Heterogeneity Involved Network-based Algorithm Leads to Accurate and Personalized Recommendations

Tian Qiu¹*, Tian-Tian Wang¹, Zi-Ke Zhang²,³,⁴, Li-Xin Zhong⁵, Guang Chen¹

¹ School of Information Engineering, Nanchang Hangkong University, Nanchang, 330063, P. R. China
² Institute of Information Economy, Hangzhou Normal University, Hangzhou, 310036, P. R. China
³ Web Sciences Center, University of Electronic Science and Technology of China, Chengdu, 610054, P. R. China
⁴ Beijing Computational Science Research Center, Beijing, 100084, P. R. China
⁵ School of Finance, Zhejiang University of Finance and Economics, Hangzhou, 310018, P. R. China

* E-mail: tianqiu.edu@gmail.com

Abstract

Heterogeneity of both the source and target objects is taken into account in a network-based algorithm for the directional resource transformation between objects. Based on a biased heat conduction recommendation method (BHC) which considers the heterogeneity of the target object, we propose a heterogeneous heat conduction algorithm (HHC), by further taking the source object degree as the weight of diffusion. Tested on three real datasets, the Netflix, RYM and MovieLens, the HHC algorithm is found to present a better recommendation in both the accuracy and personalization than two excellent algorithms, i.e., the original BHC and a hybrid algorithm of heat conduction and mass diffusion (HHM), while not requiring any other accessorial information or parameter. Moreover, the HHC even elevates the recommendation accuracy on cold objects, referring to the so-called cold start problem, for effectively relieving the recommendation bias on objects with different level of popularity.

Introduction

The development of internet has made it easy to access information, which also brought about great convenience for our daily life. On the other hand, when facing various information, one is also puzzled how to get what he/she really wants. As a powerful tool, recommender system emerges to help people out of the overloaded information, which therefore attracts great interest of scientists from different disciplines [1], including physicists [2].

Different algorithms have been proposed, and achieved considerable progress. One of the most widely applied algorithm is the so-called collaborative filtering algorithm [4, 5], which can also be divided into the memory-based [6–8] and model-based collaborative filtering [9–11]. Another Line is the content-based algorithm [12]. Various extensive algorithms have been comprehensively investigated [13–17], such as the hybrid algorithms of collaborative filtering and content-based algorithm.

A recommender system can be taken as a complex system composed of some interactive units [18–24]. The complex interactions may affect the user’s activity. For example, it makes a movie well known by advertising for it via medias like internet or television. After watching the movie, audiences may feedback favorable comments or negative criticism, which on the other hand should have an impact on the potential audiences, as well as the popularity of the movie. The comments made by users to an extent reflect the individual preference. Now, it has accumulated a large amount of historical data of the users’ past activities, which makes it possible to design effective recommendation algorithms, and provide personalized recommendation for users by analyzing the data.
Physicists devote to the scientific program of pushing the personalized recommendation by applying the concepts from statistical physics. Different physical-concept-based recommendation algorithms have been proposed [25–32], and have presented great advantages in personalized recommendation. A typical example is the standard heat conduction algorithm (HC) [27], which generates a highly personalized but less accurate recommendation. To improve the recommendation accuracy, many improved algorithms have been studied from different perspectives [26,32,33], e.g., introducing accessorials information such as social tags [19,34,35]. Among those improved algorithms, a biased heat conduction algorithm (BHC) [36] is reported as an excellent one, for prominently enhancing both the recommendation accuracy and diversity, well resolving the long-standing challenge of accuracy-diversity dilemma, without any aid of accessorials information.

Most heat-conduction-incorporated algorithms describe the recommender system as a user-object bipartite network, where objects affect each other via the linked users. Similarly as the directional characteristic of the heat flow, the influence between objects are also found to be directional. If a resource is assigned to each object, the object disseminating resource is named as source object, while the object receiving resource is called target object. The standard heat conduction method does not consider the object heterogeneity. However, most network-based recommender systems show heterogeneous structures. Without effectively eliminating the recommendation bias on objects with different level of popularity induced by the system heterogeneity, algorithms generally cannot achieve an excellent performance in both accuracy and diversity. To alleviate the recommendation bias, the biased heat conduction algorithm is proposed [36], which takes the object heterogeneity into account, but unilaterally for the target objects, whereas overlooks the heterogeneity of the source objects. In fact, the source object also presents heterogeneous characteristic. For example, some objects are popular with large degrees, while others are cold with small degrees. Obviously, the popular objects would contribute to much more users than the cold objects. Therefore, it is essential to take the heterogeneity of both the source and target objects into consideration in the algorithm design.

In this article, based on the biased heat conduction algorithm [36], we further introduce the source object degree as the diffusion weight, and propose the heterogeneous heat conduction algorithm (HHC). Tested on three real datasets, the Netflix, RYM, and MovieLens, the HHC is found to present an excellent performance in both the recommendation accuracy and diversity, even more advantageous than the original biased heat conduction algorithm, and also an excellent hybrid algorithm of heat conduction and mass diffusion (HHM) [26], for effectively relieving the recommendation bias induced by the system heterogeneity. Moreover, the HHC meanwhile elevates the recommendation accuracy on cold objects, better resolving the so-called cold start problem.

