



# Homework 1: Where to Go Now?

## Time-aware? Type-aware? or Geo-distance Location Recommender

H.P.  
NCKU EE

# Announcement

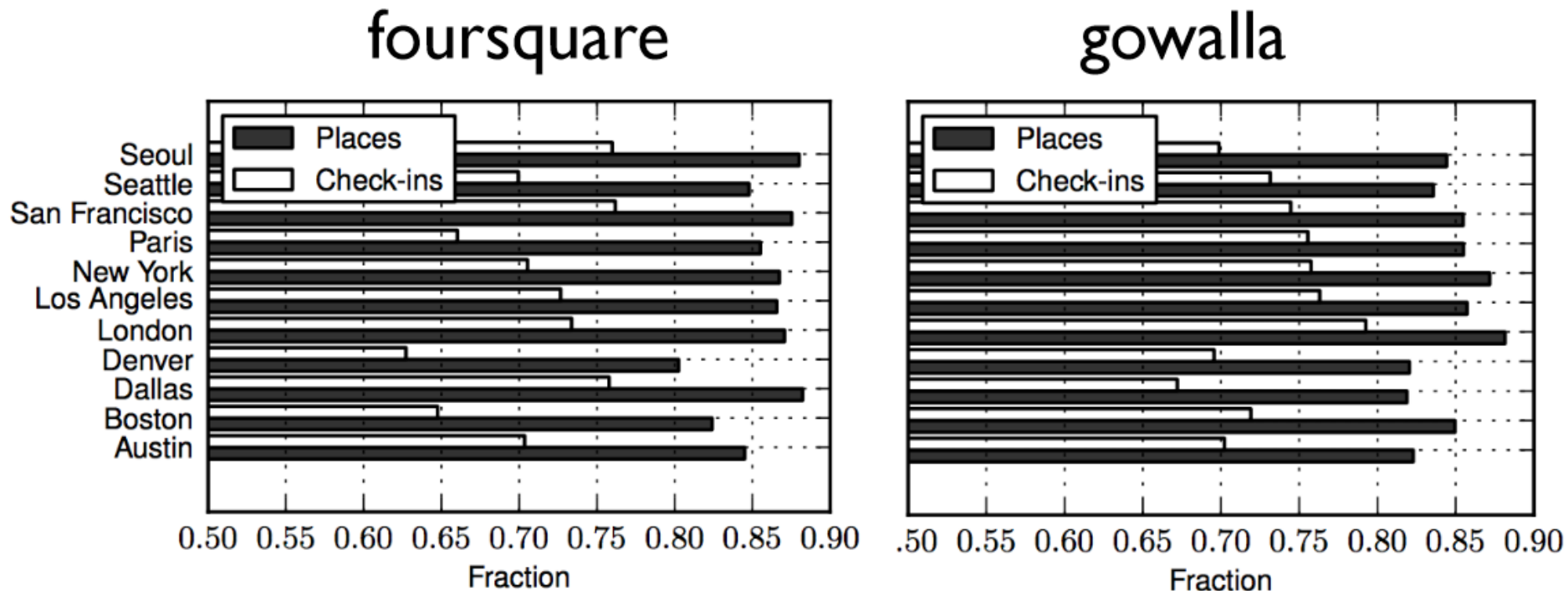
- 11/7 I have a business trip so.....

# Location Recommendation vs. Location Prediction

- Location Recommendation
  - Recommend **NEW** locations (**never visited before**)
- Location Prediction
  - Predict the next **existing** locations (**had ever visited**)
- General considered factors
  - Current location info
  - Current time
  - User history/preference
  - Social interaction

Route Planning can be viewed as the **successive** applications of location recommendation.

# Check-in at New Venues



- People seek to discover new locations
  - 80%-90% of visited places are new
  - 60%-80% of check-ins occur at new places

# Homework1: Location Recommendation/ ~~Prediction~~ for POI-Sequence data

- **Goal:** Given users' historical and destination checkins, recommend personalized locations for users:
- **Motivation**
  - For user
    - Recommend the locations the user did not visit before
  - For service provider
    - Location based services or commercials would profit by a location prediction engine.
    - Location package for finding target users
- **Point of Interest (POI)** :A certain location/attraction/place
- **Route**: The visiting location sequence of a certain day.

