A generalized model via random walks for information filtering

Zhuo-Ming Ren [a,b,*], Yixiu Kong [a], Ming-Sheng Shang [b,*], Yi-Cheng Zhang [a]

[a] *Department of Physics, University of Fribourg, Chemin du Musée 3, CH-1700, Fribourg, Switzerland*
[b] *Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, ChongQing, 400714, PR China*

There could exist a simple general mechanism lurking beneath collaborative filtering and interdisciplinary physics approaches which have been successfully applied to online E-commerce platforms. Motivated by this idea, we propose a generalized model employing the dynamics of the random walk in the bipartite networks. Taking into account the degree information, the proposed generalized model could deduce the collaborative filtering, interdisciplinary physics approaches and even the enormous expansion of them. Furthermore, we analyze the generalized model with single and hybrid of degree information on the process of random walk in bipartite networks, and propose a possible strategy by using the hybrid degree information for different popular objects to toward promising precision of the recommendation.

1. Introduction

Over the last decade, the rapid growth of information in both online and offline leads to an information overload problem [1,2]. All of surfers would have the feeling of confusion which one is the best when searching online, reading online, shopping online, entertaining online, or even dating online [3,4]. To address these problems, the information filtering has become a promising and effective way to filter out the irrelevant information and provides personalized suggestions according to the track of past purchases of users as well as other information of products and users [5,6]. Due to its significance for the economy and society, designing efficient information filtering algorithms has received wide attractions in many branches of science such as computer science, information science and interdisciplinary physics. One of the most promising information filtering algorithms is the collaborative filtering (CF) [7,8]. The CF makes work according to the database of the users' past history of purchases and the product searching records to offer the personalized recommendation. Breese et al. [9] classified CF into two broad groups which were memory-based and model-based methods. The memory-based methods predict missing information and recommend products based on similarity measures between users and products [7,8,10]. The model-based algorithms use the collection of the user and object information to learn an information filtering model by clustering [11],

bayesian [12], matrix factorization [13,14] and other machine learning techniques [15,16]. Being different from the perspective computer sciences, the interdisciplinary physics approaches adapted the complex network theory and various classical physics processes have provided some new insights and solutions for the challenges in the active field of the information filtering [6,17,18], for instance, a diffusion process analogous to the heat conduction process across a bipartite complex network [19], a network-based inference method considering the resource allocation dynamics on bipartite complex networks [20], the opinion diffusion [21] and the gravity principle [22] being extended in the information filtering, the information core and information backbone [23,24] shed some light on the in-depth understanding of information filtering. Further, the review [6] highlighted a prospect of physicists to a comprehensive guide to information filtering algorithms.

In summary, the CF and interdisciplinary algorithms have already been successfully applied to many well-known online e-commerce platforms. Meanwhile, many recent works have been devoted to study the expansion of both algorithms, for instance hybrid method [25,26], biased-heat conduction [27,28], multi-channel diffusion [29], preferential diffusion [30,31], hybrid diffusion [32], direct random walks method based on CF [33], hypergraph model with social tag [34,35], multi-linear interactive matrix factorization [36]. These algorithms would further improve the efficiency of the information filtering. By referring to them, there could exist a simple general formula behind CF, interdisciplinary physics algorithms as well as the extension methods. Motivated by this idea, we propose a simple general model in which employing the dynamics of the random walk in bipartite networks and then derive an analytical expression for tunable parameters of the transition

* Corresponding authors.
E-mail addresses: zhuomingren@gmail.com (Z.-M. Ren), msshang@cigit.ac.cn (M.-S. Shang).
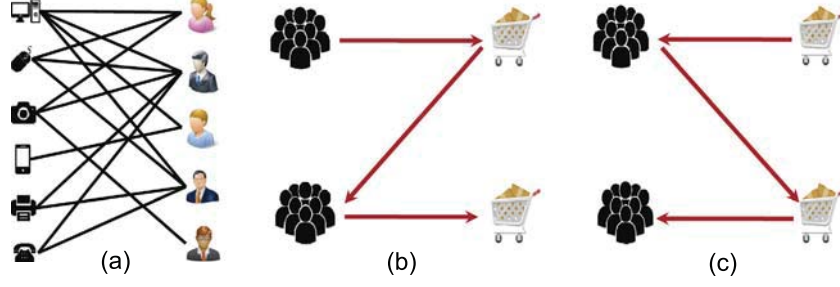
Fig. 1. (Color online.) Illustration of a bipartite network which is constructed by the tracks of the consumer purchases, which can be described as the process of the random walker model. (a) A small example of a bipartite network which keeps the track of the consumer purchases. (b) The random walker starts from the user side. (c) The random walker starts from the object side.

probability matrix. When taking into account the degree information, the process of random walkers can be equivalent to the representative information filtering algorithms such as the CF [7,8], heat conduction method [21], network-based inference method [20], hybrid method [25]. These above methods are exceptional cases from the proposed generalized model. Consequently, we analyze the generalized model with single and hybrid of degree information on the process of random walk in bipartite networks, and the effect of degree information on the different popular level objects in the process of the information filtering. Finally, we suggest a possible strategy for different popular objects in the recommendation when using the single and hybrid of degree information on the process of random walk in bipartite networks.

