Using Location-Based Social Media for Ranking Individual Familiarity with Places

A Case Study with Foursquare Check-in Data

Wangshu Wang

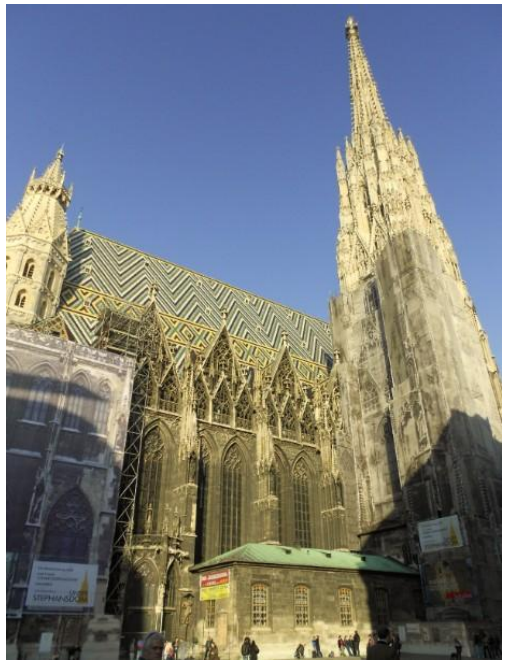
Outline

- **Introduction and research objective**
- **Methodology**
  - Identifying individually meaningful places
  - Ranking individual familiarity with places
- **Conclusions and future work**

1. Introduction and research objective

1.1 Introduction

- **Meaningful place**
  - A place that is associated with certain activities and meanings
- **Individual familiarity with a place**
  - How familiar is a place to an individual
  - Can be inferred as visiting frequency, the extensity and intensity of the experiences
- Location-based services adapted to individual priori spatial knowledge



1.2 Research objective

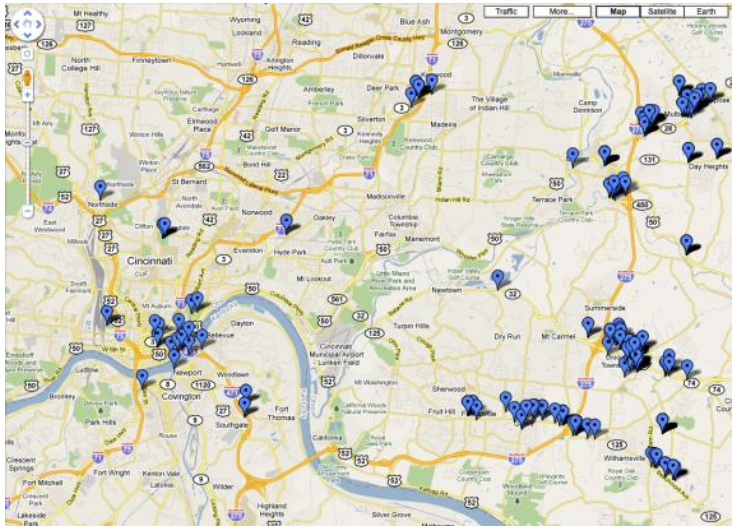
- **This research aims at ranking individual familiarity with places using Foursquare check-in data.**
  - Identifying individually meaningful places
  - Ranking individual familiarity with places

2. Identification of individually meaningful places:

comparison of existing clustering algorithms

2.1 Methodology

- **Clustering user's check-ins to find out potential personal meaningful places**
- **Comparison of existing clustering algorithms**
  - SLINK
  - K-means
  - DBSCAN
  - EM algorithm for Gaussian mixture model



2.2 Evaluation

- **Study design:**
  - 12 participants
  - Users' check-in histories
  - Their provided meaningful places lists , each place is a cluster of check-ins
- **Evaluation metrics:**
  - Precision, recall, tolerance factor and F1-score