# Challenges

- Difficult cases in real data
  - **Problem 1:** Unobserved nature:
    - Frequency does not work. (New POI-user)
    - data sparsity
  - **Problem 2:** Complex user preference
    - geographical coordinates
    - category information (semantic)
    - Time-sensitive behaviors
  - **Problem 3:** Different role in a city
    - Visitors or non-visitors

# Task

- Given:
  - Users' check-in records (route-based) in New York City with
    - time information (hour)
  - A list of candidate locations
  - The category information of all POIs
  - The position (geographical coordinates) of all POIs
- You need to
  - For each unknown location, guess a ranked location list
    - The list is ranked by visiting probabilities of locations

# checkins\_missing.txt

(row: one-day trajectory )

- **Uid:hour1, location\_id1, hour2, location\_id2, hour3, location\_id3, ...**
- 9448:12,3fd66200f964a520c6e61ee3,14,4d8774c140a7a35d4d4648be,18,47f39422f964a5209b4e1fe3,19,4393dd78f964a520782b1fe3,19,3fd66200f964a52003e71ee3
- 9448:12,4e78ec28483bc8fe840a40d9,18,?,20,4b089510f964a520da0e23e3,22,4a96f1fcf964a520a82720e3
- 9448:12,4cc5efac91413704e185cc55,13,?,13,4b968e08f964a52047d434e3,17,49edce76f964a520f9671fe3,18,?,23,4adfb640f964a520e57c21e3
- 9448:20,4fae5867e4b091b462a96176,21,4f54e133e4b036244ce58fc7,21,3fd66200f964a520bbe61ee3
- 9448:3,41e46880f964a520d41e1fe3
- 9448:11,4db84a97f7b15ca52ce70e7f,12,4f8ca5e0e4b0cbdd4aee48cd,14,4d30b3de2748b60c4d50dea0,19,431e2d80f964a52079271fe3
- 9448:21,43a52546f964a520532c1fe3,22,4aa3bf7ff964a520114420e3,23,46797032f964a520f7471fe3
- 9448:12,49c046f1f964a5205b551fe3,13,4c3c8d3d7f49c9b658036ce3
- 9448:12,4df01a9ce4cdfec0ed2c2b05,12,4f80a871e4b0a6d0b654f7dc



# loc\_id\_info.txt (all locations in NYC)

Location_id	latitude	longitude	location_type	country
3fd66200f964a52000e71ee3	40.733596	-74.003139	Jazz Club	US
3fd66200f964a52000e81ee3	40.758102	-73.975734	Gym	US
3fd66200f964a52000ea1ee3	40.732456	-74.003755	Indian Restaurant	US
3fd66200f964a52000ec1ee3	42.345907	-71.087001	Indian Restaurant	US
3fd66200f964a52000ee1ee3	39.933178	-75.159262	Sandwich Place	US
3fd66200f964a52000f11ee3	40.652766	-74.003092	Bowling Alley	US
3fd66200f964a52001e51ee3	40.726961	-73.980039	Dive Bar	US

# candidate\_100\_places.txt

## Location\_id

- 43a52546f964a520532c1fe3
- 4ace6c89f964a52078d020e3
- 42911d00f964a520f5231fe3
- 49b7ed6df964a52030531fe3
- 42829c80f964a5206a221fe3
- .
- .
- .
- 4af5a46af964a520b5fa21e3
- 43695300f964a5208c291fe3
- 3fd66200f964a520d7f11ee3
- 3fd66200f964a520def11ee3
- 4297b480f964a52062241fe3
- 4840fe6bf964a52030501fe3
- 49c046f1f964a5205b551fe3
- 40f1d480f964a5206a0a1fe3
- 49d60947f964a520a85c1fe3
- 4b992b04f964a520726635e3
- 4a43bcb7f964a520bba61fe3
- 508af9ccf13627e878935855
- .
- .

# Effectiveness

- Hit counter
  - For each user, we retrieve the hit position  
(第幾次猜中)
  - Average all users

# Report format (1)

- No formal format, you just can find a way to introduce your (but not limited to)
  - methodologies & thinking
  - or experimental process/framework
  - or experimental results

# Report format (2) Questions

- What are the physical meanings of your proposed methods? Why do you want to do that?
- If same approach results in much different performance in other cities, what are the possible reasons you think?
- Do you have other idea for solving this problem? (but no time to try)

# Testing

- We will provide an opportunity to let your try your answer before 10/29 23:59pm
- Please submit your files to our server (the link will be provided later)
  - Two testing results

