University of Reading

The effect of product distance on the eWOM in recommendation network

Article

Accepted Version

Pan, X., Hou, L. and Liu, K. (2022) The effect of product distance on the eWOM in recommendation network. Electronic Commerce Research, 22 (3). pp. 901-924. ISSN 1572-9362 doi: 10.1007/s10660-020-09432-1 Available at https://centaur.reading.ac.uk/91908/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1007/s10660-020-09432-1

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

University of Reading

Reading's research outputs online

The Effect of Product Distance on the eWOM in Recommendation Network

Xue Pan[1,2], Lei Hou[1,2], Kecheng Liu[2]

1. School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nanjing, China
2. Informatics Research Centre, Henley Business School, University of Reading, UK

Email: x.pan@nuist.edu.cn; l.hou@nuist.edu.cn; k.liu@henley.ac.uk

Abstract

The online Product Recommendation Networks (PRNs), connecting similar products with hyperlinks, have been widely implemented in user-generated content websites and ecommerce systems. With the PRNs as the virtual shelves, this paper explores the impact of the distance between products on the formation of product electronic Word-of-Mouth (eWOM). Employing an empirical book recommendation network of Amazon, the study one explores the effect of a focal product's neighborhood (nearby others) on its eWOM, and study two explores the eWOM similarity between product pairs that are at one, two and three clicks away from each other. The results reveal the significant role played by the product distance on the association of their eWOM. On one hand, a focal product's eWOM is largely influenced by that of its neighborhood. On the other hand, the good connectivity between two products, which is defined as the number of paths connecting them, is closely associated with the eWOM similarity between them. The findings suggest that the products should be considered as interactive collectives rather than separated individuals particularly in the eWOM studies.

Keywords: product recommendation network, eWOM similarity, electronic commerce, network distance

1. Introduction

In traditional brick-and-mortar grocery stores, the product position is physically represented by shelf display, which has been found significantly influencing product selection [1]. Products with more facings or placed on more prominent shelf positions will get more attentions first, which is normally referred to as shelf effect in the literature [2-3].

Nowadays, with the dramatic development of online commercial marketplaces, "virtual shelves" have been constructed by putting products at some specific online positions, to provide consumers a flexible and personalized experience [4]. Among these virtual shelves, the Product Recommendation Network (PRN) is one of the most popular forms, which is organized by connecting similar products with hyperlinks. An example is Amazon's "Customers who bought this item also bought" list. Such PRN can navigate consumers to explore different yet relevant products [5]. Hence, a consumer browsing products in a PRN is largely similar to the scenario of walking down an aisle in an offline store. It has been found

that PRNs play an important role on consumers' choices [6] and the shelf effect can also be found in the online stores.

Since every product has a position in the PRN, any pair of products thus have a specific distance between them, such as one can reach another in two clicks. The literature has suggested the primacy effect that products encountered earlier during the online search process or products next to a focal item will receive more attentions [7]. Thus, the connected products in the PRN normally have similar attributions, leading to similar demands and sales [8]. As another focus of marketing studies, the electronic Word-of-Mouth's (eWOM) impact on consumer purchase intention [9-10] or consumer engagement of posting [11] has been widely studied. With the increasing attentions on the PRNs in recent year, the link between PRN and eWOM has been explored, suggesting that two directly-connected products in a PRN would have similar eWOM rating with each other [12]. However, whether the indirect connections have the similar effect, and what is the influence of the number of paths connecting two products are still unknown.

In addition, the anchoring bias theory suggests that individuals tend to make decisions based on earlier information (anchor) [13]. For example, in an auction, people would overestimate a painting's value if the former one was bided by a high price. Such effect has also been found in online voting, that a user tends to give a high rating on a subsequent movie after watching a high-quality movie [14], or may consecutively give high/low ratings to movies s/he watched [15]. Considering the navigation function of PRNs, the structure of the PRNs may largely determine the sequence of browsing behavior of consumers. Hence, the anchoring bias could be applied to the eWOM correlation (such as ratings) of products that are near to each other in the PRN.

Yet, such link between the PRN and eWOM is still not fully explored. Most product pairs in the PRNs, though are not directly connected with each other, have connections of paths with certain distances (larger than 1). Besides, the purpose of PRNs is to distribute web traffic to appropriate products, and users are normally surfing on such network, from one product to another following hyperlink paths. Therefore, whether a pair of products with a certain distance have similar eWOM is a key question that may help to further understand the consumer behavior and product performances in the PRNs.

To explore such question, we employ an empirical book recommendation network collected from Amazon along with the eWOM information of every book. The analysis is carried out from two levels, namely the neighborhood-to-product level and dyadic product-to-product level. An nth-order neighborhood of a focal book is defined as the collective of books who have shortest distance of exactly n to the focal book in the PRN, i.e. it takes at least n clicks on the recommendation hyperlinks to reach the focal book. For the neighborhood-to-product level, we analyze the association between the average volume and rating [10-11] of the neighborhood and that of the focal books. For the dyadic product-to-product level, we focus on the connectivity of product pairs which is defined as the easiness for a consumer to visit from one to another by following the recommendations, measured as the number of paths

between the two books. Accordingly, we investigate the effect of such connectivity on the eWOM rating difference between the books. According to the results, the eWOM of focal books is largely related to the neighbors' eWOM, and such impact can reach three clicks away. Products that are close to each other indeed have similar eWOM in the recommendation network, and the more paths connecting two products, the more similar their eWOM will be.

The rest of the paper is organized as follows. Section 2 reviews the relevant literature on eWOM and PRN respectively, and develops the hypotheses based on the literature review and theoretical foundation. Section 3 introduces the data collection method of this paper and the model specification. The results are presented in section 4. Section 5 concludes and discusses the findings of the study. Section 6 summarizes the contributions including theoretical and practical implications. At last, the limitations and future work are presented in section 7.

2. Related Work and Hypotheses Development

2.1 Electronic Word-of-Mouth (eWOM)

Given the rapid development of ecommerce platforms, eWOM has become an important component in the marketplace environment [16]. Particularly, being one of the most common forms of eWOM, online product reviews have become important sources of information for consumers to compose the intensions and impressions towards the products. Thus, eWOM research has attracted a lot of attentions.

Some studies believed that the eWOM can significantly influence consumers' purchase decisions [17-18] and the products' subsequent sales [17, 19-21]. Volume, valance and variance of online reviews are the most important attributes that the related literature concerns. Chevalier and Mayzlin [19] reported that the volume and valance of consumer reviews have positive impact on the book's subsequent sales in Amazon. Li and Hitt [21] found similar results for the website of Barnesand-Noble.com. Liu [22] and Duan [17] found that the review volume has significant positive impact on the box sales of movies. Dellarocas et al. [23] showed that the valance of online ratings posted during a movie's opening weekend is the most important indicator for its revenue trajectory in the following weeks. Similarly, Chintagunta et al. [20] found that the rating valance of pre-release advertising is the main driver of box office performance. Clemons et al. [24] found that the beer sales grow faster for the brands with higher variance of consumer ratings. Godes and Mayzlin [25] reported that the variance of opinions on weekly TV shows across the Internet communities is positively related to the evolution of viewership.

Another focus of eWOM research examines why consumers would like to engage in the eWOM activities, i.e. the motivations for eWOM. Hennig-Thurau et al. [26] concluded that social benefits, economic incentives, concern for others, and extraversion are the main motivations for consumer to participate in eWOM on the Internet. Berger and Iyengar [27] examined the psychological factors that are related to the immediate and ongoing eWOM,

thereby found that more interesting products could get more immediate eWOM but do not receive more ongoing eWOM over time. In addition, the motivations of positive and negative eWOM have also been found to be associated with different antecedents [28]. For example, consumers who have a favorable attitude towards the eWOM communications intend to post positive reviews, whereas negative eWOM is more driven by social pressure.

