Research Article

The Definition of Novelty in Recommendation System

Liang Zhang[1,2,*]

[1]*School of management Xiamen university Xiamen China*
[2]*Guizhou normal university Guiyang, China*

___

Abstract

With the development of information technology and application of the Internet, People gradually entered the time of information overload from information scarcity. User 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. Novelty is one of the important metrics of customer satisfaction. There is an increasing realization in the Recommender Systems (RS) field that novelty is fundamental qualities of recommendation effectiveness and added-value. This paper combed research results about definition and algorithm of novel recommendation, and starting from the meaning of "novel", defined novelty of item in recommendation system. Experiment proved using the definition of novelty to recommend can effectively recognize the item that the user is familiar with and ensure certain accuracy.

Keywords: Novelty; Novel Recommendation; Definition; Algorithm

___

1. Introduction

With the development of information technology and application of the Internet, People gradually entered the time of information overload from information scarcity. Recommendation system creates user preferences model through the analysis of the historical behavior of the user, and recommends information meeting the user's demands. The Tapestry is one of the earliest recommendation system. So far, there are many recommendation algorithms in recommendation system, such as Collaborative Filtering [2-4], Content-based Recommendation[5-6]，Latent Factor Model[7], Heat Conduction[8], Mass diffusion [9]，Tag-based Filtering[10], Visualization Techniques[11]and so on.

Starting from the birth of recommendation system, more than 90% relevant papers discussed about the accuracy metrics. Using MAE and RMSE calculates the accuracy of prediction score, and using precison, recall and F-metric evaluates the accuracy of Top-N recommendation. Some recommenders produce highly accurate results that are still useless in practice, e.g., suggesting chewing gum to customers in grocery stores. Though being highly accurate, their recommending appears far too obvious and of little help to the shopper. More frequently than not, recommendation lists contain similar items in them. The Item-Item collaborative filtering algorithm can trap users in a 'similarity hole' [12].

One major issue with accuracy metrics is their inability to capture the broader aspects of user satisfaction, hiding several blatant flaws in existing systems[13]. Torres suggested that differences in language and cultural background influenced user satisfaction[14]. Swearingern and Sinha examined how usefulness, novelty and usability are related to user satisfaction by means of a questionnaire survey and reported that they are significantly correlated[15]. User 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 [16]. In 2011 ACM recommendation systems Conference, held a special seminar to discuss recommended diversity and novelty. Nowadays novel recommendation is one of the hottest topics in recommendation system field.

The remainder of this paper is structured as follows. Section 2 review related work on novelty in recommendation system. Section 3 presents the definition of novelty in recommendation system according to the meaning of word 'novel' and Section 4 describes our experimental design and evaluation. We conclude in Section 5 with final remarks and our plan for future research.

2. Related work

As early as 1999, Baeza-Yates and Ribiero-Neto briefly discussed the novelty in information retrieval, the novelty of a retrieval set has been defined with respect to the end-user as the proportion of known and unknown relevant items in the recommended list[17]. That is, given $L \subseteq R$ is the set of items in R that the user likes, L can be partitioned as $L = L_k \cup L_U$ into those items, L_k is already known items to the user and L_U is unknown items to the uer. Then the novelty is

$$Novelty(R) = \frac{|L_U|}{|L|}$$

$$L = L_k \cup L_U$$

Novel recommendations are recommendations for items that the user did not know about [18]. In the O'scar Celma's doctoral dissertation"Music Recommendation and Discovery in the Long Tail"，the novelty for a given user u defined the ratio of unknown items in the list of top–N recommended items.

$$Novelty(u) = \frac{\sum_{i \in R}(1 - knows(u,i))}{|R|}$$

Where R is the set of top–N recommended items, and knows(u, i) is a binary function that returns 1 if user u already knows item i, and 0 otherwise. While we can obviously measure novelty in a user study, by asking users whether they were already familiar with a recommended item [19-20].

