A Collaborative Filtering Recommendation Algorithm Incorporated with Life Cycle

Zheng Chonglin[1, a], Hao Kuangrong[1,2,b] and Ding Yongshen[1,2,c]

[1]College of Information Sciences and Technology, Donghua University, Shanghai, P. R. China

[2]Engineering Research Center of Digitized Textile & Fashion Technology, Ministry of Education

[a]jeff_airu@126.com, [b]krhao@dhu.edu.cn, [c]ysding@dhu.edu.cn

Keywords: Collaborative filtering, Time information, User interest, Life cycle

Abstract. Collaborative filtering recommendation algorithm is the most successful technology for recommendation systems. However, traditional collaborative filtering recommendation algorithm does not consider the change of time information. For this problem，this paper improve the algorithm with two new methods:Predict score incorporated with time information in order to reflect the user interest change; Recommend according to scores by adding the weight information determined by the item life cycle. Experimental results show that the proposed algorithm outperforms the traditional item in accuracy.

Introduction

With the rapid development of information technology and the internet, recommendation system exists to solve the highlighted problem of information overload. Recommendation system connects the user and information. It can help the user search for interest information in mass information. As the most successful recommendation system, Collaborative filtering recommendation system is widely used in reality. As a basic data set in recommendation algorithm, the user-item score set has a serious problem of data sparseness. Score predicting exists to solve the problem by predicting the unknown scores with history information. The item-based collaborative filtering algorithm calculates the item similarity according to the user-item score set, and then predicts scores with k-neighbors formed by k items having higher similarity. Finally it recommends items to the user with scores incorporated with content information.

The traditional collaborative filtering algorithm does not consider the change of time information, which includes the change of user interest and item life cycle. The change of user interest reflects the inconsistency in interest items for the user in different period, and item life cycle reflects the regularity of popularity. Therefore the change of time information greatly affects the quality of recommender system. To reflect interest change, Literature [1] presents time-based data weight and item-based data weight. Literature [2] tracks and studies user interest according to the Ebbinghaus forgetting curve in psychology. It divides user interest into short-term interest and long-term interest and proposes weight function based on time window. This paper considers the effect of time information change in score predicting, also considers the effect of item life cycle in recommendation process, then proposes a collaborative filtering recommendation algorithm incorporated with life cycle.

Score Predicting

In traditional collaborative filtering algorithm, score predicting does not consider the change of user interest with the item. In real life, the demand of items for the user is changing with time information, thus the traditional recommendation algorithm diverges from the user's demand, leading to the inaccurate of recommendation results. At present, research on user interest change is devoted to fit the forgetting curve and recommend based on time weight. The algorithm is difficult to quantify. This paper considers the effect of user interest change by establishing time information

model in score predicting, which can predict more accurately.

In a collaborative filtering system, given user set $U = \{u_1, u_2, \ldots, u_n\}$ and item set $P = \{p_1, p_2, \ldots, p_m\}$, there is a user-item score matrix $R = (r_{ij})_{n \times m}$. In which, r_{ij} is the score which user u_i rates item p_j. The problem of data sparseness due to serious missing data reduces accuracy of the recommendation. To predict missing scores becomes the central research of recommendation system. To improve the accuracy of recommendation, firstly standardize user-item score matrix with Gauss regularization method. Then get standardized user-item score matrix $R' = (r'_{ij})_{n \times m}$, where

$$r'_{ij} = \frac{r_{ij} - \bar{r}_i}{\sqrt{\sum_{R_{u_i}} (r_{ij} - \bar{r}_i)^2}} \tag{1}$$

In which, R_{u_i} is a score set which user u_i rated, $\bar{r}_i$ is the average score of the set R_{u_i}. Time information model in score predicting presents user interest change. Simplified to consider the time difference $\Delta t = t_{ui} - t_{uj}$ between item i and item j which user u rated. Then

$$\text{sim}_{ij} = \text{sim}(p_i, p_j) = \frac{\sum_{u \in U_{p_{i,j}}} r'_{ui} \cdot r'_{uj}}{\sqrt{\sum_{u \in U_{p_{i,j}}} r'^2_{ui} \cdot \sum_{u \in U_{p_{i,j}}} r'^2_{uj}}} \tag{2}$$

$$\hat{r}_{up} = \bar{r}_u + \frac{\sum_{i \in TOP(p,K) \cap P_u} f(\text{sim}_{pi}, \Delta t) \cdot r'_{ui}}{\sum_{i \in TOP(p,K) \cap P_u} f(\text{sim}_{pi}, \Delta t)} \tag{3}$$