Pan et al. [11] explored the impact of the social network on the posting behaviors. They distinguished the impact of the reviews from friend and crowd and showed that friend reviews are more determinant to drive a consumer's engagement to eWOM behaviors. Recent study from Dixit et al. [29] suggest that perceived behavioral control, subjective norms, ego involvement, and taking vengeance are significant predictors for consumers to engage in posting reviews for restaurants.

By far, the majority of eWOM studies regard product as isolated entities and explore the eWOM of each product as individual. As suggested by the anchoring bias theory [13-14], the browsing/purchasing sequence of consumers will have significant influence on their evaluation on different products. Hence, the eWOM of products have potential interactions with each other. However, such interaction among products in terms of their eWOM has not been fully explored.

2.2 Product Recommendation Network (PRN)

Many ecommerce websites evaluate the similarities among products, in terms of how often they are purchased together by the same consumers. Accordingly, recommendations can be made for each product by uncovering the products that have been most frequently "co-purchased" with the target product. Examples can be found in various websites, such as the "Customers who bought this item also bought…" list of Amazon, the "People also viewed" list in Yelp and so on. Regarding the products as nodes, the recommendation hyperlinks thus connect the nodes as a massive network, which is normally referred to as the product recommendation network (PRN), or co-purchase network.

The PRNs enable users to browse diverse products by clicking and following the hyperlinks, thereby influencing the consumption pattern [30]. Assuming the revenue of a product as the summation of its intrinsic value and incoming value, Oestreicher-Singer et al. [31] estimated the network value of products, which consists the value generated by itself, and the value it contributes to other products. Hou et al. [5] explored the users' surfing behavior in Amazon and found that the recommendation network of Amazon tends to rapidly navigate users to very popular books, leading to the monopoly of web traffic by these blockbusters. Goldenberg et al. [32] integrated the social and product network as a dual-network and found it to be able to facilitate the process of content exploration.

The PRN is also found with significant effect on the sales and demand of products in it. Oestreicher-Singer and Sundararajan [8] quantified the incremental correlation between book sales and the visibility in the book network. The PageRank centrality, which is a measure for a node's position in a network, and some other quantities such as the in-degree, are

closely associated with the books' demand as measured by the sale rank in Amazon [33]. Leem and Chun [34] further examined the other centrality measures, including degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality, and confirmed that a book's demand is vastly influenced by its position in the PRNs. Lin et al. [35] explored the impact of network diversity and stability on product demand. Huang et al. [36] found the products with more connected links to have greater demands.

In summary, most literature explores the relationship between a product's demand or sales and its position in the PRN, such as the visibility, centrality measures etc. Actually, one of the most important functions of PRN is to navigate consumers to browse relevant products, through which, sequences of browsing may be resulted according to the local structure of the PRNs. Thus, the paths connecting two products would be of great importance. There are two recent studies made efforts towards such direction. Lin and Wang [12] studied the effect of direct paths, e.g. one product recommends another, on the rating similarity between the two products. However, most product pairs have indirect paths. For example, A recommends B while B recommends C, and then products A and C have an indirect path with length of two clicks. Whether such indirect paths have effects on the eWOM convergence is still unknown. On the other hand, Carmi et al. [37] reported that the boost in sale for one product may spread to its nearby neighbors up to four clicks away. It is a clear sign that the paths between products have significant effect on product performance. Despite the great significance, they ignored the fact that two products may have multiple paths connecting them, which can well describe the possibility for consumers to surf from one to another. How would the number of paths influence the product demand is still unknown.

2.3 Theoretical Foundation and Hypotheses Development

As discussed in section 2.1 the eWOM studies lack of investigations where the products are considered interactive. PRNs provide a very good scenario for such studies, where the products are connected as a network by recommendation hyperlinks. Meanwhile, section 2.2 discussed the gap of in-depth exploration on the effect of paths on product performance. Hence, in this paper, we study the impact of product distance in the PRN, i.e. the length and number of paths, on their eWOM similarity.

An apparent question thus raised that why would the product distance have an effect on their eWOM? The literature of shelf display and management provides good theoretical foundation to justify such relationship. It has been demonstrated in the marketing research that shelf allocation or product display in traditional brick-and-mortar stores, has a great influence on consumer purchasing behavior and substantial demand of products [7,26,38-39]. Previous literature suggested that product display such as categorizing the same set of products by keeping actual assortment size equal or organizing product based on the physical position, could influence consumer perceptions and inferences about the product popularity [40-41]. The product distance based on an effective layout is shown to stimulate the demands and sales of products by making them more visible. Chen et al. [38] identified and classified the effect of relationships between relative spatial distance of displayed products and the

items' sales in a retail store. As suggested by proximity effect [7], products located next to the focal item have a higher probability of selection than isolated ones.

To examine whether and how such offline shelf effect translate to an online grocery context, Breugelmans et al. [7] found that products encountered earlier during the online search process will receive more attention and would be selected with a higher probability (primacy effect), and product located next to the focal item have a higher probability of being selected as well (proximity effect). The study of Ert and Fleischer [42] revealed the significant influence of the display position of hotels in a list of relevant hotels on the likelihood for them to be selected. In-store display of products in online context has a positive effect on the sales of the displayed products [4]. These studies suggest that the selection of a product in an online store is largely related to where it is displayed.

PRN can be regarded as a "virtual shelf" that displays products. Thus, we believe that the products distance in a PRN will have an effect on their performance. According to the Stimulus-Organism-Response (SOR) paradigm, the online environmental stimuli influence consumers' internal states, which in turn influence consumes' overall responses such as consumer satisfaction or trust that can be reflected by eWOM [43].

We aim to study the distance effect from two levels, namely the neighborhood-to-product level, and the product-to-product level. Firstly, we believe that the eWOM of a focal product is largely influenced by the "neighborhood" it belongs to, which is the collective of products that have very short distance to the focal product. By "short distance", in this paper we mean the distance within three clicks away. For example, if a focal product gets recommended by a lot of highly-rated others (who have hyperlinks connecting it), it is more likely that the focal product would also get a high rating. The previous findings on the spreading of sale in PRN [33] also inspired such assumption. As such, we develop the following hypotheses:

***H1a**: The review volume of a focal product is positively influenced by that of its neighborhood of products who have paths of length up to three clicks connecting to it.*

***H1b**: The average rating of a focal product is positively influenced by that of its neighborhood of products who have paths of length up to three clicks connecting to it.*

Secondly on the product-to-product level, we focus on product pairs. Since Lin and Wang [12] have reported the influence of direct recommendation relationships on the rating similarity, which is in accordance with the anchoring bias theory [13], here we re-test such observation to confirm whether it applies also to the Amazon book network context. Accordingly, we have the following hypothesis:

***H2a**: The directly connected products have similar average rating with each other.*

As we discussed in section 2.2, there lacks the study on the indirect paths' impact on product performances in the literature. Considering the fact that most product pairs actually are not

directly connected, we study the effect of indirect connections in a detailed manner by introducing the concept of connectivity. The connectivity measures the number of paths with a certain length connecting two products, thereby describing the easiness of surfing from one product to another. In other words, we not only examine the rating similarity between two products with certain distance, but also explores the effect of the strength of their connection. Thus, we hypothesize the following.

