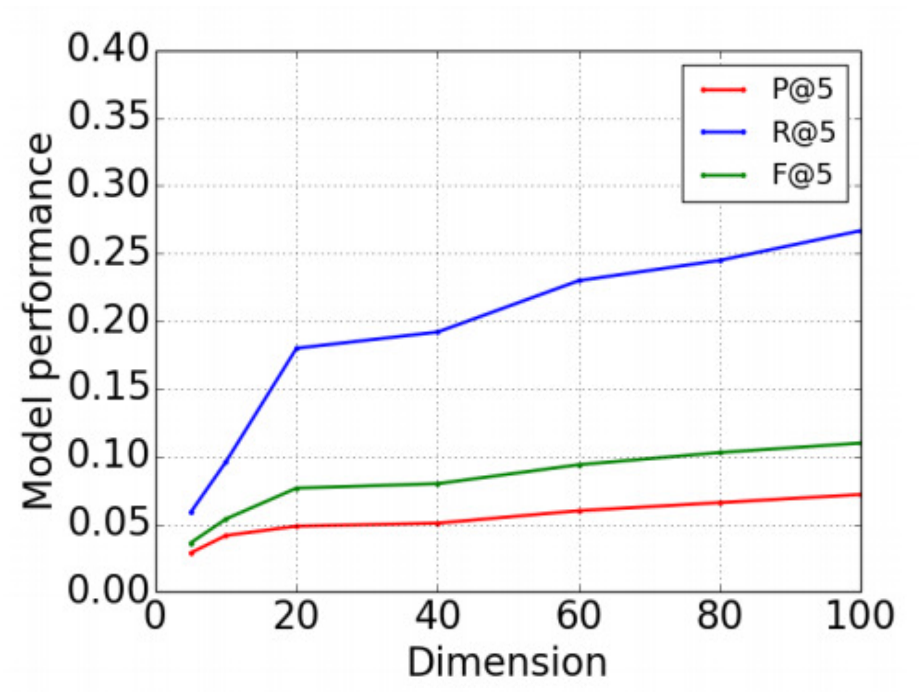
(b) Gowalla

The effect of regularization parameter λ

Parameter Effect



(a) Foursquare



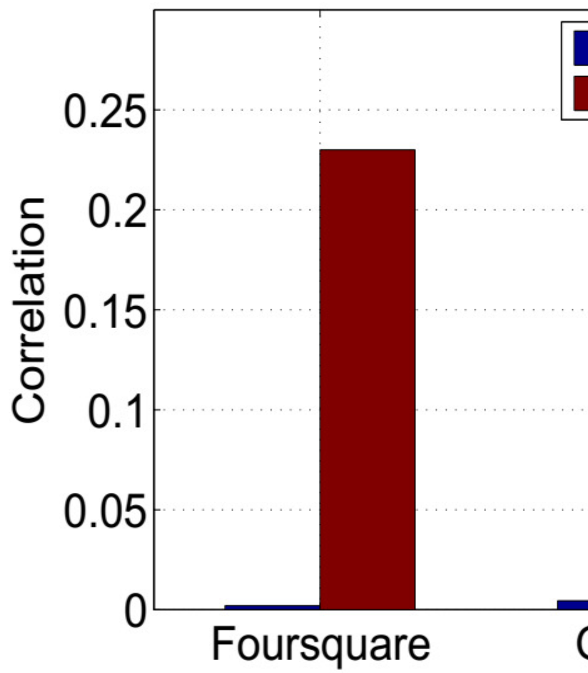
(b) Gowalla

The effect of latent factor dimension

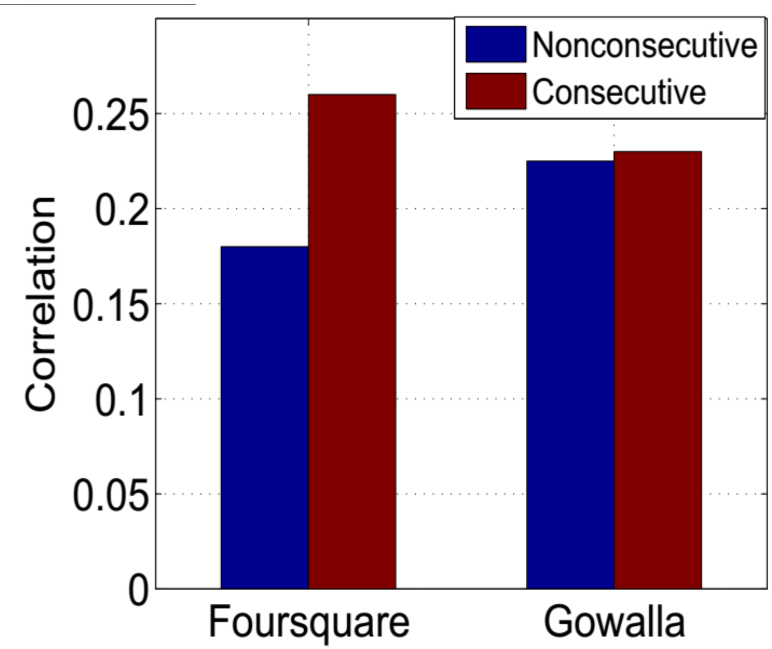
Geo-Teaser

Data Analysis

- How POIs in sequences correlate each other



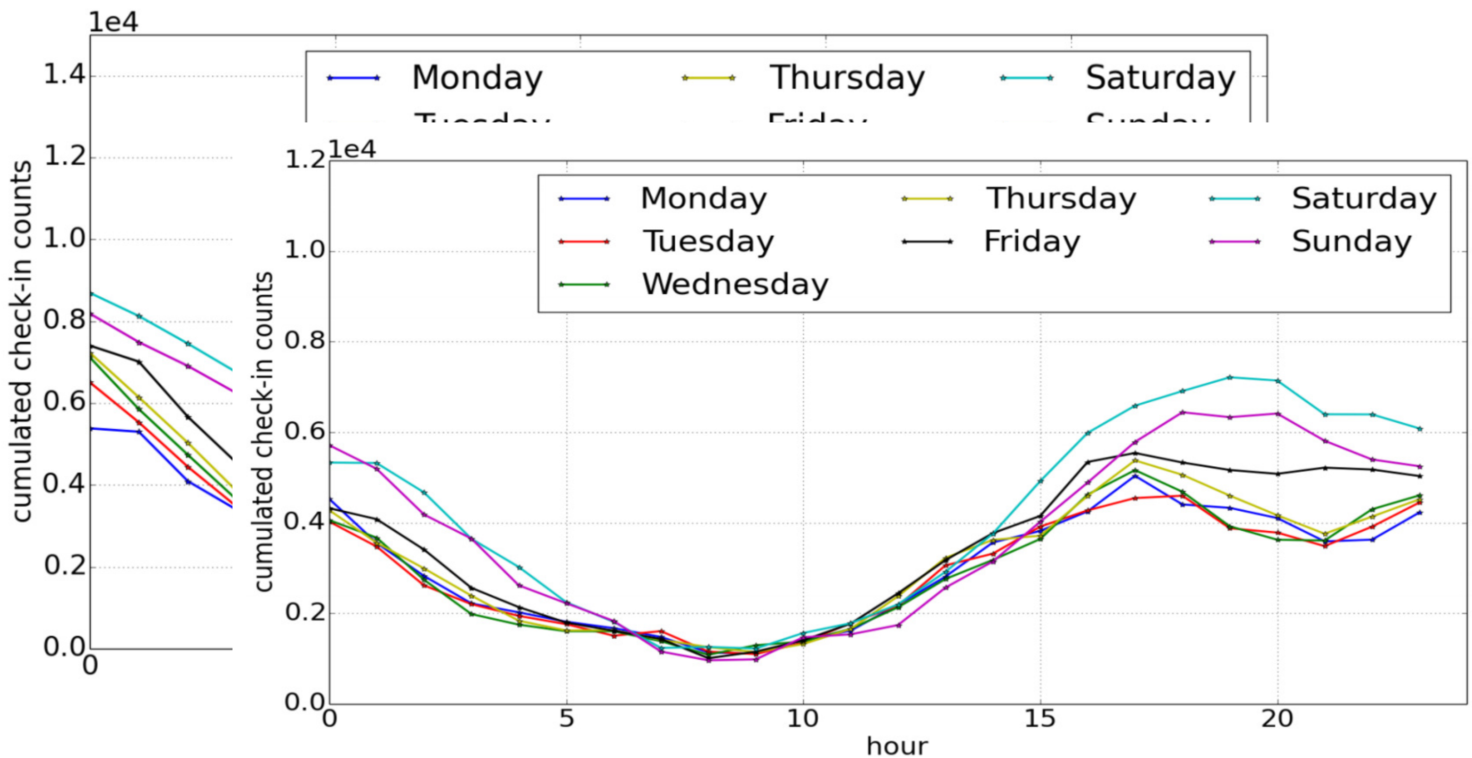
(a) Sequence vs. I



(b) Consecutive vs. Nonconsecutive

Data Analysis

- Temporal variance between weekday and weekend



(b) Gowalla

Sequence Pair Explanation

- A sequence: I_1, I_2, I_3
  - Consecutive pairs: (I_1, I_2), (I_2, I_3)
  - Non-consecutive pairs: (I_1, I_3)
- A sequence: I_1, I_2, I_3, I_4
  - Consecutive pairs: (I_1, I_2), (I_2, I_3), (I_3, I_4)
  - Non-consecutive pairs: (I_1, I_3), (I_1, I_4), (I_2, I_4)

Context Window Demonstration

		l_1	l_2	l_3	l_4	l_5	l_6		

- When k = 2,
  - For l_1, context POIs are l_2, l_3
  - For l_2, context POIs are l_1, l_3, l_4
  - For l_3, context POIs are l_1, l_2, l_4, l_5