2. Models

In Fig. 1(a), we present an example of a bipartite network which is constituted by the records of consumer purchases. We would recommend an object to a user based on his/her similar past purchases with others or the purchase records of those who have some resemblance with him/her. In a broad view, we can make personalized recommendations for each consumer depending on the past purchases of the consumer as well as information relating to the similarity of other consumers or items. With this concept, many online business platforms such as Alibaba, Amazon, Netflix, Digg are reported to develop sophisticated information filtering systems to boost their online sales. Therefore, the information filtering problem is reduced to the problem that estimating the valuation for products that have not been seen by consumers. Considering the fact that millions of products exist in online business platform, the information filtering problem could be thought as identifying only the highest ranked products for each user instead of predicting the rating of each product for each user.

2.1. The random walker model

Let us consider an unweighted bipartite network $G(U, O, E)$, where $U = \{u_1, u_2, \cdots, u_m\}$, $O = \{o_1, o_2, \cdots, o_n\}$ and $E = \{e_1, e_2, \cdots, e_l\}$ are the set of users, products and links respectively. The network is described by A_{M*N}, adjacency matrix with $A_{i\alpha}$ if there is a track between user i and product α, and zero otherwise. The dynamics of a random walker on the bipartite network is encoded by a transition probability matrix with elements of the form,

$$P(u_i \rightarrow o_\alpha) \text{ or } P(o_\alpha \rightarrow u_i), \tag{1}$$

measuring the probability that a walker passes from u_i to o_α (or o_α to u_i).

Firstly, we can assume that a random walker starts from the user side at the beginning, and the walker's initial state assigns as $X_0 = u$. We represent the probability $P(u, o, t)$ as the walker

passes from the user side u_i to the object side o_α during the t steps and corresponding to transition state which interprets as $X_0, X_1, \cdots, X_{t-1}, X_t$. Consequently, the probability $P(u, o, t)$ that the walker reaches the object side o_α during t steps under the initial state $X_0 = u$ is,

$$P(u, o, t) = P(X_t = o \mid X_0 = u) \tag{2}$$

and then, from bayesian theory, we can get,

$$
\begin{aligned}
P(X_t = o \mid X_0 = u) \\
= \sum_{o' \in O} P(X_t = o \mid X_{t-2} = o', X_0 = u) P(X_{t-1} = o' \mid X_0 = u) \\
= \sum_{o' \in O} P(X_t = o \mid X_{t-2} = o', X_0 = u) P(u, o, t-1).
\end{aligned}
\tag{3}
$$

We now define the walkers from the user side u and along the object side o to the next pathway. The each pathway is selected by the probability ψ.

$$P(X_t = o \mid X_{t-2} = o', X_0 = u) = \psi. \tag{4}$$

Finally, combing the equation (2), (3), (4), we can get

$$P(X_t = o \mid X_0 = u) = \sum_{o' \in O} P(u, o, t-1)\psi. \tag{5}$$

The recommendation is to develop an effective way to provide personalized suggestions according to the track of past purchases of users and objects. We can assume the walkers can walk along with the track of past purchases of users and objects. For simplicity, Fig. 1(b) and (c) represent the process of the walker when considering the situation $t = 3$. Fig. 1(b) shows that the walker's road map (users–objects–users–objects) which starting from the user side, while as shown in Fig. 1(c), the walker's road map appears objects–users–objects–users which starting from the object side. Starting from the user side, the random walker must walk three steps at least to the object side, and starting from the object side, the random walker also must walk three steps at least to the user side. So, if an object is selected to recommend to a user, the random walker needs to walk t ($t = 2*j+1$, $j = 1, 2, \ldots$) steps. In summary, there are two pathways for random walkers in one step.

1) The walkers pass from the user side to the object side with the probability matrix,

$$P(u \rightarrow o) = \Psi(u)A. \tag{6}$$

2) The walkers pass from the object side to the user side with the probability matrix,

$$P(o \rightarrow u) = A\Psi(o), \tag{7}$$

where, $\Psi(u)$ and $\Psi(o)$ are $M*M$ and $N*N$ diagonal matrix respectively corresponding to the user and product profiles, for

example, degree, eigenvector, cluster coefficient and so on in the bipartite networks. Apparently, we could also present the probability matrix as $A^T \Psi(u)$ or $\Psi(o)A^T$ respectively.