Materials and Methods

A recommender system can be characterized by a bipartite graph composed of the user set $U$ containing $m$ users, and the object set $O$ containing $n$ objects. If an object $o_\alpha$ is collected by a user $u_i$, then add a link between them. The relation between the user and the object can be described by an adjacent matrix $A = \{a_{i\alpha}\}$, with $a_{i\alpha}$ to be 1 if there is a link between the user-object pair, otherwise, to be 0.

The standard heat conduction method is firstly proposed by Zhang et al [27], by introducing the heat conduction analogous process into the recommender systems. Assume each object has an initial resource. The resource would flow between different objects directionally like the heat. That is to say, the resource would spread from the source object to target object through the linked user, and all the objects would achieve a final resource. The resource diffusion process can be described by,

$$f' = Wf$$

where $W$ is the transformation matrix, characterizing the resource diffusion process from the source object to target object. $f$ is the initial resource of the object, and the $f'$ is the final resource. For each
user, rank his/her uncollected objects in the decreasing order of the final resource, and then recommend the top $L$ objects to the user. The simplest way to assign the initial resource to object is to set 1 or 0, on the basis of whether the object is collected by the user or not. The initial resource of user $u_i$ to object $o_{\alpha}$ can be described as,

$$\bar{f}_{i0(\alpha)} = a_{i\alpha}$$  \hspace{1cm} (2)

If an object is collected by the user $u_i$, its initial resource is assigned to be 1, otherwise, to be 0. All the following algorithms are based on the simplest assigning way of initial resource.

The process to redistribute the resources, here represented by the transformation matrix $W$, therefore plays a key role in the recommendation algorithm. An illustration of the diffusion process of the standard heat conduction algorithm (HC) is shown in Fig. 1 (a). At first, the particular user $i$ indicated by the solid circle receives an average level resource from his/her neighboring objects. Here the user $i$ has two neighbors, the first and the fourth neighbors, therefore can get the average resource of 1. Then the objects again get the average resources from all their neighboring users. The transformation matrix of the HC method can be formulated by,

$$W_{\alpha\beta}^{HC} = \frac{1}{k_{\alpha}} \sum_{j \in U} \frac{a_{\alpha j} a_{\beta j}}{k_j}$$  \hspace{1cm} (3)

where $k_{\alpha}$ is the degree of object $o_{\alpha}$, and $k_j$ is the degree of user $u_j$. The transformation matrix $W$ is found to be asymmetrical, i.e., the influence of the object $o_{\beta}$ to $o_{\alpha}$ is different from that of the object $o_{\alpha}$ to $o_{\beta}$. For the directional influence from the object $o_{\beta}$ to $o_{\alpha}$ in $W_{\alpha\beta}$, we notate the object $o_{\beta}$ as the source object, and $o_{\alpha}$ as the target object, respectively. The HC method assigns more priority to the low-degree objects, which leads to a highly personalized, but less accurate recommendation.

Generally, enhancing recommendation accuracy inhibits recommendation diversity. Effectively solving the accuracy-diversity dilemma has been a long-standing challenge of recommender systems. Recently, Liu et al proposed a biased heat conduction (BHC) method [36], by taking the heterogeneity of the target object into account, with its transformation matrix formulated by,

$$W_{\alpha\beta}^{BHC} = \frac{1}{k_{\alpha}} \sum_{j \in U} \frac{a_{\alpha j} a_{\beta j}}{k_j \gamma}$$  \hspace{1cm} (4)

where $\gamma$ is a tunable parameter. Compared with a number of network-based recommendation algorithms, the BHC shows a great advantage in both the recommendation accuracy and diversity [36]. In our study, based on the BHC, we further consider the heterogeneous effect of the source objects, and propose the heterogeneous heat conduction method (HHC).

In the network-based recommender systems, due to the heterogeneity of the source objects, the contribution of the source objects to users should be quite different. For instance, the popular objects usually have a big degree for widely collected by users, while the cold objects generally own a small degree. Hence, the popular objects should contribute to much more users than the cold objects. Using the source object degree as the diffusion weight, the transformation matrix of the HHC can be formulated by,

$$W_{\alpha\beta}^{HHC} = \frac{1}{k_{\alpha}} \sum_{j \in U} \frac{a_{\alpha j} a_{\beta j}}{k_j k_{\beta}}$$  \hspace{1cm} (5)

In order to show the advantages of the HHC method, we compare it with another excellent method, i.e., the hybrid method of heat conduction and mass diffusion (HHM) [26], which is outstanding in both the recommendation accuracy and diversity. So far, the network-based recommendation algorithm that outperforms the HHM is still rarely reported. Although there are few algorithms are reported to
outperform the HHM in some aspect, they usually introduce additional accessoril information or parameter. Therefore, comparing the HHC with the HHM can give a solid evidence of evaluating the algorithm performance, since the HHC does not require any accessoril information or extra parameter.

