Time-Aware Point-of-interest Recommendation

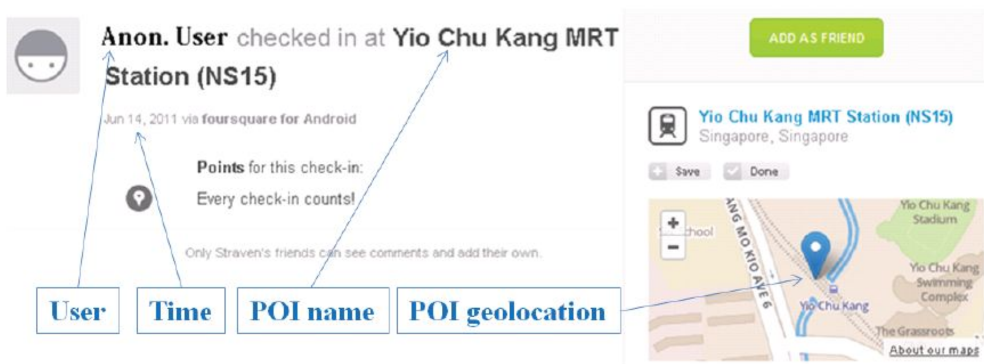
presented by

Yuan Quan
IMI & SCE
Supervisor: Asst Prof. Gao CONG
Co-Supervisor: Prof. Nadia Magnenat Thalmann

06-Nov-2012

Motivation

- Location-based social networks (LBSNs) such as Foursquare, Gowalla, grow rapidly in recent years.
- Users can post physical locations via check-ins.



- **Point-of-interest recommendation**: Recommend unvisited POIs to users
  - Explore new places
  - Know cities better

Motivation

- Two user's checkin behaviors
  - Spatial behavior: Visit nearby POIs
  - Temporal behavior: visit different POIs at different time (have not been used)
- **Time-aware POI recommendation**:

  Given a user u and time t, recommend unknown POIs that u will visit at t.

NANYANG TECHNOLOGICAL UNIVERSITY

Previous work

- Collaborative filtering (CF)
  - Memory-based
  - Model-based
- POI recommendation
  - M. Ye et al. Exploiting geographical influence for collaborative point-of-interest recommendation. In SIGIR 2011.
    - Social information: user-based CF, insignificant
    - Geographical influence: Bayesian model, significant

Preliminary: user-based CF (U)

- *u, v*: two users, *l*: candidate POI
- Calculate similarities between users.
- Produce prediction.

score: $\hat{c}_{u,l} = \frac{\sum_v w_{u,v} c_{v,l}}{\sum_v w_{u,v}}$ similarity: $w_{u,v} = \frac{\sum_l c_{u,l} c_{v,l}}{\sqrt{\sum_l c_{u,l}^2}\sqrt{\sum_l c_{v,l}^2}}$

$C_{u,l}$	l_1	l_2	l_3	l_4
u_1	1	1		
u_2	1	1	1	
u_3		1		1

$$w_{u1,u2} = \cos\left(\langle 1,1,0,0\rangle, \langle 1,1,1,0\rangle\right) = 0.82$$

$$w_{u1,u3} = \cos\left(\langle 1,1,0,0\rangle, \langle 0,1,0,1\rangle\right) = 0.5$$

$$\hat{c}_{u1,l3} = \frac{w_{u1,u2}}{w_{u1,u2} + w_{u1,u3}} = 0.62$$

$$\hat{c}_{u1,l4} = \frac{w_{u1,u3} \quad 5}{w_{u1,u2} + w_{u1,u3}} = 0.38$$

NANYANG
TECHNOLOGICAL
UNIVERSITY

User-based CF with Time Preference (UT)

- Split a day into 24 slots based on hour.
  - *Tony, IMI seminar room,* **Nov. 6 2012, 10:30 ➡ 10**
- Introduce time dimension into u-l matrix: $c_{u,l} \rightarrow c_{u,t,l}$
- Leverage time factor when
  - Computing the similarities between users over all time.

$$w_{u,v}^{(t)} = \frac{\sum_t \sum_l c_{u,t,l} c_{v,t,l}}{\sqrt{\sum_t \sum_l c_{u,t,l}^2} \sqrt{\sum_t \sum_l c_{v,t,l}^2}}$$

  - Making predictions.

