User Check-In History Based Venue Recommendation

Boya Yu Jiamin Xuan Yuxiang Zhang New York University, CUSP

Contribution

We present an application of collaborative filtering for the venue recommendation in New York City. In the first part, we implement traditional memory-based collaborative filtering method to recommend venues when we train on user check-in history. Next, we combine with user relationship or clustering labels as weighting function that surpasses traditional memory based collaborative filtering method.

Data

- **Foursquare-Twitter Check-In Dataset**

We get data indirectly from Foursquare check-in records shared on Twitter. The data were a sample of record from Feb. 2014 to Feb. 2015, parsed using PySpark from 113 GB geotagged tweets around NYC. Next we used the Foursquare Venue Search API to scrape venue information denoted by latitude and longitude. After preprocessing, we got 838 venues in which 4,365 users have checked-in 54,017 times in total.

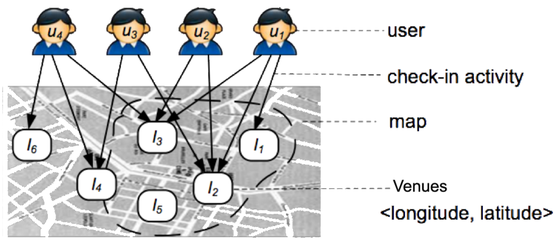
- **Venue Similarity**

For all the venues, we use Cosine Similarity and Jaccard Distance Similarity to transform the sparse matrix of user-venue check-in history into venue similarity scores.

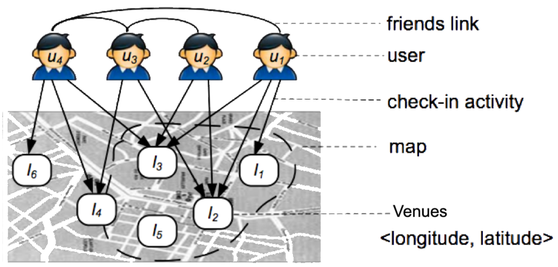
- **User Relationship**

From a social media user's standpoint, following certain users may indicate a tendency of referring to their online activities. We scraped all the Twitter IDs that each of our users follows and discarded those IDs which are not in our set of 4,365 users.

Memory-Based Collaborative Filtering



Trust-Based Collaborative Filtering



Spectral Clustering

) can be seen as a trainable feature extractor coupled with a learning model[2]. Parameters of the system include th

Results

Add text here

Acknowledgments

Add text here