Learning Geo-Teaser Algorithm

Learn the embedding model

Learn the geo pairwise ranking

Algorithm 1: Learning algorithm for the Geo-Teaser model

Input: S

Output: $\mathbf{U}, \mathbf{L}, \mathbf{T}$

Initialize $\mathbf{U}, \mathbf{L}, \mathbf{L}'$, and $\mathbf{T}$ (uniformly at random)

for *iterations* **do**

  for $S_u \in S$ **do**

    for $\langle u, l_i \rangle \in S_u$ **do**

      for *each context POI* l_c **do**

        $\mathbf{l}_i \leftarrow \mathbf{l}_i + \alpha\eta(1 - \sigma(\hat{\mathbf{l}}_c' \cdot \mathbf{l}_i^t))\mathbf{l}_c'$

        $\mathbf{t}_i \leftarrow \mathbf{t}_i + \alpha\eta(1 - \sigma(\hat{\mathbf{l}}_c' \cdot \mathbf{l}_i^t))\mathbf{l}_c'$

        $\mathbf{l}_c' \leftarrow \mathbf{l}_c' + \alpha\eta(1 - \sigma(\hat{\mathbf{l}}_c' \cdot \mathbf{l}_i^t))(\mathbf{l}_i + \mathbf{t}_i)$

        for $\{k' \sim P_{nc_c}\}$ **do**

          $\mathbf{l}_i \leftarrow \mathbf{l}_i - \alpha\eta\sigma(\hat{\mathbf{l}}_{k'}' \cdot \mathbf{l}_i^t)\mathbf{l}_{k'}'$

