Recommender Systems in Future Intelligence

Concordia University
Montreal, Canada
11 Jan. 2016

Kaveh Bakhtiyari

PhD Student at
University of Duisburg-Essen, Germany
&
The National University of Malaysia

Universiti Kebangsaan Malaysia
The National University of Malaysia

Universität Duisburg Essen

Index

Information Overload

- "People read around 10 MB worth of material a day, hear 400 MB a day, and see 1 MB of information every second" - The Economist, November 2006.

- In 2015, consumption will raise to 74 GB a day - UCSD Study 2014

Recommendation

- **CNN Money, “The race to create a 'smart' Google”:**
- *The Web, they say, is leaving the era of search and entering one of discovery. What's the difference? Search is what you do when you're looking for something. Discovery is when something wonderful that you didn't know existed, or didn't know how to ask for, finds you.*

Introduction

What is Recommender System?

- Recommender Systems (RS) generate a list of items (or people) to be recommended to the users. These systems predict the rating of the item which the user would give.

- Estimate a ***utility function*** to ***predict*** how a user will ***like*** an item.

Introduction

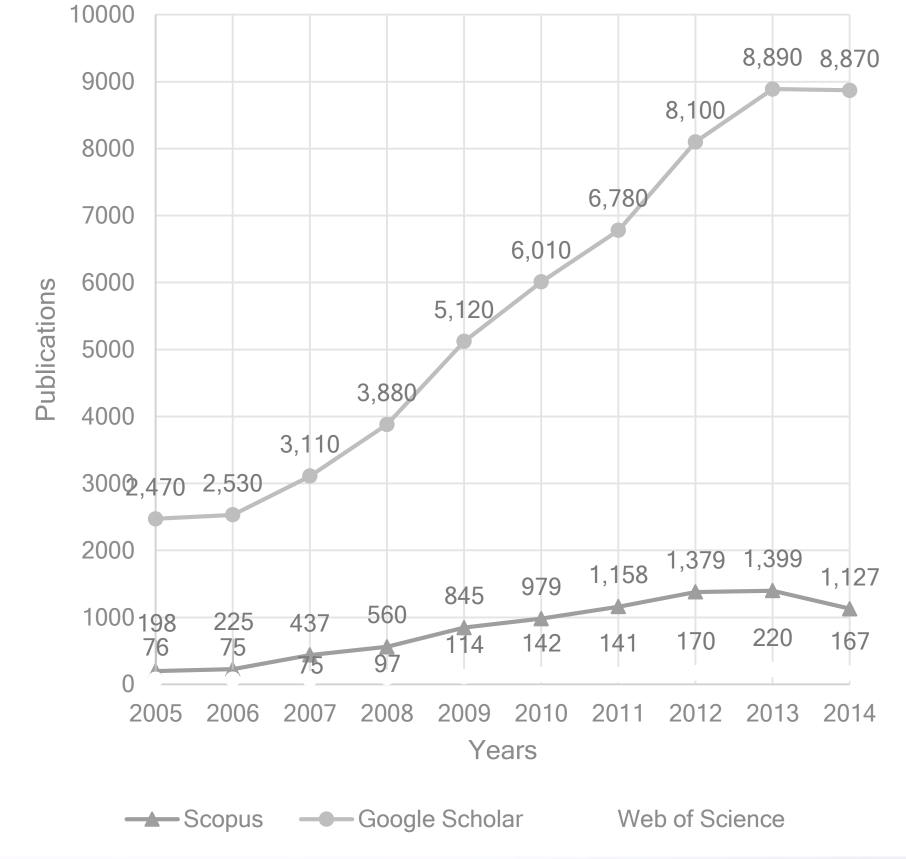
Why is Recommender System important?

- Netflix: 2/3 of the movies watched are recommended
- Google News: recommendations generate 38% more click-through
- Amazon: 35% sales from recommendations
- Choicestream: 28% of the people would buy more music if they found what they liked.

Introduction

RS as a research area

- Recommender Systems (RS) was being discussed in Data Mining and Information Filtering (Information Retrieval) areas, but it has been chosen as a separate research area in 1990s and it is becoming very popular.



UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

UNIVERSITÄT DUISBURG ESSEN

Approaches

Common Recommender Systems Approaches

- Collaborative Filtering
- Content-based Filtering
- Context-aware
- Demographic
- Social Recommendation (trust-aware)
- Hybrid

Index

- Introduction
- Collaborative Recommender System
  - User-based Filtering
  - Item-based Filtering
  - Sparsity
  - Netflix Prize
  - Clustering
  - Association Rules
  - Matrix Factorization
  - Pros & Cons
- Content-Based Filtering
  - TF-IDF
  - Pros & Cons
- Context-Aware RS
  - Tensor Factorization
- Other Approaches
  - Knowledge-Based RS
  - Demographic RS
  - Social & Trust RS
  - Ranking
- Hybridization
- Evaluation Criteria
- New Research Areas
  - HCI & AC
  - Look, Think, Feel
  - New features overview
- Future
- References
- Q&A

Collaborative Filtering

Steps:

- 1. Identify set of ratings for the **target/active user**
- 2. Identify set of users most similar to the target/active user according to a similarity function (**neighborhood** formation)
- 3. Identify the products these similar users liked
- 4. **Generate a prediction** - rating that would be given by the target user to the product - for each one of these products
- 5. Based on this predicted rating recommend a set of top N products

CF Approaches

- Memory-Based
  - User-based CF
  - Item-based CF
- Model-Based
  - Clustering (Classification)
  - Association rules
  - Matrix Factorization
  - Restricted Boltzmann Machines (RBM)
  - Probabilistic Latent Semantic Analysis

User-Based CF

- Target user u, ratings matrix Y
- $y_{v,i} \longrightarrow$ rating by user v for item i
- Similarity Pearson correlation sim(u,v) between users u & v

$$sim(u,v) = \frac{\sum_{i \in I_{uv}} (y_{u,i} - \hat{y}_u)(y_{v,i} - \hat{y}_v)}{\sqrt{\sum_{i \in I_{uv}} (y_{u,i} - \hat{y}_u)^2 \sum_{i \in I_{uv}} (y_{v,i} - \hat{y}_v)^2}}$$

- Predicted rating y*(u,i)

$$y^*(u,i) = \hat{y}_u + \frac{\sum_{j \in I_{y_{*j} \neq 0}} sim(v_j,u)(y_{v_j,i} - \hat{y}_{v_j})}{\sum_{j \in I_{y_{*j} \neq 0}} |sim(v_j,u)|}$$

User-Based CF: Example

						sim(u,v)
2			4	5		NA
5		4			1	
		5		2		
	1		5		4	
		4			2	
4	5		1			NA

User-Based CF: Example

						sim(u,v)
2			4	5		NA
5		4			1	0.87
		5		2		1
	1		5		4	-1
3.51*	3.81*	4	2.42*	2.48*	2	
4	5		1			NA

Item-Based CF

- Target item i
- $y_{u,j}$ → rating of user u for item i
- Similarity sim(i,j) between item i and j (Pearson correlation).

$$sim(i,j) = \frac{\sum_{u \in I_{ij}} (y_{u,i} - \hat{y}_i)(y_{u,j} - \hat{y}_j)}{\sqrt{\sum_{u \in I_{ij}} (y_{u,i} - \hat{y}_i)^2 \sum_{u \in I_{ij}} (y_{u,j} - \hat{y}_j)^2}}$$

- Predicted rating y*(u,i)

$$y^*(u,i) = \hat{y}_i + \frac{\sum_{v \in I_{y_{u*} \neq 0}} sim(i,j_v)(y_{u,j_v} - \hat{y}_{j_v})}{\sum_{v \in I_{y_{u*} \neq 0}} |sim(i,j_u)|}$$

Item-Based CF: Example

2			4	5	
5		4			1
		5		2	
	1		5		4
		4			2
4	5		1		

Item-Based CF: Example

	2			4	5	2.94*
	5		4			1
			5		2	2.48*
		1		5		4
			4			2
	4	5		1		1.12*
sim(i,j)	-1	-1	0.86	1	NA	

Similarity Computation

- Pearson correlation-based similarity
  - *does not account for user rating biases*
- Cosine based similarity
  - *does not account for user rating biases*

$$cos(i,j) = \frac{\sum_{u \in I_{ij}} y_{u,i} y_{u,j}}{\sqrt{\sum_{u \in I_{ij}} y_{u,i}^2 \sum_{u \in I_{ij}} y_{u,j}^2}}$$

