Iterative group-based and difference ranking method for online rating systems with spamming attacks

Fu, Quan-Yun; Ren, Jian-Feng; Sun, Hong-Liang

University of Nottingham Ningbo China, 199 Taikang East Road, Ningbo, 315100, Zhejiang, China.

First published 2021

Iterative group-based and difference ranking method for online rating systems with spamming attacks

Quan-Yun Fu[a], Jian-Feng Ren[b], Hong-Liang Sun[a]

[a]*Nanjing University of Finance and Economics,Nanjing 210023, People's Republic of China*
[b]*University Nottingham Ningbo China. Ningbo, 315100, Zhejiang,People's Republic of China*

Abstract

It is significant to assign reputation scores to users and identify spammers in the bipartite rating networks. How to evaluate the reputation of users accurately in the presence of spamming attacks is essential in practice. In this paper, we propose an Iterative Group-based and Difference Ranking (IGDR) method, which is based on the original Iterative Group-based Ranking (IGR) method. The IGR method considers users' grouping behaviours, but it ignores the characteristics of the individual ratings. Therefore, in the IGDR method, we introduce the users' rating deviation score. The user with a smaller rating deviation will be given a higher reputation score. The IGDR method is more accurate than the IGR method tested on three real data sets. It also can be applied to deal with big data in a short time.

Keywords: Complex Networks,Rating Networks,Spamming Attacks

1. Introduction

With the rapid development of the modern society , customers face many choices in e-commerce online systems[1]. The ratings of objects play an important role in users' decision [2] makings. Therefore, many online websites (such as MovieLens, Netflix, etc.) have launched rating systems that allow users to rate objects [3, 4]. However, there are some spammers who give unreasonable ratings to distort the rating system [2, 5, 6]. Generally speaking, there are two types of spammers, one of which is random spammers and another one is malicious spammers. The random spammers give ratings randomly to affect system [7, 8], while the malicious spammers give extreme ratings [9, 10]. Therefore, how to identify spammers quickly and accurately is essential for the online rating system in e-commerce[11, 12, 13].

To solve this problem, a variety of methods have been proposed [14, 15]. Using these methods, each user's reputation score is calculated according to their ratings of the objects in different ways [16, 17]. Then the users are ranked by their reputation scores. Users with low reputation scores are defined as spammers. The existing methods can be divided into the following categories:

Quality-based methods: This kind of methods include Iterative Ranking (IR) [18], Correlation-based Ranking (CR) [19] and Reputation Redistribution Ranking (RR) [20]. In these methods, each object gets a predicted quality based on the score they received. The reputation of users is determined by the similarity between their ratings and the predicted quality of the objects [21]. The IR method is proposed by Laureti *et al.*, where a user's reputation is calculated by the mean of the inverse of difference between user's rating and object's quality. Then Zhou *et al.* propose CR method which uses Pearson correlation instead. These methods are well-performed in user reputation evaluation. However, in these methods, the predicted object quality calculated by the average ratings may deviate from the actual product quality seriously due to the existence of noisy ratings. Therefore, how to accurately calculate the quality of objects is an urgent problem for these methods [22, 23].

*Group-based methods:*This kind of methods include Group-based Ranking(GR) method [24], and Iterative Group-based Ranking(IGR) method [25]. In these methods, users are grouped by their rating behaviours. The users are divided into the same group if they give an object the same rating. Users are assigned a high reputation if they always fall into large rating groups. The GR method is proposed by Gao *et al.*, where the users' reputation is calculated once, while the users' reputation is calculated iteratively until being stable in the IGR method. Unlike quality-based methods, these methods don't calculate object quality, but consider whether users' rating behaviour is the same as most users. However, these methods are not always robust when there are a large number of spammers[26, 27].

Distribution-based methods: This kind of methods include Deviation-based Ranking (DR) method [28], and Bayesian Ranking (BR) method [29]. These methods assume that the ratings of normal users follow a certain distribution, such as normal distribution, beta distribution and etc. However, the rating characteristics of users cannot be accurately summarized with one distribution. For example, if the user has given only one or two ratings, we can't sure what distribution these ratings follow [30].

