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

Irwin King with Chen Cheng

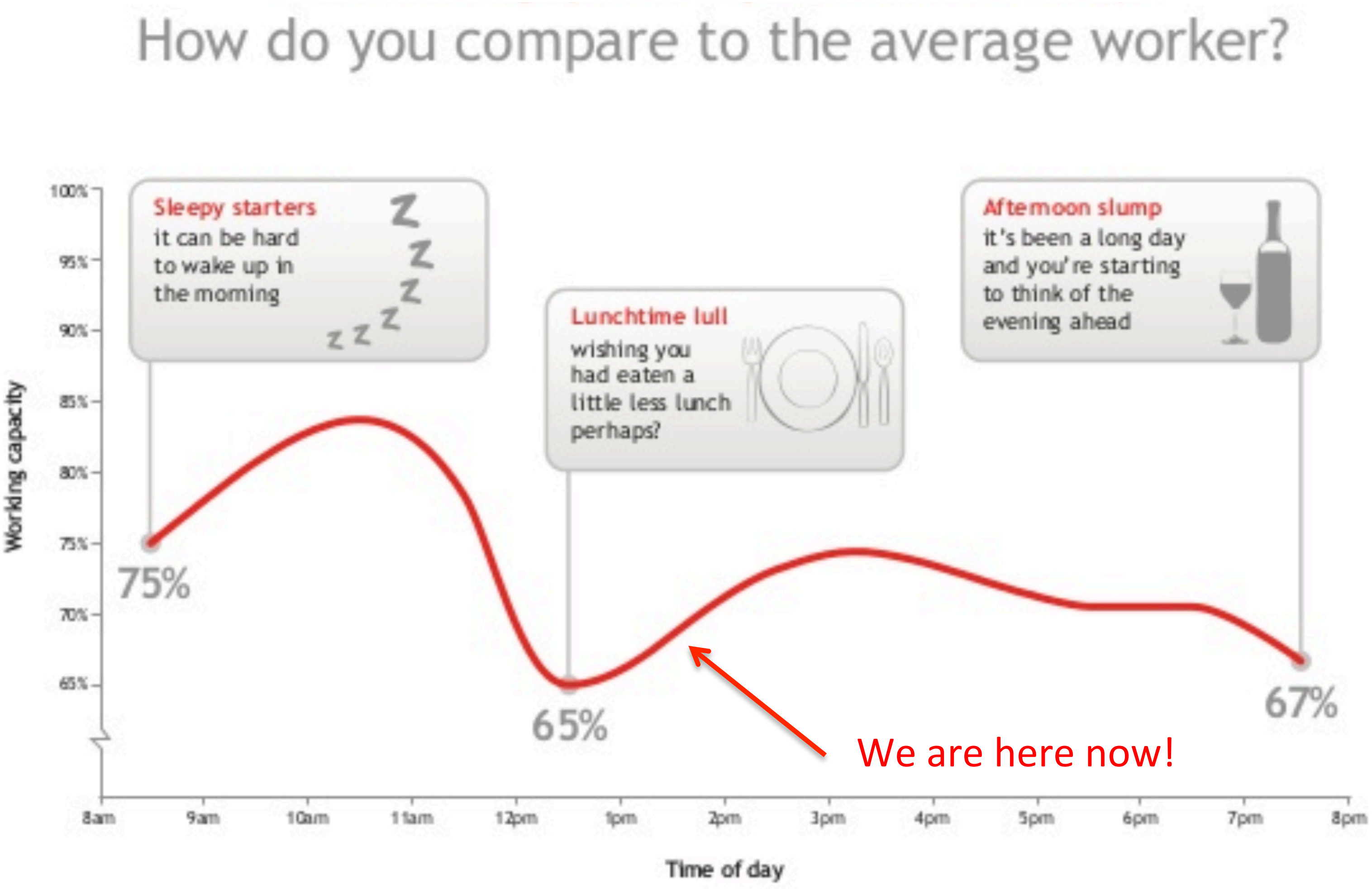
Department of Computer Science and Engineering
The Chinese University of Hong Kong

king@cse.cuhk.edu.hk
http://www.cse.cuhk.edu.hk/~king

LARC

to conduct research on behavioural and social network analytics and behavioural experiments so as to discover and harness the laws of information network evolution for networks of people, organisations and businesses

Our Productivity Plot



We are here now!

http://www.slideshare.net/RobCubbon/24343104-productivitychart

Outline

- Introduction & motivations
- POI recommendation in LBSNs
- Successive POI recommendation
- Conclusion

http://scobleizer.com/2010/01/29/the-foursquare-squeeze-will-it-survive-to-check-in-on-2011/

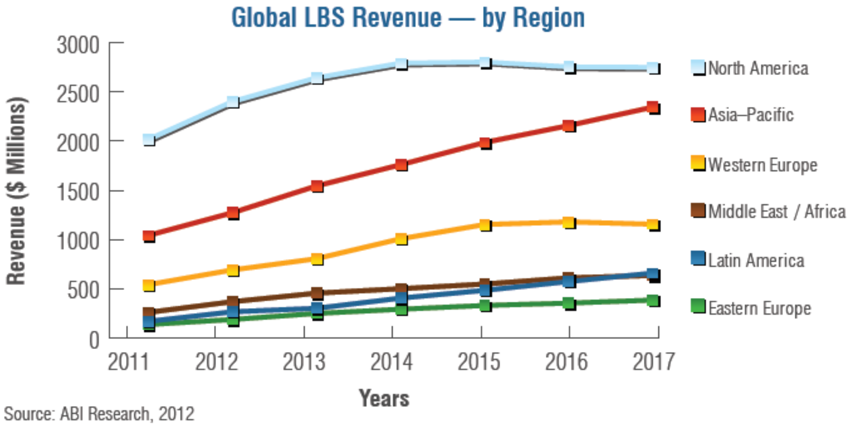
Location is a $17B Industry

		Total	
		Revenue ($B)	Jobs (K)
Geo-applications & devices	• Develops and manufactures devices and software for creating, visualizing, sharing, and analyzing geographic information	54	175
Location-based geo-data	• Collects, manages, and distributes spatial information and imagery • **Provides navigational aides and other location- finding services**	**17**	200
		$70.2 B	375K

http://www.slideshare.net/Locaid/locaid-location-based-services-industry-stats-nov2013pdf

Growth of Location-based Services

- Almost one fifth **(19%) of the world's six billion mobile users are already using LBS**
  - Navigation via maps and GPS is currently the most popular application, used by 46%
- One in five (**22%**) of LBS users are using applications designed to **help them find their friends nearby**
- **26% use** the technology **to find restaurants** and entertainment venues
- **74%** of smartphone owners use location-based services.

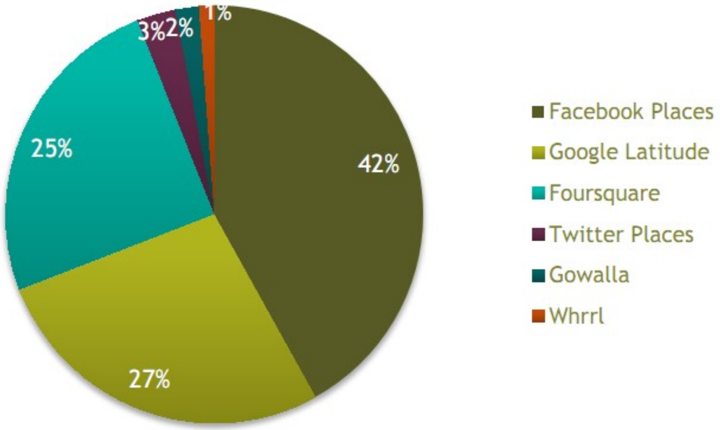


Source: ABI Research, 2012

Figure 2. Projected LBS services revenue by region (2011-2017)[6]

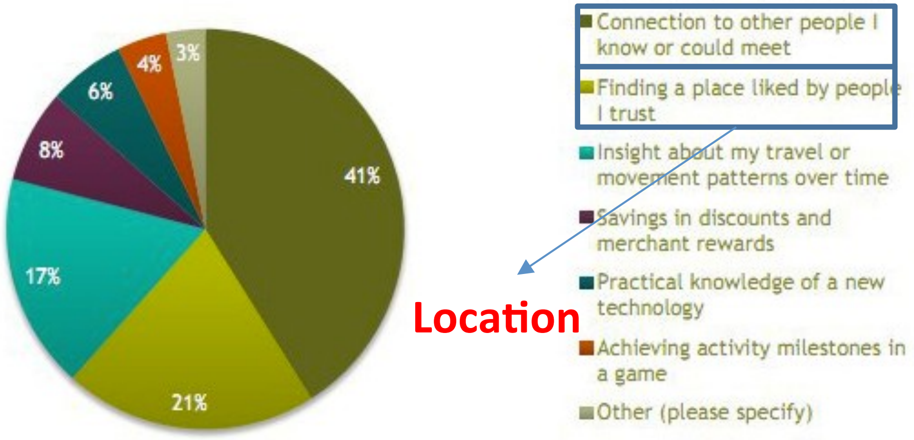
Check-in Becomes a Life Style…

"Which of these apps do you use most frequently?" (n=169)



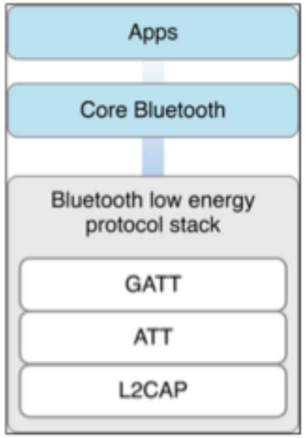
"What is the most important benefit of these apps to you, personally?" (n=253)

Social Networks



Location

iBeacon
Indoor and Micro-location Positioning


Apps

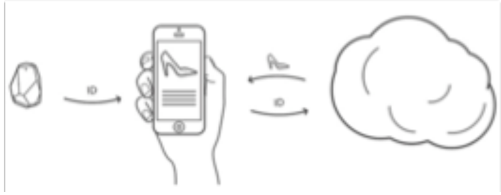
Core Bluetooth

Bluetooth low energy protocol stack

GATT

ATT

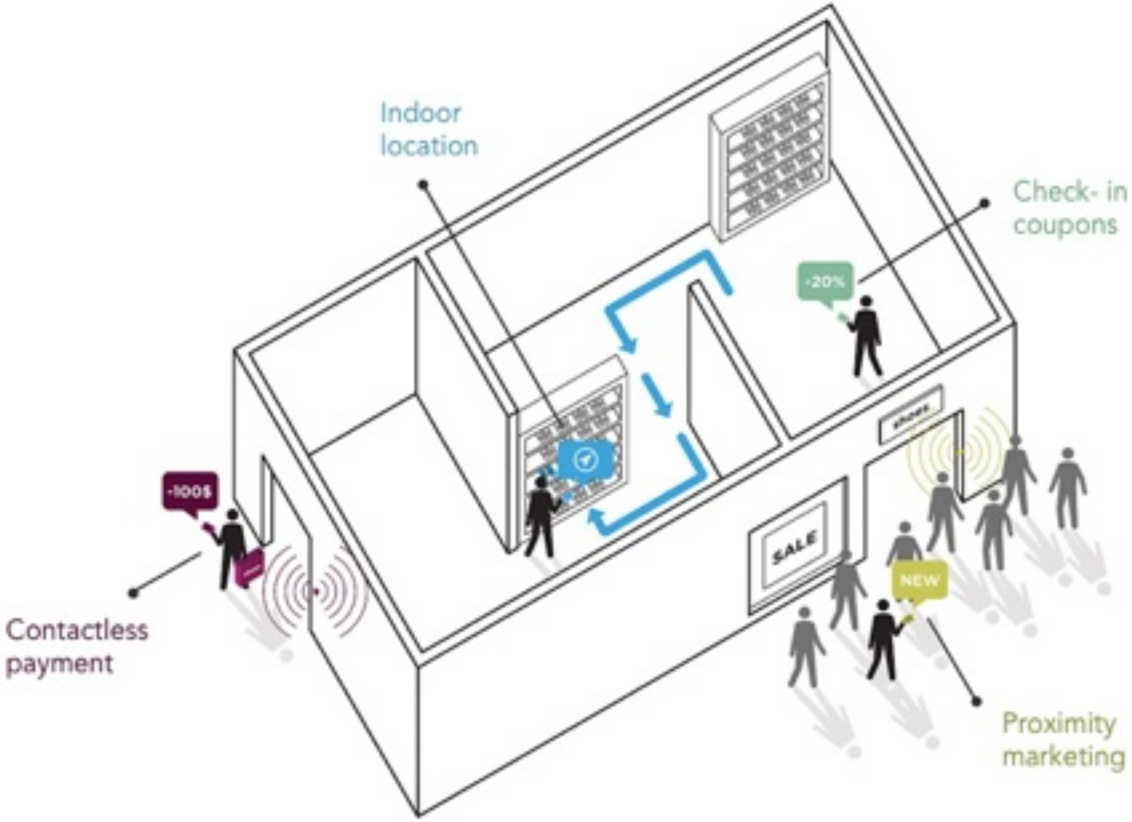
L2CAP



Apps for iBeacon

iBeacon

Get In-Store Notifications



http://www.ubergizmo.com/2014/02/mlb-completes-rollout-of-ibeacon-to-two-stadiums/

http://www.fanengagement.nl/news/social-media/apple-ruling-location-awareness-with-new-ibeacon/

http://www.tuaw.com/2013/12/06/apple-now-using-ibeacon-technology-in-its-us-retail-stores/

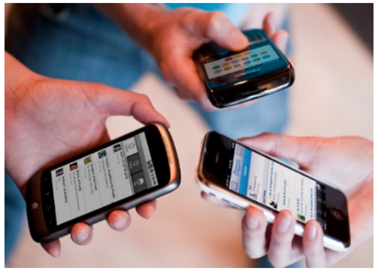
Categories of LBSN Services

- Geo-tagged-media-based
- Point-of-interest driven
- Trajectory-centric

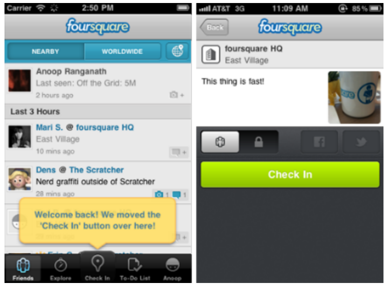
Chapter 8 and 9 of the book
Computing with Spatial Trajectories by Yu Zheng and Xing Xie

Location + Social Networks

- Add a new dimension to social networks
  - Geo-tagged user-generated media: texts, photos, and videos, etc.
  - Location history of users recorded
- Location is a new object in the network
- Bridging the gap between the virtual and physical worlds
  - Sharing real-world experiences online
  - Consume online information in the physical world



Physical world



Anoop Ranganath
Last seen: Off the Grid: 5M
2 hours ago

Last 3 Hours

Mari S. @ foursquare HQ
East Village
10 mins ago

Dens @ The Scratcher
Nerd graffiti outside of Scratcher
28 mins ago

Welcome back! We moved the 'Check In' button over here!

