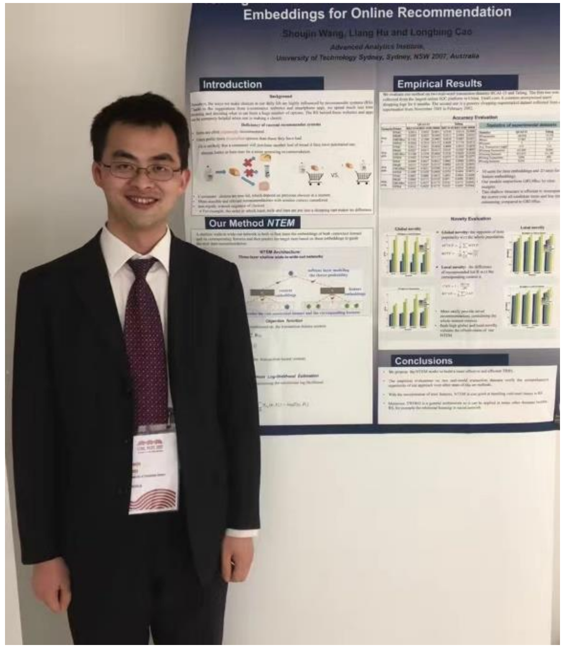
Complement, Composite and Context: The 3C-Law to Build Multidomain Recommender Systems

Authors: Dr. Liang Hu, Dr. Shoujin Wang, Dr. Qi Zhang, Dora. D. Liu,
Dr. Zhong Yuan Lai, and Prof. Longbing Cao

Dr. Shoujin Wang (shoujin.wang@mq.edu.au)

RMIT
UNIVERSITY



- Shoujin Wang has been working as a research fellow at RMIT University, Australia.
- He obtained his PhD in Data Analytics from University of Technology Sydney in 2018.
- His main research interests include data mining, machine learning and recommender systems.

https://sites.google.com/view/shoujinwanghome/home

Dr. Qi Zhang (qi.zhang-13@student.uts.edu.au)

- Qi Zhang received his first Ph.D. from the Department of Computer Science and Engineering, Beijing Institute of Technology, China in 2020.

- Currently, he is an AI scientist in DeepBlue Academy of Sciences, and a Ph.D. candidate in Analytics at University of Technology Sydney, Australia.

- His research interests include recommender systems, learning to hash, machine learning and general artificial intelligence.

Dr. Zhong Yuan Lai (abrikosoff@yahoo.com)

- Dr. Zhong Yuan Lai obtained his PhD from the University of Bonn, Germany in 2017.
- He was subsequently postdoc researcher at Fudan University before assuming his current position as researcher at the DeepBlue Academy of Sciences.

Dr. Liang Hu

- Liang Hu received dual Ph.D. degrees from Shanghai Jiao Tong University, China and University of Technology Sydney, Australia.
- His research interests include recommender systems, machine learning, data science and general intelligence.
- He has published more than 40 papers in top-rank international conferences and journals, including WWW, IJCAI, AAAI, ICDM, ICWS, TOIS, IEEE-IS, and etc.

Goal

Providing a comprehensive understanding of how to build state-of-the-art multidomain recommender systems with complex data and relations across various domains

Agenda

- Overview of multidomain recommender systems (Liang Hu, 25 mins)
- Multi-item domain recommender systems (Z. Y. Lai, 20 mins)
- Multi-user domain recommender systems (Qi Zhang, 20 mins)
- Multi-data domain recommender systems (Z. Y. Lai, 20 mins)
- Multi-spatial domain recommender systems (Qi Zhang 20 mins)
- Multi-temporal domain recommender systems (Z. Y. Lai, 20 mins)
- Multi-goal domain recommender systems (Liang Hu, 20 mins)
- Summary (Liang Hu, 5 mins)

Outline

Overview of multidomain recommender systems

- Overview of recommender systems
- Challenges of recommender systems
- A case study of multidomain recommender systems
- The 3C-law to build multidomain recommender systems

What are recommender systems

- Recommender systems (push information) are the evolution of information retrieval systems (pull information).

Information Age → Recommendation Age

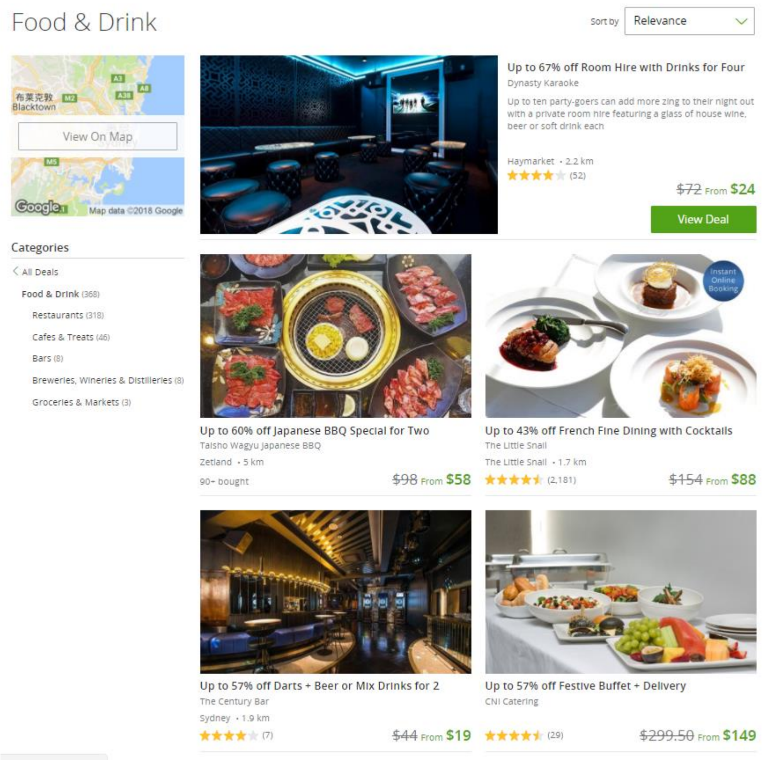
Pull mode (IRS):
Query → Matched Results → Manual Filtering

Push mode (RS):
Potential Requirement → Machine Filtering → Recommendation

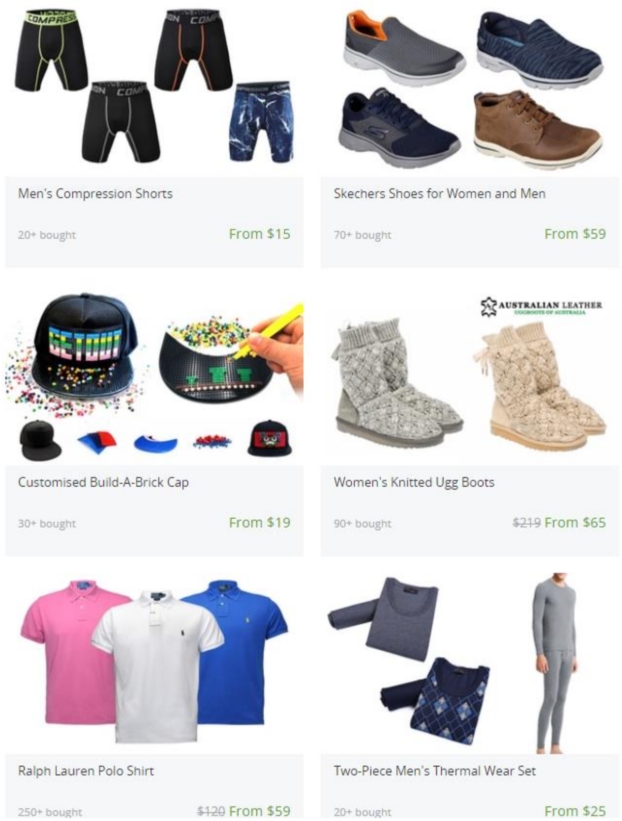
Anderson, C. (2006). *The long tail: Why the future of business is selling less of more*

