

RESEARCH ARTICLE

Promoting Cold-Start Items in Recommender Systems

Jin-Hu Liu^{1,2}, Tao Zhou^{1,2*}, Zi-Ke Zhang³, Zimo Yang^{1,4}, Chuang Liu³, Wei-Min Li^{1,5}



CrossMark
click for updates

OPEN ACCESS

Citation: Liu J-H, Zhou T, Zhang Z-K, Yang Z, Liu C, et al. (2014) Promoting Cold-Start Items in Recommender Systems. PLoS ONE 9(12): e113457. doi:10.1371/journal.pone.0113457

Editor: Luis A. Nunes Amaral, Northwestern University, United States of America

Received: April 17, 2014

Accepted: October 28, 2014

Published: December 5, 2014

Copyright: © 2014 Liu et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: This work was partially supported by the National Natural Science Foundation of China under grant nos. 11222543 and 61433014, and the Special Project of Sichuan Youth Science and Technology Innovation Research Team under grant no. 2013TD0006. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: In particular, the authors declare that no competing interests exist. There are two authors (Zimo Yang and Wei-Min Li) employed by a commercial company. They are both master students studying in the Web Sciences Center. They started the current work while at the university, and finished some minor computations and revisions while employed by the company. These two companies (Alibaba Group and Beijing Baifendian Information Technology Co., Ltd.) both encourage their employees to produce high-quality academic articles but never give any financial supports to the lab, and they will not apply any patents on related results, and the anonymous data have been uploaded. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

1. Web Sciences Center, University of Electronic Science and Technology of China, Chengdu, People's Republic of China, **2.** Big Data Research Center, University of Electronic Science and Technology of China, Chengdu, People's Republic of China, **3.** Alibaba Research Center for Complexity Sciences, Hangzhou Normal University, Hangzhou, People's Republic of China, **4.** Alibaba Group, Hangzhou, People's Republic of China, **5.** Beijing Baifendian Information Technology Co., Ltd., Beijing, People's Republic of China

*zhutou@ustc.edu

Abstract

As one of the major challenges, cold-start problem plagues nearly all recommender systems. In particular, new items will be overlooked, impeding the development of new products online. Given limited resources, how to utilize the knowledge of recommender systems and design efficient marketing strategy for new items is extremely important. In this paper, we convert this ticklish issue into a clear mathematical problem based on a bipartite network representation. Under the most widely used algorithm in real e-commerce recommender systems, the so-called item-based collaborative filtering, we show that to simply push new items to active users is not a good strategy. Interestingly, experiments on real recommender systems indicate that to connect new items with some less active users will statistically yield better performance, namely, these new items will have more chance to appear in other users' recommendation lists. Further analysis suggests that the disassortative nature of recommender systems contributes to such observation. In a word, getting in-depth understanding on recommender systems could pave the way for the owners to popularize their cold-start products with low costs.

Introduction

Thanks to the blazing development of Internet, e-commerce has flourished over the past decades. With the online buy-and-sell platforms getting increasingly more available products (e.g., more than a billion products in *taobao.com*), shopping online has become a fashionable style of living and more people choose to purchase on the Internet rather than go to stores. E-commerce makes our life

much more convenient, meanwhile, it throws us into a dilemma of information overloads. Facing millions of items online, finding out favourites is rather difficult. As an effective information filtering tool, recommender system is thus of particular significance nowadays [1, 2]. In fact, it has already made considerable contributions to the socioeconomic fields in the past decade. For example, 60% of DVDs rented by *Netflix* are selected based on personalized recommendations, and about a half of sales in *Amazon* are brought by recommendations [2]. Consequently, recommender systems have received huge attentions from both physicists and computer scientists, and many advanced recommendation algorithms are proposed recently, including collaborative filtering [3–8], content-based analysis [9–11], dimensionality reduction techniques [12–14], diffusion-based methods [15–22], and so on.

One long-standing challenge, called cold-start problem, has plagued almost all recommender systems. Namely, when new users or items enter the system, there is usually insufficient information to produce reasonable recommendation [23]. Considering this fact, several potential solutions have been raised. The additional content information [23–26], tagging information [27–29] and cross-domain information [30] can be used to marginally relieve this problem, but they don't work in a purely cold-start setting, where no information is available to form any basis for recommendations. Furthermore, improving diversity and novelty of recommended lists can help new items be pushed out [19, 31, 32].

Practically speaking, as a holder of the recommender system, one can ask for extra information to generate initial profiles on users or items [24], or probe users' preferences by pushing to them some carefully selected items according to complicated algorithms [33]. Both methods are costly and risky. In contrast, an owner would like to popularize his/her new items. An improper method, called "shilling attacks", injects a number of mendacious users into the system to raise predicted ratings of new items, and thus enhances the possibility of these new items to appear in the recommendation lists [34, 35]. But, it is easily to be detected [36–38]. Furthermore, as a wide-spreading market strategy, advertisements are generally preferred and become more and more prosperous [39]. However, to popularize new items costs a lot and imposes an unbearable financial burden for small businesses [40]. As mentioned above, how to promote new items under limited marketing resources is a nontrivial challenge and the knowledge of recommendation algorithm may be helpful. Putting aside operational details, if the marketing activities can bring some purchases of certain users, a smart marketing manager will carefully choose the target users so that these purchases can lead to more exposures in the recommendation lists afterwards.

Taking a stand as a marketing manager, in this paper, we focus on how to promote cold-start items by utilizing the knowledge of recommender systems. The main contributions are threefold: (i) We convert this ticklish problem into a clear mathematical model that ignores some insignificant details. (ii) We show that to push new items to active users, a straightforward strategy that will jump into our mind at the first time, is an unexpectedly poor-performed strategy. (iii) We propose a degree-based solution that outperforms some baseline methods.