2.2. The random walker model with degree information

The transition probability matrix allows us to capture many attributes of the bipartite networks. Among these attributes, the degree information which is the very simple way of quantifying attributes of bipartite networks would be considered in this paper. Firstly, we define $D(u, \lambda)$ and $D(o, \lambda)$ as the diagonal degree matrix which is calculated detailedly,

$$D(u, \lambda) = \begin{cases} \frac{1}{(\sum_j A_{ij})^\lambda}, & i = j \\ 0, & i \neq j \end{cases} \tag{8}$$

$$D(o, \lambda) = \begin{cases} \frac{1}{(\sum_i A_{ij})^\lambda}, & i = j \\ 0, & i \neq j \end{cases} \tag{9}$$

We then consider the simplest situation of the random walker dynamics when $t = 3$. As we can see from Fig. 1(b) and (c), there are two ways in one step, which the walk starts from the user side or the abject side respectively,

1) The walkers pass from the user side to the object side with the probability matrix,

$$P(u \to o) = D(u, \lambda)A. \tag{10}$$

2) The walkers pass from the object side to the user side with the probability matrix,

$$P(o \to u) = AD(o, \lambda). \tag{11}$$

Now, we introduce the CF, interdisciplinary physics algorithms from the prospect of the dynamics of random walkers by equation (10) and (11). The common procedure of the collaborative filtering technique is firstly computing similarity between users (user-based) or objects (object-based) and then making a prediction based on the similarity in the first step. The basic idea in the quantification of similarity between two users is based on the number of objects which have been chosen by both users in the past. It is also possible to define a similarity between two objects based on the number of users who have chosen them. Thus, taking account of these two steps, the simple collaborative filtering procedure can be expressed by AA^T. In the HC, the resource is redistributed via an averaging procedure with users receiving a level of resource equal to the mean amount possessed by their neighbouring objects, and objects then receiving back the mean of their neighbouring users' resource levels. By contrast, in the NBI, the initial resource placed on objects is first evenly distributed among neighbouring users and then evenly redistributed back to those users neighboring objects. Both HC and NBI are redistributed resource in a manner akin to a random walker process. Whereas, HC employs a row-normalized transition matrix, that of NBI is a column-normalized transition matrix. Thereupon, the hybrid method (HM) can be achieved by incorporating the a tunable parameter λ into the transition matrix normalization. The detailed matrices are expressed as $D(o, 1-\lambda)A^T D(u, 1)AD(o, \lambda)A^T$. When $\lambda = 0$ the equation equals HC, and $\lambda = 1$ the equation equals NBI. As discussing above, these recommendation procedures can be expressed according to the random walker process as shown in Table 1, from which the above methods are exceptional cases from equation (10) and (11). When we consider three steps in random walks, we can achieve a simple recommendation. While in the three or more steps, there exists a great number of combinations which can develop a large amount of recommendation algorithms.

Table 1
Deducing the proposed information filtering methods recently in the literature by the random walker model.

Method	First step	Second step	Third step
CF	A	A^T	A
HC	A	$A^T D(u, 1)$	$AD(o, 1)$
NBI	A	$D(o, 1)A^T$	$D(u, 1)A$
HM	$D(o, 1-\lambda)A^T$	$D(u, 1)A$	$D(o, \lambda)A^T$

3. Materials and metrics

3.1. Materials

In this paper, we test the general model in a representative real network which is sampled from MovieLens [37]. MovieLens is an online video recommendation web site which invites users to rate videos. The rating records in the MovieLens range from one (i.e., worst) to five (i.e., best). If a user i selects an video α and rates it, a link between the user i and the video α would be established as shown in Fig. 1(a). In this paper, we only consider the ratings larger than 3 as links. Based on those selected links, we can construct a user-video bipartite network. MovieLens-1 data set (M1) contains more than 8 million records, and MovieLens-2 (M2) is a subset of M1 dataset which only contains 800 thousand records. M2 data set is usually used as benchmark data set of the information filtering. In the following we report the results corresponding to the M2 data set that is randomly divided into the training set and the test set. The training set contains 90% of the data, while the test set consists of the remaining 10% of data. It is notable that in results, we consider that the length of a recommendation list L which can be proposed to each user is 20. This value has been already used in other works and can be considered a valid reference to compare the performance of different recommendation measurements. In addition, we run 100 independent realizations in each experiment.

For Movielens data sets, we first display the degree distribution of objects in Fig. 2(a). The distribution exhibits relatively broad tails which indicates most objects' degree is small, and a few objects' degree is large. This suggests that we should take the degree long tail effect into account. In this paper, we classify the objects' degree into three different levels. In the high level, we consider the objects whose degree ranks top 1%. In the medial level, the objects' degree ranks from top 1% to top 10%. The low level contains the rest 90% of objects. For each targeted level object set, the user degree distribution is shown in Fig. 2(b) and (c). We can find in the high degree object group, the user degree distribution shows a strong inclination, which suggests the users with small degree are more likely to select the objects in the high level, while for the low degree object group, the inclination of the user degree drops down clearly. It is notable that these objects accounted for 90%. Thereupon, when designing an efficient information filtering algorithms, the inclination of the user preference to objects at different popularity level should be well-considered.