Recommender cannot accurately judge item whether the user knows if the user never interacted with it. So it is reasonable to assume the less popular items, the more novel to users[21]. The simplest way to quantify novelty is the average popularity of top–N recommended items. The popularity of item is defined the self-information of the item.

$$Novelty(R) = \frac{1}{|R|}\sum_{i \in R} POP_i = \frac{1}{|R|}\sum_{i \in R} -\log(p(i))$$

Where $p(i)$ is the probability that I is observed. Depending on the availability and type of observation data, Pablo Castells distinguished three main categories of user-item relationships：choice, discovery and relevance. The essence of measurement method is based on popularity, but joined the utility function to measure the user's satisfaction[22]. Saúl Vargas and Pablo Castells also considered the rank of recommended list[23].

Whether user survey or based popularity only considers the possibility that the user known the item, but does not reflect the "new" of item. Content-based recommendation algorithm, some researchers suggested that novel item should be similar to previously delivered relevant documents in the sense of having the same topic, but also dissimilar to the previously delivered documents in the sense of containing new information[24-25]. Zhang,Y suggested three measures to capture redundancy document: Set Difference，Geometric Distance and Distributional Similarity.

$$R(d_t|d_i) = \|Set(d_t) \cap \overline{Set(d_i)}\|$$

$$R(d_t|d_i) = \cos(d_t, d_i) = \frac{\sum_{k=1}^{n} w_k(d_t)w_k(d_i)}{\|d_t\|\|d_i\|}$$

$$R(d_t|d_i) = -\sum_{w_i} p(w_i|d_t)\log\frac{p(w_i|d_t)}{p(w_i|d_i)}$$

Where d_i represents a relevant document to a new document d_t, and $Set(.)$ is the set of keywords in the document. Formula (5) Calculates the number of new words in d_t with respect to d_i. If we represent document d as a vector $d = (w_1(d), w_2(d), ..., w_n(d))^T$, formula (6) calculates the cosine distance between d_t and d_i. In the language model approach, a document d is represented by a unigram word distribution θ_d. Formula (7) calculates Kullback-Leibler divergence between d_t and d_i, a distributional similarity measure, is one way to measure the redundancy of one document given another. But Kullback-Leibler divergence does not meet three conditions: negative, symmetry and triangle inequality.

Novel recommenddation have just recently attracted more and more attention from academia and industry. There is not a unified definition about novelty，therefore, novel recommendation algorithms are different with respect to the definition of novelty. The research most closely related to novelty or redundancy detection in adaptive information filtering is perhaps the First Story Detection task associated with Topic Detection and Tracking (TDT) research[27]. A TDT system monitors a stream of chronologically-ordered documents, usually news stories. The First Story Detection (FSD) task is defined as detecting the first story that discusses a previously-unknown event. FSD is an event-based task, whereas Information filtering tend to be more subject-oriented, recommender tend to be user-ended[24].

Novel recommendation simply understood as recommended items users don't know, 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 is 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[28]. Through such adjustment can increase the possibility of recommended unpopular items. Ziegleret al. [29]and Celma and Herrera [30] also give accuracy measures that take popularity into account.

$$\widetilde{R_{u,i}} = \log(\frac{|U|}{|U_i|})R_{u,i}$$

Where $R_{u,i}$ represents the prediction score for a user u and each item i in a list of recommendations, $|U|$ is the number of users and $|U_i|$ the number of users who liked item i. Jinoh, Oh proposes 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) [31].

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 [32].

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 dis-advantage of this approach is that the list of obvious items might be different for 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 [12]. Hijikata, Y suggeated that using the ratings of acquaintance calculates the probability that a user knows an unrated item [33].

Zhang, Y suggested that a system that delivers documents that are novel and relevant must identify documents that are similar to previously delivered relevant documents in the sense of having the same topic, but also dissimilar to the previously delivered documents in the sense

of containing new information[25]. Zhang, Y assumed that novelty/redundancy detection is performed on a stream of documents that are presumed to be relevant, and used three approached(Set Difference，Geometric Distance and Distributional Similarity) to redundancy detection.

Y. Yang and J. Z proposed a novel recommendation algorithm for document based on the principle of TF-IDF, used cosine distance to find relevant documents, calculated the weight vector difference of keywords to find novel documents[25]. The TF-IDF method can also be used recommendation for non-document and based-tag recommendation.