Results

Recommendation can be considered as a variant of link prediction in bipartite networks [41] and thus the better understanding of network structures can in principle improve the quality of recommendations [42–45]. We denote a recommender system by a user-item bipartite network $G(U, O, E)$, where $U = \{U_1, U_2, \dots, U_n\}$ and $O = \{O_1, O_2, \dots, O_m\}$ are respectively the sets of users and items, and E is the set of links connecting users and items. Consequently, we use the adjacent matrix, A , to describe the user-item relations: if user U_i has purchased item O_α , $a_{i\alpha} = 1$, otherwise $a_{i\alpha} = 0$ (throughout this paper we use Latin and Greek letters, respectively, for user- and item-related indices). Figure 1(a) illustrates a small bipartite network that consists of eight users (gray squares) and eight items (blue circles). k_i , the degree of user U_i , is defined as the number of items linked to U_i . Analogously, the degree of item O_α , denoted by k_α , is the number of users connected to O_α . For example, as shown in Figure 1(a), $k_i = 3, k_j = 1$ and $k_\alpha = 2$. The user degree distribution $p_u(k)$, is the probability that the degree of a randomly selected user, is equal to k , and the survival function, $P_u(k)$, denotes the probability that the degree of a randomly selected user, is no less than k . The item degree distribution $p_o(k)$ and survival function $P_o(k)$ are defined in a similar way. Degree distribution reflects the network heterogeneity [46].

We consider two real data sets with anonymous users in this paper (datasets are free to download as **Dataset S1**), including (a) *Tmall.com* (TM): an open business-to-consumer (B2C) platform where enrolled businessmen can sell legal items to customers; (b) *Coo8.com* (Coo8): a well established online retailer mainly trading in electrical household appliances and a leading supplier to daily necessities. In order to avoid the isolate nodes in the data sets, each user has bought at least one item, and each item has been purchased at least once. Table 1 shows the basic statistics of the two data sets. Due to the different types of products, these networks have much different average item degrees. As shown in Figure 2, all degree distributions are heavy-tailed and the item degree distributions are generally more heterogenous than the corresponding user degree distributions. These observations complement previous empirical analyses on user-item bipartite networks [47–50].

The nearest neighbors' degree for user U_i , denoted by $d_{nn}^u(i)$, is defined as the average degree over all items connected to U_i [50]. For example, in Figure 1(a), $d_{nn}^u(i) = \frac{k_\alpha + k_\beta + k_\gamma}{k_i} = \frac{10}{3}$. Furthermore, the degree-dependent nearest neighbors' degree, $\overline{d_{nn}^u(k)}$ is the average nearest neighbors' degree over all users of degree k , namely $\overline{d_{nn}^u(k)} = \frac{1}{N_k} \sum_{k_i=k} d_{nn}^u(i)$, where N_k is the number of users with degree k . Corresponding concepts for items, $d_{nn}^o(\alpha)$ and $\overline{d_{nn}^o(k)}$ are defined in a similar way and thus omitted here. The degree-dependent nearest neighbors' degree is an appropriate index to characterize the network assortativity [51]. As shown in Figure 2, both the two networks are disassortative.

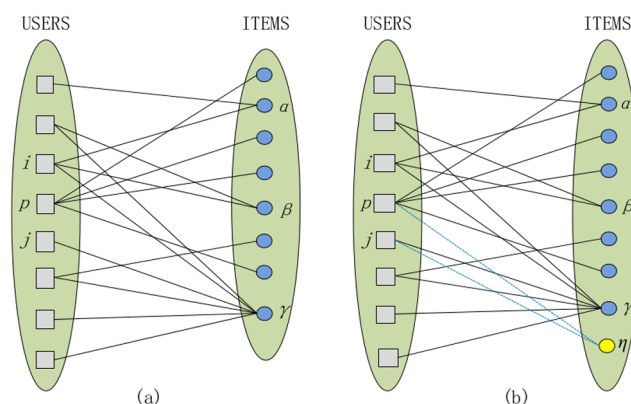


Figure 1. How to add a cold-start item to the user-item bipartite network. Users and items are represented by squares and circles respectively, and solid lines represent the existent links between them. Plot (a) is the original network, and plot (b) is the network after adding the item η (the yellow circle). The dotted lines are new links connecting η with two existent users.

doi:10.1371/journal.pone.0113457.g001

Recommender systems typically produce a given-length list of unpurchased items for each user based on his/her historical purchases. Of nothing comes nothing, that is to say, it is impossible to predict links for an isolate user or item. So only after having been purchased by some users, an item could have the chance to appear in some other users' recommendation lists. In real e-commerce web sites, to get a new customer is highly costly, and thus under the limited investment, choosing users with considerable coming influence on further recommendations is absolutely critical. Concretely speaking, this problem is described as follow. Given a bipartite network containing n users, m items and w links. A novel item O_η enters this network, and it can at most establish R links to users. Given the recommendation algorithm, we need to answer the question that how to choose such R users to maximize the frequency that O_η appears in other $(n - R)$ users' recommendation lists. For example, in [Figure 1\(b\)](#), item O_η (the yellow circle) comes and needs to link to some existent users. If $R = 1$, then to choose which user, U_p (most active user), U_j (one of the most inactive users) or another one, can make O_η be recommended more times?

Table 1. Basic statistical properties of the two data sets.