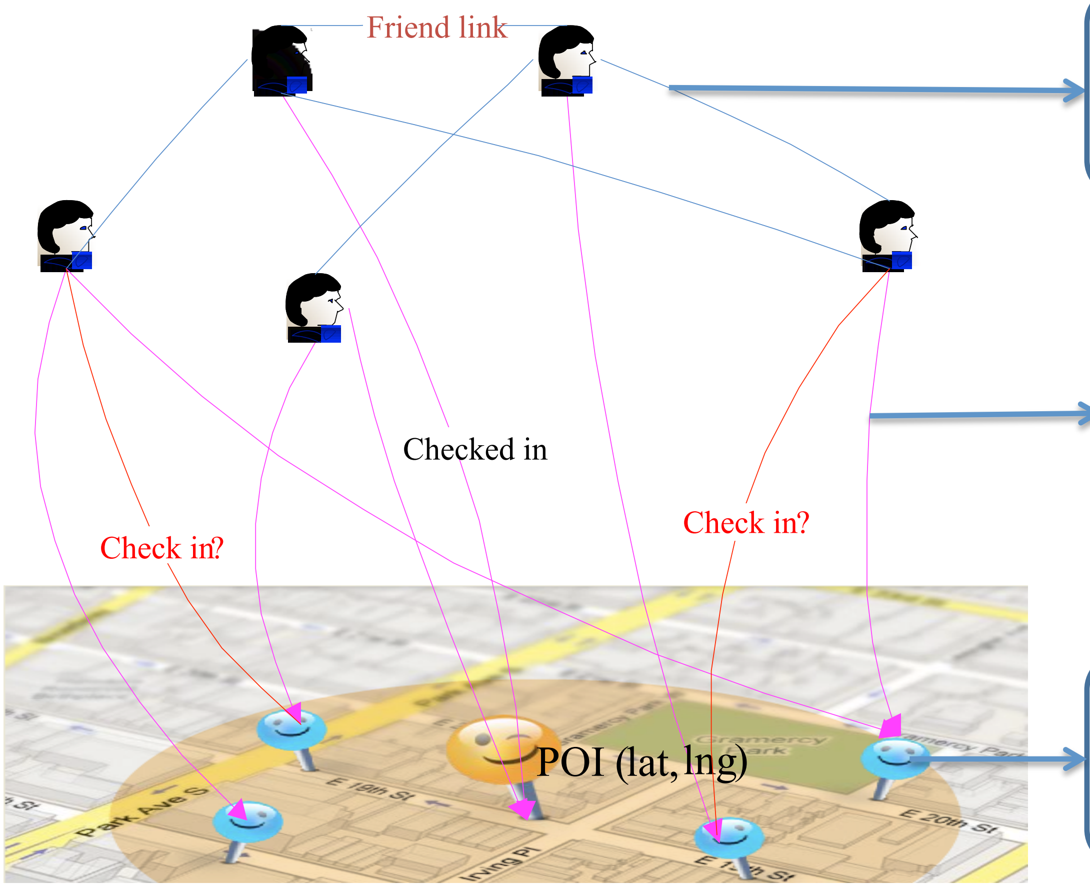
foursquare HQ
East Village

This thing is fast!

Virtual world

Social sharing

Graph Illustration of Location-based Social Networks (LBSNs)



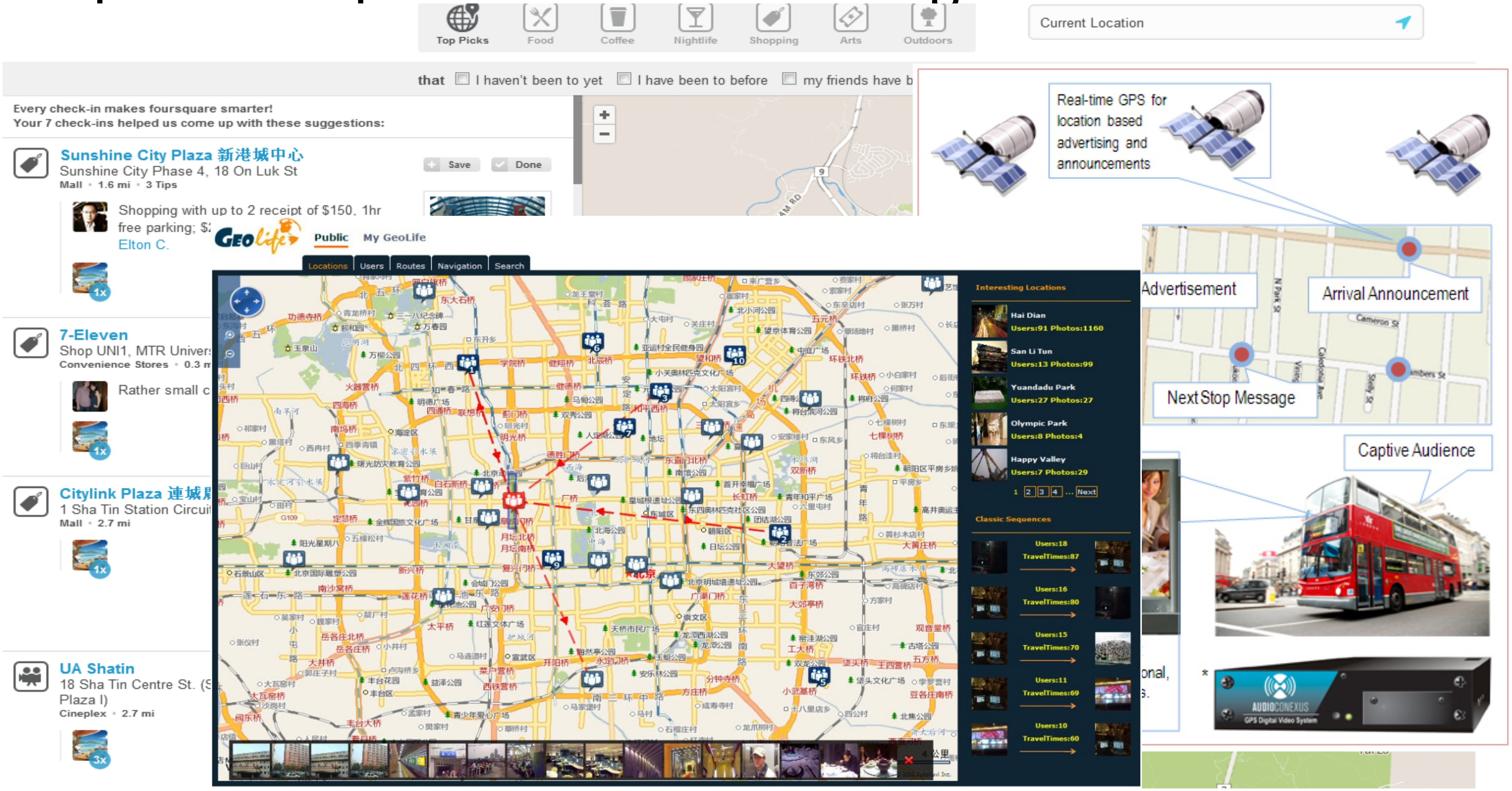
- Community detection
- User behavior learning

- POI recommendation
- Next place prediction

- Travel sequence detection
- Trip recommendation
- Location semantic meaning

Our Focus: POI Recommendation

- Help users explore their surroundings
-
-

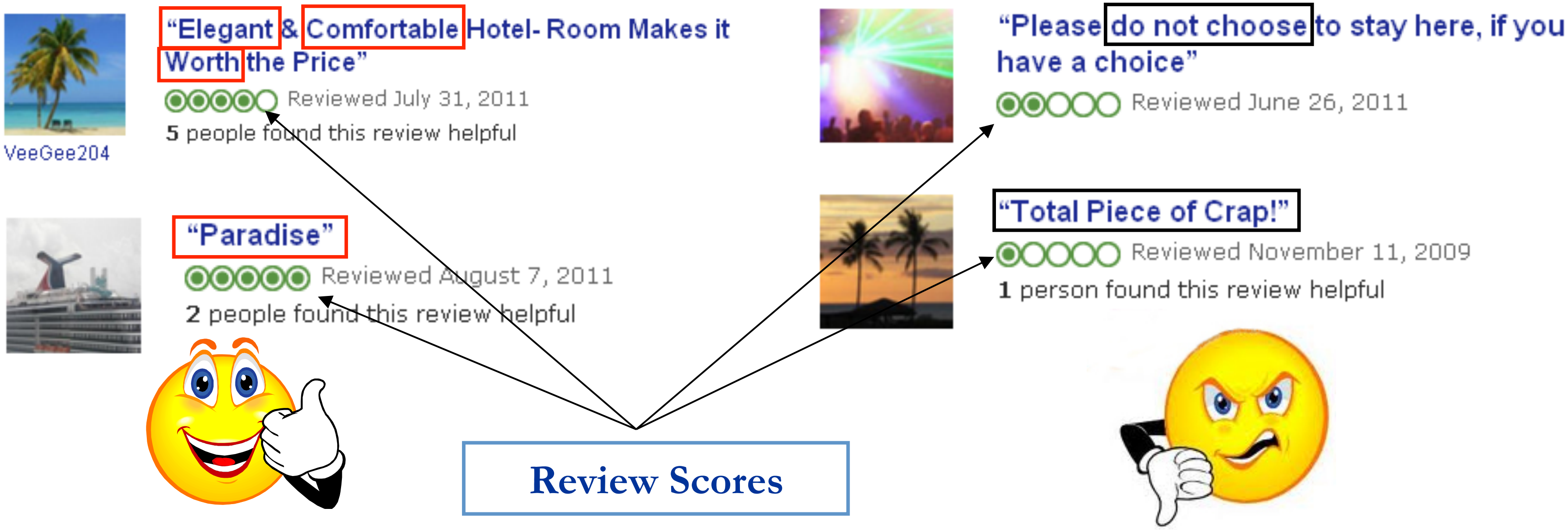


POI Recommendation

- Non-personalized recommendation
  - Tree-based Hierarchical Graph + HITS [Zheng et al., WWW’09]
  - Location-feature-activity factorization [Zheng et al., WWW’10]
- Personalized recommendation
  - Model-based method: UCLAF [Zheng et al., AAAI’10]
  - Item-based method: Community Location Model (CLM) [Leung et al., SIGIR’11], User+Loation+Social fused model [Ye et al., SIGIR’11]

Recommendation

- From contents



"Elegant & Comfortable Hotel- Room Makes it Worth the Price"
●●●●○ Reviewed July 31, 2011
5 people found this review helpful
VeeGee204

"Paradise"
●●●●● Reviewed August 7, 2011
2 people found this review helpful

"Please do not choose to stay here, if you have a choice"
●●○○○ Reviewed June 26, 2011

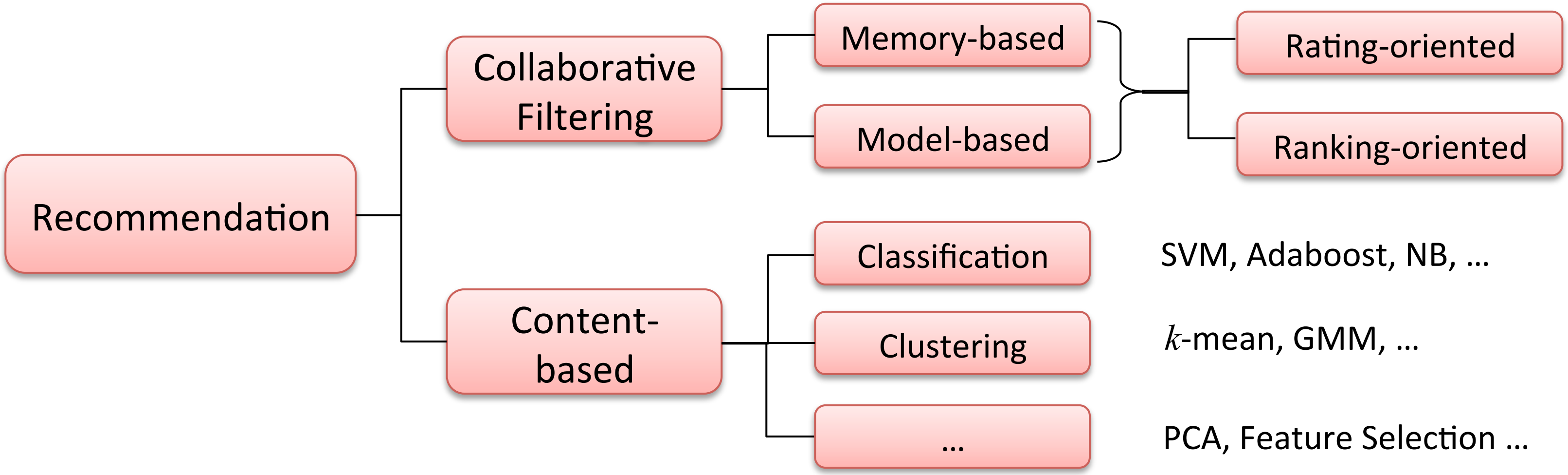
"Total Piece of Crap!"
●○○○○ Reviewed November 11, 2009
1 person found this review helpful

Review Scores

- From collaborative filtering
  - Form user-item matrix

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3

Learning Techniques in Recommendation

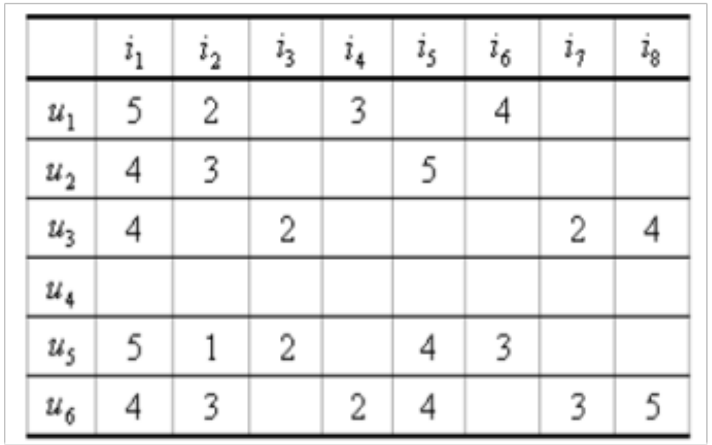


- Collaborative filtering
  - Use user-item matrix to predict rating/ranking
  - Simple in data collection
- Content-based learning
  - Users' preference expressed in intrinsic features
  - Difficult in feature representation

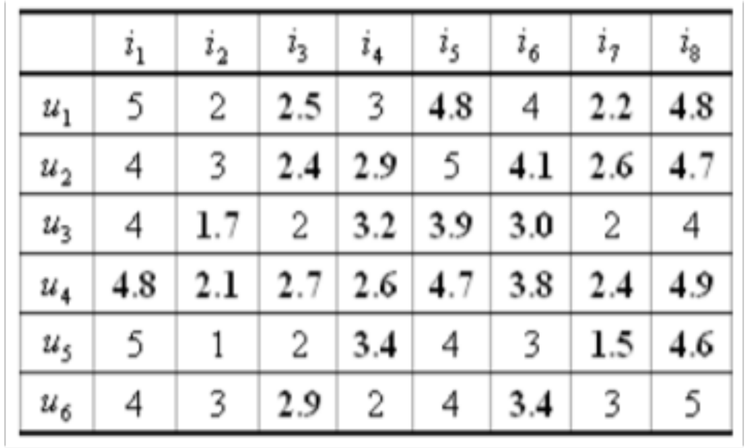
Social Recommendations with Matrix Factorization

- Model-based Collaborative Filtering
  - Clustering Methods [Hkors et al, CIMCA '99]
  - Bayesian Methods [Chien et al., IWAIS '99]
  - Aspect Method [Hofmann, SIGIR '03]
  - **Matrix Factorization** [Sarwar et al., WWW '01]
- Social Recommendations
  - Social recommendation using probabilistic matrix factorization [CIKM'08]
  - Learning to recommend with social trust ensemble [SIGIR'09]
  - Recommend with social distrust [RecSys'09]
  - Website recommendation [SIGIR'11]

Matrix Factorization



	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8
u_1	5	2		3		4		
u_2	4	3			5			
u_3	4		2				2	4
u_4								
u_5	5	1	2		4	3		
u_6	4	3		2	4		3	5



	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8
u_1	5	2	2.5	3	4.8	4	2.2	4.8
u_2	4	3	2.4	2.9	5	4.1	2.6	4.7
u_3	4	1.7	2	3.2	3.9	3.0	2	4
u_4	4.8	2.1	2.7	2.6	4.7	3.8	2.4	4.9
u_5	5	1	2	3.4	4	3	1.5	4.6
u_6	4	3	2.9	2	4	3.4	3	5

$$U = \begin{bmatrix} 1.55 & 1.22 & 0.37 & 0.81 & 0.62 & -0.01 \\ 0.36 & 0.91 & 1.21 & 0.39 & 1.10 & 0.25 \\ 0.59 & 0.20 & 0.14 & 0.83 & 0.27 & 1.51 \\ 0.39 & 1.33 & -0.43 & 0.70 & -0.90 & 0.68 \\ 1.05 & 0.11 & 0.17 & 1.18 & 1.81 & 0.40 \end{bmatrix}$$

$$V = \begin{bmatrix} 1.00 & -0.05 & -0.24 & 0.26 & 1.28 & 0.54 & -0.31 & 0.52 \\ 0.19 & -0.86 & -0.72 & 0.05 & 0.68 & 0.02 & -0.61 & 0.70 \\ 0.49 & 0.09 & -0.05 & -0.62 & 0.12 & 0.08 & 0.02 & 1.60 \\ -0.40 & 0.70 & 0.27 & -0.27 & 0.99 & 0.44 & 0.39 & 0.74 \\ 1.49 & -1.00 & 0.06 & 0.05 & 0.23 & 0.01 & -0.36 & 0.80 \end{bmatrix}$$