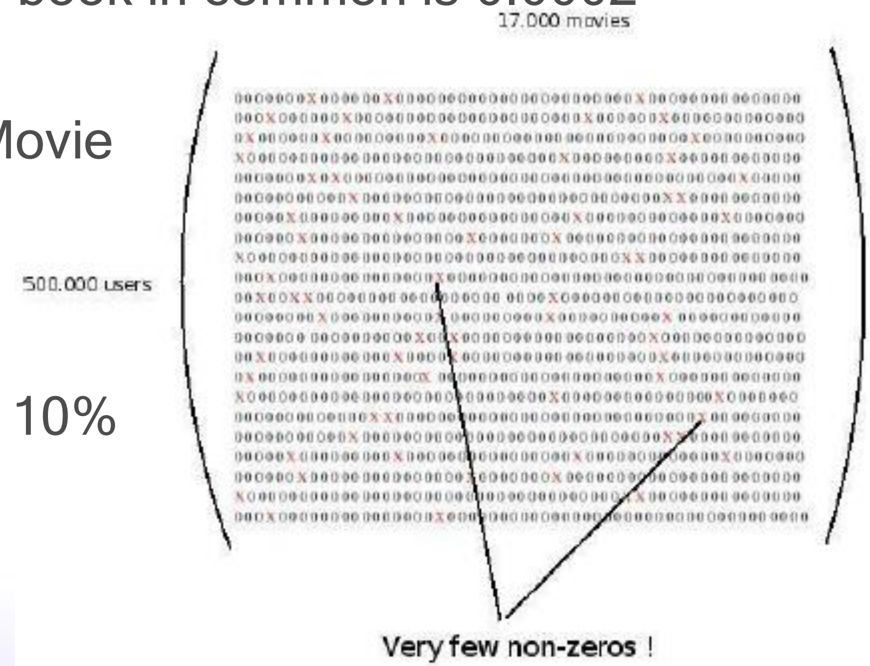
- Adjusted cosine similarity
  - *takes care of user rating biases as each pair in the co-rated set corresponds to a different user.*

$$sim(i,j) = \frac{\sum_{u \in I_{ij}} (y_{u,i} - \hat{y}_u)(y_{u,j} - \hat{y}_u)}{\sqrt{\sum_{u \in I_{ij}} (y_{u,i} - \hat{y}_u)^2 \sum_{j \in I_{uv}} (y_{u,j} - \hat{y}_u)^2}}$$

Collaborative Filtering

Sparsity Problem:

- Typically large product sets & few user ratings e.g. Amazon:
- in a catalogue of 1 million books, the probability that two users who bought 100 books each, have a book in common is 0.01
- in a catalogue of 10 million books, the probability that two users who bought 50 books each, have a book in common is 0.0002



- Netflix Prize rating data in a User/Movie matrix:
  - 500,000 x 17,000 = 8,500 M positions
  - Out of which only 100M are not 0's!
- CF must have a number of users ~ 10% of the product catalogue size

Netflix Prize

- Looking for: High quality recommendation
- Evaluation metric: RMSE

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}$$

Accuracy Improvement by 10%

=

1,000,000$

Netflix Prize

- **2007 Top Algorithms:**
- SVD: RMSE = 0.8914
- RBM: RMSE = 0.8990
- Linear Blend: RMSE = 0.88
- **2008 Top Algorithm:**
- SVD++ RMSE = 0.8567
- **Limitations:**
- Designed for 100M ratings (the actual number of ratings was 10B ratings)
- Not adaptable as users add new ratings
- Performance issues

Solution to Sparsity

Model-Based Collaborative Filtering:

- Clustering
- Association rules
- Matrix Factorization
- Restricted Boltzmann Machines (RBM)

Clustering

	Book1	Book2	Book3	Book4	Book5	Book6
Customer A	X			X		
Customer B		X	X		X	
Customer C		X	X			
Customer D		X				X
Customer E	X				X	

- Customers B, C and D are « clustered » together.
- Customers A and E are clustered into another separate group

- «Typical» preferences for CLUSTER are:
- Book 2, very high
- Book 3, high
- Books 5 and 6, may be recommended
- Books 1 and 4, not recommended at all

Clustering

	Book1	Book2	Book3	Book4	Book5	Book6
Customer A	X			X		
Customer B		X	X		X	
Customer C		X	X			
Customer D		X				X
Customer E	X				X	
Customer F			X		X	

- Any customer that shall be classified as a member of CLUSTER will receive recommendations based on preferences of the group:

- Book 2 will be highly recommended to *Customer F*
- Book 6 will also be recommended to some extent

Association Rules

Past purchases used to find relationships of common purchases

	BOOK 1	BOOK 2	BOOK 3	BOOK 4	BOOK 5	BOOK 6
CUSTOMER A	X			X		
CUSTOMER B		X	X		X	
CUSTOMER C		X	X			
CUSTOMER D		X				X
CUSTOMER E	X				X	
CUSTOMER F			X		X	

