# The Application of Transfer Learning on E-Commerce Recommender Systems *

JiuHong TANG, ZhiHong ZHAO, Jia BEI, WeiQing WANG
State Key Laboratory for Novel Software Technology, Nanjing University
Software Institute, Nanjing University
Nanjing, China
zhaozh@software.nju.edu

**Abstract—Nowadays, recommender systems occupy an increasingly important position in people's lives. Recommender systems are widely applied in e-commerce websites, they discover users' potential consuming habits by analyzing their behaviors, and then recommend users with what they may purchase. However, recommender systems on e-commerce sites are facing the problem of data sparsity. Data sparsity may cause poor recommendations, thereby reducing users' shopping satisfaction. In order to alleviate this problem, we propose a new approach based on the idea that combines user-based collaborative filtering techniques with transfer learning. The method alleviates the data sparsity problem by transferring the knowledge learned from dense data set to sparse ones. We use the data from a glasses site as the dense data set and the data from an underware site as the sparse one, experiments are conducted for evaluating the proposed method in this paper. Results show that our method can alleviate the data sparsity problem and improve the effect of user based collaborative filtering method.**

**Keywords—E-commerce; Recommender System; User-based Collaborative Filter; Transfer Learning**

## I. INTRODUCTION

Recommender systems have been applied in many areas, e.g. e-commerce recommender system. E-commerce recommender systems help customers find merchandises they need. However, recommender systems often face the problem of data sparsity, and data sparsity may cause poor recommendations. Antonia Dattolo [1], Aleksandra Klasnia Milicevic [2], etc put out some methods that are insteading sparse rating data with dense tag data. Since tag data are sparse and not able to reflect users' potential consuming habits,e-commerce recommender systems generally only use order data as input data.

Currently, mainstream methods of alleviating data sparsity problem include using tag data instead of rating data, combining rating data with tag data, or fill in data through some ways. In [3] it presented a method that using topic data based on rating data to predict user's rates. In [4] it proposed a new approach of using tag data reflecting users' interests in products. In [5] it used clustering method that predicts user's rating by averaging other users' rating to a product in the same cluster. In [6] it suggested that we can fill in data with dimension reduction. In [7][8] it solvedata sparsity problem with transfer learning .

Since resources of some e-commerce recommender systems are more abundant in users, orders, clicks and so on, meanwhile resources of other ones are poor. In order to solve this problem-data sparsity, this paper proposes a novel approach based on the idea that combine user-based collaborative filtering with transfer learning. This method improves the accuracy of recommendation by transfer the knowledge learned from dense data sets to sparse ones.

The remainder of this paper is organized as follows is: In section 2 we review some related work, including e-commerce, e-commerce recommender systems and transfer learning. Section 3 describes our method applied to e-commerce recommender system based transfer learning. Experimental results and analysis are given in section 4,and we make an conclusion finally.

## II. RELATED WORK

Nowadays, e-commerce has become part of everyday life and people are familiar with e-commerce, like e-commerce website, telephone shopping, etc. According to different products or services sold from e- commerce websites, e-commerce websites can be divided into vertical e-commerce website and integrated e-commerce websites. Vertical e-commerce websites sell products generally belongs to the same category, e.g. glasses website only sell a variety of glasses and sunglasses. The integrated e-commerce websites sell various products across multiple areas, such as taobao.com, this site not only sell food, clothing, etc., also offers recharge services.

E-commerce recommender system in accordance with used techniques can be divided into three categories, content-based filtering, collaborative filtering and hybrid recommendation. Content-based filtering [9] [10] is based on information retrieval. Recommender system firstly analyse products in order to obtain a set of features, and then build products' feature vector. Afterwards recommender systems calculate the similarity between users and products, recommend users product according to the similarity. Besides content-based filtering can also use technologies related with machine learning recommended, such as clustering based on product characteristics. Collaborative filtering recommendation can be

978-1-4799-3219-1/13 $31.00 © 2013 IEEE
DOI 10.1109/WISA.2013.96

subdivided into two categories: memory-based collaborative filtering [12] [13] and model-based collaborative filtering [14] [15]. Memory-based algorithms can be divided into user-based collaborative filtering. User-based collaborative filtering based on user ratings over the common product calculates the similarity between users, and predicts user's preference in a particular product based on similar users' rating to the product. Item-based collaborative filtering calculates the similarity between products, and then predicts target product rating according to similar products' score. Unlike content-based filtering, collaborative filtering is based on similar products' or similar users' ratings, rather than contents or features of the products. Unlike memory-based algorithms, model-based algorithms require firstly learn the model according to rating records and then use the model to predict target user's rating. Hybrid recommendation is a fusion of content-based recommendation and collaborative filtering recommendation and could result in better performance since overcoming some shortcomings of content-based recommendation and collaborative filtering recommendation.

