

An investigation of Social Support Exchange and Communication Patterns among Chinese on Online Discussion Forum

Xi Wang
Central University of Finance
and Economics
xiwang@cufe.edu.cn

Xing Tong
George Mason University
xtong@masonlive.gmu.edu

Yushan Zhu
Central University of Finance
and Economics
1096048233@qq.com

Tianyi Tan
Central University of Finance
and Economics
antonio21tan@163.com

Bowen Zheng
Tsinghua University
bowenzheng@mail.tsinghua.edu.cn

Abstract

Online Health Communities (OHCs), frequently adopted as online discussion forums for online users to communicate on health issues, have been used worldwide. By analyzing a representative breast-cancer-related OHC from mainland China—Baidu Discussion Forum, this study attempts to investigate social support exchange and communication patterns through user-generated content by data mining approaches. According to the outcomes, emotional support seeking and providing presents itself to be a more critical theme among Chinese users than other types of social support. In addition, almost half of the users on Baidu Discussion Forum have simple patterns of involvement, and a fairly small proportion of highly active Chinese users are quite influential in shaping the connections of the social support network. Meanwhile, the off-topic discussions which are not directly on health concerns are not frequently touched by Chinese people. This may impact the longevity of both users and threads, and undermine the foundation of OHCs in the long term. The findings have practical implications for researchers and health practitioners targeting on the Chinese population.

1. Introduction

The speedy development of social media has fundamentally changed the means of seeking and acquiring knowledge on health issues. Online Health Communities (OHCs) are becoming prevalent channels for the public to communicate health-related concerns. In China, eighty percent of Internet users have reported experiences of online healthcare services [1]. As a part of digital healthcare, OHCs integrate scientific achievements from multiple disciplines.

The OHCs are able to generate great benefits for individuals, the public, and the market. First, OHCs are beneficial for the involved users to obtain therapeutic information, understand their medical circumstances, heighten their emotional comfort, advance personal empowerment, and acquire resources offline [2]–[5]. Although physicians play a crucial role in providing medical information, OHCs are also vital resources for newcomers to refer to other patient users' personal experiences [6]. Second, OHCs are valuable in informing the public during health crisis and establishing interventions. Timely and appropriate analyses on OHC user-generated content may assist in early detection of large-scale public health events and developing appropriate countermeasures. Third, a well-designed OHC can be beneficial for stakeholders, including website operators, medical marketers, and other groups seeking monetary gain from commerce with the OHC communities. For example, operators have opportunities to configure their websites to attract newcomers and maintain senior members to sustain OHCs [7], and insurance representatives are offered a platform of online advertisements to maximize the profits of their medical products [8].

Users' continuous participation in an OHC can be helpful to both the users and the sustainability of the community [9]. While advancements in modern medical science can be indiscriminately applied in different countries to cure patients worldwide, it is not clear whether the research findings of OHCs based on the Western culture can be generalized to Chinese context. Therefore, understanding how people from Chinese cultural backgrounds use OHCs is prerequisite to designing user-friendly OHCs to facilitate computer-mediated-communication among patients and to help them to benefit from the development of information technology.

This empirical study combined text mining methods and social network analysis to identify user behavior patterns through analyzing user-generated content in a representative Chinese OHC, Baidu Discussion Forum. The outcomes of this study are able to help decision makers to better understand users' behavior patterns and maximize the potential value of the community.

2. Related work

Social support refers to an exchange of resources between individuals in order to enhance the well-being of the receiver [10]. Community psychologists have identified different types of social support, which mainly including informational support, emotional support and companionship [9]. Informational support is the transmission of information, suggestion, or guidance to the community users. The content is usually related to advice, referrals, education, and personal experience on disease or health problems. Emotional support mainly contains the expression of understanding, encouragement, empathy, etc. Companionship, also known as network support, consists of chatting, humor, and topics of offline activities such as birthday wishes, holiday plans that are not necessarily related to one's health problems. Thus, the companionship is sometimes referred to as off-topic discussions.

Users' demands for social support are varied depending on the theme of OHCs and users' features. For example, compared to emphasizing informational support in a diabetes OHC, companionship is more frequently mentioned in an Amyotrophic Lateral Sclerosis OHC, and emotional support is more intensely required by users from the breast cancer sites [11]–[14].

Not only type of disease and status of disease may impact users' demand of social support type, but also users' cultural backgrounds. First, culture may shape the individuals' conceptualization and perception of disease, especially the chronic disease with very complicated and diversified causes, like cancer. In Chinese culture, cancer is in many occasions, regarded as a mysterious and clueless disease that nobody knows for certain how it progresses [15]. Individuals are frequently reported unwilling to be told of their risk of experiencing its threat or refuse to admit their chances of getting cancer. Such accumulated non-scientific perceptions on cancer among Chinese are likely to trigger their illusion of invulnerability [16] and make them resistant or disregarding of the messages on cancer. Information providers have expressed caution and hesitancy to address the topic of cancer in carefully avoided using the word "cancer"

during conversations because they strongly believe the simple act of mentioning the name of the disease has been powerful enough to increase their likelihood of getting cancer [15]. The stigma attached to cancer also impose their scary of the disease, and their potential necessity of emotional support in OHC communication with other patients.

