The reinforcing influence of recommendations on global diversification

An Zeng[1], Chi Ho Yeung[1], Ming-Sheng Shang[2] and Yi-Cheng Zhang[1,2(a)]

[1] *Department of Physics, University of Fribourg - Chemin du Musée 3, CH-1700 Fribourg, Switzerland*
[2] *Web Sciences Center, School of Computer Science and Engineering, University of Electronic Science and Technology of China - Chengdu 610054, PRC*

Abstract – Recommender systems are promising ways to filter the abundant information in modern society. Their algorithms help individuals to explore decent items, but it is unclear how they distribute popularity among items. In this paper, we simulate successive recommendations and measure their influence on the dispersion of item popularity by Gini coefficient. Our result indicates that local diffusion and collaborative filtering reinforce the popularity of hot items, widening the popularity dispersion. On the other hand, the heat conduction algorithm increases the popularity of the niche items and generates smaller dispersion of item popularity. Simulations are compared to mean-field predictions. Our results suggest that recommender systems have reinforcing influence on global diversification. Finally, the study of the hybrid method of mass diffusion and heat conduction reveals that the influence of recommender systems is actually controllable.

Introduction. – Due to the rapid expansion of the internet, we are overloaded by an increasing amount of information from the World Wide Web [1]. For instance, one has to choose among millions of candidate commodities to shop online. Comprehensive exploration is infeasible [2]. Various recommendation approaches have thus been proposed to help filtering the relevant information [3,4]. The recommendation algorithms include popularity-based (PR) method, collaborative filtering (CF) method [5,6], mass diffusion (MD) method [7], heat conduction (HC) method [8], the hybrid method of mass diffusion and heat conduction [9] and so on. In general, they use the activity record and available personal profiles of users to uncover their potential preferences.

Though recommendation algorithms are helpful in filtering information, they may significantly influence the distribution of items' popularity (*i.e.* the degree distribution of items). This is caused by guiding people's choices, which influence subsequent recommendations and hence the choices of others. The influence is amplified with successive recommendations. For example, if a recommendation algorithm always recommends popular items, gradually only the most popular items survive, causing the market to be further dominated by these items. On the other hand, if a recommendation algorithm tends to recommend less popular items, item popularity will become homogeneous. In these cases, recommendations impose a reinforcing influence on the dispersion of items' popularity, *i.e.* the *diversity*, which affects subsequent choices.

Actually, understanding the reinforcing influence of recommender system is of great significance. From the theoretical point of view, it presents a physics perspective and utilizes microscopic interactions to explain macroscopic behaviors of recommender systems [10,11], unlike most existing works which are devoted to improve recommendation accuracy [6]. It is also worth noting that similar studies on evolution of movie popularity [12,13] have resulted in consistent predictions compared with the observed data. In a practical sense, one can control the diversity of commodities in online retailers with the help of recommender systems. In particular, when one considers the reinforcing influence on global diversity as an undesired side-effect of recommending systems, theoretical understanding may provide a solution to minimize the effect.

(a) E-mail: `Yi-Cheng.Zhang@unifr.ch`

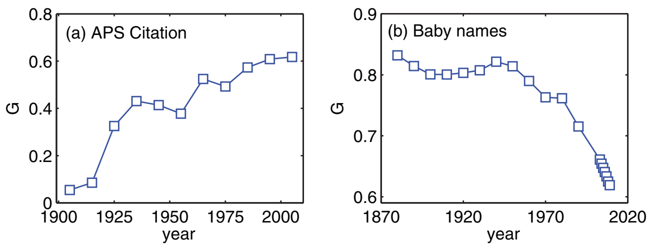

Fig. 1: (Color online) The change of Gini coefficient with time in the APS citation data and the baby name data.

In this paper, we use the Gini coefficient to measure the diversity of the system which is actually the dispersion in item popularity [14]. We note that a small dispersion implies similar popularity among items, and hence diverse recommendations for users. If the dispersion is large, some items dominate in popularity and users have limited choices. Our result indicates that MD and CF reinforce the popularity of already popular items, as similar to PR. On the other hand, HC increases the popularity of the niche items and generates smaller dispersion in item popularity. Our results suggest that recommender systems indeed have reinforcing influence on global diversification.