3.2. Metrics

As we known, the discussed above each method will generate a recommendation list L for users. In principle, the recommendation list from an effective recommendation method should be as close as possible to the user purchase records in the real world. While, the proposed methods allow us to employ many attributes of the bipartite networks and tunable parameters to capture the individual interests of users. As a result, we need to evaluate that how the attributes of the bipartite networks and tunable parameters effect the recommendation list. So, we define the **preference**
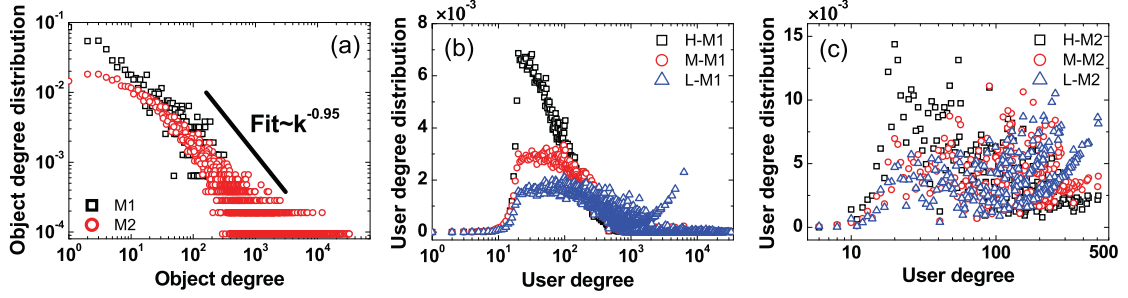
Fig. 2. (Color online.) The degree distribution of objects in two movielens data sets. (a) The degree distribution of objects. The two sets have the same slope −0.95. (b) and (c) The distribution of the user degree with objects which are in high, medial, low level respectively.
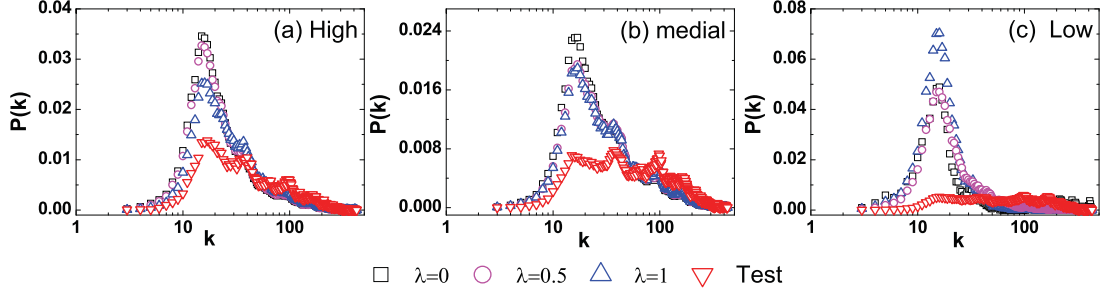

Fig. 3. (Color online.) The preference of recommendation when considering single factors of the user degree information. We employ the random walker process which is described as $D(U, \lambda)AA^T D^2(u, \lambda)A$. In order to compare the prediction with the realization, we consider the tunable parameters $\lambda = 0$, 0.5, 1 and the test set.

of the recommendation to analyze that for a specific group of objects, the users by which the recommendation algorithms tend to recommend to. Meanwhile, we use the **precision of the recommendation** to analyze how many objects in the recommendation list are actually in user purchase records in the real world. Firstly, we limit to the objects in the high level and assume that for a targeted user i, there are $H(i)$ high level objects in the recommendation list, and $C(i)$ high level objects belong to the test and the recommendation list in common. We define the preference of recommendation as follows,

$$P(k) = \frac{\sum_{i \in U_k} H(i)}{\sum_{i \in U} H(i)}, \qquad (12)$$

where U_k is the set of users whose degree equals k. The precision is defined as the ratio of the relevant objects in the real world selected to the number of objects recommended accurately. The precision equals,

$$Precision = \frac{1}{m_H} \sum_{i \in U} \frac{C(i)}{H(i)}, \qquad (13)$$

where m_H is the total number of objects in the high level. Thus, a higher the precision value means a more accurate prediction of the recommendation. In addition, the corresponding objects in the medial and low level are taken the same processes.

4. Validations

4.1. The effect of single factors

For a better understanding of the generalized model, we consider the single degree information in the process of the random walk. Firstly, we consider the process of the random walk only with the probability matrix $(D(u, \lambda)A)$ during the three steps. The equation for the three steps of random walks can be expressed as $D(U, \lambda)AA^T D^2(u, \lambda)A$. The preference of recommendation is shown in Fig. 3. With λ increasing, the recommendation algorithm

is less likely to recommend objects which are in the high and medial level, but for objects which are in the low level as shown in Fig. 3(c), with λ increasing, the corresponding algorithm is more inclined to recommend these objects. It suggests that the prediction of recommendation for objects which are in high and medial level is contrary to that for objects in low level with λ increasing.