***H2b**: The product pairs with strong indirect connectivity tend to have similar average rating with each other.*

3. Methodology and Model Development

3.1 Data collection

The data applied in this paper is a book recommendation network which we collected from Amazon via a self-developed Python-based web crawler. In the Amazon system, each book has a unique ID, and the webpage of the book is composed as "http://www.amazon.com/dp/**ID**", where the "ID" should be replaced by a real ID. Consequently, to collect the book network is actually to collect the corresponding webpages and the recommendation hyperlinks connecting them.

We firstly selected 5 books as the seeds of the crawling from the Amazon's bestseller list (www.amazon.com/gp/bestsellers/books). Note that, the list may change from time to time and in our collection, the seeds were collected on 1st January 2016. For each of the seed books, we collect books from its recommendation list known as "Customers who bought this item also bought" list. We follow such recommendation hyperlinks originated from each of the seeds to collect the recommended books based on a width-first search. While normally 100 similar recommendations are offered in each book's recommendation list, there are generally 5 to 10 recommendations in the first page depending on the window size of the web browser. Assuming the recommendations displayed in the first page would get most attentions, we collect only the 10 books ranking at the top of the list as the current one's recommendations. The search continued for 8 steps. And for the 8th-order books, we collect the first ten recommendations that have already been collected in previous steps out from the list as their recommended books, in purpose of avoiding new books. Accordingly, 157,856 books are collected. In addition, the reviews of each collected books are also downloaded. There are in total of 4,520,102 reviews for the collected books.

The structure of the PRN largely depends on the recommendation algorithm applied by Amazon, which is not exactly known to us. While we collected the list of "customers who bought this item also bought", there are also lists in Amazon pages such as "customers who viewed this item also viewed" etc. Hence, according to the title of the list, the collected Amazon book network is organized according to a certain collaborative filtering algorithm [44] (association rule) based on the co-purchase records. The viewing data, i.e. consumers' view sequence, shall have been considered in a separate recommendation network which is not in

this paper's concern. Given the constantly evolving purchase data, the structure of Amazon PRN may be updated from time to time. The PRN in this paper was collected in January 2016. Thus, we mainly focus on the review posted in December 2015, and assume that our collected network can well represent the structure of the recommendations in this month. The selected period covers 34,100 books that have at least one review and have been recommended at least once, and a total of 92,405 associated reviews.

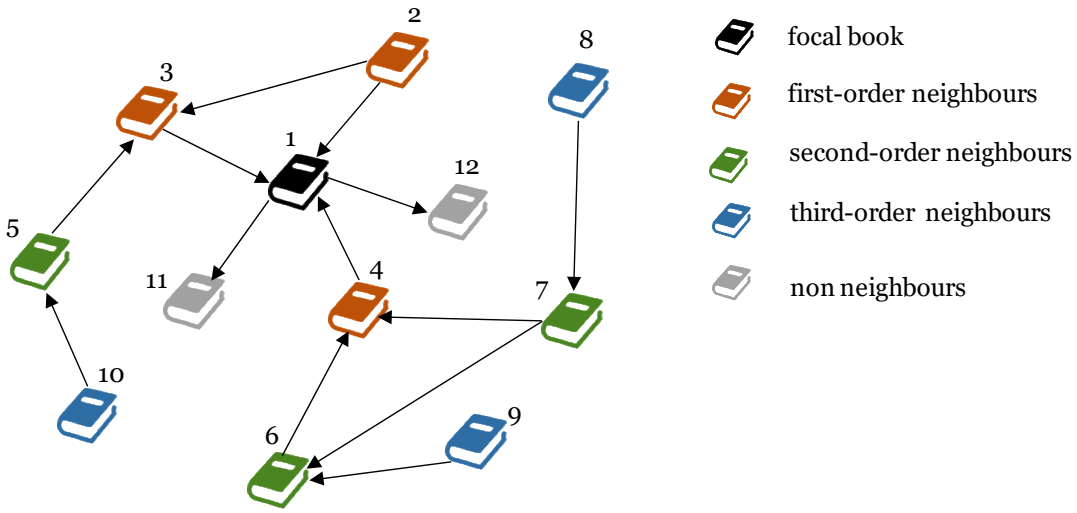

Figure 1. A Toy Network to Illustrate Amazon Book Recommendation Network

Figure 1 shows a simple recommendation network that the books are connected with each other via directed hyperlinks. Between every pair of books, there would normally be a certain distance, which represents how many clicks at minimum it takes for a user to start from a book to visit another. For example, book 3 needs one click to reach book 1, while book 10 needs at least three clicks to reach book 1. For a focal book, we define another book as one of its nth-order neighbors if this book has a distance of n to reach the focal book. Accordingly, as shown in Figure 1, taking the book 1 as the focal book, books 2, 3, 4 are its first-order neighbors; books 5, 6, 7 are its second-order neighbors; and books 8, 9, 10 are its third-order neighbors. Note that, we only consider books from which a user can reach the focal book as the neighbors, while ignore these books that can only be reached by the focal book. Accordingly, book 11 and 12 will not be considered as the focal book's neighbors.

3.2 Variable Operationalization

The analysis in this paper will be addressed from two levels to test the hypotheses: neighborhood-to-product level and dyadic product-to-product level. The research models, the variable operationalization and the results will be introduced respectively in the following.

3.2.1 Neighborhood-to-product level

The neighborhood level examines the correlation between a focal book's eWOM and the average level of its neighbors. Thus, it mainly addresses the hypothesis H1, answering the question that to what extent a product's eWOM is determined by its nearby neighbors.

Dependent variables

The dependent variables in this part measures the eWOM of the focal books. We operationalize the eWOM with two most direct measurements known as the review volume and the average rating (valance). Accordingly, the dependent variables are the number of reviews at day t, $RV_i(t)$ and the average rating at day t, $R_i(t)$, for each focal book i.

Independent variables and control variables

The independent and control variables are basically the eWOM information of the focal books' neighbors. We only consider the neighbors that are within three clicks from the focal books, because a lager distance may include most of the population as the neighbors, which will be of less value and requires much more computational cost. Accordingly, there are three categories of independent variables namely the variables for the first-order neighbors, second-order neighbors and third-order neighbors, denoting with (1), (2) and (3) respectively.

We consider the average value of daily review volume and daily average rating of focal books' neighbors as the independent variables. For a neighbor book j, its review volume and average rating at the day t are denoted with $V_j(t)$ and $R_j(t)$. Note that, not every neighbor book has reviews every day in the studied period. If a neighbor book j does not have reviews at the day t, we regard its volume as $V_j(t) = 0$. But for the rating, a value of 0 would suggest an extremely low rating. Therefore, we regard the average rating of such books as the system average of all reviews $R_j(t) = 4.299$ so that it represents a neutral rating. In this way, for a focal book i, there are two independent variables for an order (n), namely the average review volume of the nth order neighbors,

$$Ave_RV_i^{(n)}(t) = \frac{1}{Indeg_i^{(n)}} \sum_{j \in \mathbf{\Gamma}_i^{(n)}} V_j(t), \quad (1)$$

and the average rating of its neighbors,

$$Ave_Rating_i^{(n)}(t) = \frac{1}{Indeg_i^{(n)}} \sum_{j \in \mathbf{\Gamma}_i^{(n)}} R_j(t), \quad (2)$$

respectively, where $\mathbf{\Gamma}_i^{(n)}$ is the set of books which are the focal book i's nth-order neighbors with a population of $Indeg_i^{(n)}$. Accordingly, we have the independent variables $Ave_RV_i^{(n)}(t)$ and $Ave_Rating_i^{(n)}(t)$, with $n = 1, 2, 3$.