	BOOK 1	BOOK 2	BOOK 3	BOOK 4	BOOK 5	BOOK 6
BOOK 1				1	1	
BOOK 2			2		1	1
BOOK 3		2			2	
BOOK 4	1					
BOOK 5	1	1	2			
BOOK 6		1				

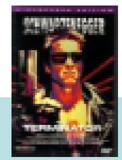
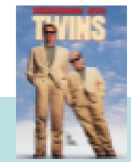
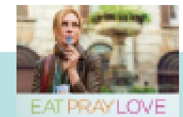
Universiti Kebangsaan Malaysia
The National University of Malaysia

Universität Duisburg Essen

Matrix Factorization

- **SVD:** $$M_k = U_k \times \Sigma_k \times V_k^T$$

U_k	Dim1	Dim2
Alice	0.47	-0.30
Bob	-0.44	0.23
Mary	0.70	-0.06
Sue	0.31	0.93

V_k^T					
Dim1	-0.44	-0.57	0.06	0.38	0.57
Dim2	0.58	-0.66	0.26	0.18	-0.36

Σ_k	Dim1	Dim2
Dim1	5.63	0
Dim2	0	3.23

- **Prediction:** $$\hat{r}_{ui} = \bar{r}_u + U_k(Alice) \times \Sigma_k \times V_k^T(EPL)$$

= 3 + 0.84 = 3.84

Collaborative Filtering

Pros:

- Requires **minimal knowledge** engineering efforts
- Users and products are symbols without any internal structure or characteristics
- Produces good-enough results in most cases

Cons:

- Sparsity Problem: Requires a large number of **reliable** "user feedback data points" to bootstrap
- Requires products to be standardized (users should have bought **exactly** the same product)
- Assumes that **prior behavior determines current behavior** without taking into account "contextual" knowledge (session-level)

Index

Content-Based Filtering

- A pure content-based recommender system makes recommendations for a user based solely on the profile built up by **analyzing the content** of items which that user has rated in the past.

- **What is content?**
- It can be explicit **attributes** or **characteristics** of the item. For example for a film:
  - Genre: Action / adventure
  - Feature: Bruce Willis
  - Year: 1995
- It can also be **textual content** (title, description, table of content, etc.)
  - Several techniques to compute the distance between two textual documents
  - Can use NLP techniques to extract content features
- Can be extracted from the signal itself (audio, image)

Content-Based Filtering

- “Importance” (or “informativeness”) of word kj in document dj is determined with some weighting measure wij
- One of the best-known measures in IR is the term frequency/inverse document frequency (TF-IDF)

- TF-IDF encodes text documents as weighted term vector
- TF: Measures, how often a term appears (density in a document)
  - Assuming that important terms appear more often
  - Normalization has to be done in order to take document length into account
- IDF: Aims to reduce the weight of terms that appear in all documents

$$\mathrm{tf}(t,d) = 0.5 + \frac{0.5 \times \mathrm{f}(t,d)}{\max\{\mathrm{f}(w,d) : w \in d\}} \qquad \mathrm{idf}(t,D) = \log \frac{N}{|\{d \in D : t \in d\}|}$$

$$\mathrm{tfidf}(t,d,D) = \mathrm{tf}(t,d) \times \mathrm{idf}(t,D)$$

Content-Based Filtering

Pros:

- No community required
- No sparsity problem
- Can recommend new and unpopular items
- Easier to be explained

Cons:

- Content descriptions necessary
- Cold start for new users
- No surprises
- Suitable only for same type of items

Index

- Introduction
- Collaborative Recommender System
  - User-based Filtering
  - Item-based Filtering
  - Sparsity
  - Netflix Prize
  - Clustering
  - Association Rules
  - Matrix Factorization
  - Pros & Cons
- Content-Based Filtering
  - TF-IDF
  - Pros & Cons
- Context-Aware RS
  - Tensor Factorization
- Other Approaches
  - Knowledge-Based RS
  - Demographic RS
  - Social & Trust RS
  - Ranking
- Hybridization
- Evaluation Criteria
- New Research Areas
  - HCI & AC
  - Look, Think, Feel
  - New features overview
- Future
- References
- Q&A