# Policy

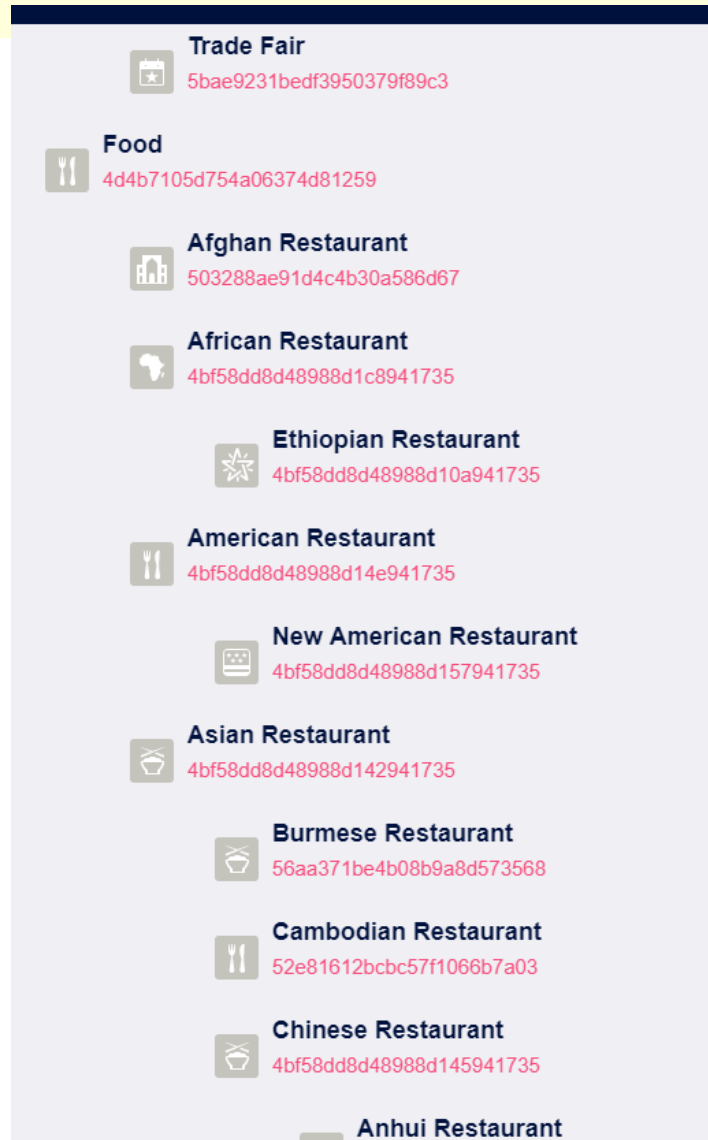
- 1-3 members for each team
- Testing deadline: 10/29 23:59pm
- Final Deadline: 11/12 23:59pm(submit one result)
  - Penalty: each day late -5
  - Submit your file(with your student ids) to our server (the link will be provided later)
- Submit your source code of your methods ([source\\_code.rar](#))
- Submit your report ([report.docx](#))
- Submit your answer([result.txt](#))
  - One user one row
    - Format(no space for each line):
      - userID1:rank1\_location\_id,rank2\_location\_id,rank3\_location\_id...,rank100\_location\_id
      - userID2:rank1\_location\_id,rank2\_location\_id,rank3\_location\_id ...,rank100\_location\_id
      - userID3:rank1\_location\_id,rank2\_location\_id,rank3\_location\_id ...,rank100\_location\_id
      - .
      - .
      - .
    - Total number of rows: 100
    - Ranked by visiting probabilities of locations (from high to low)
- Compress all files into a rar or zip file
- 15~20% of your final grade
- Grade: Hit count: 50%, report: 50%
  - Normalized score
- Encourage propose your own idea
  - Reporting some failing methods is ok (if it spend much time)
- Discussion but no plagiarism

# Hints

- Capture personal behaviors from some perspectives (some are mutually influenced)
  - Collaborative preference
  - Distance preference
  - Temporal preference(hour)
  - Category preference (category hierarchy)
    - <https://developer.foursquare.com/docs/resources/categories>
  - Victors or non-visitors
- Bonus: try your methods on validation data (other location) and see the effectiveness
- Recommend unseen locations!