Although e-commerce recommender systems have made great progress, there still exist some problems, especially the data sparsity problem. This paper proposes a novel method which combine transfer learning and collaborative filtering recommendation for alleviating data sparsity problem of e-commerce recommender systems. Transfer learning is a new framework which transfer knowledge learned from the source domain and the source task to target domains and target tasks. It is not a new concept, in the real world, we can observe many examples of transfer learning. Here is a simple example in terms that the children learning how to play the piano helps him learning how to play the piano. Transfer learning is based on the fact that people will unconsciously use knowledge learned earlier for solving new problems faster. In the study of transfer learning, many people are concerned about following three main issues. These three issues are what to transfer, how to transfer and when to transfer. What to transfer, mainly solve which knowledge can be transferred in different domains or tasks. Some knowledge is unique for specific domains or tasks, while some knowledge is shared between different domains. Only the knowledge can be transferred from source domains to the target domain to help improve performance. The second issue is concerned with how to transfer between different domains and tasks. When to transfer refers to the case that transfer learning can bring improvements, while there exists a research called robust migration, it studies under which circustances, transfer learning is inefficient or even negative.

## III. THE APPLICATION OF TRANSFER LEARNING ON E-COMMERCE RECOMMENDER SYSTEM

In this section, we introduce our approach which combines transfer learning with user-based collaborative filtering recommendation in detail.

Transfer learning method we adopt is put forward in [7]. The method obtain rating pattern between user class and product class in source domain and apply this pattern for fill in valid data into target domain. The algorithm consists of two steps: compress rating data in source domain into rating pattern—codebook, i.e. codebook construction, reconstruct target domain by extending codebook with an efficient algorithm, i.e. codebook transfer.

Collaborative filtering recommendation assumes that similar users or similar products have similar behaviors. Hence, we can cluster respectively users and products at the same time and obtain a cluster-level codebook.

In this paper we apply orthogonal nonnegative matrix tri-factorization algorithm put forward in [16] to gain cluster indicators with Eq. (1),

$$\min_{U\geq 0, V\geq 0, S\geq 0} \|X_{aux} - USV'\|_F^2$$
$$\text{s.t. } U'U = I, V'V = I \quad (1)$$

then construct codebook based on Eq. (2).

$$B = [U'_{aux} X_{aux} V_{aux}] \oplus [U'_{aux} 11' V_{aux}] \quad (2)$$

After constructing codebook, we transfer cluster-level order pattern from source domain to target domain. The precondition of transfer learning is that there exists relevance between domains. We presume that these two domains have implicit correspondence, thereby domain can be reconstructed by extending codebook. We reconstruct order matrix of target domain with Eq. (3).Now is the filled order matrix, then traditional collaborative filtering recommendation can be used to predict users' consuming behavior.

$$\widetilde{X_{tgt}} = W \circ X_{tgt} + [1 - W] \circ [U_{tgt} B V'_{tgt}] \quad (3)$$

where W is weight matrix, Eq. (4) is its computational process.

$$W_{ij} = \begin{cases} 1, & \text{if } [X_{tgt}]_{ij} \neq 0, \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

and are respectively user cluster indicator and product indicator and satisfy conditions in Eq. (5).Finally we can learn from target domain and predict users' intention.

$$\min_{U_{tgt}\in\{0,1\}^{p\times k}, V_{tgt}\in\{0,1\}^{q\times l}} \left\| [X_{tgt} - U_{tgt} B V'_{tgt}] \circ W \right\|_F^2,$$
$$\text{s.t. } U_{tgt} 1 = 1, V_{tgt} 1 = 1 \quad (5)$$

## IV. EXPERIMENTS AND RESULTS

### A. Dataset

We conducted our experiments under two datasets: the dense dataset and the sparse dataset. We gathered data from a glasses websites as the dense dataset and the data from an underware site as the sparse one. After preprocessing, the dense dataset contained 624 users, 797 items, 67,804 records and its ration of sparsity was 13.6%. The sparse dataset contained 116 users, 105 kinds of products, 206 records, its ratio of sparsity was 1.69%.