$$\hat{c}_{u,t,l} = \frac{\sum_v w_{u,v}^{(t)} c_{v,t,l}}{\sum_v w_{u,v}^{(t)}}$$

The Sparsity Problem

- UT's performance is sensitive to data sparsity.

checkin history of **u** and **v**

c	l_1	l_2
u	1	1
v	1	1

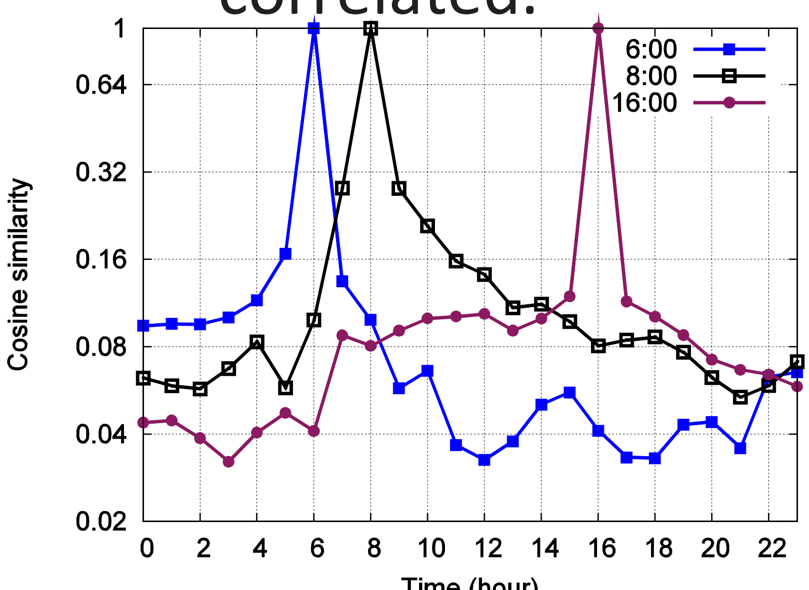
temporal checkin history of **u**

$c_{u,l}$	l_1	l_2
t_1	1	0
t_2	0	1

temporal checkin history of **v**

$c_{v,l}$	l_1	l_2
t_1	0	1
t_2	1	0

- Checkin behaviors of users at different time are correlated.



- For each user, calculate similarity between every pair of temporal checkin vectors.
- Average the temporal similarities of all user's.

NANYANG TECHNOLOGICAL UNIVERSITY

Enhancing UT by Smoothing

- Smooth $c_{u,t,l}$ based on the checkin similarity between different time slots.

temporal checkin history of **u**

$c_{u,l}$	l_1	l_2
t_1	1	0
t_2	0	1

temporal checkin history of **v**

$c_{v,l}$	l_1	l_2
t_1	0	1
t_2	1	0

Smoothed temporal checkin history of **u**

$c_{u,l}$	l_1	l_2
t_1	0.8	0.3
t_2	0.2	0.7

Smoothed temporal checkin history of **v**

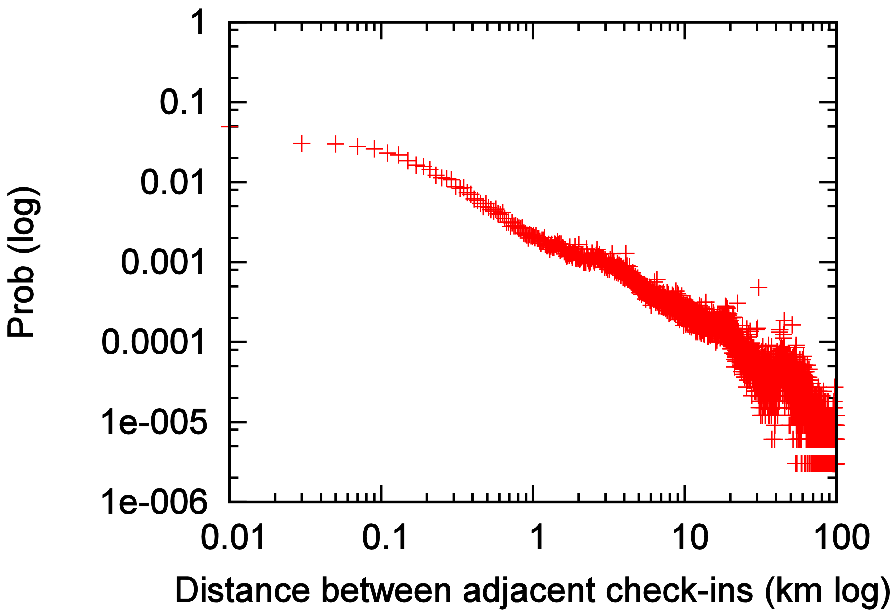
$c_{v,l}$	l_1	l_2
t_1	0.2	0.7
t_2	0.8	0.3

- Calculate the similarity.
- Consider other similar time slots when making recommendation.