Dispersion of item popularity. – To quantify global diversity, we use Gini coefficient G [14] which measures the dispersion of item popularity, as in the case of individual wealth. The Gini coefficient has also been used to measure dispersion in sociology, science and engineering. Mathematically, it is given by

$$G = 1 - 2\int_0^1 C(x)\,\mathrm{d}x, \qquad (1)$$

where $C(x)$ is the normalized cumulative popularity when items are ranked in ascending order of popularity, with x being the normalized rank. Specifically, $G=0$ corresponds to uniform popularity among items, while $G=1$ corresponds to maximal dispersion.

When the dispersion of popularity changes with time, the changes can be well captured by the Gini coefficient [15]. As an example, we show such changes on data of scientific citations and baby names[1]. The results are reported in fig. 1, from which we can see the G increases in APS citation system while it decreases for baby names. One possible reason is the improving information accessibility: good papers have wider spread among the research communities and are cited more, which leads to a larger G; on the other hand, parents know more candidate names for babies to avoid overlap, resulting in a smaller G. In the next section, we will study the influence of the recommender systems on the dispersion of item popularity.

[1]The scientific citation data is based on the citation relation in the APS (American Physics Society) journals from 1893 to 2009 [16], and the baby name data is based on the first names taken from US Social Security Administration, and contain the top 1000 boy and girl names from 1880 to 2009 [17].

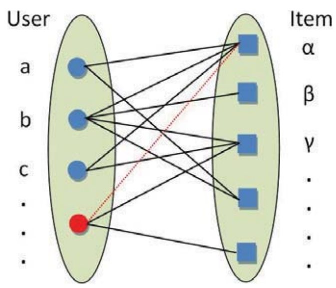

Fig. 2: (Color online) An illustrative example of the evolution of the bipartite network. The red node corresponds to the active user, and the red link corresponds to the choice made by the user according to recommendation results.

Table 1: Abbreviations for each recommendation method.

Method	Acronym
Mass Diffusion	MD
Heat Conduction	HC
User-based Collaborative Filtering	UCF
Item-based Collaborative Filtering	ICF
Popularity-based Recommendation	PR
Random Recommendation	RR

The reinforcing influence of recommender systems. – We investigate in this section the influence of recommender systems on the global diversity. Here we consider four simple and fundamental recommendation algorithms, including mass diffusion (MD), heat conduction (HC), user-based collaborative filtering (UCF) and item-based collaborative filtering (ICF). In addition, we consider two benchmark algorithms including popularity-based recommendation (PR) and random recommendation (RR), corresponding to the recommendations of respectively most popular and random items. Abbreviations for each recommendation method are given in table 1.

Online commercial systems can be well described by bipartite networks as shown in fig. 2, where circles represent users and squares represent items. If a user collects an item, a link is drawn between them. Specifically, we consider a system of N users and M items represented by a bipartite network with adjacency matrix A, where the element $a_{i\alpha}=1$, if a user i has collected an object α, and $a_{i\alpha}=0$, otherwise (throughout this paper we use Greek and Latin letters, respectively, for object- and user-related indices).

For a target user i, the MD algorithm starts by assigning one unit of resource to each object collected by i, and redistributes the resource through the user-item network. We denote the vector $\mathbf{f}^i$ as the initial resources on items, where the α-th component f^i_α is the resource possessed by object α. Recommendations for the user i are obtained by setting the elements in $\mathbf{f}^i$ to be $f^i_\alpha = a_{i\alpha}$, in accordance with the objects the user has already collected. The redistribution is represented by $\widetilde{\mathbf{f}}^i = W\mathbf{f}^i$, where

$$W_{\alpha\beta} = \frac{1}{k_\beta}\sum_{j=1}^{N}\frac{a_{j\alpha}a_{j\beta}}{k_j}, \quad (2)$$

is the diffusion matrix, with $k_\beta = \sum_{l=1}^{N} a_{l\beta}$ and $k_j = \sum_{\gamma=1}^{M} a_{j\gamma}$ denoting the degree of object β and user j respectively [7]. The resulting recommendation list of uncollected objects is then sorted according to $\widetilde{f}^i_\alpha$ in descending order. Physically, the diffusion is equivalent to a three-step random walk starting with k_i units of resources on the target user i. The *recommendation score* of an item is taken to be its amount of gathered resources after the diffusion.

