Correcting Popularity Bias by Enhancing Recommendation Neutrality

Toshihiro Kamishima, Shotaro Akaho, and Hideki Asoh
National Institute of Advanced Industrial Science and Technology (AIST)
AIST Tsukuba Central 2, Umezono 1-1-1, Tsukuba, Ibaraki, 305-8568 Japan
mail@kamishima.net, s.akaho@aist.go.jp, h.asoh@aist.go.jp

Jun Sakuma
University of Tsukuba
1-1-1 Tennodai, Tsukuba, 305-8577 Japan
jun@cs.tsukuba.ac.jp

ABSTRACT

In this paper, we attempt to correct a popularity bias, which is the tendency for popular items to be recommended more frequently, by enhancing recommendation neutrality. Recommendation neutrality involves excluding specified information from the prediction process of recommendation. This neutrality was formalized as the statistical independence between a recommendation result and the specified information, and we developed a recommendation algorithm that satisfies this independence constraint. We correct the popularity bias by enhancing neutrality with respect to information regarding whether candidate items are popular or not. We empirically show that a popularity bias in the predicted preference scores can be corrected.

Keywords

recommender system, neutrality, fairness, popularity bias, probabilistic matrix factorization, information theory

1. RECOMMENDATION NEUTRALITY AND POPULARITY BIAS

We proposed the notion of *recommendation neutrality* with respect to a specified viewpoint if no information about the viewpoint is exploited when generating the recommendation results [3]. If we use terms of information theory, this notion can be formalized as the condition that the mutual information between a recommendation result and a viewpoint is zero, and it further implies statistical independence between them. We developed *information-neutral recommender systems* (INRS) that predict users' preference scores while satisfying the constraint of statistical independence [3, 4]. This INRS could be useful for the avoidance of biased recommendation, fair treatment of content providers, or adherence to laws and regulations. In this paper, we use the proposed INRS to avoid a well-known *popularity bias*, which is the tendency for popular items to be recommended more frequently [1]. When users have no interest in the popularity of items and wish to ignore this information, they can obtain recommendations that are neutral with respect to the popularity of items by specifying the volume of their consumption as a viewpoint.

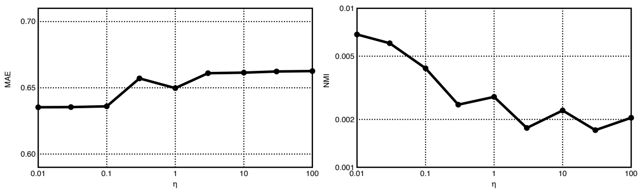
The popularity bias has previously been corrected by diversifying recommended items [5]. Specifically, instead of the most popular and preferred items, slightly less preferred and diverse kind of items are recommended. This diversification approach is different from our approach of enhancing recommendation neutrality. While diversity is a property of a set of recommendations, neutrality is a relation between recommendations and a specified viewpoint. Many notions of diversity have been proposed, but all of them target a set of recommendations; thus, it is impossible to correct a bias with a single recommendation. On the other hand, a single recommendation can be neutral in its prediction of ratings with respect to a specified viewpoint. This is useful, for example, when attaching a list of items with predicted ratings that match a user's query. Therefore, our INRS can be used for correcting the popularity bias in each predicted score.

2. EXPERIMENTS

We applied our INRS, `mean-match` [4], to show that our approach is effective in correcting a popularity bias. Simply speaking, this algorithm is a variant of the probabilistic matrix factorization model [6] that adopts a constraint term for enhancing neutrality.

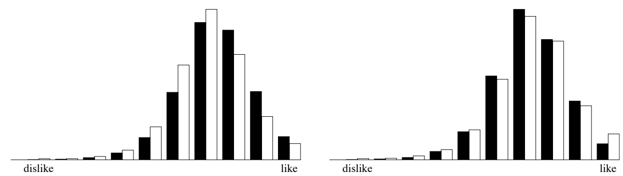
We evaluated our experimental results in terms of prediction errors and degree of neutrality. Prediction errors were measured by the mean absolute error (MAE). This index was defined as the mean of the absolute difference between the observed rating values and predicted rating values. A smaller value of this index indicates better prediction accuracy. To measure the degree of neutrality, we adopted normalized mutual information (NMI) [4]. The NMI is defined as mutual information between the predicted ratings and viewpoint values, normalized into the range $[0, 1]$. A smaller NMI indicates a higher level of neutrality. Note that the distribution of scores is modeled by a multinomial distribution after discretizing prediction scores. We performed a five-fold cross-validation procedure to obtain evaluation indices.

