



Recommender Systems

Intelligent User Interfaces

Professor Charles Rich
Computer Science Department
rich@wpi.edu

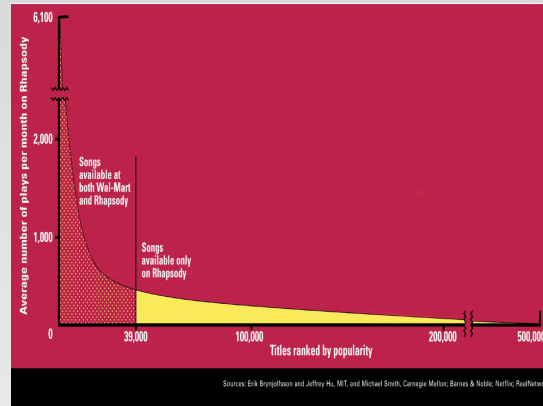
(Based in part on lecture by Rajaraman & Ullman, Stanford U., CS 345, 2007)

Readings

- Pazzani & Billsus, Content-Based Recommendation Systems, LNCS 2007
- Schafer et al, Collaborative Filtering Recommender Systems, LNCS 2007

Motivation

- The “*long tail*”
 - people not just interested in most popular items
 - a lot of area (market) under tail of the curve



Motivation (cont'd)

- In traditional *shelf-based* retailer, it was not possible to stock or advertise all items
- For *internet stores*, the cost of offering the whole tail dramatically less (approaching zero)
- For consumer, however, *more choice* necessitates better *filters*

Types of Recommender Systems

- Editorial
 - CNET reviews
 - TravelAdvisor
- Simple aggregates
 - Top 10 lists
 - Recent uploads
- **Tailored to individual users**
 - Amazon
 - NetFlix

Basic Concept -- Rating Matrix

| | King Kong | LOTR | The Matrix | National Treasure |
|-------|-----------|------|------------|-------------------|
| Alice | 1.0 | | 0.2 | |
| Bob | | 0.5 | | 0.3 |
| Carol | 0.2 | | 1.0 | |
| David | | | | 0.4 |

Key Problems

| | | | |
|-----|-----|-----|-----|
| 1.0 | | 0.2 | |
| | 0.5 | | 0.3 |
| 0.2 | | 1.0 | |
| | | | 0.4 |

1. How to *gather* “known” values
2. How to *extrapolate* unknown values from known values
 - sparse matrix; most people have not rated most items
 - mainly interested in high unknown values
 - many different extrapolation algorithms

Gathering Ratings

- Explicit approach
 - ask people to rate items
 - doesn't work well in practice--people can't be bothered
- Implicit approach
 - learn ratings from user actions (e.g., purchase implies high rating)
 - but how do you learn low ratings?

Extrapolation Approaches

- Content-based
 - recommend *similar items* to previous items rated highly by *same user*
 - e.g., movie with same actor(s), director, genre...
 - limitations:
 - finding appropriate features for similarity
 - over-specialization (never broadens/surprises user)
 - no information for new users

Extrapolation Approaches (cont'd)

- Collaborative (Filtering)
 - find set of other *similar users* based on their ratings of items
 - estimate user's rating of item based on other users' ratings of *same item*
 - limitations:
 - no information for new items
 - privacy issues

Hybrid Approach

- Add content-based methods to collaborative filtering
 - content-based approach handles new items
 - demographics to handle new users

Readings

- Pazzani & Billsus, Content-Based Recommendation Systems, LNCS 2007
 - survey/review paper
 - content-based recommenders
- Schafer et al, Collaborative Filtering Recommender Systems, LNCS 2007
 - survey/review paper
 - collaborative filtering