NANYANG TECHNOLOGICAL UNIVERSITY

Incorporating Spatial Influence

- We assume that human tend to visit nearby POIs to their previous locations.
- We calculate the distance between two POIs of every two adjacent checkins, and plot the number as a function of distance.



- Power law distribution
- Users are more willing to visit nearby POIs
- The willingness of a user to visit dis km far away POI:

$$wi(dis) = a \cdot dis^{k}$$

Spatial Influence based Recommendation (S)

- The probability u at l_i will checkin l_j is proportional to the willingness:

$$p(l_j \mid l_i) = \frac{wi(dis(l_i, l_j))}{\sum_{l_k \in L, l_k \neq l_i} wi(dis(l_i, l_k))}$$

- Given user u and historical POIs L_u, we calculate $P(l|L_u)$ as the ranking score for each candidate POI l:

$$\hat{c}_{u,l}^{(s)} = P(l \mid L_u) \propto P(l)P(L_u \mid l) = P(l)\prod_{l' \in L_u} P(l' \mid l)$$

- Enhanced by Temporal popularity (SE)
  - $P(l)$ is the prior probability that POI l is checked in by all users
  - Change $P(l)$ to temporal checkin popularity of POI

A Unified Framework

- Given a user *u* and a candidate POI *l*, we can get:
  - Temporal recommendation score $\overline{c}_{u,t,l}^{(t)}$
  - Spatial recommendation score $\overline{c}_{u,t,l}^{(s)}$

- Linear interpolation:

$$c_{u,t,l}^{(t)} = \alpha \times \overline{c}_{u,t,l}^{(t)} + (1 - \alpha) \times \overline{c}_{u,t,l}^{(s)}$$

NANYANG TECHNOLOGICAL UNIVERSITY

Experimental Setup

- We use two real-world datasets.

	Foursquare	Gowalla
Region	Singapore	California
Time	Aug. 2010-Jul. 2011	Feb. 2009-Oct. 2010
#user	2,321	10,162
#POI	5,596	24,249
#checkin	183,330	423,362
Spasity (24 bins)	$2.65*10^{-4}$	$4.10*10^{-5}$

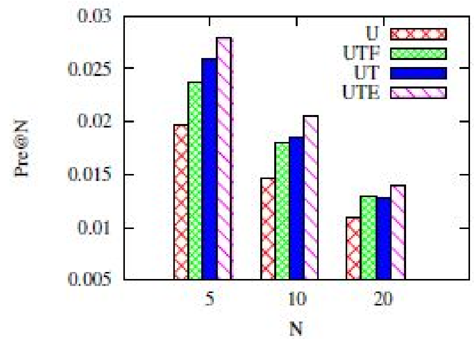
- Split historical POIs of a user into two parts:
- |training set| : |testing set| = 4 : 1
- Metrics
  - Study how well the methods can recover the hold-off POIs in the testing set given active user and active time.
  - Precision@N, Recall@N. N=5,10,20

NANYANG TECHNOLOGICAL UNIVERSITY

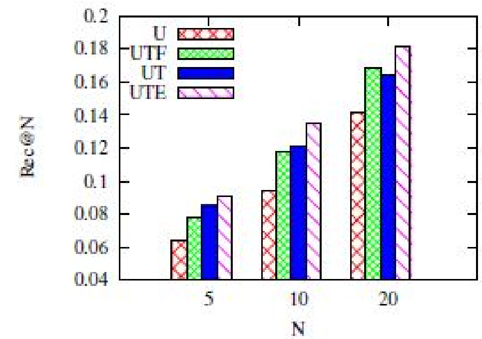
Performance of Methods Utilizing Temporal Influence

Method	Description
U	User-based CF
UTF	U with Time Function
UT	U with Temporal Preference
UTE	UT with Smoothing Enhancement

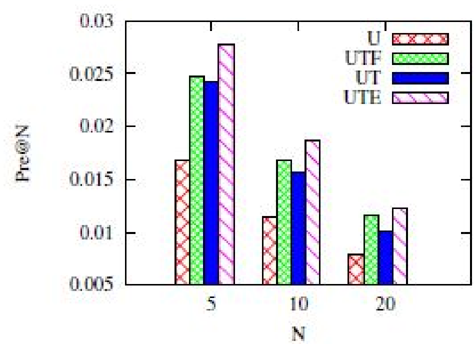
- Time plays a significant role in POI recommendation.
  - UTE outperforms U by 37%-51%
- UT performs worse on Gowalla data.
  - Much sparser
- Smoothing enhancement is effective to address the sparsity problem.