In which, $U_{p_{i,j}}$ is a user set which rates item p_i and item p_j. $TOP(p, K)$ is an item set with the top K neighbors similarity to item p. P_u is an item set which user u rated. $f(\text{sim}_{pi}, \Delta t)$ is a similarity function considering time decay, which improve the effect of the user's most recent rating data in recommended results. Then it reduces the prediction error effectively. where

$$f(\text{sim}_{pi}, \Delta t) = \sigma\left(\delta \cdot \text{sim}_{pi} \cdot \exp\left(\frac{-|\Delta t|}{\beta}\right) + \gamma\right) \tag{4}$$

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \tag{5}$$

From the above, the weight will gradually become smaller with Δt expansion, that is, the more distant the user rated, the smaller the recommended results affect. It is in line with the change in the characteristics of user interest. In short term, the change of user interest maintains stability, and after a period of long term will be greatly different. Recommendation system tends to recommend the recent interest to the user.

Recommendation Algorithm Incorporated with Life Cycle

This paper proposes the algorithm incorporated with item life cycle to improve recommendation quality. Harvard University professor Raymond Vernon first proposed Product Life Cycle Theory. Generally a product includes the introduction period, the growth period, the maturity period and the recession period in the market from the entry to the elimination. Product sales grow slowly in the introduction period, and grow fast in the growth period, then stabilize in the maturity period, finally decline sharply until demise in the recession period. According to the nature of the product life cycle, the life cycle curve in Fig.1 below.
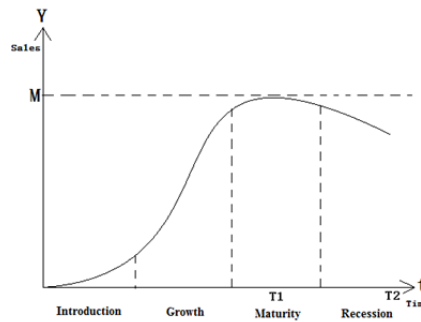

Fig.1 Life Cycle Curve

Through the fitting of the life cycle curve, it can get the mathematical function meeting the curve in Fig.1:

$$Y_t = \begin{cases} \frac{M}{1+Ae^{-k_1 t}}, & t \in (0, T_1) \\ -k_2(t-T_1)^2 + M, & t \in [T_1, T_2] \\ 0, & t \in (T_2, +\infty) \cup \{0\} \end{cases} \tag{6}$$

In the formula, the parameter M is the peak sales, the parameter A descripts position of the curve of the first paragraph, the parameters k_1、k_2 are respectively used to control the morphology of the two curves. T_1 is the time when falls after stable in the peak for some time, and T_2 is the extinction moment. It can be estimated based on historical sales data for each parameter for that similar products are in line with the same parameters of the curve.

The life cycle curve shows a process including slow growth, accelerating, stabilizing, rapid decline and demise. For recommendation system, the products in the accelerated phase inclined to recommend to the user, because the products of this stage are fresh to users; the products in the slow grew phase is also fresh to users but questioned with unpopular and risky; the products in the stabilized phase are so popular that recommendation loses the significance as well as hard to realize the personalized features. Users will buy, if not recommended, and it causes serious long tail effect; the products in the rapid decline have been loss of market competitiveness and freshness at all. From the above analysis, the tangent slope of the point on the curve can reflect the recommended level i.e. the recommended weighting factor. Derivative function as follow:

$$\dot{Y}_t = \begin{cases} \frac{Mk_1 Ae^{-k_1 t}}{(1+Ae^{-k_1 t})^2}, & t \in (0, T_1) \\ -2k_2(t-T_1), & t \in [T_1, T_2] \\ -\text{MAX}, & t \in (T_2, +\infty) \cup \{0\} \end{cases} \tag{7}$$

In the recommendation process, recommendation algorithm incorporated with item life cycle can improve the accuracy and freshness of the recommended results. Integrating into the weighting information generated by the item life cycle, interest user i for item j can be expressed as $\hat{r}_{ij} \cdot \dot{Y}_t$, in which the time of the item is estimated based on historical sales of similar items. Finally, recommend Top-N items.