          $\mathbf{t}_i \leftarrow \mathbf{t}_i - \alpha\eta\sigma(\hat{\mathbf{l}}_{k'}' \cdot \mathbf{l}_i^t)\mathbf{l}_{k'}'$

          $\mathbf{l}_{k'}' \leftarrow \mathbf{l}_{k'}' - \alpha\eta\sigma(\hat{\mathbf{l}}_{k'}' \cdot \mathbf{l}_i^t)(\mathbf{l}_i + \mathbf{t}_i)$

        end

      end

      Uniformly sample m unvisited POIs

      for $(u, l_i, l_{ne}) \in D_m$ **do**

        $\delta = 1 - \sigma(\mathbf{u} \cdot \mathbf{l}_i - \mathbf{u} \cdot \mathbf{l}_{ne})$

        $\mathbf{u} \leftarrow \mathbf{u} + \beta\eta\delta(\mathbf{l}_i - \mathbf{l}_{ne})$

        $\mathbf{l}_i \leftarrow \mathbf{l}_i + \beta\eta\delta\mathbf{u}$; $\mathbf{l}_{ne} \leftarrow \mathbf{l}_{ne} - \beta\eta\delta\mathbf{u}$

      end

      for $(u, l_{ne}, l_{nn}) \in D_m$ **do**

        $\delta = (1 - \sigma(\mathbf{u} \cdot \mathbf{l}_{ne} - \mathbf{u} \cdot \mathbf{l}_{nn}))$

        $\mathbf{u} \leftarrow \mathbf{u} + \beta\eta\delta(\mathbf{l}_{ne} - \mathbf{l}_{nn})$