Recommender systems have occupied our life

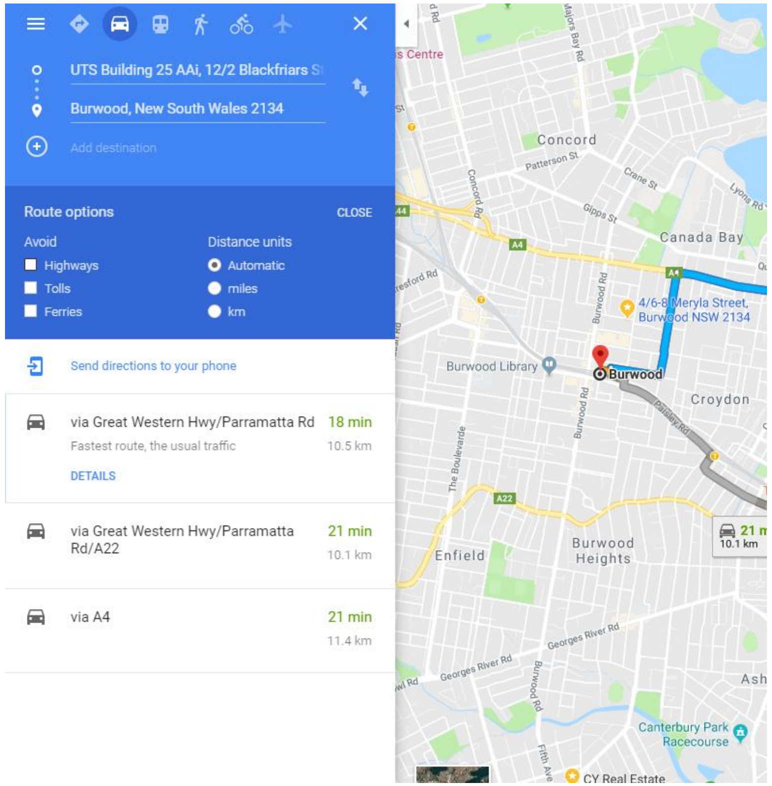
What to eat



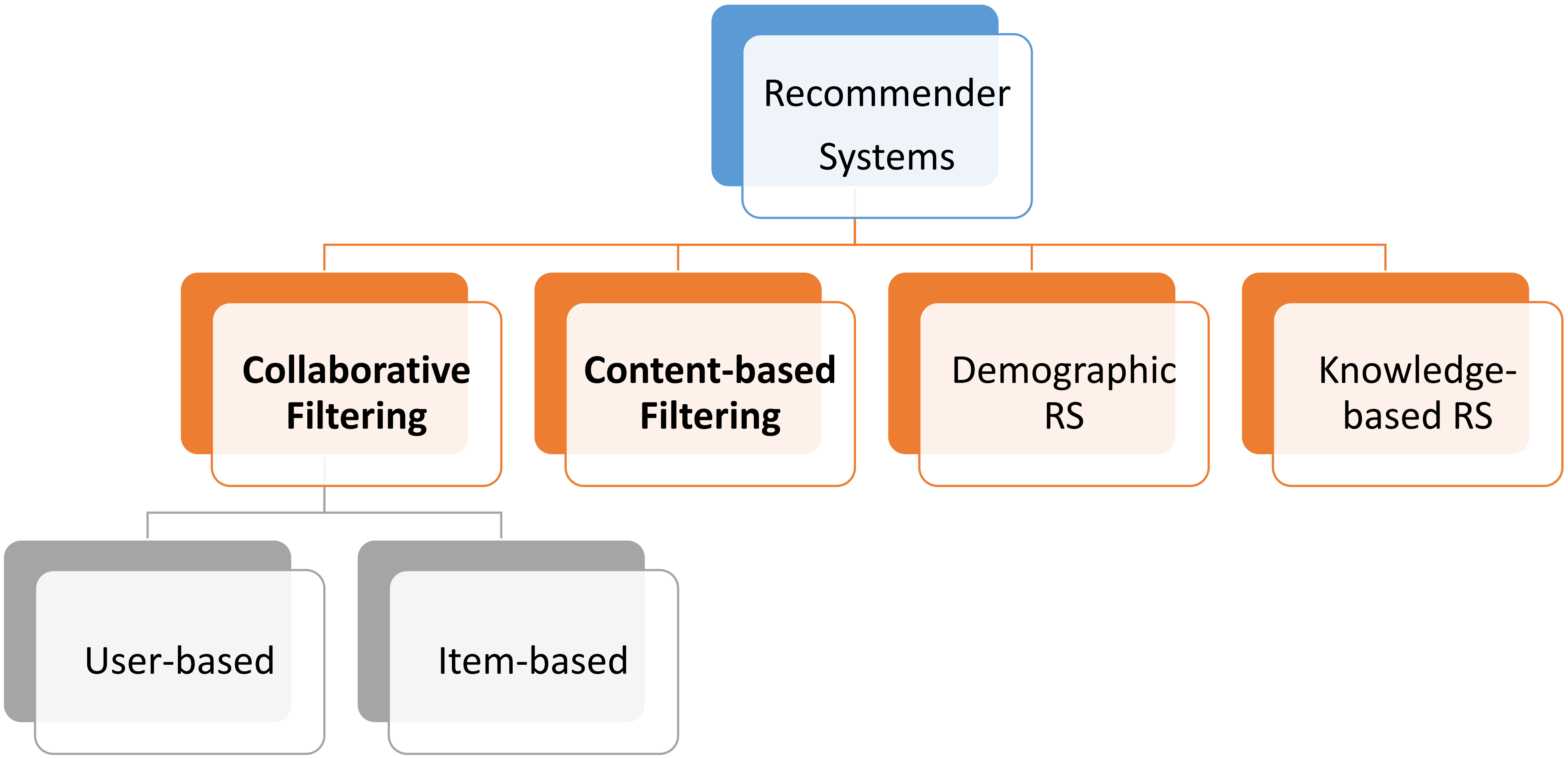
Which to dress



Where to go

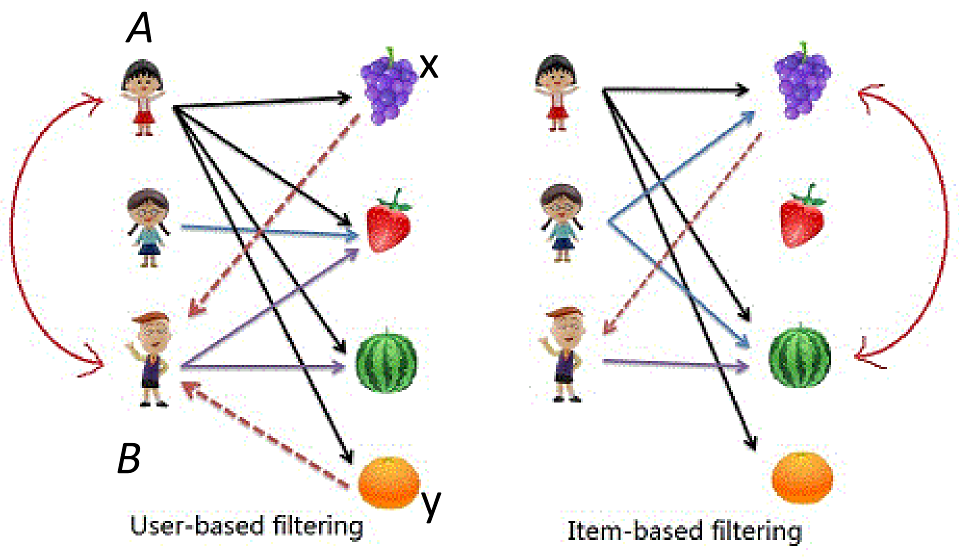


Classic recommender systems



Collaborative Filtering (CF)

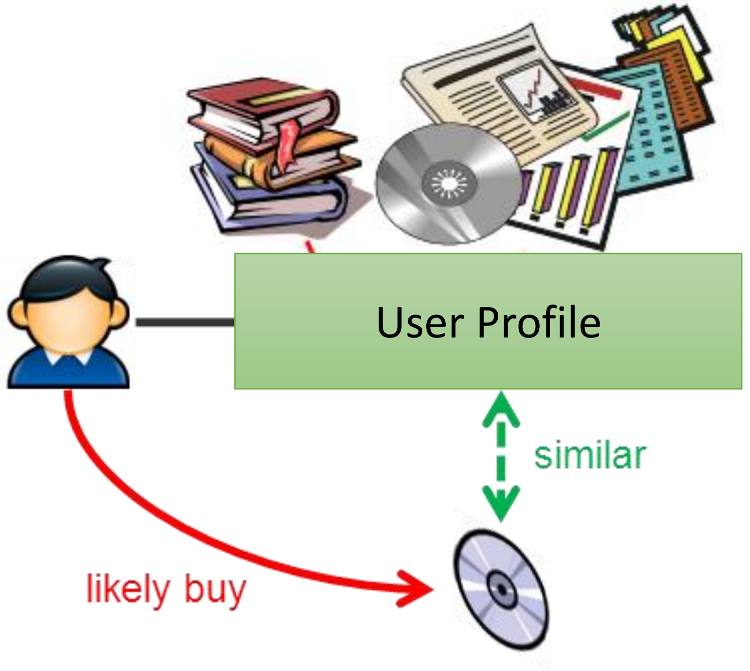
- Intuition (user-based filtering): If user ***A*** related to user ***B*** and ***A*** bought x and y, then ***B*** bought x tend to buy y.
- Famous examples(item-based filtering): Amazon.com's recommender system
- Facebook, MySpace, LinkedIn use collaborative filtering to recommend new friends, groups, and other social connections.



Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). *Item-based collaborative filtering recommendation algorithms*. Paper presented at the Proceedings of the 10th international conference on World Wide Web, Hong Kong, Hong Kong.

Content-based Filtering (CBF)