Matrix Factorization

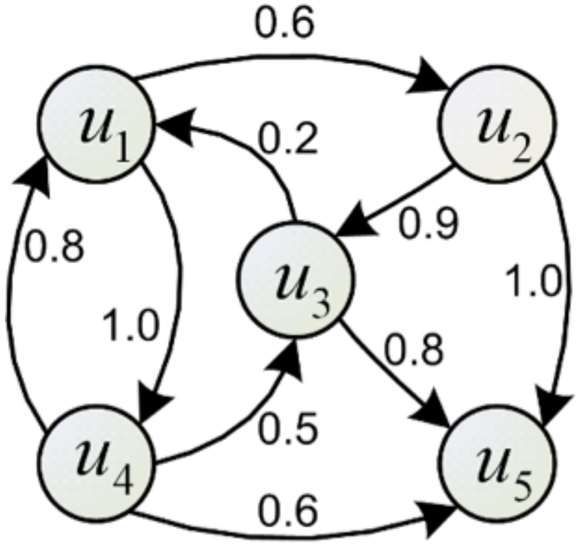
- Minimizing

$$\frac{1}{2}||R - U^T V||_F^2,$$

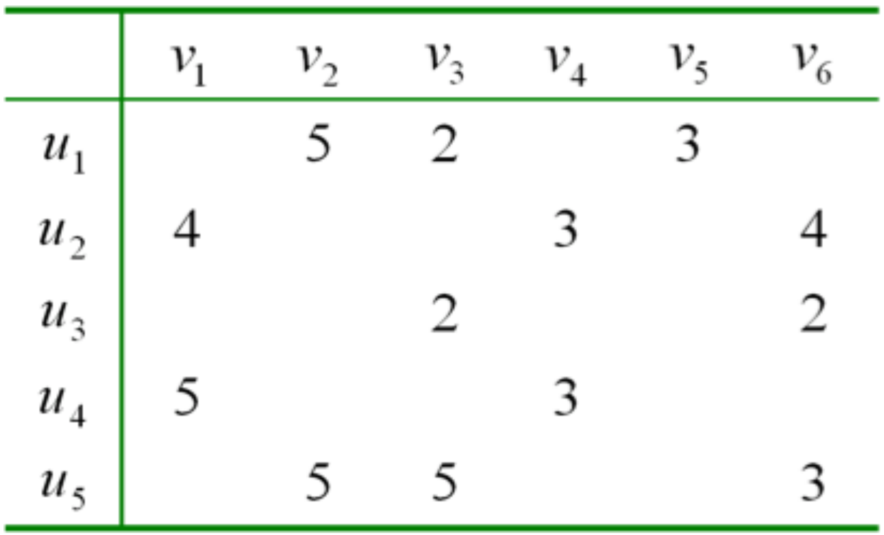
$$\min_{U,V} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}(R_{ij} - U_i^T V_j)^2$$

$$\min_{U,V} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}(R_{ij} - U_i^T V_j)^2 + \frac{\lambda_1}{2}||U||_F^2 + \frac{\lambda_2}{2}||V||_F^2$$

Social Recommendation Using Probabilistic Matrix Factorization



Social Trust Graph

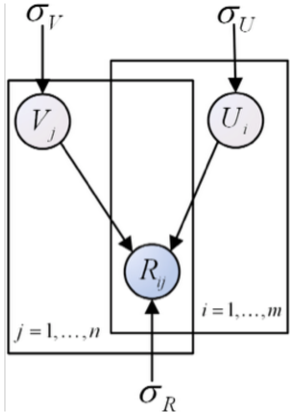


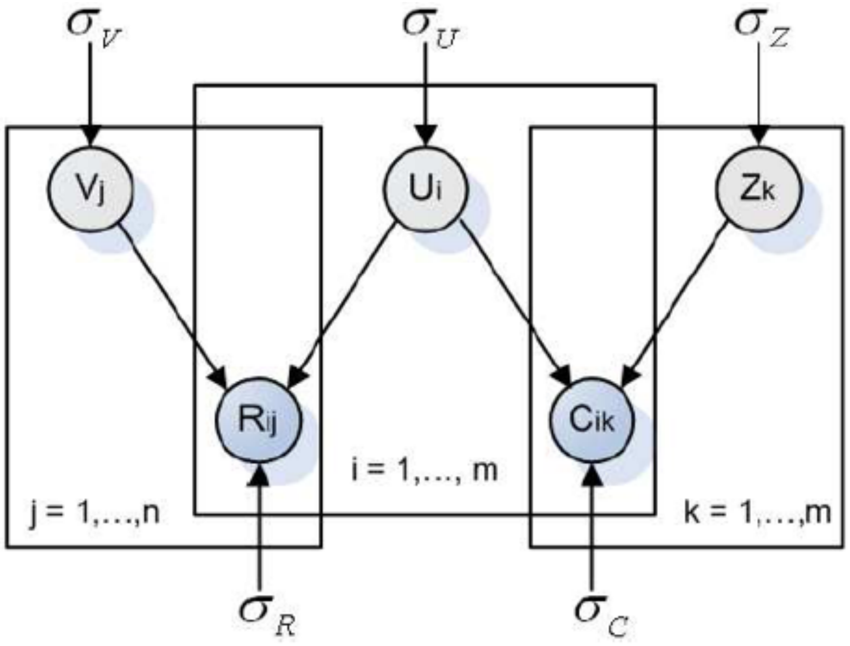
	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3

User-Item Rating Matrix

User-Item Matrix Factorization

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3



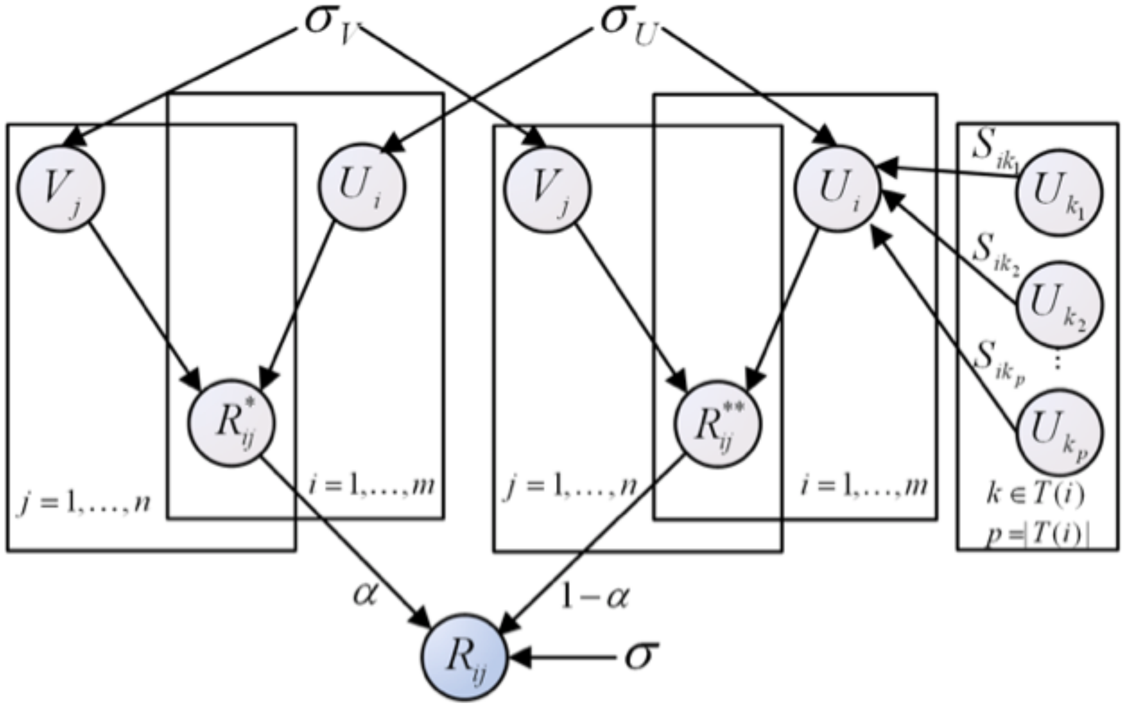


$$\mathcal{L}(R,C,U,V,Z) =$$

$$\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}^{R}(r_{ij}-g(U_i^T V_j))^2+\frac{\lambda_C}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} I_{ik}^{C}(c_{ik}^{*}-g(U_i^T Z_k))^2$$

$$+\frac{\lambda_U}{2}\|U\|_F^2+\frac{\lambda_V}{2}\|V\|_F^2+\frac{\lambda_Z}{2}\|Z\|_F^2,$$

Recommendation with Social Trust Ensemble



$$\prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathcal{N}\left(R_{ij}|g(\alpha U_i^T V_j + (1-\alpha)\sum_{k\in\mathcal{T}(i)} S_{ik}U_k^T V_j), \sigma^2\right)\right]^{I_{ij}^R}$$

Distrust

$$\max_{U} \frac{1}{2}\sum_{i=1}^{m} \sum_{d\in\mathcal{D}+(i)} S_{id}^{\mathcal{D}} \|U_i - U_d\|_F^2$$

$$\begin{aligned}\min_{U,V} \mathcal{L}_{\mathcal{D}}(R, S^{\mathcal{D}}, U, V) &= \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}^{R}(R_{ij} - g(U_i^T V_j))^2 \\ &+ \frac{\beta}{2}\sum_{i=1}^{m}\sum_{d\in\mathcal{D}+(i)} (-S_{id}^{\mathcal{D}}\|U_i - U_d\|_F^2) \\ &+ \frac{\lambda_U}{2}\|U\|_F^2 + \frac{\lambda_V}{2}\|V\|_F^2. \qquad (3)\end{aligned}$$

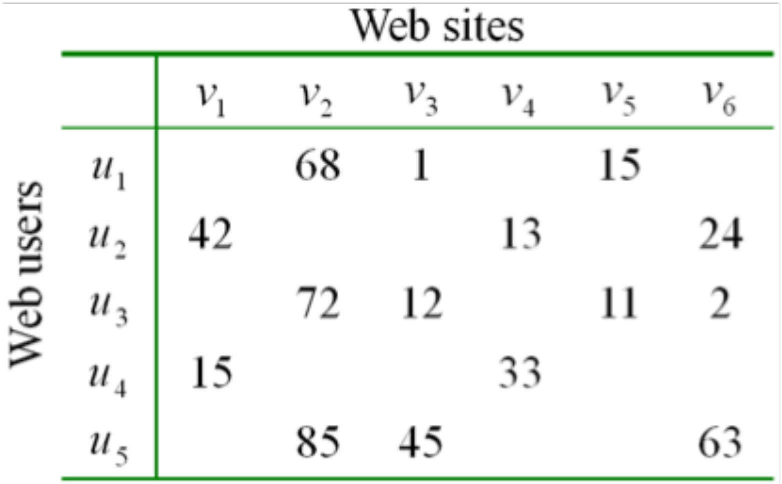
Trust

$$\min_{U} \frac{1}{2} \sum_{i=1}^{m} \sum_{t \in \mathcal{T}^{+}(i)} S_{it}^{\mathcal{T}} \|U_i - U_t\|_F^2$$

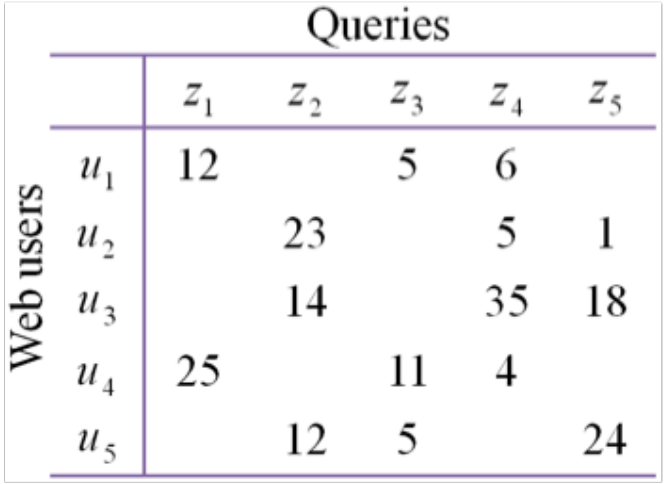
$$\begin{aligned}
\min_{U,V} \mathcal{L}_{\mathcal{T}}(R, S^{\mathcal{T}}, U, V) &= \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}^{R} (R_{ij} - g(U_i^T V_j))^2 \\
&+ \frac{\alpha}{2} \sum_{i=1}^{m} \sum_{t \in \mathcal{T}^{+}(i)} (S_{it}^{\mathcal{T}} \|U_i - U_t\|_F^2) \\
&+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2. \qquad (7)
\end{aligned}$$

Using Clicks as Ratings

ID	Query	URL
358	facebook	http://www.facebook.com
358	rww	http://www.readwriteweb.com
3968	iphone4	http://www.apple.com
3968	ipad	http://www.apple.com
...	...	...

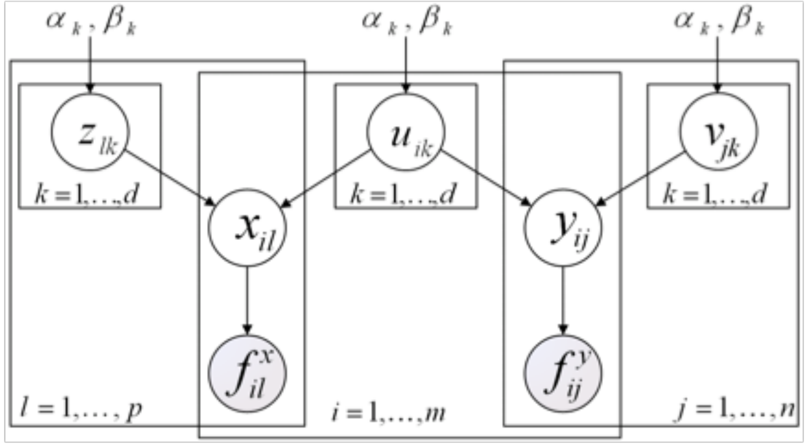


	Web sites					
Web users	v_1	v_2	v_3	v_4	v_5	v_6
u_1		68	1		15	
u_2	42			13		24
u_3		72	12		11	2
u_4	15			33		
u_5		85	45			63



	Queries				
Web users	z_1	z_2	z_3	z_4	z_5
u_1	12		5	6	
u_2		23		5	1
u_3		14		35	18
u_4	25		11	4	
u_5		12	5		24

Collective Probabilistic Factor Model



$$\begin{aligned}
&\mathcal{L}(U, V, Z; F^x, F^y) \\
&= \sum_{i=1}^{m}\sum_{l=1}^{p}(f^x_{il}\ln x_{il} - x_{il}) + \sum_{i=1}^{m}\sum_{j=1}^{n}(f^y_{ij}\ln y_{ij} - y_{ij}) \\
&+ \sum_{i=1}^{m}\sum_{k=1}^{d}((\alpha_k - 1)\ln(u_{ik}/\beta_k) - u_{ik}/\beta_k) \\
&+ \sum_{j=1}^{n}\sum_{k=1}^{d}((\alpha_k - 1)\ln(v_{jk}/\beta_k) - v_{jk}/\beta_k) \\
&+ \sum_{l=1}^{p}\sum_{k=1}^{d}((\alpha_k - 1)\ln(z_{lk}/\beta_k) - z_{lk}/\beta_k) + \text{const.}
\end{aligned}$$

$$u_{ik} \leftarrow u_{ik}\frac{\sum_{j=1}^{n}(f^y_{ij}v_{jk}/y_{ij}) + \sum_{l=1}^{p}(f^x_{il}z_{lk}/x_{il}) + (\alpha_k - 1)/u_{ik}}{\sum_{j=1}^{n}v_{jk} + \sum_{l=1}^{p}z_{lk} + 1/\beta_k}$$

$$v_{jk} \leftarrow v_{jk}\frac{\sum_{i=1}^{m}(f^y_{ij}u_{ik}/y_{ij}) + (\alpha_k - 1)/v_{jk}}{\sum_{i=1}^{m}u_{ik} + 1/\beta_k},$$

$$z_{lk} \leftarrow z_{lk}\frac{\sum_{i=1}^{m}(f^x_{il}u_{ik}/x_{il}) + (\alpha_k - 1)/z_{lk}}{\sum_{i=1}^{m}u_{ik} + 1/\beta_k}.$$