        $\mathbf{l}_{ne} \leftarrow \mathbf{l}_{ne} + \beta\eta\delta\mathbf{u}$; $\mathbf{l}_{nn} \leftarrow \mathbf{l}_{nn} - \beta\eta\delta\mathbf{u}$

      end

    end

  end

end

Complexity

Window size

Vector dimension

All check-ins

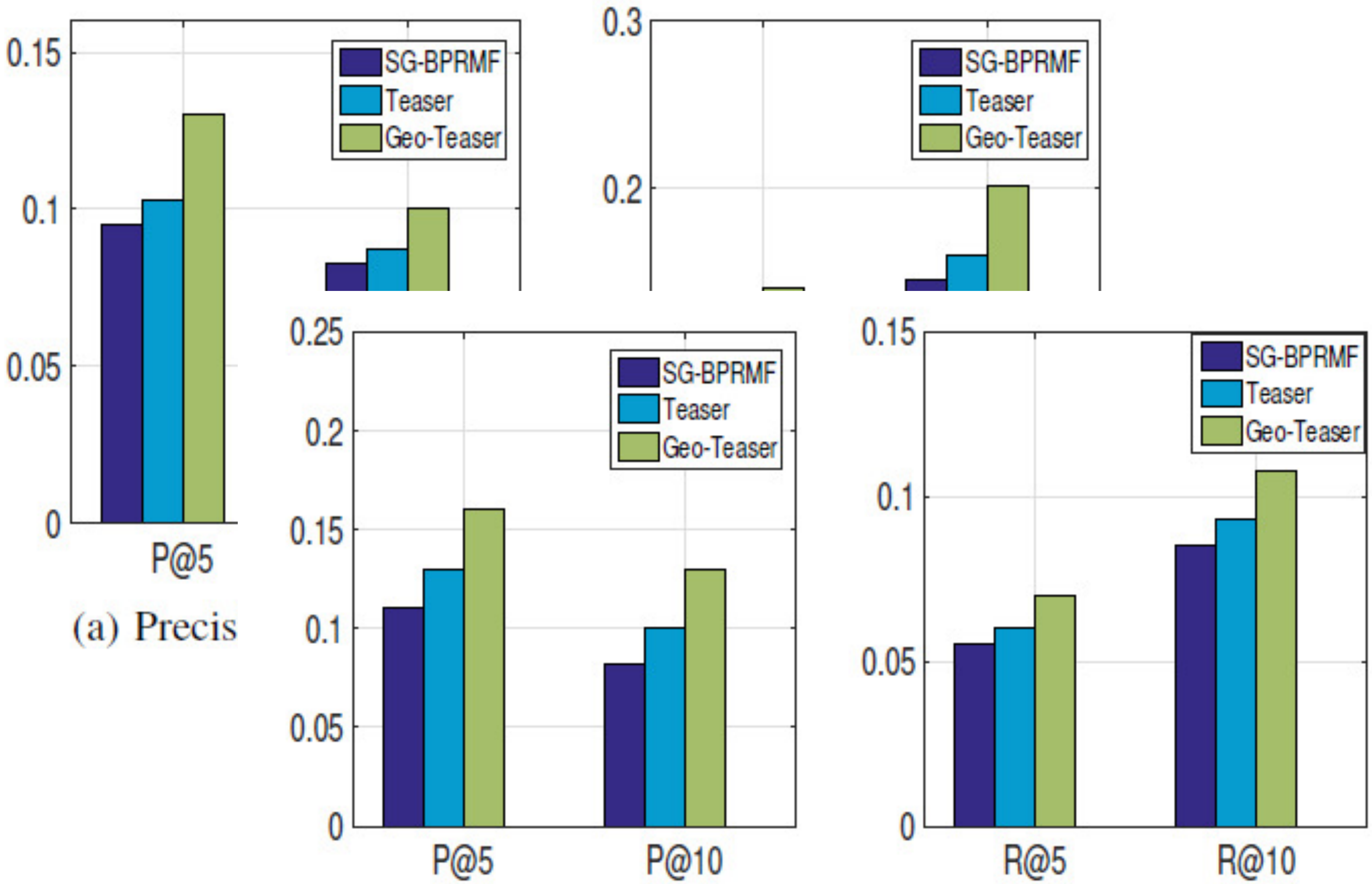
$$O\left(\left(k \cdot h + m^2\right) \cdot d \cdot |C|\right)$$