We firstly introduce the so-called mass diffusion method (MD), with an example of the diffusion process shown in Fig. 1 (b). At first, each object distributes the resource to its neighboring users with an equal probability. Then the user again redistributes all his/her resource to his/her neighboring objects, also with the equal probability. By summing up all the resources from their neighboring users, the objects then obtain their final resources. The transformation matrix of the MD algorithm is formulated by,

\[ W_{\alpha\beta}^{MD} = \frac{1}{k_\beta} \sum_{j \in U} \frac{a_{\alpha j} a_{\beta j}}{k_j}, \]

(6)

The MD method assigns more priority to the popular objects, since objects receive resources from all their neighboring users in the last step of diffusion. Such a diffusion pattern results in an excellent recommendation accuracy, yet a relatively poor diversity.

The hybrid method of heat conduction and mass diffusion (HHM) takes advantage of the two processes [26], with its transformation matrix formulated by,

\[ W_{\alpha\beta}^{HHM} = \frac{1}{k_\alpha} \left( 1 - \gamma \right) \frac{1}{k_\beta} \sum_{j \in U} \frac{a_{\alpha j} a_{\beta j}}{k_j}, \]

(7)

where \( \gamma \in [0, 1] \). When tuning the parameter \( \gamma \) to a proper value, the HHM method presents a high recommendation efficiency in both the accuracy and diversity.

**Metrics**

The most important aspect of assessing the performance of a recommendation algorithm is the recommendation accuracy. To give a solid evaluation of the accuracy, we use three indicators to present the accuracy performance, i.e., the ranking score \( \langle RS \rangle \), precision \( P \), and recall \( R \). On the other hand, to show how personalized the algorithm, we use the novelty and diversity to evaluate the personalization performance.

1. **Ranking Score** (\( \langle RS \rangle \)) [25].-In the network-based recommender system, if an object is collected by a user, i.e., there is a link between the user and the object, we take the object preferred by the user. The ranking score \( RS_{ui} \) then quantifies how the deleted link of object \( o_\alpha \) to user \( u_i \) in the test set rank in all \( u_i \)'s deleted links, which is defined as,

\[ RS_{ui} = \frac{p_\alpha}{n - k_i}. \]

(8)

where \( n \) is the number of all objects, \( k_i \) is the degree of the user \( u_i \), and \( p_\alpha \) is the position of the recommended object \( o_\alpha \) located in all the uncollected objects of the user \( u_i \). Obviously, the smaller the \( RS_{ui} \), the higher rank of the deleted link, and the more accurate the algorithm. The average ranking score \( \langle RS \rangle \) is taken an average of \( RS_{ui} \) over all the deleted links.

2. **Precision** (\( P \)) [39].-The recommendation precision \( P \) evaluates how the deleted links are recovered, which is defined as

\[ P = \frac{1}{m} \sum_{i=1}^{m} \frac{q_{iL}}{L}, \]

(9)

where \( q_{iL} \) is the number of the user \( u_i \)'s deleted links contained in the top \( L \) recommended object list. The higher the precision, the more accurate the recommendation, and vice versa.

3. **Recall** (\( R \)) [39].- The recall \( R \) is defined as
\[ R = \frac{1}{m} \sum_{i=1}^{m} q_{iL} l_i, \]  
(10)

where \( q_{iL} \) is the number of the user \( u_i \)'s deleted links contained in the top \( L \) recommended object list, \( l_i \) is the number of the user \( u_i \)'s deleted links in the test set. The higher the recall, the more accurate the recommendation, and vice versa.

4. **Novelty** (\( NL \)).- The novelty indicates how unexpected the recommended objects to the user. Here we use the average degree of the objects in the recommendation list to quantify the novelty, which is defined as,

\[ NL = \frac{1}{mL} \sum_{i=1}^{mL} \sum_{o_{i} \in O_{R}^i} k_{o_{i}^L}, \]
(11)

where \( O_{R}^i \) is the object set of the user \( u_i \)'s recommendation list. The smaller the \( NL \), the more novel the recommendation to the user, and vice versa.

5. **Hamming Distance** (\( H \)).- The recommendation diversity is quantified by the Hamming distance \( H \) of the recommendation lists of two different users, which is defined as,

\[ H = \frac{2}{m(m-1)} \sum_{i=1}^{m} \sum_{j=i+1}^{m} (1 - \frac{\emptyset_i \cap \emptyset_j}{L}), \]
(12)

where \( \emptyset_i \cap \emptyset_j \) is the number of the common objects recommended for the user \( u_i \) and \( u_j \) in the top \( L \) recommendation list. The Hamming distance \( H \) therefore evaluates how different the recommendation lists of users. The higher the \( H \), the less in common the recommended objects for different users, i.e., the more diverse the algorithm, and vice versa.

**Data**

In order to provide a comprehensive understanding of the algorithms, we employ the algorithms on three empirical datasets, i.e., the Netflix, RYM, and MovieLens. The Netflix and MovieLens are both five-level rating movie systems, and the RYM is a ten-level rating music system. The Netflix is randomly selected from the huge dataset of the Netflix Prize, the MovieLens is downloaded from the web site of GroupLens Research [http://grouplens.org], and the RYM is downloaded from the music rating web site RateYourMusic.com.

