UAM
Universidad Autónoma de Madrid

Escuela Politécnica Superior

Discovering Related Users in Location-Based Social Networks

Sergio Torrijos, **Alejandro Bellogín**, Pablo Sánchez
Universidad Autónoma de Madrid
Spain

UMAP, July 2020

Motivation

- Neighbour-based recommender systems
  - Easy to understand and implement
  - Allow straightforward explanations

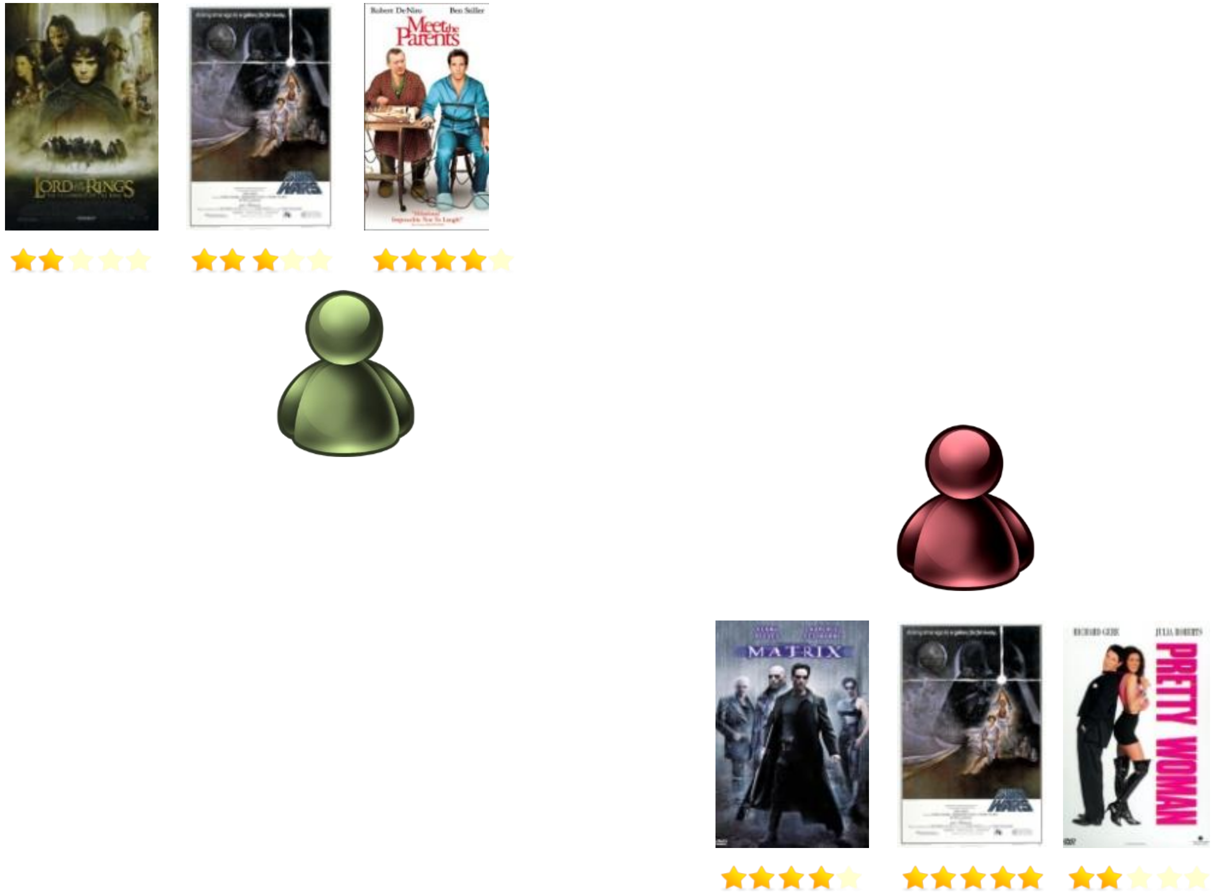
- In this work: focus on LBSN (users check-in in POIs)
  - Is it possible to adapt similarity metrics to this domain?
  - In particular: how can we integrate **sequentiality** and **geographical** information into neighbour-based recommendation?

