Novel Recommendation of User-based Collaborative Filtering

Liang Zhang[1, 2] Li Fang Peng[1], Phelan C.A[3]
[1]School of Management
Xiamen University, Xiamen,China
[2]School of Economics and Management
Guizhou Normal University
Guiyang,China
[3]School of Computer Science
ABBD Co., Ltd.
Ottawa, Canada
ftygygz@vip.sina.com, Phelan_ca@gmail.com

Journal of Digital
Information Management

ABSTRACT: *Recommendation system has been widely used in various types of e-commerce sites. One of the most successful examples is the collaborative filtering algorithm. However, the traditional algorithms only aim at accuracy and ignore these factors closely related with customer satisfaction, such as novelty etc. In this paper, we defined novelty of item from the perspective of the users, designed the corresponding offline experiment scheme and evaluation metrics. The dissimilarity and the time-popularity were embedded in the traditional collaborative filtering algorithm, the ability of predicting user's future needs and coverage of recommended list were obviously improved, and the ability of recommended long tail items were also enhanced.*

Subject Categories and Descriptors
H.2.8 [Database Applications] Data Mining; **H.5.3 [Group and Organization Interfaces]** Collaborative Computing

General Terms: Recommendation System, Data Mining

Keywords: Collaborative Filtering, Novel Recommendation, Dissimilarity, Popularity

Received: 11 November 2013, Revised 1 March 2014, Accepted 19 March 2014

1. Introduction

With the development of information technology and application of the Internet, people gradually transferred from the time of information scarcity to the time of information overload [1-3]. Nowadays, it is a very challenging problem to choose the right product to the right customer. Recommendation system is the same as search engine in order to help users quickly and efficiently find valuable information. Recommender systems do not require users to provide specific requirements. Through analysis of users' past behavior, model user preference automatically provides users' need and interested information. Now, almost all the major e-commerce systems such as Amazon, e-Bay, etc., used various recommendation systems to varying degrees. Recommendation systems also have several advantages, including increasing the probability of cross-selling, establishing customer loyalty, and fulûlling customers' needs by presenting products of possible interest to them [4].

By far, there are many recommendation algorithms in recommendation system, such as Collaborative Filtering(CF) [5-6], Content-based Recommendation [7], Latent Factor Model [8], Heat Conduction[9], Mass diffusion [10], Tag-based Filtering [11] and so on. Collaborative filtering is the most successful and widely used recommendation technology in E-commerce recommender systems, and has been widely used in many E-commerce web sites, e.g. Amazon.com [12], and other applications, e.g. mobile commerce [13], E-learning [14], digital library [15].

In most cases, recommended list contains a lot of similar projects in recommendation system. One major issue with accuracy metrics is their inability to capture the broad aspects of users' satisfaction, hiding several blatant flaws in existing systems [16].Torres suggested that differences in language and cultural background influenced users' satisfaction [17]. Swearingern and Sinha examined how usefulness, novelty and usability are related to users' satisfaction by means of a questionnaire survey and reported that they are significantly correlated [18]. Users' satisfaction with recommender systems is related not only to how accurately the system recommends but also to how much it supports the user's decision making. Users will be more satisfied if recommender system suggests unexpected items that are relevant to the users' preferences [19]. Recommending relevant items alone, though, is often not suûcient to satisfy user expectations, but other characteristics, such as diversity, novelty, serendipity and trust must be considered as well [20-22].

In this paper, we investigated novel recommendation of user-based collaborative filtering. Section 2 reviews related work on novelty in recommendation system. Section 3 defines novelty of item and designs offline experiment, section 4 evaluates novel recommendation in traditional user-based collaborative filtering algorithm, we embed dissimilarity variable into user-based collaborative filtering algorithm and evaluate recommended result in section 5; in section 6 we embed dissimilarity and unknown variables into user-based collaborative filtering algorithm and evaluate recommended result. We conclude in Section 7 with ûnal remarks and our plan for future research.

2. Related work of novel recommendation

Novel recommendation has recently attracted more and more attention from both academy and industry. There is no unified definition about noveltyÿtherefore, novel recommendation algorithms are different with respect to the definition of novelty.

Novel recommendation is simply understood as recommended items users unknown; hence the simplest way to novel recommend is to filter items in profile of the user. Although this method is simple and less system resources, but items in profile of the user is less proportion of all items, making result less effective. Ensuring a certain degree of accuracy increased novel recommendation is the focus of research scholars.

The algorithm based on popularity is the most currently used method. Shani, G suggested novelty can be taken into account be using an accuracy metric where the system does not get the same credit for correctly predicting popular items as it does when it correctly predicts non-popular items [23]. Through such adjustment can increase the possibility of recommended unpopular items. Ziegleret al. [24] and Celma etc. [25] also used accuracy measures that take popularity into account. Jinoh, Oh proposed an efficient novel-recommendation method called Personal Popularity Tendency Matching (PPTM) which recommends novel items by considering an individual's Personal Popularity Tendency (or PPT) [26].

Herlocker suggested to create a list of "*obvious*" recommendations and to remove the obvious ones from each recommendation list before presenting it to users. A disadvantage of this approach is that the list of obvious items might be different from each user, since each person has had different experiences in the past. An alternative would combine what is known about the user's tastes with what is known about the community's tastes [27]. Hijikata,*Y* suggested that using the ratings of acquaintance calculates the probability that a user knows an unrated item [28].

Weng et al. suggested a taxonomy-based RS that utilizes hot topic detection using association rules to improve novelty and quality of recommendations. To better capture user's range of tastes [29], Mi Zhang and Neil Hurley proposed to partition the user profile into clusters of similar items and compose the recommendation list of items that match well with each cluster, rather than with the entire user profile, and evaluate a number of partitioning strategies in combination with a dimension reduction strategy [30].