### B. Compared Method

User-based Collaborative Filtering (U-CF-R): no extra process.

Cluster-based Recommendation (CBS): Clustered the training data based on k-means method, and used consumers' average order data as target user's order data in the same cluster. We considered number of clusters in this experiment 20.

Transfer Learning-based Recommendation (TLR): Extended and reconstructed training dataset based on codebook. The number of User classes and the number of product class in this experiment were respectively set to 20, 15.

### C. Evaluation

When evaluating the accuracy of our approach, we choose a certain percentage of order data for each user from the sparse dataset as the training set, so that we can adjust the ration of sparsity according to the certain proportion. We evaluated the traditional User-based Collaborative Filtering Recommendation (U-CF-R), Cluster-based Recommendation (CBS), Transfer Learning-based Recommendation (TLR) under different ratio of sparsity. And We carried out experiments five times on the same situation, Table I and Table II showed results in an average of five experiments. We used the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to measure recommendation effects. The lower the value of MAE and RMSE was, the higher the accuracy of recommendation was. The formulas of MAE and RMSE were shown in Eq. (6) and Eq. (7) below, where N denoted the total number of records in the test set, and denoted the actual score and predicted score of the i-th record.

$$MAE = \frac{\sum_{i=1}^{N} |r_i - \hat{r}_i|}{N} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (r_i - \hat{r}_i)^2}{N-1}} \quad (7)$$

### D. Result and Analysis

Nu,Ns,Nr in Table I and Table II denoted the number of users, the number of the products, the number of records of the training set, Sparsity indicated the ratio of sparsity of the training set, its computational method was Eq. (8) as follows:

$$Sparsity = 1 - \frac{N_r}{N_u * N_s} \quad (8)$$

Table I and Table II showed respectively results of these three recommendations in the sparse dataset, U-CF-R, CBS, TLR. We can see it from the tables that the higher the ratio of sparsity was, the lower the accuracy of U-CF-R, CBS, TLR were. It suggested that the effect of these three recommendations has a close relationship with the ratio of sparsity. However, as expected, under the same sparsity, compared to U-CF-R, the accuracy of CBS decreased slightly, it indicated that when the data is extremely sparse, CBS can not accurately cluster users and fill in valid data.Meanwhile TLR had better effect of recommendation compared with U-CF-R and CBS, which mean that the Transfer Learning-based recommendation may alleviate the data sparsity problem more efficiently. It can be seen that between e-commerce sites there maybe exist the common knowledge, such as users' consuming habits, TLR would fill in valid data through transferring common knowledge. Thus, TLR can be applied to e-commerce recommender systems and perfom better effect.

TABLE I. MAE（U-CF-R&CBS&TLR）

| Nu | Ns | Nr | Sparsity | U-CF-R | CBS | TLR |
|---|---|---|---|---|---|---|
| 116 | 56 | 85 | 98.70% | 0.51 | 0.49 | 0.47 |
| 116 | 60 | 88 | 98.74% | 0.52 | 0.54 | 0.49 |
| 116 | 63 | 89 | 98.79% | 0.53 | 0.57 | 0.51 |
| 116 | 67 | 90 | 98.85% | 0.59 | 0.59 | 0.55 |
| 116 | 51 | 64 | 98.92% | 0.60 | 0.60 | 0.58 |

TABLE II. RMSE(U-CF-R&CBS&TLR)

| Nu | Ns | Nr | Sparsity | U-CF-R | CBS | TLR |
|---|---|---|---|---|---|---|
| 116 | 56 | 85 | 98.70% | 0.65 | 0.63 | 0.58 |
| 116 | 60 | 88 | 98.74% | 0.66 | 0.69 | 0.63 |
| 116 | 63 | 89 | 98.79% | 0.68 | 0.72 | 0.65 |
| 116 | 67 | 90 | 98.85% | 0.74 | 0.74 | 0.70 |
| 116 | 51 | 64 | 98.92% | 0.77 | 0.77 | 0.73 |

## V. CONCLUSIONS AND FUTURE WORK

### A. Conclusions

In this paper we propose a new solution for the sparsity problem of e-commerce recommender systems. We transfer knowledge from dense datasets to sparse dataset on the basis of user-based collaborative filtering method to predict the user's purchasing intentions. Results show that our method has higher accuracy than the traditional user-based collaborative filtering recommendation and transfer learning can be applied toe-commerce recommender system.