The HC algorithm works similar to the MD algorithm, but instead follows a conductive process represented by

$$W_{\alpha\beta} = \frac{1}{k_\alpha}\sum_{j=1}^{N}\frac{a_{j\alpha}a_{j\beta}}{k_j}. \quad (3)$$

Physically, the recommendation scores can be interpreted as the temperature of an item, which is the average temperature of its nearest neighborhood, *i.e.* its connected users. The higher the temperature of an item, the higher its recommendation score [8].

Unlike the above physical processes, the CF algorithms provide recommendations based on user or item similarities. It is divided into two main categories: the user-based CF and the item-based CF [6]. In UCF, the recommendation score of an item is evaluated by the similarity between the target user and the users who collected the item. Actually, the measure of similarities of two nodes in a network is subject to definition. Here we define the similarity as the number of common neighbors [18] in the bipartite networks. The final recommendation score for each item can be written as

$$\widetilde{f}^i_\alpha = \sum_{j=1}^{N} s_{ij}a_{j\alpha}, \quad (4)$$

where s_{ij} is the similarity between user i and j. While in ICF, the recommendation score of an item is evaluated based on its similarity with the collected items of the target user. Similarly, the final recommendation score for each item can be written as

$$\widetilde{f}^i_\alpha = \sum_{\beta=1}^{M} s_{\alpha\beta}a_{i\beta}, \quad (5)$$

where $s_{\alpha\beta}$ is the similarity between item α and β.

To study the effect of the above mentioned algorithms on the dispersion of item popularity, we consider a scenario of recommender systems as follows. At each step, a random user is selected as the active user, and the recommendation scores of all items are then evaluated for him/her. For simplicity, we assume that the active user would accept the recommendations by selecting the uncollected item with the highest score, *i.e.* we add a link between the active user and this item in the bipartite network. An illustrative example can be seen in fig. 2. We have also considered a similar scenario where users choose the top-20 recommended items with probability proportional to their scores and consistent results are obtained.

Table 2: Description of the data.

Network	Users	Items	Links	Sparsity
Movielens	943	1682	82520	$5.20 \cdot 10^{-2}$
Netflix	10000	6000	701947	$1.17 \cdot 10^{-2}$
Delicious	10000	232657	1233997	$5.30 \cdot 10^{-4}$
Amazon	20000	66525	258911	$1.95 \cdot 10^{-4}$

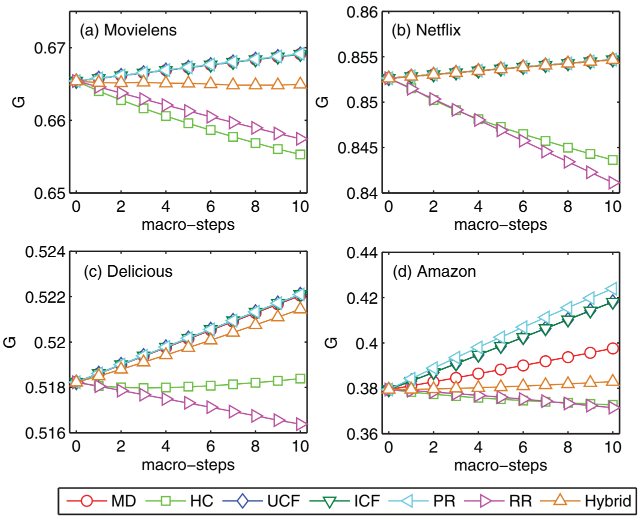

Fig. 3: (Color online) The evolution of the Gini coefficient for item popularity when different recommendation methods are implemented in real systems. Discussions on hybrid method will be given in eq. (10) and the corresponding descriptions.