- CBF is based on the features of items
  - Attributes of items
  - Description of items
  - Text of an article
- User profile is built with the features of historical items
- Recommend items according to user profile



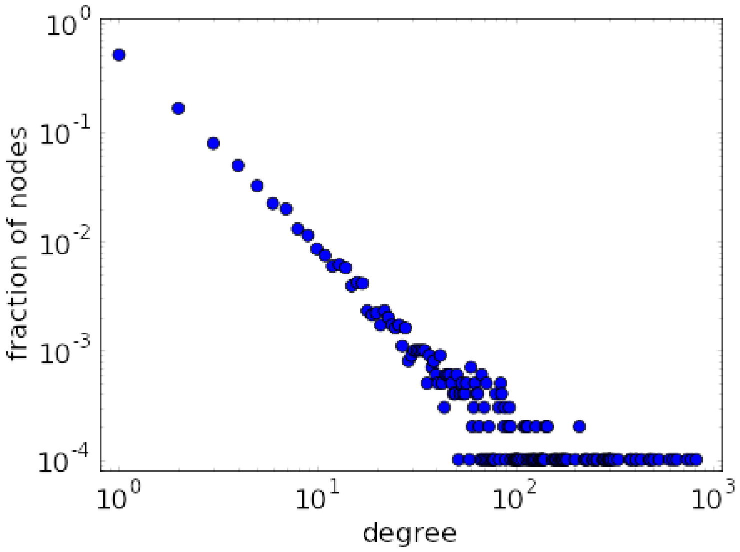
Outline

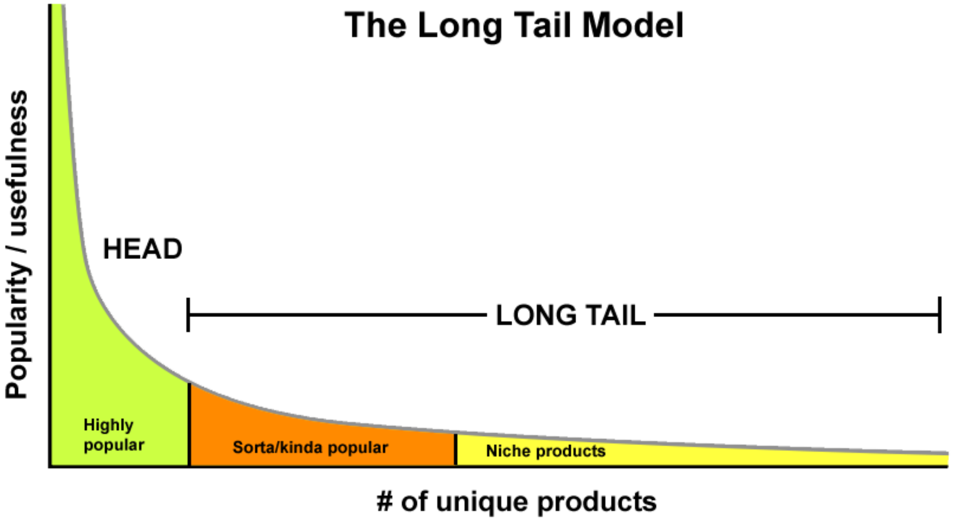
Overview of multidomain recommender systems

- Overview of recommender systems
- Challenges of recommender systems
- A case study of multidomain recommender systems
- The 3C-law to build multidomain recommender systems

Data characteristics in recommender systems

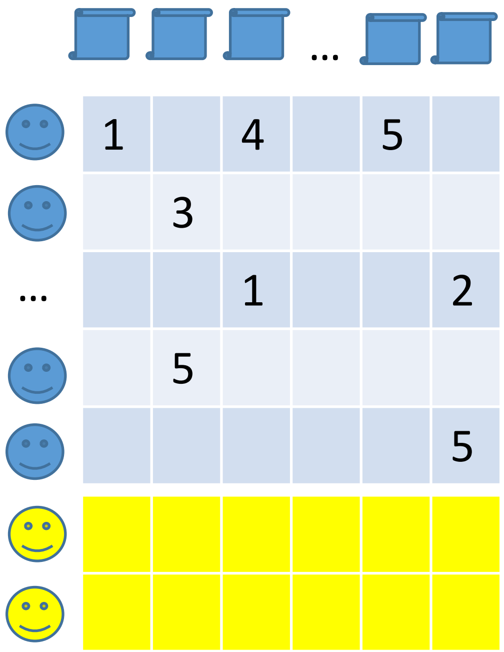
- Power law or Long tail distribution
  - Data associated with the majority of users are insufficient and even absent in real world.
  - In most recommender systems, the majority of users/items only associated with very few data while only minority of users/items have sufficient data





