ISSN 1598-0170 (Print)
ISSN 2287-1136 (Online)
http://www.jksii.or.kr

Information-Sharing Patterns of A Directed Social Network: The Case of Imhonet

Danielle Lee[1*]

ABSTRACT

Despite various types of online social networks having different topological and functional characteristics, the kinds of online social networks considered in social recommendations are highly restricted. The pervasiveness of social networks has brought scholarly attention to expanding the scope of social recommendations into more diverse and less explored types of online social networks. As a preliminary attempt, this study examined the information-sharing patterns of a new type of online social network – unilateral (directed) network – and assessed the feasibility of the network as a useful information source. Specifically, this study mainly focused on the presence of shared interests in unilateral networks, because the shared information is the inevitable condition for utilizing the networks as a feasible source of personalized recommendations. As the results, we discovered that user pairs with direct and distant links shared significantly more similar information than the other non-connected pairs. Individual users' social properties were also significantly correlated with the degree of their information similarity with social connections. We also found the substitutability of online social networks for the top cohorts anonymously chosen by the collaborative filtering algorithm.

☞ keyword : Online Social Network; Information Similarity; Social Structure; Homophily

1. INTRODUCTION

A plethora of social media sites leads us to consider various ways to utilize online sociality as users' information management aids. For example, several streams of studies take into account online social networks as a foundation to enhance information searching, knowledge learning [1], security management [2], scholarly communication [3], emergency management [4] and more. Along with the studies, one critical direction of research in the modern big-data era is to utilize users' online sociality as a foundation of personalized recommendation techniques. This research direction is collectively referred to as 'social recommendations.' Social recommendation techniques are generally intended to replace anonymous cohorts of the popular collaborative filtering recommendation technique ('CF' hereafter) with users' self-defined social networks. Various problems with the CF, such as the lack of user involvement, risks and attacks caused by the black-box manner, and cold-start user problems, gave a new impetus to the direction [5].

1 Adaptive Interactions, Co., Seoul, South Korea
* Corresponding author (suleehs@gmail.com)
[Received 06 March 2017, Reviewed 13 June 2017 (R2 05 July 2017), Accepted 12 July 2017]

Although various types of social networks having different topological and functional characteristics are available online [6, 7], the kinds of online sociality considered in social recommendations are highly restricted to a couple of kinds: online friendship and trust-based networks [5]. According to the survey of the field, among existing 40 studies about social recommendations published through April 2016, 46% and 39% of them are based on friendship and trust-based networks, respectively [5]. Therefore, we should find ways to expand the scope of social recommendations into more diverse and less explored types of online social networks. As a preliminary attempt, this study aims to prove the value of a new type of online social network – the unilateral network – as a useful source of personalized recommendations. This study specifically focuses on the information sharing patterns in unilateral networks because the presence of the shared information is the inevitable condition for using the network type as a feasible source of personalized recommendations. This study explores the information-sharing patterns of unilateral networks by addressing the following questions:

- Does the interpersonal similarity of information change with increase/decrease of social distance?
- Are information similarities of individual users' social connections comparable to Top-N anonymous cohorts that are

http://dx.doi.org/10.7472/jksii.2017.18.4.07

〈Table 1〉 Studies about Information Similarity of Online Social Networks

Paper	Input of Similarity Measures	Similarity Measures	Kinds of Social Networks	Domain
Akcora & Carminati [9]	Friendship Graph	Social Structure-based Similarity	Friendship	Online SNS (Facebook)
Anderson, et al. [15]	Users' actions (i.e. editing posts, questioning and answering and evaluating reviews)	Co-occurrence of same user actions	Implicit Online Social Connections	Wikipedia, Stack Overflow & Epinions
Baartarjav, et al. [16]	Personal traits such as age, gender, religion living are, political opinions, etc.	Clustering of personal traits	Friendship	Online SNS (Facebook)
Bhattacharyya, et al. [17]	Keywords in Facebook user profile	Distance between keywords based on the 'Forest model'	Friendship	Online SNS (Facebook)
Bischoff [18]	Music Listening History (loved and banned tracks) and the Related tags	Similarity of Music Listening History and Tags	Friendship	Music (Last.fm)
Brzozowski, et al. [19]	Voting Patterns of Political Resolves	Similarity of Voting Values	Friendships, Ideological Allies and Foes	Online Forum about Political Problems
Hajian & White [20]	Users' various activities (posting, liking and commenting) & following network	The frequency of the same activities between users and the similarity of social structure	Following Network	Online SNS (FriendFeed)
Lee & Brusilovsky [21]	Bookmarks of Scientific Articles and the Social Tags	Similarity of Tags	Group membership & Watching Relations	Social Bookmarking System (Citeulike & Delicious)
Liu, et al. [22]	Ratings of Various Products	Similarity of Ratings	Trust-based Network	Product Review (Epinions.com)
Ma, H. [10]	Movie Ratings and Check-in Records of several Venues	Similarity of Ratings	Friendship and Trust-based Network	Movies & Venues
Modani, et al. [23]	5-Scale Ratings of Movies and Friendship Structure	Item Rating-based Similarity and social structure-based similarity	Friendship	Movie Review (Filmtipset)
Yu, et al. [24]	Users' following network	Social Structure-based Similarity	Following Networks	Micro-blogging Site (Weibo)
Ziegler & Golbeck [25]	Ratings of Books and Movies and Trust Values among users	Similarity of Ratings	Trust-based Network	Book and Movie

mainly used in collaborative filtering?