To construct the bipartite network, a link between a user and an object is added if the rating of the user to the object is no less than three for the Netflix and MovieLens, and no less than six for RYM. The basic statistics of the three datasets is summarized in table 1. The sparsity of the dataset is defined as the number of links proportional to the total number of the user-object links. To test the algorithm performance, all the links of the network are split into two subsets, i.e., the training set and the test set. Randomly delete 10% links as the test set to test the performance of the algorithm, and remain the rest 90% links as the training set to make predictions for users.

**Results and Discussion**

To evaluate the algorithm performance, we compare the results of the HHC with the excellent BHC and HHM algorithms, with the results being the average over six runs. The recommendation accuracy results are presented by the indicators of ranking score (\( RS \)), precision \( P \), and recall \( R \). For all the three algorithms, one tunable parameter is introduced. To obtain the optimal value of the tunable parameter \( \gamma \), we investigate the ranking score on \( \gamma \). As shown in Fig. 2, for the HHC, the minimal value of the
ranking score is obtained at $\gamma = 0.68, 0.61 \text{ and } 0.74$ for the Netflix, RYM and MovieLens, respectively. Similar procedure is also carried for the BHC and HHM method. For the BHC method, the minimal value of the ranking score is obtained at $\gamma = 0.85, 0.80 \text{ and } 0.85$ for the Netflix, RYM and MovieLens, and for the HHM method, the minimal value of the ranking score is obtained at $\gamma = 0.17, 0.25 \text{ and } 0.17$ for the Netflix, RYM and MovieLens. The following results are obtained at the optimal value of $\gamma$.

0.1 Accuracy and Personalization of Recommendation

The performance of the HHM, BHC and HHC algorithms are summarized in table 2. For the ranking score $\langle RS \rangle$ and the precision $P$, the HHC algorithm outperforms both the HHM and BHC for all the three datasets. The recall $R$ of the RYM and MovieLens of the HHC is also more advantageous than the HHM and BHC. It indicates that the HHC is more effective in recommendation accuracy.

To quantitatively examine how much the HHC outperforms the HHM and BHC, we define an improvement percentage as $\delta_{\text{ALG}} = (Q_{\text{SCL}} - Q_{\text{ALG}})/Q_{\text{ALG}}$, where the subhead ALG refers to the investigated algorithm, and the $Q_{\text{ALG}}$ is the value of the indicator. The results of the improvement percentage are summarized in table 3. Taking the Netflix as an example, the improvement percentage of the ranking score of the HHC against the BHC and HHM is found to be as much as 14.6% and 8.9%. Such an improvement is very appreciable, since the HHC algorithm does not introduce any accessoril information or additional parameter.

The novelty $NL$ and diversity $H$ are also investigated to evaluate how personalized the algorithm. As shown in table 1, taking the Netflix as an example, there is also a great improvement of the novelty for the HHC, with the improvement percentage of the HHC against the BHC and HHM as much as 8.2% and 11.5%. Similarly, the diversity of the HHC method also shows a great advantage, with the improvement percentage of the HHC against the BHC and HHM to be 10.4% and 12.5%. It suggests that the HHC is not only accurate, but also personalized.

To show how the personalized indicators evolve with the recommendation list length, we study the novelty $NL$ and diversity $H$ for different recommendation list lengths $L$. As shown in Fig. 3, for the Netflix and RYM, the novelty of the HHC is more advantageous than the HHM and BHC for a wide range of recommendation list lengths. For the MovieLens, the novelty of the HHC is also more advantageous than the BHC, and is similar as the HHM algorithm. The diversity $H$ on the recommendation list lengths is shown in Fig. 4. It is observed, for all the three datasets, the diversity of the HHC is much more advantageous than that of both the HHM and BHC for all the investigated range of the recommendation list lengths. It further confirms the results in table 2 and 3.

Our results have shown that the HHC is more advantageous than the BHC and HHM in the recommendation accuracy, novelty, as well as diversity. To further test the robustness of the results, we adjust the proportion of the training set to the total data, and investigate the corresponding ranking score $\langle RS \rangle$, Hamming distance $H$ and novelty $NL$. As shown in Fig. 5, taking the Netflix dataset as an example, it is observed that for all the investigated training set ratios, even for the sparse training data condition with its ratio as low as 0.4., the HHC outperforms the BHC and HHM, for all the three metrics, i.e., recommendation accuracy, novelty and diversity. It provides a further solid evidence of the great advantages of the proposed HHC method.

Why the HHC can achieve such an appreciable improvement in both aspects of accuracy and personalization? By analyzing the statistical property of the dataset, a great heterogeneity is observed for the data. As shown in Fig. 6 (a), a power-law-like probability distribution of the object degree is observed for the Netflix, which indicates that a large proportion of the objects owns very small degrees, whereas a small fraction of the objects has big degrees. Assume that objects with big degrees are popular objects, and with small degrees are cold objects. Focusing on the recommendation of either popular objects or cold objects cannot achieve an effective recommendation simultaneously in the accuracy and diversity. Typical examples are the standard heat conduction method and the mass diffusion method. The standard heat conduction algorithm emphasizes the recommendation of cold objects, leading to a very personalized
but less accurate recommendation, whereas the mass diffusion algorithm tends to recommend popular objects, generating an accurate yet relatively less personalized recommendation. Eliminating the recommendation bias on objects with different level of popularity induced by the system heterogeneity is essentially required for designing effective algorithms. By analyzing the degree distribution $p(k)$ for the objects in the top $L = 50$ recommendation list shown in Fig. 6 (b), we find that the BHC and HHM methods have well resolved the recommendation for both the popular and cold objects. Compared with the BHC and HHM methods, the HHC obtains a better recommendation balance between the cold objects and popular objects, which to an extent explains why the HHC can further improve both the accuracy and personalization of recommendation.

