International Conference on Advanced Electronic Science and Technology (AEST 2016)

Problems of recommendation

Juntao Liu[1,a] and Caihua Wu[2]

[1]709th Research Institute, China Shipbuilding Industry Corporation, Wuhan, China
[2]Section of Automatic Command, Huang Pi NCO School, Air Force Early warning Academy, Wuhan, China, e-mail: wucaihua2009@163.com

Abstract. In this paper, we propose a classification framework of recommendation problems. In this framework, the problems of recommendation are classified according to the input and output of the problems. Different kinds of recommendation problems are introduced under this framework. And the corresponding recommendation approaches are also introduced briefly. At the end of this paper, possible research directions are provided.
Keywords: recommendation; recommender system; collaborative filtering; implicit feedback; heterogeneous feedback;

1 Introduction

Nowadays, recommender system is one of the necessary components in online business systems. Recommendation becomes an important problem in recent years. Usually, recommender systems record the users' activities and recommend items to the users. The goal of recommender systems is to learn the users' preferences and give the recommendation accordingly. In different scenarios, the problems of recommendation are different. For example, a movie recommender system recommends a few of movies to the users, while the cloth recommender system sometimes recommend a set of clothes to the user including hat, shirt, and pants, and the recommend clothes should be matched well. It is necessary to distinguish different recommendation problems in order to select suitable recommendation approaches to resolve the problems.

In this paper, we investigate the problems of recommendation in the existed literatures, and proposed a classification framework of these problems. In the framework, problems of recommendation are classified according to the input and output of the problems. The recorded feedback activities in the recommender systems are the input of the recommendation problems. In this paper, four kinds of feedback are investigated, including rating, implicit feedback, heterogeneous feedback and time feedback. Output of the recommendation problem includes predicted ratings, ranked item list and item combination. For each kind of recommendation problem, we also list the recommendation approached resolving this problems in the literatures.

The rest of this paper is organized as follows. In section 2, related work is provided. In section 3, the classification framework of recommendation problem is introduced. And the different kinds of recommendation problems with the corresponding recommendation approaches are also introduced. Future research directions are given in section 4 followed by the conclusion in section 5.

[a] Corresponding author : prolay@163.com

2 Related work

Adomavicius and Tuzhilin [13] review the recommender systems, and divide the recommendation approaches into three kinds: content-based, collaborative and hybrid recommendation approaches. The limitations of each kind of recommendation approaches are discussed. According to these limitations, possible extensions of the recommendation approaches are given. Adomavicius and Tuzhilin [13] classify the recommender systems in the view of recommendation approaches, not in the view of recommendation problems.

Park et al. [14] review articles on recommender systems from 37 journals which were published from 2001 to 2010. The reviewed approaches are classified by recommendation fields and data mining techniques. The distribution of these approaches by recommendation fields and data mining techniques are provided. And the research trend is analyzed.

The main difference between the existed surveys on recommender systems and this paper is that we main focus on the recommendation problems, while the existed surveys review the recommendation approaches.

3 Problems of recommendation

3.1 Framework of the classification

In this paper, we classify the recommendation problem according to the input and output of the recommender systems. Recommender systems usually collect the users' activities in the systems, including rating, clicking, buying, and comment and so on. Recommender systems take this information as input of the recommendation method. The output of recommender systems is the recommendation result given to the users. Different recommender systems give different kinds of recommendation results. Some systems predict ratings, while some systems recommend a ranked item list to each user. And in particular domination, a combination of items is needed.

The classification framework of recommendation problem is shown in Figure 1. Next, we will introduce the problems of recommendation according to this framework.

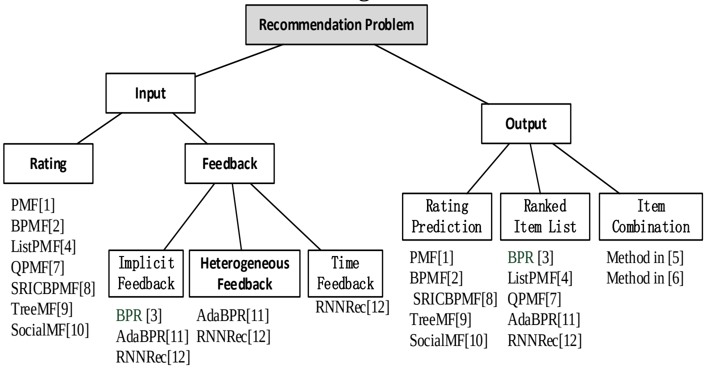

Figure 1. Classification Framework of the recommendation problems.