Meanwhile, we discuss the dynamics of the random walk which is only affected by the object degree information with the probability matrix $(D(o, \lambda)A^T)$. The random walk process can be expressed as $AD^2(o, \lambda)A^T AD(o, \lambda)$. The results of the recommendation preference are shown in Fig. 4. It is notable that when $\lambda = 1$, the prediction results that the recommendation algorithm can not recommend the objects which are in the high and medial level as shown in Fig. 4(a) and (b). As given in Fig. 4(b) and (c), the recommendation algorithm is more tendency to recommend the objects which are in medial and low level with λ decreasing. In summary, when only considering the user degree information, objects in the high and medial level are less likely to be recommended to the users with the parameter λ increasing, while the objects in the low level, the preference of recommendation is reversed. When only considering the object degree information with the parameter λ increasing, the objects in the low level are less inclined to be recommended to the users.

Fig. 5 shows that how the precision changes with the tunable parameter λ when the process of the random walk is only affected by the object or user degree information. There exist different optimal parameters leading to the highest precision value for objects in the high, medial and low level respectively. Firstly, we analyze that the precision value changes with λ varying when only considering the user degree information. The precision about the objects in the high level is closely to 1 and keeps stably. While in the medial level, the precision declines slowly when λ increases, in addition in low level, the precision value is very low, and also increases slowly. But when only consider the object degree information, the precision can reach 0.11 in the low level. In the high and medial level, when λ is more than 0.75, the precision stays at 0. The main reason is that when λ rising, the recommendation algorithm do not recommended objects in the high and medial level as shown
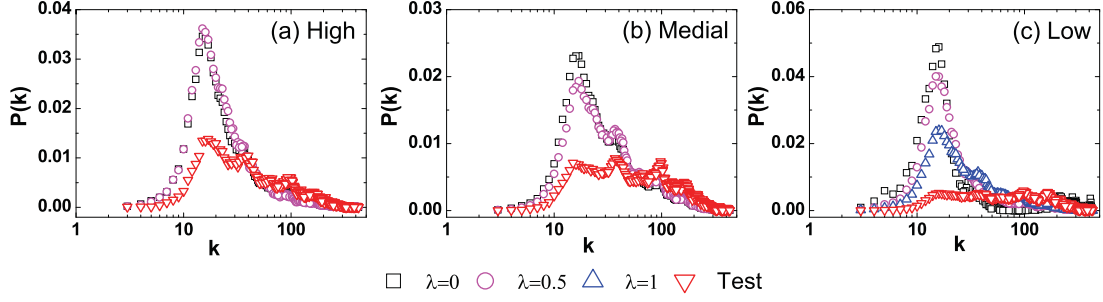
4

Fig. 4. (Color online.) The recommendation preference when considering single factors of object degree information. We employ the random walker process which can be described as $AD^2(o, \lambda)A^T AD(o, \lambda)$. In order to compare the prediction with the realization, we consider the tunable parameters $\lambda = 0, 0.5, 1$ and the test set.
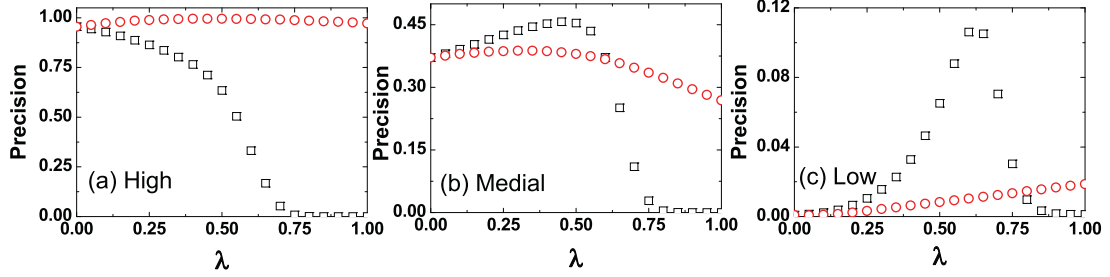

Fig. 5. (Color online.) Precision with the tunable parameter λ for the prediction of the recommendation when considering the single degree information in the process of the random walk. Squares ($\square$) indicate that the precision of the recommendation when the random walker process with the probability matrix $D(o, \lambda)A^T$ uses the object degree information. Circles ($\circ$) report that the precision of the recommendation when the dynamics of random walk affects by user degree information with the probability matrix $(D(u, \lambda)A^T)$.
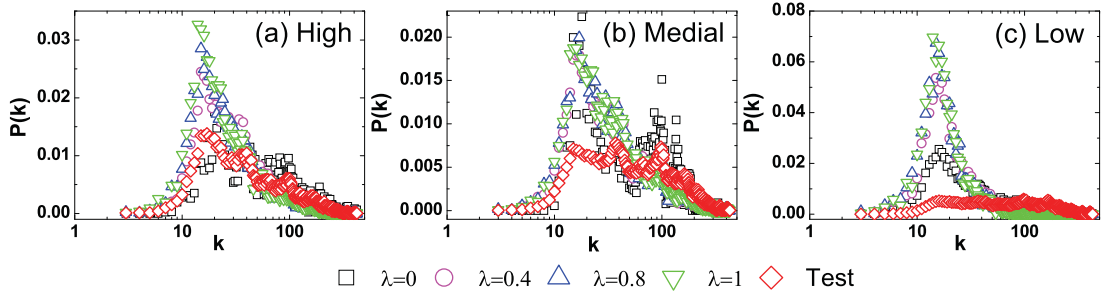