Second, people from different cultures have distinct rules and customs that shape their patterns of self-disclosure [17] and privacy. In communication studies, the individualism-collectivism dimension of culture is often adopted to explain the differences in communication style from different cultures. Chinese people, as members from a collectivistic culture, are typically represented as "more formal and cautious in expressing themselves and communicate less openly and freely" [18]. In addition, the collectivistic culture is constantly resistant in sharing internal feelings because its members think highly of privacy [19]. The social norms of such a culture often prohibit patients from narrating their own feelings and experiences thus change the users' interaction patterns in OHCs. For example, studies found that Chinese users treat OHCs as a source to approach health information, rather than a sharing platform to interact with other users [20]. Comparatively, people from a Western background tend to consider seeking social support is a way of pressure relief [19]. Such differences may shape information-seeking and providing behaviors from Western and Eastern background in quite dissimilar ways.

To sum up, OHC users from different cultures may use OHC differently in complex ways. Culture not only influences individuals' perceptions of disease, but also has an effect on what individuals may choose to disclose or withhold, to whom to disclose, and their interaction patterns with others. Prior studies about user behaviors in Chinese OHC mainly focus on motivation [21][22], to investigate why a user is involved in an OHC. The research methodologies depended more on studies such as surveys or interviews. Meanwhile, although the social network has been established in analyzing OHC users as well, most of the work mainly from the homogenous perspective, without differentiating multiple communication channels among users. However, how the macro-contextual information and cultural background work together to shape Chinese OHC users' participation and their communication patterns remain to be discussed.

3. Methodology

The study was conducted through crawling users' posts from Baidu Discussion Forum of "Breast", a peer-to-peer discussion forum with large content volume accessible to the public. There are three reasons to choose Baidu Discussion Forum as the research target. First, its format well represents a typical Chinese OHC—one user initiates a thread and others reply comments which helps to avoid confounding factors during comparison. Second, majority of its users are breast cancer patients, caregivers, and survivors. As chronic disease sufferers, users usually need continuous support during their long-term battles with the health issue. Third, due to the specific cancer perceptions among Chinese people, the topic on breast cancer provides sufficient room for the researchers to examine how cultural beliefs may impact the Chinese users' behaviors in an OHC. The dataset contains 186,291 posts contributed by 36,184 users from March 2007 to March 2018.

Different from traditional communication studies of identifying social support manually, this study introduced text mining approaches to uncover social support automatically. In order to explore social support in the posts, we trained classifiers based on machine learning techniques. Specifically, we built classification models to determine what categories of social support each post contains. Consistent with our previous study [9], [20] and aforementioned definitions of social support, we first determined whether a post was seeking informational support (SIS), providing informational support (PIS), seeking emotional support (SES), providing emotional support (PES), or simply inquiring about companionship (COM). The category labels are not mutually exclusive.

Table 1 shows the examples of each category. With annotating 2,000 randomly selected posts by five trained coders, we acquired ground-truth data as the training set. Table 2 shows the distribution of social support in the annotated training set.

Table 1. Example posts for different types of social support

| Category | Posting Content Translations |
|----------|--|
| COM | <i>I have thin hair and it grows slowly.</i> |
| SIS | <i>I was dragonized with breast hyperplasia yesterday. Any ideas for the treatment? I am so anxious...</i> |
| SES | <i>After getting the puncture report this afternoon, I was totally breakdown. I shed floods of tears, feeling helpless and devastated. When faced with disease, humans are really tiny and powerless. Thinking of the elderly parents and young children, I feel extremely</i> |

| | |
|---------------------|---|
| | <i>desperate.</i> |
| PIS | <i>I was diagnosed with breast nodules calcification through B-ultrasound test. The doctor recommended an NMR, and the results turned out it was benign.</i> |
| PES | <i>Come on, be confident, you will be fine.</i> |
| Multiple Categories | <i>Ms. Bai, I have been constipated for quite a while, causing acne on my face. I looked so bad, and diagnosed with breast nodules. I have no idea if I can take capsules to lubricate the bowel. I have too many health problems. I am wondering whether endocrine disorder is caused by long-term constipation, and results in the breast nodules. Now it is getting even worse. I need a puncture, sigh, so upset.</i> |

Due to the different writing styles and linguistic preferences of OHC users, we needed to capture the characteristics of each post and extract various types of features for training classifier. In our previous study [9], we included basic features, lexical features, sentiment features, and topic features, named as original features (Feature-set 1) in this study and received decent performance on the training of classifier models. Considering the increasing advantages of deep learning metrics in text mining research field, we applied new attempts on the crawled dataset. Specifically, we included two more standard text feature sets derived from the text of a focal post in addition to the original features. One was the bag-of-words (BOW) feature set, the most common features in text mining problems, in which we extracted unigrams after removing stop-words and stemming. The other was the Word2Vec (W2V) feature set, a word embedding technique that generates vector representations for posts. Newly added features were combined with the original features as Feature-set 2 and Feature-set 3 in order to increase the abstraction level of the focal post.

