

CCCF: Improving Collaborative Filtering via Scalable User-Item Co-Clustering

Yao Wu^{1*}, Xudong Liu^{2*}, Min Xie³
Martin Ester¹ and Qing Yang²

*Equal Contribution

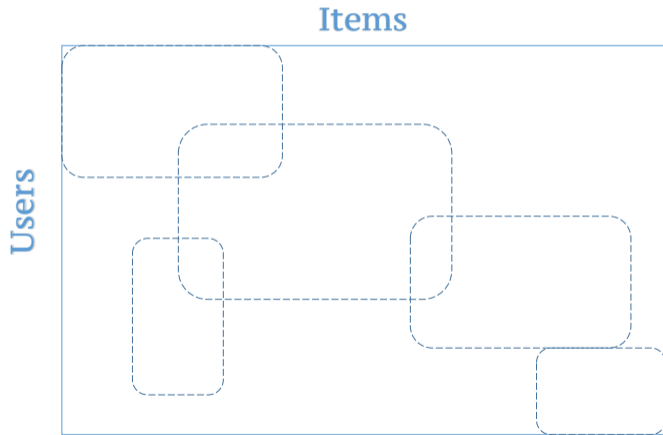
¹ Simon Fraser University

² Institute of Automation, Chinese Academy of Sciences

³ @WalmartLabs

February 23, 2016

Motivation



Overview of Proposed Methodology

- 1 Cluster users and items in to subgroups.
 - We present a probabilistic method to learn overlapping co-clusters.
 - We present a scalable inference algorithm using stochastic MCMC.
- 2 Collaborative Filtering models in subgroups.
 - Any recommender model can be used.
 - Easily parallelizable.
- 3 Aggregate recommendations from subgroups.
 - We discuss several strategies for recommendation aggregation.

CCCF v.s. Basic CF method

CCCF improves CF method by co-clustering.

Table: CCCF vs. No Co-Clustering on the Netflix data

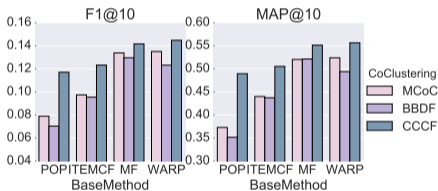
| | POP | | ITEMCF | | MF | | WARP | |
|--------|-------|-------|--------|-------|-------|-------|-------|-------|
| | NONE | CCCF | NONE | CCCF | NONE | CCCF | NONE | CCCF |
| P@10 | 0.189 | 0.283 | 0.308 | 0.312 | 0.330 | 0.354 | 0.347 | 0.364 |
| R@10 | 0.047 | 0.075 | 0.085 | 0.085 | 0.091 | 0.098 | 0.095 | 0.100 |
| F1@10 | 0.070 | 0.109 | 0.123 | 0.123 | 0.132 | 0.142 | 0.138 | 0.145 |
| MAP@10 | 0.348 | 0.471 | 0.493 | 0.505 | 0.528 | 0.551 | 0.538 | 0.556 |

Table: CCCF vs. No Co-Clustering on the Yelp data

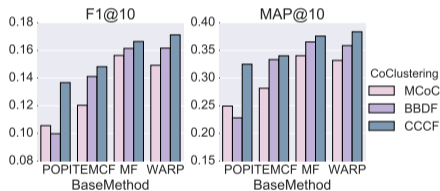
| | POP | | ITEMCF | | MF | | WARP | |
|--------|-------|-------|--------|-------|-------|-------|-------|-------|
| | NONE | CCCF | NONE | CCCF | NONE | CCCF | NONE | CCCF |
| P@10 | 0.018 | 0.046 | 0.048 | 0.057 | 0.043 | 0.046 | 0.047 | 0.056 |
| R@10 | 0.024 | 0.058 | 0.062 | 0.072 | 0.055 | 0.058 | 0.058 | 0.068 |
| F1@10 | 0.018 | 0.046 | 0.048 | 0.058 | 0.044 | 0.046 | 0.047 | 0.055 |
| MAP@10 | 0.050 | 0.124 | 0.142 | 0.159 | 0.124 | 0.131 | 0.137 | 0.152 |

CCCF v.s. Other Co-Clustering Methods

CCCF outperforms other co-clustering methods.



(a) Results on the Netlix data



(b) Results on the MovieLens data

Figure: Comparison with other two co-clustering methods on the three data sets.

Qualitative Analysis on Yelp data: Items with largest weight in subgroups

| Sample Cluster 1 | | | Sample Cluster 2 | | | Sample Cluster 3 | | |
|---------------------------|---------|--------------|----------------------------|-----------|-------------|--|-----------|------------------|
| Business Name | City | Description | Business Name | City | Description | Business Name | City | Description |
| Gallo Blanco | Phoenix | Restaurants | Bread and Butter | Henderson | Restaurants | Las Vegas North Premium Outlets | Las Vegas | Shopping |
| Lux | Phoenix | Restaurants | Snow Ono Shave Ice | Las Vegas | Restaurants | Orleans Hotel & Casino | Las Vegas | Casinos & Hotels |
| Postino Central | Phoenix | Restaurants | Buldogis Gourmet Hot Dogs | Las Vegas | Restaurants | Fremont Street Experience | Las Vegas | Entertainment |
| Maizie's Cafe & Bistro | Phoenix | Restaurants | Sweet Tomatoes | Las Vegas | Restaurants | Conservatory & Botanical Garden | Las Vegas | Entertainment |
| Cherryblossom Noodle Cafe | Phoenix | Restaurants | Strip N Dip Chicken Strips | Las Vegas | Restaurants | Planet Hollywood Las Vegas Resort & Casino | Las Vegas | Casinos & Hotels |
| Pizza a Metro | Phoenix | Restaurants | Patisserie Manon | Las Vegas | Restaurants | Flamingo Las Vegas Hotel & Casino | Las Vegas | Casinos & Hotels |
| SideBar | Phoenix | Bars | Island Flavor | Las Vegas | Restaurants | New York - New York | Las Vegas | Casinos & Hotels |
| Churn | Phoenix | Coffee & Tea | Japanese Curry Zen | Las Vegas | Restaurants | Miracle Mile Shops | Las Vegas | Shopping |
| Carly's Bistro | Phoenix | Restaurants | Slidin' Thru | Las Vegas | Restaurants | Palms Casino Resort | Las Vegas | Casinos & Hotels |

Conclusion

- We propose a new co-clustering method CCCF and a scalable inference algorithm.
- CCCF improves the performance of several popular Collaborative Filtering methods, and outperforms other co-clustering methods.

Thanks!

yaowu.co