Experimental Results and Analysis

This paper uses the Netflix data set, which contains not only the user sparse scores also time information consistent with this research about collaborative filtering algorithm based on time information change.

Predicting accuracy is the most important evaluation indicators in offline experiments of recommendation system. For score predicting, calculate the mean absolute error (MAE):

$$\text{MAE} = \frac{\sum_{r_{ij} \in T} |r_{ij} - \hat{r}_{ij}|}{|T|} \tag{8}$$

In which, T is an item set not rated. The lower MAE is, the smaller score prediction error is, that is the better the quality of recommendation is.

First, compare the predictive accuracy between traditional collaborative filtering algorithm and algorithm based on time information model. For each certain K in K-neighbors model, divide experimental data set into five different divisions of the training set and the testing set. Obtain average MAE calculated based on each division to reduce random factor and improve the accuracy of the MAE calculations corresponding with K.
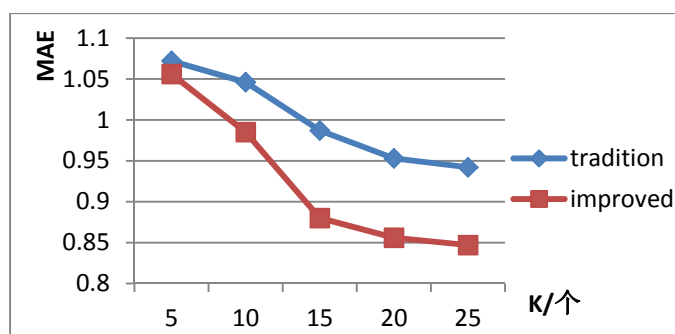
Fig.2 Comparison of recommendation algorithm

Experimental results about comparison between the traditional item-based collaborative filtering recommendation algorithm and the proposed improved algorithm in a number of different neighbors are shown in Fig.2. Under the conditions of different value of K, the proposed improved recommended algorithm incorporated with time information has a smaller MAE. It is visible to increase the score prediction accuracy, and thereby improve recommendation accuracy.

Conclusion and Future Work

For the problem that traditional collaborative filtering algorithm does not consider the impact of the time information, this paper proposes score predicting with time information change model and recommendation adding weight information determined by item life cycle. It overcomes the drawbacks of the traditional algorithm and improves recommendation precision and freshness. The next step:with the increasing users and items, the increased complexity of improved recommendation algorithm brings a burden to system, then how to be simpler and more effective becomes a real problem; better item life cycle curve fitting and parameter estimation, and special item life cycle.

Reference

[1] Xing Chunxiao, Gao Fengrong, Zhan Sinan. A Collaborative Filtering Recommendation Algorithm Incorporated with User Interest Change. Journal of computer research and development, 2007, 44(2) : 296-301 .

[2] Yu Hong, Li Zhuanyun. A Collaborative Filtering Recommendation Algorithm Based on Forgetting Curve. Journal of Nanjing University (Natural Sciences), 2010,46(5) : 520-527.

[3] Li Kechao, Liang Zhengyou. Exponential Forgetting Collaborative Filtering Recommendation Algorithm Incorporated with User Interest Change. Computer Engineering and Applications, 2011,47(13) : 154-156 .

[4] Liu Xudong, Chen Deren, Wang Huimin. A User-based and Item- based Collaborative Filtering Recommendation Algorithm. Journal of WUT (Information &Management Engineering), 2010,32(4):550-553.

[5] Jiang Wei, Pang Xiuli. Research on Personal Hybrid Recommendation Overcoming Data Sparse Problem. Computer Engineering and Applications, 2012,17(12):1-6.

[6] J. Herlocker, J. Konstan, L. Terveen, and J. Riedl.Evaluating Collaborative Filtering Recommender Systems. ACM Transactions on Information Systems, pp.5-53, ACM Press,2008.

[7] Wujian Yang, Zebing Waug and Mingyu You, An improved Collaborative Filtering Method for Recommendiations' Generation, In Proceedings of IEEE International Conierence on Systems, Man and Cybernetics, 2009.