- Are social properties of individual users significantly related to their interest similarities with online social connections?

The first question is to empirically verify that the associations in unilateral network are rooted in the utility of shared information and to substantiate the homophily theory in the unilateral network. The second question aims to examine the substitutability of unilateral connections for anonymous top cohorts of CF. The last question is to briefly investigate how the information-sharing patterns vary with individual users' social status. One typical social networking system, Imhonet, where users are actively participating in both a unilateral network and book- rating activities, serves as the context of this study.

The remainder of this study is organized as follows. Section 2 introduces several existing studies about the information similarity of online social networks and preeminent social science theories related to users' information-sharing with their online connections. Section 3 provides detailed descriptions of the data set used in this study and shows how to calculate information similarities and various social properties. Section 4 explains the analysis results. The conclusion and implications of this study will be summarized in Section 5.

2. RELATED WORK

2.1 User Interest Similarity of Online Social Networks

Table 1 summarizes existing studies that investigated the information similarities of online social networks. As shown, most of them focused on just limited type of online sociability and neglected the diversity of current online sociability [8]. Moreover, the studies considering the relations between users' social structures and their information similarity are rare. For instance, in a study based on Facebook, Akcora and Carminati (2011) examined network-only graph-based similarities, thus focusing on how much similarity in social structure would predict existing social connections. In Akcora and Carminati's study, the similarity of users' information preferences was not examined [9]. As the closest work to the current study, Ma [10] demonstrated that information similarity varies according to social properties such as the number of co-friends, sub-graph topology, and connected components. Two friends shared more similar interests when they were connected with more edges and higher connection density. In the study, the author focused on reciprocal online friendship [10]. The scholarly attempts to explore how information similarity is relevant to social connections' properties need to be expanded to more diverse types of online social networks in various social media systems. This study is one of the early attempts to focus on a unilateral network.

2.2 Homophily and Object-Centered Sociality

Many social scientists have suggested that we feel attractions to other people who are similar to us; thus, we selectively make social connections with them due to ease of communication, shared knowledge, and other factors that make the interactions comfortable [11]. The principle articulating this social selection made by a person's perceived similarity is referred to as homophily [12, p.416; 13]. Cumulative studies of homophily have demonstrated that the similarity has been traditionally associated with personal status: for instance, age, sex, religion, ethnicity, educational and occupational class, social positions, etc. [14].

However, contemporary society driven by information and knowledge has led to new dimensions of homophily. We feel attracted to people whose information or knowledge is useful (i.e. high utility) and whose information preferences are similar to our own. This is because people tend to perceive high intellectual value in those who can provide information that they seek [13]. We use others as a reference group and compare ourselves with them to obtain information or make a decision as a social comparison process [26].

Singla and Richardson (2008) tested the relationship between instant messenger logs andthe similarity of search queries. The authors were able to demonstrate that search queries of people who exchanged instant messages frequently shared more similar interests than those of random pairs. Moreover, the longer the people talked, the more similar were their queries [27].

In contemporary society, which is highly driven by information and knowledge, the theory of object-centered sociality also insists that knowledge cultures are inter-stitched with current social structures. The main idea of the theory is that information-driven culture weaves the fabric of modern society, and information objects are critical social interaction triggers and anchors of communications [28, p. 40 - 41]. The type of social networks used in our study ― the unilateral network ― is considered to have a high degree of object-centered sociality [21].

3. DATASET AND THE ANALYSIS

3.1 Imhonet Dataset

As the context of this study, one online social networking system, Imhonet, was used. It is one of the most famous SNSs in Russia and a famous personalized recommender system. The system allows users to make social connections called 'friends' and to review and rate a variety of products (such as books, movies, songs, games, TV shows, etc.). This study uses a dataset assembled and provided by Imhonet administrator, containing lists of 216,240 users, 125,657 books, and 8,810,045 ratings on books. Each user has 45.17 book ratings on average (σ= 117.55), and each book has received 70.11 ratings (σ= 797.63) on average. Users assign Likert-type ratings from 1 (very negative) to 9 (very positive) to books.