Table 2. Distribution of social support in the training set

| | Posts | % |
|-----|-------|-------|
| COM | 220 | 11.00 |
| PES | 1,019 | 50.95 |
| PIS | 612 | 30.60 |
| SES | 161 | 8.05 |
| SIS | 125 | 6.25 |

We evaluated the feature-sets in identifying social support by leveraging seven algorithms, including Logistic Regression, Support Vector Machine, Naïve Bayes, K-Nearest Neighbors, Decision Tree, Random

Forest, and AdaBoost with the Decision Tree as the weak learner. In terms of addressing the unbalanced type of social support, such as SES, we leveraged under-sampling to decrease the impact of microvolume of the positive instances during the training process. We used weighted F1-score and weighted AUC to evaluate the performance of classification models with 10-fold cross-validation. F1-score measures the mean of precision and recall. AUC (the area under the ROC curve) measures the probability that a positive sample is ranked higher than a negative sample and provides robust measurements of classification performance. After trying different combinations of algorithms and feature sets, the best performance classification model for each type of social support has been selected in Table 3.

OHC users publish posts to provide or receive social support, which formats a social network. Similar to the social networks in our daily lives, an individual's daily network can be connected by different social relations, such social support networks are multi-relational as well, which means individuals should be able to be connected by multiple types of social support. For example, one user is connected to another by providing information support through answering medical question, while two other users can be connected by a COM post describing a favorite recipe. Since the social support behavior is multi-relational composing with five category-connections, the topological analysis of a social support network would present us a better picture of communication patterns of users in different types of social support.

Table 3. Evaluations of classification model

| Category | Algorithm | Features | F1-Score |
|----------|---------------------|---------------|----------|
| | | | AUC |
| COM | Logistic Regression | Feature-set 3 | 0.871 |
| | | | 0.868 |
| PIS | Logistic Regression | Feature-set 3 | 0.824 |
| | | | 0.886 |
| PES | Adaboost | Feature-set 2 | 0.778 |
| | | | 0.851 |
| SIS | Logistic Regression | Feature-set 3 | 0.872 |
| | | | 0.856 |
| SES | Logistic Regression | Feature-set 3 | 0.891 |
| | | | 0.894 |

We constructed a social support behavior network of users from the OHC with one node representing an individual user. An undirected link between two users A and B means that A or B published a comment within the thread both of them participated in. For example, user A initiated a thread to seek informational support and emotional support, user B posted a PIS

comment followed by user C, who provided emotional support. User A and user B are connected by SIS, SES and PIS, user A and user C are connected by SIS, SES, and PES, while user B and user C are connected by PIS and PES. We emphasize “reading to receive”, therefore B and C are assumed receiving social support by reading each other's post. Meanwhile, the number of the same type of links is summarized as the weight of the tie. For example, if user B posted another PIS after user C's comment, the weight of the PIS link between user B and user C would be 2.

With aggregating the five subnetworks of social support behaviors, the aggregated network emerges. The links of the aggregated network represent some types of social support behavior between two users in the community, and the degree of a node is the total number of its neighbors in the network. By incorporating different types of social support, we can better capture how users co-participated in the OHC.

To identify the users' communication patterns, we conducted a social network analysis. First, we conducted a topological analysis at the network level to illustrate the characteristics of the aggregated network and its subnetworks. We calculated the number of nodes, the number of edges, density (defined as the number of actual ties divided by the number of possible ties) and summarized the degree of nodes and weight of edges of the subnetworks. Meanwhile, we measured structural similarities by calculating Jaccard coefficients of the connections between two users in different subnetworks. A high Jaccard coefficient between two subnetworks signaled the high correlation and similarity of communication patterns of two types of social support. Second, we calculated the correlations of nodes' degree and edges' weight at the individual level. Specifically, we removed the nodes or edges only appearing in one subnetwork during the intersected calculation—we correlated each node's degree or edge's weight in one subnetwork to the same node's degree or edge's weight in another subnetworks. A strong correlation between two subnetworks suggested that individuals with more activities in one subnetwork tended to have similar patterns in another.

4. Results

4.1. Social Support Exchange

After selecting the best performance classification algorithms and feature sets, we leveraged the models to automatically assign social support labels to the entire dataset. Table 4 shows the distribution of social support across the entire dataset. Note that some posts might not be labeled by any type of social support, resulting in the sum of the types of social support less than 100.