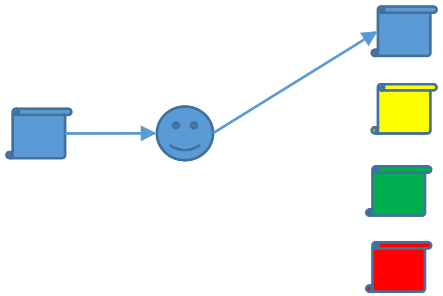
Challenges in collaborative filtering

- Data Sparsity
  - In real-world recommender systems, the user-item matrix is very sparse.
- Cold Start
  - When new users or new items are added, the system cannot recommend to these users and these items.
- Scalability
  - There are millions of users and products in real systems.
    - Large amount of computation
    - Large storage



Challenges in content-based filtering

- Limited Content Analysis
  - System has a limited amount of information on its users or the content of its items.
- Over-specialization
  - The system can only recommend items that highly similar with user’s profile, the user is limited to be recommended items similar to those already rated.



Question: what's the main cause of these challenges?

- Data Sparsity
- Cold Start
- Limited Content Analysis
- Over-specialization

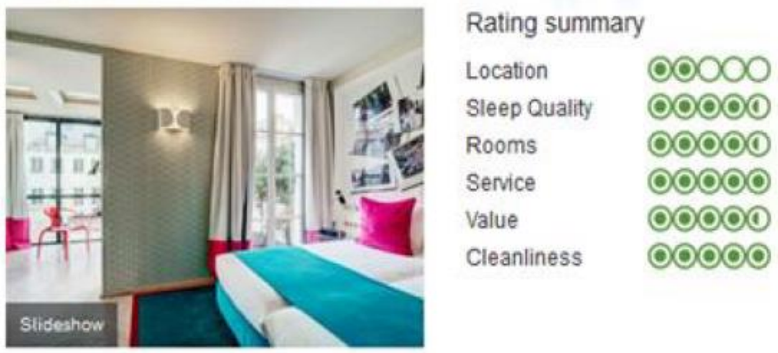
Data insufficiency in a single domain

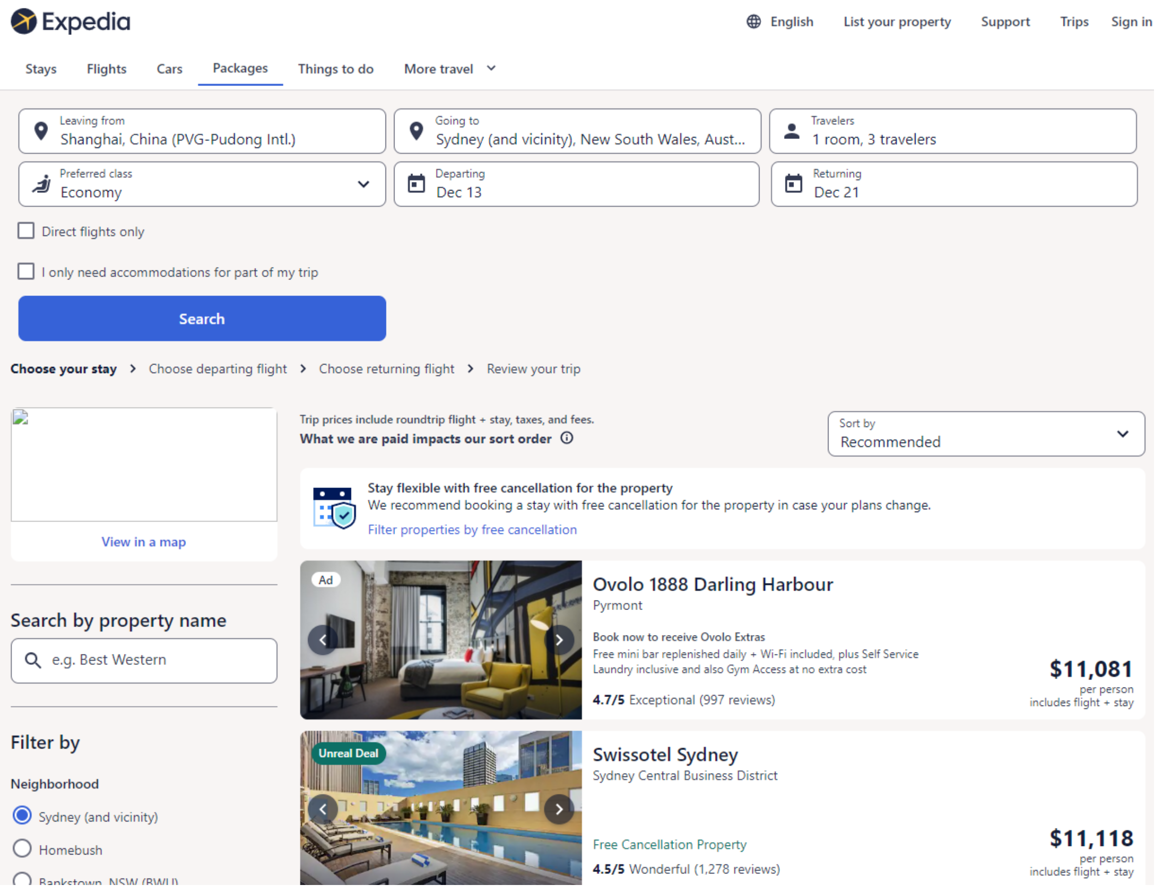
Outline

Overview of multidomain recommender systems

- Overview of recommender systems
- Challenges of recommender systems
- A case study of multidomain recommender systems
- The 3C-law to build multidomain recommender systems