3.2 Output

3.2.1 Rating prediction based recommendation

In some recommender systems, users are allowed to rate the items that they have accessed. Suppose there are M users and N items. Let matrix R denote the rating matrix, where r_{ij} represents the rating of user i for item j. Typically, rating value r_{ij} is a k-point integer.

The problem of rating prediction is to predict the rating of item for a user according to the observed ratings. Suppose the matrix $\hat{R}$ is the predicted rating matrix. The goal of rating prediction is learning some parameters Ω from the observed rating matrix R to make the predicted rating matrix $\hat{R}$ is close to R as much as possible. Formally, in rating prediction problem, the following objective function is minimized

$$f_P(R,\hat{R}) + reg(\Omega) \tag{1}$$

where f_P is the loss function measuring the difference between observed ratings and predicted ratings. f_P is typically the sum-of-squares of prediction error. That is

$$f_P = \sum_{i,j}(r_{ij} - \hat{r}_{ij})^2 \tag{2}$$

$reg(\Omega)$ is the regulation term of parameters Ω, which is used to prevent over-fitting.

After the parameters Ω is learned, the recommender system use these parameters to predict the ratings for the items that the users have not accessed. And the items with high predicted ratings are recommended to the users.

Probabilistic matrix factorization (PMF) [1] is the most popular method resolving the rating prediction problem. In PMF, it is assumed that the predictive error of matrix factorization follows the Gaussian distribution. Gradient descent algorithm is used to find the local maximal of the posterior probability over user and item latent matrices with parameters. Several rating prediction methods based on PMF are proposed, such as BPMF [2], SRICBPMF [8], TreeMF [9], Social MF [10]. It's worth noting that the conditional preference of user is represented by a tree model in TreeMF.

3.2.2 Ranking based recommendation

In some recommender system, the preference orders of the users are collected. The preference order reflects the user preference. That's mean the user prefers the items in the front of the list to those in the end of the list. The preference order may obtain by sorting the user's ratings, or learned by observing the user's behavior. For example, in online busyness system, the preference order is obtained by comparing the bowering times of the products. The collected preference orders are used as training samples to learn the users' preference models. Suppose the collected preferences are $\Pi = \{\pi_i\}$, here π_i is the preference order of user i. The recommender system learns the users' preference according to the collected preference orders Π to predict the preference order accurate. The task of the learning algorithm is to minimize the following objective function

$$f_L(\Pi,\hat{\Pi}) \tag{3}$$

Where $\hat{\Pi} = \{\hat{\pi}_i\}$ are the predicted preference orders. $\hat{\pi}_i$ is the predicted preference order of user i. f_L Measures the difference between the collected preference orders Π and the predicted preference orders $\hat{\Pi}$. f_L can be pair-wise or list-wise.

Pair-wise objective function counts the number of different comparison pairs between collected preference orders Π and the predicted preference orders $\hat{\Pi}$. The example is BPR [3], in which the objective function is the weighted sum of comparison error. AdaBPR [11] modifies the original BPR to make it can handle different kinds of implicit feedback at the same time.

List-wise objective function takes the preference order as a whole instance and measure the difference. Liu et al. [4] use P-L model based function and cosine based function to measure the difference between two preference orders. Objective function in QPMF [7] is very similar to that in ListPMF [4]. But the preference of user is represented by quartic function to make the method can handle conditional preference.

RNNRec[12] is proposed to solve the time heterogeneous feedback recommendation problem. A recurrent neural network is used to embed user, item and feedback. The recurrent structure makes the model can remember the historical feedback.

3.2.3 Combination of items recommendation

In some scenarios, the user may want to buy a combination of products. For example, a customer may want to buy a jacket as well as the matched pant. In these cases, the recommender systems should recommend a suitable combination of items to the users.

In the problem of combination recommendation, the relation among items should be investigated. Some items are complementary, and these items should be bought together. Some items are replaceable, and they can be interchanged by each other. Some items are conflict, and customers may not buy them simultaneously.

