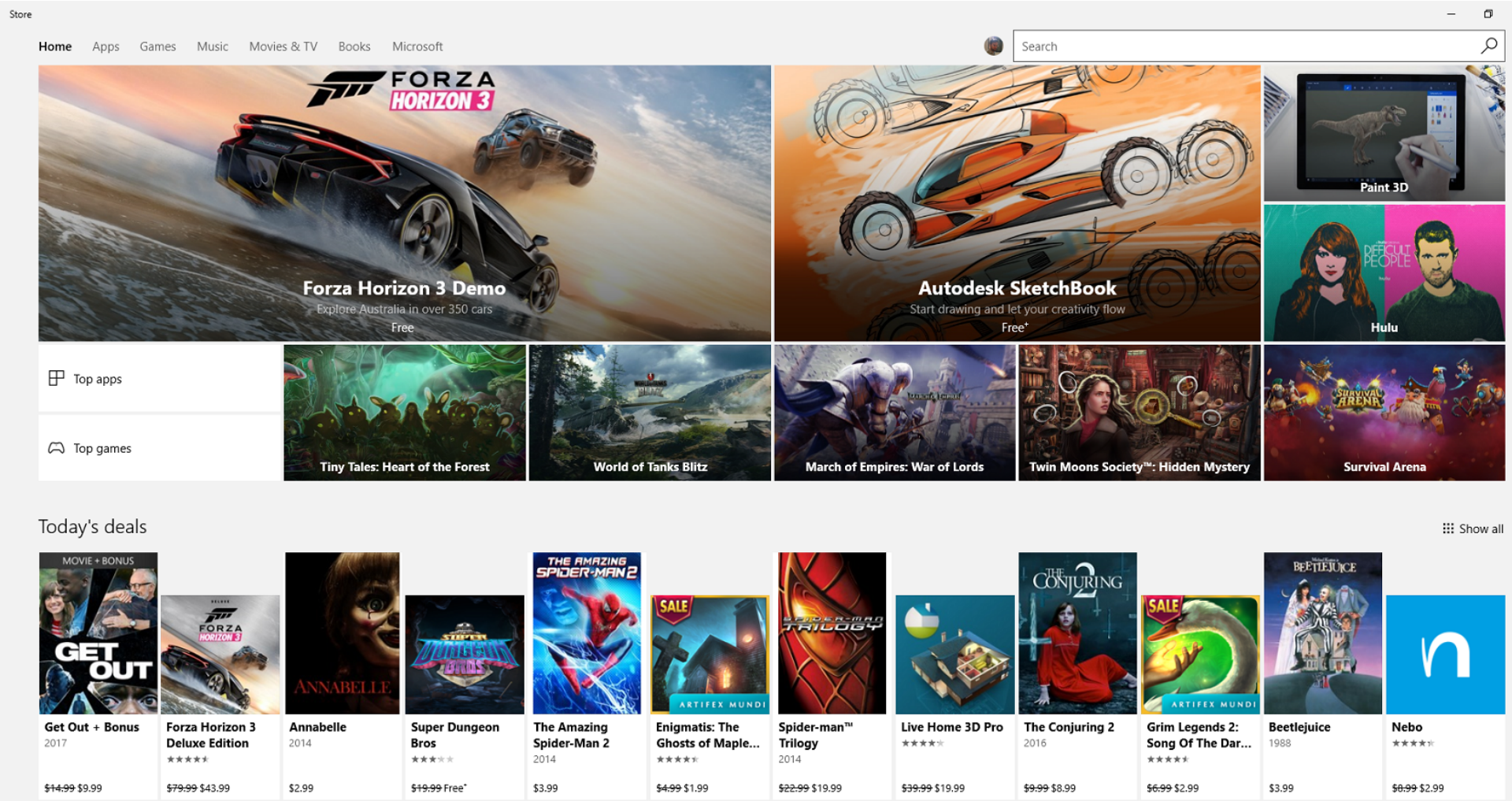
Rethinking Collaborative Filtering: A Practical Perspective on State-Of-The-Art Research Based on "Real-World" Insights and Challenges