In one *macro-step* of our simulation, we randomly choose 10% of users to be active. After each macro-step, we evaluate the dispersion of the item popularity by Gini coefficient. Note that we do not consider the growth of the system since introducing new users or items may involve the cold start problem for them [19]. At the beginning of the simulation, we use real data as the initial bipartite networks. The datasets we examined are the subsets of data obtained from four online systems: Movielens, Netflix, delicious and Amazon[2] (see table 2).

We show in fig. 3 the evolution of Gini coefficient in simulations as a function of macro-step. As one can see, the Gini coefficient increases with MD, UCF and ICF algorithms. This is because of their reinforcing influence on the system, which leads to a wider dispersion of item popularity after successive recommendations. A further

[2]Movielens: sampled from 19 Sep 1997 to 22 Apr 1998. Netflix: sampled from Netflix prize data. Delicious: downloaded from Delicious.com. Amazon: sampled from 28 Jul 2005 to 27 Sep 2005. We consider a link exists when rating is greater than 2 in Movielens and Netflix.

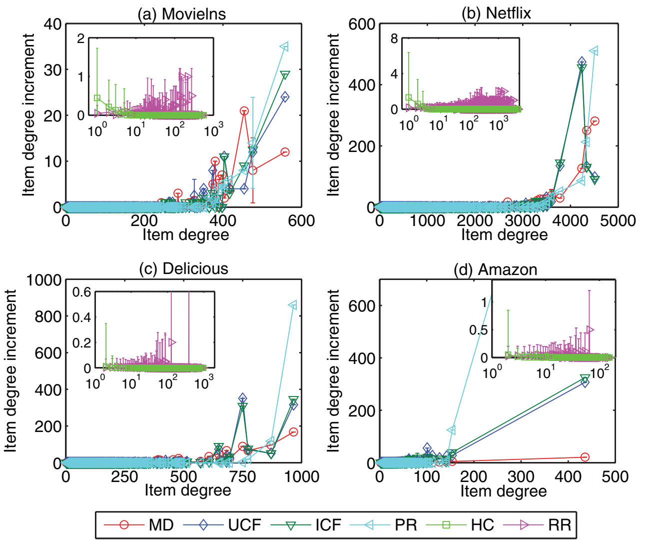

Fig. 4: (Color online) The increment in item popularity with MD, UCF, ICF and PR implemented in real systems. Insets: The increment in item popularity with HC and RR implemented (horizontal axis in log-scale).

evidence can be seen in fig. 4, which shows that popular items become more popular, while the rest of the items remain undiscovered. This corresponds to an undesired influence, as choices for users become more limited with these recommendation algorithms.

We can further understand the reinforcing influences of the MD, UCF and ICF recommender systems by comparing their Gini coefficients with that of the unpersonalized PR algorithm. The results in figs. 3 and 4 imply that these four algorithms only recommend popular items, leading to similar changes in the distribution of item popularity. Some differences are observed between PR and the rests in fig. 4(c) and (d). It is because MD, UCF and ICF only recommend the relevant items according to former choices of users. For each user, the number of relevant items is small, especially in sparse datasets like Delicious and Amazon. This personalization in MD, UCF and ICF inevitably guide the popularity to some not so hot but relevant items, which is different from PR, though the resultant G is not significantly smaller than that of PR.

On the other hand, the HC algorithm behaves quite differently from the other algorithms. As we can see in fig. 3(a), (b) and (d), it decreases G. This implies that HC does not reinforce the popularity of hot items as MD, UCF and ICF. While in fig. 3(c), G increases slightly. This may due to the high sparsity of delicious dataset, such that the three-step conduction process in HC can only reach some items with large degree, and inevitably add links to them many times. On the other hand, HC is different from the uniform addition of links in RR, as it inclines to add links to items with small degree as seen from the insert of fig. 4.

We emphasize that similar results are observed with entropy or second moment of item degree distribution to be the measures of diversity, which indicates the robustness of our findings here.