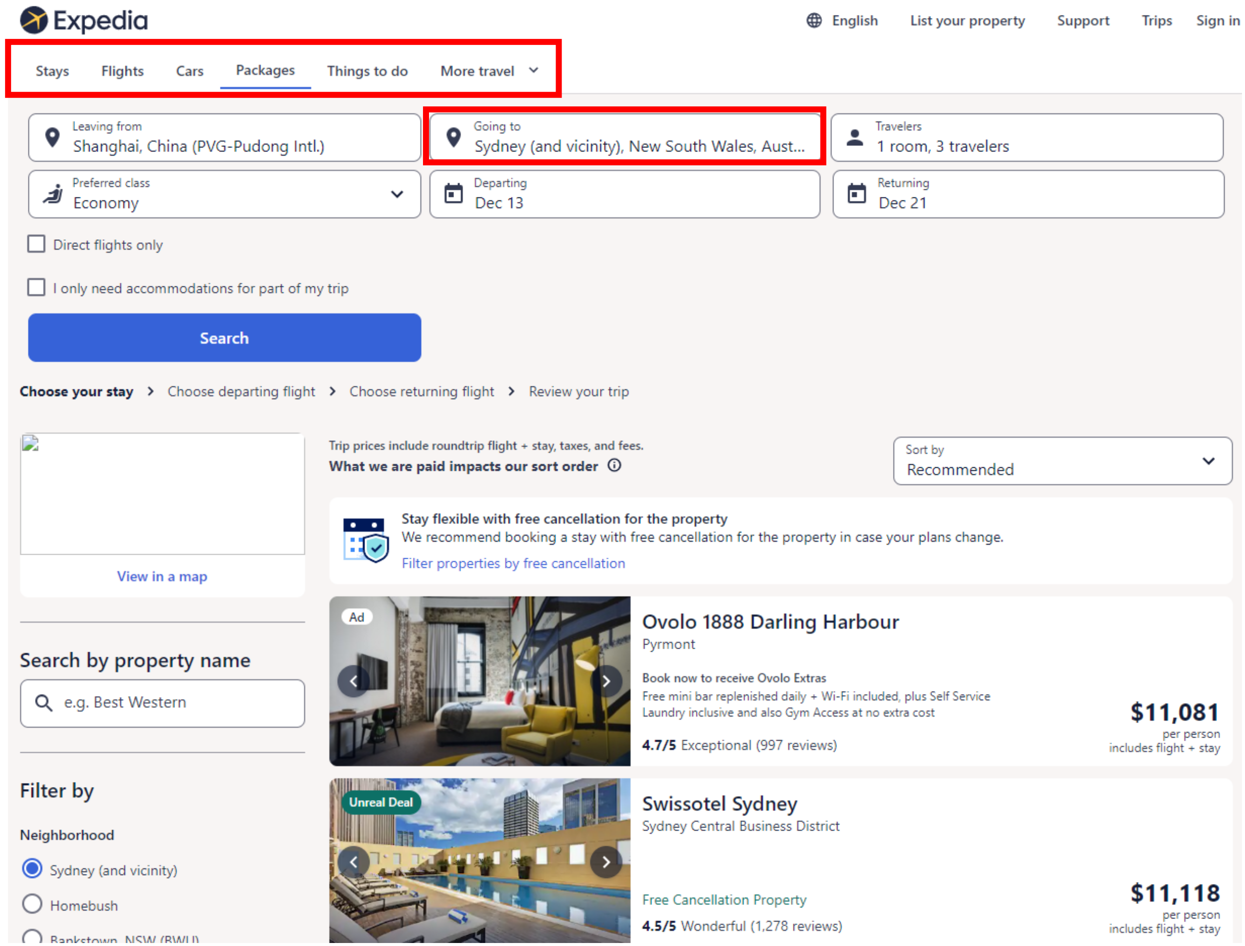
Case study: recommendations for a travel





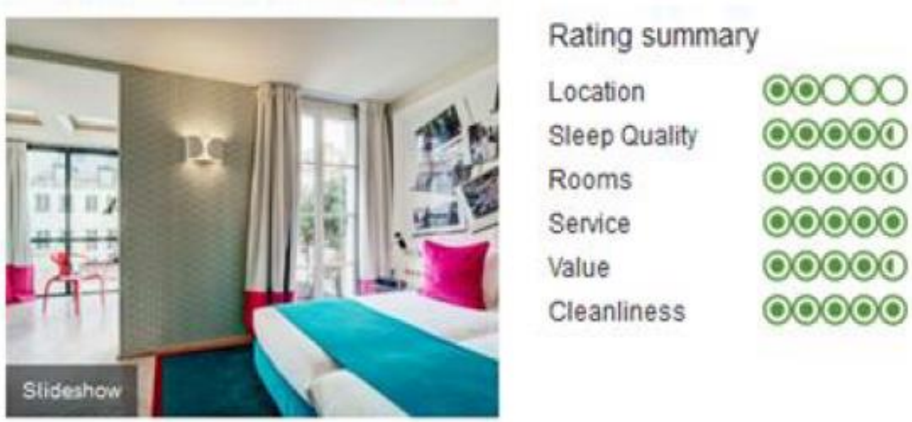
Multi-item domain recommendation

- Item types
  - E.g., flights, stays, cars
- Item relations:
  - E.g., destination (flight), location (hotel)



Multi-user domain recommendation

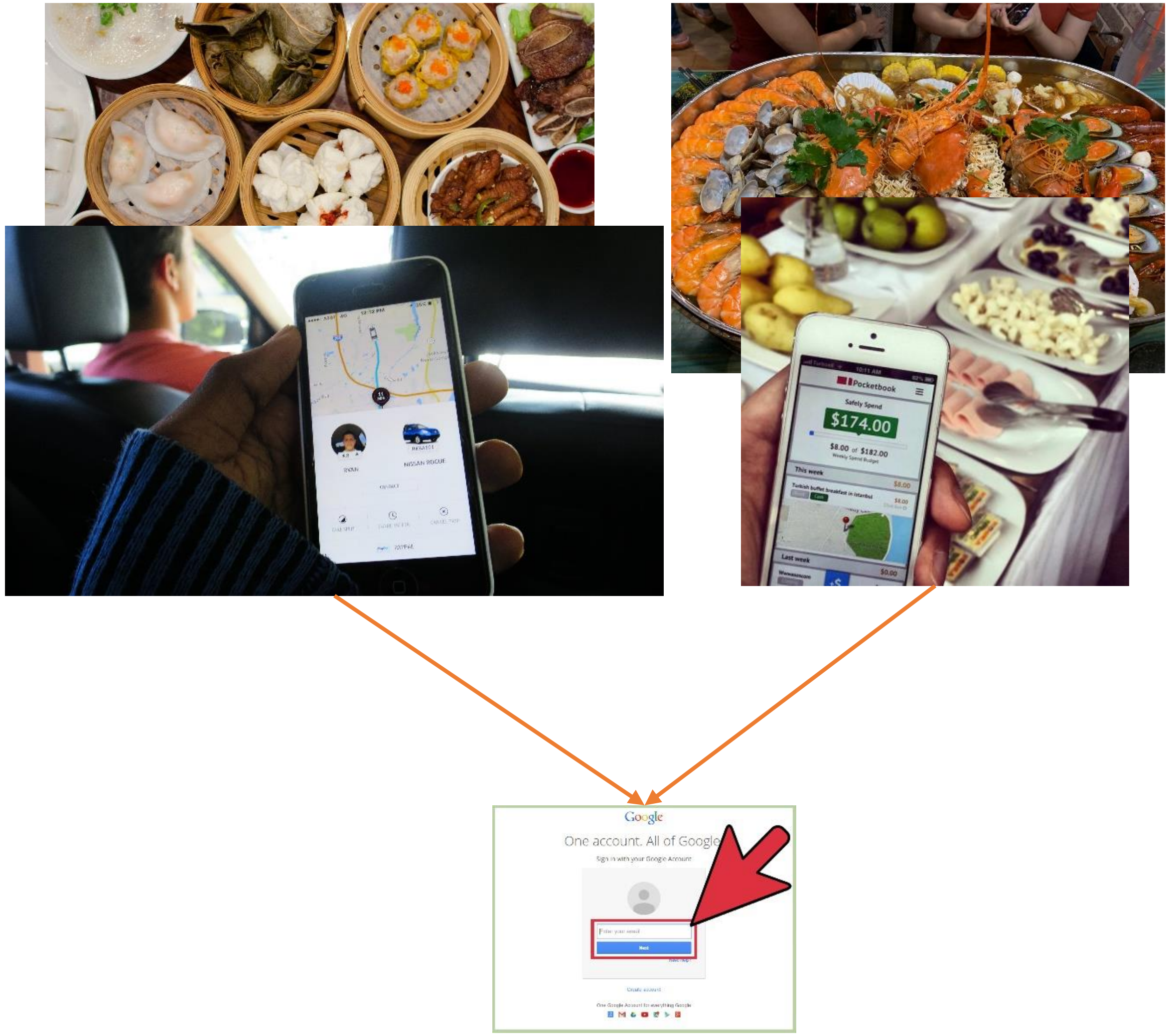
- User groups
  - E.g., family vs. colleagues for stays