Noam Koenigstein

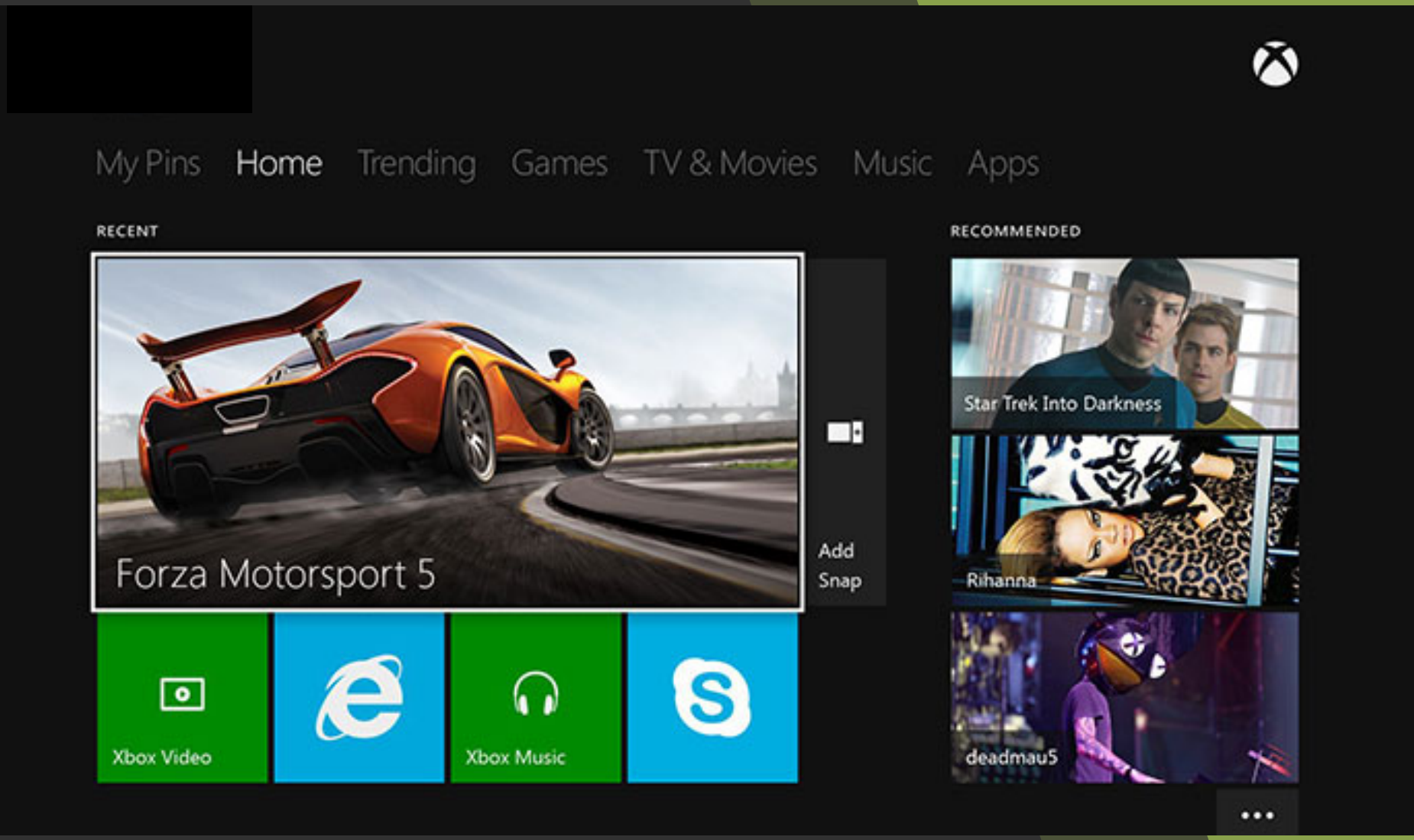
Microsoft

RECOMMENDATIONS IN MICROSOFT STORE

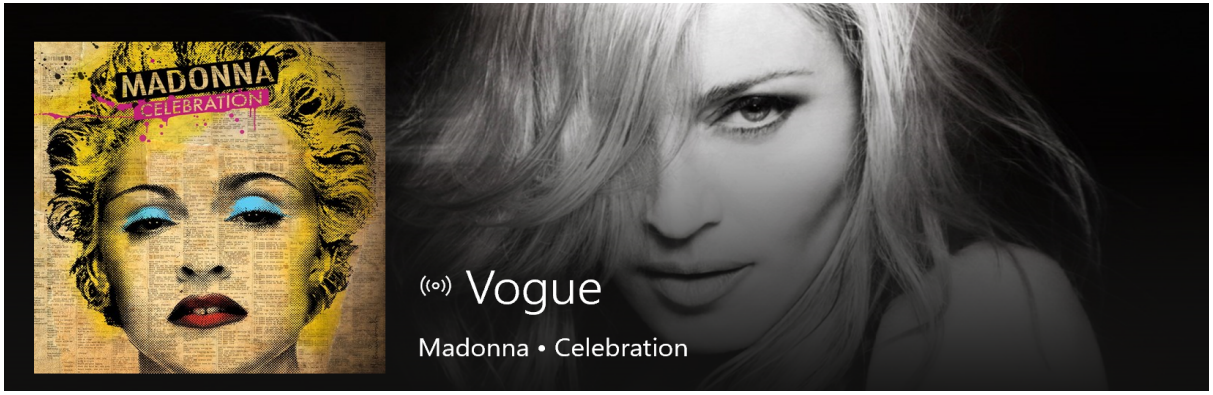
Windows Store



The Xbox Marketplace



Groove Radio



Vogue	Madonna	Celebration (Double Disc	2009	Pop	5:16
Well, Well	Nelly Furtado	Whoa, Nelly!	2000	Pop	2:59
1,2,3 (Remix)	Gloria Estefan	Greatest Hits	1992	Pop	3:34
Viva Forever	Spice Girls	Spiceworld	1997	Rock	5:10
Man in the Mirror (2012 Rem	Michael Jackson	Bad (Remastered)	1987	Pop	5:18