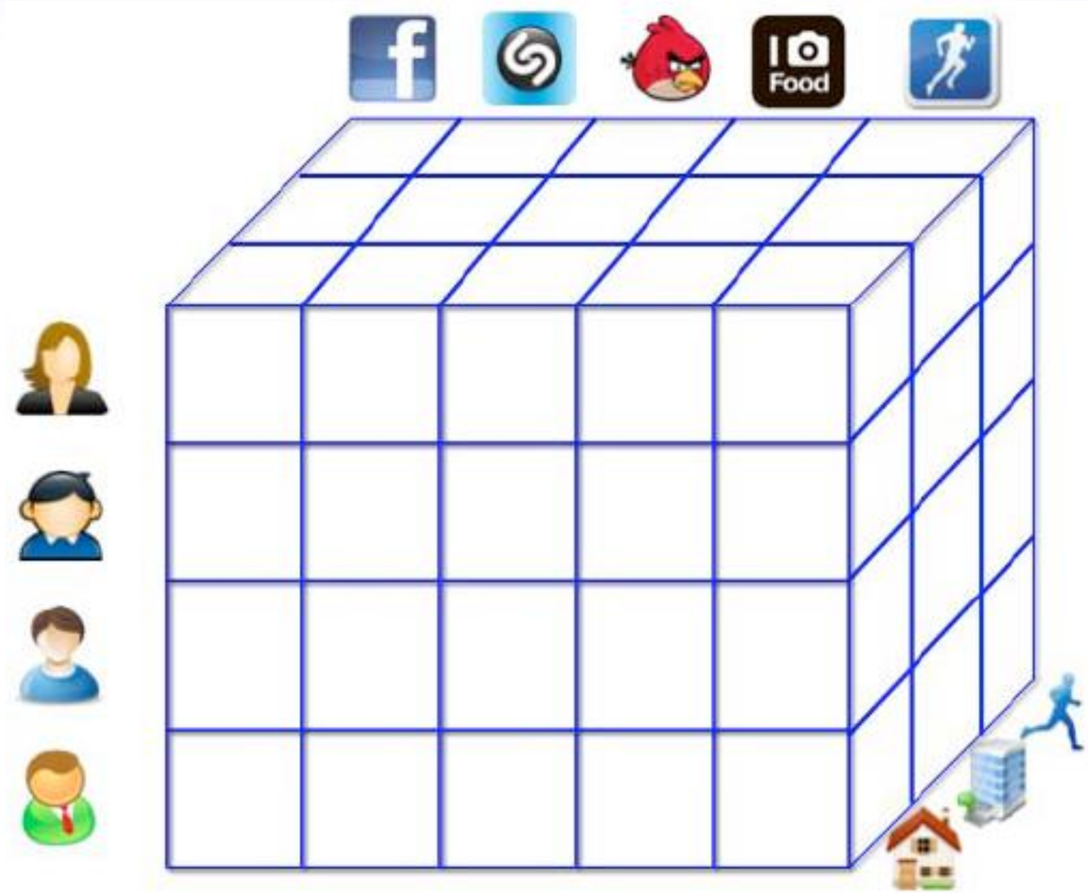
Context-Aware RS

- Context is a dynamic set of factors describing the state of the user at the moment of the user's experience
- Context factors can rapidly change and affect how the user perceives an item

Type of Context:

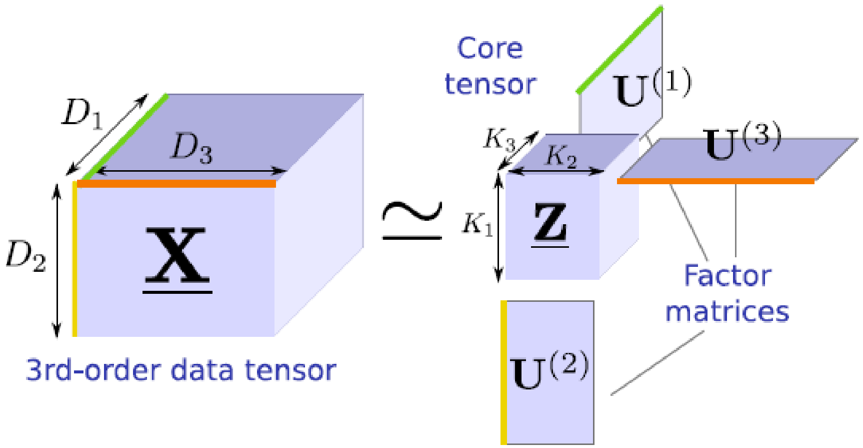
- **Temporal:** Time of the day, week / weekend
- **Spatial:** Location, Home, Work, etc.
- **Social:** With friends, Family