- Social relations:
  - E.g., local food recommendations according to your local friends

Multi-spatial domain recommendation

- Geographic points
  - E.g., Chinese food restaurants (Hongkong), Seafood restaurants (Sydney)
- Systems/Sites:
  - E.g., Uber (google account), Eats Apps (google account)



Multi-temporal domain recommendation

- Sessions:
  - E.g. a shopping transaction may help travel package recommendation

Recent transaction

Travel package

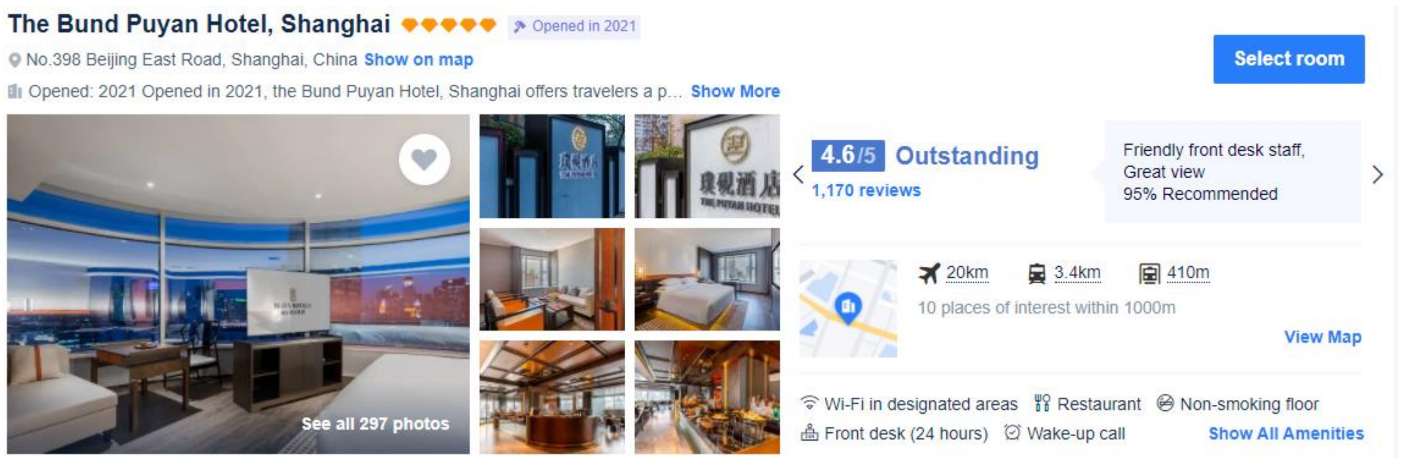
- Time period:
  - E.g., preferences of breakfast (morning) help to recommend dinner (evening)

Breakfast

Dinner

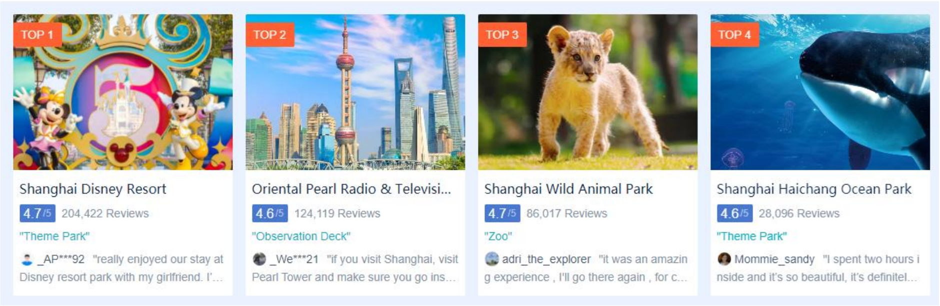
Multi-data domain recommendation

- Modalities:
  - E.g. rating, description, photos, geo information
- Distributions:
  - Ratings have different distributions for different domains, e.g., food vs stays
  - The content of image distribution is also different, e.g., color distributions between food and stays.