In this paper, we propose an Iterative Group-based and Difference Ranking (IGDR) method by introducing a rating devi-

Email address: hlsun84@mail.ustc.edu.cn (Hong-Liang Sun) ation into the original IGR method. Considering the grouping of users' ratings, the IGR method is superior to other methods when there are few malicious users, and can better detect the outlying spammers. However, the IGR method only considers users' grouping behaviors and ignores the characteristics of the individual ratings. Therefore, in the IGDR method, we consider that normal users usually have their rating preferences. Some prefer to give higher ratings, and some prefer to give medium scores, while others prefer to give lower ratings. Thus their ratings may be more focused than spammers. For example, the users prefer to give higher ratings might often rate objects higher than their real quality, so they rarely rate the lowest. Therefore, their rating deviation might be smaller. Considering that, the users with a small rating deviation should be assigned a high reputation score. The IGDR method has better performance and is more robust compared with the original GR method and IGR method on three real data sets in our study.

2. Method

We firstly introduce some basic notations for the user reputation evaluation methods. The online rating system is naturally described by a weighed bipartite network $G = \{U, O, E\}$, where $U = \{U_1, U_2, ..., U_m\}$, $O = \{O_1, O_2, ..., O_n\}$ and $E = \{E_1, E_2, ...E_l\}$ are sets of users, objects and ratings, respectively [31]. Here, we use Greek and Latin letters, respectively, for object-related and user-related indices to distinguish them. The degree of a user i and an object α are denoted as k_i and k_α. Considering a discrete rating system, the bipartite network can be represented by a triad B, where $B_{i\alpha} = \{U_i, O_\alpha, \omega_{i\alpha}\}$. In order to describe the calculation process clearly, here we use a rating matrix A, where the element $A_{i\alpha} \in \Omega = \{\omega_{1\alpha}, \omega_{2\alpha}, ..., \omega_{z\alpha}\}$ is the weight of the link connecting node U_i and node O_α, i.e., the rating given by user i to object α. In a reputation system, each user i will be assigned a reputation score, denoted as R_i. The users with very low reputation scores are detected as spammers. In the following parts, we will introduce the proposed user reputation evaluation method. The IGDR method works as follows.

Firstly, a matrix Λ is created to calculate the group size. Based on the idea that users with high reputation scores are more trustworthy and should have greater impacts in the group, the matrix Λ considers users' reputation. The matrix Λ contains n rows and r columns. Here, n indicates the number of objects and r denotes the number of ratings. Mathematically, the group size $\Lambda_{s\alpha}$ is defined as

$$\Lambda_{s\alpha} = \sum_{i \in O_\alpha} R_{is}. \tag{1}$$

Where $i \in O_\alpha$ represents user i who rates object α. $\Lambda_{s\alpha}$ calculates the sum of users' reputation who rate the object α as s.

Secondly, a group size proportion matrix Λ^* is established by normalizing matrix Λ by column. The group size proportion matrix Λ^* is defined as

$$\Lambda^*_{s\alpha} = \frac{\Lambda_{s\alpha}}{\sum_\alpha \Lambda_{s\alpha}}. \tag{2}$$

Thirdly, referring to the group size proportion matrix Λ^*, the original rating matrix A is mapped to a proportion matrix A'. More specifically,the proportion of user i's rating group for object α $A_{i\alpha}$ is defined as $A'_{i\alpha} = \Lambda^*_{s\alpha}$, where $A_{i\alpha} = \omega_s$. A' is also a triad like B. However, here we use a matrix A' to explain more clearly. $A'_{i\alpha}$ is defined as

$$A'_{i\alpha} = \begin{cases} \Lambda^*_{s\alpha} & if A_{i\alpha} = \omega_s \\ - & otherwise \end{cases} \tag{3}$$

where the symbol "-" indicates a non-value. The IGR method define user i's reputation as

$$R_i = \frac{\mu(A'_i)}{\sigma(A'_i)}, \tag{4}$$

where μ is the function of mean value and σ is the standard deviation, respectively. On the one hand, if the average value of the group size proportion of a user's rating is small, it means that the user's rating often falls into the small group, that is, it deviates from the rating of most users. And the user is untrustworthy. On the other hand, if the varies of the group size proportion of a user's rating is large, it means that his rating behavior is unstable. The user is also untrustworthy.