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 the user's range of tastes, 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[27].

3. 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 "refreshing–pleasantly new or different". Likewise, familiar (adj.) is defined as "well known or easily recognised". 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 information 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 $dis(i, pref_u)$ is dissimilarity between item i and the set of items in user's profile, the definition of novelty as follows:

$$Novelty(i,u) = p(i|unknown, u) \times dis(i, pref_u) \times p(i|like, u)$$

Recommendation system creates user preferences model through the analysis of the historical behavior of the user, and recommends information meeting the user's demands. Recommendation system is unable to determine whether the user known items outside user's profile. In thispaper, we use the popularity of item to measure the possibility of unknown.

$$p(i|unknown, u) = -\log(1 - pop_i)$$

Recommendation system creates user preferences model through the analysis of the historical behavior of the user. Assuming that normal user tastes are well reflected in the distance function, if $Prof_u$ is the set of items that a user likes, $d(i, j)$ is distance function. Representative distance functions include Pearson correlation coefficient, cosine similarity and Jaccard coefficient and so on. In this paper, we use cosine similarity to calculate dissimilarity.

$$dis(i, pref_u) = \min_{j \in Prof_u} d(i, j) = \min_{j \in Prof_u}(1 - cosine(i, j))$$

In recommender systems, generally assume that high rating value items are satisfied to the user. The item is more similar to these items and more high probability of it being liked. Estimation of satisfactory and dissimilarity are based similarity metrics, therefore, there is a paradox between satisfactory and dissimilarity. We use slope one predictor to predict score because the algorithm does not explicitly use similarity metrics. Supposing $\widetilde{R_{u_i}}$ is predicted score and $U(i, j)$ is the set of users who rated items i and j, the probability of the user u likes the item i is defined as:

$$p(i|unknown, u) = \begin{cases} 0, & x < \theta \\ \widetilde{R_{u_i}}, & x \geq \theta \end{cases}$$

$$\widetilde{R_{u_i}} = \frac{1}{|Prof_u|}\sum_{j \in Prof_u}\left(R_{uj} + \frac{\sum_{\alpha \in U(i,j)}(R_{\alpha i} - R_{\alpha j})}{|U(i,j)|}\right)$$

4. Experiment

4.1 Experimental design and evaluation

The best way to measure novelty is asking users whether they were already familiar with a recommended item[19-20],but it is impracticable in offline experiments.Therefor we split the data set on time, hide all the user ratings that occurred after a specific point in time. In addition, we hide some ratings that occurred prior to that time, simulating the items that the user is familiar with, but did not report ratings for. When recommending, the system is rewarded for each item that was recommended and rated after the split time, but would be punished for each item that was recommended but rated prior to the split time[26]. Due to the number of the user ratings that occurred after a specific point in time is not same as the number of the user ratings that occurred before a specific point in time, and each user is not different also, we evaluate the result as follows:

$$Novelty\_score = \frac{1}{|U|}\sum_{u \in U}\left(\frac{|Recomm_u \cap hide_{au}|}{|hide_{au}|} - \frac{|Recomm_u \cap hide_{bu}|}{|hide_{bu}|}\right)$$

Where $Recomm_u$ is set of recommended items to the user u, $hide_{au}$ is set of items which are rated after the specific point in time, $hide_{bu}$ is set of items which are rated before that time.

4.2 Experimental result

We have tested our definition of novelty in different metrics on two datasets with two baseline recommender methods. Datasets included MoviesLen 10M and Netflix. We use definition of novelty to recommend, and evaluation matrices included average popularity and coverage besides the aforementioned Novelty_score. The coverage defined as follows:

$$Coverage = \frac{|\cup_{u \in U} Recomm_u|}{|I|}$$

Where I is set of all items. Baseline recommender methods are random recommendation (RND) and the topN popularity recommendation (TPR). The aforementioned definition of novelty has three facors, First of all, we only considered the probability of items being liked, then successively considered dissimilarity and popularity. Table 1 shows experiment results on MoviesLen data set and Netflix Dataset. A denoted the proportion of recommended items which are rated after the specific point in time, and B denoted the proportion of recommended items which are rated before that time. PL denoted only considering the probability of items being liked(PL), L&D denoted considering PL and dissimilarity, L&D&P denoted considering three factors in definition of novelty. Fig.1 shows dependence of every metric on the number of recommended items.