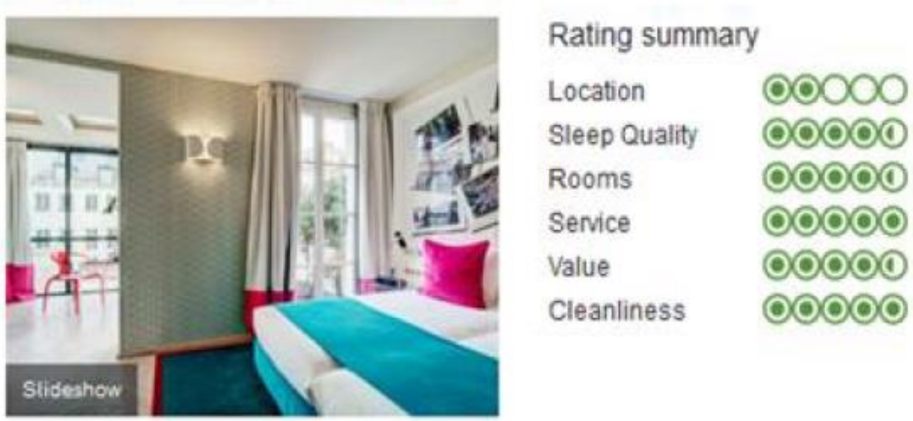


Multi-goal domain recommendation

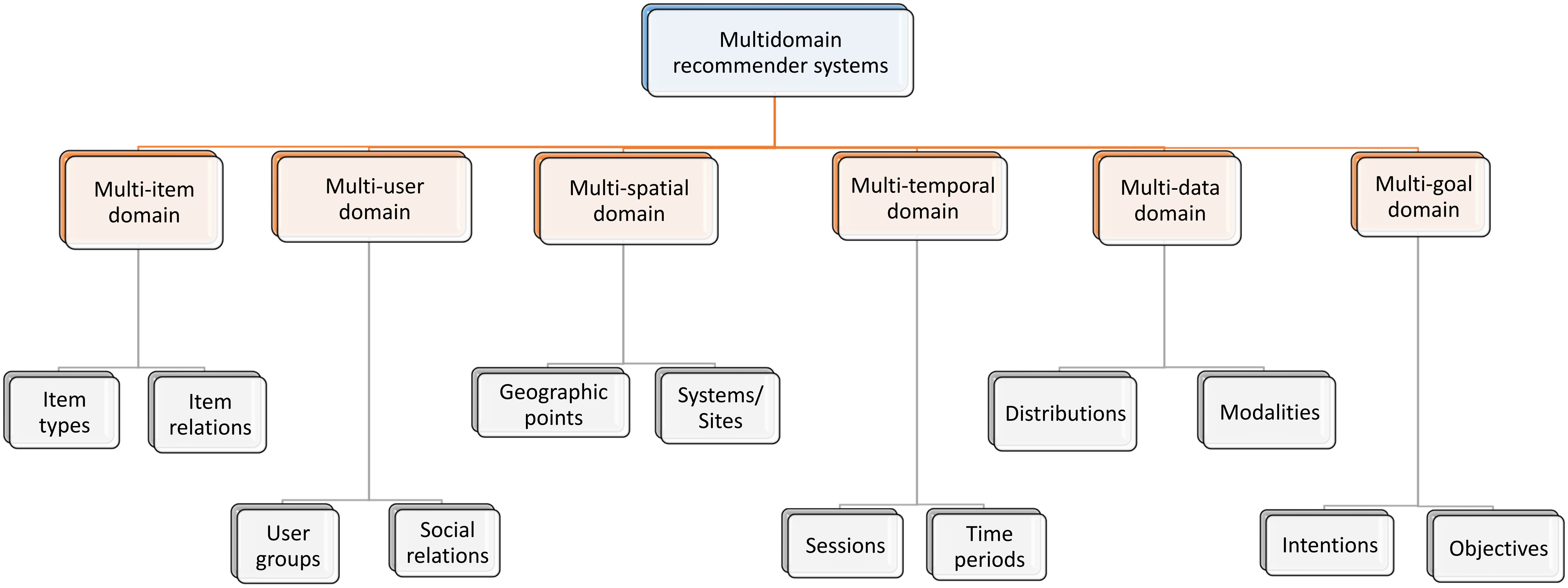
- Intentions:
  - A travel often consists of multiple intentions：
  - E.g. visiting attractions, enjoy local food.



- Objectives:
  - Multiple criteria:
    - E.g. location, cleanliness, services
  - Multiple tasks:
    - E.g. booking the cheapest flight and a sea view hotel



Multidomain RSs covered in this tutorial

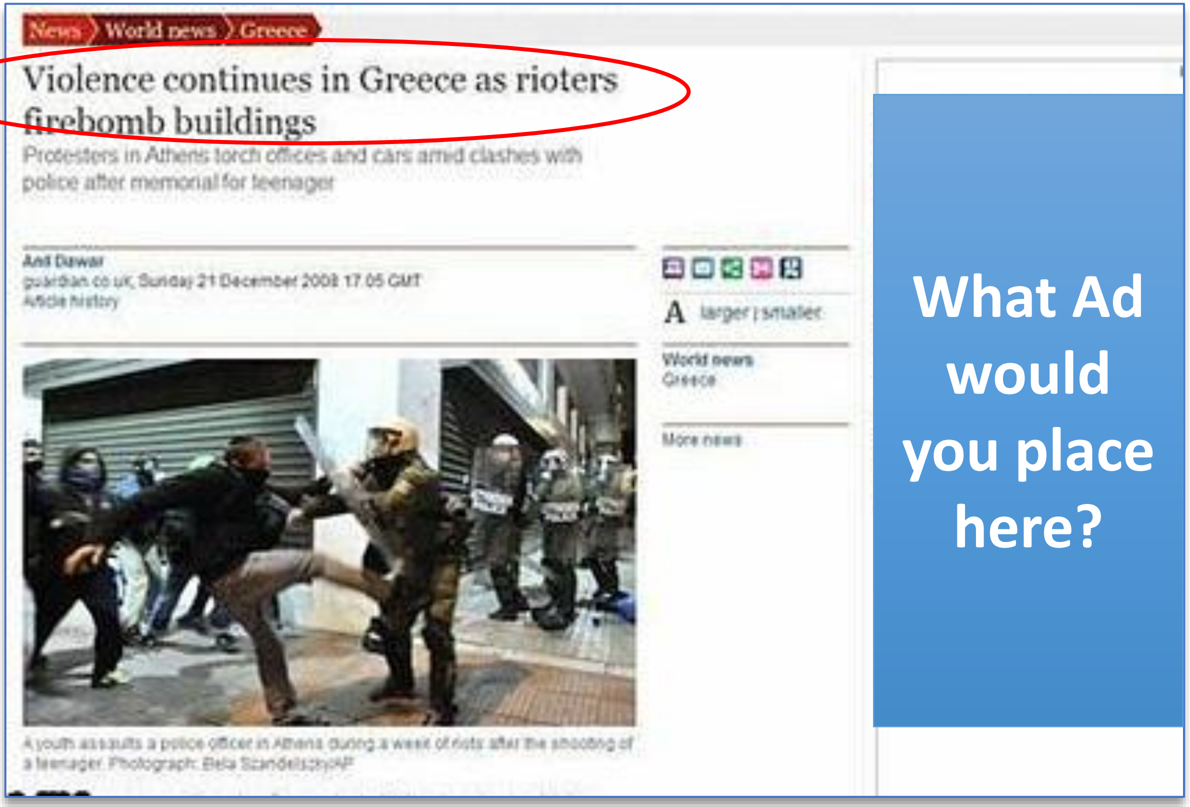


Outline

Overview of multidomain recommender systems

- Overview of recommender systems
- Challenges of recommender systems
- A case study of multidomain recommender systems
- The 3C-law to build multidomain recommender systems

Data complexity challenges to build multidomain recommender systems



Irrelevant and Damaging to Brand

Non-IIDness in multidomain RS modeling

- Heterogeneities:
  - Data types, attributes, sources, aspects, …
  - Formats, structures, distributions, relations, …
  - Learning outcomes

Not identically distributed.

- Coupling relationships:
  - Within and between values, attributes, objects, sources, aspects, …
  - Structures, distributions, relations, …
  - Methods, models, objectives…
  - Outcomes, impact, …

Not independent distributed.

Non-IIDness

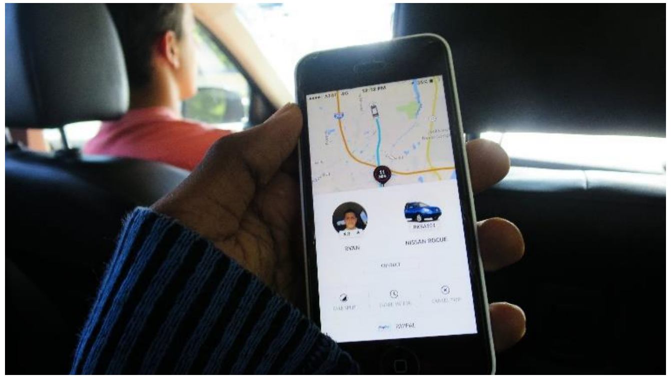
Longbing Cao. *Non-IIDness Learning in Behavioral and Social Data*, The Computer Journal, 57(9): 1358-1370 (2014).