Number of negative context POIs

Number of sampled unchecked POI

- Linear in $O(|C|)$
  - k, h, m, d are fixed hyper-parameters

Experiment Discussion

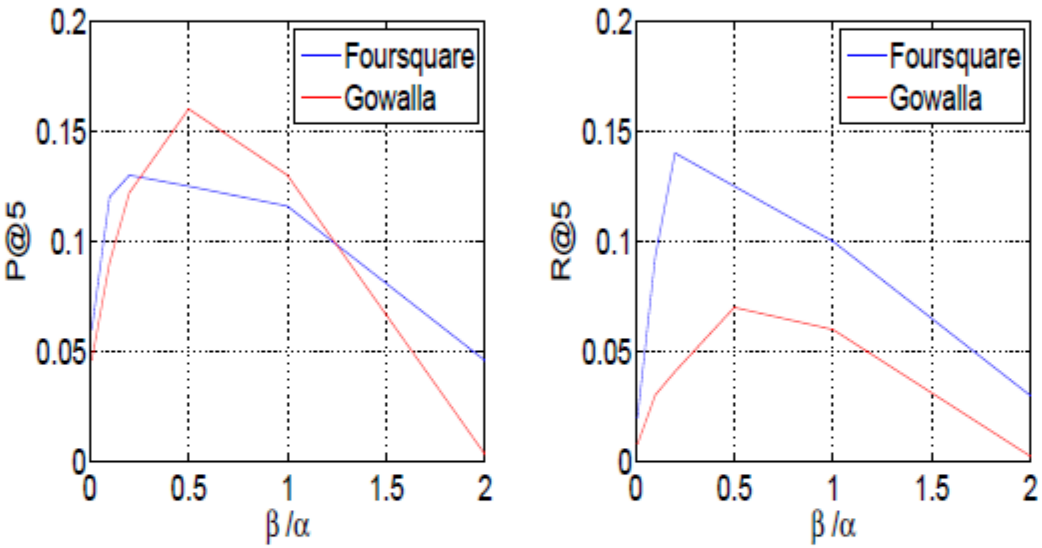


(a) Precis

(c) Precision-Gowalla