To better understand the obtained results, we then analyze the eigenvalues of the transformation matrix $W$ for the BHC, HHM, and HHC. As shown in Fig. 7, the eigenvalues are displayed in the decreasing order of the values for the three methods. The large eigenvalues indicate that there exist advantageous components in the $W$. For the BHC method, several large eigenvalues are observed for the Netflix, RYM and MovieLens. Similar large eigenvalues are also observed in the HHM method, but the values of the first several eigenvalues are smaller than those of the BHC. However, for the HHC method, nearly all the eigenvalues show very close values, i.e., all the components have a near effect. That is to say, regardless of the cold or popular objects, the HHC assigns an approximate weight for all the objects. The frequency distribution of the eigenvalues of the transformation matrix $W$ is shown in Fig. 8 for the Netflix. The eigenvalue distribution of the HHC method indicates an apparently more homogeneous distribution than that of the BHC and HHM, with the largest eigenvalue of the HHC to be 0.35, far less than 2.68 of the BHC. It well accounts for the success of the HHC in relieving the recommendation bias. The results of the eigenvalue analysis of the transformation matrix $W$ are consistent with the observations from the degree distribution of the top $L$ recommended objects, which provides a deeper insight into understanding the underlying mechanism of the algorithm.

### 4.2 Accuracy on Cold Objects

From the analysis of the object degree distribution and the eigenvalues of the transformation matrix, the cold objects are observed to get more priority in the HHC, compared with the BHC and HHM. It reminds us of the so-called the ‘cold-start’ problem, referring to how to recommend the new objects or cold objects, i.e., the cold object start [10], or how to recommend objects to newly added users, i.e., the cold user start [11]. Here we focus on the cold object start.

Due to lack of information, it is hard for users to be aware of the cold objects, resulting in the difficulty to recommend them effectively. However, the cold objects usually occupy a big proportion of the total objects. If we define the object with its degree no more than 10 as the cold object, there are as much as 49.59%, 21.73%, and 41.26% cold objects for the Netflix, RYM and MovieLens. Hence, how to solve the cold start problem is an important but challenging problem in recommender systems.

To present the recommendation accuracy on the cold objects, we investigate an object-dependent ranking score $\langle RS \rangle_k$, which is defined as the average ranking score over objects with the same value of degrees [18]. As shown in Fig. 9, whereas showing similar recommendation accuracy on the popular objects with large degrees, the $\langle RS \rangle_k$ of the low-degree objects of the HHC is found to be smaller than that of both the BHC and HHM for all the three datasets. Focusing on objects whose degree are no more than 10, the $\langle RS \rangle_{k \leq 10}$ of the HHC is found to be much smaller than that of both the BHC and HHM, as shown in table 2. Taking the Netflix as an example, the improvement percentage is as much as 23.0% against the BHC, and 18.2% against the HHM. Similar improvement is also found for the RYM and MovieLens.

Further, we investigate an object-dependent precision $P_k$ and an object-dependent precision $R_k$. $P_k$ is defined as $P_k = \frac{1}{m} \sum_{i=1}^{m} \frac{q_i^k}{l_i^k}$, and $R_k$ is defined as $R_k = \frac{1}{m} \sum_{i=1}^{m} \frac{q_i^k}{l_i^k}$, where $q_i^k$ is the number of the user $u_i$’s deleted links for objects with degree $k$ in the top $L$ recommended object list, and $l_i^k$ is the number
of the user $u_i$’s deleted links for objects with degree $k$ in the test set. The larger the $P_k$ or $R_k$, the more accurate the recommendation on cold objects, and vice versa. As shown in Table 2, both the $P_{k \leq 10}$ and $R_{k \leq 10}$ of the HHC greatly outperform those of the BHC and HHM. For the Netflix, the improvement percentage of $P_{k \leq 10}$ is as much as 1019.4% against the BHC, and 508.8% against the HHM, and of $R_{k \leq 10}$ is 900.0% against the BHC, and 511.1% against the HHM. Similar improvement is also found for the RYM and MovieLens. It gives a strong evidence that the HHC better resolves the cold start problem than both the BHC and HHM.

By studying the three indicators of recommendation accuracy on cold objects, i.e., the $\langle RS \rangle_{k \leq 10}$, $P_{k \leq 10}$ and $R_{k \leq 10}$, for different ratios of the training set to the total data, the robustness of the results can be further confirmed. As shown in Fig. 10, taking the Netflix as an example, the $R_{k \leq 10}$ of the HHC is much smaller than that of the BHC and HHM, and the $P_{k \leq 10}$ and $R_{k \leq 10}$ of the HHC are much higher than those of the BHC and HHM, for all the investigated ratios, even for the sparse training data with the ratio as low as 0.3. Similar behavior is also found for the RYM and MovieLens.