Traditional recommendation algorithms only take into account the contributions of similar users, thus, they tend to recommend popular items for users. W. Zeng proposed a recommendation algorithm by considering both the effects of similar and dissimilar users under the frame work of collaborative filtering. The algorithm to some extent is to remove popular items [31].

3. Define novelty and design offline experiment

3.1 Definition of novelty

Definition of novelty is a crucial role and basis for novel recommendation. From the analysis of research status described above, researchers for the definition of novelty are not the same according to research field and data characteristics, and algorithm design of novel recommendation is also different.

According to Wordnet dictionary (http://wordnet.princeton.edu), novel(adj.) has two senses: "*new–original and of a kind not seen before*"; and "r*efreshing–pleasantly new or different*". Likewise, familiar (adj.) is defined as "*well known or easily recognized*". According to the definition of novel words,"*Novel item*" Should have the following three characteristics:

(1) **Unknown:** The item is unknown to the user;

(2) **Satisfactory:** The item is satisfied for the user;

(3) **Dissimilarity:** The item is dissimilar to items in profile of the user.

Recommendation system explicitly obtaining nformation about unknown and satisfactory of items will seriously destroy user experience, so we only infer the possibility of unknown and satisfactory of items through profile of user. Suppose $\text{dis}\,(i, prof_u)$ is dissimilarity between item i and the set of items in user's profile, the definition of novelty as follows [32]:

$$Novelty\,(i, u) = p\,(i \mid like, u) \times p\,(i \mid unknown, u) \times dis\,(i, prof_u) \quad (1)$$

Traditional recommendation algorithm is mainly forecast the possibility of the user like the item by modeling user's preferences. It considered the first aspect of the novelty. In this work, we gradually embed dissimilarity and unknown variables in the traditional algorithm to explore methods of how to enhance the novel recommendation of user-based collaborative filtering.

3.2 Offline experiment of novel recommendation

In Traditional experiment, researchers hide a part of rated data and use recall and precision metrics to evaluate the result of recommendation system. It is reasonable to suppose hided items meet user's preferences, but ignores the following points: the item meeting the user's preferences is not always the user's need; hided items were known to users. From the essence of the recommendation system, the ideal state is that the recommended item is the user's next need. Therefore the experiment design and evaluation metrics must consider the factor of time. Therefore, we use user behavior data set with a timestamp, and set a time point to divide into two subsets. For users, items rated before the time point are known to users, recommending these items is meaningless. It is reasonable to suppose these items that rated high score by the user after the time point are novelty to the user. According to the above ideas we designed evaluation metric (Formula 2).

$$Novelty = Recall_a - Recall_b = \frac{|\, Recomm_u \cap Hide_{after} \,|}{|\, Hide_{after}} - \frac{|\, Recomm_u \cap Hide_{before} \,|}{|\, Hide_{before}} \quad (2)$$

Where $Recomm_u$ is recommended items set of user u, $Hide_{after}$ and $Hide_{before}$ respectively is hided items set of rated after and before the time point. $Recall_b$ is the recall of traditional experiment, $Recall_a$ represents accuracy metric of predicting the user's demand in the future. The number of recommended list is very limited, we want to the recommended list include more items meeting the user's future need and less items user known, it is implication of the metric '*Novelty*'. Meanwhile, the average popularity (formula 3) and coverage (formula 4) act as reference metrics of novelty.

$$Avg\_pop = \frac{\sum_{u \in U} \sum_{i \in Recomm_u} pop_i}{\sum_{u \in U} |\, Recomm_u \,|} \quad (3)$$

$$Coverage = \frac{|\sum_{u \in U} Recomm_u|}{|\, I \,|} \quad (4)$$

Where pop_i is rated times of item i, U is user set and l is item set. The two metrics measure the ability of recommended long tail items, also are reflections of novel recommendation.

In our experiment, we used MovieLens RecSys2011 data set which includes 855598 rated data of 2113 users to 10197 films with timestamp. There are 405 rated data of each user and 84.6 rated data of each movie on average, rated score range is from 1to 5 points, the higher score represents that the user more liked the movies. We used the last 300 days of data as R_a (items set rated after the time pont), the rest data as R_b(items set rated before the time pont). The R_b included 641600 rated data, The R_a which be weed out users and movies not be in R_b included 36266 rated data. We randomly hided 36266 rated data in R_b and used it as training data, used the hided data in R_a and R_b as test data in our experiment. The following test results are mean values after repeat 5 times experiment.

4. Novel recommendation of the traditional user-based CF

In the traditional user-based CF algorithm, first convert $\{R\,(u, i, t)\}$ to UI matrix, afterwards calculate similarity between user to user, finally recommend items which the nearest neighbor of the user liked. In this paper, we used pearson correlation coefficient (formula 5) to calculate similarity between users, and used formula 6 to predict score, finally recommended N items to the user according to rank of predict score.

$$Sim\,(x, y) = \text{Pearson}\,(x, y) = \frac{\frac{1}{I\,(x, y)} \sum_{i \in I\,(X, Y)} (x_i - \bar{x})\,(y_i - \bar{y})}{\sqrt{\frac{1}{I\,(x, y)} \sum_{i \in I\,(x, y)} (x_i - \bar{x})^2}\,\sqrt{\frac{1}{I\,(x, y)} \sum_{i \in I\,(x, y)} (y_i - \bar{y})^2}} \quad (5)$$

$$Predict\,(i - u) = \sum_{n \in neigh\,(u, K)} sim\,(u, n)\, r_{ik} \quad (6)$$

$$I\,(x, y) = |\, I_x \cup I_y \,| \quad (7)$$

$$Top\,(N, u) = \{i \in I \mid Rank\,(i, Predict\,(i, u)) \leq N\} \quad (8)$$

$$Recomm\,(N, u) = Top\,(N, u) \quad (9)$$

Where $neigh\,(u, K)$ represents K users are most similar to the user u, lx is rated items of user x, $Rank\,(i, Predict\,(i, u))$ is predicting score ranking of item i, $Recomm\,(N, u)$ is recommended N items to the user u. The table 1, 2 and figure 2,3 are the experimental result of the traditional user-based CF algorithm.