Context

- Recommender systems
  - Users interact (rate, purchase, click) with items

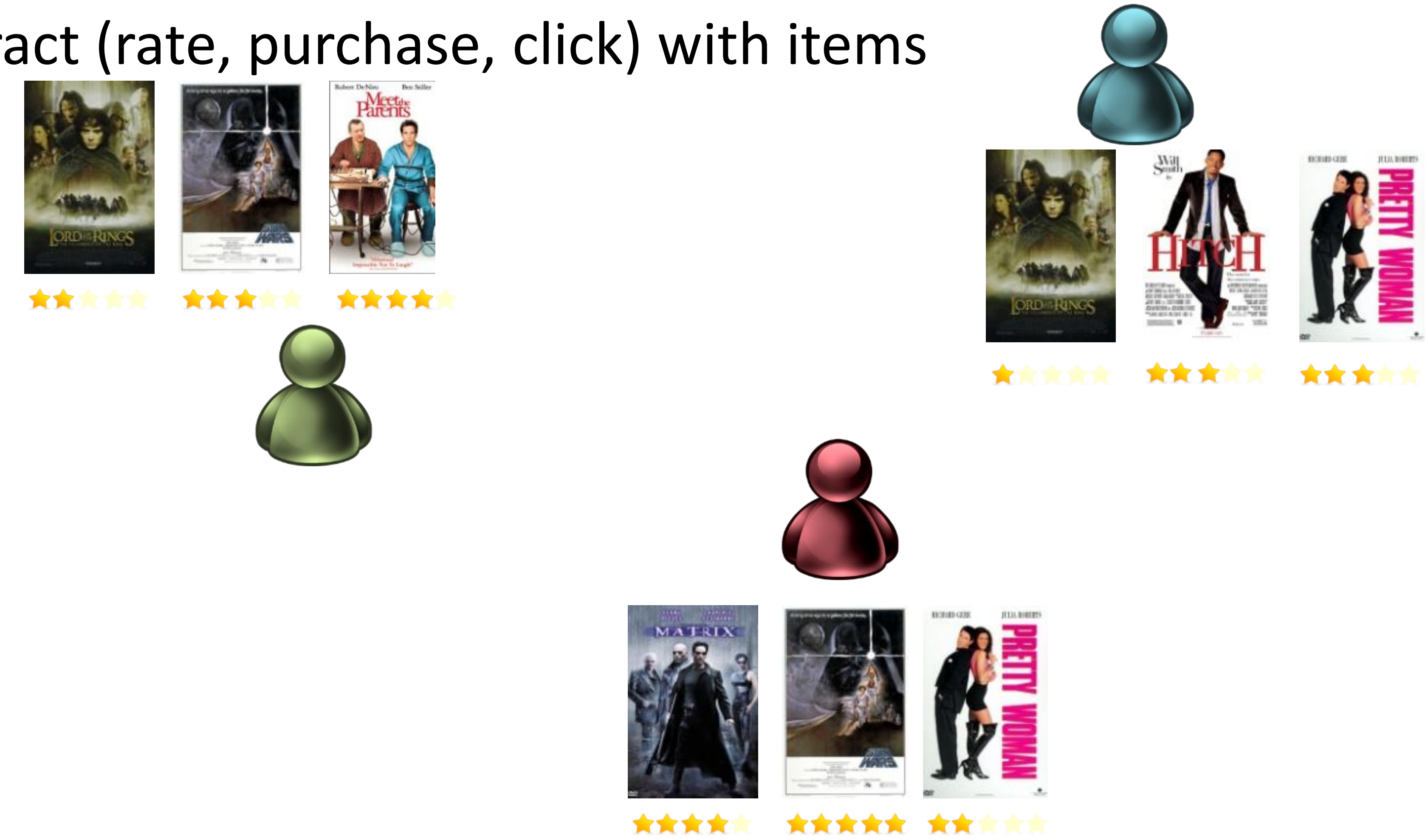
Context

- Recommender systems
  - Users interact (rate, purchase, click) with items



Context

- Recommender systems
  - Users interact (rate, purchase, click) with items



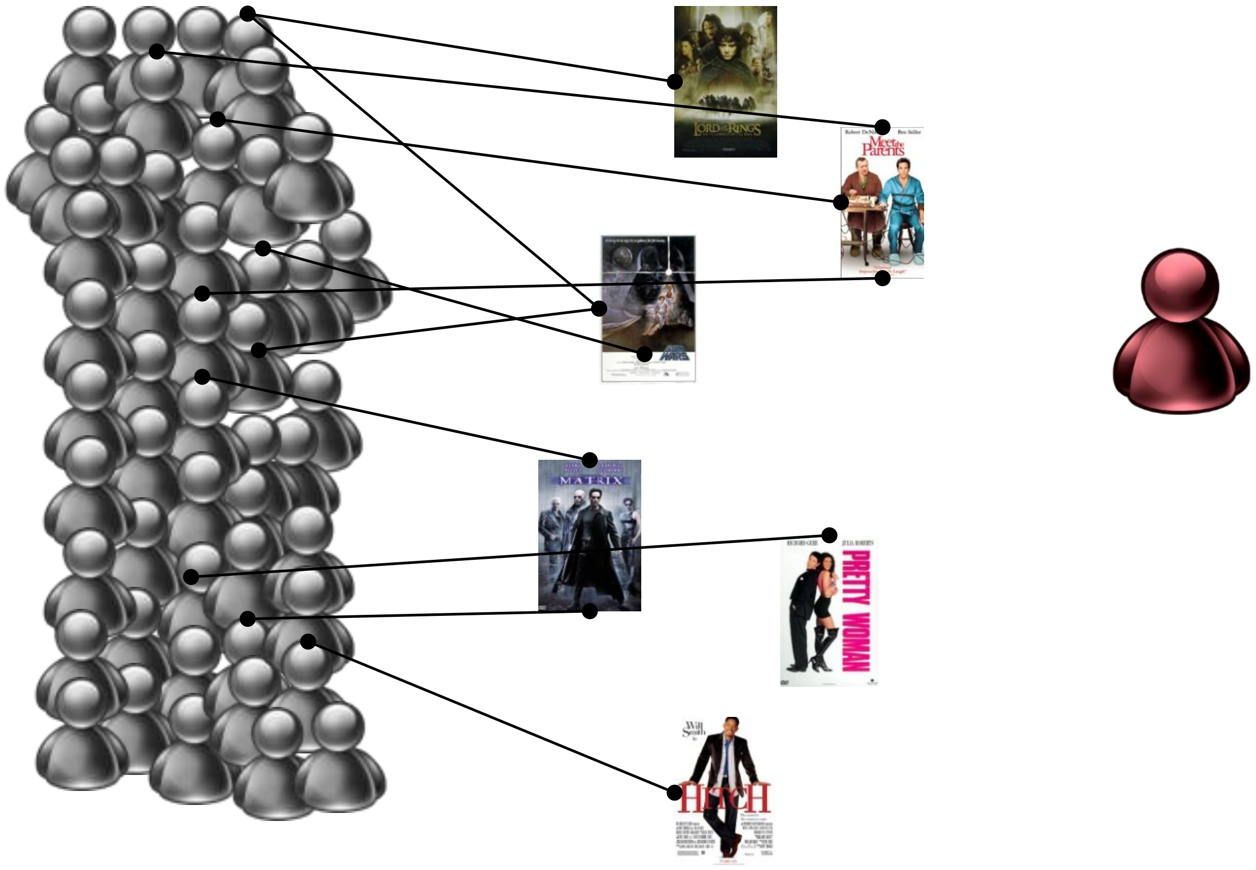
Context

- Recommender systems
  - Users interact (rate, purchase, click) with items

  - Which items will the user **like**?