(d) Recall-Gowalla

Parameter Effect



(a) P@5 on Geo-Teaser (b) R@5 on Geo-Teaser

Figure 7: Parameter effect on α and β

Parameter Effect

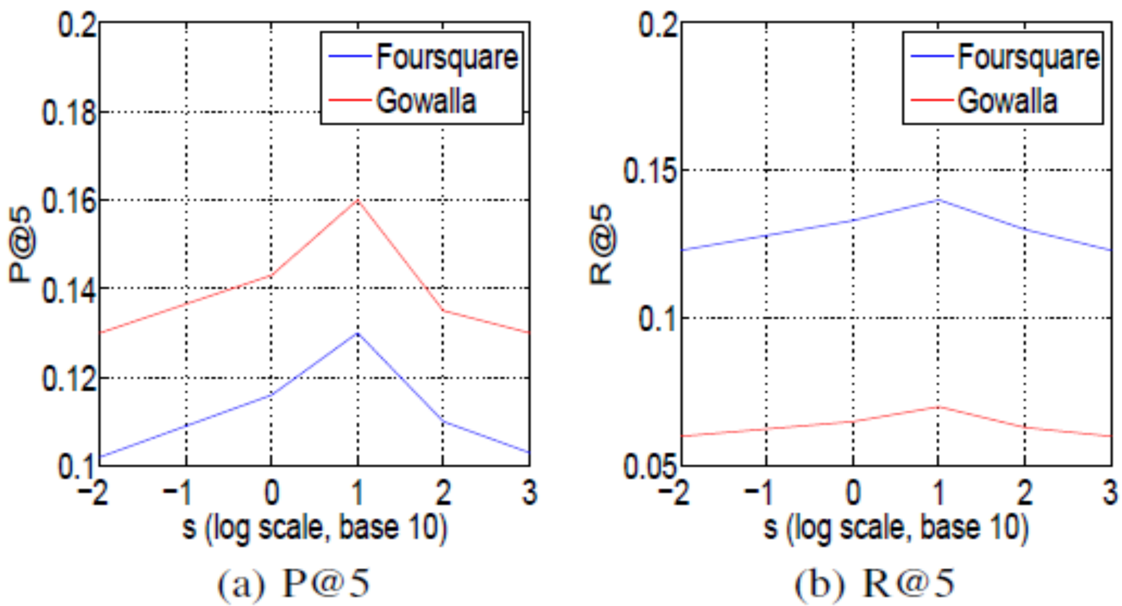
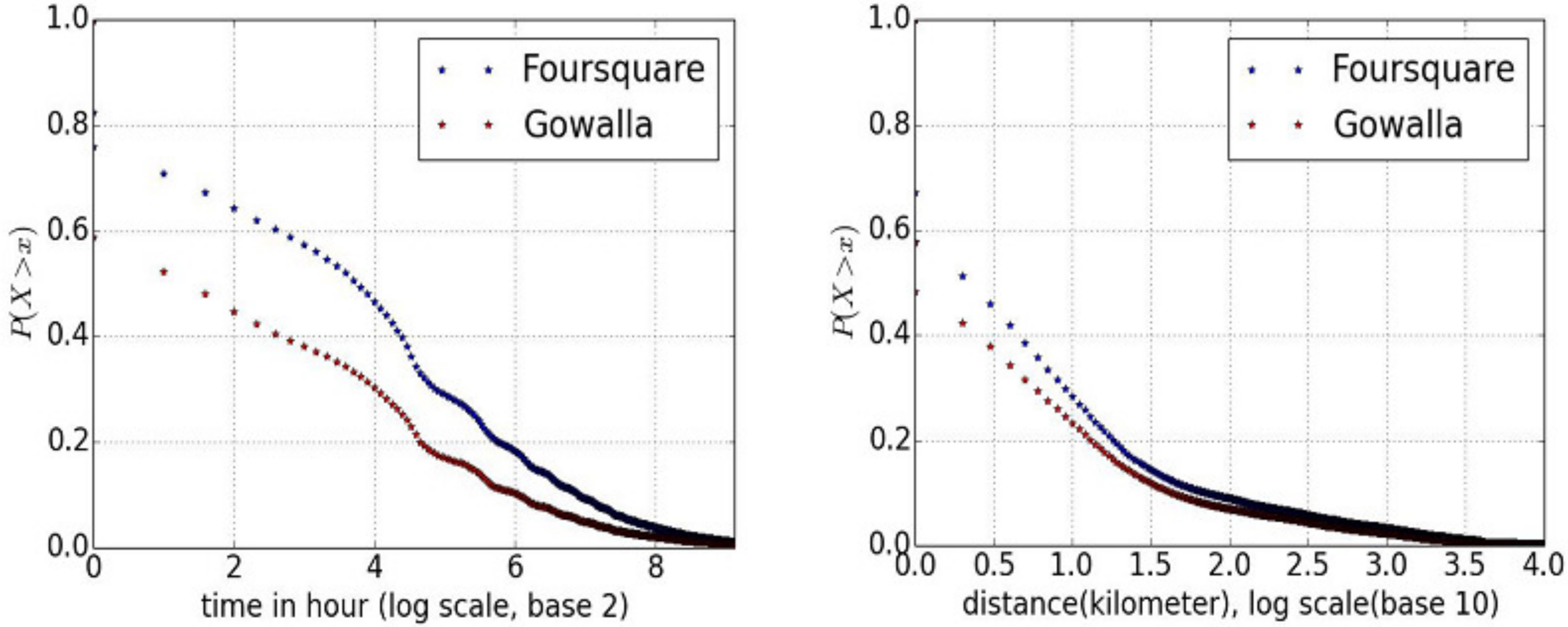