Hung Up	Madonna	Confessions On A Dance	2005	Pop	5:37
So What (Main Version)	P!nk	Funhouse	2009	Pop	3:35

Microsoft’s Web-Store

Microsoft Store Products Support Search Microsoft.com 0

Store Devices Software & Apps Games & Entertainment Deals More Store Locations

Students save 10% on new Surface Pro. SHOP NOW >

BUNDLE AND SAVE

New Surface Pro i5 128GB + Type Cover

★★★★½ (4)

Build your bundle

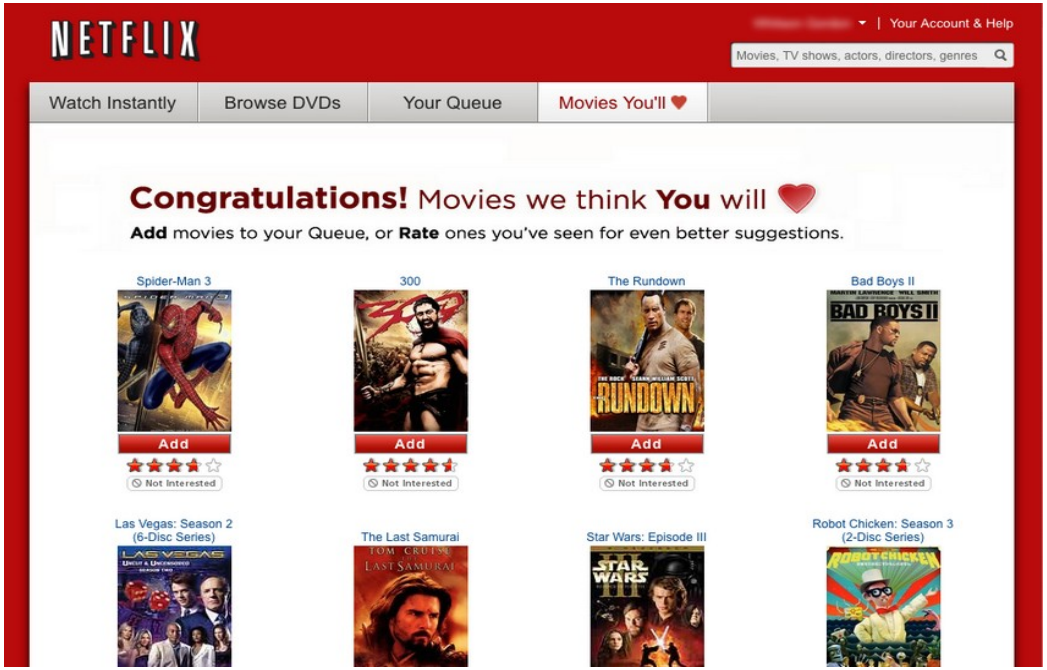
Free shipping. Free returns.