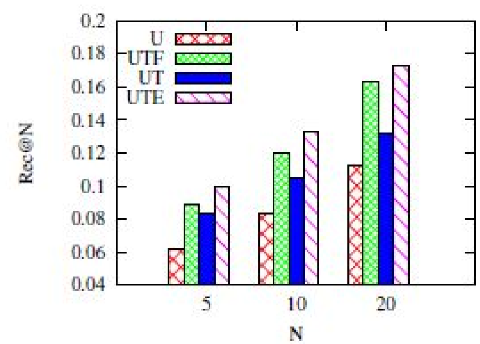
(a) Pre@N - Foursquare



(b) Rec@N - Foursquare



(c) Pre@N - Gowalla

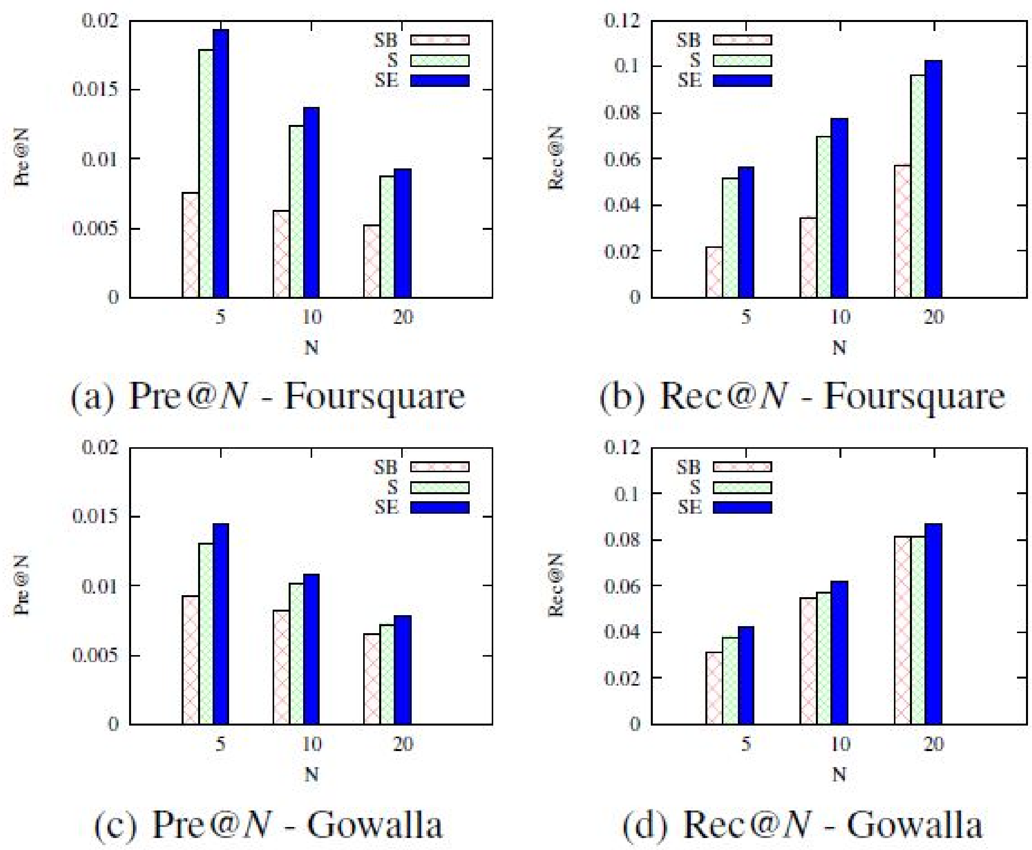


(d) Rec@N - Gowalla

Performance of Methods Utilizing Spatial Influence

Method	Description
SB	Spatial influence-based baseline
S	Spatial influent-based method, ***SIGIR 2011***
SE	S with temporal popularity enhancement

- S outperforms SB on both dataset.
- The integration of temporal popularity (SE) further improves the precision and recall.