Although the traditional collaborative filtering algorithm can better reflect the user's preferences, but items user known occupies a large part. The ability of predicting user's need is relatively weak, and such gap is bigger with the increase of number of recommended. These phenomena can't be observed if using traditional accuracy metric. Avg_pop also reflects traditional algorithms seriously tend to recommend hot items. The duplication of the recommended items for different is very high, thus led to coverage is low. The accuracy has a certain degree of improvement with the increase of the number of neighbor, but novelty has no obvious change and coverage comes down very much. We used the result of the traditional algorithm as reference in these following experiments, The result of following experiments will be present to use formula 10 to calculate it's amplitude of change. Where M_n and M_s respectively represents new experimental data and reference test data.

$$M = \frac{M_n - M_s}{|M_n|} \tag{10}$$

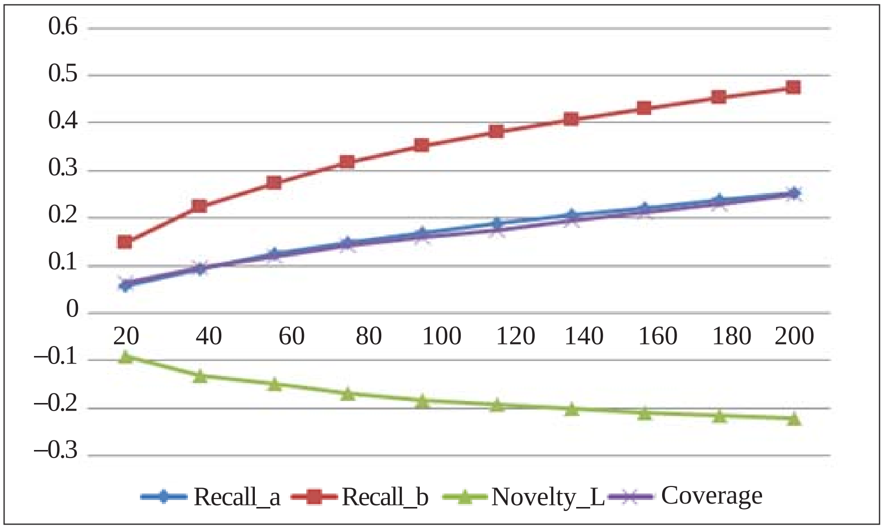

Figure 2. Novel recommendation of the traditional user-based CF plotted against the number of recommended items ($K = 20$)

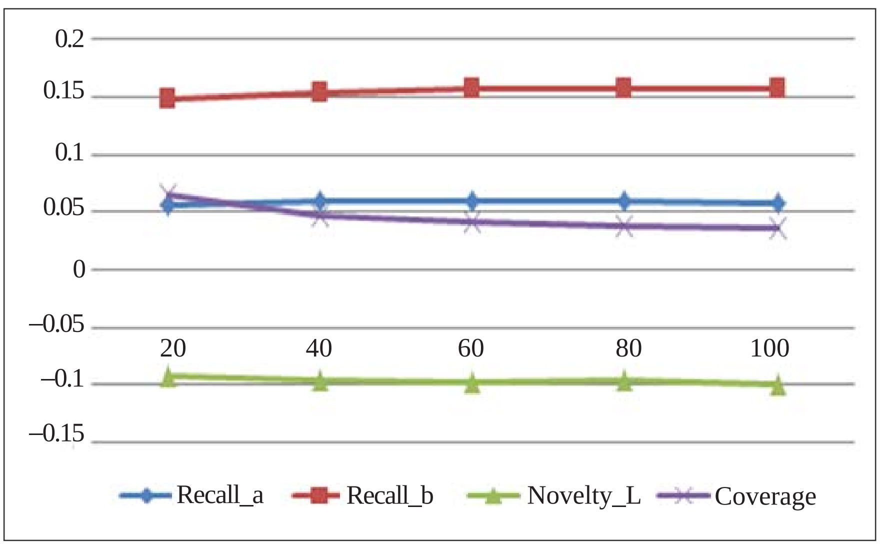

Figure 3. Novel recommendation of the traditional user-based CF plotted against the number of neighbor ($K = 20$)

5. Novel recommendation embedded Dissimilarity

There are many methods to calculate dissimilarity, such as Zhang.*Y* listed three way to measure dissimilarity: Set Difference, Geometric Distance and Distributional Similarity [36]. Measuring novelty of the item to the user can adopt average distance (formula 11) or minimum distance (formula 12) between the item to items in user's profile [33-37], we adopt minimum distance in this work. We amplified the threshold of $p(i \mid like, u)$ to $2N$. The results of the experiment are shown table 3, 4 and figure 4, 5.

$$dis(i, u) = min_{j \in porf_u} d(i, j) = min_{j \in porf_u} (1 - \text{cosine}(i, j)) \tag{11}$$

$$dis(i, u) = mean_{j \in porf_u} d(i, j) = mean_{j \in porf_u} (1 - \text{cosine}(i, j)) \tag{12}$$

N	Recall_a	Recall_b	Novelty	Avg_pop	Coverage
20	0.0570	0.1485	–0.0915	750.97	0.0644
40	0.0932	0.2245	–0.1313	687.00	0.0946
60	0.1254	0.2735	–0.1481	643.28	0.1186
80	0.1478	0.3176	–0.1698	609.37	0.1420
100	0.1684	0.3525	–0.1840	581.36	0.1605
120	0.1885	0.3799	–0.1914	557.30	0.1758
140	0.2062	0.4083	–0.2021	536.64	0.1938
160	0.2224	0.4319	–0.2094	518.17	0.2108
180	0.2380	0.4539	–0.2158	501.78	0.2294
200	0.2535	0.4752	–0.2217	486.69	0.2489

Table 1. Novel recommendation of the traditional user-based CF ($K = 20$)