In spite of the conventional naming practice of ‘friend’, the Imhonet operates the online social networking service as an innovative form. It is a unilateral and un-weighted network. For instance, when user *A* finds another user *B* to be interesting and worthy of connection, he can connect with user *B* as a friend without B’s consent. This means that user *B*’s decision to be connected to user *A* is not necessarily reciprocal. Additionally, unlike trust-based networks, there is no explicitly defined weight about the degree of user *A*’s associations to user B. “Following” on Twitter and Instagram, “watching” on CiteULike, “network” on Delicious and “contacts” on Flickr are other examples of unilateral networks. This type of relationship has emerged as one of the new and less bounded relationships in the booming era of social systems, as social scientists predicted [29]. Once a user makes a connection to other users on Imhonet, he can monitor his connections’ online activities in the system. Thus, we hypothesize that users are likely to copy their connections’ favorite products and subsequently to be influenced by their connections’ interests.

3.2 Information Similarity

This sub-section explains how to compute information similarity among users. For the rest of this paper, we use the following notations: R is the user-item rating matrix, $R=\{R_{ui}\}_{L\times N}$ where L and N denote the number of users and items, respectively. r_{ui} is the user u’s rating on item i. R_i is the set of ratings on item i, and R_u is the item ratings of user u.

Imhonet users express their book preferences via numeric ratings. In order to measure the similarity of users’ preferences based on these numeric ratings, this study considered two types of similarity measures. The first type was the number of co-rated items [30], and the second type was the Pearson Correlation Coefficient (PCC hereafter) [31]. These two types of measures investigated different perspectives of user preferences. The first measure — the number of co-rated items — is to evaluate how many common books two users pay attentions to. The second measure — PCC — is to quantify how differently two users feel about a given book. Among various similarity measures for numeric ratings, these measures were chosen because of the popularity of personalized recommendations [32]. The number of co-rated items literally counts items rated by both given users. The eq. (1) shows how to calculate the PCC.

$$PCC(u,v) = \frac{\sum_{i\in R_u\cap R_v}(r_{ui}-\overline{R_u})(r_{vi}-\overline{R_v})}{\sqrt{\sum_{i\in R_u\cap R_v}(r_{ui}-\overline{R_u})^2}\sqrt{\sum_{i\in R_u\cap R_v}(r_{vi}-\overline{R_v})^2}} \quad \text{eq. (1)}$$

The PCC similarity ranges ±1 inclusive, where a value greater than 0 indicates positively similar preferences between two users, and a value less than 0 indicates negatively similar preferences. In spite of the popularity, however, a major shortcoming of the PCC measure is its absence of a tendency to compute the similarity proportionally to the number of co-rated items between two users [30]. Let’s assume that there are two user pairs. One pair (e.g., user *A* and *B*) rated 150 items in common, and another pair (e.g., user *C* and *D*) rated 5 items in common. When the PCC similarity of both pairs are 0.5, even though the similarities are the same, the former pair’s agreement on ratings is much more significant than the latter pair. However, the PCC of the eq. (1) did not properly reflect the different degrees of agreement. In order to alleviate this shortcoming, we set a threshold λ on the number of co-rated items and scaled down the similarity when the number of co-rated items falls below the threshold [30]. In this study, by following the convention of other studies on personalized recommendations, we set the threshold λ as 50 [33, 34]. The eq. (2) shows the modified PCC with the weight of the number of co-rated items. In other words, the original PCC similarity value was adjusted accordingly, like the eq. (2) when the number of co-rated books was less than 50.

$$PCC_\lambda(u,v) = \min\left(\frac{R_u\cap R_v}{\lambda, 1}\right)\times PCC(u,v) \quad \text{eq. (2)}$$

3.3 Social Property

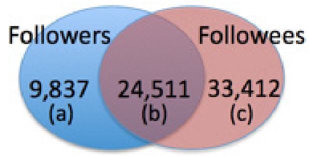
This study calculated users’ social properties on two levels: edge and node. The edge-level properties were used to investigate how users’ information similarities changed according to the increase/decrease in social distances and strength of social associations (the first question in the section 1). The node-level properties were used to examine how users' personal traits affect their information similarity with their links (the third question in the section 1).

As the edge-level properties, the number of co-friends and Dijkstra's shortest path were explored. The number of co-friends between two users is one of the most popular measures to compare the equivalence of these users' social structures [10]. The Dijkstra's shortest path counts how many social connections a user has to traverse to reach another user [28, p. 40 – 41]. When we calculate the shortest paths for a user *A*, the main idea of the algorithm is to put all nodes of a social network in a priority queue, and that queue is initially keyed by the distances of all the other nodes from the starting user *A*, although the distances are initially unknown. Iteratively, the closest neighbors of user *A* are visited, and the queue is updated with the newly calculated distances of the visited neighbors. The update continues until all the other nodes are visited [35].