Table 4. Distribution of social support in the entire dataset

| | Posts | % |
|-----|--------|-------|
| COM | 16,477 | 8.84 |
| PES | 81,525 | 43.76 |
| PIS | 47,251 | 25.36 |
| SES | 8,589 | 4.61 |
| SIS | 2,038 | 1.09 |

The distribution of social support shows that the average longevity of a thread in Baidu Discussion Forum on Breast was 6.83 days, with 4.54 users involved in one thread on average. Figure 1 plots the boxplot distribution of social support on the basis of threads in the community. Since 23.49% of threads sustained only 1 day, and 93.74% lasted less than 15 days, we binned the longevity of threads into five categories to observe the variety of social support components over threads. For example, the third set of boxplots in Figure 1 reflects all the threads lasted from 6 to 10 days, containing various types of social support in their comments. Among all types of social support, PES was published on an average of 52.5% and PIS was published on an average of 30.2%. The trend of social support types does not change substantially over time.

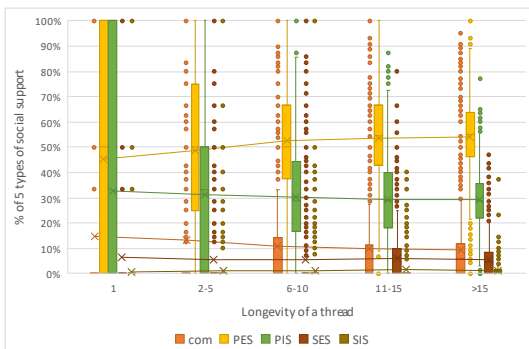


Figure 1. Distribution of social support on threads

At the individual level, the time span of a user is calculated as the days from users' first post to the last one within our dataset. Longer time span indicates higher loyalty of a user to the OHC. We binned users into 5 groups by their time span—less than 1 week (1-7 days), less than 1 month (8-30 days), less than 3 months (31-90 days), less than 1 year (91-365 days), and more than a year (>365 days). Figure 2 shows the relationships between types of social support users published and time span of the users in the

communities. On Baidu Discussion Forum, the posting behavior does not change much between temporary users and long-term users.

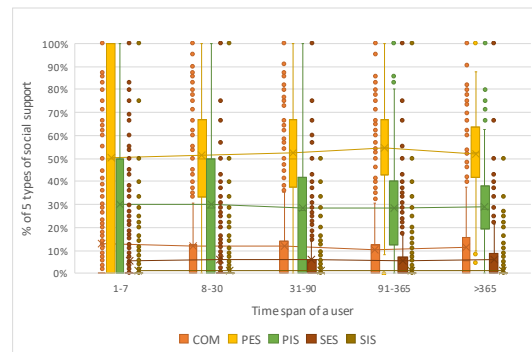


Figure 2. Distribution of users' posts

4.2. Social Support Network

To conduct the topological analysis, we built a multi-relational network of social support. Figure 3 shows the distributions of degrees of the nodes and weights of the edges in the aggregated network and five social support subnetworks. The log-log plots of degrees show that the minority of the users are well-connected to others, while the majority linked with only a few peers. Similar patterns exist in the distribution of weight of edges. Specifically, the minority of user pairs are connected by huge amounts of overlapped social support links, while the majority of the user pairs are connected dispersedly. In other words, a few user pairs are highly interactive and co-participate a lot in threads, while most of the users are connected to only a few of discussion. Intuitively, there are sudden drops that occur in the plot. Taking the degree distribution of the aggregated network as an example, the sudden drop indicates lots of users having the same and large value of degree, which is caused by a few threads having surprisingly long longevity. This is a very interesting finding, and we will discuss further on the influential users in discussion.

The plots of distribution of degree and weight in five different social support networks suggest that each subnetwork features its own topological characteristics. On one hand, the degree distribution of subnetworks does not follow the linear trend but conforms to the general pattern of scale-free network—nodes with higher degrees appear less frequently. On the other hand, the flat distribution for the low degree values and sudden drop for high degree values suggest the imbalanced length and quality of social support in the community. Meanwhile, the trend of weight's distribution, especially SIS and PES, suggests the different communication patterns of users' behavior in two different types of social support. Specifically, the

weight of edges represents the repetition of social support occurring between two users. The scattered distribution of weights confirmed the relatively simple social support exchange in the community, which actually has been shown in Figure 1 and 2.

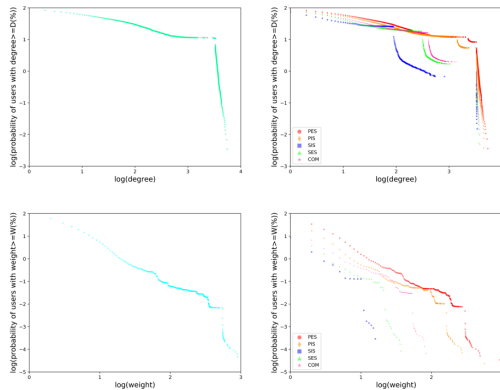


Figure 3. Distributions of degree and weights of the aggregated network and social support subnetworks

Table 5 summarizes the descriptive statistics of the aggregated network and each of the five subnetworks. In the OHC, among all the social support subnetworks, the PES has the most nodes, edges, and highest density, followed by the PIS subnetwork.