Considering the visibility of focal books largely varies from each other, we take the number of neighbors at each order as the control variables. Hence, $Indeg_i^{(1)}$, $Indeg_i^{(2)}$ and $Indeg_i^{(3)}$ represent the number of first-order neighbors, second-order neighbors and third-order neighbors of the focal book i respectively. In addition, we control the analysis with the historical number of reviews of the focal book. To be specific, for a focal book i, the number of reviews that are posted before 1st December 2016, denoting with NOR_i is also a control variable. The control variables do not change over time, as we have assumed that the structure of recommendations in the studied time period is fixed.

The operationalization and descriptive statistics of all variables are shown in Table 1, while the Table 2 reports the correlations among these variables.

Table 1. Descriptive Statistics of Variables for the Neighborhood-to-Product Level Study

Variable Type	Variable Name	Descriptions	Mean	Std. Dev.	Min	Max
Dependent Variables	$RV_i(t)$	review volume of the focal book i at day t.	1.205	0.996	1	107
	$R_i(t)$	average rating of the focal book i at day t.	4.449	1.043	1	5
Independent Variables	$Ave_RV_i^{(1)}(t)$	average review volume of first-order neighbors of focal book.	0.073	0.150	0	7.1
	$Ave_RV_i^{(2)}(t)$	average review volume of second-order neighbors of focal book.	0.065	0.076	0	3
	$Ave_RV_i^{(3)}(t)$	average review volume of third-order neighbors of the focal book.	0.028	0.035	0	2.22
	$Ave_Rating_i^{(1)}(t)$	average review rating of first-order neighbors of the focal book.	4.428	0.083	1	5
	$Ave_Rating_i^{(2)}(t)$	average review rating of second-order neighbors of the focal book.	4.423	0.061	2.71	5
	$Ave_Rating_i^{(3)}(t)$	average review rating of third-order neighbors of the focal book.	4.422	0.016	3.74	4.81
Control Variables	NOR_i	number of historical reviews.	188.9	376.5	1	5263
	$Indeg_i^{(1)}$	number of first-order neighbors.	29.60	51.0	1	911
	$Indeg_i^{(2)}$	number of second-order neighbors	198.1	478.5	1	10204
	$Indeg_i^{(3)}$	number of third-order neighbors.	823.3	1842	1	26414

Table 2. Correlations among Variables for Neighborhood-to-Product Level Study

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. $RV_i(t)$	1											
2. $R_i(t)$	0.02	1										
3. $Ave_RV_i^{(1)}(t)$	0.17	0.04	1									
4. $Ave_RV_i^{(2)}(t)$	0.14	0.01	0.28	1								
5. $Ave_RV_i^{(3)}(t)$	0.05	0.00	0.20	0.33	1							
6. $Ave_Rating_i^{(1)}(t)$	0.00	0.12	0.16	0.06	0.01	1						
7. $Ave_Rating_i^{(2)}(t)$	0.00	0.50	0.06	0.08	-0.01	0.14	1					
8. $Ave_Rating_i^{(3)}(t)$	0.00	0.05	0.02	0.04	0.08	0.03	0.04	1				
9. NOR_i	0.27	0.02	0.16	0.00	0.06	0.00	0.01	0.01	1			
10. $Indeg_i^{(1)}$	0.34	0.01	0.06	-0.13	-0.02	-0.02	-0.0	-0.0	0.47	1		
11. $Indeg_i^{(2)}$	0.31	0.00	0.11	-0.09	-0.02	-0.02	-0.01	-0.01	0.44	0.92	1	
12. $Indeg_i^{(3)}$	0.28	-0.01	0.12	-0.08	-0.01	-0.03	-0.01	-0.01	0.45	0.83	0.95	1

3.2.2 Dyadic product-to-product level

The dyadic level explores the correlation between the connectivity and eWOM similarity between two individual products. Addressing hypothesis H2, it explores the association between the paths connecting two products and their eWOM similarity. Theoretically, every

pair of books can be regarded as an observation. However, given the population of the data, such number of book pairs would be very large. Since we only explore the books with distances of within three clicks, here we consider the pairs of books, between which there is at least one path of either length of 1, 2, or 3 connecting them.

Dependent variable

In this study, we use the difference of eWOM rating during the studied time month for each product pair consisting of book i and book j as the dependent variable, denoting with $WOM_diff_{ij} = |R_i - R_j|$. Accordingly, a smaller (larger) value of WOM_diff_{ij} indicates that the books i and j have more similar (different) eWOM.

Independent variables and control variables

Since the two products being connected or not can only partially represents their connectivity, here in this study, we use the number of paths between the measured two products as the independent variables. For two books i and j, we count how many paths of length n are connecting them, denoted with $NOP_{ij}^{(n)}$. Therefore, such value can well describe the traffic flow between the two products, i.e. how easily can consumers surf from one to the other. A larger value of $NOP_{ij}^{(n)}$ represents better connectivity between the two products. In respect to the neighborhood level analysis, here we also consider distances up to three, i.e. the number of paths $NOP_{ij}^{(1)}$, $NOP_{ij}^{(2)}$, $NOP_{ij}^{(3)}$. While $NOP_{ij}^{(1)}$ can only take values of 1 or 0, $NOP_{ij}^{(2)}$ and $NOP_{ij}^{(3)}$ could be any integer values. Actually, Lin and Wang (2018) have already proved that the direct connection ($NOP_{ij}^{(1)} = 1$) between two products can lead to the convergence of their eWOM rating. In this work, we consider in addition the indirect connections, namely $NOP_{ij}^{(2)}$ and $NOP_{ij}^{(3)}$, to explore the effect of distance between two products in the network on their eWOM similarity. In addition, we define the connectivity between two products by combining $NOP_{ij}^{(1)}$, $NOP_{ij}^{(2)}$, $NOP_{ij}^{(3)}$ to measure the likelihood for consumers to surf from the homepage of product i to j within three clicks,

$$Connectivity_{ij} = \frac{NOP_{ij}^{(1)}}{L} + \frac{NOP_{ij}^{(2)}}{L^2} + \frac{NOP_{ij}^{(3)}}{L^3}, \quad (3)$$

where L is the number of recommendations per product, i.e. out-going degree of each product. In our collected data, $L = 10$.

We use the differences between the basic information of two products as the control variables including the difference of in-degree of two books, $Indeg_diff_{ij} = |Indeg_i - Indeg_j|$; and the difference of review volume at the studied time period, $RV_diff_{ij} = |RV_i - RV_j|$.

The operationalization and descriptive statistics of all variables for the dyadic level are shown in Table 3, while the Table 4 reports the correlations among these variables.