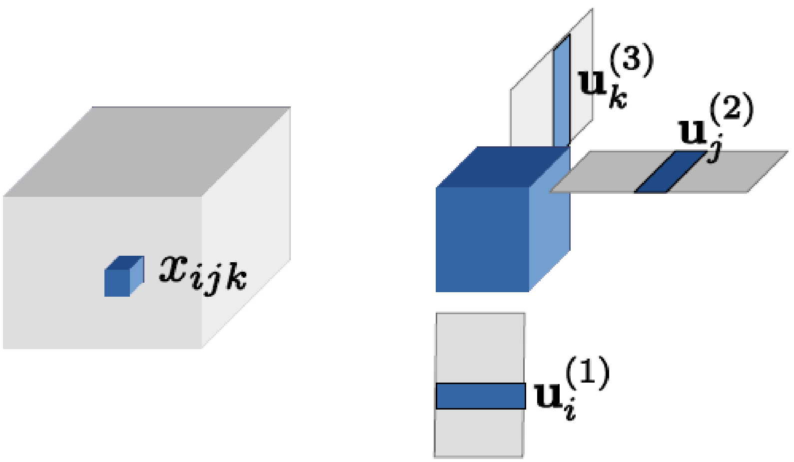
Context-Aware RS



$$F_{ijk} = \langle U_i, M_j, C_k \rangle$$

Tensor Factorization





$$x_{ijk} = \sum_{q=1}^{K_1} \sum_{r=1}^{K_2} \sum_{s=1}^{K_3} u_{iq}^{(1)} u_{jr}^{(2)} u_{ks}^{(3)} z_{qrs} + \varepsilon_{ijk}$$

Universiti Kebangsaan Malaysia
The National University of Malaysia

Universität Duisburg Essen

Index

Knowledge-Based RS

Knowledge-based: "Tell me what fits based on my needs“

Views:

- Case-based: Similarity functions
- Utility-based
- Constraint-based:
  - **IF** purpose=“*on travel*” **THEN** lower focal length < *28mm*

WHY:

- Low number of available rankings
- Timespan plays an important roles
- Customers want to define their requirements explicitly
  - “The color of the car should be black"

Demographic RS

- Aim to categorize the user based on personal attributes and make recommendation based on demographic classes

	gender	age	area code	education	employed	Dolce
Karen	F	15	714	HS	F	+
Lynn	F	17	714	HS	F	–
Chris	M	35	714	C	T	+
Mike	F	40	714	C	T	–
Jill	F	10	714	E	F	?