The IGR method merely considers users' grouping behaviours and ignores the characteristics of the individual ratings. Therefore, in the IGDR method, we consider that normal users usually have their rating preferences. Some prefer to give higher ratings, some prefer to give medium scores, while others prefer to give lower ratings. So their ratings may be more stable than spammers. Based on these considerations, we define user i's reputation score as

$$R_i = \sqrt{\mu(A'_i)} + \frac{1}{5\sqrt{\sigma(A'_i)} + \sigma(\omega_i)}, \tag{5}$$

where $\sigma(\omega_i)$ indicates the standard deviation of user i's ratings. Considering that different parameters may have different specific gravity effects, we change their proportion. It is found that Eq. 5 performs better when there are fewer spammers. Specifically, the mean value of A'_i is defined as

$$\mu(A'_i) = \frac{\sum_{\alpha \in O_i} A'_{i\alpha}}{k_i}, \tag{6}$$

$\alpha \in O_i$ denotes the object α rated by user i. The standard deviation A'_i is defined as

$$\sigma(A'_i) = \sqrt{\frac{\sum_{\alpha \in O_i} (A'_{i\alpha} - \mu(A'_i))^2}{k_i - 1}}, \tag{7}$$

and the standard deviation of user i's rating is defined as

$$\sigma(\omega_i) = \sqrt{\frac{\sum_{\alpha \in O_i} (\omega_{i\alpha} - \mu(\omega_i))^2}{k_i - 1}}, \tag{8}$$