Description

- Save $150
- Intel Core i5, 128GB SSD, 4GB RAM
- Choice of Surface Type Cover (Burgundy, Cobalt Blue, or Platinum)
- A best-in-class laptop, with the versatility of a studio and tablet
- Powerful Intel Core processor
- The fastest startup and resume of any Surface Pro yet
- Up to 13.5 hours of battery life[2]
- Surface Pen sold separately

A DECADE AGO…

A Decade Ago… The Netflix Prize



The goal: 10% improvement in RMSE over Netflix's Cinematch

$$RMSE=\sqrt{1/n \sum_{i=1}^{n} (y_i - y_i)^2}$$

It took tens of thousands of participants over 2 years….

Microsoft

The Problem with Ratings

- They do not exist!

- Missing items not at random
  Collaborative Filtering and the Missing at Random Assumption

  B. M. Marlin, R. S. Zemel, S. Roweis, M. Slaney

- Ratings are fuzzy and influenced by the order of items
  Rating vs. Preference: A comparative study of self-reporting
  G. N. Yannakakis, J. Hallam

- Learning ratings is very different from personalization!
  Yahoo! Music Recommendations: Modeling Music Ratings with Temporal Dynamics and Taxonomy
  Gideon Dror, Noam Koenigstein and Yehuda Koren

IF NOT RMSE THEN WHAT?

Implicit Feedback and Ranking

- Collaborative Filtering for Implicit Feedback Datasets
  Y. Hu, Y. Koren, C. Volinsky

- Implicit-to-Explicit Ordinal Logistic Regression
  D. Parra, A. Karatzoglou, X. Amatriain, I. Yavuz

- BPR - Bayesian Personalized Ranking
  S. Rendle, C. Freudenthaler, Z. Gantner, and L. S. Thieme

- RankALS – Alternating Least Squares for Personalized Ranking
  G. Takacs, D. Tikk

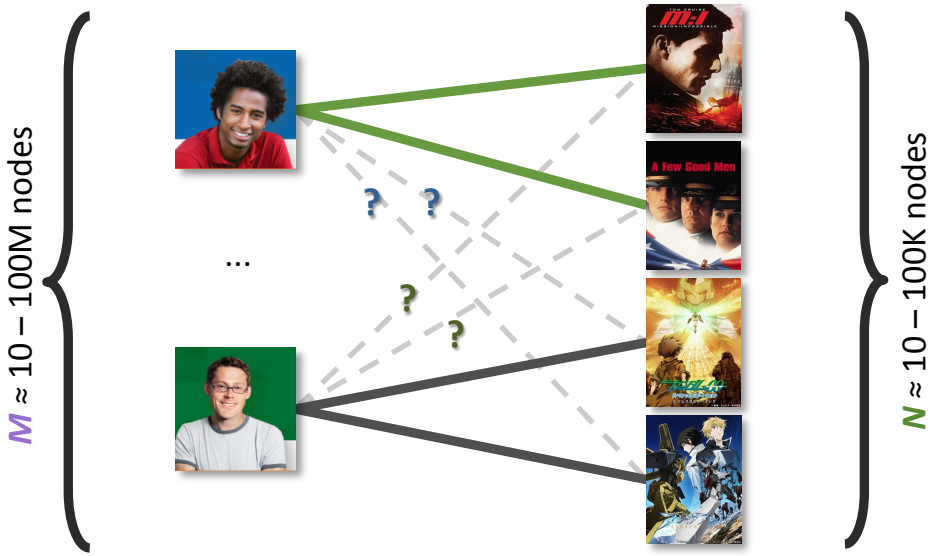
- CLiMF – Reciprocal Rank Optimization
  Y Shi, A. Karatzoglou, L. Baltrunas, M. Larson, N. Oliver, A. Hanjalic