### B. Future work

In actual life, relationship maybe always has influence on people's decision on something, i.e. consuming behavior. The relationship is called trust network .so we can combine e-commerce recommender system with social network which owns abundant resources related with trust network for recommendation.

## ACKNOWLEDGMENT

The authors are partially supported by the National Natural Science Foundation of China under Grant No.11171148 and 61003024;the Humanities and Social Sciences Foundation of Ministry of Education of China under Grant No. 10YJC870020 and 10YJC630283.We thank the anonymous reviewers for their greatly helpful comments.

## REFERENCES

[1] Antonia Dattolo, Felice Ferrara, and Carlo Tasso. The role of tags for recommendation: A survey. In Proceedings of the 3rd International Conference on Human System Interaction (HIS'10), pages 548-555, Rzeszow, Poland, Mary 2010.

[2] Aleksandra Klasnja Milicevic, Alexandros Nanopoulos, and Mirjana Ivanovic.Social tagging in recommender systems: A survey of the state-of-the-art and possible extensions. Artificial Intelligence Review, 33(3)187-209, March 2010.

[3] Z. Wang, Y. Wang, and H. Wu. Tags meet ratings: Improving collaborative filtering with tag-based neighborhood method. In Proceedings of the Workshop on Social Recommender Systems (SRS'10), HongKong, China, 2010.

[4] Sen, S., Vig, J., Riedl, J.: Tagommenders: Connecting users to items through tags. In: Proceedings of the 18th international conference on World wide web, New York, NY, USA, ACM (2009) 671-680.

[5] GuiRong Xue, ChenXi Lin, Qiang Yang, Wensi Xi, Hua-Jun Zeng, Yong Yu,Zheng Chen. Scalable collaborative filtering using cluster-based smoothing. In Proc. of the 28th SIGIR Conf., pages 114-121, 2005.

[6] Nathan Srebro, Tommi Jaakkola. Weighted low-rank approximations. In proc. of the 20th Int'l Conf. on Machine Learning,pages 720-727,2003.

[7] Bin Li, Qiang Yang and Xiangyang Xue. Can movies and books collaborate? Cross-domain collaborative filtering for sparsity reduction. In Proceedings of the 2009 International Joint Conference on Artificial Intelligence, pages 2052-2057,2009.

[8] W. Dai, Q. Yang, G. Xue, Y. Yu, "Boosting for Transfer Learning," Proc. 24th Int'l Conf. Machine Learning, pp. 193-200, June 2007.

[9] K.Lang.Newsweeder: Learning to filtering netnews. In Proceedings of the 12th International Conference on Machine learning, 1995.

[10] R.J.Mooney, L. Roy. Content-based book recommending using learning for text categorization. In Proceedings of ACM SIGIR'99 Workshop on Recommender Systems: Algorithms and Evaluation, 1999.

[11] G.Salton, Automatic text processing. Addison-Wesley, 1989.

[12] John S.Breese, David Heckerman, Carl Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence (UAI98), pages 43-52, 1988.

[13] A.Nakamura, N.Abe. Collaborative filtering using weighted majority prediction algorithms. In Proceedings of the 15th International Conference on Machine Learning, 1998.

[14] DanielBillsus, MichaelJ.Pazzani. Learning Collaborative Information Filters.

[15] L.Getoor, M.Sahami. Using probabilistic relational models for collaborative filtering. In Proceedings of the Workshop on Web Usage Analysis and User Profiling, August 1999.

[16] [Chris Ding, Tao Li, Wei Peng, and Haesun Park.Orthogonal nonnegative matrix tri-factorizations for clustering. In proc. of the 21th ACM SIGKDD Int'l Conf. on Data Mining, pages 625-628,2005