where the $\mu(\omega_i)$ is defined as

$$\mu(\omega_i) = \frac{\sum_{\alpha \in O_i} \omega_{i\alpha}}{k_i}. \tag{9}$$

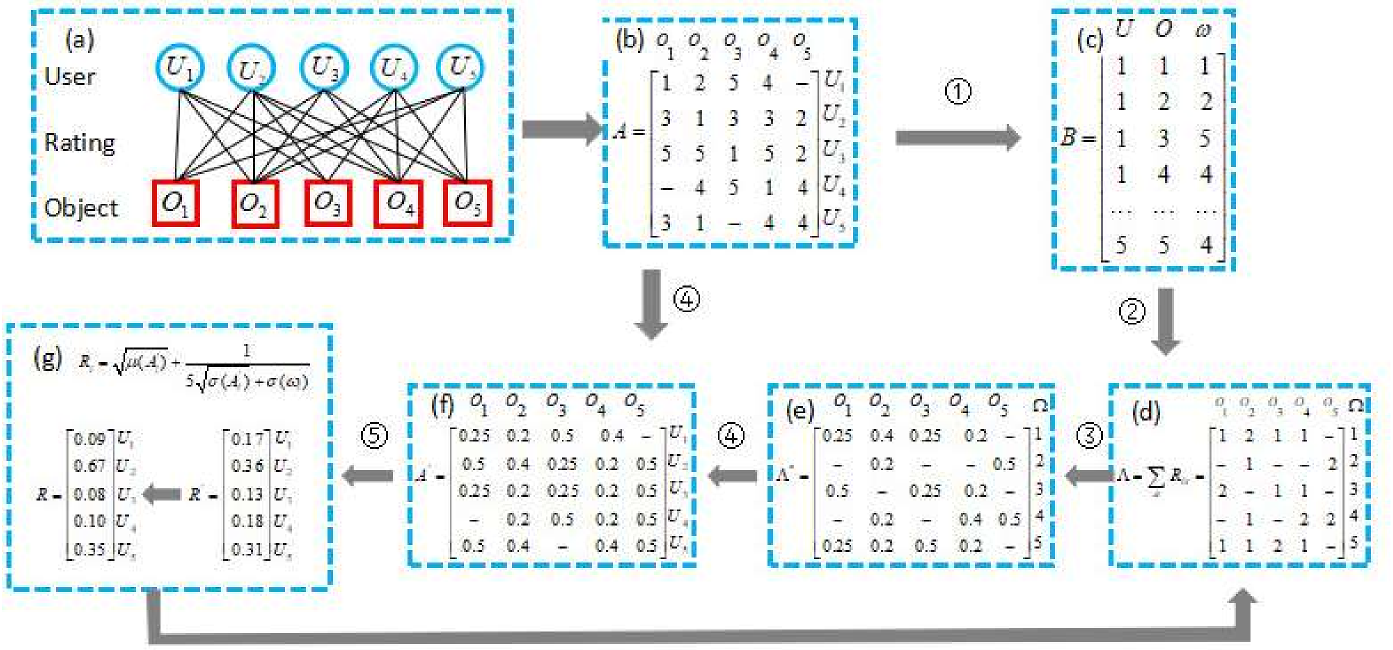

Figure 1: Illustration of the IGDR method.The number besides the arrow marks the order of the procedure. The symbol ”-” in matrix indicates a non-value, which should be ignored in the calculation. (a) The original weighed bipartite network G. (b) The corresponding rating matrix A. (c) The rating triad B. (d) The reputation-weighted group size matrix Λ. Taking O_2, Ω=4 as an example, $\Lambda_{4,2} = R_4 = 1$. (e) The rating-rewarding matrix Λ^* is constructed by normalizing Λ by column. e.g, $\Lambda^*_{4,2} = 1/(2 + 1 + 1 + 1) = 0.2$. (f) The rewarding matrix A' is obtained by mapping matrix A referring to Λ^* .e.g, $A'_{4,2}$=0.2. (g) The reputation of users R. R' is temporal reputation in the previous iteration step. In IGDR method,Λ and R' are iteratively updated according to (d),(e),(f) and (g), as indicated by the grey arrows. Finally, a stable reputation score R is obtained.

In IGDR, the reputation R and the group size Λ are iteratively updated according to Eqs. 2, 3 and 4 until the change of the reputation $\left|R - R'\right| = \sum_i (R_i - R'_i)^2/m$ is smaller than the given threshold value $\Delta = 10^{-4}$. Here, R' denotes the reputation vector at the previous iteration step. A visual representation of the IGDR method is shown in Fig. 1.

Table 1: Basic statistics of the three real data sets. m is the number of users, n is the number of objects, and l is the number of ratings. U is the average degree of users, O is the average degree of objects, and $S = l/mn$ is the sparsity of the bipartite network.

Data set	m	n	l	U	O	S
MovieLens_10	943	1682	100000	106	59	0.06305
MovieLens_100	7120	130642	1048575	147	8	0.00113
Netflix	5000	17768	3496614	699	196	0.03936

3. Data and metrics

3.1. Rating data

We consider three commonly studied real data sets. Two of them are *MovieLens* [1] and another is *Netflix* [2]. The two MovieLens data are recorded as *MovieLens_10* and *MovieLens_100* separately. *MovieLens_10* and *Netflix* use a 5-point rating scale with 1 being the worst and 5 being the best. Here, we sampled and extracted a small data set from the original Netflix data set, by choosing users who have at least 50 ratings and objects. The basic statistics of data sets are summarized in table 1.

3.2. Generating artificial spammers

Two types of distorted ratings are widely found in real rating systems, namely, malicious ratings and random ratings. The malicious ratings are from spammers who always give minimum (maximum) ratings to push down (up) certain target objects. The random ratings are from spammers who give ratings randomly to disturb the rating system.

In the data sets, we randomly select d users as spammers and change their ratings. The ratings are changed to 1 or 5 for malicious spammers and changed to 1,2,..., max randomly for random spammers.

3.3. Metrics for evaluation

We adopt two commonly used metrics to evaluate the performance of ranking, namely recall [32] and AUC [33]. We should rank the users from low to high scores according to their reputation firstly. Then choose the first L users form a top-L ranking list. The recall is defined as

$$R_c(L) = \frac{d'(L)}{d} \tag{10}$$

[1] https://grouplens.org/datasets/movielens
[2] https://grouplens.org/datasets/movielens

where $d'(L)$ is the number of spammers in top-L ranking list.The larger value of R_c the higher accuracy of the method.

