

# A Study of Methods for Normalizing User Ratings in Collaborative Filtering

Rong Jin

Dept. of Computer Science and Engineering  
Michigan State University  
East Lansing, MI 48824  
rongjin@cse.msu.edu

Luo Si

School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA 15213  
lsi@cs.cmu.edu

## ABSTRACT

The goal of collaborative filtering is to make recommendations for a test user by utilizing the rating information of users who share interests similar to the test user. Because ratings are determined not only by user interests but also the rating habits of users, it is important to normalize ratings of different users to the same scale. In this paper, we compare two different normalization strategies for user ratings, namely the Gaussian normalization method and the decoupling normalization method. Particularly, we incorporated these two rating normalization methods into two collaborative filtering algorithms, and evaluated their effectiveness on the EachMovie dataset. The experiment results have shown that the decoupling method for rating normalization is more effective than the Gaussian normalization method in improving the performance of collaborative filtering algorithms.

## Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Information Search and retrieval—*Information Filtering*

## General Terms

Algorithms, Experimentation

## Keywords

Collaborative filtering, rating normalization

## 1. INTRODUCTION

The task of collaborative filtering is to predict the utilities of items for a particular user based on the rating information for the same set of items given by many other users. In general, methods for collaborative filtering can be categorized into two classes: memory-based algorithms and model-based algorithms [1]. Memory-based algorithms predicate the rating for a test user on an item by averaging the ratings of the training users who share the similar interests as the test user. This category includes the Pearson- correlation coefficient approach [2] and the vector space similarity approach [1]. Model-based approaches group different training users into a small number of classes based on their rating patterns. The rating of a test user on a particular item is predicated as the rating of the user class that the test user fits in best.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '04, Month 1–2, 2004, City, State, Country.  
Copyright 2004 ACM 1-58113-000-0/00/0004...\$5.00.

Algorithms such as the aspect model [3] and the flexible mixture model [4] belong to the category.

One common problem with many collaborative filtering models is that different users adopt different criteria in determining the ratings of items. As a result, users with similar interests can still assign different ratings for the same item. For example, some users are more “tolerant” and therefore their ratings of items tend to be higher than other users even though they may share very similar tastes of the items. In an early study of collaborative filtering by Resnick et al., this problem is addressed through a normalization method that essentially normalizes the ratings of individual users to Gaussian distributions [2]. This Gaussian normalization method has been successfully applied to not only the memory-based algorithms but also the model-based algorithm for collaborative filtering. In the recent paper [5], a different rating normalization method is proposed which converts the rating of an item into likelihood for the item to be favored by the user. The conversion is computed based on the rating distribution of each individual user. Although there is positive evidence for both normalization methods for collaborative filtering, there have not been any studies on the comparison of these two methods. In this paper, the two normalization methods are compared for the same set of collaborative filtering algorithms and under the same configurations.

## 2. Related Work

Two methods for normalizing ratings are introduced here: the Gaussian normalization method and the decoupling normalization method.

**Gaussian normalization method.** This method is first proposed in [2]. It considers two factors that can lead to the rating variance among users with similar interests.

1) *Shift of average ratings.* This problem is related to the fact that some users are more “tolerant” and tend to give higher ratings than others. As a result, the average ratings for those “tolerant” users are usually higher than the rigorous users. This factor can be accounted by subtracting ratings of each user from the average rating of the user.

2) *Different rating scales.* This problem is related to the fact that “conservative” users tend to assign items to a narrow range of rating categories whereas “liberal” users tend to assign items to a wide range of rating categories. To account for this factor, the ratings of each user are divided by the variance in his ratings.

Combining the above two ideas, for each user, the normalized rating for item  $x$  by user  $y$   $\hat{R}_y(x)$  is computed as:

$$\hat{R}_y(x) = \frac{R_y(x) - \bar{R}_y}{\sqrt{\sum_x (R_y(x) - \bar{R}_y)^2}} \quad (1)$$

where  $R_y(x)$  stands for the rating for item  $x$  by user  $y$ , and  $\bar{R}_y$  stands for the average rating for user  $y$ . Since this normalization method essentially normalizes the rating distribution of a user to a Gaussian distribution, we refer to it as the ‘‘Gaussian normalization method’’.

**Decoupling normalization method.** The approach converts the rating of an item by a user into a probability for the item to be favored by the user. The probabilistic measurement is determined based on the following two assumptions:

- 1) When a large portion of items are rated by a user as no more than category  $r$ , items in the rating category  $r$  are likely to be favored by the user.
- 2) When more items are rated as category  $r$ , it becomes less likely for the user to favor items in the category  $r$ .