Data	n	m	w	$\langle k_{user} \rangle$	$\langle k_{item} \rangle$	S
TM	103,867	83,342	113,624	1.09	1.36	1.31×10^{-5}
Coo8	77,947	18,751	94,457	1.21	5.04	6.46×10^{-5}

n , m , and w represent the number of users, items and links, $\langle k_{user} \rangle$ and $\langle k_{item} \rangle$ stand for the average degrees of users and items, and $S = \frac{w}{n \times m}$ denotes the data sparsity.

doi:10.1371/journal.pone.0113457.t001

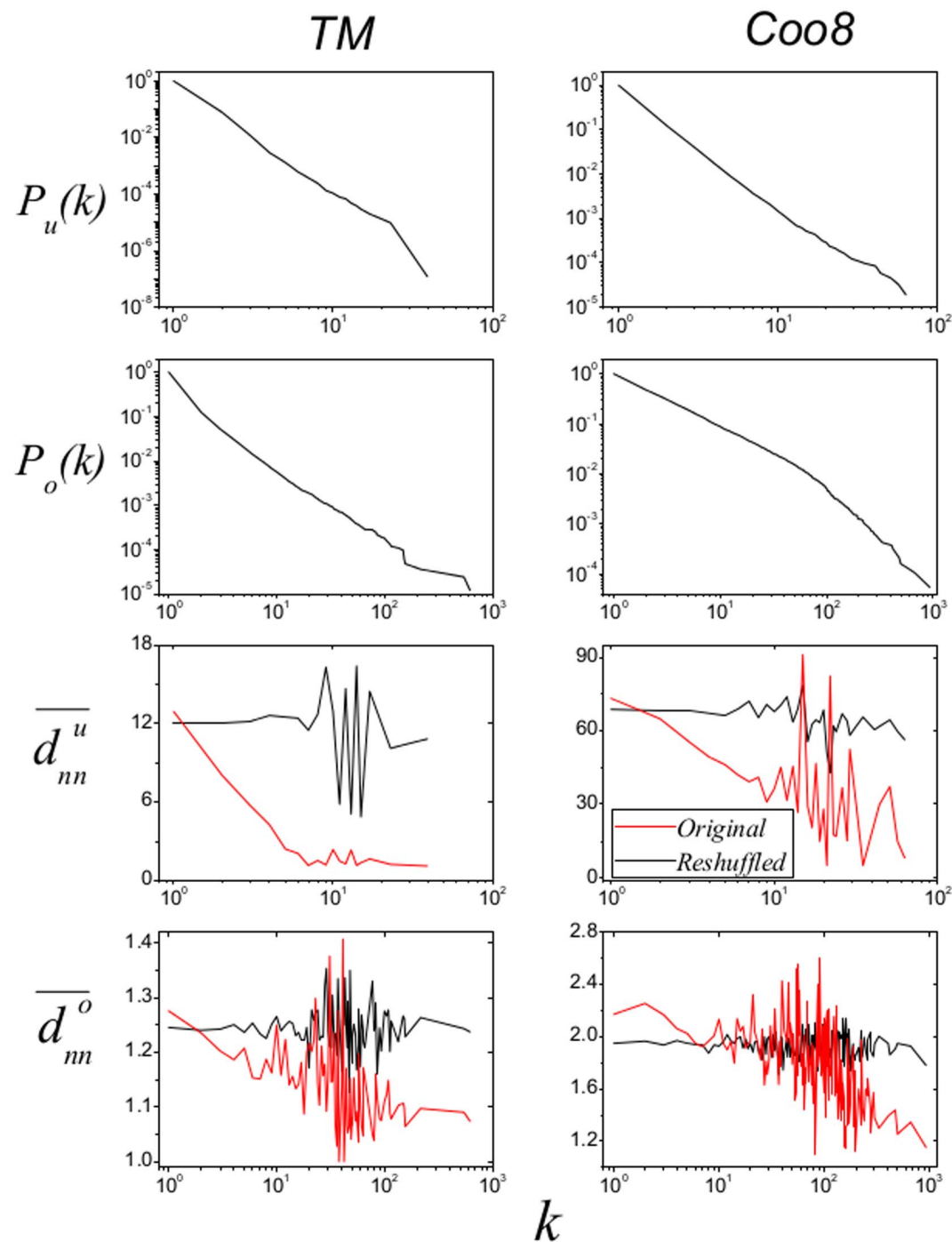


Figure 2. Degree distributions and degree correlations. All degree distributions are power-law-like. $\overline{d_{nn}^u}$ and $\overline{d_{nn}^o}$ are respectively showed in the 3rd and 4th rows, where red and black lines representing the results from original and reshuffled networks. Results of reshuffled networks are obtained by averaging over five independent realizations.

doi:10.1371/journal.pone.0113457.g002

We consider four strategies to choose those R users: (I) Maximum-degree strategy (MaxD). To rank all users in the descending order of degree, and select the top- R users, where users with the same degree are ranked randomly. (II) Minimum-degree strategy (MinD). To rank all users in the ascending order of degree, and select the top- R users, where users with the same degree are ranked randomly. (III) Preferential attachment strategy (PA). Each user's probability to be selected is proportional to his/her degree. (IV) Random strategy (RAN). The R users are selected completely randomly. Actually, all strategies above can be unified by a selecting probability on U_i as $p(U_i) \propto \frac{k_i^\tau}{\sum k_i^\tau}$, where τ is a tunable parameter. More specifically, the strategies MaxD, MinD, PA and RAN correspond to the cases of $\tau = +\infty$, $\tau = -\infty$, $\tau = 1$ and $\tau = 0$, respectively.

Among existent recommendation algorithms, item-based collaborative filtering (ICF) has found the widest applications in real e-commerce platforms for its accuracy, stability, scalability and robustness [5, 6, 38]. Here, we apply cosine similarity for each pair of items, say

$$\text{sim}(\alpha, \beta) = \sum_{s=1}^n \frac{a_{s\alpha} a_{s\beta}}{\sqrt{k_\alpha k_\beta}}, \quad (1)$$

where k_α and k_β are degrees of items O_α and O_β , respectively. In fact, the main results are not sensitive to the specific choices of common neighborhood based similarity indices [41], except for some very different indices irrelevant to the common neighbors between two nodes, such as preferential attachment index $\text{sim}(\alpha, \beta) = k_\alpha k_\beta$. For the target user U_i , we calculate the accumulative score $w_{i\alpha}$ for each item O_α by

$$w_{i\alpha} = \sum_{\gamma \neq \alpha} a_{i\gamma} \text{sim}(\alpha, \gamma), \quad (2)$$

and then rank all the unpurchased items in descending order according to their scores in Eq. (2). The top- L items will be recommended to U_i , where L is the length of recommendation list.