The mean-field approximation. – To better understand the influences of recommendation systems, we derive analytically the distribution of item scores after the recommendation processes. The major difficulty in the analysis comes from the particular network topology of each dataset, which embeds the non-trivial correlations between users and items [20,21]. Here we focus on the recommendation influences, and assume a simple topology where users and items are randomly connected [22]. This corresponds to a crude mean-field approximation, but such assumption facilitates the analysis and the illustration of physical behaviors underlying the recommendation algorithms.

To begin our analysis, we derive the probability $p_{i\alpha}$ for a user i to connect to an item α in a random graph, with predefined user and item degrees identical to that of the real network. Suppose we start with k_i cavities on user i and k_α cavities on item α, which are, respectively, the degree of i and α. We then randomly connect the cavities of users to the cavities of items to establish links. If one cavity is picked randomly among the items, the probability that α being picked is $\frac{k_\alpha}{\sum_{\beta=1}^{M} k_\beta}$. It implies that $p_{i\alpha} = 1-(1-\frac{k_\alpha}{\sum_{\beta=1}^{M} k_\beta})^{k_i}$, where $(1-\frac{k_\alpha}{\sum_{\beta=1}^{M} k_\beta})^{k_i}$ is the probability that i is not connected to α. As $\sum_{\beta=1}^{M} k_\beta \gg k_\alpha$, expansion to the first order of k_α leads to $p_{i\alpha} \approx 1-(1-k_i\frac{k_\alpha}{\sum_{\beta=1}^{M} k_\beta}) = \frac{k_i k_\alpha}{c}$, where $c = \sum_{\beta=1}^{M} k_\beta$ is the total number of links in the bipartite network.

We then derive the mean-field expression of recommendation scores in the MD recommender system. As mentioned above, MD is based on the three-step diffusion. The resource vector for items in the first step and the last step are denoted, respectively, by $\mathbf{f}$ and $\widetilde{\mathbf{f}}$. In the second step, the resources are in the users' side and the corresponding vector is denoted as $\mathbf{e}$. By considering the last step of the diffusion process, the score of α from user i is given by $\widetilde{f}^i_\alpha = (1-p_{i\alpha})\sum_{j=1}^{N} \frac{e^i_j p_{j\alpha}}{k_j}$. Substitution of $p_{j\alpha} = \frac{k_j k_\alpha}{c}$ leads to

$$\widetilde{f}^i_\alpha = \left(1-\frac{k_i k_\alpha}{c}\right)\sum_{j=1}^{N}\left(\frac{e^i_j}{k_j}\frac{k_j k_\alpha}{c}\right) = \left(1-\frac{k_i k_\alpha}{c}\right)\frac{k_i k_\alpha}{c}, \qquad (6)$$

as $\sum e^i_j = \sum f^i_\alpha = k_i$.

Next we derive the scores for HC by again considering the last step of the conduction process. However, the total "resources" does not conserve in heat conduction but instead the temperature of user j is given by $e^i_j = \frac{k_i}{M}\sum_{\gamma=1}^{M}\frac{p_{j\gamma}}{k_j} = \frac{k_i}{M}$, where $\frac{k_i}{M}$ corresponds to the random choices of initial collected item for i. Therefore,

$$\widetilde{f}^i_\alpha = \left(1-\frac{k_i k_\alpha}{c}\right)\sum_{j=1}^{N}\left(\frac{k_i}{M k_\alpha}\frac{k_j k_\alpha}{c}\right) = \left(1-\frac{k_i k_\alpha}{c}\right)\frac{k_i}{M}, \qquad (7)$$