Fig. 6. (Color online.) The recommendation preference when considering the hybrid factors of the constant user degree information and the variable object degree information. We employ the random walker process which is described as $(D(o, 1 - \lambda)A^T D(u, 1)AD(o, \lambda)A^T)$. In order to compare the prediction with the realization, we consider the tunable parameters $\lambda = 0, 0.4, 0.8, 1$ and the test set.

in Fig. 4(a) and (b). As a result, in order to improve the precision of the recommendation, we should take different strategies to recommend the objects at different popularity levels.
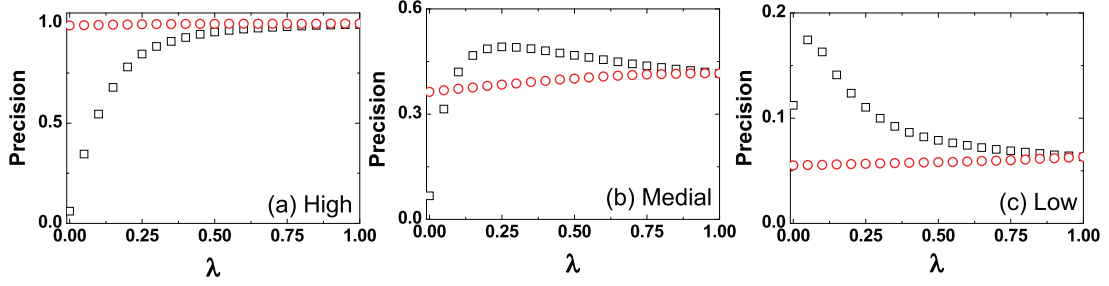
4.2. The effect of hybrid factors

In this section, we analyze the effect of the hybrid factors about the degree information of users and objects. Firstly, we consider the formula $(D(o, 1 - \lambda)A^T D(u, 1)AD(o, \lambda)A^T)$, in which the object degree information and user degree information are both taken account in the information filtering systems. In the system, the user degree information is constant, while the object degree information is tunable. We can describe the dynamics of random walk in detail. In the first step, the random walkers pass from the object side to the user side who are only affected by the object degree information with the probability matrix $(D(o, 1 - \lambda)A^T)$. In the second step, the random walkers move through the access from the user side to the object side with the constant probability matrix $(D(u, 1)A)$. While in the third step, the random walkers go to the user side from the object side with the tunable probability matrix $(D(o, \lambda)A^T)$ which just includes the object information. In addition, we consider the parameter $\lambda = 0, 0.4, 0.8$ and 1. We

can see from Fig. 6(a) in the high level, when $\lambda = 0$, the recommendation preference fits with the test set well. However, when λ increases to 1, the preference of recommendation to objects in the high level becomes clearer, which suggests that the recommendation algorithm is more inclined to recommend the objects in the high level. From Fig. 6(b) and (c) in the medial and low level, the preference of recommendation to objects also becomes more evident as λ increasing.

After discussing the random walker dynamics with the constant user degree information and the tunable object degree information, we then analyze the process of the random walk with the tunable user degree information and the constant object degree information. In the first step, walkers start from the user side to object side, we only consider the user degree information with the tunable probability matrix $(D(u, 1 - \lambda)A)$. In the second step, the random walkers pass from the object side to the user side determined by the constant object degree information with the probability matrix $(D(o, 1)A^T)$. In the last step, the walkers track from the user side to the object side, the user degree information is also taken into account with the tunable probability matrix $(D(u, \lambda)A)$. The results of the recommendation preference are revealed in Fig. 7, it is surprising to us that the correspond-
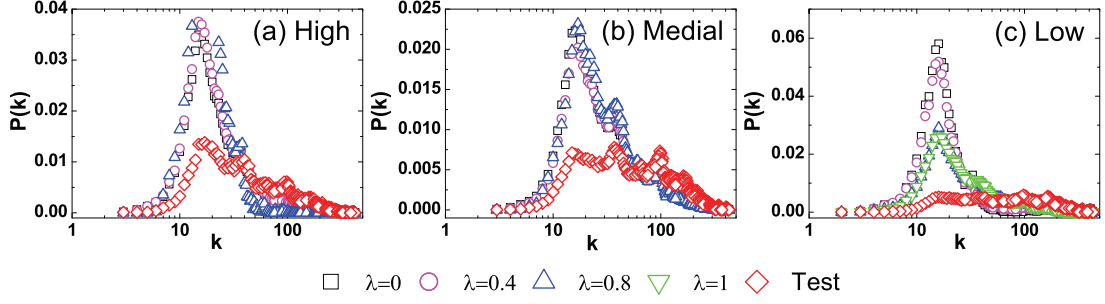
Fig. 7. (Color online.) The recommendation preference when considering the hybrid factors of the constant object degree information and the variable user degree information. We employ the random walker process which is described as $(D(u, 1 - \lambda)AD(o,1)A^T D(u,\lambda)A)$. In order to compare the prediction with the realization, we consider the tunable parameters $\lambda = 0, 0.4, 0.8, 1$ and the test set.