Next, we introduce the AUC. The value of AUC can be seen as the probability a randomly chosen spammer's reputation is lower than a normal user's reputation. To calculate AUC, each time a pair of spammer and normal user are picked up and their reputation scores are compared. The AUC value is defined as

$$AUC = \frac{N' + 0.5N''}{N} \tag{11}$$

Here N represents the number of compared times. And N' means the number of times that spammer's reputation is lower than normal user's reputation. N'' means the number of times that their reputation scores are the same. The normal user's reputation score should be higher than spammer's, so the larger the value of AUC is, the better the ranking method performs.

4. Results

4.1. Effectiveness and Robustness

To test the effectiveness of the ranking methods using the three real data sets, we generate artificial data set with 50 spammers and 100 spammers. Each data set is only with one type of spammers: malicious or random. On the generated data sets, we calculate the recall of different methods as a function of the spammer list's length L.

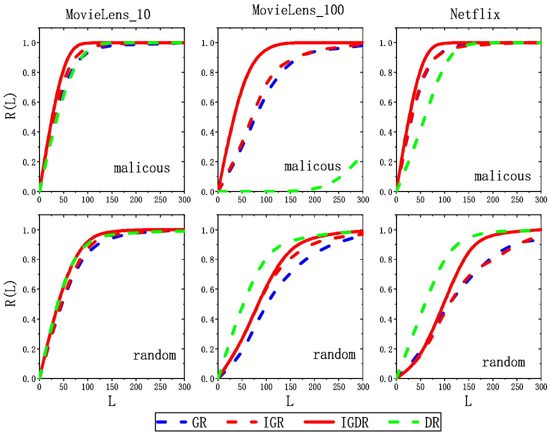

Figure 2: (Color online) The recall R_c of different methods as a function of the length of the list L. Panels with malicious indicates malicious spammers; panels with random denotes random spammers with d = 50 being fixed. The results are averaged over 50 independent runnings.

As shown in Fig. 2 and Fig. 3. The IGDR method outperforms IGR method and GR method on detecting both types of spammers. The DR method performs better than the IGDR method for random spammers. But the R_c of the DR method is lower than that of the IGDR method for malicious spammers, especially on *MovieLens*_100. The R_c of ranking malicious spammers in IGDR is higher than others. The results with 50 spammers and 100 spammers are similar in three real data sets.

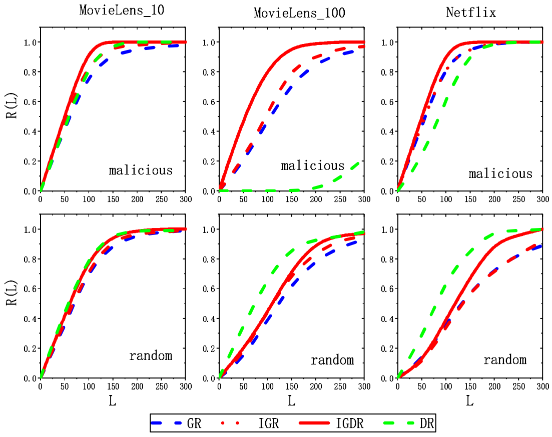

Figure 3: (Color online) The recall R_c of different methods as a function of the length of the list L. Panels with malicious stand for malicious spammers; panels with random stand for random spammers with d=100 being fixed. The results are averaged over 10 independent realizations.

The AUC of IGDR is more stable than the IGR method and GR method, especially in random spammers. For random spammers, the DR method and the IGDR perform similarly. For malicious spammers, the IGDR performs a little better than the DR when there are fewer spammers. But the IGDR is not as stable as the DR method.

