

CSE 8803RS: Recommendation Systems

Lecture 25: Group Recommendation

Hongyuan Zha

School of Computational Science & Engineering
College of Computing
Georgia Institute of Technology

Group Recommendation

Recommended items are for group of users rather than a single user

- Key difference
 - a single user: possible temporal variation of preferences
 - a group of users: variation of preferences
- Problem formulation
 - trying to satisfy, *as much as possible*, the individual preferences of all the group's members
- Applications
 - web/news pages: group modeling in public space
 - tourism/vacation package
 - music
 - TV programs and movies

Group Recommendation: Evaluation Metrics

Comparing the generated recommendations for a group with the true preferences of the individual members

- User studies
 - Acquire the users' individual evaluations and aggregate
 - Conduct group a joint evaluation of the group: emotional contagion
- Off-line evaluations: using traditional CF data sets
 - generate synthetic groups
 - evaluate group recommendations simultaneously against the preferences of all the users in the group

Group Recommendation: Evaluation Metrics

Comparing the generated recommendations for a group with the true preferences of the individual members

- User studies
 - Acquire the users' individual evaluations and aggregate
 - Conduct group a joint evaluation of the group: emotional contagion
- Off-line evaluations: using traditional CF data sets
 - generate synthetic groups
 - evaluate group recommendations simultaneously against the preferences of all the users in the group

Group Recommendation: Two Classes of Methods

- Construct a joint user profile for all the users in the group
 - performing a recommendation for this artificial user represented by the group profile
- Recommendations for each individual member
 - aggregation into a group recommendation
 - existing user studies regarding methods used for aggregating individual recommendations: average and least misery
 - main methodology: rank aggregation

Group Recommendation: Two Classes of Methods

- Construct a joint user profile for all the users in the group
 - performing a recommendation for this artificial user represented by the group profile
- Recommendations for each individual member
 - aggregation into a group recommendation
 - existing user studies regarding methods used for aggregating individual recommendations: average and least misery
 - main methodology: rank aggregation

Rank Aggregation

- Kendall tau distance: between two ranked lists
 - number of pairwise disagreements between two lists
 - usually normalized by $n(n - 1)/2$
- Kemeny optimal aggregation of a set of ranking lists
 - ranking list that minimizes the average Kendall tau distances
 - NP-hard to compute

Rank Aggregation with Scoring

Assume g the set of ranking lists of N items

— $s_u(i)$ is the score (predicted value) for user u

- Each item has a score which generates the ranking list
 - Average aggregation

$$s(i) = \frac{1}{|g|} \sum_{u \in g} s_u(i)$$

then generate ranking using $\{s(i)\}$

— Least Misery

$$s(i) = \min_{u \in g} s_u(i)$$

Rank Aggregation w/o Scoring

Assume $g = \{\sigma_u\}$ the set of ranking lists of N items

— each ranking list σ_u a permutation of $\{1, \dots, N\}$

— $\sigma(i)$ is the position of item i in the list

- Spearman footrule distance

$$d(\sigma_1, \sigma_2) = \sum_i |\sigma_1(i) - \sigma_2(i)|$$

- Spearman footrule aggregation: ranking list that minimizes the average Spearman footrule distances
- Equivalent to finding a minimum cost perfect matching in a particular bipartite graph
- 2-approximation of Kemeny optimal aggregation

Rank Aggregation w/o Scoring

- Borda count: $score_u(i)$ number of items below item i in ranking list σ_u
- Borda count aggregation:

$$s(i) = \frac{1}{|g|} \sum_{u \in g} s_u(i)$$

- 5-approximation of Kemeny optimal aggregation

DCG (Discounted Cumulated Gain)

- Relevance labels converted to **gain values**
 - Perfect \Rightarrow 20, Excellent \Rightarrow 10, Good \Rightarrow 5, Fair \Rightarrow 1, Bad \Rightarrow 0.
- List of ranked documents with labels \Rightarrow gain vector G .
- DCG for the ranked list is computed as

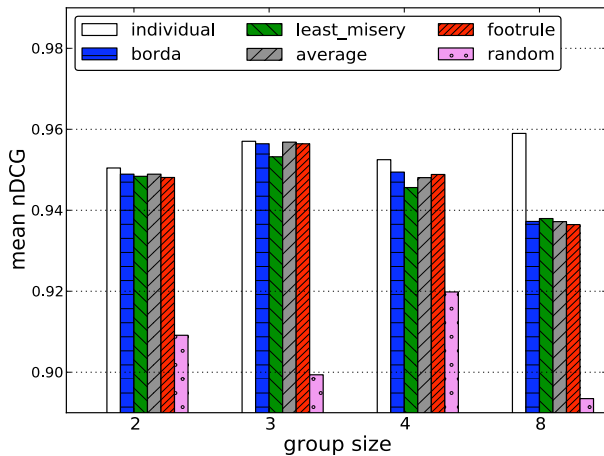
$$\text{DCG}_K = \sum_{i=1}^K \frac{G(i)}{\log_2(1+i)},$$

for $K = 10 \Rightarrow$ DCG-10. Often, we use **normalized** DCGs.

DCG (Discounted Cumulated Gain)

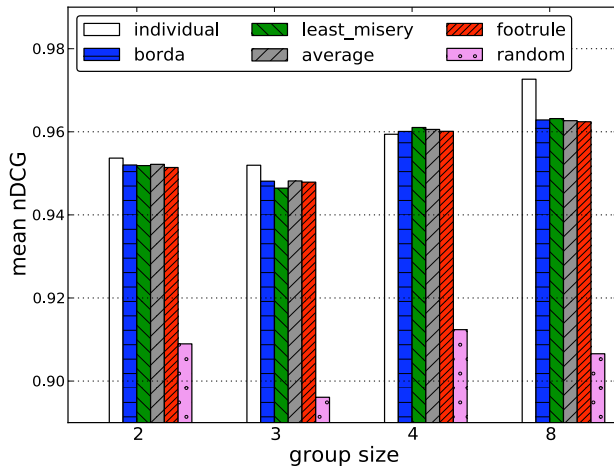
- Not all recommended items rated by the users
- $[1, 4, 5, 8, 3, 7, 6, 2, 9]$ is a ranked list of recommendations for a group
- user u test set consists of eight items $\{1, 4, 7, 8, 9, 12, 14, 20\}$
- nDCG is computed on the ranked list $[1, 4, 8, 7, 9]$

Random Groups



(a) random groups

High Similarity Groups



(b) high similarity groups