Figure 8: Parameter effect on distance threshold s

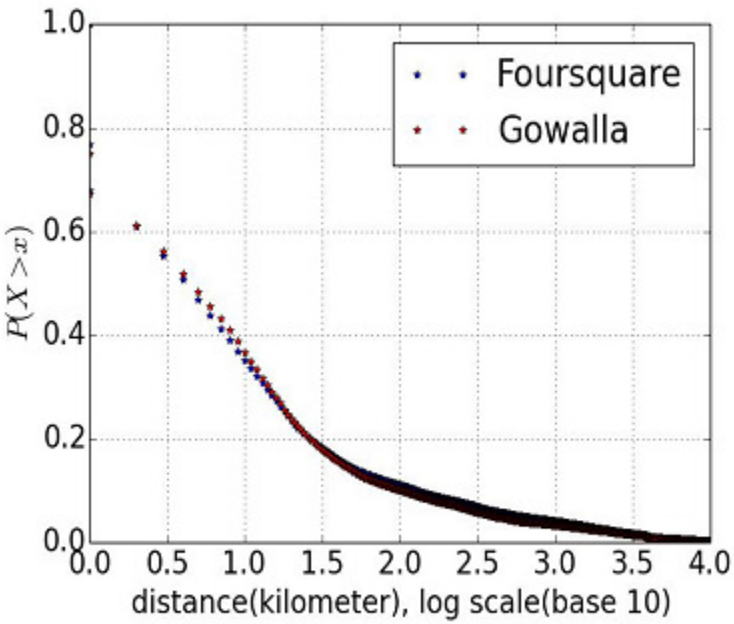
STELLAR

Data Analysis

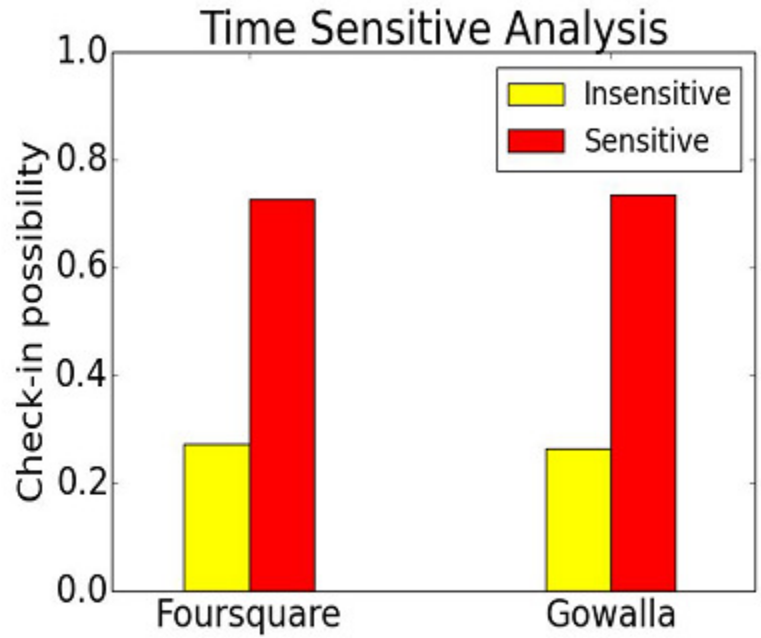


(a) CCDF of intervals in successive check-ins (b) CCDF of distances in successive check-ins