RecSys 2014 Poster Proceedings, October 6–10, 2014, Foster City, Silicon Valley, USA.



(a) Prediction error (MAE) (b) Degree of neutrality (NMI)

Figure 1: Changes in the accuracy and degree of neutrality accompanying an increase in the neutrality parameter



(a) standard (b) neutrality enhanced

Figure 2: Distribution of the predicted ratings for short-head and long-tail items

The data set was the Flixster data set[1] [2]. The total numbers of users and movies were 147,612 and 48,794, respectively, and the data set consisted of 8,196,077 ratings. Ratings are represented by a ten-point-scale whose domain is 0.5 to 5.0 in 0.5 increments. To correct a popularity bias, we adopted the popularity of items as a viewpoint. Candidate movies were first sorted by the number of users who rated the movie in a descending order, and a viewpoint represented whether or not a movie was in the top 1% of this list. We called the group of top 1% items the *short-head* items, and the group containing the rest the *long-tail* items.

Figure 1(a) shows the change of prediction errors measured by the MAE in a linear scale. Figure 1(b) shows the change in NMI in a logarithmic scale. The X-axes of these figures represent the values of a neutrality parameter, η, which balances the prediction of accuracy and neutrality. These parameters were changed from 0.01, at which the neutrality term was almost completely ignored, to 100, at which neutrality was strongly enhanced.

We first compared these with two baseline results. The MAE was 0.871 when the rating being offered was held constant at 3.61, which is the mean rating over all sample ratings in the training data. This approximately simulated the case of randomly recommending items, and can be considered the most unbiased and neutral recommendation. However, this prediction error was clearly worse than those in Figure 1(a). On the other hand, when the original probabilistic matrix factorization model was applied, the MAE was 0.652. Although the trade-off for enhancing neutrality generally worsened prediction accuracy, the errors in 1(a) were not significantly worse. This was very positive, indicating that prediction accuracies were not degraded even if a popularity bias was corrected.

We then observed the changes of MAE and NMI accompanying an increase in the neutrality parameter, η. Overall, the increase of MAEs as increase of η was not great. Turning to Figure 1(b), we see that recommendation neutrality was successfully enhanced. This means that predicted scores were less influenced by the factor of whether candidate items were short-head or long-tail. In summary, our INRS successfully corrected a popularity bias without seriously sacrificing prediction accuracy.

To illustrate the influence of correcting a popularity bias, Figure 2 shows the distributions of predicted ratings for short-head and long-tail items. Black and white bars show the distributions of ratings for short-head and long-tail items, respectively. In Figure 2(a), ratings are predicted by a standard recommendation algorithm, and short-head items are highly rated. After correcting the popularity bias ($\eta = 100$) as in Figure 2(b), the distributions of ratings for short-head and long-tail items become much closer; that is to say, the predicted ratings are less influenced by items' popularity. It follows from this figure that our INRS successfully corrected a popularity bias.

3. CONCLUSIONS

We corrected a popularity bias by enhancing recommendation neutrality and empirically showed the effectiveness of our approach. We plan to improve the efficiency of our information-neutral recommendation algorithm and to adopt a more sophisticated model for expressing popularity.

4. ACKNOWLEDGMENTS

We would like to thank for providing a data set for Dr. Mohsen Jamali. This work is supported by MEXT/JSPS KAKENHI Grant Number 16700157, 21500154, 24500194, and 25540094.

5. REFERENCES

[1] Ò. Celma and P. Cano. From hits to niches?: or how popular artists can bias music recommendation and discovery. In *Proc. of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition*, 2008.

[2] M. Jamali and M. Ester. A matrix factorization technique with trust propagation for recommendation in social networks. In *Proc. of the 4th ACM Conf. on Recommender Systems*, pages 135–142, 2010.

[3] T. Kamishima, S. Akaho, H. Asoh, and J. Sakuma. Enhancement of the neutrality in recommendation. In *Proc. of the 2nd Workshop on Human Decision Making in Recommender Systems*, pages 8–14, 2012.

[4] T. Kamishima, S. Akaho, H. Asoh, and J. Sakuma. Efficiency improvement of neutrality-enhanced recommendation. In *Proc. of the 3rd Workshop on Human Decision Making in Recommender Systems*, pages 1–8, 2013.

[5] M. Levy and K. Bosteels. Music recommendation and the long tail. In *WOMRAD 2010: RecSys 2010 Workshop on Music Recommendation and Discovery*, 2010.

[6] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems 20*, pages 1257–1264, 2008.

[1] http://www.sfu.ca/~sja25/datasets/