2.3 Results

Algorithm	Precision	Recall	Tolerance Factor	F1-score	Precision + Tolerance Factor
SLINK	0.516	**0.647**	0.281	**0.574**	0.797
K-means	**0.672**	0.441	0.134	0.533	0.806
DBSCAN	0.631	0.520	**0.286**	0.570	**0.917**
EM	0.537	0.353	0.119	0.426	0.656

- The algorithms' performances vary on different data sets, i.e., different distribution of the data points
- **In general, DBSCAN balanced the best among the four algorithms.**

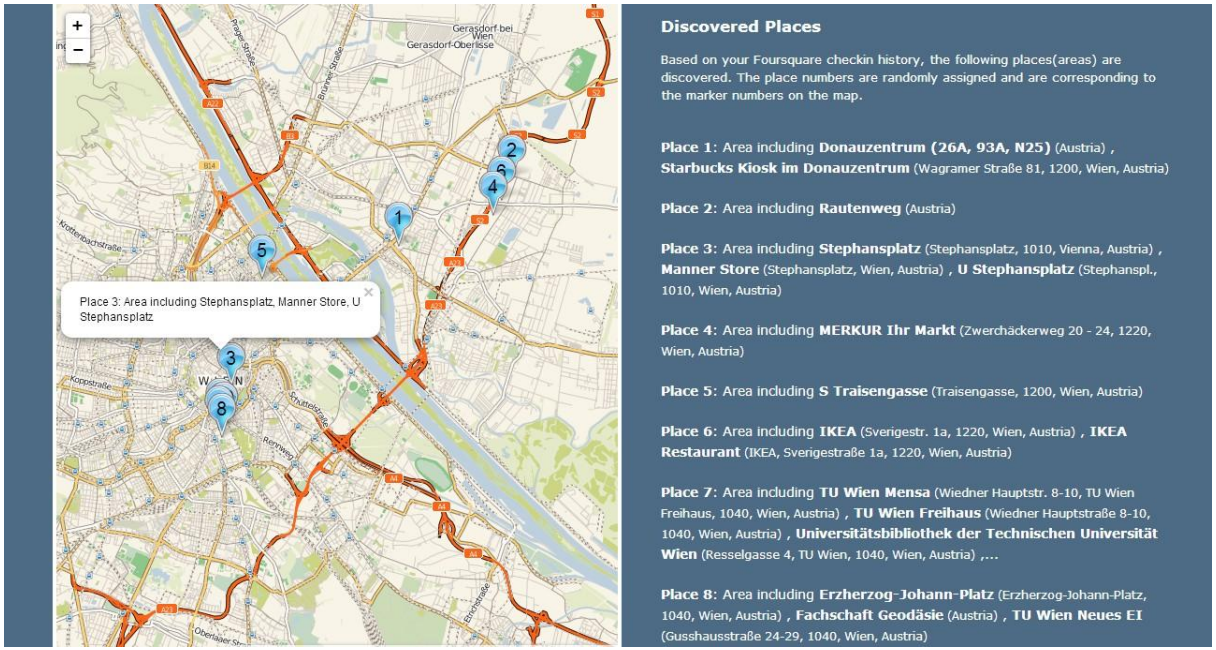
3. Ranking individual familiarity with places