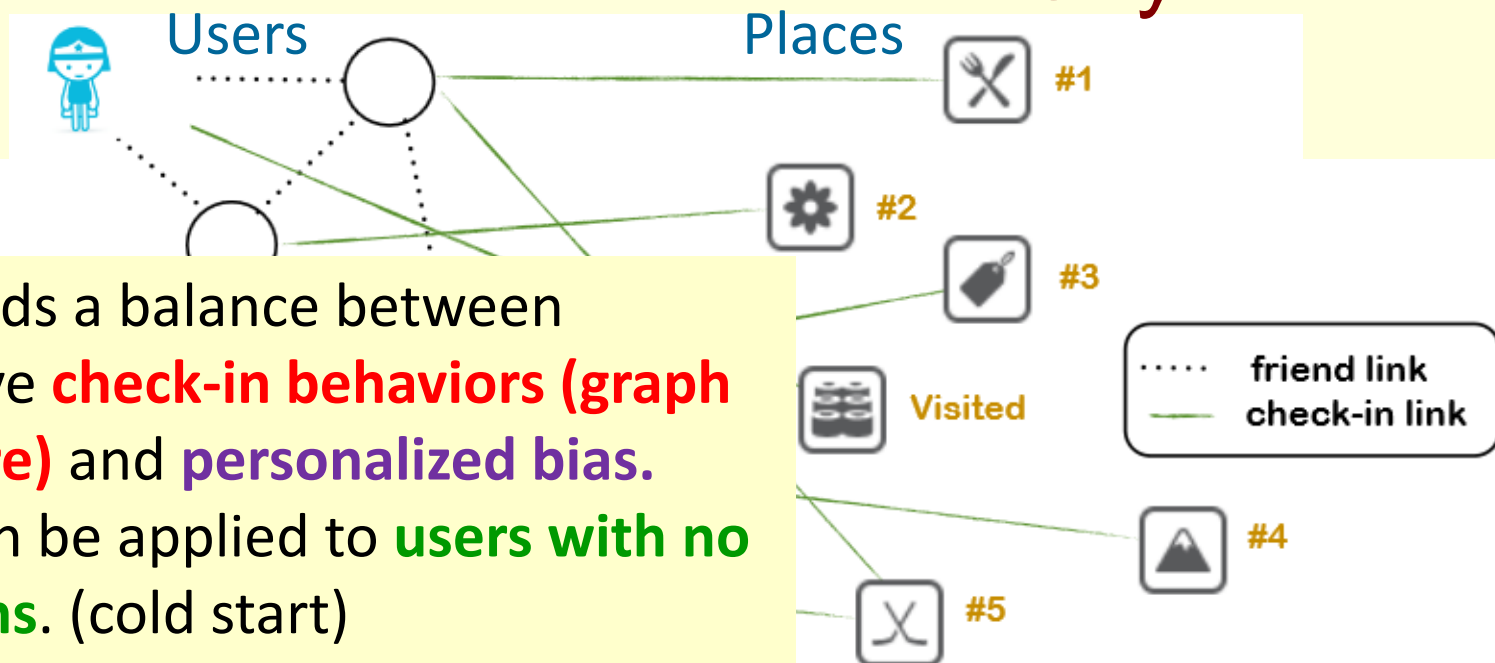
# category hierarchy



# General Location Rec Strategies

- **Popularity**: rank locations using # of check-ins
- **Content Filtering**: using venue type preference
- **Social Filtering**: rank locations using # of check-ins by friends
- **Home Distance**: geo-distance from home
- **K-NN User Similarity** (CF)
- **Place Network** (Item Similarity)
- **Matrix Factorization**

# A Random Walk Around The City



1. RWR finds a balance between collective **check-in behaviors (graph structure)** and **personalized bias**.
2. RWR can be applied to **users with no check-ins**. (cold start)

Method	APR	Precision@10	Recall@10	Foursquare			Gowalla	
				City	popularity	rwr	popularity	rwr
Random	0.500	0.000	0.003	Austin	0.235	0.222	0.175	0.144
Popular	0.228	0.026	0.089	Boston	0.204	0.196	0.313	0.258
Activity	0.228	0.025	0.087	Dallas	0.247	0.232	0.248	0.197
Home	0.383	0.008	0.026	Denver	0.233	0.200	0.285	0.236
Social	0.392	0.015	0.049	London	0.264	0.262	0.311	0.244
kNN	0.443	0.003	0.011	Los Angeles	0.212	0.196	0.281	0.242
PlaceNet	0.337	0.026	0.077	New York	0.192	0.185	0.280	0.242
MF	0.281	0.004	0.014	Paris	0.265	0.256	0.271	0.204
RW	0.217	0.028	0.094	San Francisco	0.208	0.200	0.220	0.183
				Seattle	0.238	0.218	0.264	0.226
				Seoul	0.210	0.226	0.410	0.381
				Average	0.228	0.217	0.278	0.232

# Tutorials on Location Recommendation

- Jiliang Tang, Jie Tang, Huan Liu.  
Recommendation in Social Media. @ ACM KDD 2014.
- Huiji Gao, Jiliang Tang, and Huan Liu.  
Personalized Location Recommendation on Location-based Social Networks. @ ACM RecSys 2014.
- Yu Zheng.  
Location-based Social Networks. @ WWW 2012.

## Survey Papers:

A. Noulas, S. Scellato, N. Lathia, C. Mascolo: A Random Walk around the City: New Venue Recommendation in Location-Based Social Networks. SocialCom 2012.

Jie Bao, Yu Zheng, David Wilkie, and Mohamed F. Mokbel. A Survey on Recommendations in Location-based Social Networks. Geoinformatica 2014.

Yonghong Yu and Xingguo Chen. A Survey of Point-of-Interest Recommendation in Location-Based Social Networks. AAAI Trajectory-based Behavior Analytics Workshop 2015.