As node-level properties, in-degree, out-degree, betweenness centrality, and PageRank are explored. While in-degree is the number of incoming links to given users, out-degree is the number of outgoing links from given users. In the Imhonet unilateral network, the in-degree and out-degree indicate the number of the users' followers and the number of followees, respectively. Betweenness centrality measures to what extent a user is located in the center of the network. Specifically, it calculates the frequency of a user to be located in the middle of any other two users' shortest path. The higher a user's centrality is, the more inclined the user is to mediate the information-oriented communications of his neighbors [28, p. 29 ~ 77]. Next, PageRank is the famous method for link analysis of Web pages and, in social network analysis, is widely used to show a user's degree of social influence. Its main idea is that important users should be connected with other important users. The algorithm computes the importance/authoritativeness of a given user in the network by iteratively updating the importance of the user's social connections [35]. Taken together, users with high centrality and high PageRank are more likely to be in the core of the networks and have a good perception about "know-who" so as to acquire the desired information [28, p. 302].

3.4 Experimental Setting

In our Imhonet dataset, approximately one-third of the user population (31.34%, n = 67,760) participated in Imhonet social network and consisted of 234,789 relationships. Depending on the direction of the Imhonet social network, there are two groups of users: followers who initiated the social connections, and followees who are followed by the others. Our Imhonet dataset has 34,348 followers and 57,923 followees (24,511 followees followed other users, as well) as Figure 1 describes. In Figure 1, the numbers of the (a), (b), (c) areas indicate the number of users who belong to each sub-group, respectively. Among the followers, 5,000 followers were randomly chosen as our target users, and the information similarity of the target users were calculated with all of the other users.



(Figure 1) Imhonet Users' Social Associations

4. THE RESULTS

The aim of this study was to prove the feasibility of unilateral social networks as useful information source by assessing the following hypotheses:

- H1. The interpersonal similarity of information is positively correlated with the social distance between two users.
- H2. Information similarities of users' online social connections are comparable to the information similarities of Top-N anonymous cohorts that are mainly used in CF.
- H3. Social properties of individual users have significant correlations with their interest similarities.

We started our analysis by examining global information collection patterns in Imhonet social network, and we tested the hypotheses in subsequent sub-sections.

4.1 Patterns of Information Collection

First, we separated users into two groups: one group consisting of users who participated in Imhonet social network and another group consisting of users outside of the network. Then, we compared the difference in the size of the information collections between the groups. The former group (M = 63.7, σ= 183.4) has rated more than twice number of books than the latter group (M = 30.3, σ= 52.1). According to the independent t-test, the difference was statistically significant (t = -46.6, p <.001). Generally, users who are involved in the Imhonet social network maintained richer information collections than the users outside the network.

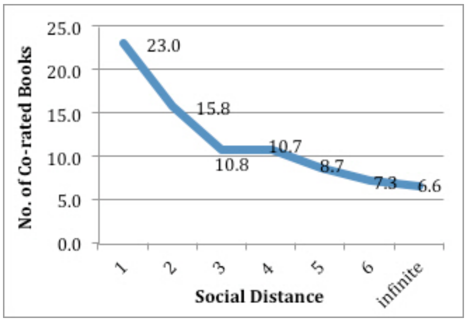
In addition, in order to investigate how the degree of social participations relates to the size of users' rating collections, the correlations between the number of users' book ratings and the number of social associations were computed. We found a small degree of positive correlation with a statistical significance (r = .063, p <.001). Put differently, the more a user follows, the more book ratings he has, even though the tendency was marginal. A similar pattern was observed in the opposite direction of relations: the more book ratings a user had, the more users followed him (r = .263, p <.001). Through these results, we can determine that users' motivation to create social associations in the Imhonet system could be to acquire useful information. Users possessing richer information collections tend to be followed by other users, and users who possessed relatively poorer information collections could not yet draw sufficient attentions from others. If the users' motivation were to foster social relationships with their acquaintances or to enjoy online spaces for fun, the number of social connections would have nothing to do with the size of their information collection. In the following section, we explored users' information-sharing patterns at a detailed level to prove the utility of the Imhonet social network as a source of useful information.

4.2 Social Distance and Information Similarities

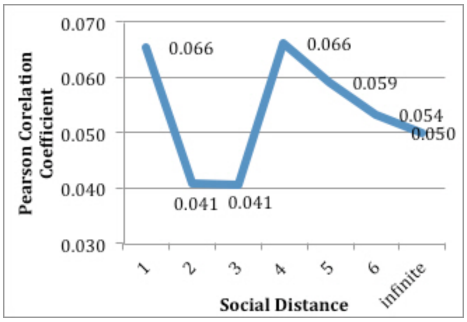
As the first information-sharing pattern, we examined how the number of co-friends is related with the information similarity. The number of co-friends is to describe the degree of equivalence in the social structure [28, p. 29 ~ 77]. The user pairs that have common friends are known to reside in more overlapped social structures than other pairs that do not share any links in common. A correlation test was performed using the number of co-friends and the information similarity. The test revealed that the number of co-friends has positive correlations with the number of co-rated books (r = .035, p <.001) and the PCC similarity (r = .020, p <.001), respectively, albeit marginally. We interpreted the results to mean that user pairs in similar social structures shared more comparable information preferences and became useful information sources to each other.