ONE-CLASS COLLABORATIVE FILTERING WITH RANDOM GRAPHS

Ulrich Paquet and **Noam Koenigstein**

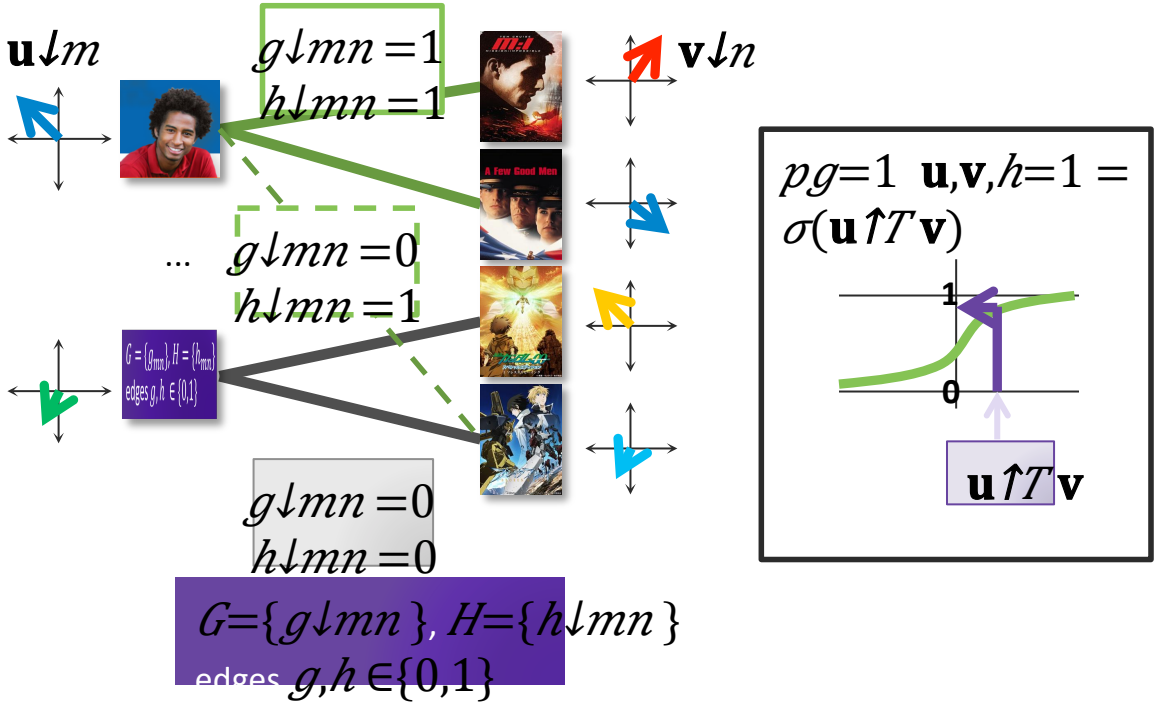
International World Wide Web Conference (WWW'13), May 2013, Rio de Janeiro, Brazil.

Problem Formulation



Bipartite graph → We care about ? = *p*(***link***)

The Hidden Graph



BESIDES FEEDBACK: COLD START, META-DATA, HYBRID, CONTEXTUAL…

XBOX MOVIES RECOMMENDATIONS: VARIATIONAL BAYES MATRIX FACTORIZATION WITH EMBEDDED FEATURE SELECTION

Noam Koenigstein and **Ulrich Paquet**

ACM Conference on Recommender Systems (RecSys'13), October 2013, *Hong Kong, China.*

Microsoft

Movie Features (tags)

Harry Potter and the Philosopher's Stone

Categories:

- Plot
- Mood
- Audience
- Time Period

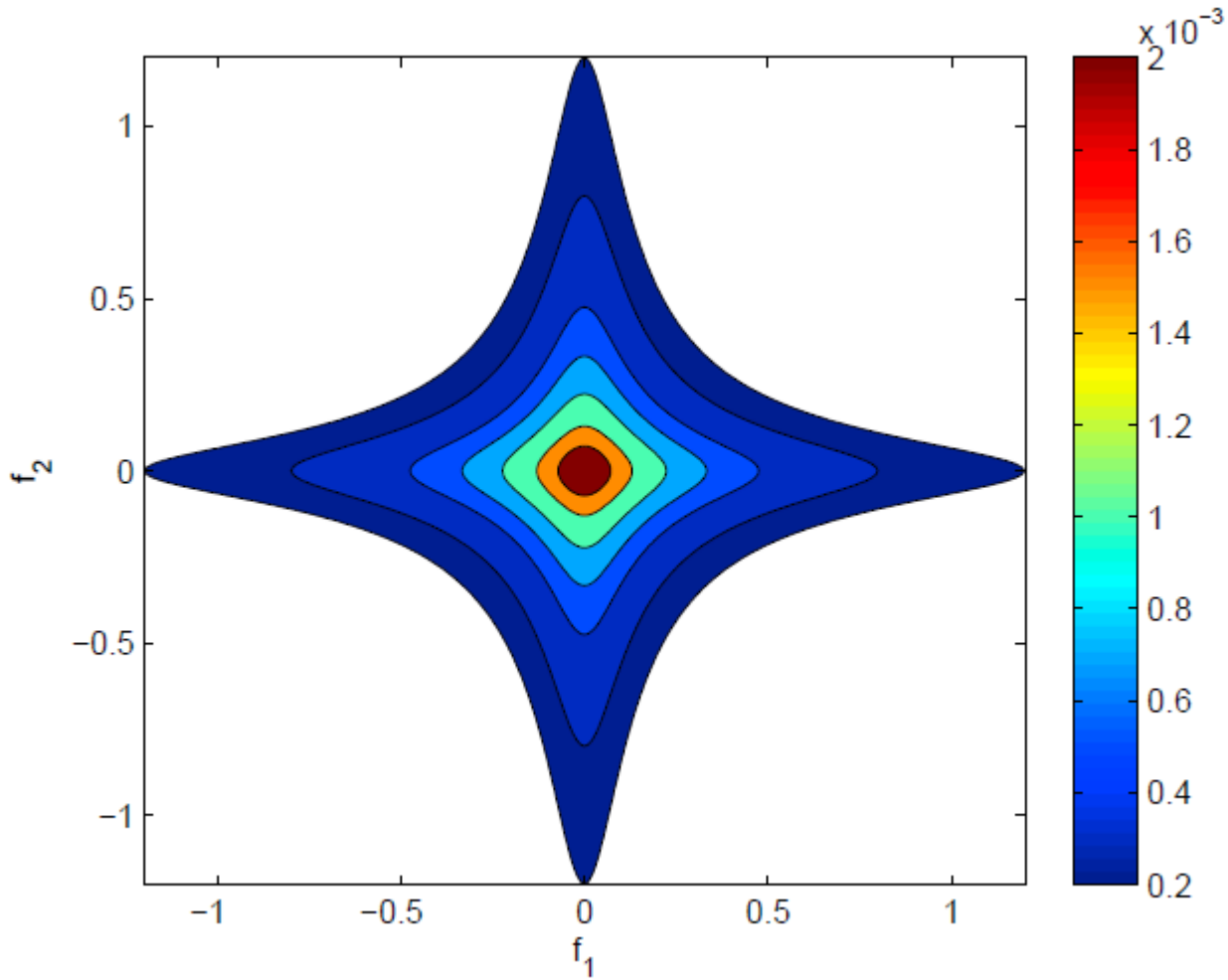
- Imaginary
- Wizards and Magicians
- Best Friends