3.1 Methodology

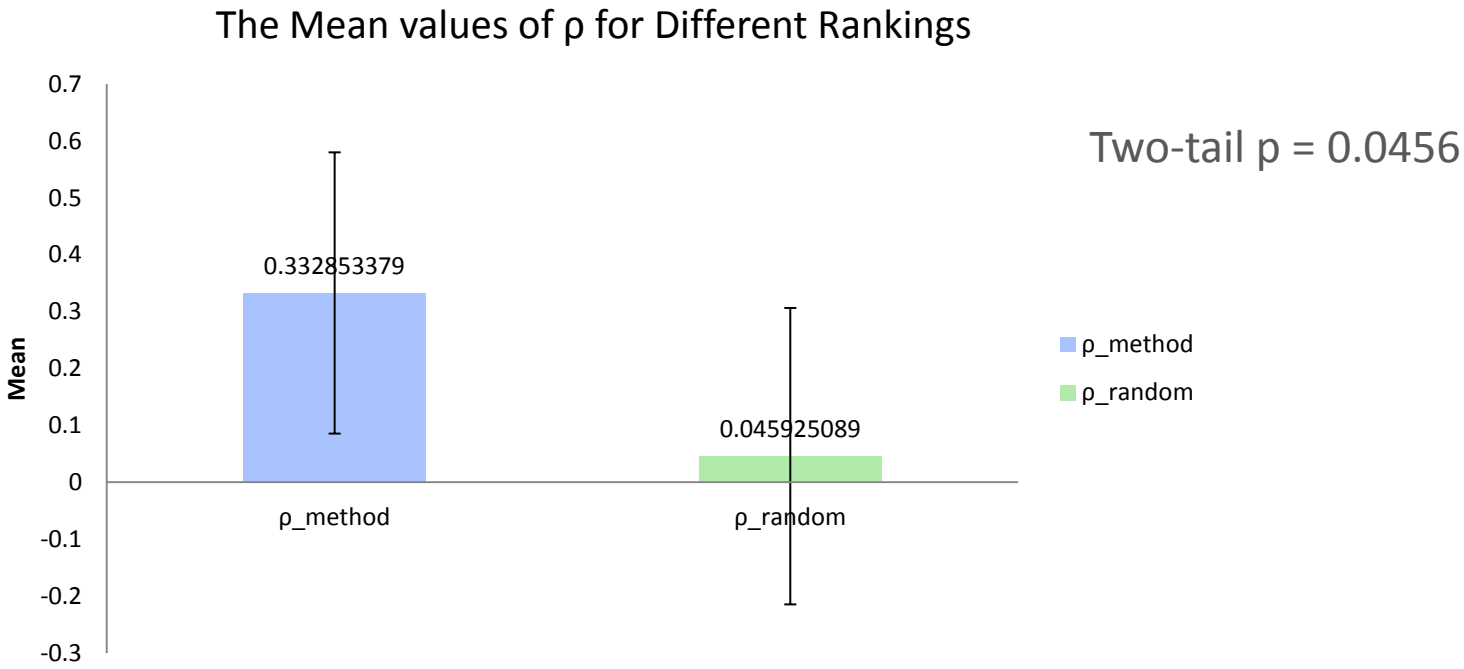
- Influencing factors of the familiarity with a place: visiting frequency, the extensity and intensity of the experiences
- **Frequency**: Foursquare check-ins
- **Extensity**: impractical to estimate
- **Intensity** : weight each check-in by user's tagging activities ("shout", "photos" and "like")
  - Each check-in has an initial weighting of 1;
  - "shout", "photos" and "like" inside a check-in item each weights 1;
  - The weighting of a check-in is at least 1 and can sum up to at most 4.
- Ranking the discovered places according to their weightings

3.2 Evaluation

- **Study design:**
  - 23 participants
  - User's ranking of the discovered places
  - Comparison with random rankings
- **Evaluation metrics**
  - Spearman's rank correlation coefficient (ρ)
  - Paired t-test



3.3 Results



- **Able to rank individual familiarity with places**
- **Positive association, but not strong**
  - Simple weighting scheme: differentiated weightings for tagging activities, different intensities of a same tagging activity
  - Limitation of Foursquare API: combination of different lcation-based social media

4. Conclusion and future work

- **Using location-based social media is able to rank individual familiarity with places.**
- **Refining the weighting scheme**
  - A deeper insight on the influencing factors of human familiarity with places
  - Natural Language Processing on the text descriptions
  - Training set and validation set
- **Combination of different location-based social media**
- **A categorization of individual importance instead of a ranking**
- **…**

Thank you very much for your attention!