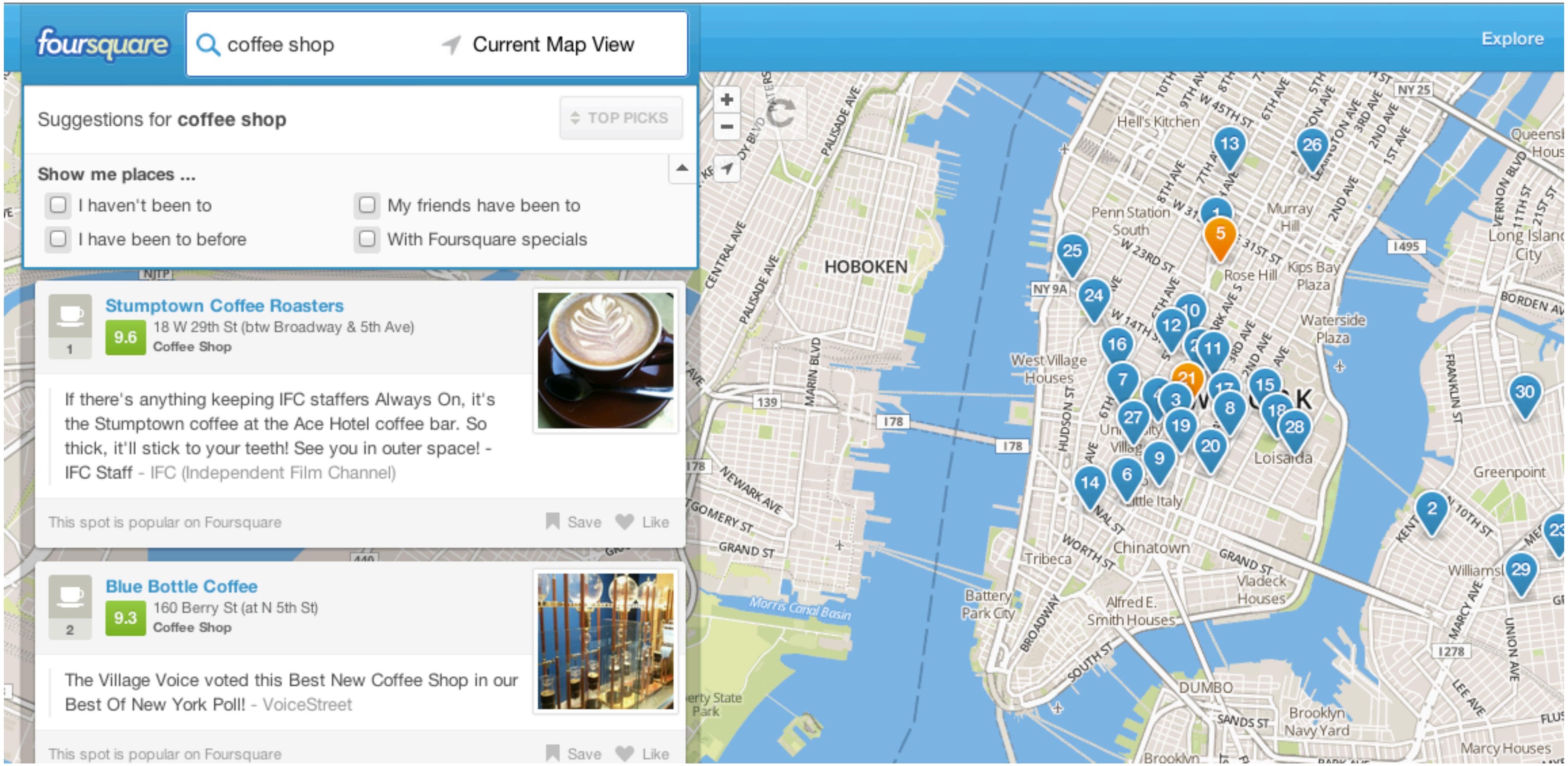
$$u_{ik} \leftarrow u_{ik}\frac{\theta\sum_{j=1}^{n}(f^y_{ij}v_{jk}/y_{ij}) + (1-\theta)\sum_{l=1}^{p}(f^x_{il}z_{lk}/x_{il}) + (\alpha_k - 1)/u_{ik}}{\theta\sum_{j=1}^{n}v_{jk} + (1-\theta)\sum_{l=1}^{p}z_{lk} + 1/\beta_k}$$

Fused Matrix Factorization with Geographical and Social Influence in Location-based Social Networks

Chen Cheng, Haiqin Yang, Irwin King and Michael R. Lyu

AAAI'12, Toronto, Canada

POI Recommendation



coffee shop

Current Map View

Suggestions for **coffee shop**

☐ I haven't been to

☐ My friends have been to

☐ I have been to before

☐ With Foursquare specials

1 **Stumptown Coffee Roasters** 9.6
18 W 29th St (btw Broadway & 5th Ave)
Coffee Shop

If there's anything keeping IFC staffers Always On, it's the Stumptown coffee at the Ace Hotel coffee bar. So thick, it'll stick to your teeth! See you in outer space! - IFC Staff - IFC (Independent Film Channel)

This spot is popular on Foursquare

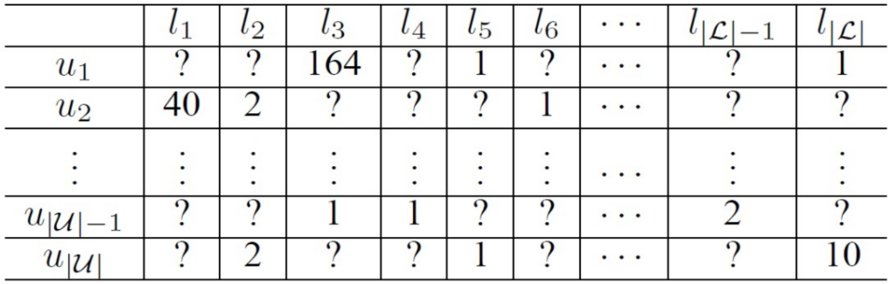
2 **Blue Bottle Coffee** 9.3
160 Berry St (at N 5th St)
Coffee Shop

The Village Voice voted this Best New Coffee Shop in our Best Of New York Poll! - VoiceStreet

This spot is popular on Foursquare

Challenges

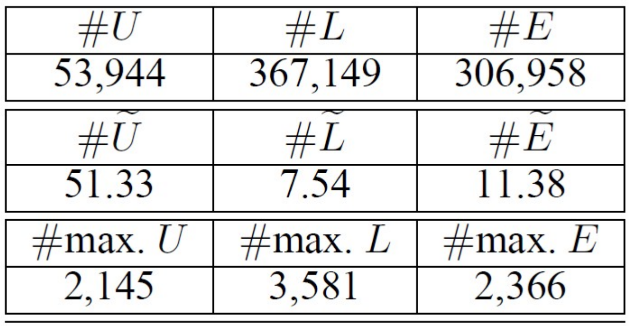
- Large dataset
  - Crawled from Gowalla from Feb. 2009 to Sep. 2011
  - 4,128,714 check-ins from 53,944 users on 367,149 locations
- Only positive data is seen
- Sparsity : density of our dataset is only 0.0208%



	l_1	l_2	l_3	l_4	l_5	l_6	$\cdots$	$l_{\lvert\mathcal{L}\rvert-1}$	$l_{\lvert\mathcal{L}\rvert}$
u_1	?	?	164	?	1	?	$\cdots$	?	1
u_2	40	2	?	?	?	1	$\cdots$	?	?
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$u_{\lvert\mathcal{U}\rvert-1}$	?	?	1	1	?	?	$\cdots$	2	?
$u_{\lvert\mathcal{U}\rvert}$	?	2	?	?	1	?	$\cdots$	?	10

Figure 1: User-location check-in frequency matrix.

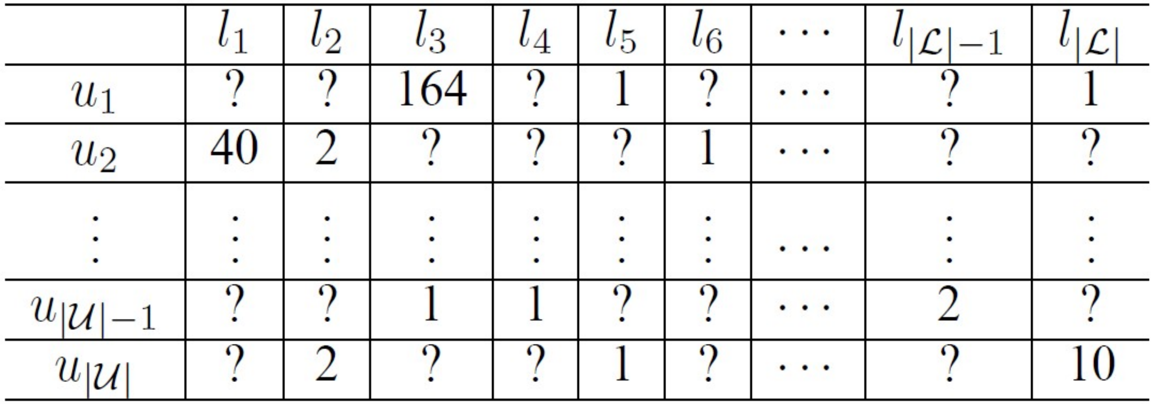
Table 1: Basic statistics of the Gowalla dataset.



$\#U$	$\#L$	$\#E$
53,944	367,149	306,958
$\#\widetilde{U}$	$\#\widetilde{L}$	$\#\widetilde{E}$
51.33	7.54	11.38
$\#$max. U	$\#$max. L	$\#$max. E
2,145	3,581	2,366

POI Recommendation in LBSNs

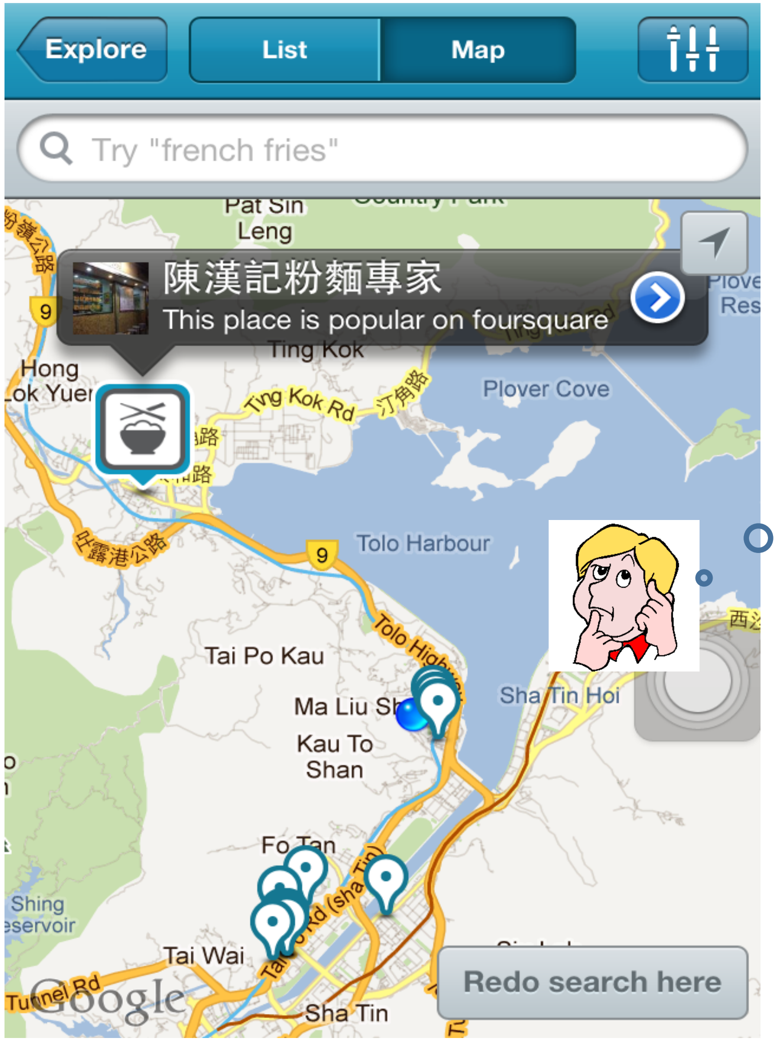
- **Matrix Factorization** can be a promising tool
- However, **Geographical influence** is ignored!



	l_1	l_2	l_3	l_4	l_5	l_6	$\cdots$	$l_{\lvert\mathcal{L}\rvert-1}$	$l_{\lvert\mathcal{L}\rvert}$
u_1	?	?	164	?	1	?	$\cdots$	?	1
u_2	40	2	?	?	?	1	$\cdots$	?	?
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$u_{\lvert\mathcal{U}\rvert-1}$	?	?	1	1	?	?	$\cdots$	2	?
$u_{\lvert\mathcal{U}\rvert}$	?	2	?	?	1	?	$\cdots$	?	10

Figure 1: User-location check-in frequency matrix.

Geographical Influence is Important

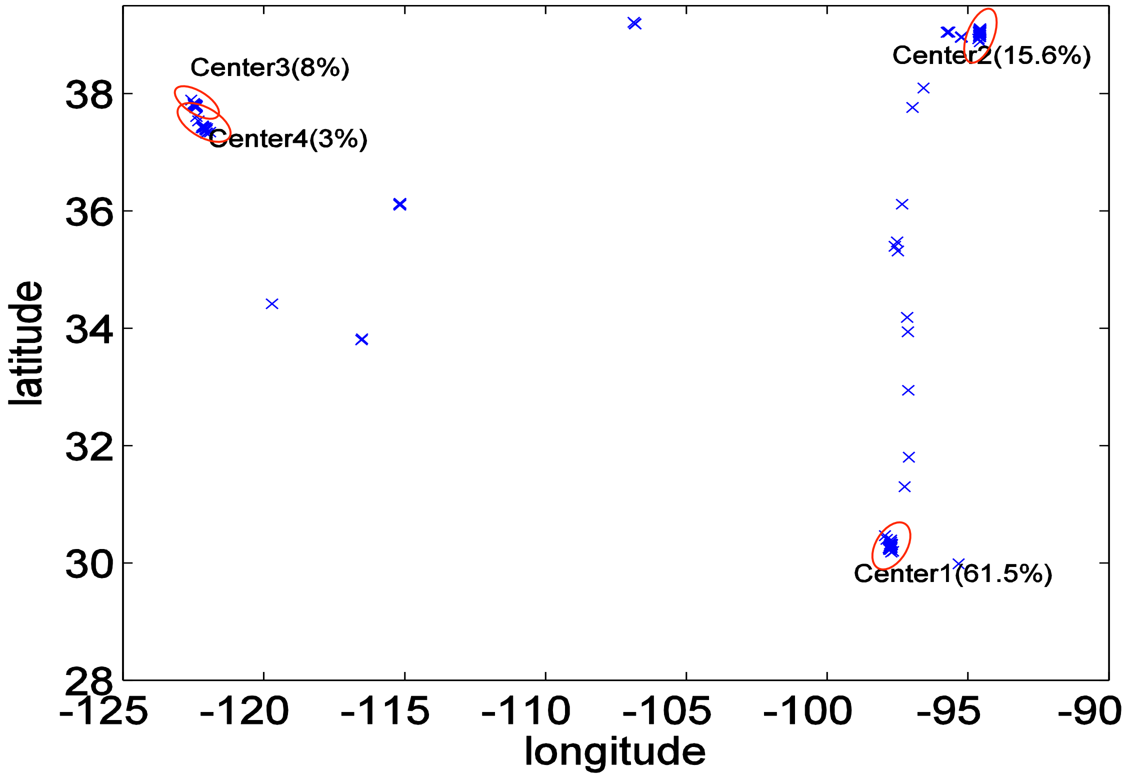


陳漢記粉麵專家

This place is popular on foursquare

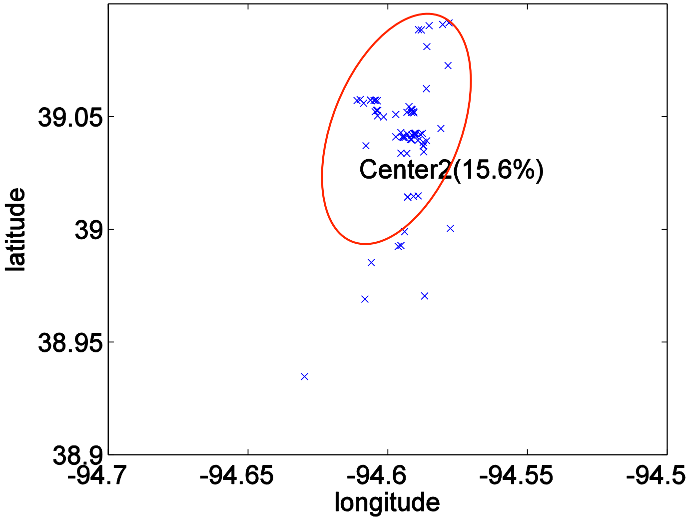
Er... a little far..