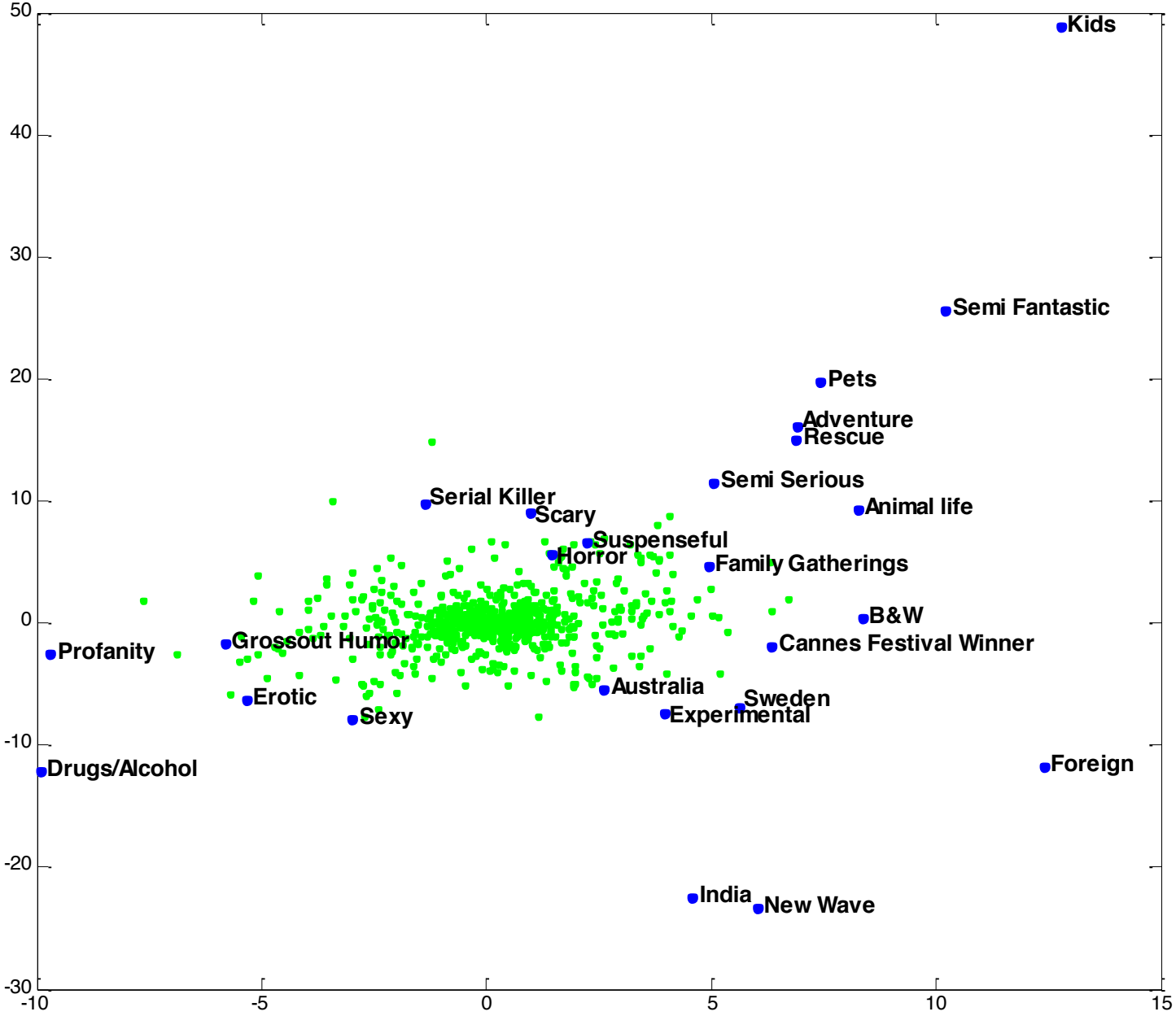
- Exciting
- Humorous
- Danger

- Kids
- Teens

- Contemporary
- 21st Century

$p(f_1, f_2 \mid \alpha=0.01, \beta=0.01)$




Kids
Semi Fantastic
Pets
Adventure
Rescue
Semi Serious
Serial Killer
Scary
Animal life
Suspenseful
Horror
Family Gatherings
B&W
Grossout Humor
Cannes Festival Winner
Profanity
Australia
Erotic
Sweden
Sexy
Experimental
Foreign
Drugs/Alcohol
India
New Wave
50
40
30
20
10
0
-10
-20
-30
-10
-5
0
5
10
15

GROOVE RADIO: A BAYESIAN HIERARCHICAL MODEL FOR PERSONALIZED PLAYLIST GENERATION

Shay Ben-Elazar, Gal Lavee, Noam Koenigstein, Oren Barkan, Hilik Berezin, Ulrich Paquet, Tal Zaccai

ACM Conference on Web Search and Data Mining (WSDM'17), Cambridge UK, February 2017.

Microsoft

THE GAP BETWEEN COLLABORATIVE FILTERING RESEARCH AND REAL WORLD RECOMMENDATIONS

The Gap Between Collaborative Filtering and Real Recommenders

- Diversity vs. accuracy - tradeoff??
- Popularity vs. personalization
- Item fatigue / freshness – repeating items
- Serendipity – when and how much to "surprise" the user
- List Recommendations / page optimization
- Predicting the future vs. influencing the user
- Metrics and Evaluation