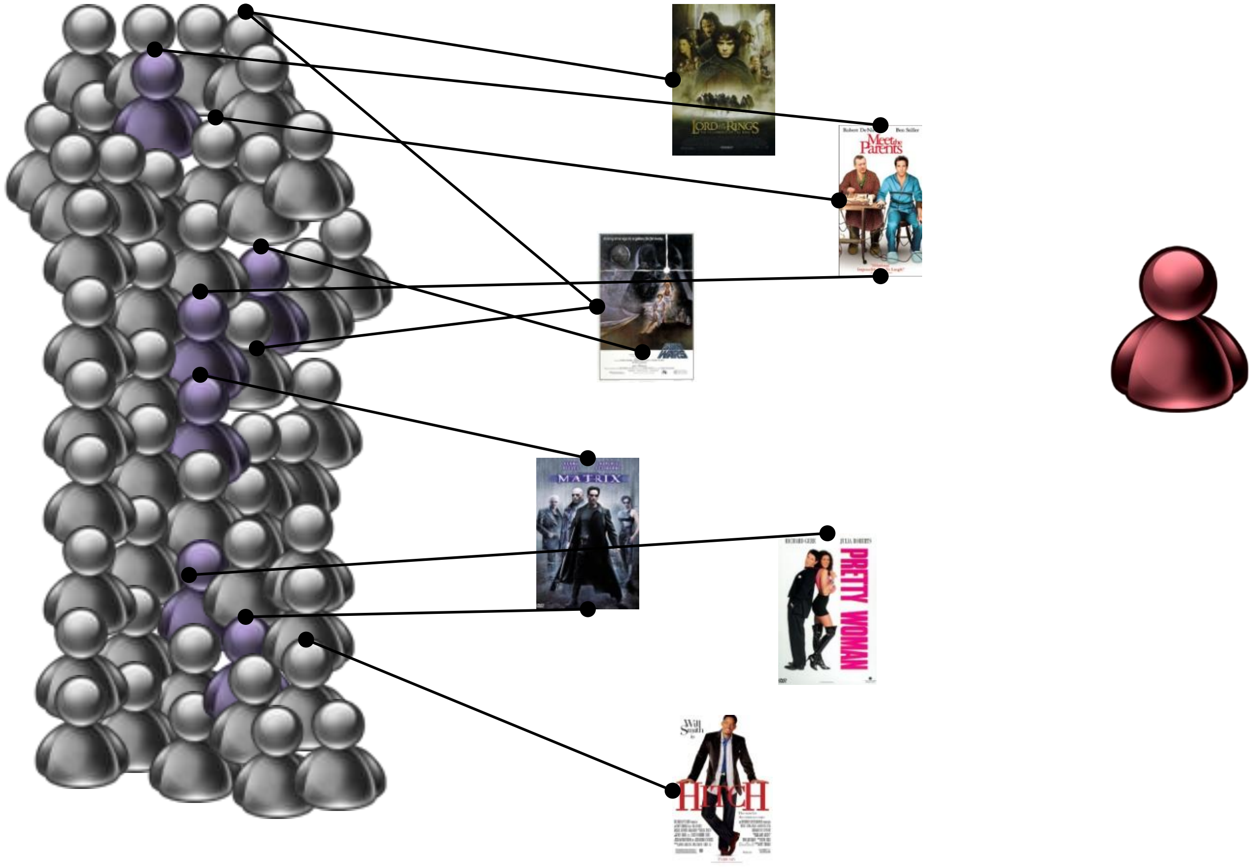
Context

- Nearest-neighbour recommendation methods
  - The item prediction is based on "similar" users