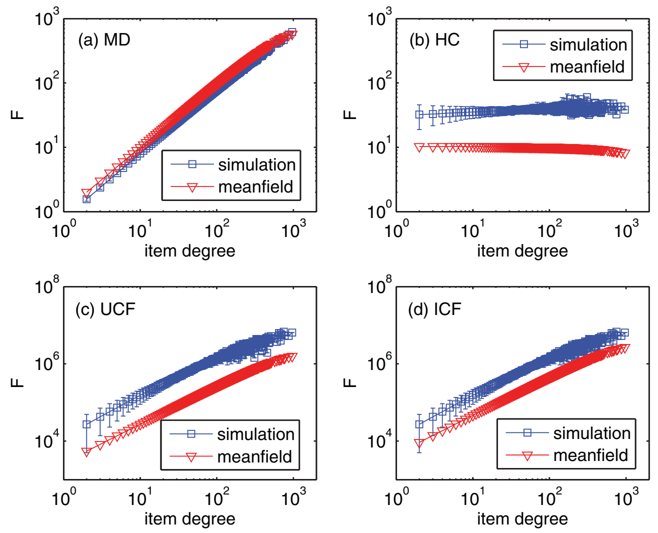

Fig. 5: (Color online) The simulation result and the theoretical result of the total recommendation score *vs.* the original item degree in different recommendation engines. The results are based on Delicious dataset.

In UCF, scores of an item are evaluated by the similarity between the target user and the users who have collected it. The user similarity is given by the number of common neighbors. Therefore, $\widetilde{f_\alpha^i} = (1-p_{i\alpha})\sum_j^N s_{ij}p_{j\alpha}$, where $s_{ij} \approx \frac{k_i k_j}{M}$ in the mean-field approximation. The score for object α is then approximated by

$$\widetilde{f_\alpha^i} = \left(1-\frac{k_i k_\alpha}{c}\right)\sum_{j=1}^{N}\left(\frac{k_i k_j}{M}\frac{k_j k_\alpha}{c}\right) = \left(1-\frac{k_i k_\alpha}{c}\right)\frac{k_i k_\alpha b}{cM}, \tag{8}$$

where $b = \sum_{j=1}^N k_j^2$ is a constant for a given network.

As similar to UCF, the item similarity in ICF can be approximated by $s_{\alpha\beta} \approx \frac{k_\alpha k_\beta}{N}$ in the mean-field approximation. The score for object α is then approximated by

$$\widetilde{f_\alpha^i} = \left(1-\frac{k_i k_\alpha}{c}\right)\sum_{\beta=1}^{M}\left(\frac{k_i k_\beta}{c}\frac{k_\alpha k_\beta}{N}\right) = \left(1-\frac{k_i k_\alpha}{c}\right)\frac{k_i k_\alpha d}{cN}, \tag{9}$$

where $d = \sum_{\beta=1}^M k_\beta^2$ is a constant for a given network.

In order to compare the simulated results and the mean-field predictions, we evaluate the corresponding total scores $F_\alpha = \sum_i^N \widetilde{f_\alpha^i}$ that an item receives from all the users. As shown in fig. 5, the mean-field approximation effective captures the trend of the recommendation scores. The difference of the magnitude between the simulation and mean-field result comes from the non-trivial correlation between users and items in real networks.

Further insights are drawn by noting $c \gg k_i k_\alpha$ in most systems, which implies $\widetilde{f_\alpha^i} \propto k_i k_\alpha$ in eqs. (6), (8) and (9). Since we assume that users always accept the item with highest recommendation scores, the recommendation scores in MD, UCF and ICF are thus similar to that of PR which recommends the most popular items. This again

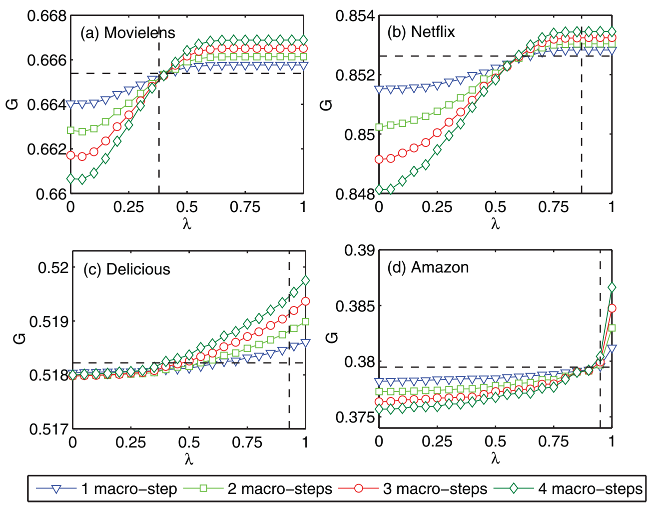

Fig. 6: (Color online) The Gini coefficient as a function of λ in the hybrid recommendation algorithm implemented on real systems. The vertical and horizontal dashed lines are, respectively, the values of λ^* with optimal recommendation accuracy and the Gini coefficients before recommendation algorithms are implemented.