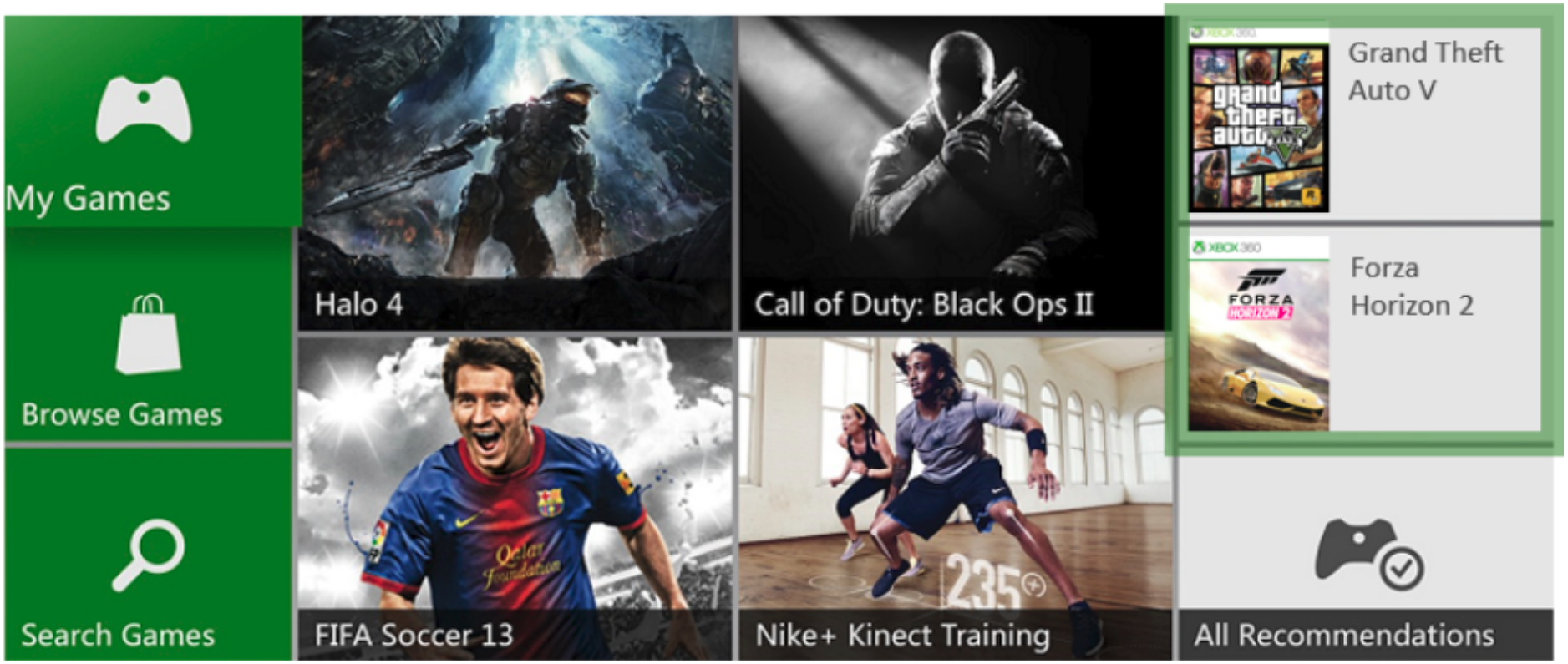
The Salesperson Analogy

BEYOND COLLABORATIVE FILTERING: THE LIST RECOMMENDATION PROBLEM

Oren Sar Shalom, Noam Koenigstein, Ulrich Paquet, Hastagiri P. Vanchinathan

International World Wide Web Conference (WWW'16), April 2016, Montreal, Canada.

List Recommendations in Xbox 360



Conclusions

- There is still a gap between most CF models and the actual goal of recommender systems

- Learning individual user-item tuples or ranking preferences is problematic because:
  - Can't handle the diversity vs. accuracy "tradeoff"
  - List recommendations / Page optimization

- Learning to predict future events from historical data is insufficient because:
  - Can't handle balancing popularity and personalization
  - Freshness / Item Fatigue
  - Serendipity

- RL alone is not the ultimate solution because:
  - The abundance of implicit data
  - Representing the "taste space"

- Offline evaluation metrics are insufficient
  - They measure our ability to predict the future but not our ability to change it (influence the user)

- <u>Bottom line</u>: We still have a lot to work in the RecSys community!

Thank You!

Microsoft

We are looking for postdocs in Israel!!!

Interested?
Find me during the coffee break….