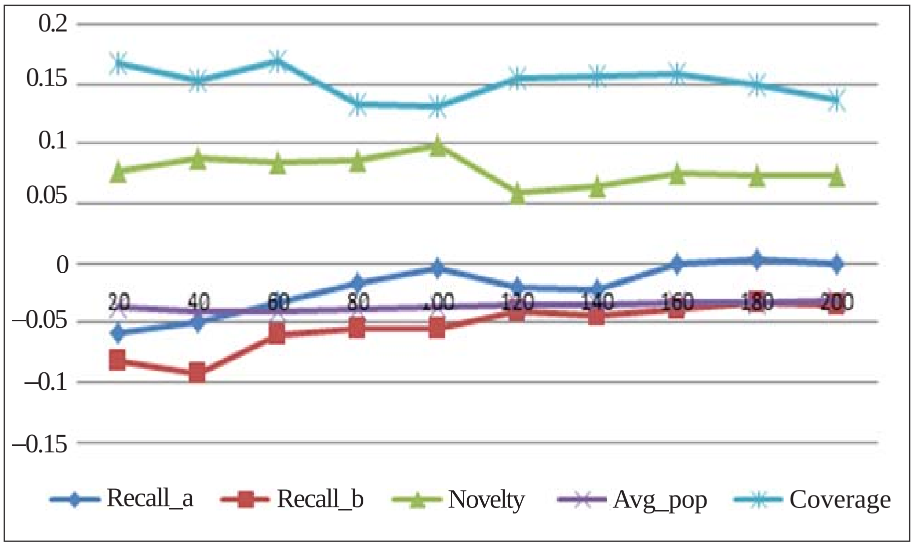

Figure 4. Novel recommendation of the algorithm embedded dissimilarity plotted against the number of recommended items ($K = 20$)

K	Recall_a	Recall_b	Novelty	Avg_pop	Coverage
20	0.0570	0.1485	–0.0915	750.97	0.0644
40	0.0591	0.1543	–0.0952	779.90	0.0474
60	0.0603	0.1582	–0.0978	790.35	0.0422
80	0.0605	0.1566	–0.0960	796.65	0.0382
100	0.0584	0.1567	–0.0984	800.46	0.0367

Table 2. Novel recommendation of the traditional user-based CF ($N = 20$)

$$score\,(i, u)_{\,i \in \ Top\,(2N,\, u)} = Predict\,(i, u) \times dis\,(i, u) \quad (13)$$

$$Recomm\,(N, u) = \{i \in I \mid Rank\,(i, score\,(i, u)) \leq N\} \quad (14)$$

The improvement of novelty is greater than the decline of the ability of predicting user's need when dissimilarity embedded in the algorithm, the coverage and the avg_pop get certain of improvement. Novelty has no improve and coverage quickly declines with the increase of the number of neighbor.

figure 6 shows novelty tend to be stable when the threshold of $p\,(i \mid like, u)$ more than 100 because $predict\,(i, u)$ quickly declines along with the drop in ranking and the change of dissimilarity does not affect the recommendation results.

6. Novel recommendation embedded unknown and dissimilarity

At present many researchers adopt popularity to measure the unknown of the item, the lower popularity of the item the less probability of it is known [38-40]. In this paper,

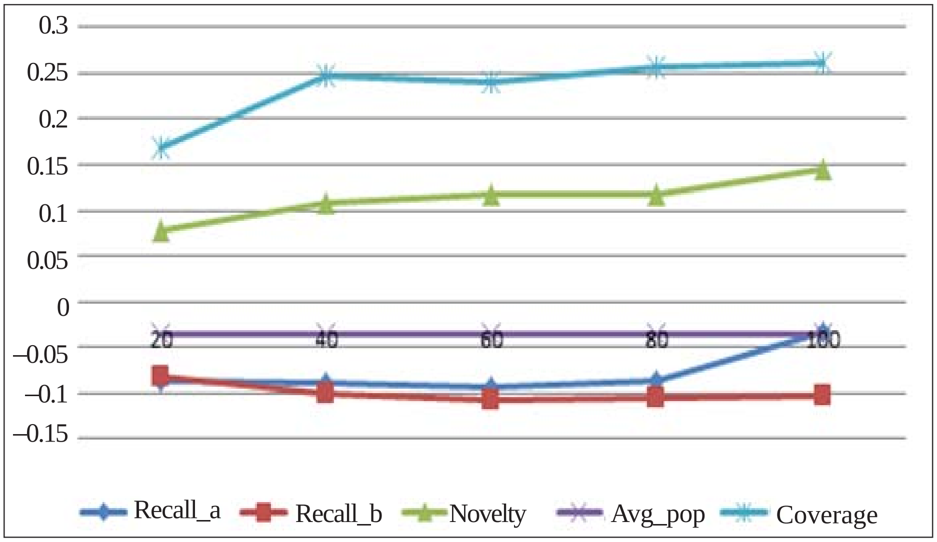

Figure 5. Novel recommendation of the algorithm embedded dissimilarity plotted against the number of neighbor (N = 20)

N	Recall_a	Recall_b	Novelty	Avg_pop	Coverage
20	–0.0577	–0.0815	0.0776	–0.0362	0.1677
40	–0.0490	–0.0922	0.0879	–0.0400	0.1533
60	–0.0331	–0.0607	0.0849	–0.0399	0.1686
80	–0.0169	–0.0545	0.0872	–0.0390	0.1331
100	–0.0048	–0.0545	0.0995	–0.0371	0.1321
120	–0.0202	–0.0397	0.0590	–0.0355	0.1553
140	–0.0228	–0.0429	0.0638	–0.0346	0.1569
160	–0.0009	–0.0377	0.0759	–0.0330	0.1584
180	0.0025	–0.0335	0.0728	–0.0323	0.1504
200	–0.0008	–0.0354	0.0744	–0.0310	0.1374

Table 3. Novel recommendation of algorithm embedded dissimilarity (K = 20)