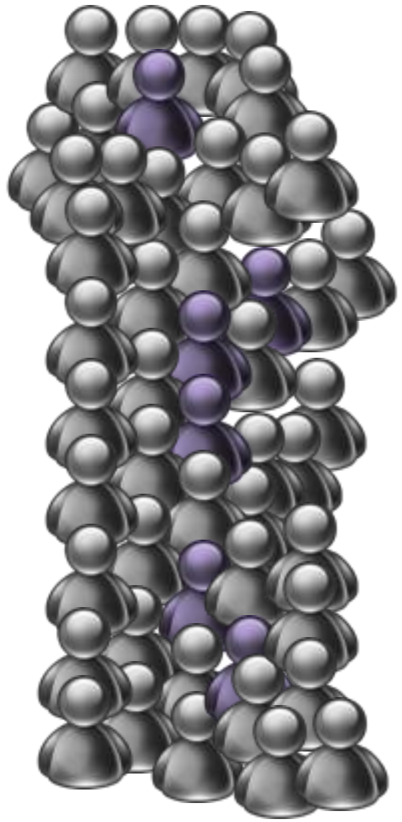


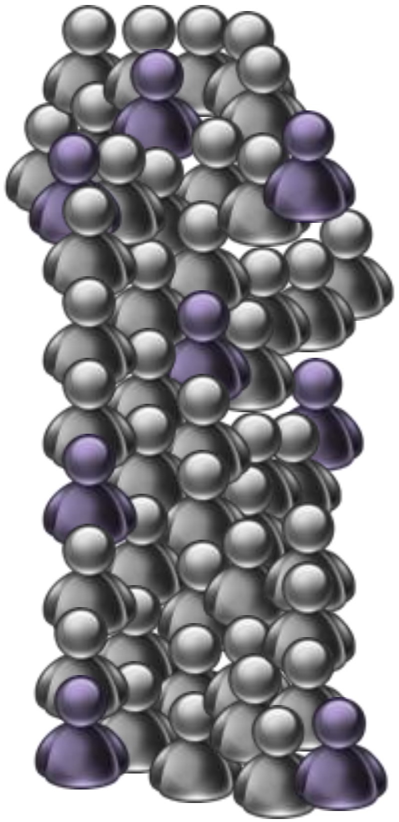
Context

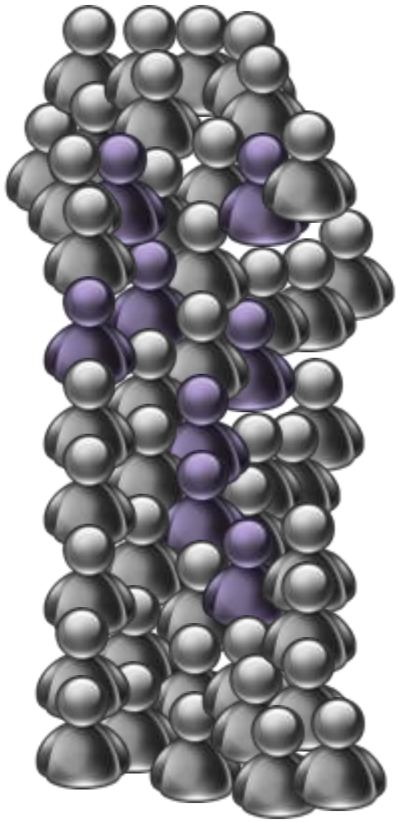
- Nearest-neighbour recommendation methods
  - The item prediction is based on “similar” users



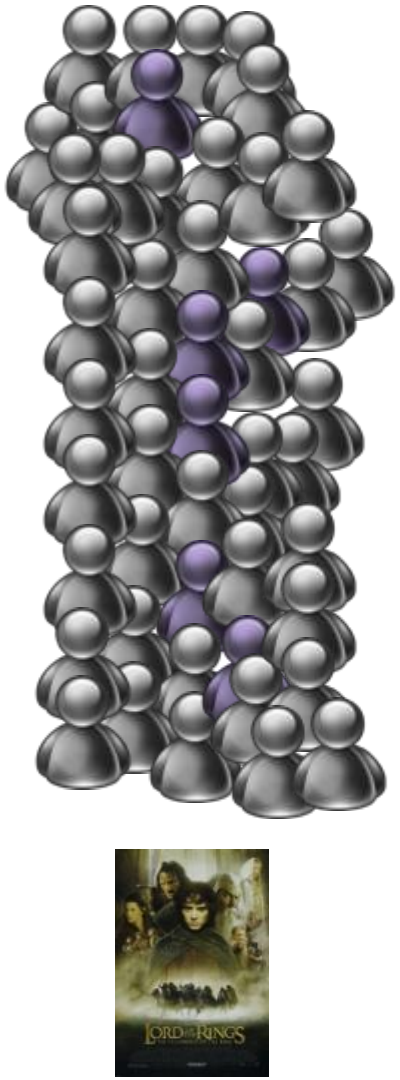
Different similarity metrics – different neighbours