In recent literatures, the problem of combination recommendation is usually treated as an information retrieval problem. Recommender system takes the items that the user selected as query and return items matching to them by some standards.

McAuley et al. [5] model the relation between two items by the visual feature. The probability of one kind of relation between two items is a function of the distance between the visual feature vectors of this two items. Visual feature vector of item is also transform to the style space by Mahalanobis transform. There are only two kinds of relations are considered in [5], that is substitute and complement.

In [6], the co-occurrence relation of clothes is learned from online fashion images and their meta-data, which is used for recommendation. The fashion images are separated into several parts. Each part is represented by a histogram. A predictive model is learned, which can predict the histogram of a missing part when given the histograms of the rest parts. Complementary Nearest Neighbor Consensus, Gaussian Mixture Models, Texture Agnostic Retrieval and Markov Chain LDA are used to represent the relation among visual feature of the parts.

3.3 Input

3.3.1 Rating

Rating is the most used input in recommendation systems. In most recommender systems, the users are allowed to rate the item they have accessed. Usually, the rating value is a k-point integer. For example, in Taobao website (http://www.taobao.com), rating values are 5-point integers, 1-point means ’very bad’ and 5-point means ’excellent’.

Rating is regarded as explicit feedback. Many recommendation approaches taking rating as input are proposed [1, 2, 4, 7-10].

3.3.2 Implicit feedback

Implicit feedback includes clicks, view, etc. Different from the explicit feedback, such as ratings, comments, etc., the users do not to express their taste explicitly through implicit feedback. But implicit feedback can be recorded by the web servers automatically. It is easy to collect implicit feedback. So, in many recommender systems, implicit feedback is used to learn the users’ preference.

BPR [3] can handle only one kind of implicit feedbacks. The binary preference relation of a user between items is infer from the feedback of this user.

3.3.3 Heterogeneous feedback

In most recommender systems, there are several kinds of feedback. Recommendation approaches should give user recommendation according to the heterogeneous feedback. For example, in 2015, a recommendation competition in mobile scenario is hold by ALI, which is the biggest P2P online business website in China. It is required that the purchase behaviors of users are predicted according to the historical behaviors, such as clicking, adding to favorite, adding to cart and purchase.

Instead of learning from only one kind of feedback, some recommender systems lavage multi-kind of feedback [11, 12]. In these systems, several kinds of feedback, including explicit feedback and implicit feedback, are collected, and the collected feedbacks are used as training samples to learn the users' preference.

3.3.4 Time feedback

Feedback activities as well as the time stamps are recorded in many recommendation systems. But little recommender system uses this information. In fact, feedback with time stamp reflects the trend of the users' preference, which is valuable information for recommendation, especially in the time-sensitive fields, such as fashion recommendation and news recommendation.

RNNRec [12] is proposed to resolve the problem of time heterogeneous feedback recommendation. To our best knowledge, it is the first time to use recurrent neural network to solve this problem.

In Table 1, we list different kinds of recommendation problems. Each cell in the table represents one kind of recommendation problem considering both input and output. And in the cell, the recommendation approaches resolving this problem are listed.

Table 1. Different Kinds of Recommendation Problems and Recommendation Approaches

Input / Output	Rating Prediction	Ranked Item List	Item Combination
Rating	BPMF[2] SRICBPMF[8] TreeMF[9] SocialMF[10]	ListPMF[4] QPMF[7]	
Implicit Feedback		BPR [3]	Method in [5] Method in [6]
Heterogeneous Feedback		AdaBPR[11]	
Time Feedback		RNNRec[12]	

4 Research directions

4.1 More suitable evaluation metrics

How to evaluate the performance of the recommender systems is the fundamental problem in this field. The existed evaluation standards can be divided into 3 categories, point-wise; pair-wise and list-wise metrics. Point-wise evaluation standards include RMSE, MAE, etc. This kind of evaluation metrics are designed for rating based recommender systems. And they measure the rating prediction error. Pair-wise evaluation metrics measure the pair-wise comparison error. List-wise evaluation metrics, such as ERR, Precise, Recall, F1 score, et al., are usually used in the field of information retrieval.