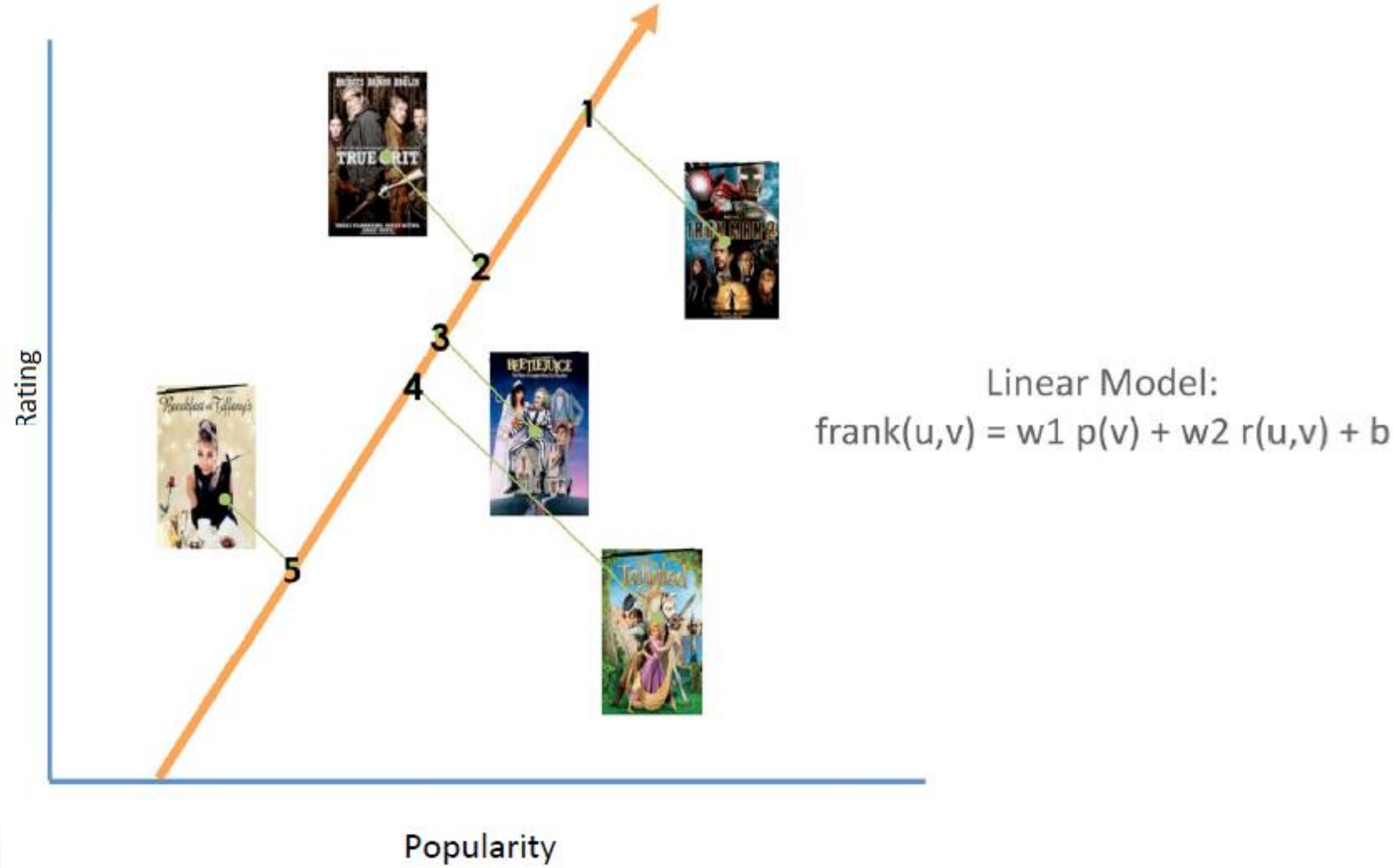
Social & Trust RS

- A social recommender system recommends items that are “popular” in the social proximity of the user
- In the context of recommender systems, trust is generally used to describe similarity in opinion
- Use trust to give more weight to some users
- Use trust in place of (or combined with) similarity

- Publicly available dataset: epinions dataset

Ranking

- Most recommendations are presented in a sorted list



Index

Hybridization

Hybridization Method	Description
Weighted	Outputs (scores or votes) from several techniques are combined with different degrees of importance to offer final recommendations
Switching	Depending on situation, the system changes from one technique to another
Mixed	Recommendations from several techniques are presented at the same time
Feature combination	Features from different recommendation sources are combined as input to a single technique
Cascade	The output from one technique is used as input of another that refines the result
Feature augmentation	The output from one technique is used as input features to another
Meta-level	The model learned by one recommender is used as input to another

Index

Evaluation Criteria

- Accuracy
  - RMSE, MAE, Precision, Recall, F1
- Coverage
- Novelty
- Diversity
- Reliability
- Serendipity
- Utility
- Robustness & Stability
- Privacy
- Adaptivity
- Scalability
- Trust
- Confidence
- Risk

Every single criterion is a research challenge!!!

Index

New research areas

Human-Computer Interactions

Affective Computing

Implicit data collection:

- Analysis of human-computer interaction features such as mouse movements, keyboard keystroke dynamics, and touch-screen interactions.
- Analysis of users’ cultural backgrounds.
- Analysis of users’ emotional states (Emotional Intelligence)

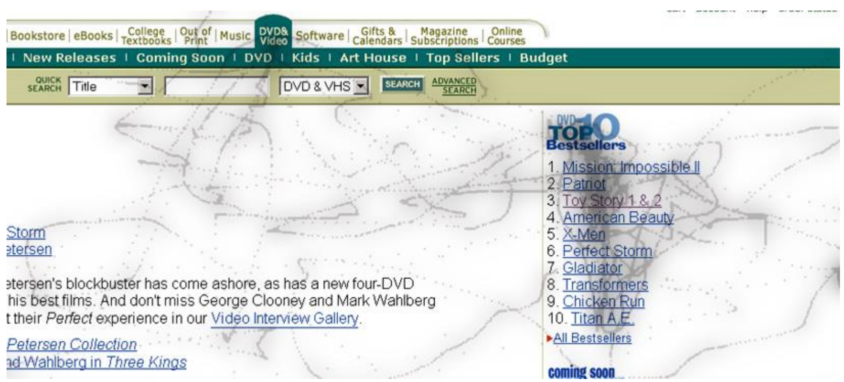


Image Source: Cheese Project, MIT University

Look, Think, Feel

- What is the user looking at?
  - Eye Tracking:

  Image Processing High Computational Cost

  Requires Web Cam

  Privacy Issues

- What is the user thinking about?
  - EEG Device

- How is the user feeling about?
  - Affective Computing:
  - (Facial Expressions)
  - (Skin Conductance)
  - (Heart Rate)
  - (Wearable devices)
  - etc.