To compare the degree-dependent strategies, we employ a metric H that counts the number of users whose recommendation lists contain the target item (the R selected users linked to the target item are excluded), say

$$H = \sum_{i=1}^n \delta_i(R), \quad \delta_i(R) = \begin{cases} 1, & r_i \leq L \\ 0, & r_i > L \end{cases}, \quad (3)$$

where r_i is the position of the target item among all U_i 's unpurchased items. Obviously, $0 \leq H \leq (n - R)$, since the target item's degree equals R , and the larger value of H means better performance. The number of recommended items, L , is limited by the user interface, with typical size no larger than 6 (see real recommendation engines of Alibaba Group and Baifendian Inc. as examples).

In our simulation, we only consider R ranging from 1 to 1000 to see the influence of different R on promoting strategies. It is because too large R will

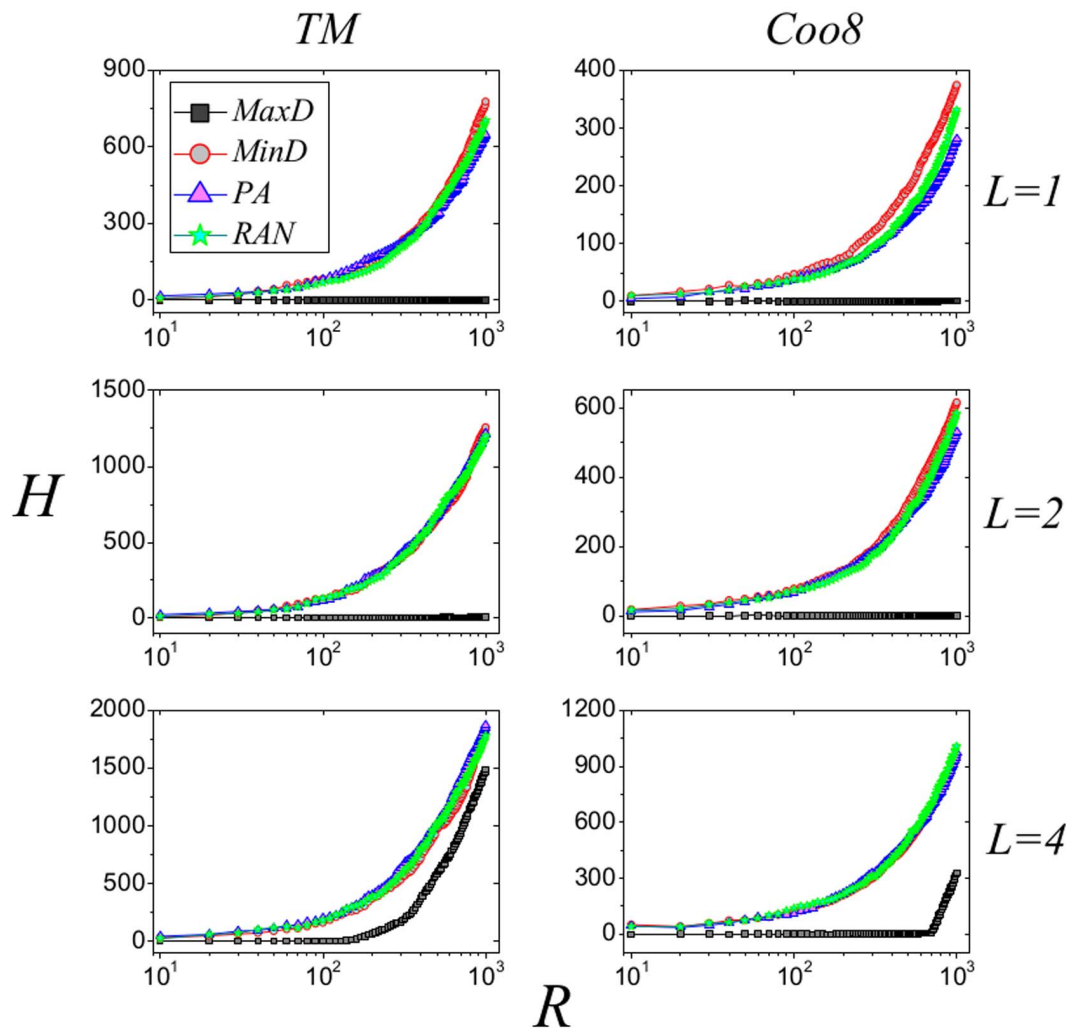


Figure 3. Performance of the four strategies for original TM and Coo8 bipartite networks. The results of MaxD, MinD, PA and RAN are represented by black squares, red circles, blue triangles and green pentagons, respectively. Data points are obtained by averaging over 100 independent realizations.

doi:10.1371/journal.pone.0113457.g003

result in very high cost and indeed $R=1000$ can make the item among the most popular ones. Unexpectedly, as shown in [figure 3](#), MaxD hardly makes new items recommended while MinD usually shows better performance. Consider the general case where the target item O_η has established a link to user U_i , and O_α and O_β are two of U_i 's collected items before O_η . For another user U_j who is not connected with O_η . If U_j has collected O_α but not O_β , then both O_β and O_η have the chance to be recommended to U_j . Since in the ICF algorithm, item similarities play the major role, let's compare the similarities $\text{sim}(\alpha, \beta)$ and $\text{sim}(\alpha, \eta)$. Statistically speaking, if U_i is a very active user selected by the MaxD strategy, O_α and O_β are probably less popular as indicated by the disassortative nature of the networks, therefore k_η (i.e., R) may be much larger than k_β and then $\text{sim}(\alpha, \eta)$ is probably smaller than $\text{sim}(\alpha, \beta)$, resulting in less probability of O_η to be