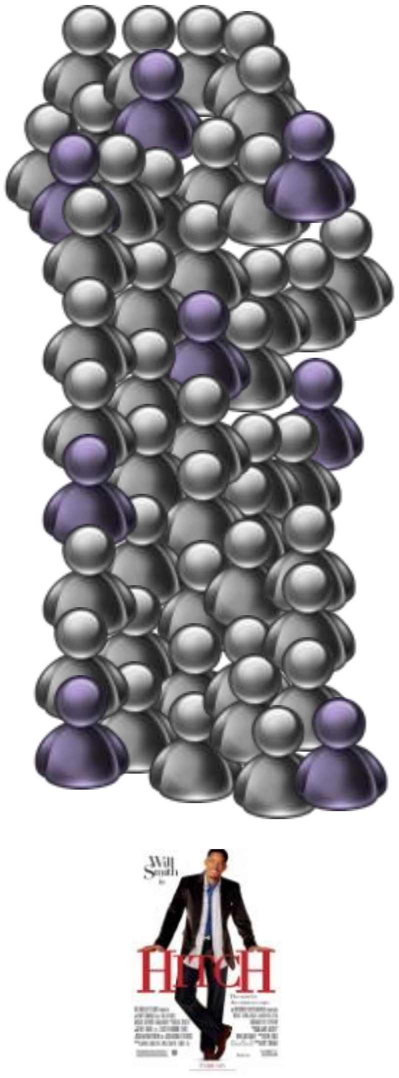


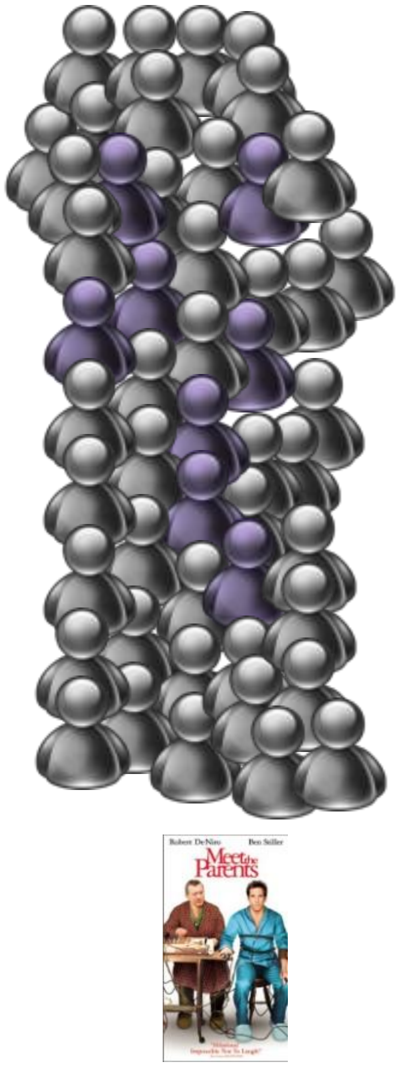




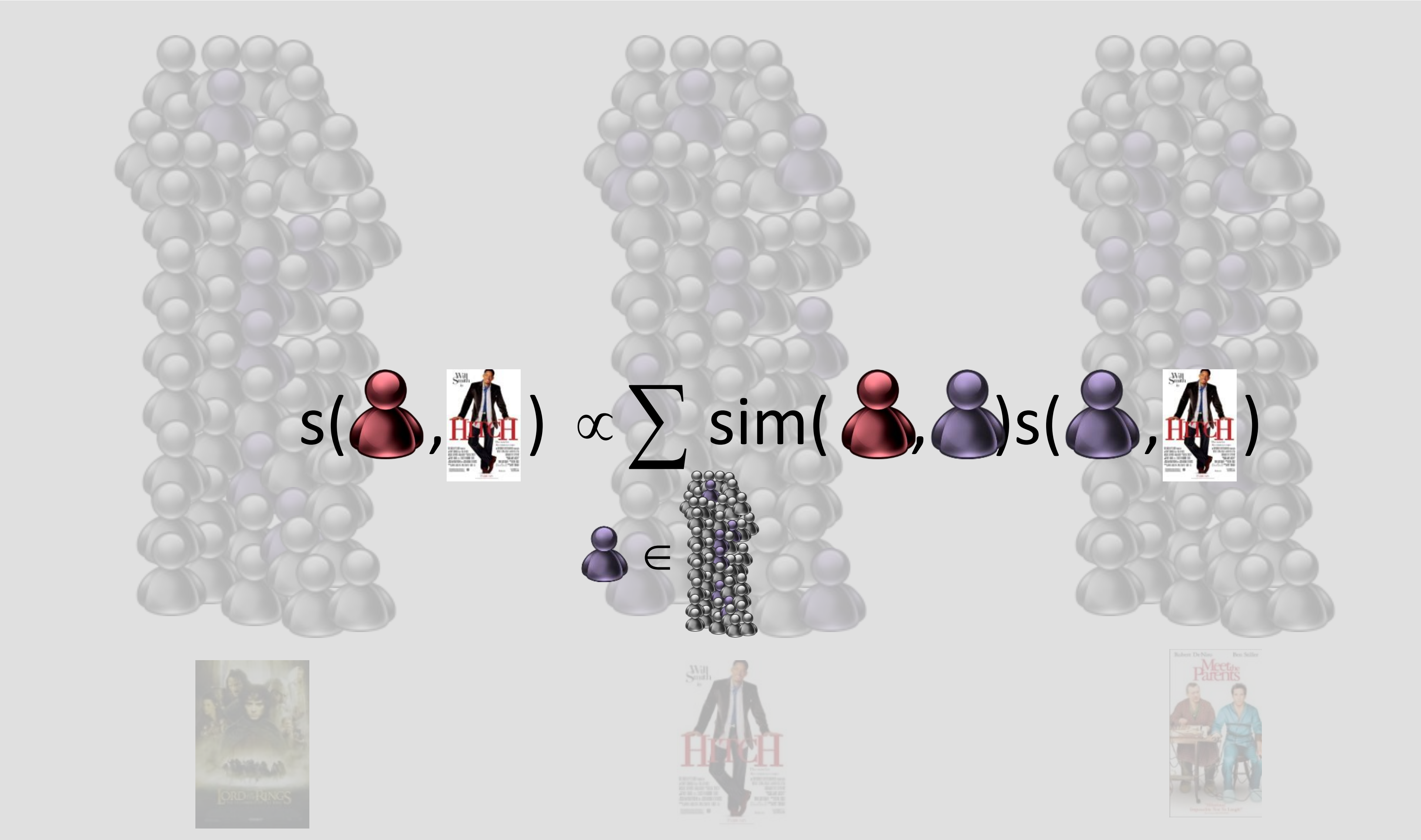
Different similarity metrics – different neighbours







