PROJECT PROPOSAL OF COMP621U

GROUP MEMBER: WAN Pengfei, PANG Chao, SUN Wenxiu

TOPIC: KDDCUP 2011 Task1

DATASET: Yahoo Music

SUGGESTED APPROACH:

Clustering-Based Evolutionary Collaborative Filtering

1. Coarse clustering using genre/album/artist for the whole dataset (songs).
2. Within each music group, apply K-means/spectral clustering [5] to further form sub-clusters of songs using user-item interactions.
3. Different weights will be given to different sub-clusters based on user's degree of preference (larger weight for lower and higher ratings).
4. Apply Neighborhood-Model CF to predict the rating (the basic idea is from Nathan's paper [1]):

$$\widehat{r_{u,i}}(t) = \overline{r_u}(t) + \frac{\sum_{j \in \mathcal{N}_i^t \cap \mathcal{I}_u^t} S_{i,j}(t) \cdot f_{u,j}^{\beta}(t) \cdot r_{u,j}}{\sum_{j \in \mathcal{N}_i^t \cap \mathcal{I}_u^t} S_{i,j}(t) \cdot f_{u,j}^{\beta}(t)}$$

Where $r_{u,j}$ is the observed rating by user u to item j.

$\overline{r_u}(t)$ is the overall means rating score.

$\mathcal{N}_i^t$ is the set of similar items with item i. (sub-cluster)

$\mathcal{I}_u^t$ is the set of items u has rated before time t.

$S_{i,j}(t)$ is the similarity of item I and j.

$f_{u,j}^{\beta}(t)$ is an exponential temporal relevance function.

5. Note that weighted averaging is applied on all members in the sub-cluster item i belongs to (zero weight for items haven't rated by u), which makes full use of the ratings of items sharing abundant properties with i.

   We believe this scheme is more appropriate than simply adopting Pearson correlation coefficient between items which may actually share low semantic correlation.

EVALUATION METRICS:

RMSE (Root Mean Square Error)

RELATED WORKS:

[1] Nathan Liu, et al, Online Evolutionary Collaborative Filtering. RecSys'10

[2] Y. Koren, Collaborative Filtering with Temporal Dynamics, KDD'09

[3] H. Ungar, et al. Clustering Methods for Collaborative Filtering, AAAI Workshop on Recommendation Systems'98

[4] Z. Huang, et al, A Comparison of Collaborative Filtering Recommendation Algorithms for E-commerce, IEEE Intelligent Systems'04

[5] U. Luxburg, A Tutorial on Spectral Clustering