We also examined whether user pairs that are socially connected have more similar information than non-connected pairs. Here, the pairs of which either user did not participate in Imhonet social network, or of which the distance between two given users is too far away to be computed constitute the non-connected pairs. Moreover, we examined how the change of social distance between two users has altered their information similarity. Users' social distance was calculated by the Dijkstra's shortest path. In the Imhonet social network, the shortest distance (i.e. geodesic) is one (i.e. directly linked connections), and the longest distance (i.e. diameter) is 15. However, only a handful of pairs have a great distance. The user pairs whose distance is equal to or greater than six are just 3.6% of all pairs. For an effective comparison, we thus formulated the distances equal to or beyond six as one type of distance – six or more – by the theory of six degrees of separation [36]. Figures 2 and 3 depict the changes of information similarity depending on the increase/decrease of social distance. In the below figures, the non-connected pairs of which the social distance is unknown and uncountable was labeled as 'infinite'.

Figure 2 displays the linearity of changes in the number of co-rated books by the social distance. The closer two users were, the more common items they shared. A one-way ANOVA test found the statistical significance in the differences of co-rated books made by the social distance (F = 836286.0, p <.001). According to Scheffé's pairwise post-hoc test, the user pairs with direct links shared the largest number of common books, and non-connected pairs shared the smallest number of common books. The post-hoc test also indicated that the average number of co-rated books decreased along with the increase in social distance with statistical significance, except there was no significant difference between two groups of pairs whose distance is 3 and 4.



(Figure 2) The Change of Co-rated Books by Social Distance

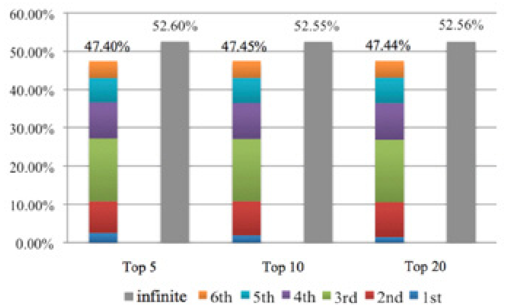


(Figure 3) The Change of PCC by Social Distance

On the other hand, Figure 3 showed that the change in PCC similarity does not necessarily coincide with the increase in distance. To determine if statistical differences existed among user groups with different social distances on the information similarity, a one-way ANOVA test was conducted. The results showed the statistical significance in the differences in PCC similarity by the social distance (F = 322402.1, p <.001). The post-hoc pairwise test indicated that the PCC similarity of the directly connected pairs have significantly higher similarity than non-connected pairs. However, the same test showed that PCC similarity does not vary in proportion to the increase in social distance. Taken together, these results suggested that users did pay attention to what their online connections read, but did not naturally concur with the opinions of their connections. Therefore, the H1 hypothesis was partially accepted only for the number of co-rated books.

4.3 Social Cohorts vs. Anonymous Cohorts

As explained in the introduction, the ultimate goal of this study is to investigate the feasibility of the Imhonet network as a useful foundation for the most popular recommendation technique, CF [37]. Therefore, we tested how many anonymous top cohorts are automatically chosen by CF approach are socially connected. Specifically, we chose the top N (N = {5, 10, 20}) cohorts of each target user by their PCC similarity, since PCC similarity is the most often used similarity measure in memory-based CF [37]. Then we counted the number of connections with direct or distant links in the top N anonymous cohorts. Figure 4 depicts the percentage of social connections found in the top N cohorts.



(Figure 4) Percentage of Social Connections in Top N Cohorts

In Figure 4, there are two bars at each top N; the left bars represent the percentage of target users' social cohorts, and the right bars represent the percentage of target users' anonymous cohorts without any social associations. Compared with the cohorts in top 20 rank, the cohorts in the top 5 rank share more similar tastes with our target users and serve as more important sources of information in personalized recommendation. The results depicted in Figure 4 showed that target users' social connections (i.e. social cohorts) were consistently comprised of the half of top anonymous cohorts, regardless of the top N ranks. Moreover, 60% of the social cohorts were closely associated with our target users within the social distance of three.

In order to see the difference in PCC similarity between the social cohorts and anonymous cohorts, the users chosen as the top N cohorts were classified into two groups: one group who are associated with our target users within the social distance of three, and another group of the remaining cohorts. The social distance of the latter group with our target users is more than 4 hops or unknown (because of the anonymity). The t-tests found insignificant differences in PCC similarities between the two groups, regardless of the top N ranks (t = 0.56, p = .46 for top 20; t = 0.08, p = .73 for top 10; t = 0.27, p = .60 for top 5).