Table 1 experiment results on MoviesLen and Netflix Dataset

MoviesLen Data Set						
method		A	B	Novelty_score	Avg_pop	coverage
RND		0.0362	0.0354	0.0008	173.4076	1
TPR		0.3277	0.5346	-0.2069	1423.8507	0.1313
Novlety	PL	0.0548	0.0278	0.0270	88.5442	0.2329
	L&D	0.0674	0.0336	0.0338	75.0284	0.2656
	L&D&P	0.0554	0.0123	0.0431	70.8179	0.2841

Netflix Data Set						
methods		A	B	Novelty_score	Avg_pop	coverage
RND		0.0210	0.0212	-0.0002	33.6353	1
TPR		0.3576	0.4168	-0.0592	459.6959	0.0494
Novlety	PL	0.0335	0.0302	0.0033	82.7118	0.6514
	L&D	0.0358	0.0163	0.0194	78.4744	0.7231
	L&D&P	0.0349	0.0108	0.0241	68.7087	0.7980

We first focus on the Novelty_score metric, the method RND cannot distinguish the item whether the user is familiar with, L&D&P can achieve better results. A first interesting observation is novelty_score of method TPR is negative, therefore it is reasonable to assume novelty is inversely proportional to the popularity.

Focusing on Method novelty, with more factors considered, the novelty_score is higher, avg_pop is lower and coverage is higher, so the definition of novelty can better find items that users are interested in future, and promote diversity of recommendation. To some extent, A and B are regarded as precision of recommendation. Focusing on A metric, Novelty was better than RND, so slope one predictor could better find user's preference, L&D was better than L&D&P, so it is rough that using traditional popularity of item measures the possibility of unknown. The traditional popularity does not consider time effect of popularity and User's sensitivity to popular.
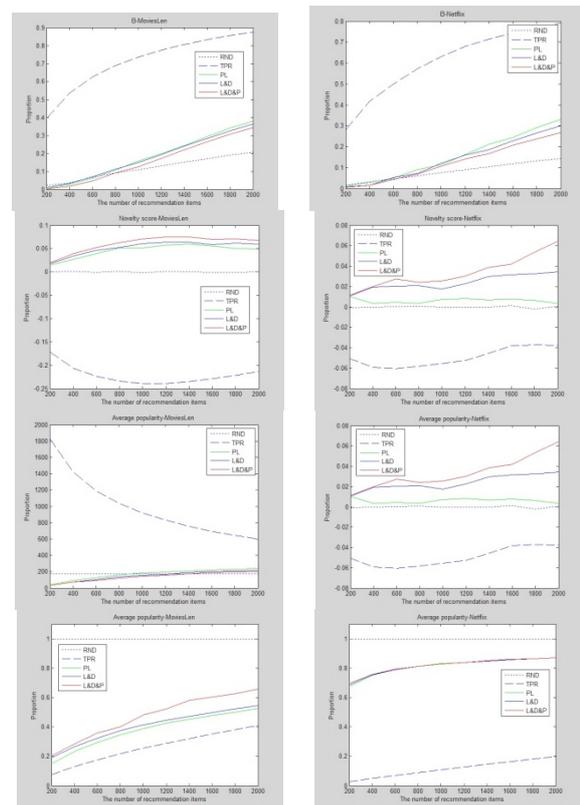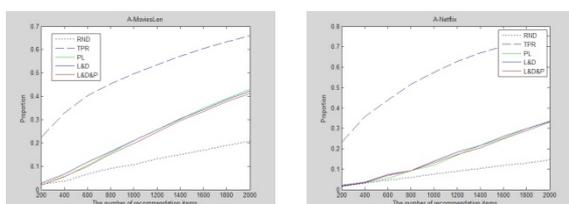

Fig.1 The dependence of every metric on the number of recommended items on two data sets

5. Conclusion and future study

The main purpose of the recommendation system is recommending user satisfied items based on user characteristics of demand. Novelty as one of the important metrics of customer satisfaction, become the focus of current research in recommendation system field. This paper combed research results about definition and algorithm of novel recommendation, and starting from the meaning of "novel", defined novelty of item in recommendation system. Experiment proved using the definition of novelty to recommend can effectively recognize the item that the user is familiar with and ensure certain accuracy.