3C law: complement, composite, context

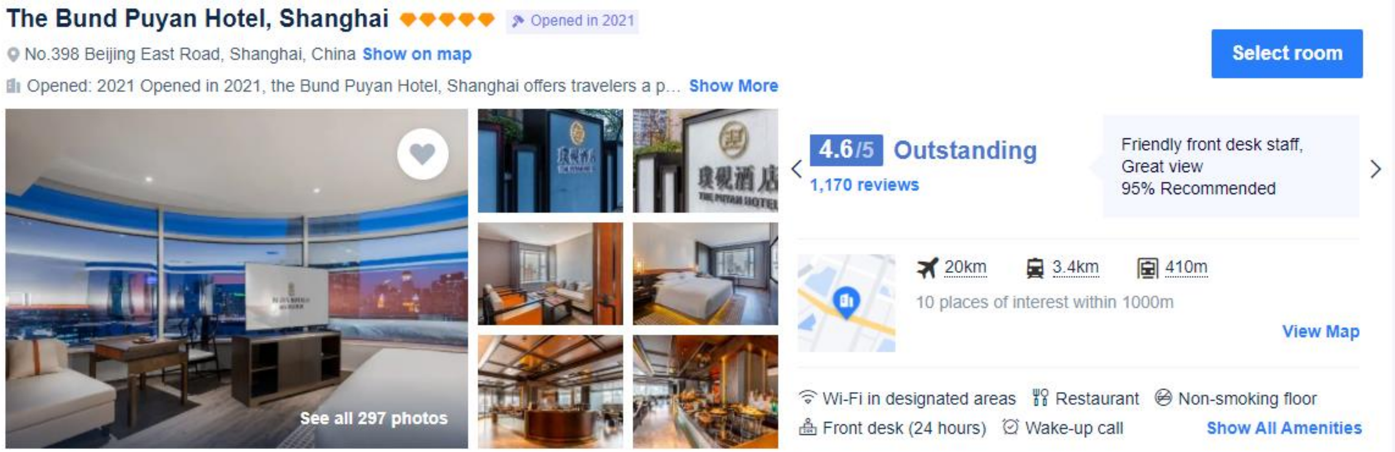
- Complement
  - To provide a complete view for recommendation by leveraging complement information over multiple domains.
- Composite
  - To provide comprehensive views for recommendation by integrating composite information over multiple domains.
- Context
  - To provide a conscious view for recommendation by considering context information over multiple domains.

Examples: complement information

- Transfer Uber histories for local restaurant recommendation

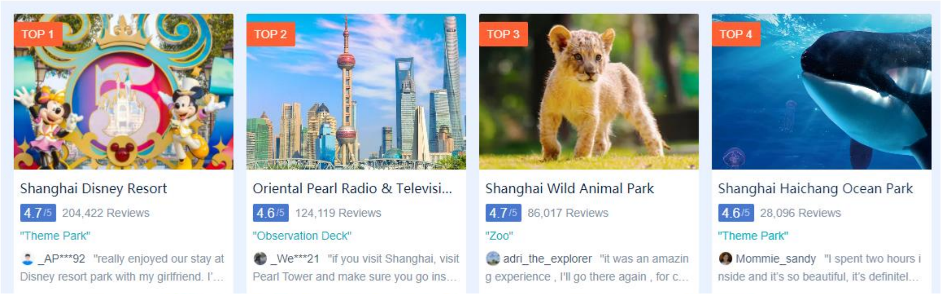


- Multimodal modeling for describing a hotel



Examples: composite information

- A travel consists of multiple requirements, including attraction, food, and living



- A recommendation for a hotel needs to comprehensively consider multiple aspects

Examples: context information

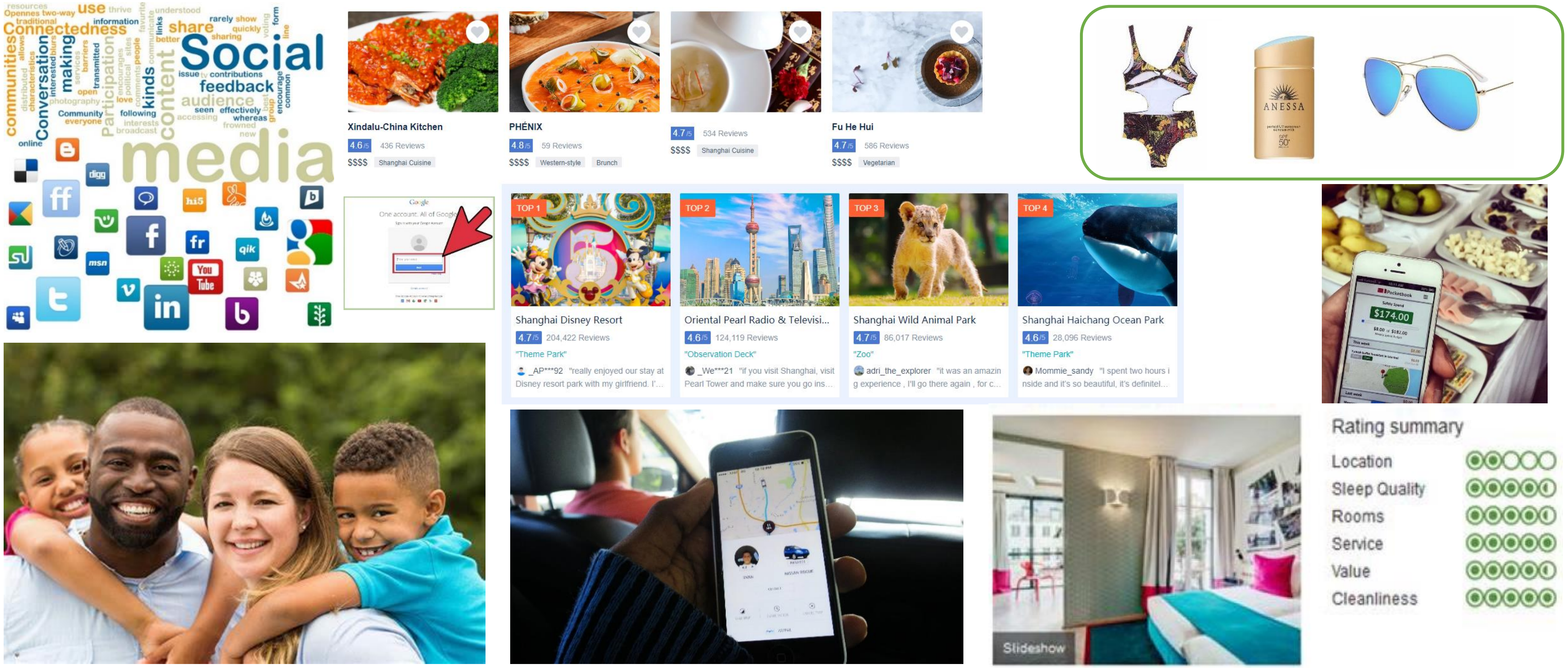
- A recommendation needs to consider the time and location

Breakfast

Dinner

- A recommendation needs to consider the preferences of all members in a user group

Multidomain recommender systems:
The way to the metaverse of recommendations



Next chapter

- Overview of multidomain recommender systems (Liang Hu, 25 mins)
- **Multi-item domain recommender systems (Z. Y. Lai, 20 mins)**
- Multi-user domain recommender systems (Qi Zhang, 20 mins)
- Multi-data domain recommender systems (Z. Y. Lai, 20 mins)
- Multi-spatial domain recommender systems (Qi Zhang 20 mins)
- Multi-temporal domain recommender systems (Z. Y. Lai, 20 mins)
- Multi-goal domain recommender systems (Liang Hu, 20 mins)
- Summary (Liang Hu, 5 mins)