Therefore, these results provide positive evidence of the feasibility of users' unilateral social network as a useful information source and led us to accept our second hypothesis. Users' social connections have comparable degrees of information similarity to the anonymous like-minded top cohorts.

4.4 Users' Social Properties and Information Similarity

The last step was to examine whether individual users' social properties have altered their information similarity with their social partners or not. As explained in Section 3.3, four kinds of properties – in-degree, out-degree, betweenness centrality, and PageRank – were examined for each target user. Partial correlational analyses were executed to examine the relationship between the information similarity and each of the social properties while controlling for the number of users' book ratings. Section 4.1 reported significantly positive correlations of in-degree and out-degree with the number of book ratings. Hence, the partial correlation analyses indicate to what extent the four kinds of social properties are linearly related with the number of co-rated books and PCC similarity when the effects of the number of book ratings were removed. Table 2 shows these results.

〈Table 2〉 The Correlations between Information Similarity and Individual Users' Social Properties

	No. of Co-rated Books	PCC
In-degree	**.150***	**.118***
Out-degree	**.043***	.009
Betweenness	**.065***	**.059***
PageRank	**.124***	**.114***

Note. Coefficients with asterisks and printed in bold are statistically significant (p <.001)

The in-degree (i.e., how many users a user follows) was consistently associated with the number of co-rated books and PCC similarity, whereas we failed to find any meaningful correlations between out-degree and the information similarity. The result indicates that participants in the Imhonet network actually are watching the activities of their connections. On the other hand, since the direction of the social links is one-sided, the insignificant correlation of out-degree with information similarity explained that the users being followed by others did not pay attention to their followers. The significantly positive correlations of betweenness centrality with the number of common books and PCC similarity were also well observed. That is, the users in the core of their local networks (i.e. high betweenness centrality) tend to have similar information as their social connections. The PageRank was also positively correlated with the number of co-rated books and PCC similarity with a statistical significance. Through the result, it seems that users with high social influence (i.e. high PageRank) actually affected their followers' collections. These results led us to accept the third hypothesis, except for the out-degree property. In summary, users who are following many other users bear a high resemblance in their book rating collection to their connections. Moreover, it seems that users who are actively engaged in online social networks find ways to connect with others who possess useful and interesting information more easily.

5. CONCLUSION

This study argued that users' unilateral social networks might be a valuable source of useful information, especially for personalized recommendations where information similarity among users is the core of the technology. To prove the feasibility of this argument, we examined how users' information similarities are correlated with their social structures.

In this study based on the Imhonet unilateral social network, we discovered that user pairs with direct and distant links shared significantly more similar information than the other non-connected pairs. We also found that users tend to refer to what their followers' read but do not automatically agree with the followers' opinions of those same books. The result corresponds to the conclusion of the study by Brzozowski and colleagues [19] in which, based on an online political forum to vote on various controversial political topics, users were influenced by their friends in the choice of resolves to vote on.

However, they did not necessarily agree with how their friends voted on those resolves [19].

In addition, we found the substitutability of unilateral networks for top anonymous cohorts of the collaborative filtering algorithm. Even though the top N cohorts were anonymously chosen in a black-box manner, half of the cohorts were socially connected with our target users. Moreover, 60% of the social cohorts were closely related with our target users with regard to social distance.

Finally, individual users' social properties were significantly correlated with the degree of their information similarity. The pairs who are connected with mutual links and more common connections have greater similarity than the other pairs with no mutual links and fewer co-connections. This paper also showed that users located in the center of their local networks and possessing high social influence are more likely to share similar information with their social connections. In conclusion, we observed the homophilous interactions oriented by users' information preferences in the Imhonet unilateral social network.

Overall, this paper's contributions are twofold: (1) this paper presented the similarity in information preferences in a new type of online sociality ― unilateral network ― and demonstrated the feasibility of sociality as a useful information source; (2) this paper also discovered valuable social properties that can be used in personalized recommendations. In ongoing research in social recommendations, the attempts to fuse information similarity with users' social properties are rare [5].

In terms of the potential future direction of this study, we plan to investigate temporal changes in users' information sharing patterns. In this study, we neglected one important piece of information: timestamps of users' book ratings. We will explore not only the homophily of the Imhonet social network, but also the presence of social influences and detailed patterns. Another future research direction is to find ways to generate good quality recommendations using users' online social networks. While this study determined the utility of online connections as a useful information source, personalized recommendations require sophisticated approaches to combine users' information preferences and their social properties. Finally, in order to examine the generalizability of our results, we will explore unilateral social networks of other social media systems in different domains.