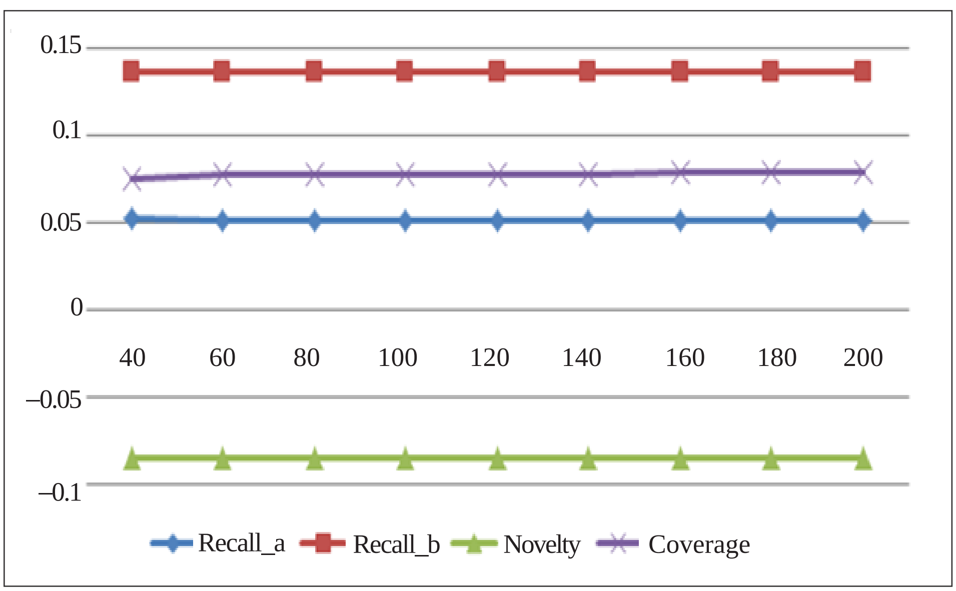

Figure 6. Novel recommendation of the algorithm embedded dissimilarity plotted against the threshold of $p(i \mid like, u)$ (N = 20, K = 20)

K	Recall_a	Recall_b	Novelty	Avg_pop	Coverage
20	–0.0577	–0.0815	0.0776	–0.0362	0.1677
40	–0.0527	–0.1011	0.1071	–0.0367	0.2468
60	–0.0507	–0.1081	0.1166	–0.0366	0.2417
80	–0.0488	–0.1060	0.1177	–0.0366	0.2565
100	–0.0401	–0.1040	0.1463	–0.0367	0.2616

Table 4. Novel recommendation of algorithm embedded dissimilarity ($N = 20$)

popularity pop_i is defined as the number of item rated, $p\,(i \mid unknown, u)$ is calculated using formula 15. Still using the above ideas, we amplified the threshold of $p\,(i \mid like, u)$ to $2N$, and then embedded the dissimilarity and the unknown into the calculation of recommended score (formula 16). The results of the experiment are shown in table 5 and figure 7.

$$p\,(i \mid unknown, u) = -\log\,(1 + pop_i) \tag{15}$$

$$score\,(i, u)_{\,i \in Top(2N,u)} = Predict\,(i, u) \times dis\,(i, u) \times (i \mid unknown, u) \tag{16}$$

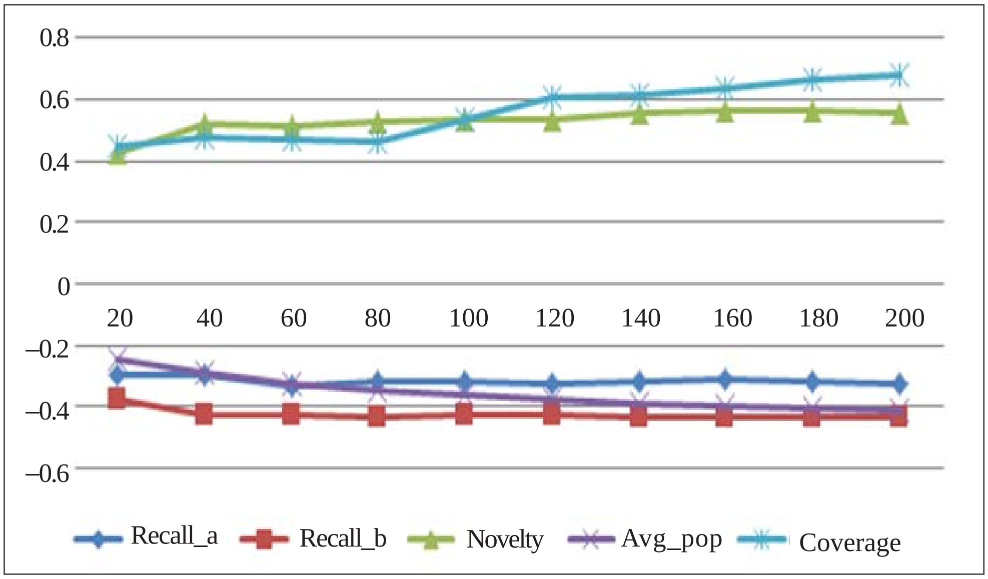

Figure 7. Novel recommendation of the algorithm embedded dissimilarity and unknown plotted against the number of recommended items ($K = 20$)

N	Recall_a	Recall_b	Novelty	Avg_pop	Coverage
20	–0.2965	–0.3771	0.4284	–0.2444	0.4519
40	–0.2972	–0.4285	0.5217	–0.2918	0.4757
60	–0.3301	–0.4282	0.5111	–0.3233	0.4680
80	–0.3207	–0.4304	0.5253	–0.3458	0.4662
100	–0.3153	–0.4287	0.5321	–0.3632	0.5321
120	–0.3220	–0.4285	0.5334	–0.3764	0.6109
140	–0.3177	–0.4347	0.5542	–0.3874	0.6146
160	–0.3112	–0.4353	0.5669	–0.3961	0.6357
180	–0.3143	–0.4329	0.5635	–0.4032	0.6665
200	–0.3254	–0.4348	0.5598	–0.4092	0.6770