As mentioned above, users are seldom aware of the cold objects. So far, most studies try to solve the cold start problem with the aid of other accessorial information, such as the trust relationship [42], and tags [3, 19, 34, 37], or introducing additional parameter [38]. However, these accessorial information or new parameter would make the recommender system more complex. Employing the source object degree as the weight of diffusion, and without introducing any accessorial information, the HHC further elevates the recommendation accuracy on cold objects, better resolving the cold start problem for fully taking the heterogeneity of objects into account.

**Conclusion**

In conclusion, based on an excellent biased heat conduction method (BHC), we consider the heterogeneity of both the source and target objects, and propose the heterogeneous heat conduction recommendation algorithm (HHC), by taking the source object degree as the weight of diffusion. Without employing any other accessorial information or additional parameter, the HHC outperforms the original BHC method, and even a highly accurate and personalized hybrid method of heat conduction and mass diffusion (HHM), in both the recommendation accuracy and personalization. Moreover, the HHC further enhances the recommendation accuracy on cold objects, referring to the so-called cold start problem.

Due to the difficulty to alleviate the recommendation bias on objects with different level of popularity, improving recommendation accuracy usually inhibits recommendation diversity, causing the great challenge of accuracy-diversity dilemma. Especially, large amounts of new objects emerge in the online applications, which not only even intensifies the system heterogeneity, but also brings about the cold start problem. The HHC not only further improves both the recommendation accuracy and personalization, but also enhances the recommendation accuracy on cold objects, for greatly relieving the recommendation bias induced by the system heterogeneity. Our work might shed some new light on better understanding and designing effective recommendation algorithms.

**References**

1. Adomavicius G, Tuzhilin A (2005) Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Trans Knowl Data Eng 17: 734–749.
2. Lü L, Medo M, Yeung CH, Zhang YC, Zhang ZK, et al. (2012) Recommender systems. Phys Rep 519: 1–49.
3. Zhang ZK, Zhou T, Zhang YC (2011) Tag-aware recommender systems: a state-of-the-art survey. J Comput Sci Technol 26: 767–777.
4. Goldberg D, Nichols D, Oki BM, Terry D (1992) Using collaborative filtering to weave an information tapestry. Commun ACM 35: 61–70.

5. Schafer JB, Frankowski D, Herlocker J, Sen S (2007) Collaborative filtering recommender systems. In: The adaptive web, Springer. pp. 291–324.

6. Breese JS, Heckerman D, Kadie C (1998) Empirical analysis of predictive algorithms for collaborative filtering. In: Proc. 14th Conf. Uncertainty Artif. Intel. Morgan Kaufmann Publishers Inc., pp. 43–52.

7. Nakamura A, Abe N (1998) Collaborative filtering using weighted majority prediction algorithms. In: Proc. 5th Intl. Conf. Mach. Learn. pp. 395–403.

8. Delgado J, Ishii N (1999) Memory-based weighted majority prediction. In: SIGIR Workshop Recomm. Syst. Citeseer.

9. Getoor L, Sahami M (1999) Using probabilistic relational models for collaborative filtering. In: Workshop Web Usage Anal. User Profil. Citeseer.

10. Hofmann T (2003) Collaborative filtering via gaussian probabilistic latent semantic analysis. In: Proc. 26th Ann. Intl. SIGIR Conf. Research Devel. Infor. Retr. ACM, pp. 259–266.

11. Marlin B (2003) Modeling user rating profiles for collaborative filtering. Adv Neural inf Process Syst 16.

12. Pazzani MJ, Billsus D (2007) Content-based recommendation systems. In: The adaptive web, Springer. pp. 325–341.

13. Balabanović M, Shoham Y (1997) Fab: content-based, collaborative recommendation. Comm ACM 40: 66–72.

14. Goldberg K, Roeder T, Gupta D, Perkins C (2001) Eigentaste: A constant time collaborative filtering algorithm. Infor Retr 4: 133–151.

15. Hofmann T (2004) Latent semantic models for collaborative filtering. ACM Trans Inf Syst 22: 89–115.

16. Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3: 993–1022.

17. Laureti P, Moret L, Zhang YC, Yu YK (2006) Information filtering via iterative refinement. EPL 75: 1006.

18. Zhou T, Jiang LL, Su RQ, Zhang YC (2008) Effect of initial configuration on network-based recommendation. EPL 81: 58004.

19. Zhang ZK, Zhou T, Zhang YC (2010) Personalized recommendation via integrated diffusion on user–item–tag tripartite graphs. Physica A 389: 179–186.

20. Shang MS, Zhang ZK, Zhou T, Zhang YC (2010) Collaborative filtering with diffusion-based similarity on tripartite graphs. Physica A 389: 1259–1264.

21. Liu RR, Jia CX, Zhou T, Sun D, Wang BH (2009) Personal recommendation via modified collaborative filtering. Physica A 388: 462–468.

22. Liu JG, Shi K, Guo Q (2012) Solving the accuracy-diversity dilemma via directed random walks. Phys Rev E 85: 016118.
23. Liu JG, Guo Q, Zhang YC (2011) Information filtering via weighted heat conduction algorithm. Physica A 390: 2414–2420.