Reference

[1] R. Bapna, A. Umyarov, "Do Your Online Friends Make You Pay? A Randomized Field Experiment on Peer Influence in Online Social Networks" Management Science, vol. 61, issue 8, pp. 1902-1920, 2015.

[2] Y. Zheng, B. Wang, W. Lou, & Y. T. Hou, "Privacy-Preserving Link Prediction in Decentralized Online Social Networks," Proceedings of the 20th European Symposium on Research in Computer Security, pp. 61-80,2015.

[3] M. Thelwall, K. Kousha, "ResearchGate: Disseminating, communicating, and measuring Scholarship?," Journal of the Association for Information Science and Technology, vol. 66, issue 5, pp. 876-889, 2015.

[4] M. Imran, C. Castillo, F. Diaz, S. Vieweg, "Processing Social Media Messages in Mass Emergency: A Survey" ACM Computing Surveys, vol. 47, issue 4, pp. 1-38, 2015.

[5] D. Lee, P. Brusilovsky, "Social Link-based Recommendations: A Review," In P. Brusilovsky & D. He (Eds.), Social Information Access. Heidelberg: Springer, Forthcoming.

[6] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, A. Arenas, "Diffusion Dynamics on Multiplex Networks," Physical Review Letters, vol. 110, issue 2, 028701, 2013. http://dx.doi.org/10.1287/mnsc.2014.2081

[7] J. Gao, D. Li, S. Havlin, "From a single network to a network of networks," National Science Review, vol. 1, issue 3, pp. 346-356, 2014. http://dx.doi.org/10.1145/2579993

[8] T. Tiropanis, W. Hall, J. Crowcroft, N. Contractor, L. Tassiulas, "Network science, web science, and internet science," Communications of the ACM, vol. 58, issue 8, pp. 76-82, 2014. http://dx.doi.org/10.1007/s13278-010-0006-4

[9] C. G. Akcora, B. Carminati, E. Ferrari, "Network and profile based measures for user similarities on social networks," Proceedings of IEEE International Conference on Information Reuse and Integration (IRI), pp. 292-298, 2011.

[10] H. Ma, "On measuring social friend interest similarities in recommender systems," Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval, pp. 465-474, 2014
[11] M. Yavaş, G. Yücel, "Impact of Homophily on Diffusion Dynamics Over Social Networks," Social Science Computer Review, vol. 32, issue 3, pp. 354-372, 2014. doi:
http://dx.doi.org/10.1016/j.socscimed.2014.05.019
[12] M. McPherson, S. Lovin, J. Cook, "BIRDS OF A FEATHER: Homophily in Social Networks" Annual Review of Sociology, vol. 27, pp. 415-445, 2001. http://dx.doi.org/10.1016/j.cosrev.2016.05.002
[13] A. Anderson, D. Huttenlocher, J. Kleinberg, J. Leskovec, M. Tiwari, "Global Diffusion via Cascading Invitations: Structure, Growth, and Homophily," Proceedings of the 24th International Conference on World Wide Web, pp. 66-76, 2015.
http://dx.doi.org/10.1093/nsr/nwu020
[14] V.-J. Ilmarinen, J.-E. Lönnqvist, S. Paunonen, "Similarity-attraction effects in friendship formation: Honest platoon-mates prefer each other but dishonest do not," Personality and Individual Differences, vol. 92, pp. 153-158, 2016.
[15] A. Anderson, D. Huttenlocher, J. Kleinberg, J. Leskovec, "Effects of user similarity in social media," Proceedings of the fifth ACM international conference on Web search and data mining, pp. 703-712, 2012. http://dx.doi.org/10.1145/2503792.2503797
[16] E.-A. Baatarjav, S. Phithakkitnukoon, R. Dantu, "Group Recommendation System for Facebook," Proceedings of the OTM Confederated International Workshops and Posters on On the Move to Meaningful Internet Systems: 2008 Workshops: ADI, AWeSoMe, COMBEK, EI2N, IWSSA, MONET, OnToContent + QSI, ORM, PerSys, RDDS, SEMELS, and SWWS, pp. 211 - 219, 2008.
[17] P. Bhattacharyya, A. Garg, S. Wu, "Analysis of user keyword similarity in online social networks," Social Network Analysis and Mining, vol. 1, issue 3, pp. 143-158. 2011.
http://dx.doi.org/10.1016/j.paid.2015.12.040
[18] K. Bischoff, "We love rock 'n' roll: analyzing and predicting friendship links in Last.fm," Proceedings of the 3rd Annual ACM Web Science Conference, pp. 47-56, 2012. http://dx.doi.org/10.1145/2771588
[19] M. J. Brzozowski, T. Hogg, G. Szabo, "Friends and foes: ideological social networking," Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, pp. 817-820, 2008.
[20] B. Hajian, T. White, "Modelling Influence in a Social Network: Metrics and Evaluation," Proceedings of IEEE third international conference on Privacy, security, risk and trust (passat) and 2011 ieee third international conference on social computing (socialcom), pp. 497-500, 2011.
[21] D. Lee, P. Brusilovsky, "Social Networks and Interest Similarity: The Case of CiteULike," Proceedings of the 21th ACM conference on Hypertext and hypermedia, pp. 151-156, 2010.
http://dx.doi.org/10.1016/j.knosys.2013.11.06
[22] 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, vol. 56, pp. 156-166, 2014.
[23] N. Modani, R. Gupta, S. Nagar, S. Shannigrahi, S. Goyal, K. Dey, "Like-Minded Communities: Bringing the Familiarity and Similarity together," the Proceedings of Web Information Systems Engineering, pp. 899-919, 2012. http://dx.doi.org/citeulike-article-id:201696
[24] Y. Yu, L. Mo, J. Zhou, "Social Friend Interest Similarity in Microblog and its Implication," International Journal of Control and Automation, vol. 8, issue 11, pp. 21-32, 2015.
[25] C.-N. Ziegler, J. Golbeck, "Investigating interactions of trust and interest similarity," Decision Support Systems, vol. 43, issue 2, pp. 460-475, 2007.
http://dx.doi.org/10.1007/978-3-642-35063-4_28
[26] D. Centola, A. van de Rijt, "Choosing your network: Social preferences in an online health community," Social Science & Medicine, vol. 125, pp. 19-31, 2015.
[27] P. Singla, M. Richardson, "Yes, there is a correlation: - from social networks to personal behavior on the web," the Proceeding of the 17th international conference on World Wide Web, pp. 655-664, 2008.
http://dx.doi.org/10.1111/coin.12041