Table 5. Descriptive statistics of aggregated network and social support subnetworks

| Attributes | # of nodes | # of edges | Density |
|------------|------------|------------|---------|
| Aggregated | 28,968 | 6,501,355 | 0.0155 |
| PIS | 26,434 | 4,294,816 | 0.0123 |
| PES | 27,645 | 5,826,551 | 0.0152 |
| SIS | 13,283 | 344,964 | 0.0040 |
| SES | 18,559 | 1,154,277 | 0.0067 |
| COM | 20,781 | 1,505,594 | 0.0070 |

Two nodes might be connected in more than one subnetwork in the community. Hence, besides summarizing the subnetwork independently, we also computed the overlap of the subnetworks (Table 6). 12.2% user pairs are connected in four. Almost half of the user pairs are connected in only two subnetworks.

To further measure the similarities between the five subnetworks, we correlated the nodes and edges by computing Jaccard coefficients and degree/weight correlations at the individual level. The Jaccard coefficients [23] are shown in Table 7. The high coefficient of the nodes between PIS and PES subnetworks suggests the high users' similarity of two subnetworks. In other words, the users' involvement related to PIS and PES are highly intersected. The

values 0.477 and 0.493 of the SIS subnetwork towards both PES and PIS reflect that the users' SIS is relatively independent of the provision of social support.

Table 6. The number of node pairs with ties

| Networks | Pairs of nodes | % |
|----------|----------------|------|
| 1 | 1,646,740 | 25.3 |
| 2 | 2,937,609 | 45.2 |
| 3 | 1,028,519 | 15.8 |
| 4 | 795,317 | 12.2 |
| 5 | 93,170 | 1.4 |

Table 8 further shows the correlations between subnetworks at the individual level, in which we calculated the correlation among the degree of nodes and weight of edges across subnetworks. The low value suggests users' inconsistent involvement levels across different types of social support.

Table 7. Jaccard coefficients of node (gray shading) and edge

| Node & Edge | PES | PIS | SIS | SES | COM |
|-------------|-------|-------|-------|-------|-------|
| PES | -- | 0.596 | 0.054 | 0.168 | 0.207 |
| PIS | 0.912 | -- | 0.058 | 0.164 | 0.193 |
| SIS | 0.477 | 0.493 | -- | 0.104 | 0.065 |
| SES | 0.653 | 0.657 | 0.636 | -- | 0.165 |
| COM | 0.698 | 0.694 | 0.577 | 0.727 | -- |

Table 8. Pearson correlation of sub-network degree of nodes (gray shading) and weight of edges

| Degree & Weight | PES | PIS | SIS | SES | COM |
|-----------------|-------|-------|-------|-------|-------|
| PES | -- | 0.932 | 0.775 | 0.807 | 0.840 |
| PIS | 0.898 | -- | 0.776 | 0.792 | 0.856 |
| SIS | 0.292 | 0.283 | -- | 0.758 | 0.828 |
| SES | 0.525 | 0.494 | 0.271 | -- | 0.782 |
| COM | 0.574 | 0.534 | 0.212 | 0.389 | -- |

5. Discussion

5.1. Support seeking and providing

The descriptive statistics showed that Baidu Discussion Forum users sought emotional support much more frequently than informational support. According to a comparative study [9], the requirement for emotional support in US OHC was frequently seen but not particularly intense and informational support seemed to be a more critical theme for user interaction. This result is in line with previous cross-cultural studies, which have found that during times of stress,

people from Asian cultures are more likely to cope through changing the individual's feelings and thoughts to adjust to the objective environment, while people from Western cultures are more inclined to cope through changing the objective environment to fit the individual's needs. Additionally, some cultural beliefs among Chinese people, especially the irrational ones like fatalistic beliefs, could enable them to accept their fate to be afflicted with cancer and make them believe they can do nothing about it. Holding these cultural perceptions, they would be more likely to seek emotional support, as informational support seems less important, or even not necessary. Simply put, when faced with breast cancer risks, Chinese users are more inclined to adopt emotional regulation strategies to change the way they think or feel about the stressful situation, seeking social support from other users' understanding, concern, encouragement, sympathy, and validation.

5.2. Companionship

Interestingly, we found that not a large number of Baidu Discussion Forum users were engaged in the off-topic discussions, which was quite different from the users' discussion topics in similar US OHCs [20]. Companionship has been a frequently reported topic of OHCs based on the Western population, and prior research reported that the seemingly off-topic discussions may help OHC users make each other's acquaintance personally during their communication beyond health topics [9]. It remains to be unknown why Chinese OHC users are less enthusiastic about daily life topics. One potential explanation based on self-disclosure theory is that Chinese people tend to feel uncomfortable discussing family life and personal habits online in public, or they are more willing to turn to family members, friends, or smaller health communities to share such information with a handful of people they perceive to be more trustworthy.