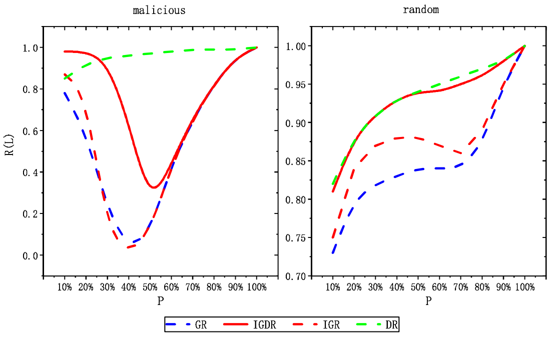

Figure 4: (Color online)The recall R_c change of different type spammers as a function of the spammer proportion. Panels with malicious stand for malicious spammers; panels with random stand for random spammers with d=L being fixed. The results are averaged over 10 independent realizations.

In Fig. 4, we show the recall R_c changes with spammer quantities in *MovieLens*_10 data set. And in Fig. 5, we show the AUC changes with the spammer quantities in *MovieLens*_10 data set. From the figures, it is known that the IGDR performs better than the IGR and GR all the time. And the GR method and the IGR method perform similarly no matter for the malicious spammers or the random spammers.

4.2. Identify Excellent Objects

How to decide whether a movie is a good one? Everyone has a different point of view. Here we think that movies win Academy Awards or with high ranking in IMDB (Internet Movie Database) should be highly recognized. Firstly, we use

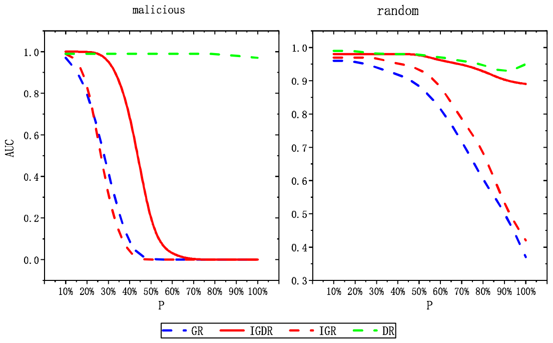

Figure 5: (Color online)The AUC change of different type spammers as a function of the spammer proportion. Panels with malicious stand for malicious spammers; panels with random stand for random spammers with d=L being fixed. The results are averaged over 10 independent realizations.

the IGR and the IGDR to rank movies in the $MovieLens\_10$ and $MovieLens\_100$ dataset and find out the top-L movies. Then we count how many of these movies have won Academy Awards [3] or is in the top-250 in IMDB [4]. Since the dataset of $MovieLens\_100$ is sparse, we only consider movies that receive ratings more than 15 times.

Table 2: The number of nominated movies in the top 100 movies of the methods under the *MovieLens* datasets. *Both* in the tile means Academy Awards and IMDB are considered at the same time.

Data set	Academy Awards	IMDB	Both
MovieLens_10-IGDR	24	49	52
MovieLens_10-IGR	21	45	48
MovieLens_100-IGDR	18	47	52
MovieLens_100-IGR	18	47	49

Table 2 shows the results when $L = 100$, and Table 3 shows the results when $L = 150$. In most cases, the performance of the IGDR method is slightly better than that of the IGR method. The IGDR method and GR method identify more movies belong to top-250 IMDB than to Academy Awards.

Table 3: The number of nominated movies in the top 150 movies of the methods under the MovieLens datasets. *Both* in the tile means Academy Awards and IMDB are considered at the same time.

Data set	Academy Awards	IMDB	Both
MovieLens_10-IGDR	32	66	73
MovieLens_10-IGR	31	61	69
MovieLens_100-IGDR	23	63	68
MovieLens_100-IGR	22	68	70

As can be seen from table 4, there are two movies belong to both Academy Awards and top-250 IMDB. And there are five movies win other awards. This shows that the IGDR method is able to discover excellent movies.

Table 4: Top 10 movies ranked by the IGDR method under $MovieLens\_100$ dataset. "Y" means that the movie won an Academy Award "A" or is in the top 250 of IMDB "I". "N" means the movie has not won an Academy Award or is not in the top 250 of imdb. "Others" lists other awards for the movie, while "-" means that other awards are unknown.

