DiRec: Diversified Recommendations for Semantic-less Collaborative Filtering

**Rubi Boim**  
**Tova Milo**  
**Slava Novgorodov**

**Goal**
Diversify recommendations of CF systems

**Top-K relevance** (default)

A more Diverse and Wider list with “Zoom-in”

**Collaborative Filtering**
Predicts ratings based on how similar users rated similar items

- User 𝑈 rated item 𝑖 “4 stars”
- Each item is viewed as a vector of ratings
- Pearson’s correlation measure the distances
- Naïve approach presets the top-

**Difficulties**
Measuring diversity

- Semantic-less environment
- When exists, choosing the right features

Balancing ranking and diversity

- Previous solutions use threshold/weights
- Problematic to choose (context dependent)
- May need to be refined when data changes

**Priority Medoids**
A generalization of the standard medoids:

- Minimize the distances between the items and their medoids
- Medoids must have higher priorities than their cluster members

- Proven to be NP-Hard

**Approx. using Priority Cover-Tree**
A generalization of the standard cover-tree:

- Nesting
- Separation (all nodes are at least 1/2)
- Covering (children are within 1/2)

- Priority (children have lower rating)

**Experimental Results**
Semantic (sequel) Diversity  
Running Time  
DiRec Screenshots