Finding more suitable evaluation standard is an important future research direction. The evaluation standards should reflect the satisfaction of the users. On the other hand, the metrics should also reflect the profit improvement of the online busyness systems, which is one of the purposes deploying the recommender systems. Using this kind of metrics can help to decide whether to deploy the recommender systems. But, in the existed literatures, there are no metrics about the profit improvidence of the online business systems.

4.2 Personal combination recommendation

In recent years, combination recommendation is receiving more and more attentions. In real life, people will select several items simultaneously. For example, when a person has dinner, he (or she) will choose several dishes from the menu. And these dishes should not only complement but also satisfy his (or her) taste.

In the existed literature, the combination recommendation problem is treated as an information retrieval problem, and only the relation among the items are considered. In future research, the user's preference should also be taken into consideration to obtain the personal recommendation results.

4.3 Content aware recommendation

Content in recommender systems include location of the user, the current user activity, time, season, weather, public event and so on. The content can help to give satisfying recommendation. In scenario of content aware recommendation, the input is not only the user feedback but also the content information. How to use content information is a valuable research direction.

5 Conclusion

In this paper, we propose a classification framework for recommendation problems. We consider both the input and the output of the recommender systems. The problems of recommendation are classified according to the framework. And the recommendation approaches resolving each kind of recommendation problem are also reviewed. Finally, possible research directions are given.

Acknowledgment

This work is supported by National Natural Science Foundation of China (No.61403350 and No.61401228).

References

1. R. Salakhutdinov, A. Mnih, Probabilistic matrix factorization, in: ProceedingsofAdvancesinneuralinformationprocessingsystems, NIPS’08, 2008, pp. 1257–1264.
2. R. Salakhutdinov, A.Mnih, Bayesianprobabilisticmatrix factorization using markov chain MonteCarlo, in: Proceedings of the 25th international conference on Machine learning, ACM, 2008, pp. 880–887.
3. Z. Gantner, L. Drumond, C. Freudenthaler, L. Schmidt-Thieme: Bayesianpersonalized ranking for non-uniformly sampled items, in: Proceedings of Knowledge Discovery and Data Mining (KDD) Cup and Workshop 2011, 2011.
4. J. Liu, C. Wu, Y. Xiong, W. Liu, List-wise probabilistic matrix factorization for recommendation, Information Sciences 278 (2014) 434 – 447.
5. J. McAuley, C. Targett, Q. Shi, A. Hengel. Image-based Recommendations on Styles and Substitutes, SIGIR 2015, 2015.
6. V. Jagadeesh, R. Piramuthu, A. Bhardwaj, W. Di, N. Sundaresan. Large Scale Visual Recommendations from Street Fashion Images, KDD2014, 2014.

7. W. Liu, C. Wu, B. Feng, J. Liu, Conditional preference in recommender systems, Expert Systems with Applications 42 (2) (2015) 774 – 788.
8. J. Liu, C. Wu, W. Liu, Bayesian probabilistic matrix factorization with social relations and item contents for recommendation, Decision Support Systems 55 (3) (2013) 838 – 850.
9. J. Liu, C. Sui, D. Deng, J. Wang, B. Feng, W. Liu, C. Wu, Representing conditional preference by boosted regression trees for recommendation, Information Sciences 327 (2016) 1 – 20.
10. H. Ma, T. Zhou, M. Lyu, I. King, Improving recommender systems by incorporating social contextual information, ACM Transactions on Information Systems (TOIS) 29 (2) (2011) 9:1–9:23.
11. W. Pan, H. Zhong, C. Xu, Z. Ming, Adaptive Bayesian personalized ranking for heterogeneous implicit feedbacks, Knowledge-Based Systems, 73(0) (2015) 173 – 180.
12. C. Wu, J. Wang, J. Liu, W. Liu. Recurrent neural network based recommendation for time heterogeneous feedbacks, Knowledge-Based Systems, (2016)
13. Adomavicius, G., Tuzhilin, A., 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. Knowledge and Data Engineering, IEEE Transactions on 17 (6), 734–749.
14. Park, D. H., Kim, H. K., Choi, I. Y., Kim, J. K., 2012. A literature review and classification of recommender systems research. Expert Systems with Applications 39 (11), 10059–10072.