Different similarity metrics – different neighbours

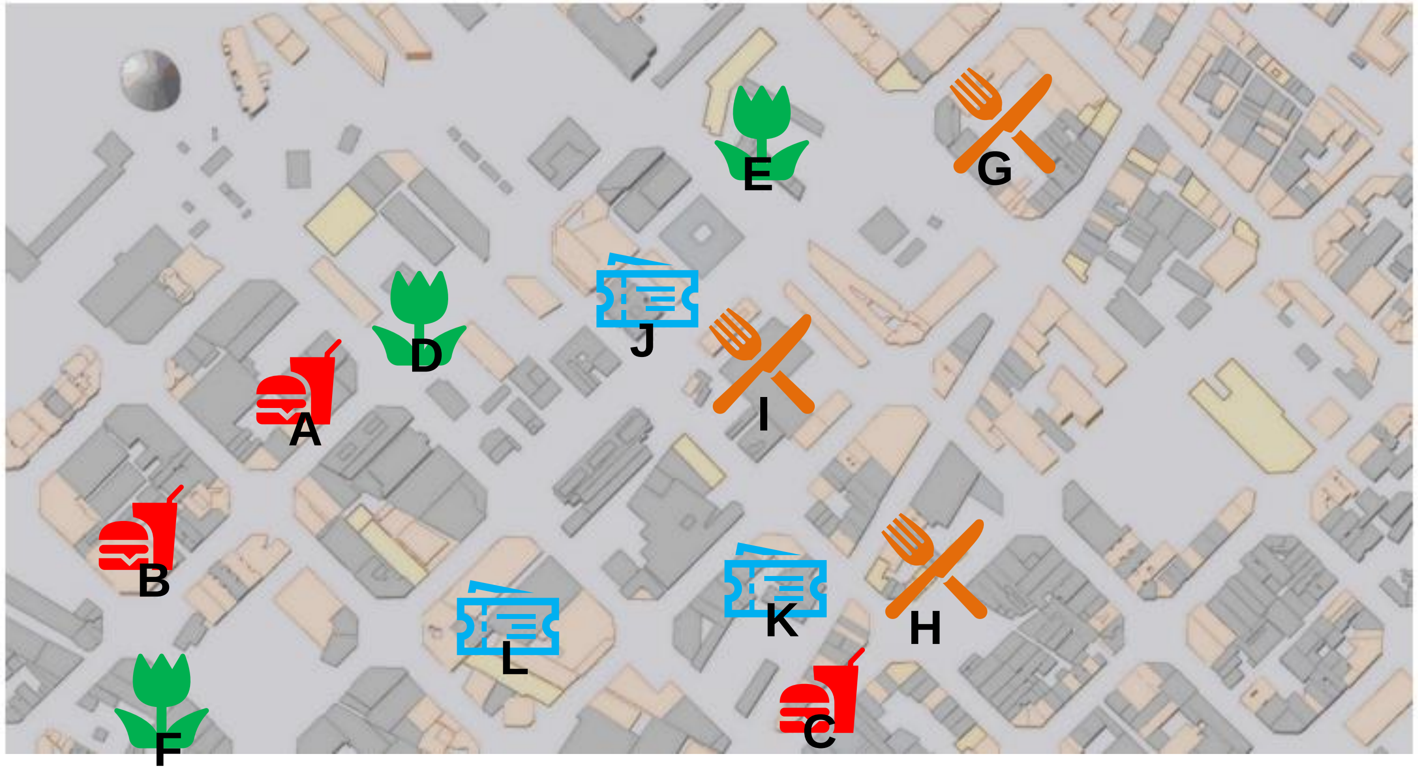


UAM Universidad Autónoma de Madrid

Escuela Politécnica Superior

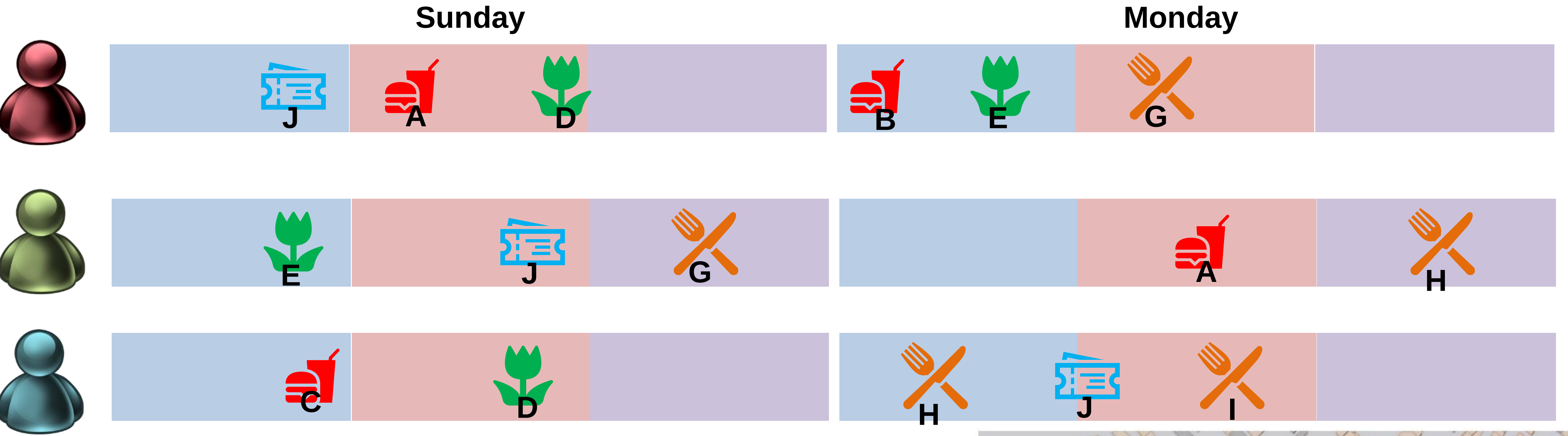
Research question

- Based on typical interactions in Location-Based Social Networks…



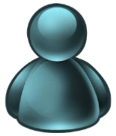
can we identify different types of users and select the most relevant ones as neighbours?