It is rough that using popularity of item measures the possibility of unknown. The aforementioned methods do not consider the following 3 factors：

(1) Time effect of popularity. The aforementioned popularity is defined as self-information of the item. Obviously, popularity of the item is proportional to the existence time of the item. From the eye of the user, the popular item is sought after by users in the current period.

(2) User's sensitivity to popular. Due to different users in such aspects as social economy, personality, and communication with their respective characteristics, cognitive and adoption speed for new project is not the same. Some users knew the item just beginning to be popular, and some will be delay for a long time.

(3) User's activity. Many researchers generally believed that the more active user more unpopular items. We may assume that these active users is what Rodgers called "the innovator" Mentioned in the "diffusion of innovation" [43].

Unpopular items active users liked have greater reference value for other users.

In future research, We will consider time effect of popularity and User's sensitivity to popular in novelty recommendation algorithm. In our experiment, we only use slope one algorithm to predict rated value, the precision is lower and this algorithm is not commonly used. we will definition of novelty integrated into more classic recommendation algorithms.

6. Acknowledge

References

1. Goldberg, D., Nichols, D., Oki, B. M., & Terry, D.. Using ollaborative filtering to weave an information apestry. Communications f the ACM (1992), 35, 12.
2. Shardanand U, Maes P.Cocial information fltering:Algorithms for automating"Word of Mouth". In:Proc.of the Conf.on Human Factors in Computing Systems. New York:ACM Press,1995.pp.210-217
3. Hill W, Stead L, Rosenstein M, Furnas G. Recommending and evaluating choices in avirtual community of use. In:Proc.of the Conf.on Human Factors in Computing Systems. New York:ACM Press,1995.pp.194-201.
4. Resnick P, Iakovou N, Sushak M, Bergstrom P, Riedl J. GroupLens: An open architecture for collaborative fitering of netnews. In:Proc.of the Conf.on Human Factors in Computing Systems. New York:ACM Press,1995.pp.175-186.
5. Ting-Peng Liang, Yung-Fang Yang, Deng-Neng Chen, Yi-Cheng Ku. A semantic-expansion approach to personalized knowledge recommendation[J]. Decision Support Systems 45 (2008) pp.401–412.
6. Ming Li , Lu Liu, Chuan-Bo Li. An approach to expert recommendation based on fuzzy linguistic method and fuzzy text classification in knowledge management systems[J]. Expert Systems with Applications 38 (2011) pp.8586–8596.
7. Y. Koren, R. Bell, and C. Volinsky, Matrix factorization techniques for recommender systems, Computer 42 (2009), no. 8,pp. 30–37.
8. Zhang YC, Blattner M, Yu YK. Heat conduction process on community networks as a recommendation model [J]. Phys Rev Lett, 2007, 99(15): 15430121543054.
9. Zhou, Tao etc. Mass diffusion Solving the apparent diversity-accuracy dilemma of recommender systems. Proceedings of the National Academy of Sciences of the USA-PNAS 107(10). 2010.pp.4511-4515,
10. Shilad Wieland Sen, Jesse Vig, John Riedl. Tagommenders:Connecting Users to Items through Tags. ACM 978-1-60558-487-4/09/04.
11. M. Michalos, P. Tselenti and S. L. Nalmpantis. Visualization Techniques for Large Datasets Journal of Engineering Science and Technology Review 5 (1) (20 12) pp. 72 - 76
12. Herlocker, J., et al.: 2004, 'Evaluating collaborative filtering recommender systems'. ACM Transactions on Information Systems 22(1), pp.5–53.