In addition, emotional support and companionship topics work well in connecting one user with the other and contribute to the maintenance and expansion of online networks. That said, even though on a breast cancer OHC, people are connected because of a disease and a common identity, discussions of non-breast-cancer related topics are critical to keeping users actively engaged in the community. This result corresponds with a previous case study of coping with breast cancer through online discussion group messages [24], finding that the most frequent interaction theme among the users is about building friendship with peers through communication of validation, encouragement, appreciation, and life sharing. On the one hand, longtime users may have received sufficient informational support and do not

require much more discussion on the disease itself. On the other hand, they may want to stay in the community and interact over time with users who have shared purposes, interests, or needs. As a way of building personal bonds, off-topic discussions in the form of sharing personal life events may help strengthen the connections among users more than informational support [9]. The lack of companionship topics on the OHC could explain why some users engage only for a short time of user engagement and with a small number of online network connections. The reasons for OHC users' preference of staying on the health-related issues deserve further investigation.

5.3. Network analysis

In terms of network analysis, a most important and interesting finding is that a large number of Baidu Discussion Forum users were interacting with a fairly small proportion of very productive users, which led the opinion leaders to be quite influential in the community. Back to the original dataset, it is notably seen that one user—Ms. Bai—has been extremely influential in shaping the entire network pattern. To the best of researchers' knowledge, there is no detailed accessible introduction to Ms. Bai on her expertise of breast cancer physician. According to her profile on Baidu Discussion Forum, she claims herself to be a breast cancer survivor for 5 years and has been very motivated to share her personal knowledge and experience in breast cancer treatment, encouraging other users to bravely face the reality of cancer. Until early 2019, she has been in this forum for 5 years 10 months, posting and commenting on approximately 46,000 messages. Based on the researcher's observation, most of the posts are informative, well-organized, focused, and represent timely responses. One of her posts has had surprisingly long longevity. It was published in 2016, stayed open for 720 days, attracted over 3,700 users, and received 8,832 comments. This explained the sudden drop noticeable in Figure 3. As this thread was initially posted with the purpose of answering other users' questions, most comments are initiated by users who are seeking social supports on breast cancer topics. Quite a number among them post their own, or their important others' breast cancer test results, or scanned medical certificates for Ms. Bai, asking for her interpretation and follow-up treatments advice, or expressing emotions such as fear, anxiety, or confusion, expecting to get emotional support. Some users are seen to register primarily aiming to post a question on Ms. Bai's stream, as no records were identified on their interaction with other users. As such, some particularly productive users shape the entire network in its unique way.

The follow-up analyses indicated that forty-five percent of the users in the OHC were connected in only two subnetworks, which constituted the highest portion of subnetwork numbers. Perhaps Chinese users treat OHCs as a source to approach cancer information, rather than a sharing platform for obtaining or providing other types of support. In such circumstances, users will leave the community after their needs being satisfied, rather than stick to the OHC for a longer period of time. This is further confirmed by the correlation of the nodes' degree and the edges' weight—only a few individuals involved in multiple social support subnetworks, and highly interactive with each other.

Another finding is that PES was especially powerful in the expansion of the community, as it features high densities and the largest numbers of edges among five subnetworks, meaning that it spread more extensively than other types of social support and cultivated more connections among users than other subnetworks. One possible explanation is that emotional support itself is a critical theme in cancer-related communication, and has been shown to increase participants' sense of social support, personal empowerment, and self-esteem and to reduce depression, cancer-related trauma, and social isolation [25]–[29]. When individuals perceive other users' responses as understanding, validating, and caring, they would be more willing to remain active in the community [30]. In light of earlier research, social support providers themselves can get benefits by providing support to others in terms of psychosocial well-being. Thus, the providers would be continuing to participate in a community. The user retention helps in expanding a user's subnetworks. Although the subnetwork of PIS has a very similar network pattern as the PES in either community, compared with other PIS, PES requires a relatively low threshold for posting, and such a lower barrier potentially enables a larger group of users to be engaged. Theoretically, any users in an OHC are capable of providing emotional support to varying degrees, no matter to what extent they are knowledgeable on breast cancer. In comparison, other post themes require an "entry level," such as personal needs, knowledge, experience, or willingness of interaction. Therefore, emotional support topics may work as an umbrella to unite everyone in communication. In addition, once diagnosed with this chronic disease, a user's demand for informational support may be strong at the very beginning or at a certain time during the treatment, but the needs for emotional support could be continuous. As such, the networks based on this topic may accumulate and grow over a longer interval.

5.4. Content description

When investigating the details of each particular posting on Baidu Discussion Forum, we also found the users' postings were simple and targeted, leaving fairly limited room for others to respond, elaborate or initiate an interaction. The users' responses to other users' postings, if any, were rarely detailed as well. As both senders and receivers play critical roles in the communication process, users' participation behaviors in a community are not only based on their own behaviors but on other users' reactions toward this individual as well [31]. Insufficient responses would undermine the active users' contributions which may finally destroy the most prominent foundation of OHC—users' interactions in a community. As such, the users may not have a strong sense of belonging towards this community and, therefore, are less likely to be willing to contribute to it [32]. Reciprocity, which requires informational exchange between the users, especially emphasizing on users' return of the favor, specifically to those who have helped them [33], serves as a critical predictor for the users to offer follow-up social support [22], [34]. Therefore, to keep a community healthy, functional, lively, and dynamic, the improvement of the entire Chinese OHCs' environment is expected.