Fig. 8. (Color online.) Precision with the tunable parameter λ for the prediction of recommendation when considering the effect of the hybrid factors about the degree information of users and objects. Squares (□) indicate that the precision of the recommendation when the random walker dynamics with the constant user degree information and the tunable object degree information. Circles (○) report that the precision of recommendation when the process of random walk with the tunable user degree information and the constant object degree information.

ing can not recommend the objects in the high and medial level to the users when $\lambda = 1$ as shown in Fig. 7(a) and (b). When $\lambda = 0, 0.4, 0.8$, the preference of recommendation is similar, which suggests the tunable parameter λ could not affect the recommendation for the objects in the high and medial level considering the constant object degree information and the tunable variable user degree information. While in Fig. 7(c) in the low level, we could see that with λ increasing, the recommendation algorithm is less likely to recommend objects, which is different from that considering the constant user degree information and the tunable variable object degree information.

We also perform the analysis of the prediction precision when taking different hybrid factors into account as revealed in Fig. 8. When the process of random walk with the tunable user degree information and the constant object degree information as the parameter λ increases, the precision of recommendation always equals to 1 in the high level, and increases slowly in the medial level, while in the low level increases more slowly. That is to say, the precision of recommendation does not change much with the tunable parameter λ varying, which is different from the process of random walk affected by the tunable object degree information and the constant user degree information, in which as the parameter λ increases, the precision also rises to the highest value 1 and then keeps stable in the high level, while in the medial level the precision can reach the highest value 0.49 and then declines slowly. In the low level, the precision value always descends with the tunable parameter λ increasing. In general, it is easy to find optimal strategy for objects in high and medial level to recommend them accurately and efficiently, but for objects in low level, we should consider a better method to improve the accuracy of the recommendation.

5. Conclusions and discussions

So far, the collaborative filtering techniques and the interdisciplinary physics approaches attract a wealth of empirical and the-

oretical research and produce many successful E-commerce platforms. There could exist a simple general formula behind the collaborative filtering, interdisciplinary physics algorithms and even the extension methods. In agreement with the previous studies, we employed the dynamics of the random walks, and the framework of the general model had three key steps, which were (1) building the initial transition probability matrix with tunable parameters; (2) obtaining the track of random walkers; and (3) figuring out the walker tracks to predict the preference of users. We considered the simplest degree information in bipartite networks and proposed a generalized formula, which could deduce and develop the proposed well-known information filtering methods recently in the literatures such as collaborative filtering, network-based inference, heat conduction and hybrid method. Furthermore, we studied the generalized method with the single and hybrid degree information on the process of the random walk in the bipartite networks. The results stated clearly that the user degree information affected the recommendation preference which is in contrast with only considering the object degree information, further the performance of recommendation was affected by the process of random walk when taking account different single and hybrid degree information. As a result, we could suggest that using the two hybrid factors adjust the recommendation preference for different popular objects to toward promising precision of recommendation. Our findings could provide several pragmatic lessons for the efficient recommendation strategy in the E-business platforms and efficient interactive information prediction in online social networks. For instance, the different popular level objects could develop an appropriate strategy to make recommendations for users, as well as considering the preference of users with different activity. Meanwhile, this work leads to many extension, for instance, the other information of bipartite networks such as clustering, eigenvector, and assortative coefficient could also be applied in the random walk dynamics in the further research. In addition, we just develop the process of the random walk to information filtering considering simplest situation in our proposed generalized method, and we

will further study the machine learning technique to train the optimal road map of random walkers for offering the personalized recommendation.

Acknowledgements

References

[1] D.J. Watts, A twenty-first century science, Nature 445 (7127) (2007) 489.

[2] J.B. Schafer, J.A. Konstan, J. Riedl, E-commerce recommendation applications, in: Applications of Data Mining to Electronic Commerce, Springer, 2001, pp. 115–153.

[3] A. Vespignani, Predicting the behavior of techno-social systems, Science 325 (5939) (2009) 425.

[4] Z.-M. Ren, Y.-Q. Shi, H. Liao, Characterizing popularity dynamics of online videos, Physica A 453 (2016) 236–241.

[5] P.B. Kantor, L. Rokach, F. Ricci, B. Shapira, Recommender Systems Handbook, Springer, 2011.

[6] L. Lü, M. Medo, C.H. Yeung, Y.-C. Zhang, Z.-K. Zhang, T. Zhou, Recommender systems, Phys. Rep. 519 (1) (2012) 1–49.

[7] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl, Grouplens: an open architecture for collaborative filtering of netnews, in: Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, ACM, 1994, pp. 175–186.