24. Zeng A, Yeung CH, Shang MS, Zhang YC (2012) The reinforcing influence of recommendations on global diversification. EPL 97: 18005.

25. Zhou T, Ren J, Medo M, Zhang YC (2007) Bipartite network projection and personal recommendation. Phys Rev E 76: 046115.

26. Zhou T, Kuscsik Z, Liu JG, Medo M, Wakeling JR, et al. (2010) Solving the apparent diversity-accuracy dilemma of recommender systems. Proc Natl Acad Sci USA 107: 4511–4515.

27. Zhang YC, Blattner M, Yu YK (2007) Heat conduction process on community networks as a recommendation model. Phys Rev Lett 99: 154301.

28. Zhou T, Su RQ, Liu RR, Jiang LL, Wang BH, et al. (2009) Accurate and diverse recommendations via eliminating redundant correlations. New J Phys 11: 123008.

29. Ren J, Zhou T, Zhang YC (2008) Information filtering via self-consistent refinement. EPL 82: 58007.

30. Zhou YB, Lei T, Zhou T (2011) A robust ranking algorithm to spamming. EPL 94: 48002.

31. Qiu T, Zhang ZK, Chen G (2013) Information filtering via a scaling-based function. PLoS ONE 8: e63531.

32. Lü L, Liu W (2011) Information filtering via preferential diffusion. Phys Rev E 83: 066119.

33. Qiu T, Chen G, Zhang ZK, Zhou T (2011) An item-oriented recommendation algorithm on cold-start problem. EPL 95: 58003.

34. Kim HN, Ji AT, Ha I, Jo GS (2010) Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. Electron Commerce Research Appl 9: 73–83.

35. Zhang ZK, Liu C (2012) Hybrid recommendation algorithm based on two roles of social tags. Int J Bifurcat Chaos 22: 1250166.

36. Liu JG, Zhou T, Guo Q (2011) Information filtering via biased heat conduction. Phys Rev E 84: 037101.

37. Zhang ZK, Liu C, Zhang YC, Zhou T (2010) Solving the cold-start problem in recommender systems with social tags. EPL 92: 28002.

38. Liu C, Zhou WX (2012) Heterogeneity in initial resource configurations improves network-based hybrid recommendation algorithm. Physica A 391: 5704-5711.

39. Herlocker JL, Konstan JA, Terveen LG, Riedl JT (2004) Evaluating collaborative filtering recommender systems. ACM Trans Inf Syst 22: 5–53.

40. Park YJ, Tuzhilin A (2008) The long tail of recommender systems and how to leverage it. In: Proc. 2008 ACM Conf. Recomm. Syst. ACM, pp. 11–18.

41. Lam XN, Vu T, Le TD, Duong AD (2008) Addressing cold-start problem in recommendation systems. In: Proc. 2nd Intl. Conf. Ubiquitous Infor. Manag. Commun. ACM, pp. 208–211.

42. Jamali M, Ester M (2009) Trustwalker: a random walk model for combining trust-based and item-based recommendation. In: Proc. 15th ACM SIGKDD Intl Conf. Knowl. Disc. Data Mining. ACM, pp. 397–406.
Figure 1. The illustration of the resource transformation. (a) for the heat conduction process, and (b) for mass diffusion process.

Tables

Table 1. The statistics of the datasets.

| Dataset   | Users | Objects | Links    | Sparsity |
|-----------|-------|---------|----------|----------|
| Netflix   | 9999  | 5870    | 815917   | 1.39 %   |
| RYM       | 10159 | 5250    | 559634   | 1.05 %   |
| MovieLens | 943   | 1682    | 100000   | 6.30 %   |

The statistics of the Netflix, RYM, and MovieLens is displayed. From left to right, the columns correspond to the name of the data, the number of users, objects, and links, and the sparsity of the data.
Figure 2. The ranking score $\langle RS \rangle$ on the tunable parameter $\gamma$ of the HHC algorithm. (a) for the Netflix, (b) for the RYM and (c) for the MovieLens.

Table 2. The performance of the HHM, BHC and HHC methods.

|        | $\langle RS \rangle$ | $P$   | $R$   | $NL$ | $H$   | $\langle RS \rangle_{k\leq 10}$ | $P_{k\leq 10}$ | $R_{k\leq 10}$ |
|--------|----------------------|-------|-------|------|-------|-------------------------------|----------------|----------------|
| Netflix| HHM 0.045            | 0.062 | 0.470 | 1844 | 0.672 | 0.417                         | 0.00057        | 0.018          |
|        | BHC 0.048            | 0.061 | 0.456 | 1777 | 0.685 | 0.443                         | 0.00031        | 0.011          |
|        | HHC 0.041            | 0.062 | 0.453 | 1632 | 0.756 | 0.341                         | 0.00347        | 0.110          |
| RYM    | HHM 0.048            | 0.050 | 0.557 | 330  | 0.940 | 0.250                         | 0.00241        | 0.092          |
|        | BHC 0.050            | 0.049 | 0.542 | 317  | 0.942 | 0.255                         | 0.00200        | 0.081          |
|        | HHC 0.045            | 0.051 | 0.571 | 324  | 0.947 | 0.221                         | 0.00478        | 0.172          |
| MovieLens| HHM 0.083          | 0.085 | 0.527 | 157  | 0.839 | 0.408                         | 0.00113        | 0.044          |
|        | BHC 0.084            | 0.084 | 0.515 | 155  | 0.839 | 0.413                         | 0.00107        | 0.042          |
|        | HHC 0.079            | 0.088 | 0.544 | 153  | 0.859 | 0.368                         | 0.00202        | 0.062          |