File name	A	I	Others
The Decalogue	N	N	-
The Shawshank Redemption	N	Y	Japan Film Academy
Day for Night	N	N	-
The Godfather	Y	Y	-
Cosmos	N	N	Golden Leopard
Band of Brothers	N	N	AFI Awards USA
The Usual Suspects	Y	Y	The Usual Suspects
The Devil and Daniel Johnston	N	N	-
The World of Apu	N	N	National Board of Review
After the Thin Man	N	N	-

5. Conclusions and discussion

In summary, we have proposed an Iterative Group-based and Difference Ranking (IGDR) method in user reputation evaluation by introducing rating deviation into the original IGR method. In real systems, the normal users usually have rating preferences. For example, some users prefer to give low ratings, some prefer to give high ratings, others prefer to give moderate ratings. Their ratings are more concentrated. However, no matter malicious spammers or random spammers, their ratings are more dispersed. Therefore, when calculating users' reputation by the IGDR method containing the standard deviation of users' ratings. Users with small rating deviation are assigned with high reputation scores while users with large rating deviation are assigned with low reputation scores. Results in real data sets suggest that the IGDR method performs better in accuracy and robustness compared with the original IGR and GR. For malicious spammers, the recall rate of the IGDR method is up to 14% higher than that of the IGR method on $MovieLens\_10$ dataset, up to 163% higher on $MovieLens\_100$ dataset, and up to 17% higher on $Netflix$ dataset. For random spammers, the recall rate of the IGDR method is up to 17% higher than that of the IGR method on MovieLens_10 dataset, up to 8% higher on MovieLens_100 dataset, and up to 16% higher on Netflix. Because of the introduction of triples, the IGDR can deal with big data in a short time.

The proposed method not only considers users' grouping behaviours, but also considers users' own rating characteristics. However, the proposed method also can't perform well when the spammer's number is more than the normal user. In future work, we could consider how to design a method which can suit for changeable data sets. Because when the data set changes with time, all the previous methods need to calculate again which is not suitable for real situations.

Acknowledgements

This work is jointly supported by the National Natural Science Foundation of China (with Grant Nos. 71901115.). It

[3] https://www.kesci.com/home/dataset/5e4588185f2816002cee4c26
[4] http://www.imdb.cn/imdb250/1

is also supported by Young Scholar Programme from NUFE (Nos.SHLXW19001).

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

[2] L. Muchnik, S. Aral, S. J. Taylor, Social influence bias: A randomized experiment, Science 341 (6146) (2013) 647–651.

[3] C. Dellarocas, Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior, in: Proceedings of the 2nd ACM conference on Electronic commerce, 2000, pp. 150–157.

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

[5] Z. Yang, Z.-K. Zhang, T. Zhou, Anchoring bias in online voting, EPL (Europhysics Letters) 100 (6) (2013) 68002.

[6] R. Y. Toledo, Y. C. Mota, L. Martínez, Correcting noisy ratings in collaborative recommender systems, Knowledge-Based Systems 76 (2015) 96–108.

[7] P.-A. Chirita, W. Nejdl, C. Zamfir, Preventing shilling attacks in online recommender systems, in: Proceedings of the 7th annual ACM international workshop on Web information and data management, 2005, pp. 67–74.

[8] S. Xie, G. Wang, S. Lin, P. S. Yu, Review spam detection via temporal pattern discovery, in: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, 2012, pp. 823–831.

[9] A. Zeng, G. Cimini, Removing spurious interactions in complex networks, Physical Review E 85 (3) (2012) 036101.

[10] F. Benevenuto, T. Rodrigues, V. Almeida, J. Almeida, M. Gonçalves, Detecting spammers and content promoters in online video social networks, in: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, 2009, pp. 620–627.

[11] H. Liu, Z. Hu, A. Mian, H. Tian, X. Zhu, A new user similarity model to improve the accuracy of collaborative filtering, Knowledge-Based Systems 56 (2014) 156–166.

[12] C.-J. Zhang, A. Zeng, Behavior patterns of online users and the effect on information filtering, Physica A: Statistical Mechanics and its Applications 391 (4) (2012) 1822–1830.