13. McLaughlin, M., and Herlocker, J.A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In Proceedings of the 27th Annual In-ternational ACM SIGIR Conference on Research and De-velopment in Information Retrieval (Sheffield, UK, 2004),ACM Press, pp. 329–336.
14. Torres, R., McNee, S.M., Abel, M., Konstan, J.A., and Riedl, J. Enhancing digital libraries with TechLens +. In Proc. of ACM/IEEE JCDL 2004 , ACM Press (2004) pp.228-236.
15. Swearingen, K., Sinha, R.: Beyond Algorithms: An HCI Perspective on Recommender Systems. In: ACM SIGIR Workshop on Recommender Systems(2001)
16. Tomoko Murakami, Koichiro Mori, and Ryohei Orihara. Metrics for evaluating the serendipity of recommendation lists. New Frontiers in Artificial Intelligence, , 2008. pp. 40–46
17. BAEZA -YATES,R. AND RIBIERO -NETO, B. Modern Information Retrieval. Addison-Wesley Longman, Boston, Mass. 1999.
18. Konstan, J.A., McNee, S.M., Ziegler, C.N., Torres, R., Kapoor, N., Riedl, J.: Lessons on applying automated recommender systems to information-seeking tasks. In: AAAI (2006).
19. 19] Konstan, J.A., McNee, S.M., Ziegler, C.N., Torres, R., Kapoor, N., Riedl, J.: Lessons on applying automated recommender systems to information-seeking tasks. In: AAAI (2006).
20. Jones, N., Pu, P.: User technology adoption issues in recommender systems. In: Networking and Electronic Conference (2007).
21. L. Lü, et al., Recommender systems, Physics Reports(2012), doi:10.1016/j.physrep. 2012.02.06.
22. P. Castells, S. Vargas, J. Wang, Novelty and diversity metrics for recommender systems: choice, discovery and relevance, in: Proceedings of International Workshop on Diversity in Document Retrieval,DDR,ACMPress,NewYork,2011,pp.29–37.
23. Saúl Vargas and Pablo Castells. Rank and Relevance in Novelty and Diversity Metrics for Recommender Systems. RecSys'11, October 23–27, 2011, Chicago, Illinois, USA. Copyright 2011 ACM 978-1-4503-0683-6/11/10.
24. Zhang, Y., Callan, J., Minka, T.: 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, pp. 81–88. ACM, New York, NY, USA (2002). DOI http://doi.acm.org/10.1145/564376.564393 ）
25. Y. Yang and J. Z. Li. Interest-based recommendation in digital library. Journal of Computer Science, 1(1) 2005. pp.40–46,.
26. Francesco Ricci. Lior Rokach. Bracha Shapira. Paul B. Kantor. recommender system handbook. Springer New York Dordrecht Heidelberg London. 2011
27. J. Allan, J. Carbonell, G. Doddington, J. Yamron, and Y. Yang. Topic detection and tracking pilot study. In Topic Detection and Tracking Workshop Report. 2001.
28. Shani, G., Chickering, D.M., Meek, C.: Mining recommendations from the web. In: RecSys '08: Proceedings of the 2008 ACM Conference on Recommender Systems.2008. pp. 35–42
29. Ziegler, C.N., McNee, S.M., Konstan, J.A., Lausen, G.: Improving recommendation lists through topic diversification. In: WWW 05: Proceedings of the 14th international conference on World Wide Web, ACM, New York, NY, USA (2005). pp. 22–32.
30. Celma, O., Herrera, P.: A new approach to evaluating novel recommendations. In: RecSys '08:Proceedings of the 2008 ACM conference on Recommender systems, ACM, New York, NY, USA (2008). pp. 179–186.
31. Jinoh Oh, SunPark, etc. Novel Recommendation based on Personal Popularity Tendency. 11th IEEE International Conference on Data Mining. 2011
32. W. Zeng, M.-S. Shang, Q.-M. Zhang, L. Lü, T. Zhou, Can dissimilar users contribute to accuracy and diversity of personalized recommendation? International Journal of Modern Physics C21(2010) pp.1217–1227
33. Hijikata, Y., Shimizu, T., & Nishida, S. Discovery-oriented collaborative filtering for improving user satisfaction. IUI (2009), pp. 67-76.