Data Analysis

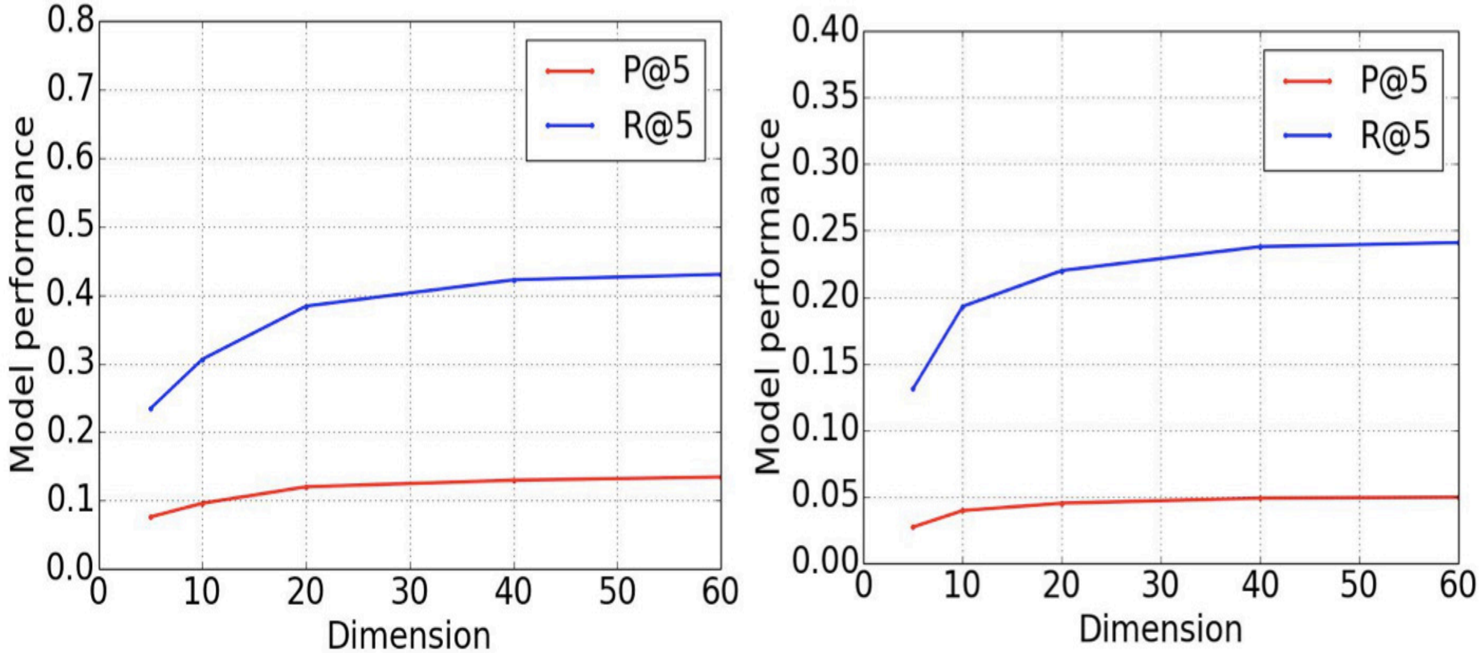


(c) CCDF of distances in successive check-ins beyond 4 hours



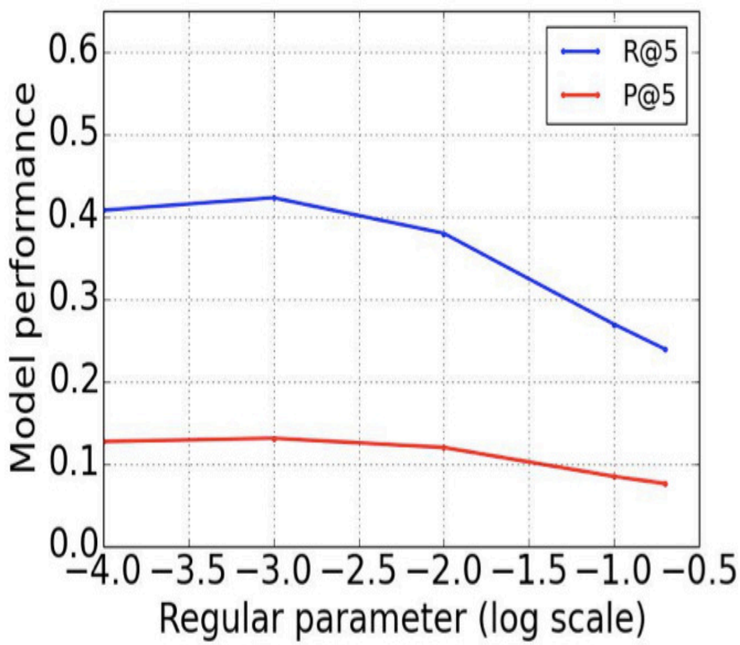
(d) Time sensitive analysis of successive POI check-ins

Parameter Effect

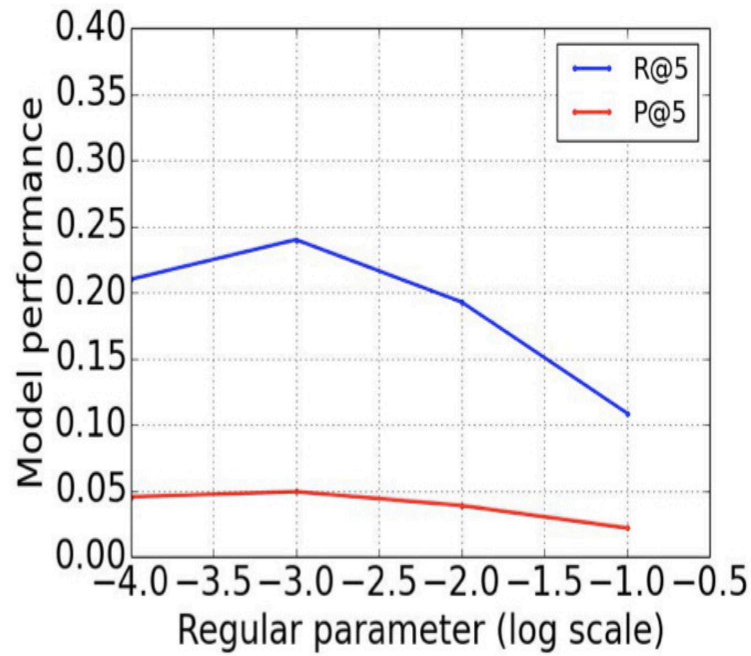


(a) Foursquare (b) Gowalla

Parameter Effect



(a) Fousquare



(b) Gowalla