Table 5. Novel recommendation of algorithm embedded dissimilarity and unknown ($K= 20$)

After embedded unknown into the algorithm, the novelty of recommendation result gets obvious improvement, the coverage and the avg_pop are also greatly improved, but the accuracy drops relatively large. Using popularity as a measure of user unknown although has certain rationality, but is too rough. From the theory of product life cycle can be seen, most users are only known and accepted after the products entering the mature stage, nevertheless, in the growth stage of the product is perceived slow. Therefore, we will consider time factor into the calculation

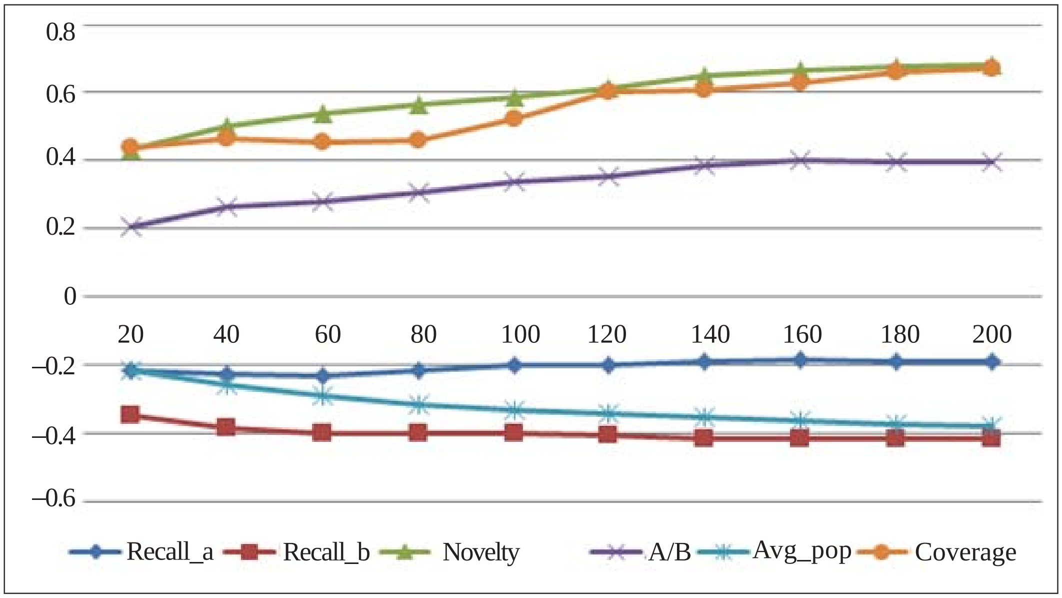

Figure 8. Novel recommendation of the algorithm embedded dissimilarity and time-popularity plotted against the number of recommended items (K = 20)

N	Recall_a	Recall_b	Novelty	Avg_pop	Coverage
20	–0.2175	–0.3508	0.4339	–0.2170	0.4363
40	–0.2264	–0.3880	0.5027	–0.2614	0.4672
60	–0.2329	–0.3996	0.5409	–0.2925	0.4553
80	–0.2192	–0.4033	0.5636	–0.3148	0.4620
100	–0.2013	–0.4034	0.5880	–0.3319	0.5252
120	–0.1989	–0.4085	0.6149	–0.3453	0.6041
140	–0.1925	–0.4178	0.6477	–0.3567	0.6084
160	–0.1848	–0.4191	0.6676	–0.3657	0.6300
180	–0.1895	–0.4199	0.6738	–0.3729	0.6587
200	–0.1913	–0.4200	0.6820	–0.3795	0.6697

Table 6. Novel recommendation of the algorithm embedded dissimilarity and time-popularity (K=20)

N	Recall_a	Recall_b	Novelty	Avg_pop	Coverage
20	0.0947	–0.7825	1.3301	–0.7096	3.4425
40	0.0300	–0.7416	1.2894	–0.6955	3.4693
60	–0.0518	–0.7133	1.2735	–0.6831	3.3702
80	–0.0392	–0.6880	1.2527	–0.6714	3.1275
100	–0.0523	–0.6633	1.2223	–0.6605	2.9526
120	–0.0663	–0.6423	1.2095	–0.6497	2.8294
140	–0.0718	–0.6294	1.1979	–0.6392	2.6465
160	–0.0710	–0.6147	1.1920	–0.6291	2.4924
180	–0.0798	–0.5948	1.1627	–0.6193	2.3191
200	–0.0828	–0.5795	1.1470	–0.6093	2.1394

Table 7. Novel recommendation of optimization calculation using maximum novelty as target ($K = 20$)

of popularity, increase the weight of the past rated (formula 17), this can be a good judge of mature popular goods. We call pop_{iT} as Time-popularity. The results of the experiment are shown in table 7 and figure 8.

$$pop_{iT} = \sum_{t \le T} (1 + \alpha\,(T - t)) \mid \{r(u, i, t) \in R\} \qquad (17)$$

As shown in the above experiment results, accuracy and novelty have been significantly improved. With the recommended amount increased, the degree of the increase is more obvious, which shows accuracy and novelty maybe directly proportional to the threshold of $p\,(i \mid like, u)$. We used maximum novelty as target, and the threshold of $p\,(i \mid like, u)$ as variable to optimization calculate. The results of the experiment are shown in table 7 and figure 9.