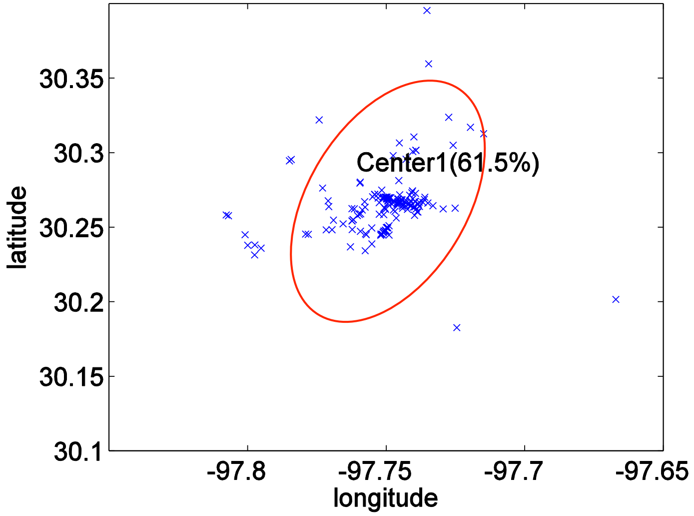
Multi-centers and Normal Distribution

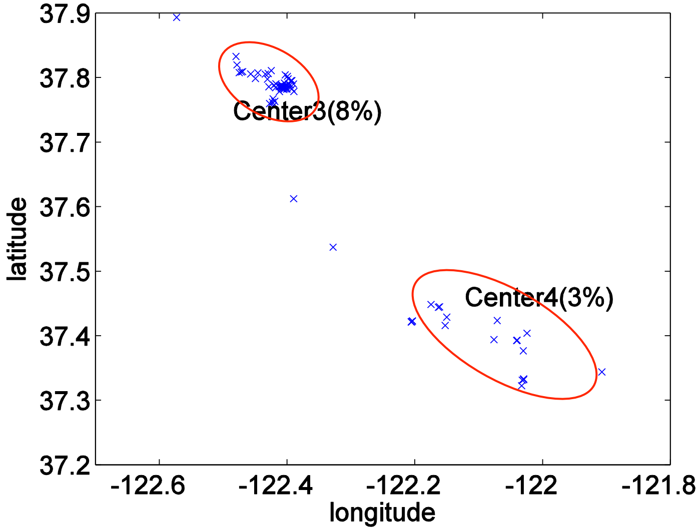


- Two centers (home & office) in [Cho et al., ’11]
- Several centers proposed in our paper

Multi-centers and Normal Distribution

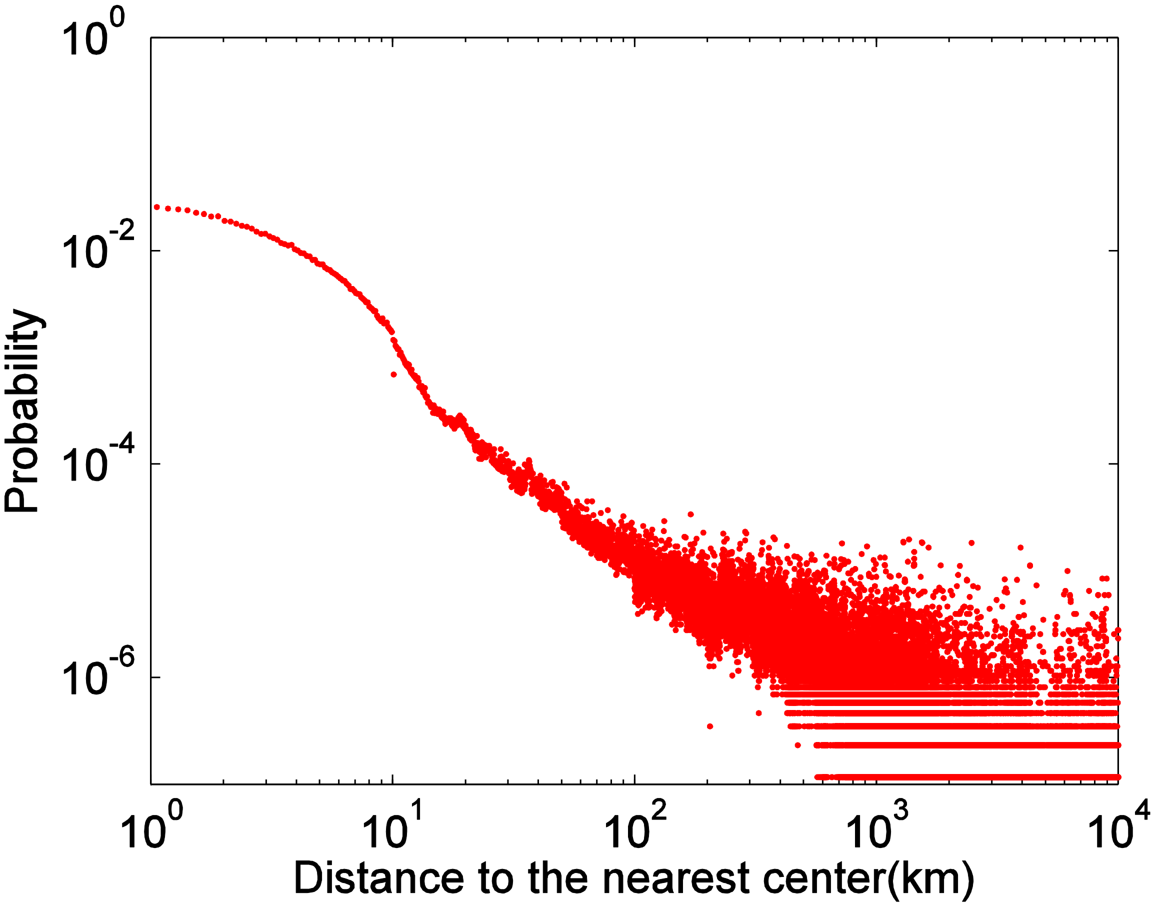






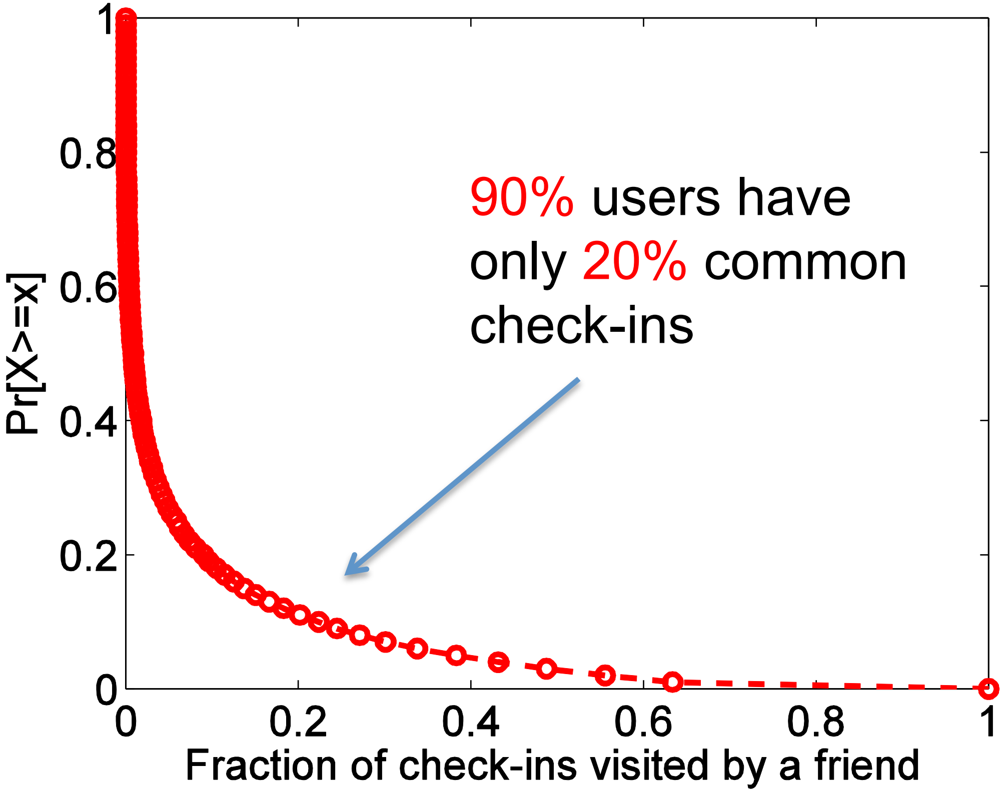
Similar to [Brockmann, ’06; Gonzalez, ’08] , we assume each center follow the norm distribution

Inverse Distance Rule



Social Influence

- On average, overlap of a user's check-ins to his friends only about 9.6%



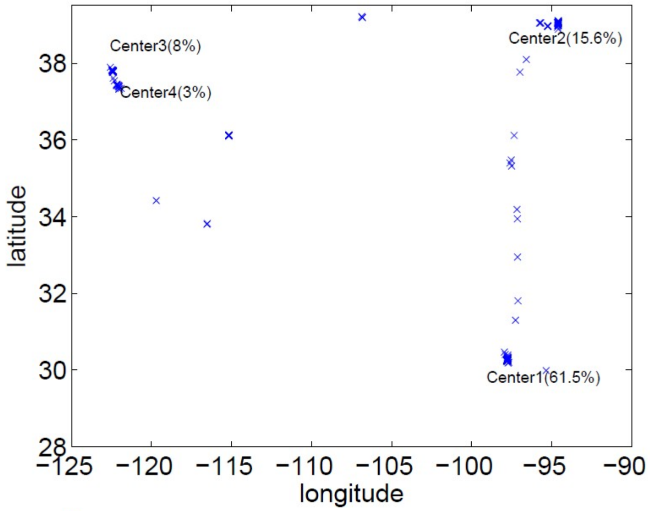
90% users have only 20% common check-ins

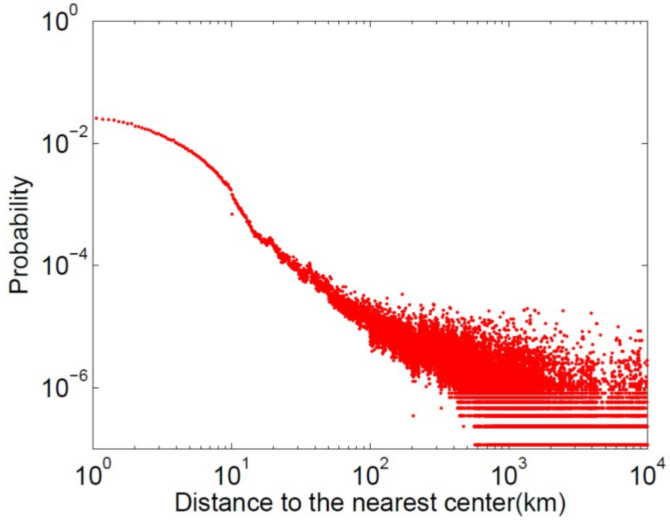
Our Proposal

- Multi-center Gaussian Model (**MGM**) to capture geographical influence
- Propose a generalized fused matrix factorization framework to include social and geographical influences
- Conduct thorough experiments on large-scale Gowalla dataset

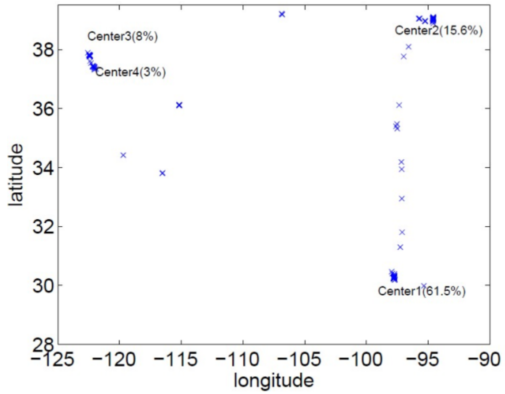
Multi-center Gaussian Model

- Recall check-in locations are located around several centers
- The probability a user visiting a location is inversely proportional to the distance from its nearest center
- MGM is proposed to model users' check-in behavior





Multi-center Gaussian Model

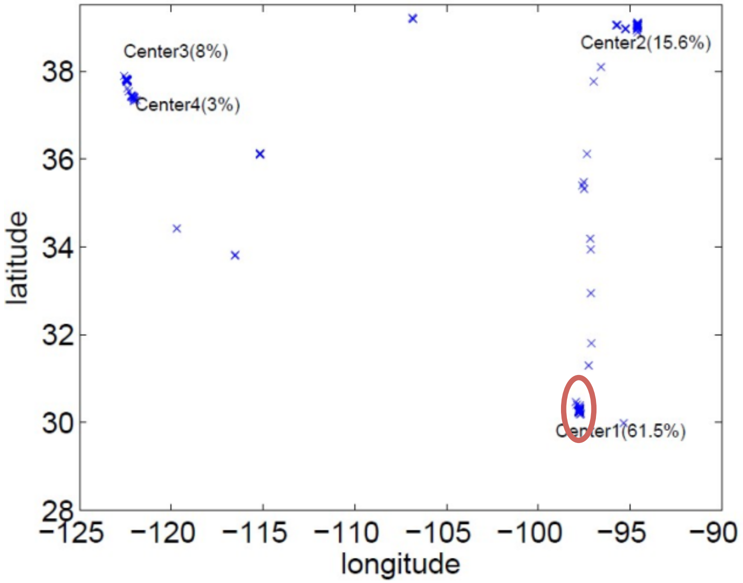


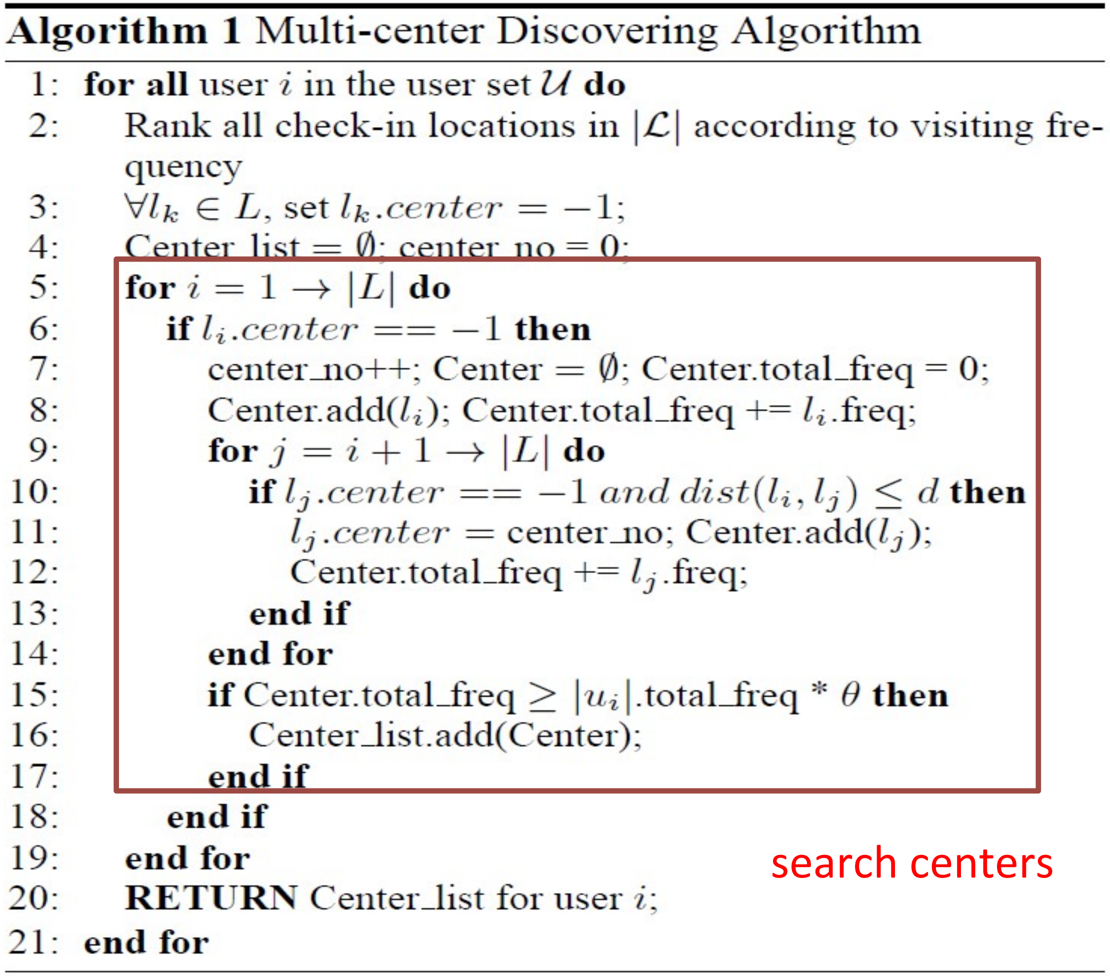
- Notation
  - C_u : multi-center set for user *u*
  - f_{c_u} : total frequency at center c_u for user *u*
  - $\mathcal{N}(l|\mu_{c_u}, \Sigma_{c_u})$ is : the pdf of Gaussian distribution, μ_{c_u} and Σ_{c_u} denote the mean and covariance matrices of regions around center c_u
- The probability a user *u* visiting a location l given c_u defined as:

$$P(l|C_u) = \sum_{c_u=1}^{|C_u|} \underbrace{P(l \in c_u)}_{\propto 1/dist(l,c_u)} \underbrace{\frac{f_{c_u}^{\alpha}}{\sum_{i \in C_u} f_i^{\alpha}}}_{\text{norm effect of check in freq on center } c_u} \frac{\mathcal{N}(l|\mu_{c_u}, \Sigma_{c_u})}{\sum_{i \in C_u} \mathcal{N}(l|\mu_i, \Sigma_i)}$$