Table 3. Descriptive Statistics of Variables for the Dyadic Level Study

Variable Type	Variable Name	Descriptions	Mean	Std. Dev.	Min	Max
Dependent Variables	WOM_diff_{ij}	Difference of eWOM rating between product i and j.	0.80	0.94	0	4
Independent Variables	$NOP_{ij}^{(1)}$	Number of paths from product i to j by one click.	0.05	0.21	0	1
	$NOP_{ij}^{(2)}$	Number of paths from product i to j by two clicks.	0.44	1.10	0	10
	$NOP_{ij}^{(3)}$	Number of paths from product i to j by three clicks.	4.60	8.91	0	91
	$Connectivity_i$	The easiness of getting to product j from i in within three clicks.	0.013	0.036	0.001	0.281
Control Variables	$Indeg_diff_{ij}$	Difference of in-degree between product i and j.	44.58	85.4	0	910
	RV_diff_{ij}	Difference of review volume between product i and j.	9.91	28.6	0	411

Table 4. Correlations among Variables for the Dyadic Level Study

Variables	1	2	3	4	5	6	7
1. WOM_diff_{ij}	1						
2. $Indeg_diff_{ij}$	-0.018	1					
3. RV_diff_{ij}	-0.027	0.551	1				
4. $NOP_{ii}^{(1)}$	-0.015	-0.011	-0.007	1			
5. $NOP_{ii}^{(2)}$	-0.020	0.029	0.028	0.574	1		
6. $NOP_{ii}^{(3)}$	-0.024	0.066	0.065	0.667	0.901	1	
7. $Connectivity_{ij}$	-0.020	0.019	0.021	0.910	0.852	0.890	1

3.3 Model Specification

For the analysis on the neighborhood level, we develop two regression models to explore how the neighbors' eWOM that are up to three clicks away, influence the eWOM of the focal books, in terms of the review volume and average rating respectively. The detailed models can be described as,

$$\begin{aligned}\boldsymbol{RV_i(t)} = \beta_0 + \beta_1 NOR_i + \beta_2 Indeg_i^{(1)} + \beta_3 Indeg_i^{(2)} + \beta_4 Indeg_i^{(3)} + \\ \beta_5 Ave_RV_i^{(1)}(t) + \beta_6 Ave_Rating_i^{(1)}(t) + \beta_7 Ave_RV_i^{(2)}(t) + \\ \beta_8 Ave_Rating_i^{(2)}(t) + \beta_9 Ave_RV_i^{(3)}(t) + \beta_{10} Ave_{Rating}{}_i^{(3)}(t) + \varepsilon_i, \end{aligned} \quad (4)$$

$$\begin{aligned}\boldsymbol{R_i(t)} = \beta_0 + \beta_1 NOR_i + \beta_2 Indeg_i^{(1)} + \beta_3 Indeg_i^{(2)} + \beta_4 Indeg_i^{(3)} + \\ \beta_5 Ave_RV_i^{(1)}(t) + \beta_6 Ave_Rating_i^{(1)}(t) + \beta_7 Ave_RV_i^{(2)}(t) + \\ \beta_8 Ave_Rating_i^{(2)}(t) + \beta_9 Ave_RV_i^{(3)}(t) + \beta_{10} Ave_Rating_i^{(3)}(t) + \varepsilon_i. \end{aligned} \quad (5)$$

We use the logarithms of each control variable in the model.

For the analysis on the dyadic product-pair level, given the high correlation (see Table 4) among the variables of number of paths and the connectivity, we explore their effect in turns with the following model:

$$\boldsymbol{WOM_diff_{ij}} = \beta_0 + \beta_1 RV_diff_{ij} + \beta_2 Indeg_diff_{ij} + \beta_3 Rating_diff_{ij} + \beta_4 X + \varepsilon_i, \quad (6)$$

where X will operationalized as $NOP_{ij}^{(1)}$, $NOP_{ij}^{(2)}$, $NOP_{ij}^{(3)}$ and $Connectivity_{ij}$ in turns to analyze the impact of different product distance separately on the eWOM similarity.

4. Results

4.1 Study one: neighborhood-to-product level

4.1.1 Review volume

Table 5. Regression Results of Focal Books' Review Volume

DV: $RV_i(t)$				
	Model 1	Model 2	Model 3	Model 4
Variables	Coef. (Std. err.)	Coef. (Std. err.)	Coef. (Std. err.)	Coef. (Std. err.)
NOR_i	0.050*** (0.003)	0.039*** (0.003)	0.027*** (0.003)	0.029*** (0.003)
$Indeg_i^{(1)}$	0.051*** (0.007)	0.092*** (0.007)	0.067*** (0.007)	0.068*** (0.007)
$Indeg_i^{(2)}$	0.227*** (0.012)	0.189*** (0.012)	0.338*** (0.012)	0.339*** (0.012)
$Indeg_i^{(3)}$	-0.100*** (0.009)	-0.100*** (0.009)	-0.159*** (0.009)	-0.159*** (0.008)
$Ave_RV_i^{(1)}(t)$		1.008*** (0.024)	0.530*** (0.025)	0.555*** (0.025)
$Ave_Rating_i^{(1)}(t)$		-0.195*** (0.043)	-0.174*** (0.043)	-0.180*** (0.041)
$Ave_RV_i^{(2)}(t)$			3.007*** (0.052)	3.156*** (0.053)
$Ave_Rating_i^{(2)}(t)$			-0.205*** (0.057)	-0.233*** (0.055)
$Ave_RV_i^{(3)}(t)$				-1.027*** (0.104)
$Ave_Rating_i^{(3)}(t)$				0.150 (0.213)
Constant	0.450*** (0.017)	1.331*** (0.192)	1.792*** (0.294)	1.280 (0.993)
Observations	75800	75800	75800	75800
R^2	0.068	0.090	0.128	0.129
F-statistic	1387	1245	1396	1128
Prob (F-statistic)	0.00	0.00	0.00	0.00
AIC	2.098e+05	2.080e+05	2.047e+05	2.047e+05
BIC	2.098e+05	2.081e+05	2.048e+05	2.048e+05
Observations	75800	75800	75800	75800

*** p<0.01, ** p<0.05, * p<0.1

Table 5 shows the results for the review volume of focal books. Four models are analyzed. Model 1 considers only the control variables, while the rest three models include the eWOM information of first-order, second-order and third-order neighbors progressively. The results of F test in these four models suggest the significant impact of the selected variables. When

the variables are progressively included, the value of R-Squared becomes bigger and AIC and BIC become smaller. The inclusion of first-order and second-order neighbors significantly improve the explanatory power of the modes (The increments of R-Squared are 0.090, 0.128 for Model 2 and Model 3 respectively). However, the impact of the third-order neighbors is not as strong as the first- and second-order neighbors, as shown from the value of R-Squared, where BIC and AIC do not change much. It means the impact of eWOM of neighbors on the focal books' review volume become slim when the distance between books reaches three clicks.

The impact of control variables, as shown in model 1, suggest that NOR_i, $Indeg_i^{(1)}$ and $Indeg_i^{(2)}$ positively and significantly impact the review volume of focal books. However, the number of books at three clicks away, $Indeg_i^{(3)}$ has a negative influence. This indicates that it is more likely for books to get more reviews if they have more reviews in the past, or more connected first-order and second-order neighbors in the recommendation network. When we include the eWOM information of the first-order neighbors into the model, as shown in Model 2, the average review volume of neighboring books, $Ave_RV_i^{(1)}(t)$ has a positive impact on the volume of focal books ($\beta = 1.008; p < 0.01$). However, the average rating of neighbor books, $Ave_Rating_i^{(1)}(t)$, lowers the review volume of focal books ($\beta = -0.195; p < 0.001$). The similar impact also exists in the second order neighbors as shown in Model 3. $Ave_RV_i^{(2)}(t)$ is positively and $Ave_Rating_i^{(2)}(t)$ is negatively influencing the review volume of focal books. These results suggest that, if there are a lot of eWOM discussion (volume) on a focal books' neighbors, it is more likely that the focal books also get more attentions; however, if the neighbors are highly-evaluated, there might be some potential competition relations among them which leads to less attention on the focal books. Such impact becomes weaker for these neighboring books at three clicks away, as shown in Model 4. It suggests that the impact of the average volume of the third-order neighbors, $Ave_RV_i^{(3)}(t)$ is negative while the average rating, $Ave_Rating_i^{(3)}(t)$ has no significant impact.