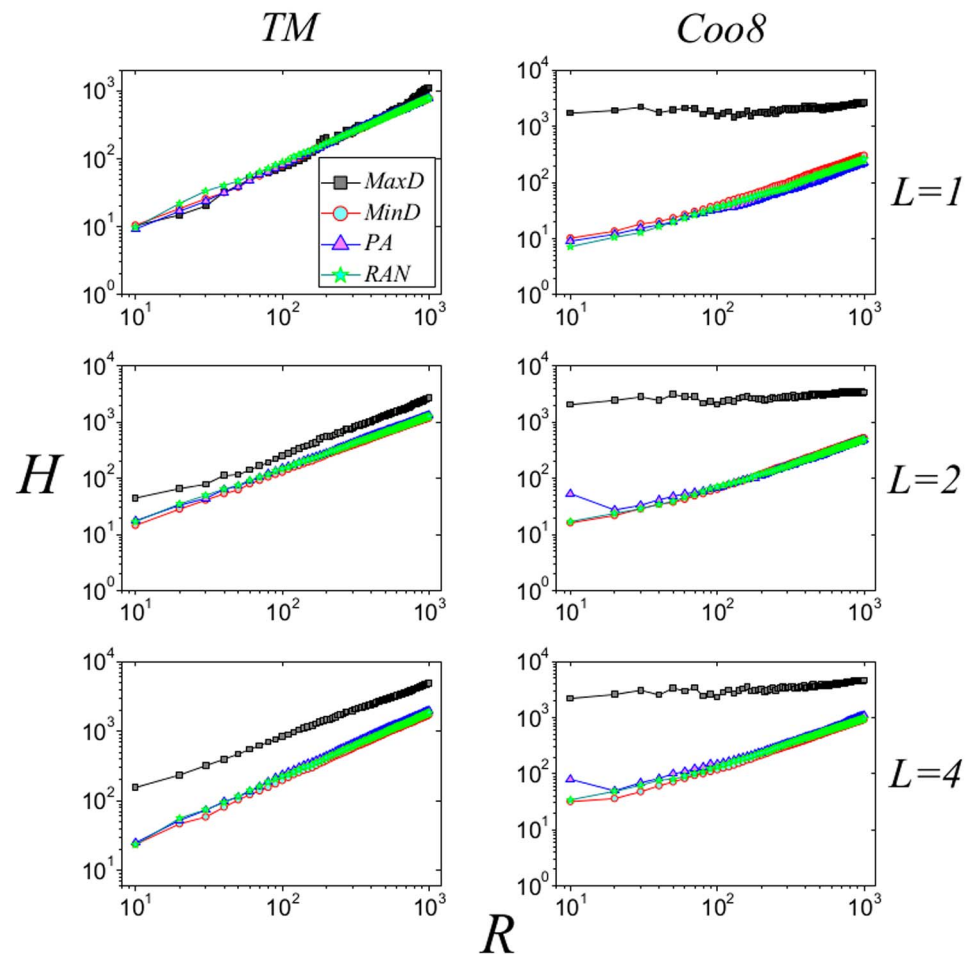


Figure 4. Performance of the four strategies for reshuffled networks. The results of MaxD, MinD, PA and RAN are represented by black squares, red circles, blue triangles and green pentagrams, respectively. Data points are obtained by averaging over 100 independent realizations.

doi:10.1371/journal.pone.0113457.g004

recommended to U_j . In contrast, if U_i is a very inactive user selected by the MinD strategy, O_α and O_β are probably of larger degrees according to the disassortative nature, resulting in smaller $\text{sim}(\alpha, \beta)$ and thus larger probability for O_η to be recommended to U_j . In addition, since U_i is very unpopular, it is also possible that $k_i = 1$ and U_i is only connected with O_α . In such case, for all other users connected with O_α , O_η will be the only recommended item related to U_i .

In a word, the disassortativity could contribute to the observations in [figure 3](#). To validate this inference, we reshuffle the original networks by link-crossing