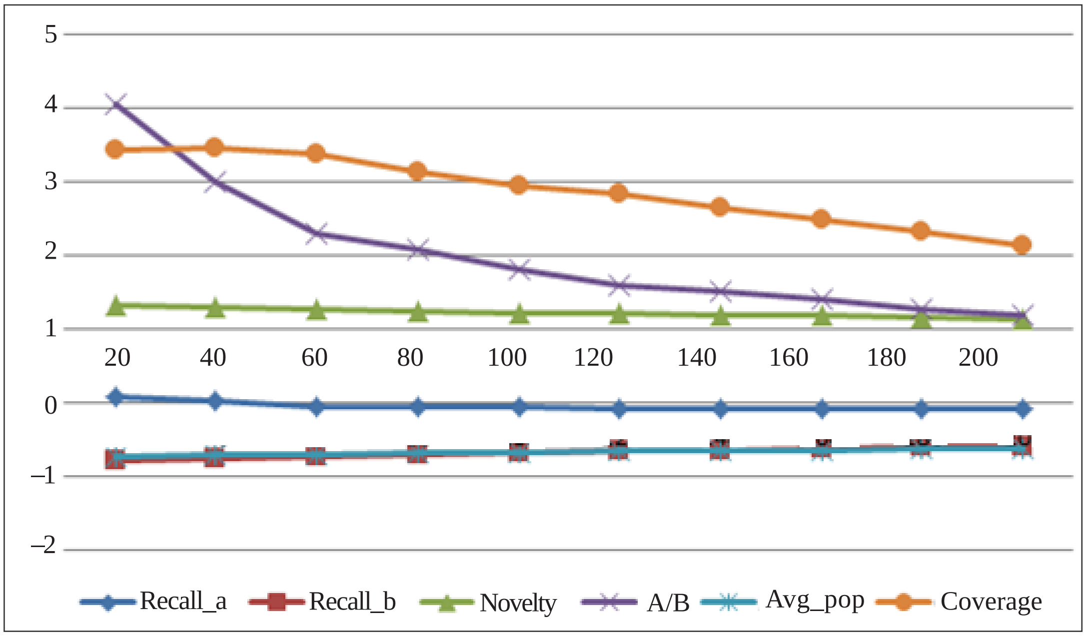

Figure 9. Novel recommendation of optimization calculation using maximum novelty as target plotted against the threshold of $p\,(i \mid like, u)$ $(K = 20)$

After optimization calculation, the accuracy of predicting user's future needs is above the traditional algorithm, novelty is from negative to positive. The result of recommendation is eliminated greatly user known item, coverage and avg_pop are significantly improved. Therefore, the proposed algorithm embedded dissimilarity and time-popularity can improve recommendation quality, so as to promote the user's satisfaction. We also experimented against another variable K (the number of neighbor), despite accuracy and novelty have a certain degree of improvement, but the decline of coverage is too obvious.

7. Conclusion

In this paper, we defined novelty of item from the perspective of the user, and designed the corresponding offline experiment scheme and evaluation metrics. The dissimilarity factor and the time-popularity factor were blended in the traditional collaborative filtering algorithm. The ability of predicting the future needs of the user's and coverage of recommended list was improved, and the ability of recommended long tail items was also enhanced.

We also studied the analysis of the item. Although time-popularity can better identify outdated popular items, but it still cannot accurately present the division of product life cycle, and this will be our future research. For different users, the same item's novel perception is not the same. For example, the user is often the leader and innovator in his/her very interested fields, conversely is the recipient. In the future we will also analyze the user types of accepting new items, so as to make the recommendation results more relevant.

References

[1] Cho, Y. H., Kim, J. K. (2004). Application of Web usage mining and product taxonomy to collaborative recommend ations in e-commerce, *Expert Systems with Applications*, 26 (2) 233–246.

[2] Lihua, W., Lu, L., Jing, L., Zongyong, L. (2005). Modeling user multiple interests by an improved GCS approach, *Expert Systems with Applications*, 29 (4) 757–767.

[3] Sarwar, B., Karypis, G., Konstan, J., Riedl, J. (2001). Item-based collaborative ûltering recommendation algorithms,*In*: Proceedings of the International World Wide Web Conference (WWW10), Hong Kong. p. 1-5.

[4] Shih, Y. Y., Liu, D. R. (2008). Product recommendation approaches: Collaborative ûltering via customer lifetime value and customer demands, *Expert Systems with Applications*, 35 (1) 350–360.

[5] Esslimani, Ilham, Brun, Armelle, Boyer, Anne Source. (2008).Behavioral similarities for collaborative recommendations. *Journal of Digital Information Management*, 6 (6) 442-448.

[6] Hill, W., Stead, L., Rosenstein, M., Furnas, G. (1995). Recommending and evaluating choices in avirtual community of use. *In*: Proc.of the Conf.on Human Factors in Computing Systems. New York:ACM Press. p. 194-201.

[7] Resnick, P., Iakovou, N., Sushak, M., Bergstrom, P., Riedl, J. (1995). GroupLens: An open architecture for collaborative fitering of netnews. *In*: Proc.of the Conf.on Human Factors in Computing Systems. New York:ACM Press. p. 175-186.

[8] Ting-Peng Liang, Yung-Fang Yang, Deng-Neng Chen, Yi-Cheng Ku. (2008). A semantic-expansion approach to personalized knowledge recommendation. *Decision Support Systems*, 45, 401–412.

[9] Zhang, YC., Blattner, M., Yu, YK. (2007). Heat conduction process on community networks as a recommendation model. *Phys Rev Lett*, 99 (15) 154-160.

[10] Mass diffusion, Tao Zhou, etc. (2010). Solving the apparent diversity-accuracy dilemma of recommender systems. *In*: Proceedings of the National Academy of Sciences of the USA-PNAS, 107 (10) 4511-4515.

[11] Shilad Wieland Sen, Jesse Vig, John Riedl. (2009). Tagommenders:Connecting Users to Items through Tags. The International World Wide Web Conference Committee (IW3C2). p. 671-680.

[12] Schafer, J. B., Konstan, J. A., Riedl, J. (2001). E-Commerce Recommendation Applications. *Data Min. Knowl. Discov.* 5 (1/2) 115-153.