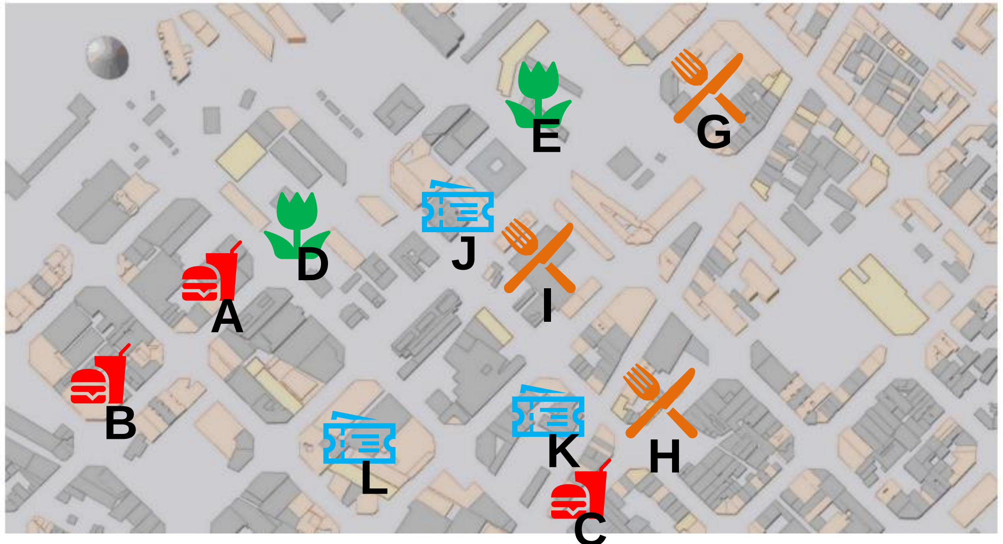
Discovering related users



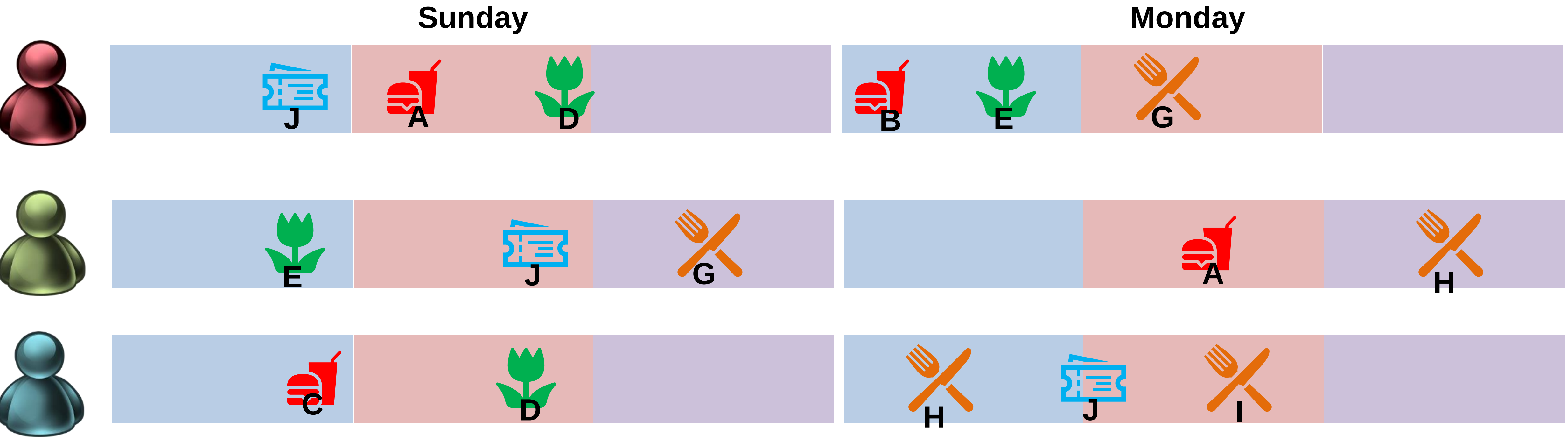
- Classical user similarities: related if users share items in common

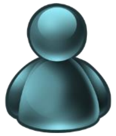
- Our approach: relatedness depends on **when** and **how near** items are