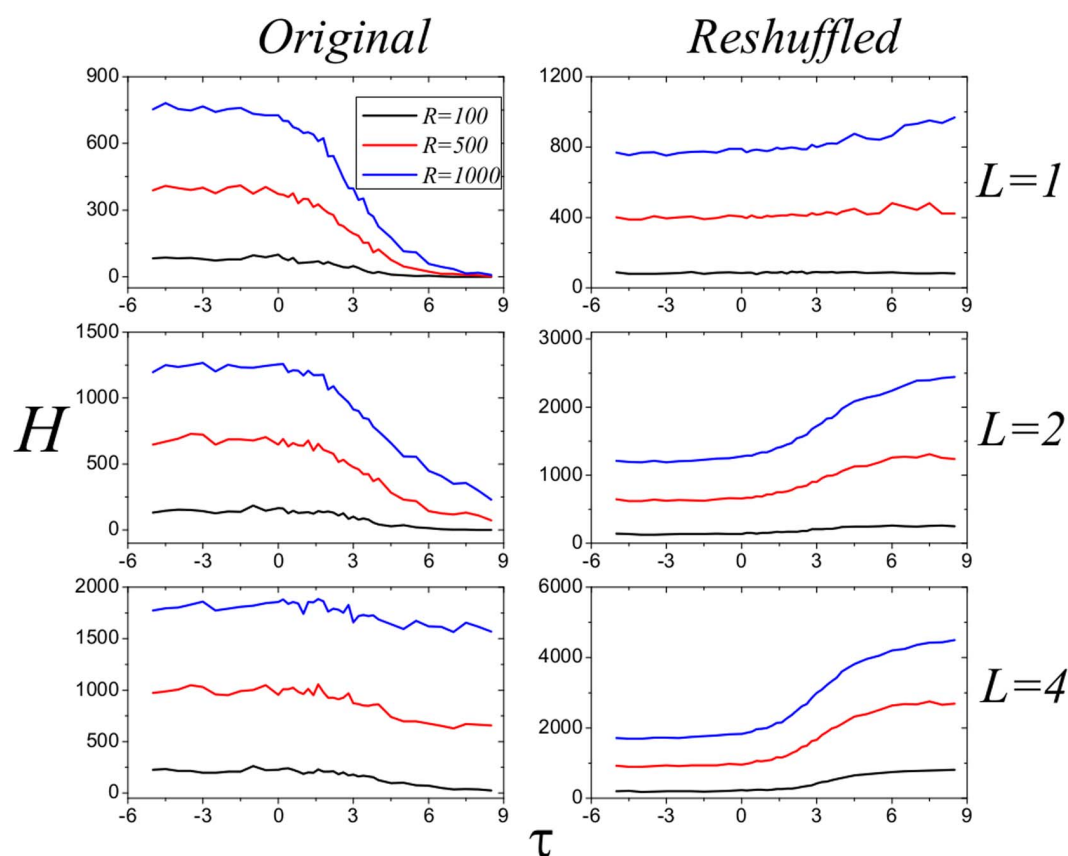


Figure 5. Performance of strategies with different τ on original and reshuffled TM networks. The black, red and blue lines represent the results for the cases $R=100$, $R=500$ and $R=1000$, respectively. Data points are obtained by averaging over 100 independent realizations.

doi:10.1371/journal.pone.0113457.g005

method to obtain the null networks [52]. Specifically speaking, in each step, two links, say (U_i, O_α) and (U_j, O_β) , are randomly picked out, and if U_i has not collected O_β and U_j has not collected O_α , these two links are rewired as (U_i, O_β) and (U_j, O_α) . In one realization, we repeat such rewiring for $3w$ times. After that, the reshuffled network has identical degree sequence as the original network but the disassortative nature is vanished as shown in figure 2. Figure 4 reports the performance of the four strategies in the reshuffled networks, from which we can see that the MaxD strategy performs the best. Comparing the results for original and reshuffled networks, we conclude that the advantage of MinD strategy results from the disassortative nature of real e-commerce user-item bipartite networks. In addition, in figure 5 and figure 6, we test the performance of strategies with different τ . For both TM and Coos8, the negative τ will lead to better performance while in the null networks, positive τ is better.

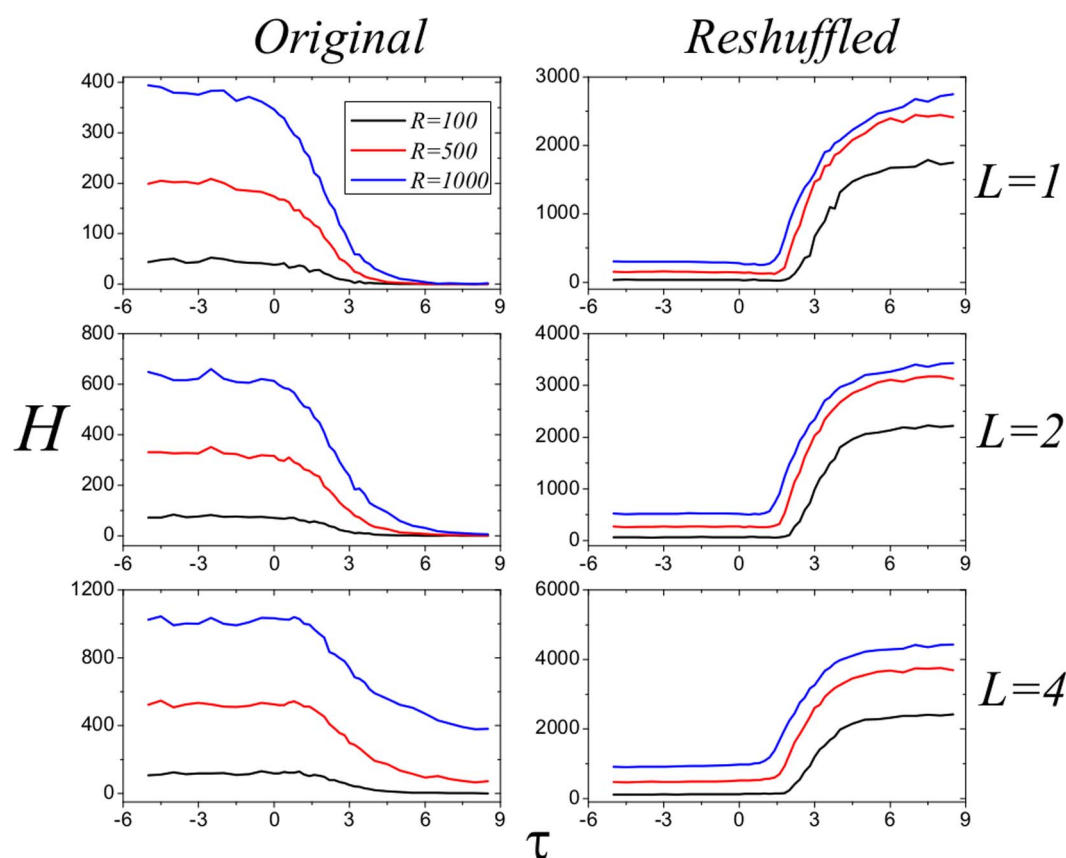


Figure 6. Performance of strategies with different τ on original and reshuffled Coo8 networks. The black, red and blue lines represent the results for the cases $R=100$, $R=500$ and $R=1000$, respectively. Data points are obtained by averaging over 100 independent realizations.

doi:10.1371/journal.pone.0113457.g006

Discussion

In this paper, we study a practical problem in e-commerce recommender systems: how to promote cold-start items? Under the item-based collaborative filtering systems, we show that the disassortative nature of real user-item networks leads to a non-trivial observation that to link a cold-start item to inactive users will give it more chance to appear in other users' recommendation lists. This observation is robust for varying recommendation length L and linking capacity R . It is also applicative to some variants of item-based collaborative filtering, such as the top- k nearest neighbors ICF [5].