In summary, the review volume of focal books is positively related to the average volume of neighbors and negatively related to the neighbors' average rating. Such impact could come from the books that are two clicks away in the recommendation network.

4.1.2 Review rating

Table 6. Regression Results of Focal Books' Average Rating

DV: $R_i(t)$				
	Model 5	Model 6	Model 7	Model 8
Variables	Coef. (Std. err.)	Coef. (Std. err.)	Coef. (Std. err.)	Coef. (Std. err.)
NOR_i	0.021*** (0.003)	0.018*** (0.003)	0.013*** (0.002)	0.013*** (0.002)
$Indeg_i^{(1)}$	0.016* (0.008)	0.020* (0.008)	0.009 (0.007)	0.009 (0.007)

$Indeg_i^{(2)}$	-0.010 (0.013)	-0.008 (0.013)	-0.013 (0.011)	-0.014 (0.011)
$Indeg_i^{(3)}$	-0.004 (0.009)	-0.006 (0.009)	0.009** (0.008)	-0.009 (0.008)
$Ave_RV_i^{(1)}(t)$		0.112*** (0.026)	0.044*** (0.024)	0.040*** (0.024)
$Ave_Rating_i^{(1)}(t)$		1.474*** (0.047)	0.635*** (0.041)	0.627*** (0.041)
$Ave_RV_i^{(2)}(t)$			-0.458*** (0.050)	-0.489*** (0.052)
$Ave_Rating_i^{(2)}(t)$			8.555*** (0.055)	8.543*** (0.055)
$Ave_RV_i^{(3)}(t)$				0.141 (0.102)
$Ave_Rating_i^{(3)}(t)$				1.993*** (0.209)
Constant	4.391*** (0.018)	-2.143*** (0.208)	-36.25*** (0.283)	-44.97*** (0.955)
Observations	75800	75800	75800	75800
R^2	0.001	0.015	0.256	0.257
F-statistic	19.85	195	3260	2621
Prob(F-statistic)	0.00	0.00	0.00	0.00
AIC	2.211e+05	2.201e+05	1.988e+05	1.987e+05
BIC	2.212e+05	2.201e+05	1.988e+05	1.988e+05
Observations	75800	75800	75800	75800

*** p<0.01, ** p<0.05, * p<0.1

Table 6 shows the results of modelling the average ratings of focal books. Four models are studied with similar manner to the analysis of the review volume of focal books. The measures of model fit are basically same to the results in the last section. The results of F test in the four models suggest the significant impact of the selected variables on the focal books' rating. The R-Squared becomes bigger when the first- and second-order neighbors are considered but remains unchanged with the addition of third-order neighbors. The impact of the neighbors on focal books' review rating become slim when the distance reaches three clicks. According to AIC and BIC, the inclusion of eWOM information of neighbors can help to fit the model, leading to lower BIC and AIC. But the improvement becomes slim when the third-order neighbor is added in model 8, where the AIC and BIC has not changed from previous model.

When analyzing the control variables in Model 5, only $Indeg_i^{(1)}$ and NOR_i have significant and positive impact on the rating of focal books. The number of books that can reach it in two or three clicks, i.e. $Indeg_i^{(2)}$ and $Indeg_i^{(3)}$, do not have significant impact on the rating of the focal books.

When considering the eWOM information of the first-order neighbors in Model 6, both $Ave_RV_i^{(1)}(t)$ and $Ave_Rating_i^{(1)}(t)$ have positive impact on focal book's rating. This indicates that the feedback of books is strongly associated with the discussions of neighbors. If a focal book's direct neighbors are popular and highly-rated, it is more likely for the focal

book to be also highly-rated. For the second-order neighbors, as shown in Model 7, their rating has a positive impact on the focal books' rating ($\beta = 8.555; p < 0.001$) which is same as the first-order neighbors. However, the review volume shows a negative effect ($\beta = -0.458; p < 0.001$) which is opposite to that of the first-order neighbors. When it comes to the third-order neighbors, their ratings show significance, which is positive ($\beta = 1.993; p < 0.001$); while the review volume do not have significant impact on the focal books' rating, due to the long distance in between.

The findings suggest that a focal book's rating is associated with the ratings of its neighborhood which could be up to three clicks away. Such influence is positive, which means if one book has a high value of average rating, it is very likely for its recommended books, including direct recommendations, to have also positive feedback. The possible reason may lie in the anchoring bias effect [13], that when consumers evaluate the book being browsed, their judgement will be positively influenced by the quality of previously browsed books. But to confirm such reason may require further examination on the consumers' profiles. Secondly, the review volumes from neighborhood show mixed impact on the focal books' ratings, where that of the first-order neighbors is positive and that of the second-order neighbors is negative, while the third-order neighbors have no impact.

To summarize the results of the neighborhood-to-product analysis, the review volume and review rating of a focal book, is indeed closely correlated to that of its neighborhood, which well supports the hypothesis H1a and H1b. On the other hand, at the distance of three (third-order neighborhood), the correlation becomes minor, and sometimes not significant, which means the neighbors that are far away have less, or no, influence over the focal book' eWOM.

4.2 Study two: dyadic product-to-product level

Table 7. Regression Results on the eWOM Rating Similarity at Dyadic Level

	Model 9	Model 10	Model 11	Model 12
Variables	Coef. (Std. err.)	Coef. (Std. err.)	Coef. (Std.err.)	Coef. (Std. err.)
RV_diff_{ij}	0.420*** (0.001)	0.419*** (0.001)	0.419*** (0.001)	0.419*** (0.001)
$Indeg_diff_{ij}$	-1.46e-05** (6.58e-06)	-1.06e-05 (6.58e-06)	-5.92e-06 (6.59e-06)	-1.19e-05* (6.58e-06)
$NOP_{ij}^{(1)}$	-0.046*** (0.002)			
$NOP_{ij}^{(2)}$		-0.011*** (0.000)		
$NOP_{ij}^{(3)}$			-0.002*** (5.29e-05)	
$Connectivity_{ij}$				-0.359*** (0.013)
Constant	0.651*** (0.000)	0.654*** (0.000)	0.656*** (0.000)	0.653*** (0.000)
Observations	3,887,462	3,887,462	3,887,462	3,887,462
R-Squared	0.029	0.029	0.029	0.029
F-statistic	2.926e+04	2.935e+04	2.935e+04	2.935e+04

Prob(F-statistic)	0.00	0.00	0.00	0.00
AIC	1.043e+07	1.043e+07	1.043e+07	1.043e+07
BIC	1.043e+07	1.043e+07	1.043e+07	1.043e+07

*** p<0.01, ** p<0.05, * p<0.1

Considering the high correlation among the variables $NOP_{ij}^{(1)}$, $NOP_{ij}^{(2)}$, $NOP_{ij}^{(3)}$, and $Connectivity_{ij}$, these variables are examined by turns, rather than progressively added, in model 9 to 12. According to the results of F-statistic in these models, the studied variables are (number of path and connectivity) significantly related to the eWOM difference of two books. The estimated value of R-Squared, AIC and BIC are unchanged, suggesting that the studied connectivity variables may have similar influence over the eWOM difference between two products.

The control variables show significant impact on the eWOM similarity between two books. The difference of review volume RV_diff_{ij}, has a positive effect ($\beta = 0.420; p < 0.001$), which suggest that book pairs with similar popularity over a period of time (in this study, the December of 2015), would have similar eWOM rating. The difference between the books' in-degrees, i.e. how many times they get recommended, show significance in model 9 and 12, while no significant impact in model 10 and 11.