shows the reinforcing influence of these recommendation algorithms. On the other hand, eq. (7) suggests $\widetilde{f_\alpha^i} \propto k_i$ in HC, which is independent of item as in the case of random recommendations.

Though the approximated scores of HC agree well with RR, their behaviors are different in terms of choices of items. According to $\widetilde{f_\alpha^i} = (1-p_{i\alpha})\sum_{j=1}^N \frac{k_i p_{j\alpha}}{k_j k_\alpha}$, users select the reachable items with lowest degree after three-step conduction, compared to the random choice. Therefore, HC and RR show different influence on the dispersion of item popularity, as we can see in fig. 3.

Steady Gini coefficient by hybrid recommendations. – As we have seen from the previous sections, MD reinforces the popularity of hot items and limits available choices, while HC recommends items with low popularity and increases global diversity. It is thus interesting to examine the influence on diversity if these two algorithms with opposite influences are combined. We thus adopt the hybrid algorithm of MD and HC proposed in [9], with the new recommendation score $\widetilde{h}_\alpha$ given by

$$\widetilde{h}_\alpha = \lambda \widetilde{f}_\alpha^{\mathrm{MD}} + (1-\lambda)\,\widetilde{f}_\alpha^{\mathrm{HC}}. \tag{10}$$

The parameter λ adjusts the relative weight between the two algorithms. When λ increases from 0 to 1, the hybrid algorithm changes gradually from HC to MD. We remark that though eq. (10) corresponds to a linear combination of scores, the hybrid algorithm is a non-linear combination of HC and MD as users select only items with highest scores.

The influence of the hybrid algorithm on the Gini coefficient is shown in fig. 6 as a function of λ. The lines with different symbols correspond to G measured after increasing macro-step. As we can see from fig. 6, G increases with λ, corresponding to a transition from

HC to MD recommender systems. It is interesting to note that G shows a significant increase in a range of λ on the Netflix and Movielens datasets, and becomes saturated afterwards. The saturated G corresponds to dominance of the MD algorithm such that only popular items are recommended, despite the presence of HC. Similar behaviors are observed in fig. 6(c) and (d) in the Delicious and Amazon datasets, in which the saturated G corresponds to the dominance of the HC algorithm such that niche items are more likely to be recommended.

Another interesting behavior is noted in fig. 6 when we compare G after different macro-steps of recommendations. As we can see, the lines with different symbols intersect at a particular value of λ, suggesting a steady G after the reinforcement of recommendations. The corresponding value of λ thus corresponds to the balance between the HC and MD algorithms, leading to steady dispersion in item popularity. This is desirable when one considers the reinforcing influence on global diversity as an undesired side-effect of recommender systems. These values of λ and the corresponding G are compared respectively to λ^* with highest recommendation accuracy [6] and to the Gini coefficient before recommendation algorithms are implemented. Specifically, λ^* is obtained by minimizing the ranking score index as in ref. [9]. We can see that the hybrid method with λ^* increases G in fig. 3. These results show that high recommendation accuracy does not always guarantee global diversity, leading to a paradox in recommendations.

Conclusion. – Recommendation is an effective way to solve the problem of excess information. However, it is unclear how it allocates popularity among items. In this paper, we simulate successive recommendations and measure their influence on the dispersion of item popularity by the Gini coefficient. The results indicate that local diffusion and collaborative filtering reinforce the popularity of hot items, widening the popularity dispersion. On the other hand, the heat conduction algorithm increases the popularity of the niche items and generates a smaller dispersion of item popularity. Simulations are compared to mean-field approximation. Our results indicate that recommender systems have a reinforcing influence on the diversity of choices of commodities.