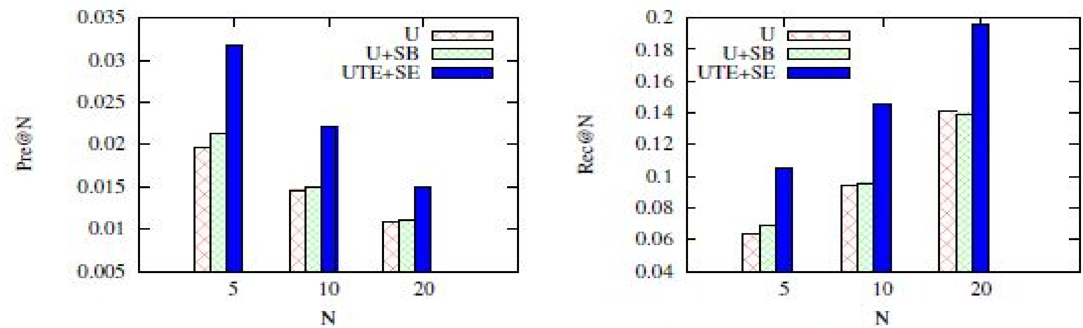


(a) Pre@N - Foursquare
(b) Rec@N - Foursquare
(c) Pre@N - Gowalla
(d) Rec@N - Gowalla

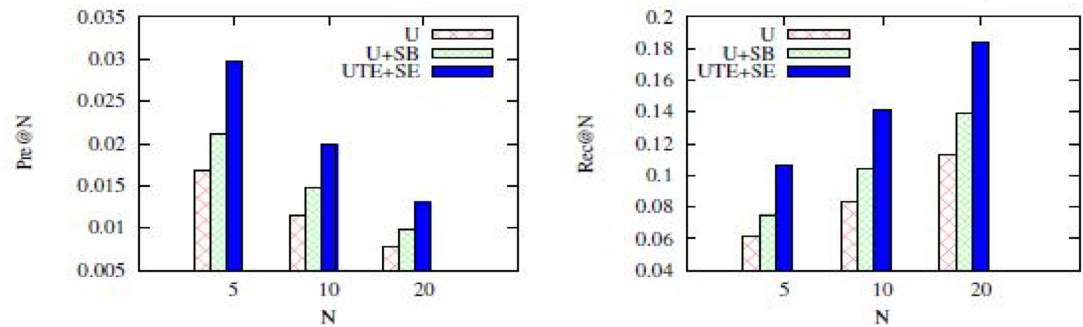
Performance of Unified Methods

Method	Description
U	User-based CF
U+SB	State-of-the-art POI recommendation, ***SIGIR 2011***
UTE+SE	Proposed unified method

- UTE enjoys the best performance
- Outperforms U+SB by 40%-50%.
- Outperforms U 60%-75%.



(a) Pre@N - Foursquare (b) Rec@N - Foursquare



(c) Pre@N - Gowalla (d) Rec@N - Gowalla

Conclusion

- User's checkin behaviors are influenced by temporal and spatial influences.
- We define a new problem: time-aware POI recommendation problem.
- We propose a solution to make use of the temporal and spatial influences.
- The experimental results show that the proposed methods beat all baselines, and improve the POI recommendation performance by 40% over the state-of-the-art method.

NANYANG TECHNOLOGICAL UNIVERSITY

Publications

- *Q. Yuan, G. Cong, A. Sun, C. Lin, NM. Thalmann*, **Category Hierarchy Maintenance: A Data-Driven Approach,** *ACM SIGIR 2012*
- *Q. Yuan, G. Cong, NM.Thalmann,* **Enhancing naive bayes with various smoothing methods for short text classification,** *WWW 2012, Poster track.*
- *X. Cao, G. Cong, B. Cui, C. S. Jensen, Q. Yuan,* **Approaches to Exploring Category Information for Question Retrieval in Community Question-Answer Archives,** *ACM Transactions on Information Systems (TOIS)*
- *Z. Ma, A. Sun, Q. Yuan, G. Cong,* **Topic-Driven Reader Comments Summarization**, *ACM CIKM 2012*

THANKS!

Q&A?