The ranking score $\langle RS \rangle$, precision $P$, recall $R$, novelty $NL$, diversity $H$, object-dependent ranking score $\langle RS \rangle_{k\leq 10}$, object-dependent precision $P_{k\leq 10}$ and object-dependent recall $R_{k\leq 10}$ of the HHM, BHC and HHC algorithms are shown for the Netflix, RYM and MovieLens, with $L = 50$. 
Figure 3. The novelty $NL$ on the recommendation list length $L$. (a) for the Netflix, (b) for the RYM and (c) for the MovieLens. The magenta, grey and cyan lines are for the BHC, HHM and HHC methods, respectively.

Figure 4. The Hamming distance $H$ on the recommendation list length $L$. (a) for the Netflix, (b) for the RYM and (c) for the MovieLens. The magenta, grey and cyan lines are for the BHC, HHM and HHC methods, respectively.
Figure 5. The ranking score $\langle RS \rangle$, novelty $NL$, and Hamming distance $H$ are displayed for different ratios of the training set to the total data in (a), (b) and (c) for the Netflix, respectively. The magenta, grey and cyan lines are for the BHC, HHM and HHC methods, respectively.

Figure 6. (a) The object degree distribution $p(k)$ of the Netflix, and (b) the degree distribution $p(k)$ for the objects in the top $L = 50$ recommendation list. In the subplot (b), the magenta, grey and cyan lines are for the BHC, HHM and HHC methods, respectively.
Figure 7. The eigenvalues of the transformation matrix $W$ are displayed in the decreasing order of the values. (a) for the Netflix, (b) for the RYM and (c) for the MovieLens. The magenta, grey and cyan lines are for the BHC, HHM and HHC methods, respectively.

Figure 8. The frequency distribution of the eigenvalues $\lambda$ of the transformation matrix $W$ is displayed for the Netflix. The inset shows the first several largest eigenvalues. (a) for the BHC method, (b) for the HHM method and (c) for HHC method.
Figure 9. The object-dependent ranking score $\langle RS \rangle_k$ vs. the object degree $k$. (a) for the Netflix, (b) for the RYM and (c) for the MovieLens. The magenta, grey and cyan lines are for the BHC, HHM and HHC methods, respectively.

Figure 10. The object-dependent ranking score $\langle RS \rangle_{k \leq 10}$, precision $P_{k \leq 10}$, and recall $R_{k \leq 10}$ are displayed for different ratios of the training set to the total data in (a), (b) and (c) for the Netflix, respectively. The magenta, grey and cyan lines are for the BHC, HHM and HHC methods, respectively.
Table 3. The improvement percentage of the HHC against the HHM and BHC methods.

| Dataset        | $\delta_{HHM}$ | $\langle RS \rangle$ | $P$  | $R$  | $NL$ | $H$  | $\langle RS \rangle_{k \leq 10}$ | $P_{k \leq 10}$ | $R_{k \leq 10}$ |
|----------------|-----------------|----------------------|------|------|------|------|-------------------------------|------------------|------------------|
| Netflix        | 8.9%            | 0.0%                 | -3.6%| 11.5%| 12.5%| 18.2%| 508.8%                       | 511.1%           |                  |
|                | 14.6%           | 1.6%                 | -0.7%| 8.2% | 10.4%| 23.0%| 1019.4%                      | 900.0%           |                  |
| $\delta_{BHC}$| 6.3%            | 2.0%                 | 2.5% | 1.8% | 0.7% | 11.6%| 98.3%                         | 87.0%            |                  |
|                | 10.0%           | 4.1%                 | 5.4% | -2.2%| 0.5% | 13.3%| 139.0%                       | 112.3%           |                  |
| RYM            | 4.8%            | 3.5%                 | 3.2% | 2.5% | 2.4% | 9.8% | 78.8%                         | 40.9%            |                  |
|                | 6.0%            | 4.8%                 | 5.6% | 1.3% | 2.4% | 10.9%| 88.8%                         | 47.6%            |                  |
| MovieLens      | 4.8%            | 3.5%                 | 3.2% | 2.5% | 2.4% | 9.8% | 78.8%                         | 40.9%            |                  |
|                | 6.0%            | 4.8%                 | 5.6% | 1.3% | 2.4% | 10.9%| 88.8%                         | 47.6%            |                  |

The improvement percentage of the HHC against the HHM and BHC in the ranking score $\langle RS \rangle$, precision $P$, recall $P$, novelty $NL$, diversity $H$, object-dependent ranking score $\langle RS \rangle_{k \leq 10}$, object-dependent precision $P_{k \leq 10}$ and object-dependent recall $R_{k \leq 10}$ is shown for the Netflix, RYM and MovieLens, with $L = 50$. To guide the eyes, if the HHC outperforms other methods, we show the improvement percentage as a positive value, otherwise, as a negative value.