Multi-center Discovering Algorithm

A greedy clustering algorithm is proposed due to Pareto principle (top 20 locations cover about 80% check-ins)





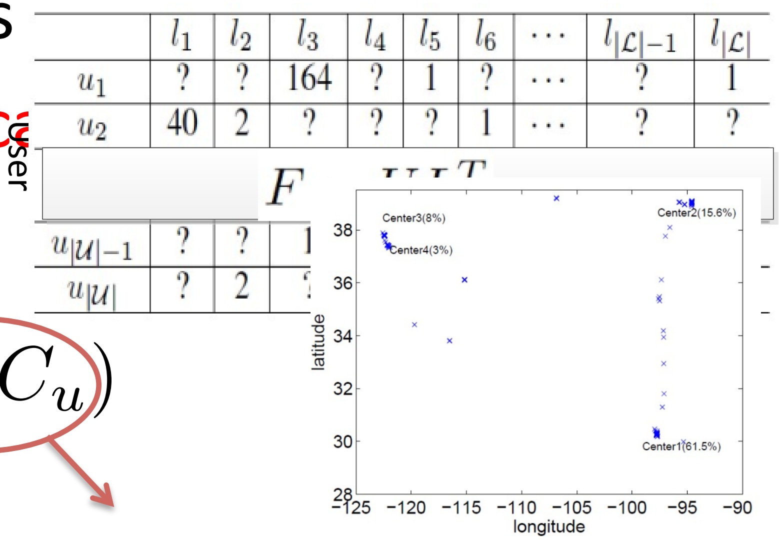
Algorithm 1 Multi-center Discovering Algorithm

```
for all user $i$ in the user set $\mathcal{U}$ do
    Rank all check-in locations in $|\mathcal{L}|$ according to visiting frequency
    $\forall l_k \in L$, set $l_k.center = -1$;
    Center_list = $\emptyset$; center_no = 0;
    for $i = 1 \to |L|$ do
        if $l_i.center == -1$ then
            center_no++; Center = $\emptyset$; Center.total_freq = 0;
            Center.add($l_i$); Center.total_freq += $l_i$.freq;
            for $j = i + 1 \to |L|$ do
                if $l_j.center == -1$ and $dist(l_i, l_j) \leq d$ then
                    $l_j.center$ = center_no; Center.add($l_j$);
                    Center.total_freq += $l_j$.freq;
                end if
            end for
            if Center.total_freq $\geq |u_i|$.total_freq * $\theta$ then
                Center_list.add(Center);
            end if
        end if
    end for
    RETURN Center_list for user $i$;
end for
```

search centers

Fused Framework

- Traditional Matrix Factorization (MF) only model users' preference on locations
- MGM only models geographic
- We can fuse both of them



	l_1	l_2	l_3	l_4	l_5	l_6	$\cdots$	$l_{\|\mathcal{L}\|-1}$	$l_{\|\mathcal{L}\|}$
u_1	?	?	164	?	1	?	$\cdots$	?	1
u_2	40	2	?	?	?	1	$\cdots$	?	?
$u_{\|\mathcal{U}\|-1}$	?	?	1						
$u_{\|\mathcal{U}\|}$	?	2							

$$\underbrace{P_{ul}}_{\text{prob. user } u \text{ visit location } l} = \underbrace{P(F_{ul})}_{\text{encode user preference based on MF}} \cdot \underbrace{P(l|C_u)}_{\text{calculated by MGM}}$$

Setup and Metric

- Split the dataset into 2 non-overlapping sets
  - Randomly select x% for each user as training data and the rest (1-x)% as the test data, x set to 70 and 80
  - Carried out 5 times independently, we report the average
- POI recommendation
  - Return top-*N* POIs for each user
  - Find out # of locations in test dataset are recovered
- Metric

$$Precision@N = \frac{\#\, of\; recovered\; POIs}{N}$$

$$Recall@N = \frac{\#\, of\; recovered\; POIs}{\#\, of\; total\; missing\; POIs}$$

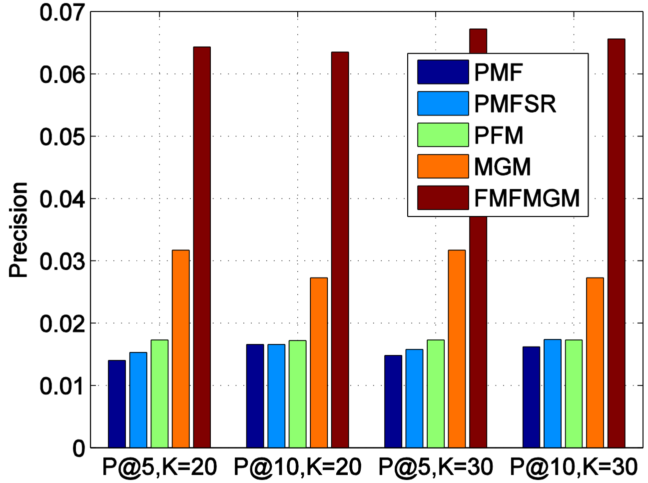
Comparison Methods

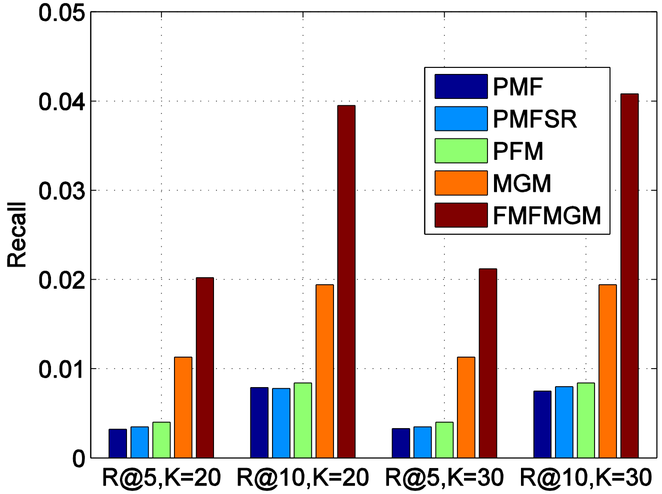
- MGM
- PMF: [Salakhutdinov and Mnih, ’07]
- PMF with Social Regularization (PMFSR): [Ma et al., ’11b]
- Probabilistic Factor Model (PFM): [Ma et al., ’11a]
- Fused MF with MGM (FMFMGM): our proposed method