[13] Liu, D. -R., Liou, C. -H. (2011). Mobile commerce product recommendations based on hybrid multiple channels, *Electronic Commerce Research andApplications*, 10 (1) 94-104.

[14] Abel, F., Bittencourt, I. I., Costa, E., Henze, N., Krause, D.,Vassileva, J. (2010). Recommendations in online discussion forums for E-learning systems, *IEEE Transactions on Learning Technologies*, 3 (2) 165-76.

[15] Gao, M., Liu, K., Wu, Z. -F. (2010). Personalisation in web computing and informatics: theories, techniques, applications, and future research, *Information Systems Frontiers*, 12 (5) 607-29. *Intelligent Agent Techology*. p. 508-515.

[16] Goldberg, D., Nichols,D., Oki, B. M., Terry, D. (1992). Using collaborative ûltering to weave an information tapestry. *Communications of the ACM*, 35 (12) 61–70.

[17] Herlocker, J., Konstan, J. A., Riedl, J. (2002). An empirical analysis of design choices in neighborhood-based collaborative ûltering algorithms. *Information retrieval*, 5 (4) 287–310.

[18] oscher, A. T¨., Jahrer, M., Legenstein, R. (2008). Improved neighborhood-based algorithms for large-scale recommender systems. *In*: Proceedings of Netûix ’08. ACM. p.451-456.

[19] McLaughlin, M., Herlocker, J. (2004). A collaborative filtering algorithm and evaluation metric that accurately model the user experience. *In*: Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (Sheffield, UK, 2004), ACM Press. p. 329–336.

[20] Torres, R., McNee, S. M., Abel, M., Konstan, J. A., Riedl, J. (2004). Enhancing digital libraries with TechLens *In*: Proc. of ACM/IEEE JCDL, *ACM Press*. p. 228-236.

[21] Swearingen, K., Sinha, R. (2001). Beyond Algorithms: An HCI Perspective on Recommender Systems. *In*: ACM SIGIR Workshop on Recommender Systems. p. 185-195.

[22] Tomoko Murakami, Koichiro Mori, Ryohei Orihara. (2008). Metrics for evaluating the serendipity of recommendation lists. New Frontiers in Artificial Intelligence, LNAI 4914. p. 40–46.

[23] Adomavicius, G., Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transaction on Knowledge and Data Engineering*, 17 (6) 734–749.

[24] Herlocker, J. L., Konstan, J. A., Terveen, L. G., Riedl, J. T. (2004). Evaluating collaborative ûltering recommender systems. *ACM Trans. Inf. Syst*. 22 (1) 5–53.

[25] McNee, S. M., Riedl, J., Konstan, J. A. (2006). Being accurate is not enough: How accuracy metrics have hurt recommender systems. In ACM SIGCHI Conference on Human Factors in Computing Systems (Montr´ eal, Qu´ ebec, Canada, 22/04/2006 2006), ACM, 1097–1101.

[26] Shani, G., Chickering, D. M., Meek, C. (2008). Mining recommendations from the web. *In*: RecSys ’08: Proceedings of the 2008 ACM Conference on Recommender Systems. p. 35–42.

[27] Ziegler, C. N., McNee, S. M., Konstan, J. A., Lausen, G. (2005). Improving recommendation lists through topic diversification. *In*: WWW 05: Proceedings of the 14th international conference on World Wide Web. ACM, New York, NY, USA. p. 22–32.

[28] Celma, O., Herrera, P. (2008). A new approach to evaluating novel recommendations. *In*: RecSys ’08:Proceedings of the 2008 ACM conference on Recommender systems. p. 179–186.

[29]Jinoh Oh, SunPark, etc. (2011). Novel Recommendation based on Personal Popularity Tendency. 2011 11th IEEE International Conference on Data Mining. p. 591-608.

[30] Herlocker, J., et al. (2004). Evaluating collaborative filtering recommender systems’. *ACM Transactions on Information Systems*, 22 (1) 5–53.

[31] Hijikata, Y., Shimizu, T., Nishida, S.(2009). Discovery-oriented collaborative filtering for improving user satisfaction. IUI'09 Proceedings of the 14th international conference on Intelligent user interfaces. p. 67-76.

[32] Liang Zhang. (2013). The Definition of Novelty in Recommendation System. *Journal of Engineering Science and Technology Review,* 6 (3) 141- 145.

[33] Weng, L. T., Xu, Y., Li, Y., Nayak, R. (2007). Improving recommendation novelty based on topic taxonomy. Web Intelligence and Intelligent Agent Technology Workshops, 2007 IEEE/WIC/ACM International Conferences on. p. 115–118.

[34] Mi Zhang, Neil Hurley. (2009). Novel Item Recommendation by User Profile Partitioning. IEEE/WIC/ACM International Joint Conference on Web Intelligence.

[35] Zeng, W., Shang, M. -S., Zhang, Q. -M., Lü, L., Zhou, T. (2010). Can dissimilar users contribute to accuracy and diversity of personalized recommendation? *International Journal of Modern Physics,* 21 (10) 1217–1227.

[36] Zhang, Y., Callan, J., Minka, T. (2002). Novelty and redundancy detection in adaptive filtering. *In*: SIGIR '02: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and development in information retrieval. p. 81–88.

[37] HURLEY, N., Z HANG, M. (2008). Analysis of methods for novel case selection. *In*: Proceedings of 20th IEEE International Conference on Tools with Artificial Intelligence. p. 217–224.

[38] Lü, L., et al. (2012). Recommender systems, *Physics Reports*, 519 (1) 1-49.

[39] Castells, P., Vargas, S., Wang, J. (2011). Novelty and diversity metrics for recommender systems: choice, discovery and relevance, *In*: Proceedings of International Workshop on Diversity in Document Retrieval, DDR, ACMPress, NewYork. p. 29–37.

[40] Saúl Vargas, Pablo Castells. (2011). Rank and Relevance in Novelty and Diversity Metrics for Recommender Systems. RecSys'11, October. p. 23–27.