We investigate the impact of independent variables ($NOP_{ij}^{(1)}$, $NOP_{ij}^{(2)}$, $NOP_{ij}^{(3)}$ and $Connectivity_{ij}$) in turns as shown by Model 9 to 12. Generally, all variables have significant negative effect on the difference of eWOM rating, indicating that both direct and indirect connections between two products can enhance the similarity of the eWOM rating between a pair of products, and the more paths connecting the two products, regardless of the length (1, 2, 3 or combined) the closer ratings (smaller WOM_diff_{ij}) the two products will have. While Lin and Wang [12] suggest that the directly connected products ($NOP_{ij}^{(1)} = 1$) tend to have similar eWOM, the results in this section generalize such finding to indirect connections and further explore the strength of the number of such connections. Any paths which can connect two products within three steps will contribute to the similarity of their eWOM.

5. Conclusions

In this paper, we apply an empirical book recommendation network from Amazon to investigate the effect of distances between products in the recommendation network on their eWOM on two levels: neighborhood-to-product level and dyadic product-to-product level.

On the neighborhood-to-product level, the objective is to study the influence of the neighbors' eWOM on the focal books' eWOM in terms of daily review volume and average rating. Firstly, the review volume of a focal book is closely associated with that of its first- and second-order neighbors which show strong positive effect. The review volume of the third-order neighbors shows negative effect, but the increment of the r-squared indicates a weak marginal explanatory power. Secondly, the average rating of the first-, second-, and third-order neighbors positively influence the rating of the focal book. The marginal explanatory power

of the third-order neighbors is also not as strong as the first- and second-order neighbors. In conclusion, a focal book's review volume and rating is strongly influenced by its neighborhood within two clicks, while is also, to a smaller degree, influenced by its neighbors at three clicks away. Hence, the hypothesizes H1a and H1b are well supported.

On the dyadic product-to-product level, Lin and Wang [12] have previously found that the eWOM rating of two products would be similar to each other if one recommends another, revealing the important role of recommendation hyperlinks. We confirm such finding in our study, thereby supporting the hypothesis H2a. In addition, we extend the recommendation links to networks, and develop a measure of connectivity to describe the easiness for users to visit from one product to another following recommendation hyperlinks. Thus, the direct recommendation hyperlink (1-click distance) is a kind of strongest connectivity, while the 2-click and 3-click distances are weaker connectivity. The results indicate that the number of paths between two products significantly influences their eWOM similarity. Thus, we can conclude that not only the direct connections, but also the indirect connections, have strong positive correlations with the eWOM similarities between two products. Hence the hypothesis H2b is also supported. An integrated measure of connectivity, which is the likelihood for a user at a product to randomly surf to another following hyperlinks, indicates that the larger the traffic flow from one product to another, the more similar their ratings are.

6. Theoretical Contributions and Practical Implications

The work contributes to the literature from several aspects. Firstly, the research extends the study of PRN by providing new insights from the view of eWOM. While most previous studies focused on the economic impact of the PRN [12,31], we examine the pattern of eWOM in it. The findings suggest that the network connections between products, including both direct and indirect connections, play significant roles on the formation of products' eWOM. In other words, the PRN is able to redistribute the eWOM among products. While previous studies show that the demand and sales of products are largely associated with their positions in the network [33, 35], this paper provides a possible explanation that the eWOM of products are correlated when they are near (short distance) to each other and thereby the demands and sales of products are enhanced accordingly.

Secondly, the work enriches the theory of eWOM and help to understand consumer eWOM behaviors. Most prior studies examined the eWOM as either a driver [17, 19-20] to influence consumer purchase decisions and product sales, or an outcome that what motivations could lead to the eWOM activities [11,27]. Our work shows that the eWOM of different products may interact with each other, where one product's rating and review volume are associated with its neighbors in the network. This also indicates the possible patterns of the underlying user behavior, that users may be influenced by what s/he just browsed when composing eWOM for a particular product, as suggested by the anchoring bias effect.

Thirdly, the work fills the gap of literature addressing the link between the studies of eWOM and PRN. Though the literature studied the impact of network connection on eWOM

convergence [12], they focused only on pairs of products with direct connections at dyadic level. Here we extend their work by examining the impact of not only first-order neighbors, but also the neighbors that are two or three clicks away, and the number of paths with two or three clicks on the dyadic level. New insights are thus provided. On one hand, not only the direct connections, but also the indirect connections, have the converging effect on the two products' eWOM. On the other hand, a product's eWOM is largely determined by the group of neighboring products as a whole collective, rather than solely by some individual products, as indicated by the neighborhood-to-product level analysis.

The paper offers several practical implications. Firstly, the results show that the feedback of a product is strongly related to the discussions of its neighbors in the PRN. Actually, the consumers' surfing behavior may not always follow the recommendations, but also sometimes other hyperlinks displayed on the page of a product, such as the "sponsored items" in Amazon, or other forms of promotions in other e-commerce website. Therefore, retailers may consider putting promotion hyperlinks on the webpages of products whose eWOM performances are desirable to further facilitate their sales. Secondly, the work provides insights for online marketplace managers regarding the construction of their PRNs. Our findings show that the eWOM of products has impact on their recommended others, even the indirectly recommended ones. Thus, the PRN designer shall carefully consider about the application of specific recommendation algorithms, thereby optimize the structure of the PRN to aid the users to make their purchasing decisions.

7. Limitations and Future Work

While we showed promising findings and contributions, our work also has limitations that could be tackled in future studies. Firstly, we only examined book recommendation network based on the co-purchase in Amazon. As a typical experience good [45], the findings based on the book network shall be applicable to other experience goods, such as movies, restaurants, hotels, clothes etc. However, future work should investigate more types of products, such as the searching goods to further generalize the findings. Whether the findings of this paper can be applied to other platforms needs also further efforts to be confirmed. Secondly, in this study, the eWOM of neighbors as independent variables are averaged over all the neighbors of the same category. While it showed significance explaining the eWOM of focal books, it is worthy to look at other operationalizations for the variables such as aggregated value over the neighbors. In addition, there are other possible measures for the eWOM besides the volume and rating applied in this paper, such as the negative/positive reviews and text mining-based measures. Due to the sparsity problem of the data, these measures have not been considered in this paper. Future studies shall further explore such measures to deepen our understanding on the network effect of the eWOM. At last, while the structure of the PRN is determined by some recommendation algorithms a website applied, whether there is any interactions between the algorithm, such as collaborative filtering or content-based approach, and the eWOM formation pattern, is also an important

question which may help the system designers to better construct the PRNs and serve consumers with better experiences. Similarly, whether the recommendation algorithms have considered the eWOM similarity among products, is unknown. As a consequence, the analysis in this paper cannot fully confirm the casual relation between the PRN distance and eWOM similarity. The full exploration on such matter may require further analysis and perhaps experiments with the recommendation algorithms controlled.

Statement

The authors declare no conflict of interest.

Acknowledgements: This work is partially supported by the startup foundation of Nanjing University of Information Science and Technology (1441182001001, 1441182001008).