[8] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, Item-based collaborative filtering recommendation algorithms, in: Proceedings of the 10th International Conference on World Wide Web, ACM, 2001, pp. 285–295.

[9] J.S. Breese, D. Heckerman, C. Kadie, Empirical analysis of predictive algorithms for collaborative filtering, in: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers Inc., 1998, pp. 43–52.

[10] D. Goldberg, D. Nichols, B.M. Oki, D. Terry, Using collaborative filtering to weave an information tapestry, Commun. ACM 35 (12) (1992) 61–70.

[11] L.H. Ungar, D.P. Foster, Clustering methods for collaborative filtering, in: AAAI Workshop on Recommendation Systems, vol. 1, 1998, pp. 114–129.

[12] L. Ungar, D.P. Foster, A formal statistical approach to collaborative filtering, CONALD'98.

[13] Y. Azar, A. Fiat, A. Karlin, F. McSherry, J. Saia, Spectral analysis of data, in: Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing, ACM, 2001, pp. 619–626.

[14] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer 8 (2009) 30–37.

[15] D.M. Blei, A.Y. Ng, M.I. Jordan, Latent Dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.

[16] R. Keshavan, A. Montanari, S. Oh, Matrix completion from noisy entries, in: Advances in Neural Information Processing Systems, 2009, pp. 952–960.

[17] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (1) (2002) 47–97.

[18] C. Castellano, S. Fortunato, V. Loreto, Statistical physics of social dynamics, Rev. Mod. Phys. 81 (2) (2009) 591–646.

[19] Y.-C. Zhang, M. Blattner, Y.-K. Yu, Heat conduction process on community networks as a recommendation model, Phys. Rev. Lett. 99 (15) (2007) 154301.

[20] T. Zhou, J. Ren, M. Medo, Y.-C. Zhang, Bipartite network projection and personal recommendation, Phys. Rev. E 76 (4) (2007) 046115.

[21] Y.-C. Zhang, M. Medo, J. Ren, T. Zhou, T. Li, F. Yang, Recommendation model based on opinion diffusion, Europhys. Lett. 80 (6) (2007) 68003.

[22] J.-H. Liu, Z.-K. Zhang, C. Yang, L. Chen, C. Liu, X. Wang, Gravity effects on information filtering and network evolving, PLoS ONE 9 (3) (2014) e91070.

[23] W. Zeng, A. Zeng, H. Liu, M.-S. Shang, T. Zhou, Uncovering the information core in recommender systems, Sci. Rep. 4 (2014) 6140.

[24] Q.-M. Zhang, A. Zeng, M.-S. Shang, Extracting the information backbone in online system, PLoS ONE 8 (5) (2013) e62624.

[25] T. Zhou, Z. Kuscsik, J.-G. Liu, M. Medo, J.R. Wakeling, Y.-C. Zhang, Solving the apparent diversity–accuracy dilemma of recommender systems, Proc. Natl. Acad. Sci. 107 (10) (2010) 4511–4515.

[26] A. Fiasconaro, M. Tumminello, V. Nicosia, V. Latora, R.N. Mantegna, Hybrid recommendation methods in complex networks, Phys. Rev. E 92 (1) (2015) 012811.

[27] A. Stojmirovic, Y.-K. Yu, Information flow in interaction networks, J. Comput. Biol. 14 (8) (2007) 1115–1143.

[28] J.-G. Liu, T. Zhou, Q. Guo, Information filtering via biased heat conduction, Phys. Rev. E 84 (3) (2011) 037101.

[29] M.-S. Shang, C.-H. Jin, T. Zhou, Y.-C. Zhang, Collaborative filtering based on multi-channel diffusion, Physica A 388 (23) (2009) 4867–4871.

[30] T. Zhou, L.-L. Jiang, R.-Q. Su, Y.-C. Zhang, Effect of initial configuration on network-based recommendation, Europhys. Lett. 81 (5) (2008) 58004.

[31] L. Lü, W. Liu, Information filtering via preferential diffusion, Phys. Rev. E 83 (6) (2011) 066119.

[32] F.-G. Zhang, A. Zeng, Information filtering via heterogeneous diffusion in online bipartite networks, PLoS ONE 10 (6) (2015) e0129459.

[33] J.-G. Liu, K. Shi, Q. Guo, Solving the accuracy–diversity dilemma via directed random walks, Phys. Rev. E 85 (1) (2012) 016118.

[34] Z.-K. Zhang, C. Liu, A hypergraph model of social tagging networks, J. Stat. Mech. 2010 (10) (2010) P10005.

[35] Z.-K. Zhang, T. Zhou, Y.-C. Zhang, Tag-aware recommender systems: a state-of-the-art survey, J. Comput. Sci. Technol. 26 (5) (2011) 767–777.

[36] L. Yu, C. Liu, Z.-K. Zhang, Multi-linear interactive matrix factorization, Knowl.-Based Syst. 85 (2015) 307–315.

[37] http://movielens.org and http://grouplens.org.