The present work raises a number of questions where investigations could further deepen our understanding of recommender systems. For instance, to incorporate in mean-field approximation the correlations between users and items is a meaningful yet non-trivial extension, as the present analysis is based only on uncorrelated networks. Moreover, though the present scenario captures a clear picture of recommendation influence, a more realistic case is to consider recommendations on evolving networks. One possibility is through rewiring and growing models [12,23], where further exploration is required.

Taken together, our work not only provides a deeper understanding of some recommendation methods, but also highlights the importance of the global diversity and may shed light on developing a new recommendation method which can directly control the global diversity.

***

This work was partially supported by the Swiss National Science Foundation (Grant No. 200020-132253), the National Natural Science Foundation of China (Grant No. 60973069) the China Postdoctoral Science Foundation (Grant No. 201003694) and the Sichuan Provincial Science and Technology Department (Grant No. 2010HH0002).

REFERENCES

[1] Broder A., Kumar R., Moghoul F., Raghavan P., Rajagopalan S., Stata R., Tomkins A. and Wiener J., *Comput. Networks*, **33** (2000) 309.

[2] Maslov S. and Zhang Y.-C., *Phys. Rev. Lett.*, **87** (2001) 248701.

[3] Adomavicius G. and Tuzhilin A., *IEEE Trans. Knowl. Data Eng.*, **17** (2005) 734.

[4] Cacheda F., Carneiro V., Fernández D. and Formoso V., *ACM Trans. Web*, **5** (2011) 1.

[5] Konstan J. A., Miller B. N., Maltz D., Herlocker J. L., Gordon L. R. and Riedl J., *Commun. ACM*, **40** (1997) 77.

[6] Herlocker J. L., Konstan J. A., Terveen K. and Riedl J. T., *ACM Trans. Inf. Syst. Secur.*, **22** (2004) 5.

[7] Zhou T., Ren J., Medo M. and Zhang Y.-C., *Phys. Rev. E*, **76** (2007) 046115.

[8] Zhang Y.-C., Blattner M. and Yu Y.-K., *Phys. Rev. Lett.*, **99** (2007) 154301.

[9] Zhou T., Kuscsik Z., Liu J.-G., Medo M., Wakeling J. R. and Zhang Y.-C., *Proc. Natl. Acad. Sci. U.S.A.*, **107** (2010) 4511.

[10] Shang M.-S., Lü L.-Y., Zeng W., Zhou T. and Zhang Y.-C., *EPL*, **88** (2009) 48006.

[11] Lambiotte R. and Ausloos M., *Phys. Rev. E*, **72** (2005) 066107.

[12] Beguerisse-Diaz M., Porter M. A. and Onnela J.-P., *Chaos*, **20** (2010) 043101.

[13] Yeung C. H., Cimini G. and Jin C.-H., *Phys. Rev. E*, **83** (2011) 016105.

[14] Gini C., *Variabilità e mutabilità* (in English: *Variability and Mutability*) (C. Cuppini, Bologna) 1912.

[15] Dragulescu A. and Yakovenko V. M., *Eur. Phys. J. B*, **20** (2001) 585.

[16] https://publish.aps.org/datasets.

[17] http://www.ssa.gov/OACT/babynames/.

[18] Lorrain F. and White H. C., *J. Math. Sociol.*, **1** (1971) 49.

[19] Zhang Z.-K., Liu C., Zhang Y.-C. and Zhou T., *EPL*, **92** (2010) 28002.

[20] Newman M. E. J., *Phys. Rev. Lett.*, **89** (2002) 208701.

[21] Peltomaki M. and Alava M., *J. Stat. Mech.* (2006) P01010.

[22] Newman M. E. J., Strogatz S. H. and Watts D. J., *Phys. Rev. E*, **64** (2001) 026118.

[23] Evans T. S. and Plato A. D. K., *Phys. Rev. E*, **75** (2007) 056101.