[28] P. R. Monge, N. S. Contractor, "Theories of communication networks," Oxford University Press, New York, USA, 2003. doi: http://dx.doi.org/10.1016/j.csi.2016.10.014
[29] A. Guille, H. Hacid, C. Favre, D. A. Zighed, "Information diffusion in online social networks: a survey," ACM SIGMOD Record, vol. 42, issue 2, pp. 17-28, 2013.
[30] M. B. Menhaj, S. Jamalzehi, "Scalable user similarity estimation based on fuzzy proximity for enhancing accuracy of collaborative filtering recommendation "Proceedings of the 4th International Conference on Control, Instrumentation, and Automation (ICCIA), pp. 220-225, 2016. http://dx.doi.org/10.1002/asi.23236
[31] C. C. Aggarwal, "Neighborhood-Based Collaborative Filtering," Recommender Systems: The Textbook, pp. 29-70, Springer International Publishing, New York, USA, 2016. http://dx.doi.org/10.1145/2699416
[32] M. Elahi, F. Ricci, N. Rubens, "A survey of active learning in collaborative filtering recommender systems," Computer Science Review, vol. 20, pp. 29-50, 2016.
[33] Bellogín, A., Castells, P., & Cantador, I. "Neighbor Selection and Weighting in User-Based Collaborative Filtering: A Performance Prediction Approach," ACM Transactions on the Web, vol. 8, issue 2, pp. 1-30, 2014. doi:http://dx.doi.org/10.1016/j.comcom.2013.06.009
[34] N. Polatidis, C. K. Georgiadis, "A dynamic multi-level collaborative filtering method for improved recommendations," Computer Standards & Interfaces, vol. 51, pp. 14-21, 2017. http://dx.doi.org/10.1177/0894439313512464
[35] R. Yan, Y. Song, C.-T. Li, M. Zhang, X. Hu, "Opportunities or risks to reduce labor in crowdsourcing translation? characterizing cost versus quality via a pagerank-HITS hybrid model," Proceedings of the 24th International Conference on Artificial Intelligence, pp. 1025-1032, 2015.
[36] S. Nepal, S. K. Bista, C. Paris, "Behavior-Based Propagation of Trust in Social Networks with Restricted and Anonymous Participation," Computational Intelligence, vol. 31, issue 4, pp. 642-668, 2015.
[37] X. Yang, Y. Guo, Y. Liu, H. Steck, "A survey of collaborative filtering based social recommender systems," Computer Communications, vol. 41, issue 0, pp. 1-10, 2014. doi: http://dx.doi.org/10.1016/j.dss.2006.11.003

◐ 저 자 소 개 ◑

이다니엘(Danielle Lee)
1996년 상명대학교 국어국문학과(문학사)
2001년 시라큐스대학교 정보대학원 정보과학과(공학석사)
2013년 피츠버그대학교 정보대학원 정보과학과(이학박사)
2013년~2016년 워싱턴대학교 시스템 & 소프트웨어 시스템학과 조교수
2016년~현재 Adaptive Interactions, Co. 연구원
관심분야 : 개인 추천 서비스, Online Social Dynamics, Social Tagging Systems, 계량서지학
E-mail : suleehs@gmail.com