References

1. Desmet, P., & Renaudin, V. (1998). Estimation of product category sales responsiveness to allocated shelf space. *International Journal of Research in Marketing*, *15*(5), 443-457.
2. Campo, K., & Gijsbrechts, E. (2005). Retail assortment, shelf and stockout management: issues, interplay and future challenges. *Applied Stochastic Models in Business and Industry, 21*(4-5), 383-392.
3. Liang, T.-P., & Lai, H.-J. (2002). Effect of store design on consumer purchases: an empirical study of on-line bookstores. *Information & Management, 39*(6), 431–444.
4. Breugelmans, E., & Campo, K. (2011). Effectiveness of in-store displays in a virtual store environment. *Journal of Retailing, 87*(1), 75-89.
5. Hou, L., Liu, K., and Liu, J., (2017). Navigated Random Walks on Amazon Book Recommendation Network. In International Workshop on Complex Networks and their Applications, Lyon, France, pp. 935-945.
6. Senecal, S., and Nantel, J., (2004). The Influence of Online Product Recommendations on Consumers' Online Choices*. Journal of Retailing, 80*(2), 159-169.
7. Breugelmans, E., Campo, K., & Gijsbrechts, E. (2007). Shelf sequence and proximity effects on online grocery choices. *Marketing Letter, 18*(1-2), 117-133.
8. Oestreicher-Singer, G., and Sundararajan, A., (2012). Recommendation Networks and the Long Tail of Electronic Commerce. *MIS Quarterly, 36*(1), 65-83.
9. See-To, E.W., Ho, K.K., (2014). Value co-creation and purchase intention in social networksites: the role of electronic Word-of-Mouth and trust–A theoretical analysis. *Computers in Human Behavior, 31*, 182–189.
10. Kudeshia, C., & Kumar, A. (2017). Social eWOM: does it affect the brand attitude and purchase intention of brands?. *Management Research Review, 40*(3), 310-330.
11. Pan, X., Hou, L., Liu, K., and Niu, H., (2018). Do Reviews from Friends and the Crowd Affect Online Consumer Posting Behaviour Differently?. *Electronic Commerce Research and Applications, 29*(3), 102-112.
12. Lin, Z., and Wang, Q., (2018). E-commerce Product Networks, Word-of-mouth Convergence, and Product Sales. *Journal of the Association for Information Systems, 19*(1), 23-39.
13. Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-economics*, 40(1), 35-42.

14. Yang, Z., Zhang, Z. K., & Zhou, T. (2013). Anchoring bias in online voting. *Europhysics Letters*, 100(6), 68002.
15. Hou, L., Pan, X., Guo, Q., & Liu, J. G. (2014). Memory effect of the online user preference. Scientific reports, 4, 6560.
16. Allsop, D. T., Bassett, B. R., & Hoskins, J. A. (2007). Word-of-mouth research: principles and applications. Journal of Advertising Research, 47(4), 398-411.
17. Duan, W., Gu, B. and Whinston, A.B., (2008). Do Online Reviews Matter? An Empirical Investigation of Panel Data. *Decision Support Systems, 45*(4), 1007–1016.
18. Shihab, M. R., & Putri, A. P. (2019). Negative online reviews of popular products: understanding the effects of review proportion and quality on consumers' attitude and intention to buy. *Electronic Commerce Research, 19*(1), 159-187.
19. Chevalier, J.A. and Mayzlin, D., (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research, 43*(3), 345–354.
20. Chintagunta, P.K., Gopinath, S. and Venkataraman, S.,(2010). The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets. *Marketing Science, 29*(5), 944–957.
21. Li, X. and Hitt, L.M., (2008). Self-Selection and Information Role of Online Product Reviews. *Information Systems Research, 19*(4), 456–474.
22. Liu, Y., (2006). Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of Marketing, 70*(3), 74–89.
23. Dellarocas, C., Awad, N., Zhang, M., (2005). Using Online Ratings as a Proxy of Word-of-Mouth in Motion Picture Revenue Forecasting. Working Paper, Smith School of Business. University of Maryland.
24. Clemons, E. K., Gao, G. G., & Hitt, L. M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of Management Information Systems, 23*(2), 149-171.
25. Godes, D. and Mayzlin, D., (2004). Using Online Conversations to Study Word-of- Mouth Communication. *Marketing Science, 23*(4), 545–560.
26. Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet?. *Journal of Interactive Marketing, 18*(1), 38-52.
27. Berger, J., and Iyengar, R., (2013). Communication Channels and Word-of-Mouth: How the Medium Shapes the Message. *Journal of Consumer Reseasch, 40*(3), 567-579.
28. Fu, J. R., Ju, P. H., & Hsu, C. W. (2015). Understanding why consumers engage in electronic word-of-mouth communication: Perspectives from theory of planned behavior and justice theory. *Electronic Commerce Research and Applications, 14*(6), 616-630.
29. Dixit, S., Badgaiyan, A. J., & Khare, A. (2019). An integrated model for predicting consumer's intention to write online reviews. *Journal of Retailing and Consumer Services, 46*, 112-120.
30. Liu, Q., Zhang, X., Zhang, L., & Zhao, Y. (2019). The interaction effects of information cascades, word of mouth and recommendation systems on online reading behavior: An empirical investigation. *Electronic Commerce Research, 19*(3), 521-547.
31. Oestreicher-Singer, G., Libai, B., Sivan, L., Carmi, E., and Yassin, O., (2013). The Network Value of Products. *Journal of Marketing, 77*(3), 1-14.
32. Goldenberg, J., Oestreicher-Singer, G., & Reichman, S. (2012). The quest for content: How user-generated links can facilitate online exploration. *Journal of Marketing Research*, *49*(4), 452-468.

33. Oestreicher-Singer, G., and Sundararajan, A., (2012). The Visible Hand? Demand Effects of Recommendation Networks in Electronic Markets. *Management Science, 58*(11), 1963-1981.
34. Leem, B., and Chun, H., (2014). An Impact of Online Recommendation Network on Demand. *Expert Systems with Applications*, *41*(4), 1723-1729.
35. Lin, Z., Goh, K. Y., and Heng, C. S., (2017). The Demand Effects of Product Recommendation Networks: An Empirical Analysis of Network Diversity and Stability. *MIS Quarterly*, *41*(2), 397-426.
36. Huang, H. J., Yang, J., & Zheng, B. (2019). Demand effects of product similarity network in e-commerce platform. *Electronic Commerce Research*, 1-31.
37. Carmi, E., Oestreicher-Singer, G., Stettner, U., and Sundararajan, A. Is Oprah Contagious? The Depth of Diffusion of Demand Shocks in a Product Network. MIS Quarterly, 41, 1 (2017), 207-221.
38. Chen, Y. L., Chen, J. M., & Tung, C. W. (2006). A data mining approach for retail knowledge discovery with consideration of the effect of shelf-space adjacency on sales. *Decision Support Systems, 42*(3), 1503-1520.
39. Chen, M. C., & Lin, C. P. (2007). A data mining approach to product assortment and shelf space allocation. *Expert Systems with Applications*, *3*2(4), 976-986.
40. Valenzuela, A., & Raghubir, P. (2009). Position-based beliefs: The center-stage effect. *Journal of Consumer Psychology, 19*(2), 185-196.
41. Valenzuela, A., Raghubir, P., & Mitakakis, C. (2013). Shelf space schemas: Myth or reality?. *Journal of Business Research, 66*(7), 881-888.
42. Ert, E., & Fleischer, A. (2016). Mere position effect in booking hotels online. *Journal of Travel Research, 55*(3), 311-321.
43. Won Jeong, S., Fiore, A. M., Niehm, L. S., & Lorenz, F. O. (2009). The role of experiential value in online shopping: The impacts of product presentation on consumer responses towards an apparel web site. *Internet Research, 19*(1), 105-124.
44. Smith, B., & Linden, G. (2017). Two decades of recommender systems at amazon.com. *IEEE Internet Computing*, 21(3), 12-18.
45. Mudambi, S. M., & Schuff, D. (2010). Research note: What makes a helpful online review? A study of customer reviews on Amazon. com. MIS quarterly, 185-200.