Results

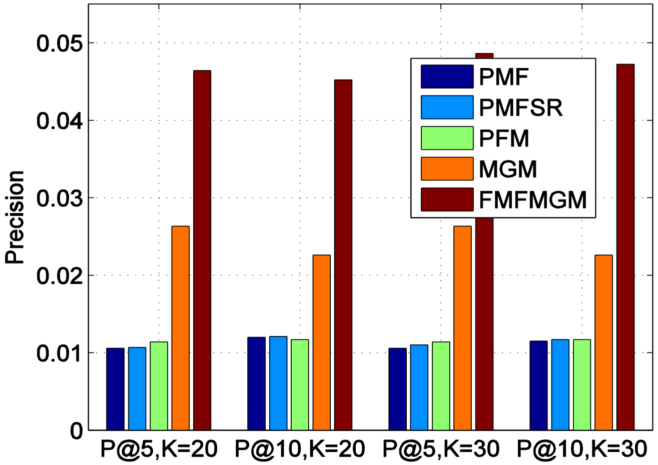
Precision

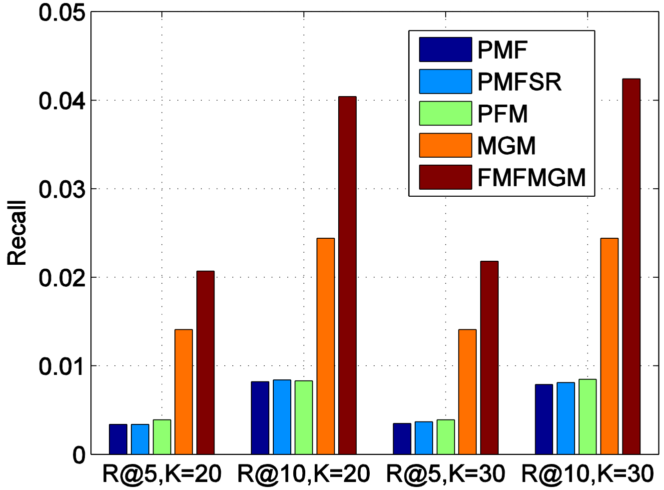
Recall





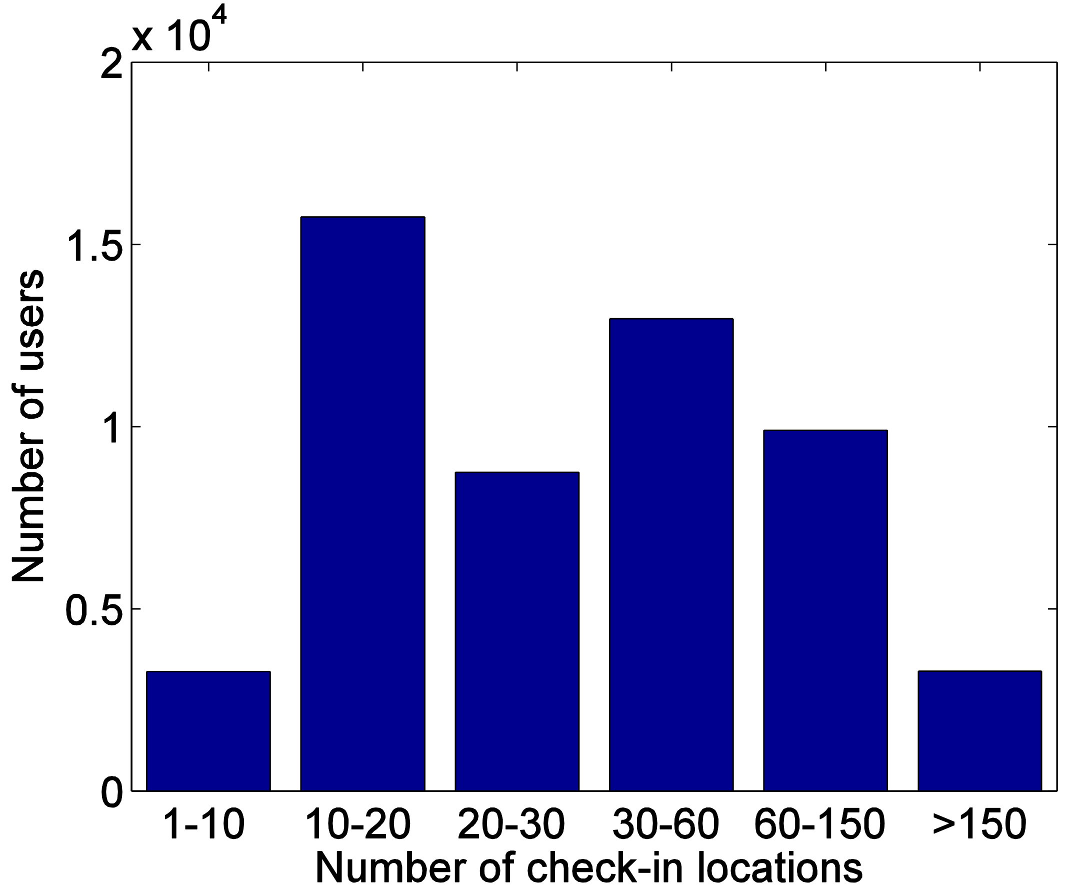
70%



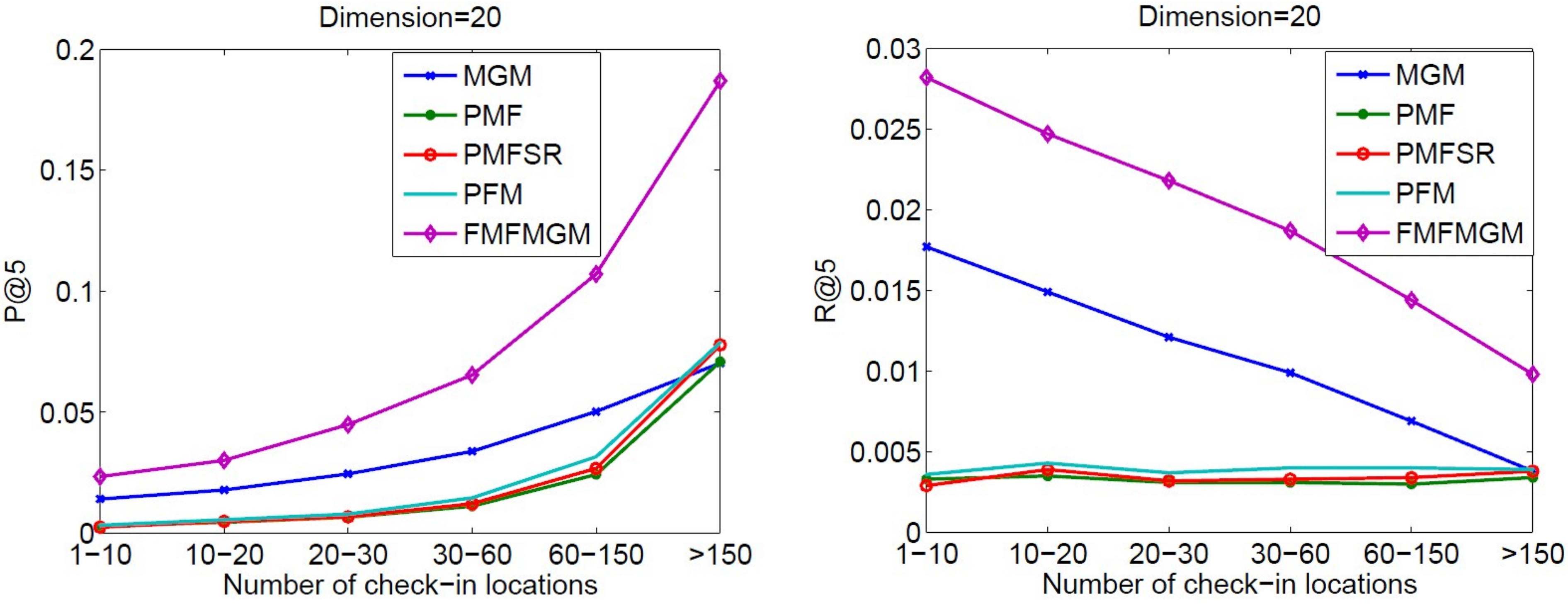


80%

User Check-in Distribution



Performance on Different Users



Conclusions

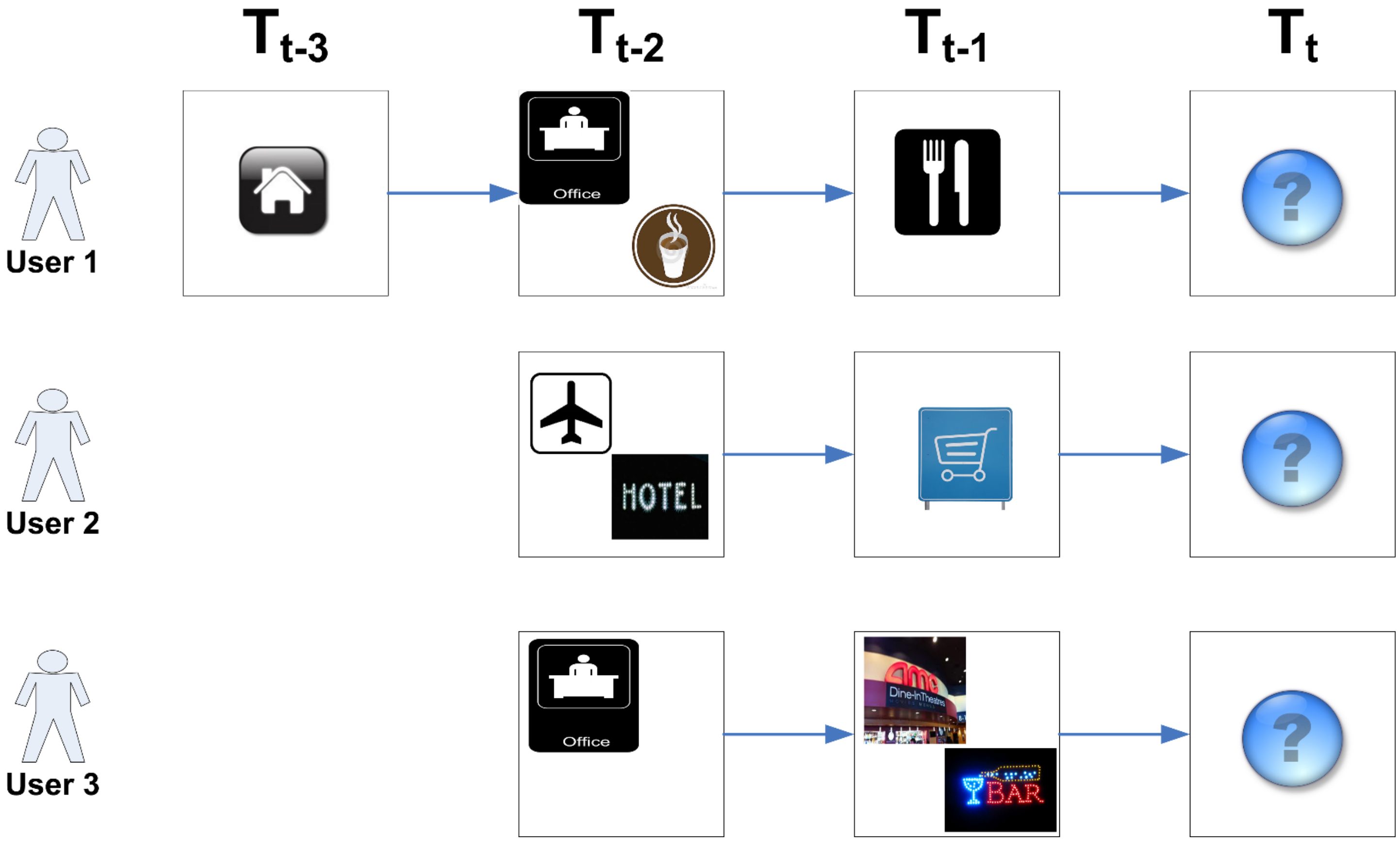
- Extract characteristics of a large dataset crawled from Gowalla
- Propose a novel Multi-center Gaussian Model (MGM) to model geographical influence
- Propose a fused MF framework which outperforms state-of-the-art methods

Where You Like to Go: Next Successive Point-of-Interest Recommendation

Chen Cheng, Haiqin Yang, Irwin King and Michael R. Lyu

IJCAI'13, Beijing, China

Successive POI Recommendation

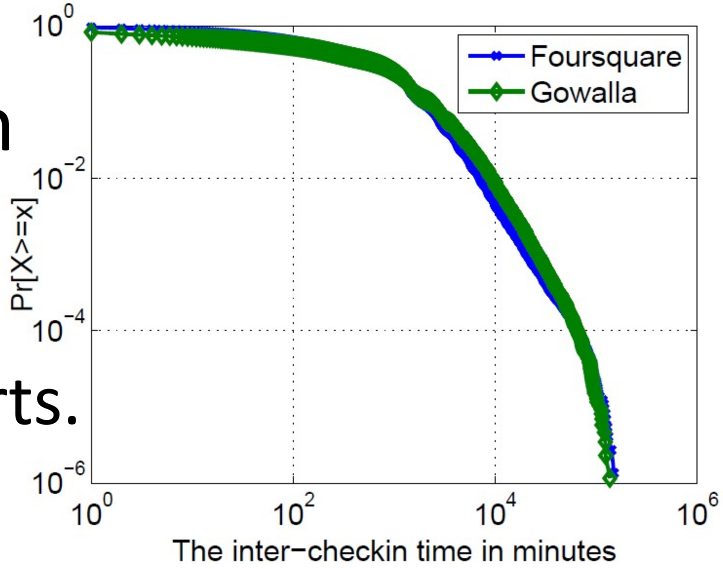


Two Main Properties in LBSNs Dataset

- Personalized Markov chain
- Localized region constraint

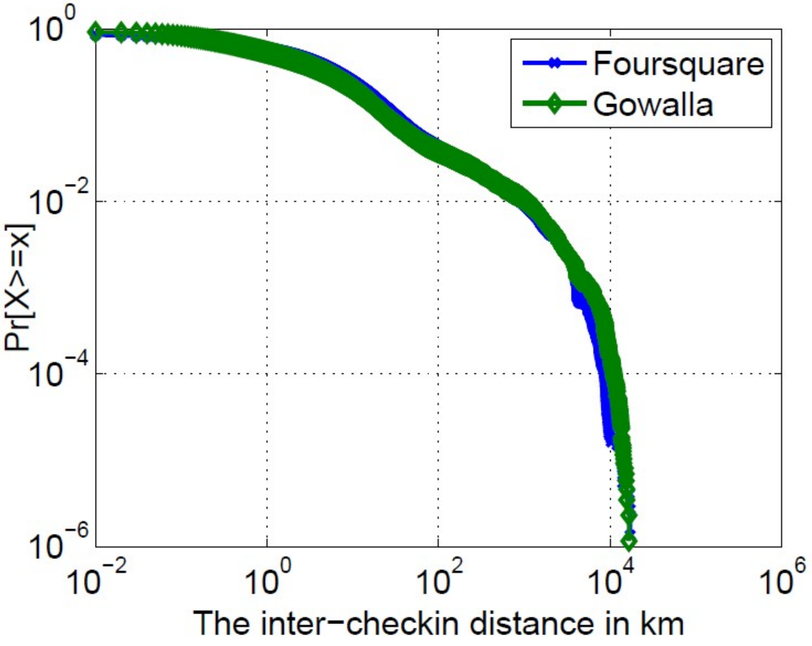
Personalized Markov Chain

- Inter check-in time
  - Around 45% successive check-ins within 2h, 70% within 12h.
- Strong connections between inter check-ins
  - E.g. cinemas or bars after restaurant, hotels after airports.
- Motivated to use transition probability



Localized Region Constraint

- Most inter check-ins occurs at nearby locations
  - 75% within 10km, less than 5% beyond 100 km.
- We can only consider the new POIs near a user's previous check-ins when providing successive POI recommendation.



Example

User 1 **Localized Region Constraint** **User 2**

Our Proposal

- We propose Factoring Personalize Markov Chain with Localized Region model (FPMC-LR).
  - Combine the personalize Markov chain and localized region constraint
  - Although borrows the idea of FPMC [Rendle et al. ’10], we emphasize on users’ movement constraint and focus on a different problem

Problem Definition

- Notation:
  - $\mathcal{U}$: users, $\mathcal{L}$: locations, $\mathcal{L}_u$: the check-in history of user u
  - *T*: slice window to construct a set check-ins, $\mathcal{T}$: time window set
  - $\mathcal{L}_u^t$: check-in time of user *u* at time t , $t \in \mathcal{T}$
- Problem:
  - Given a sequence of check-ins, $\mathcal{L}_u^1, \ldots, \mathcal{L}_u^t$, the (lat, lng) pair of locations , recommend POIs to users at *t*+1

Model

- FPMC-LR is to recommend a successive personalized POI by the prob. a user *u* will visit at time *t*:

$$x_{u,i,l} = p(l \in \mathcal{L}_u^t | i \in \mathcal{L}_u^{t-1})$$

- Base on first-order Markov chain property

$$p(l \in \mathcal{L}_u^t | \mathcal{L}_u^{t-1}) = \frac{1}{|\mathcal{L}_u^{t-1}|} \sum_{i \in \mathcal{L}_u^{t-1}} \boxed{p(l \in \mathcal{L}_u^t | i \in \mathcal{L}_u^{t-1})}$$

Prob. for user *u* from location *i* to *l*

Model

- FPMC-LR only consider the neighborhood locations of previous check-ins

$$N_d(\mathcal{L}_u^t) = \{l \in \mathcal{L} \setminus \mathcal{L}_u^{t-1} : D(l, l_0) \leq d, \forall l_0 \in \mathcal{L}_u^{t-1}\}$$

- Thus our FPMC-LR yields a transition tensor

$$\mathcal{X} \in [0,1]^{|\mathcal{U}| \times |\mathcal{L}| \times |N_d(\mathcal{L})|}$$

  – Note: $|N_d(\mathcal{L})|$ is reduced largely compared to $|\mathcal{L}|$, around 100 when d = 40 km

Model

- Use the same idea in [Rendle et al, '10], we approximate the tensor as:

$$\hat{x}_{u,i,l} = \boldsymbol{v}_u^{\mathcal{U},\mathcal{L}} \cdot \boldsymbol{v}_l^{\mathcal{L},\mathcal{U}} + \boldsymbol{v}_l^{\mathcal{L},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{L}} + \boldsymbol{v}_u^{\mathcal{U},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{U}}$$

where $\boldsymbol{v}_u^{\mathcal{U},\mathcal{L}}$ and $\boldsymbol{v}_l^{\mathcal{L},\mathcal{U}}$ model the latent features for users and the next locations, respectively.

– This gives the set of model parameters , i.e.,

$$\Theta = \{\boldsymbol{V}^{U,L}, \boldsymbol{V}^{L,U}, \boldsymbol{V}^{U,I}, \boldsymbol{V}^{I,U}, \boldsymbol{V}^{L,I}, \boldsymbol{V}^{I,L}\}$$

Model

- Model top-k recommendations as a ranking over locations:

$$i >_{u,t} j :\Leftrightarrow \hat{x}_{u,t,i} > \hat{x}_{u,t,j}$$

- The MAP estimator is

$$\arg\max_{\Theta} \sum_{u \in \mathcal{U}} \sum_{\mathcal{L}_u^t \in \mathcal{L}_u} \sum_{i \in \mathcal{L}_u^t} \sum_{j \in N(\mathcal{L}_u^{t-1}) \backslash \mathcal{L}_u^t} \ln \sigma(\hat{x}_{u,t,i} - \hat{x}_{u,t,j}) - \lambda_{\Theta} ||\Theta||_F^2$$

- Learning algorithm: Stochastic gradient descent

Data Set

- Two publicly available data sets: Foursquare and Gowalla

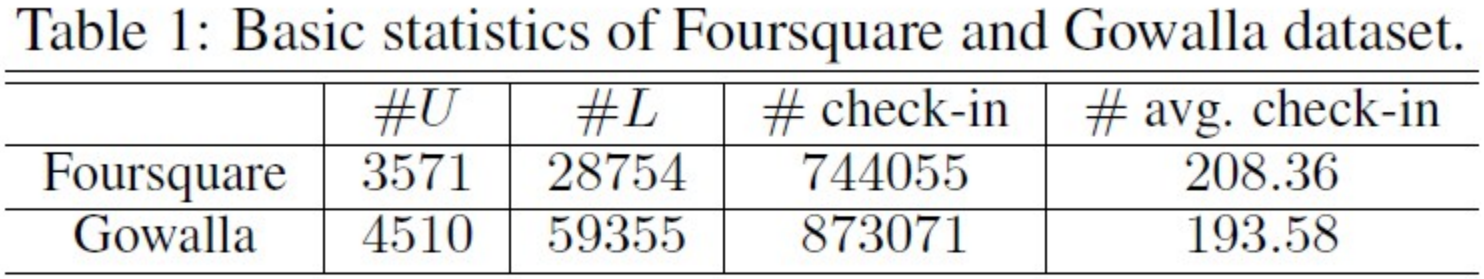


Table 1: Basic statistics of Foursquare and Gowalla dataset.

	$\#U$	$\#L$	# check-in	# avg. check-in
Foursquare	3571	28754	744055	208.36
Gowalla	4510	59355	873071	193.58

Experiment: Comparison

- Compared methods
  - – PMF: proposed by [Salakhudinov and Mnih, ’07]
  - – PTF: proposed by [Xiong et al., ’07].
  - – FPMC: proposed by [Rendle et al. ’10].
- Metric

$$\mathrm{P}@N := \frac{|S|}{N}, \;\; \mathrm{R}@N := \frac{|S|}{|\mathcal{L}_u^{t+1}|}$$

Results

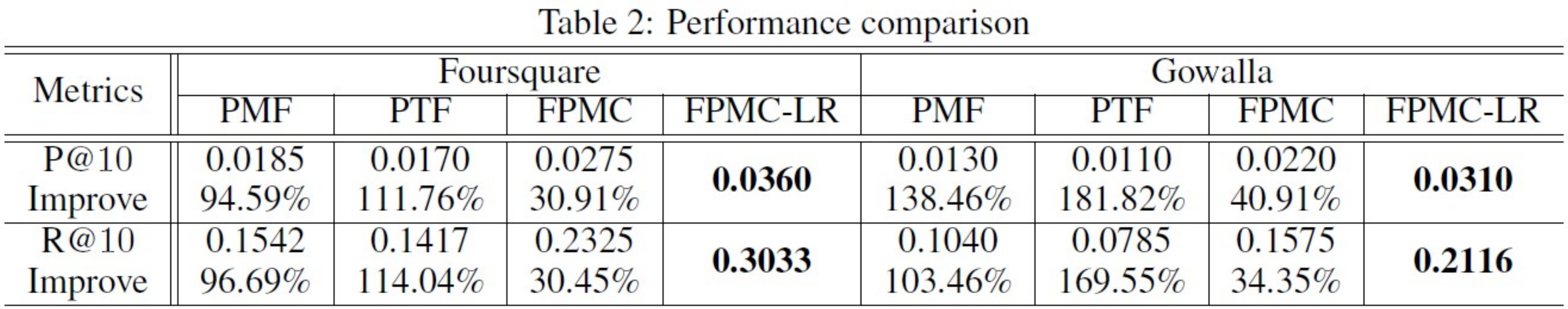
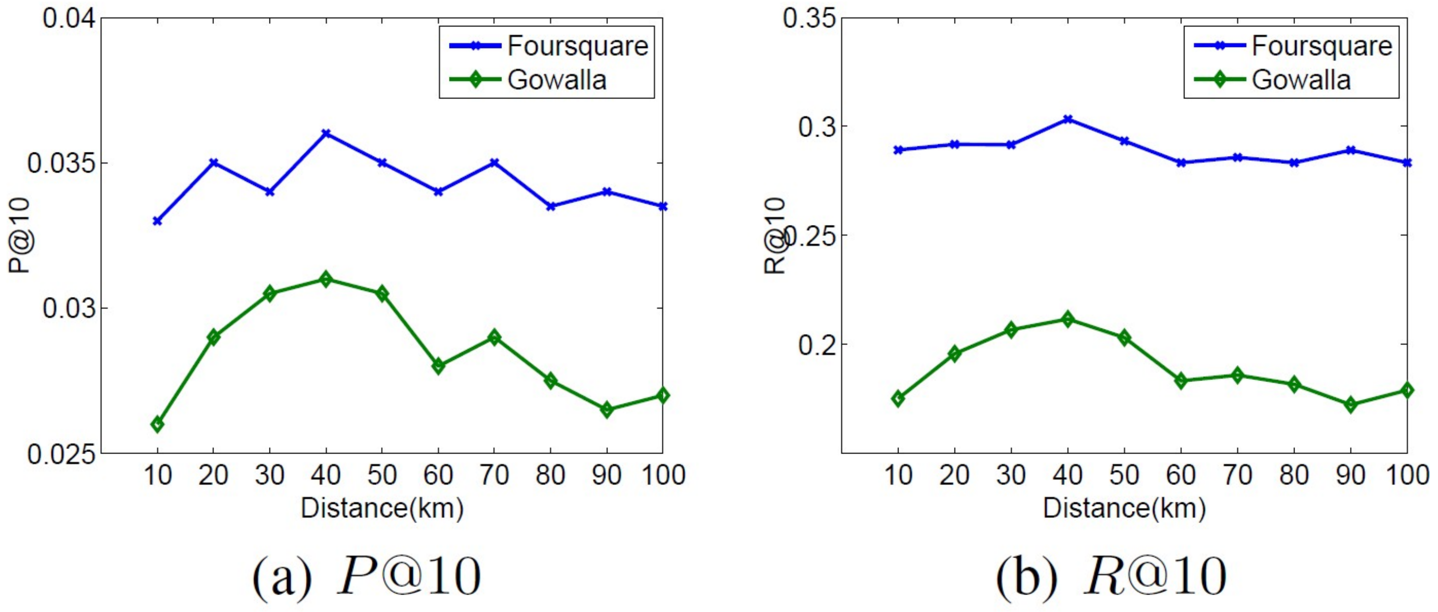