6. Implications

From the perspective of theoretical implications, the study serves a good example of combining text mining and social network analysis in identifying Chinese users' communication patterns in the OHC. The outcomes of the study further provide evidence that there are cultural differences in participating healthcare communities. Not only the motivation of participation, the users' disparities are also shown in the content they post and the social roles they act. Such conclusions may inspire new research directions in health communication studies.

In terms of practical implications, first, as OHCs have been an important platform for users to seek and obtain social support among the Chinese population, the administrators and moderators may try to organize the most representative, comprehensive, professional, and well-organized health messages, provide a list of top sticky posts, inform the users of the updates, and enable the users to archive the target information more conveniently. Second, as productive users are regarded as social support providers with expertise and sincerity, more similar influencers, such as doctors, health practitioners, and survivors should be encouraged to join this community. Their engagement could be effective in motivating users' active participation. In addition, given the PES is beneficial for the expansion

of the OHC, the community should encourage users to be more engaged in providing emotional support. For example, the community is able to recommend individuals who are seeking support to some experienced users and spur the latter's responses for providing encouragement, understanding, or comfort.

7. Limitations and Future Research

This study has limitations that point to the need for future research to be conducted. First, we encourage future research to investigate what the Chinese OHC users' concerns are in disclosing their own information and providing social support for others in order to motivate them to engage actively in OHCs. Qualitative studies can be helpful in more completely understanding and explaining their behaviors and leaving a clue to improve OHCs' sustainability.

Second, the social support network in this study is an undirected weighted network representing the co-participation relationships between users. Considering the direction of edge in providing and seeking social support would lead us a better understanding of the diffusion pattern of social support in among the Chinese population. As a matter of fact, its realization would require more user-generated content analyses in the future.

Last, this study provides some empirical evidence of the features of user participation among Chinese. Future studies are encouraged to explore the causal effects of cultural factors on users behaviors in OHCs.

8. References

- [1] Ji. Shi, "Chinese embrace online healthcare - Chinadaily.com.cn," *China Daily*, 19-Jul-2016.
- [2] P. K. H. Mo and N. S. Coulson, "Online support group use and psychological health for individuals living with HIV/AIDS," *Patient Education and Counseling*, vol. 93, no. 3, pp. 426–432, Dec. 2013.
- [3] R. S. Mano, "Social media and online health services: A health empowerment perspective to online health information," *Computers in Human Behavior*, vol. 39, pp. 404–412, Oct. 2014.
- [4] M. Lieberman, "The role of insightful disclosure in outcomes for women in peer-directed breast cancer groups: a replication study," *Psycho-Oncology*, vol. 16, no. 10, Oct. 2007.
- [5] A. Barak, M. Boniel-Nissim, and J. Suler, "Fostering empowerment in online support groups," *Computers in Human Behavior*, vol. 24, no. 5, pp. 1867–1883, Sep. 2008.
- [6] P. Wicks *et al.*, "Sharing Health Data for Better Outcomes on PatientsLikeMe," *J Med Internet Res*, vol. 12, no. 2, Jun. 2010.
- [7] C. Young, "Community Management That Works: How to Build and Sustain a Thriving Online Health Community," *J Med Internet Res*, vol. 15, no. 6, Jun. 2013.
- [8] R. K. Jayanti and J. Singh, "Pragmatic learning theory: An inquiry-action framework for distributed consumer learning in online communities," *Journal of Consumer Research*, vol. 36, no. 6, pp. 1058–1081, 2010.
- [9] X. Wang, K. Zhao, and N. Street, "Analyzing and Predicting User Participations in Online Health Communities: A Social Support Perspective," *Journal of Medical Internet Research*, vol. 19, no. 4, p. e130, 2017.
- [10] S. A. Shumaker and A. Brownell, "Toward a theory of social support: Closing conceptual gaps," *Journal of social issues*, vol. 40, no. 4, pp. 11–36, 1984.
- [11] T. O. Blank, S. D. Schmidt, S. A. Vangsness, A. K. Monteiro, and P. V. Santagata, "Differences among breast and prostate cancer online support groups," *Computers in Human Behavior*, vol. 26, no. 6, pp. 1400–1404, Nov. 2010.
- [12] J. A. Greene, N. K. Choudhry, E. Kilabuk, and W. H. Shrank, "Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook," *J Gen Intern Med*, vol. 26, no. 3, pp. 287–292, Mar. 2011.
- [13] S. S. Loane and S. D'Alessandro, "Communication That Changes Lives: Social Support Within an Online Health Community for ALS," *Communication Quarterly*, vol. 61, no. 2, pp. 236–251, Apr. 2013.
- [14] Y. Zhang, D. He, and Y. Sang, "Facebook as a Platform for Health Information and Communication: A Case Study of a Diabetes Group," *J Med Syst*, vol. 37, no. 3, p. 9942, Apr. 2013.
- [15] Y. W. C. Vivien and N. A. M. Noor, "AN EXAMINATION OF SOCIOCULTURAL FACTORS INFLUENCING LIFESTYLE, HEALTH AND HEALTH-SEEKING BEHAVIOUR – A CASE STUDY OF MALAYSIAN CHINESE CANCER SURVIVORS," p. 29, 2013.
- [16] F. J. Boster and P. Mongeau, "Fear-Arousing Persuasive Messages," *Annals of the International Communication Association*, vol. 8, no. 1, pp. 330–375, Jan. 1984.
- [17] L. K. Guerrero, P. A. Andersen, and W. A. Afifi, *Close encounters: Communication in relationships*. Sage Publications, 2017.
- [18] D. C. Barnlund, *Public and private self in Japan and the United States: Communicative styles of two cultures*. Simul Press Tokyo, 1975.