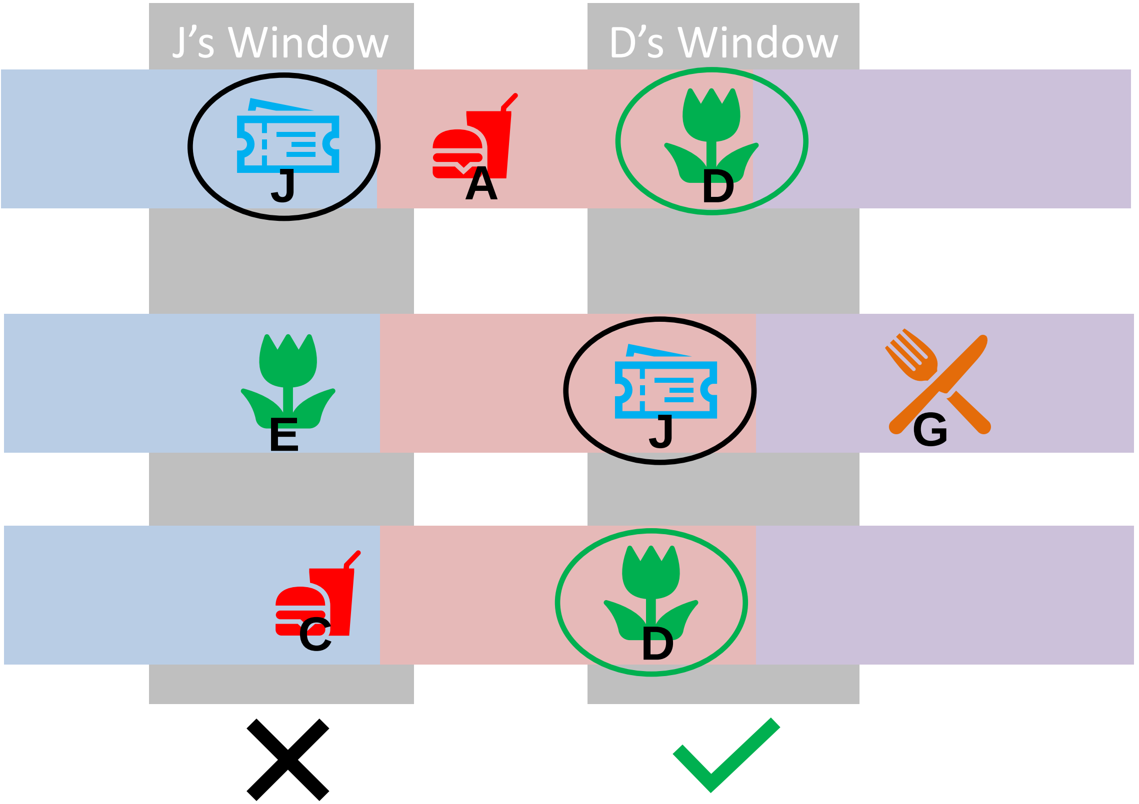
Discovering related users

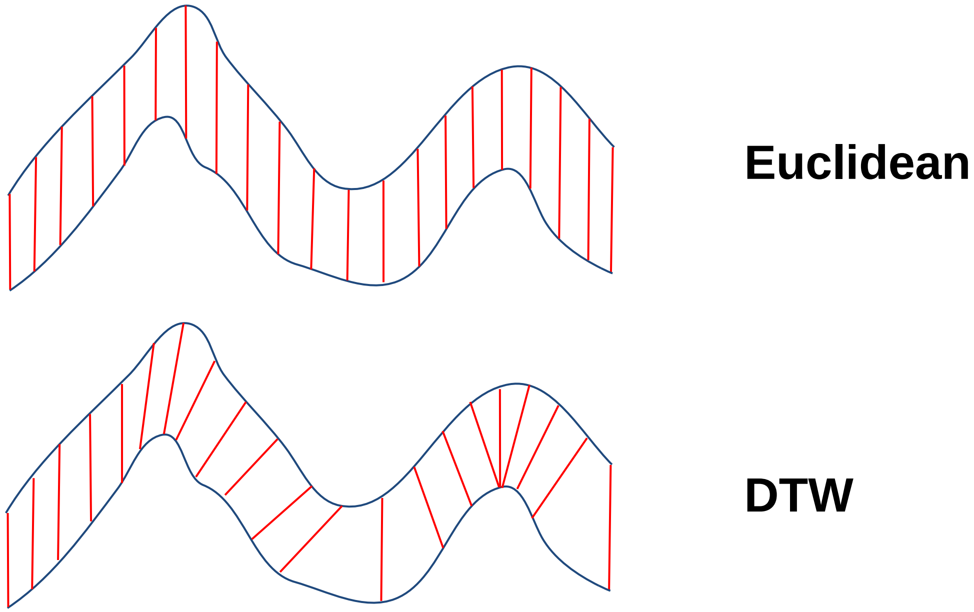


- Classical user similarities: related if users share items in common  →Captures global preferences
- Our approach: relatedness depends on **when** and **how near** items are  →Useful for contextual suggestions

Exploiting temporal and geographical information

- Exploiting check-ins within a temporal window: *ad-hoc*
  - focus on check-ins **around the same time**





- Exploiting common trajectories
  - Users are similar if their trajectories are similar
  - Trajectory similarity metrics:
    - Dynamic Time Warping (DTW)
    - Hausdorff distance

Experiments

- Foursquare data: Tokyo from global check-in dataset (33M) ~ 328K
- Temporal Split: 6 months for training, 1 month test
- Baselines
  - UB: neighbour recommender with classic user similarity
  - IB: neighbour recommender with classic item similarity
  - BPR: Bayesian Personalised Ranking using a matrix factorisation algorithm
  - IRenMF: matrix factorisation algorithm that exploits geographical influence
- Metrics
  - NDCG: accuracy of item recommendations
  - FA: feature agreement, or precision in terms of category matching (not items)
  - AD and EPC: diversity and novelty metrics

Performance comparison

- Neighbours are not competitive against MF methods in terms of accuracy

Recommender	NDCG	FA	AD	EPC
Ad-hoc	0.054	0.206	**3,314**	0.741
TS-DTW	0.036	0.203	408	0.871
TS-Haus	0.036	**0.215**	442	0.870
UB	**0.062**	0.201	1,899	**0.912**
IB	0.046	†0.264	†13,685	†0.956
BPR	0.066	0.229	28	0.872
IRenMF	†0.069	0.225	1,475	0.888

Performance comparison

- Neighbours are not competitive against MF methods in terms of accuracy
- Much better results are found for beyond-accuracy dimensions:
  - Ad-hoc is the best one for diversity (AD)
  - Similarity with Hausdorff is the best one for category accuracy (FA)

Recommender	NDCG	FA	AD	EPC
Ad-hoc	0.054	0.206	**3,314**	0.741
TS-DTW	0.036	0.203	408	0.871
TS-Haus	0.036	**0.215**	442	0.870
UB	**0.062**	0.201	1,899	**0.912**
IB	0.046	†0.264	†13,685	†0.956
BPR	0.066	0.229	28	0.872
IRenMF	†0.069	0.225	1,475	0.888

Impact on local vs tourist users

- There are different types of users in LBSNs:
  - Locals (if their check-ins span more than 21 days) vs tourists
- IRenMF is still the best approach
- But neighbour recommenders improve their performance for tourists
  - In particular, for FA

Recommender	NDCG	FA	AD	EPC
Ad-hoc	0.050	0.181	**2,829**	0.735
TS-DTW	0.032	0.179	358	**0.869**
TS-Haus	0.032	**0.191**	405	0.869
UB	**0.057**	0.177	1,440	0.704
IB	0.042	†0.250	†11,929	†0.872
BPR	0.059	0.204	26	0.608
IRenMF	†0.063	0.200	1,199	0.647

Local users

Recommender	NDCG	FA	AD	EPC
Ad-hoc	0.071	†**0.271**	**1,217**	0.770
TS-DTW	0.069	0.246	320	**0.847**
TS-Haus	0.068	0.252	333	0.839
UB	**0.095**	0.245	1,133	0.710
IB	0.064	0.262	†3,208	†0.858
BPR	0.108	0.266	18	0.566
IRenMF	†0.111	0.262	106	0.580

Tourist users

Social network analysis

- How similar are the found neighbours to explicit social connections?

Neigh. Sel.	Method	Avg.	Total	T-NDCG	T-FA
Best	Ad-hoc	2	1,008	0.071	0.271
	TS-DTW	9	4,833	0.069	0.246
	TS-Haus	9	4,833	0.068	0.252
As UB	Ad-hoc	2	1,008	0.071	0.273
	TS-DTW	2	1,074	0.056	0.271
	TS-Haus	5	2,685	0.050	0.270
	UB	4	2,140	0.095	0.245

- TS-Haus always obtains more social connections than the baseline UB
- Performance accuracy on tourist users is competitive (T-NDCG)
- Feature agreement is always better than baseline (T-FA)

Conclusions

- Two novel similarity metrics for LBSN are proposed
  - Integrating the temporal dimension and geographical information

- Competitive results in terms of beyond-accuracy metrics
  - Novelty and diversity
  - Especially positive when users are identified as tourists

- Future: explore research on mining GPS trajectories to analyse its application to check-ins from LBSNs

Thank you

Discovering Related Users in Location-Based Social Networks

Slides, code and more: http://ir.ii.uam.es/~alejandro/publications.html

Sergio Torrijos, Alejandro Bellogín, Pablo Sánchez

Universidad Autónoma de Madrid

Spain

UMAP, July 2020