Universiti Kebangsaan Malaysia
The National University of Malaysia

Universität Duisburg Essen

Index

- Introduction
- Collaborative Recommender System
  - User-based Filtering
  - Item-based Filtering
  - Sparsity
  - Netflix Prize
  - Clustering
  - Association Rules
  - Matrix Factorization
  - Pros & Cons
- Content-Based Filtering
  - TF-IDF
  - Pros & Cons
- Context-Aware RS
  - Tensor Factorization
- Other Approaches
  - Knowledge-Based RS
  - Demographic RS
  - Social & Trust RS
  - Ranking
- Hybridization
- Evaluation Criteria
- New Research Areas
  - HCI & AC
  - Look, Think, Feel
  - New features overview
- **Future**
- References
- Q&A

Future

Here is the future of personalized content (computational advertising) by Recommender System and Artificial Intelligence......

References

- Del Missier, F., & Ricci, F. (**2003**). *Understanding recommender systems: Experimental evaluation challenges.* Paper presented at the Proceedings of the Second Workshop on Empirical Evaluation of Adaptive Systems, held at the 9th International Conference.
- Fang, B., Liao, S., Xu, K., Cheng, H., Zhu, C., & Chen, H. (**2012**). A novel mobile recommender system for indoor shopping. *Expert Systems with Applications, 39*(15), 11992-12000. doi: 10.1016/j.eswa.2012.03.038
- Happel, H. J., & Maalej, W. (**2008**). *Potentials and challenges of recommendation systems for software development.* Paper presented at the Proceedings of the 2008 international workshop on Recommendation systems for software engineering.
- Hussein, T. (**2010**). Context-aware Recommendations.
- Konstan, J. A., & Riedl, J. (**2012**). Deconstructing Recommender Systems. Retrieved 06/12/2012, from http://spectrum.ieee.org/computing/software/deconstructing-recommender-systems
- Levandoski, J. J., Sarwat, M., Eldawy, A., & Mokbel, M. F. (**2012**). *Lars: A location-aware recommender system.* Paper presented at the Data Engineering (ICDE), 2012 IEEE 28th International Conference on.
- Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (**2010**). The Cultural Psychology of the Emotions - Ancient and Renewed. In J. H. Richard A. Shweder, Randall Horton, Craig Joseph (Ed.), *Handbook of Emotions* (3rd ed., pp. 19): The Guilford Press.
- Nunes, M. A. S. N. (**2010**). Towards to Psychological-based Recommenders Systems: A survey on Recommender Systems. *Scientia Plena, 6*(8).
- Park, D. H., Kim, H. K., Choi, I. Y., & Kim, J. K. (**2012**). A literature review and classification of recommender systems research. *Expert Systems with Applications, 39*(11), 10059-10072. doi: 10.1016/j.eswa.2012.02.038
- Rajaraman, A., & Ullman, J. (**2011**). Recommendation Systems *Mining of Massive Datasets*: Cambridge University Press.
- Ricci, F., Rokach, L., & Shapira, B. (**2011**). Introduction to recommender systems handbook *Recommender Systems Handbook* (pp. 1-35).
- Robillard, M., Walker, R., & Zimmermann, T. (**2010**). Recommendation systems for software engineering. *Software, IEEE, 27*(4), 80-86. doi: 10.1109/MS.2009.161
- Shani, G., Brafman, R. I., & Heckerman, D. (**2002**). *An MDP-based recommender system.* Paper presented at the Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence.
- Shani, G., & Gunawardana, A. (**2011**). Evaluating recommendation systems *Recommender Systems Handbook* (pp. 257-297): Springer US.
- Strickroth, S., & Pinkwart, N. (**2012**). *High quality recommendations for small communities: the case of a regional parent network*. Paper presented at the Proceedings of the sixth ACM conference on Recommender systems, Dublin, Ireland.
- Tkalcic, M., Košir, A., & Tasic, J. (**2011**). *Affective recommender systems: the role of emotions in recommender systems.* Paper presented at the Proceedings of the RecSys 2011 Workshop on Human Decision Making in Recommender Systems. Decisions@ RecSys11.
- Ullah, F., Sarwar, G., Lee, S. C., Park, Y. K., Moon, K. D., & Kim, J. T. (**2012**). *Hybrid recommender system with temporal information.* Paper presented at the Information Networking (ICOIN), 2012 International Conference on.
- Xu, B., Bu, J., Chen, C., & Cai, D. (**2012**). *An exploration of improving collaborative recommender systems via user-item subgroups.* Paper presented at the Proceedings of the 21st international conference on World Wide Web.

Q&A

Thank you for your attention.