- [19]D. W. Sue and D. Sue, *Counseling the culturally different: Theory and practice, 2nd ed.* Oxford, England: John Wiley & Sons, 1990.
- [20]X. Wang and Y. Zhu, "Cross-Cultural Comparison of User Engagement in Online Health Communities," in *Smart Health*, vol. 10983, H. Chen, Q. Fang, D. Zeng, and J. Wu, Eds. Cham: Springer International Publishing, 2018, pp. 96–104.
- [21]X. Zhang, S. Liu, Z. Deng, and X. Chen, "Knowledge sharing motivations in online health communities: A comparative study of health professionals and normal users," *Computers in Human Behavior*, vol. 75, pp. 797–810, Oct. 2017.
- [22]X. Zhang, S. Liu, X. Chen, L. Wang, B. Gao, and Q. Zhu, "Health information privacy concerns, antecedents, and information disclosure intention in online health communities," *Information & Management*, vol. 55, no. 4, pp. 482–493, Jun. 2018.
- [23]M. Szell, R. Lambiotte, and S. Thurner, "Multirelational organization of large-scale social networks in an online world," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 13636–13641, Aug. 2010.
- [24]K.-Y. Wen, F. McTavish, G. Kreps, M. Wise, and D. Gustafson, "From Diagnosis to Death: A Case Study of Coping With Breast Cancer as Seen Through Online Discussion Group Messages," *Journal of Computer-Mediated Communication*, vol. 16, no. 2, pp. 331–361, 2011.
- [25]J. Fogel, S. M. Albert, F. Schnabel, B. A. Ditkoff, and A. I. Neugut, "Internet use and social support in women with breast cancer," *Health Psychology*, vol. 21, no. 4, pp. 398–404, 2002.
- [26]E.-O. Im, W. Chee, H.-M. Tsai, L.-C. Lin, and C.-Y. Cheng, "Internet Cancer Support Groups: A Feminist Analysis," *Cancer Nursing*, vol. 28, no. 1, pp. 1–7, Jan. 2005.
- [27]C. Lindemalm, P. Strang, and M. Lekander, "Support group for cancer patients. Does it improve their physical and psychological wellbeing? A pilot study," *Supportive Care in Cancer*, vol. 13, no. 8, pp. 652–657, Aug. 2005.
- [28]S. Rodgers and Q. Chen, "Internet community group participation: Psychosocial benefits for women with breast cancer," *Journal of Computer-Mediated Communication*, vol. 10, no. 4, pp. 00–00, 2005.
- [29]A. J. Winzelberg *et al.*, "Evaluation of an internet support group for women with primary breast cancer," *Cancer*, vol. 97, no. 5, pp. 1164–1173, Mar. 2003.
- [30]J.-P. Laurenceau, L. F. Barrett, and P. R. Pietromonaco, "Intimacy as an interpersonal process: The importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges.," *Journal of Personality and Social Psychology*, vol. 74, no. 5, pp. 1238–1251, 1998.
- [31]J. M. Levine and R. L. Moreland, "Group Socialization: Theory and Research," *European Review of Social Psychology*, vol. 5, no. 1, pp. 305–336, Jan. 1994.
- [32]A. Cifollilli, "Phantom authority, self-selective recruitment and retention of members in virtual communities: The case of Wikipedia," *First Monday*, vol. 8, no. 12, Dec. 2003.
- [33]J. Heyman and D. Ariely, "Effort for Payment: A Tale of Two Markets," *Psychological Science*, vol. 15, no. 11, pp. 787–793, Nov. 2004.
- [34]F. Yang, B. Zhong, A. Kumar, S.-M. Chow, and A. Ouyang, "Exchanging Social Support Online: A Longitudinal Social Network Analysis of Irritable Bowel Syndrome Patients' Interactions on a Health Forum," *Journalism & Mass Communication Quarterly*, vol. 95, no. 4, pp. 1033–1057, Dec. 2018.