Based on the above two assumptions, a special formula, named halfway accumulative distribution, was proposed in [5] to convert the rating of an item into a probability that measures how likely the item will be favored by the user. The formal expression is:

$$p_y(\text{R is favored}) = p_y(\text{Rating} \leq r) - p_y(\text{Rating} = r)/2 \quad (2)$$

where  $p_y(\text{Rating} \leq r)$  stands for the percentage of items that are rated no more than category  $r$ , and  $p_y(\text{Rating} = r)$  stands for the percentage of items that are rated as  $r$ . For later reference, we call this method ‘‘decoupling normalization method’’.

### 3. Experiment

To examine the effectiveness of the two methods for normalizing ratings, we will incorporate them into different collaborative filtering algorithms. More specifically, we first apply the two methods to normalize user ratings and the normalized ratings are then used as the inputs to the collaborative filtering algorithm. Two collaborative filtering algorithms are selected for this experiment, which are representatives of the memory-based approaches and the model-based approaches, respectively. They are: Pearson Correlation Coefficient (PCC) method and Flexible Mixture Model (FMM).

EachMovie dataset is used as the testbed, which contains 1648 different movies. 2000 users were randomly selected from those who have rated more than 40 movies. The average number of rated movies for each user is 129. To compare the two normalization methods in full spectrum, we vary the number of users in the training database and the number of rated items that are provided by the test user, as what have been done in [1]. We choose the size of training database to be 200 and 400, and the number of rated items by test users to be 5, 10 and 20. In all cases, the rest of the users are used for testing. Mean Absolute Error (MAE) [1] is used for evaluation. The results for the two collaborative filtering algorithms using different normalization methods are listed in Table 1 and 2, respectively.

Compared to the Gaussian normalization method, the decoupling normalization method achieves substantially larger improvement over the baseline model in which no normalization method is used. The advantage of decoupling normalization method against the Gaussian method is more noticeable for the memory-based

**Table 1:** MAE results for Pearson Correlation Coefficient method for no normalization, Gaussian normalization, and decoupling normalization method.

	Norm.	5 Items	10 Items	20 Items
200 Training User	No	1.22	1.16	1.13
	Gaussian	1.18	1.14	1.11
	Decouple	1.04	1.01	0.98
400 Training Users	No	1.22	1.15	1.12
	Gaussian	1.18	1.14	1.11
	Decouple	1.03	1.00	0.97

**Table 2:** MAE results for Flexible Mixture Model for no normalization, Gaussian normalization, and decoupling normalization method.

	Norm.	5 Items	10 Items	20 Items
200 Training User	No	1.08	1.05	1.04
	Gaussian	1.10	1.04	1.02
	Decouple	1.06	1.02	1.00
400 Training Users	No	1.06	1.04	1.03
	Gaussian	1.08	1.03	1.00
	Decouple	1.04	1.01	0.99

approach than the model-based approach. This phenomenon can be attributed to the reason that the flexible mixture model (FMM) clusters users according to their rating patterns. By averaging the ratings of different users among the same cluster, the variance in ratings is actually reduced. As a result, the normalization methods are less effective with the model-based approaches for collaborative filtering than the memory-based approaches.

### 4. Conclusion

In this paper, we compared the Gaussian normalization method and the decoupling normalization method for collaborative filtering. Experiments with two collaborative filtering methods have shown that the decoupling normalization method is more effective in improving prediction accuracy for collaborative filtering than the Gaussian normalization method.

### 5. REFERENCES

- [1] J. S. Breese, D. Heckerman and C. Kadie, Empirical Analysis of Predictive Algorithms for Collaborative Filtering, Proc. of the 14<sup>th</sup> Conference on Uncertainty in Artificial Intelligence (UAI), 1998.
- [2] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom and J. Riedl, Grouplens: An Open Architecture for Collaborative Filtering of Netnews. In Proc. of the ACM 1994 Conference on Computer Supported Cooperative Work, pages 175-186.
- [3] T. Hofmann, Gaussian Latent Semantic Models for Collaborative Filtering, In the Proc. of the 26<sup>th</sup> Annual International ACM SIGIR Conference, 2003
- [4] L. Si and R. Jin, Flexible Mixture Model for Collaborative Filtering, In Proc. Of the 20<sup>th</sup> International Conference on Machine Learning (ICML 2003), 2003
- [5] R. Jin, L. Si, and C.X. Zhai, Collaborative Filtering with Decoupled Models for Preferences and Ratings, in Proc. Of the 12<sup>th</sup> Conference on Information and Knowledge Management (CIKM 2003), 2003