Notice that, the reported results are affected by both the topological features and underlying recommendation algorithms. We have tested the user-based collaborative filtering [3], under which the MaxD is usually better than MinD. It is because the high-degree users tend to have high similarities with others, and to connect with those high-degree users can directly benefit items. In spite of this, this work is still relevant since in most real recommender systems, ICF plays a significant role. In addition, the perspectives and methods reported here are useful for real e-commerce applications, with the core merit is that the in-depth

understanding of the structure and algorithms of recommender systems can be transferred into applicable knowledge to better market products.

Supporting Information

Dataset S1. The TM and Coo8 data sets after anonymization.

[doi:10.1371/journal.pone.0113457.s001](https://doi.org/10.1371/journal.pone.0113457.s001) (ZIP)

Acknowledgments

We acknowledge Kuan Fang, Zhidan Zhao, Jian Gao, and Ying Zhou for useful discussions. We acknowledge Alibaba Group and Baifendian Inc. to share their real business data sets after anonymization.

Author Contributions

Conceived and designed the experiments: JHL TZ ZKZ ZY. Performed the experiments: JHL. Analyzed the data: JHL TZ ZKZ ZY CL. Contributed reagents/materials/analysis tools: JHL ZKZ WML. Wrote the paper: JHL TZ.

References

1. Ricci F, Rokach L, Shapria B, Kantor PB (2011) Recommender Systems Handbook. Springer.
2. Lü L, Medo M, Yeung CH, Zhang YC, Zhang ZK, et al. (2012) Recommender systems. Physics Reports 519: 1–49.
3. Goldberg D, Nichols D, Oki BM, Terry D (1992) Using collaborative filtering to weave an information tapestry. Communications of the ACM 35: 61–70.
4. Shardanand U, Maes P (1995) Social information filtering: algorithms for automating “word of mouth”. In: Proceedings CHI'95 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM Press, pp. 210–217.
5. Sarwar B, Karypis G, Konstan J (2001) Item-based collaborative filtering recommendation algorithms. In: Proceedings WWW'01 Proceedings of the 10th international conference on World Wide Web. New York, NY, USA: ACM Press, pp. 285–295.
6. Deshpande M, Karypis G (2004) Item-based top-n recommendation algorithms. ACM Transactions on Information Systems 22: 143–177.
7. Schafer JB, Frankowski D, Herlocker J, Shilad S (2007) Collaborative filtering recommender systems. In: Lecture Notes in Computer Science. Springer-Verlag, pp. 291–324.
8. Kim HN, Ji AT, Ha I, Jo GS (2010) Collaborative filtering recommender systems. In: Lecture Notes in Computer Science. Springer-Verlag, pp. 291–324.
9. Balabanovic M, Shoham Y (1997) Content-based, collaborative recommendation. Communications of the ACM 40: 66–72.
10. Pazzani MJ, Billsus D (2007) Content-based recommender systems. In: Lecture Notes in Computer Science. Springer-Verlag, pp. 325–341.
11. Niemann K, Wolpers M (2013) A new collaborative filtering approach for increasing the aggregate diversity of recommender systems. In: Proceeding KDD'13 Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. New York, NY, USA: ACM Press, pp. 955–963.