[13] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, H. W. Lauw, Detecting product review spammers using rating behaviors, in: Proceedings of the 19th ACM international conference on Information and knowledge management, 2010, pp. 939–948.

[14] R.-H. Li, J. Xu Yu, X. Huang, H. Cheng, Robust reputation-based ranking on bipartite rating networks, in: Proceedings of the 2012 SIAM international conference on data mining, SIAM, 2012, pp. 612–623.

[15] B. Khosravifar, J. Bentahar, M. Gomrokchi, R. Alam, Crm: An efficient trust and reputation model for agent computing, Knowledge-Based Systems 30 (2012) 1–16.

[16] K. Fujimura, T. Nishihara, Reputation rating system based on past behavior of evaluators, in: Proceedings of the 4th ACM conference on Electronic commerce, 2003, pp. 246–247.

[17] X.-L. Liu, Q. Guo, L. Hou, C. Cheng, J.-G. Liu, Ranking online quality and reputation via the user activity, Physica A: Statistical Mechanics and its Applications 436 (2015) 629–636.

[18] P. Laureti, L. Moret, Y.-C. Zhang, Y.-K. Yu, Information filtering via iterative refinement, EPL (Europhysics Letters) 75 (6) (2006) 1006.

[19] Y.-B. Zhou, T. Lei, T. Zhou, A robust ranking algorithm to spamming, EPL (Europhysics Letters) 94 (4) (2011) 48002.

[20] H. Liao, A. Zeng, R. Xiao, Z.-M. Ren, D.-B. Chen, Y.-C. Zhang, Ranking reputation and quality in online rating systems, PloS one 9 (5).

[21] H. Liao, A. Zeng, Y.-C. Zhang, Towards an objective ranking in online reputation systems: the effect of the rating projection, arXiv preprint arXiv:1411.4972.

[22] Z. Bu, H. J. Li, C. C. Zhang, J. Cao, IEEE Trans. on Know. and Data Eng. 1 (2019) 123005.

[23] Z. Bu, H. J. Li, J. Cao, Z. Wang, G. Gao, IEEE Trans. on Cyb. 1 (2017) 49.

[24] J. Gao, Y.-W. Dong, M.-S. Shang, S.-M. Cai, T. Zhou, Group-based ranking method for online rating systems with spamming attacks, EPL (europhysics letters) 110 (2) (2015) 28003.

[25] J. Gao, T. Zhou, Evaluating user reputation in online rating systems via an iterative group-based ranking method, Physica A: Statistical Mechanics and its Applications 473 (2017) 546–560.

[26] D. Chen, H. L. Sun, Q. Tang, S. Z. Tian, M. Xie, Chaos 29 (2019) 033120.

[27] H. L. Sun, E. Chng, J. M. Garibaldi, S. Simon, D. B. Chen, Physica A 496 (2018) 108–120.

[28] D. Lee, M. J. Lee, B. J. Kim, Deviation-based spam-filtering method via stochastic approach, EPL (Europhysics Letters) 121 (6) (2018) 68004.

[29] Y.-Y. Wu, Q. Guo, J.-G. Liu, Y.-C. Zhang, Effect of the initial configuration for user–object reputation systems, Physica A: Statistical Mechanics and its Applications 502 (2018) 288–294.

[30] H. L. Sun, E. Chng, S. Simon, IMDS 1 (2018) 119.

[31] M.-S. Shang, L. Lü, Y.-C. Zhang, T. Zhou, Empirical analysis of web-based user-object bipartite networks, EPL (Europhysics Letters) 90 (4) (2010) 48006.

[32] J. L. Herlocker, J. A. Konstan, L. G. Terveen, J. T. Riedl, Evaluating collaborative filtering recommender systems, ACM Transactions on Information Systems (TOIS) 22 (1) (2004) 5–53.

[33] J. A. Hanley, B. J. McNeil, The meaning and use of the area under a receiver operating characteristic (roc) curve., Radiology 143 (1) (1982) 29–36.