Table 2: Performance comparison

Metrics	Foursquare				Gowalla			
	PMF	PTF	FPMC	FPMC-LR	PMF	PTF	FPMC	FPMC-LR
P@10 Improve	0.0185 94.59%	0.0170 111.76%	0.0275 30.91%	**0.0360**	0.0130 138.46%	0.0110 181.82%	0.0220 40.91%	**0.0310**
R@10 Improve	0.1542 96.69%	0.1417 114.04%	0.2325 30.45%	**0.3033**	0.1040 103.46%	0.0785 169.55%	0.1575 34.35%	**0.2116**

- Both FPMC and FPMC-LR outperforms PMF and PTF
  - Importance of personalize Markov chain
- PMF performs better than PTF
  - Latent features are similar to previous time is not valid in LBSNs data
- FPMC-LR performs better than FPMC
  - Localized region constraint can reduce noisy information and achieve better results compared to consider all locations.

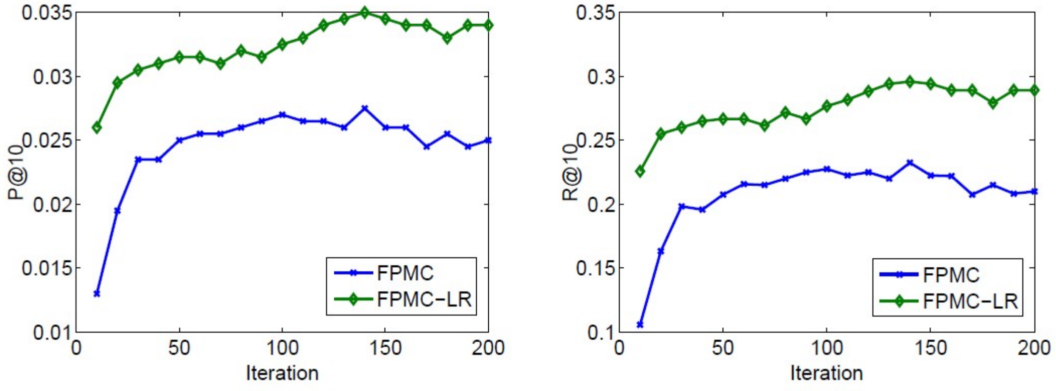
Impact of Parameter *d*



(a) $P@10$ (b) $R@10$

- *d* = 40 km is best.
  - d is too small: do not include enough information which yields suboptimal performance
  - d is too large: introduce noisy information, extreme case is FPMC

Convergence and Efficiency Analysis



(a) $P@10$(Foursquare) (b) $R@10$(Foursquare)

- Each iteration we draw 2×10^5 quadruples, FPMC-LR attains best performance around 150 iterations
- Each iteration takes around 30s, and FPMC-LR is much more efficient at recommendation time than FPMC: consider only the neighbor locations, almost 0.4% of total locations

Conclusions

- We propose FPMC-LR model to solve the successive POI recommendation in LBSNs
- FPMC-LR reduces computation cost largely compared to FPMC
- The performance on two large dataset shows the effectiveness of our model

Conclusion

- LBSs are becoming more and more important!
- Combine social and geographical information
- Indoor and outdoor LBSs
- Living analytics!

Acknowledgments

- Shouyuan Chan (Ph.D.)
- Xixian Chen (Ph.D.)
- Chen Cheng (Ph.D.)
- Junjie Hu (Ph.D.)
- Baichuan Li (Ph.D.)
- Guang Ling (Ph.D.)
- Haiqin Yang (Postdoc)
- Connie Yuen (PT Ph.D.)

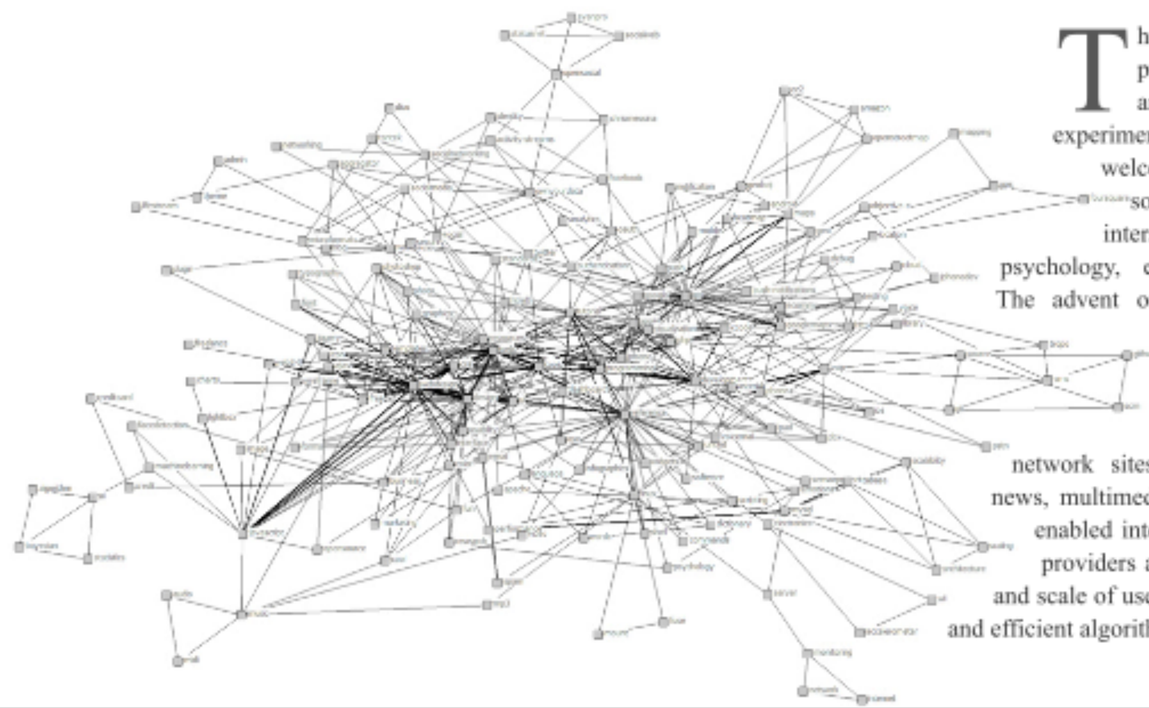
- Hongyi Zhang (Ph.D.)
- Shenglin Zhao (Ph.D.)
- Tong Zhao (Ph.D.)

- Patrick Lau
- Yuanyuan Man
- Shuai Wang
- Raymond Yeung

New Book Series

SOCIAL MEDIA & SOCIAL COMPUTING

CALL FOR PAPERS!



The Social Media and Social Computing Series focuses on publishing high quality references in the rapidly emerging area of social media and social computing. Both experimental/practical as well as theoretical investigations are welcome. The series targets both scholars and practitioners in social media and social computing for work in the intersection of computer science, information technology, psychology, economics, education and other social sciences. The advent of the Internet and the Web has resulted in social interactions and behaviors through the use of technologies and web services, e.g., hardware devices such as smart phones, tablets, RFID, etc., software services such as wikis, blogs, micro-blogs, social network sites, recommender systems, social bookmarking, social news, multimedia sharing sites, etc. Analyzing these technologically-enabled interactions in their social context will benefit information providers and information consumers. However, the large volume and scale of user-generated contents require effective modeling methods and efficient algorithms to handle these chalenging problems.

Series Editor:

Irwin King

Prof. King is Associate Editor of the IEEE Transactions on Neural Networks (TNN) and IEEE Computational Intelligence Magazine (CIM). He is a senior member of IEEE and a member of ACM, International Neural Network Society (INNS), and VP & Governing Board Member of the Asian Pacific Neural Network Assembly (APNNA) . He serves the Neural Network Technical Committee (NNTC) and the Data Mining Technical Committee under the IEEE Computational Intelligence Society.

https://www.cse.cuhk.edu.hk/irwin.king/home

kin@cse.cuhk.edu.hk

Pleas send Proposals to either the Series Editor or Directly to :

LEONG Li-Ming
Editor, CRC Press
240 Macpherson Road, #08-01,
Pines Industrial Building,
S'pore 348574
li.ming@tandf.com.sg
Tel: (65) 67415166 x 115

VeriGuide

- Similarity text detection system
- Developed at CUHK
- Promote and uphold academic honesty, integrity, and quality
- Support English, Traditional and Simplified Chinese
- Handle .doc, .txt, .pdf, .html, etc. file formats
- Generate detailed originality report including readability

Originality

VeriGuide aims to promote academic integrity with VeriGuide's similarity detection capabilities.

Readability

VeriGuide assists educators to assess students' writing ability over time with our readability features.

Accessibility

VeriGuide serves to support and maintain assignment database for students and educators.

VERIGUIDE

- Promote and uphold academic honesty, integrity, and quality
- Support similarity text detection in English, Traditional Chinese, and Simplified Chinese
- Provide class and assignment management via a graphical web interface
- Maintain assignment databases for institutions providing added security and privacy

WHAT'S NEW?

New service plans for various educational institutions have been released.

The Chinese University Plagiarism IDentification Engine (CUPIDE) is now renamed as VeriGuide.

We are pleased to introduce version 2.0 of VeriGuide: improved general processing of the system, better support of different types of file formats, speedier and more comprehensive Internet search, and many other behind the scene improvements.

WWW2014 Workshop on Web-based Education Technologies (WebET 2014)
April 9, 2014, Seoul, Korea

WebET 2014

The Web has long been recognized as a powerful platform for teaching and learning. The educational community was among the early adopters of the technology and has contributed to its evolution. We are at this point at a major inflection point for Web-based Education Technologies. The convergence ("a perfect storm") of new technologies supporting search, social media, semantics, data mining (Big Data), and others along with current interest to distributed educational pedagogies such as connectivism, behaviorism, and "the flipped classroom" promises to dramatically change Web-based Education Technologies in the near future. The interest in Massive Open Online Courses (MOOCs) has been described as "a tsunami in education" and has re-kindled valuable discussions regarding the role of WebET.

Call for Papers

Important Dates

- Submission Deadline: Jan. 14, 2014
- Author Notification: Feb. 4, 2014
- Final Manuscript: Feb. 12, 2014

The Chinese University of Hong Kong

CUHK Excellence

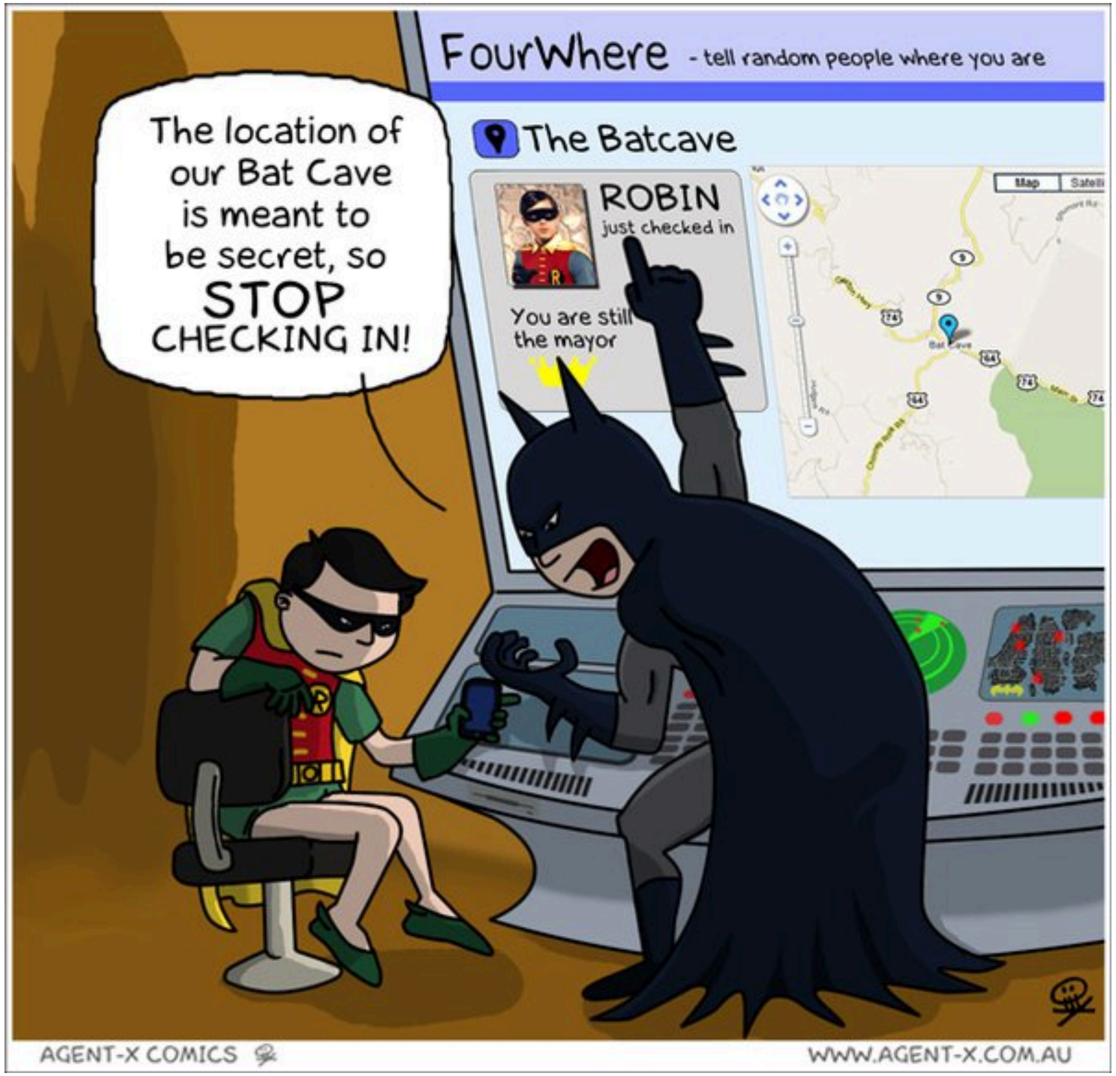
- The only university in Hong Kong having Nobel Laureates as faculty. **Four** Distinguished Professors-at-Large.
- **Nine** academicians of Chinese Academy of Sciences and Chinese Academy of Engineering


Professor Yang Chen-Ning, Nobel Laureate in Physics


Professor Sir James A. Mirrlees, Nobel Laureate in Economic Sciences


Professor Yau Shing-Tung, Fields Medalist


Professor Andrew Yao, Turing Award Winner



http://tricompr.com/blogs/prticles/4-tips/211-asideeffectoflbs

Q & A