12. **Hofmann T** (2004) Latent semantic models for collaborative filtering. *ACM Transactions on Information Systems (TOIS)* 22: 89–115.
13. **Takács G, Pilászy I, Námeth B, Tikk D** (2007) Major components of the gravity recommendation system. *ACM SIGKDD Explorations Newsletter* 9: 80–83.
14. **Bunte K, Schneider P, Hammer B, Schleif FM, Villmann T, et al.** (2012) Limited rank matrix learning, discriminative dimension reduction and visualization. *Neural Networks* 26: 159–173.
15. **Huang Z, Chen H, Zeng D** (2004) Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information System (TOIS)* 22: 116–142.
16. **Zhang YC, Medo M, Ren J, Zhou T, Li T, et al.** (2007) Recommendation model based on opinion diffusion. *Europhysics Letters* 80: 68003.
17. **Zhang YC, Blattner M, Yu YK** (2007) Heat conduction process on community networks as a recommendation model. *Physical Review Letters* 99: 154301.
18. **Zhou T, Ren J, Medo M, Zhang YC** (2007) Bipartite network projection and personal recommendation. *Physical Review E* 76: 046115.
19. **Zhou T, Kuscsik Z, Liu JG, Medo M, Wakeling JR, et al.** (2010) Solving the apparent diversity-accuracy dilemma of recommender systems. *Proceedings of National Academy of Science of the United States of America* 107: 4511–4515.
20. **Liu JG, Zhou T, Guo Q** (2011) Information filtering via biased heat conduction. *Physical Review E* 84: 037101.
21. **Lü L, Liu W** (2011) Information filtering via preferential diffusion. *Physical Review E* 83: 066119.
22. **Zhou Y, Lü L, Liu W, Zhang J** (2013) The power of ground user in recommender systems. *PLoS ONE* 8: e70094.
23. **Schein AI, Popescul A, Ungar LH, Pennock DM** (2002) Methods and metrics for cold-start recommendations. In: *Proceeding SIGIR'02 Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA: ACM Press, pp. 253–260.
24. **Park ST, Chu W** (2009) Pairwise preference regression for cold-start recommendation. In: *Proceeding Recommender Systems'09 Proceedings of the third ACM conference on Recommender systems*. New York, NY, USA: ACM Press, pp. 21–28.
25. **Cantador I, Bellogín A, Vallet D** (2010) Content-based recommendation in social tagging systems. In: *Proceeding Recommender Systems'10 Proceedings of the forth ACM conference on Recommender systems*. New York, NY, USA: ACM Press, pp. 237–240.
26. **Ronen R, Koenigstein N, Ziklik E, NICE N** (2013) Selecting content-based features for collaborative filtering recommenders. In: *Proceeding Recommender Systems'13 Proceedings of the 7th ACM conference on Recommender systems*. New York, NY, USA: ACM Press, pp. 407–410.
27. **Zhang ZK, Liu C, Zhang YC, Zhou T** (2010) Solving the cold-start problem in recommender systems with social tags. *Europhysics Letters* 92: 28001.
28. **Zhang ZK, Zhou T, Zhang YC** (2011) Tag-aware recommender systems: A state-of-the-art survey. *Journal of computer science and technology* 26: 767–777.
29. **Yin D, Guo S, Chidlovskii B, Davison BD** (2013) Connecting comments and tags: improved modeling of social tagging systems. In: *Proceeding WSDM'13 Proceedings of the 6th ACM international conference on Web search and data mining*. New York, NY, USA: ACM Press, pp. 547–556.
30. **Zhang L, Bai LS, Zhou T** (2013) Crossing recommendation based on multi-business to customer behavior. *Journal of University of Electronic Science and Technology of China* 42: 154–160.
31. **Ziegler CN, McNee SM, Konstan JA, Lausen G** (2005) Improving recommendation lists through topic diversification. In: *Proceeding WWW'05 Proceedings of the 14th international conference on World Wide Web*. New York, NY, USA: ACM Press, pp. 22–32.
32. **Zhou T, Su RQ, Liu RR, Jiang LL, Wang BH, et al.** (2011) Accurate and diverse recommendations via eliminating redundant correlations. *New Journal of Physics* 11: 123008.
33. **Zhou JL, Fu Y, Lu H, Sun CJ** (2011) From popularity to personality a heuristic music recommendation method for niche market. *Journal of Computer Science and Technology* 26: 816–822.

34. **Lam SK, Riedl J** (2004) Shilling recommender systems for fun and profit. In: *Proceeding WWW'04 Proceedings of the 13th international conference on World Wide Web*. New York, NY, USA: ACM Press, pp. 393–402.
35. **Mobasher B, Burke R, Bhaumik R, Williams C** (2005) Effective attack models for shilling item-based collaborative filtering systems. In: *Proceedings of the 2005 WebKDD Workshop*. New York, NY, USA: ACM Press.
36. **Mobasher B, Burke R, Williams C, Bhaumik R** (2006) Analysis and detection of segment-focused attacks against collaborative recommendation. In: *Lecture Notes in Computer Science*. Springer-Verlag, pp. 96–118.
37. **Mehta B, Hofmann T, Fankhauser P** (2007) Lies and propaganda: detecting spam users in collaborative filtering. In: *Proceeding IUI'07 Proceedings of the 12th international conference on Intelligent user interfaces*. New York, NY, USA: ACM Press, pp. 14–21.
38. **Mobasher B, Burke R, Bhaumik R, Williams C** (2007) Toward trustworthy recommender systems: An analysis of attack models and algorithm robustness. *ACM Transactions on Internet Technology* 7: 1–41.
39. **Verhoef PC, Leeflang PSH** (2009) Understanding the marketing departments influence within the firm. *Journal of Marketing* 73: 14–37.
40. **Edward FC, Hadlock CJ, Pierce JR** (2008) Investment, financing constraints, and internal capital markets: Evidence from the advertising expenditures of multinational firms. *The Review of Financial Studies* 22: 2361–2392.
41. **Lü L, Zhou T** (2011) Link prediction in complex networks: A survey. *Physica A* 390: 1150–1170.
42. **Huang Z, Zeng DD, Chen H** (2007) Analyzing consumer-product graphs: Empirical findings and applications in recommender systems. *Management Science* 53: 1146–1164.
43. **Koren Y, Sill J** (2011) An ordinal model for predicting personalized item rating distributions. In: *Proceedings of the 5th ACM Conference on Recommender Systems*. New York, NY, USA: ACM Press, pp. 117–124.
44. **Zhang QM, Zeng A, Shang MS** (2013) Extracting the information backbone in online system. *PLoS ONE* 8: e62624.
45. **Zeng W, Zeng A, Liu H, Shang MS, Zhou T** (2014) Uncovering the information core in recommender systems. *Scientific Reports* 4: 6140.
46. **Barabási AL, Albert R** (1999) Emergence of scaling in random networks. *Science* 286: 509–512.
47. **Lambiotte R, Ausloos M** (2005) Uncovering collective listening habits and music genres in bipartite networks. *Physical Review E* 72: 066107.
48. **Goncalves B, Ramasco JJ** (2008) Human dynamics revealed through web analytics. *Physical Review E* 78: 026123.
49. **Yang Z, Zhang ZK, Zhou T** (2012) Anchoring bias in online voting. *Europhysics Letters* 100: 68002.
50. **Shang MS, Lü L, Zhang YC, Zhou T** (2010) Empirical analysis of web-based user-object bipartite networks. *Europhysics Letters* 90: 48006.
51. **Pastor-Satorras R, Vázquez A, Vespignani A** (2001) Dynamical and correlation properties of the internet. *Physical Review Letters* 87: 258701.
52. **Zhang CJ, Zeng A** (2012) Behaviour patterns of online users and the effect on information filtering. *Physica A* 391: 1822–1830.