Full length article

Quantifying the Effect of Informational Support on Membership Retention in Online Communities through Large-Scale Data Analytics

Wanli Xing [a, *], Sean Goggins [b], Josh Introne [c]

[a] Department of Educational Psychology & Leadership, Texas Tech University, Lubbock, TX 79409, USA
[b] Department of Computer Science, University of Missouri, Columbia, MO 65211, USA
[c] Department of Media & Information, Michigan State University, East Lansing, MI 48824, USA

ABSTRACT

Participating in online health communities for informational support can benefit patients in various ways. For the online communities to be sustainable and effective for their participants, membership retention and commitment are important. This study explores how informational support requesting and providing by users holding different social roles (core user and periphery user) are related with participants' retention in the community. We first crawled six years of data in the WebMD fibromyalgia forum with around 200,000 posts and 10,000 users. Then a supervised machine learning model is trained and validated to automatically identify the requesting and providing informational support posts exchanged between the members in the community. Lastly, survival analysis was employed to quantify how the informational support requesting and providing by different social roles predicts the member's continued participation in the online community. The results reveal the different influencing mechanism of requesting and providing support from different social roles on the patients' decision to stay in the community. The findings can aid in the design of better support mechanisms to enhance member commitment in online health communities.

Published by Elsevier Ltd.

1. Introduction

Large numbers of American adults engage with online health communities to seek health-related informational support (Qiu et al., 2011). Doing so can not only help them to deal with stress, but can also teach them to manage their health conditions in new ways (Coulson & Greenwood, 2012). "Informational support" refers to information provided via online forums in which participants can ask about health problems and receive information and advice about treatments, coping with symptoms, side effects, and financial and other burdens (Chuang, 2013; McCormack, 2010). Online health communities have become important vehicles for informational support, and they assure users that they are cared for (Chuang, 2013). However, patients must continue to participate in online communities to receive all of their benefits.

Extensive research has been conducted into the informational support provided via online health communities. This research has primarily examined the information exchanged via these communities and the effects of such exchanges. For example, several researchers have identified various types of informational support provided via posts to online health communities and have tested how these types of support relate to individual health outcomes and a sense of community (Zhang, Liu, Deng, & Chen, 2017; Kirk & Milnes, 2016; Greene, Choudhry, Kilabuk, & Shrank, 2011). Such studies generally use methods like surveys, interviews, content analyses, and ethnographic cases, and they generally involve between a few and several hundred subjects.

While such studies are common, little is known about how the informational support exchanged among the members of online health communities affects their commitment to these communities. Understanding user commitment is important to both individual participants and whole communities. The benefits of becoming involved in a given online health community likely depend to a large extent on the information exchanged among the members of that community (Wang, Kraut, & Levine, 2012). If a member does not remain in a community, they are less likely to receive the benefits offered by that community. Moreover, members are primary sources of resources. Because communities

* Corresponding author.
E-mail addresses: wanli.xing@ttu.edu (W. Xing), gogginss@missouri.edu (S. Goggins).

aggregate their knowledge, a larger community is likely to know more about a given problem than a smaller one does (Butler, 2001). A sustainable level of participation cannot be taken for granted, however. Many studies have shown that large numbers of members drop out before they can contribute to or benefit from their communities (Resnick, Janney, Buis, & Richardson, 2010; Wang et al., 2012; Yang, Kraut, & Levine, 2017). For this reason, an online health community must understand the factors that affect its members' participation if it is to remain sustainable and effective.

Since members join online health communities to cope with their medical conditions, the amount and quality of informational support that they request and receive should greatly impact their commitment to their communities (Wang et al., 2012). Many sustainable online health communities exhibit a core–periphery structure (Cobb, Graham, & Abrams, 2010), in which a small but active group of core users provides most of the support requested by a much larger group of periphery users. We hypothesized that interacting with users holding different social roles have different influences on their commitment to the community. Certain members have more experience and expertise than others do in meeting participants' informational needs (Atanasova, Kamin, & Petrič, 2017), and this is likely to influence users' decisions to continue participating in their communities. Given the fact that an online health community can easily have thousands of members and tens of thousands of forum posts, understanding members' commitment to online health communities requires automated data processing methods beyond those traditionally used in the social sciences. For this reason, this study uses data mining to process posts on a large scale.

The research question for this study is, "How do requests for and provisions of informational support by members with different social roles influence members' continued participation in online health communities?" To answer this question, we crawled six years' worth of posts to the WebMD fibromyalgia forum, accumulating approximately 200,000 posts from 100,000 users. A large-scale data analysis was then used to process these posts. This analysis had two parts. First, a supervised machine-learning model was trained to automatically identify and distinguish posts requesting informational support and posts providing informational support. Second, a survival analysis was used to quantify how requests for and provisions of informational support by members with different social roles impact members' continued participation in the online community. The findings of this study could aid in the development of support mechanisms designed to increase members' commitment to online health communities.

2. Background

Members of online health communities typically do not know each other personally. Instead, they communicate online to acquire information and advice about a variety of health-related issues, including shared health conditions, treatments, and side effects. Studies show that the informational support provided via online health communities can benefit the members of these communities in a number of ways. McCormack (2010), for example, found that exchanging information and support can improve a patient's ability to deal with stress. Rodgers and Chen (2005) observed that patients who interacted frequently with other members of an online breast-cancer community had better mood profiles than those who did not. Braithwaite, Waldron, and Finn (1999) found that the informational support provided by online communities for chronic diseases is especially useful: it benefits not only the patients, but also their health providers and families. However, patients must continue to participate in online communities to receive all of their benefits.

2.1. Theoretical foundations of online-community retention

According to the resource-based model of online communities (Butler, 2001), people use their resources—e.g. their time, energy, and knowledge—to benefit the community. Äkkinen (2005) argues that if the perceived benefits of participating in a community exceed the resources sacrificed, then the community creates value for its members. Butler (2001) links perceived benefits and resources sacrificed to the number of members in the community. If the perceived benefits from a community exceed the resources sacrificed, then the number of members in the community will increase. If the perceived benefits from a community do not exceed the resources sacrificed, however, members will leave the community. Butler's study generated a comprehensive list of the benefits offered by online social communities. One of the most important of these benefits is informational support. If an online community provides to a member informational support whose benefits exceed the sacrifices made by the member, the member will likely continue to engage in the community. In contrast, if the informational support provided does not exceed the sacrifices made by the member, the member may gradually drop out of the community. Gu and Jarvenpaa (2003) echoed Butler's claim from the perspective of economic theory, stating that people will only contribute if the benefits outweigh the costs. In collective settings, such as online communities, the tangible and intangible returns valued by the members (e.g. informational support) incentivize their continued participation.

Social theories also offer arguments linking the informational support provided via online communities to user retention. Social exchange theory claims that people contribute and engage because of the benefits they expect to receive in return; this is called "future reciprocity" (Bearman, 1997). In other words, individuals engage in social interactions because they expect that doing so will result in social rewards, including status and the support of others. This theory argues that the expectation of reciprocity motivates individuals to contribute and engage beyond the equilibrium predicted by economic models of utility (Blau, 1964). Members of online communities may hold different roles (e.g. "core" and "periphery") based on their relative contributions and activity levels. Interacting with different users in requesting and providing informational support may produce different reciprocity expectations, which can, in turn, influence members' decisions to stay or to leave.

2.2. Empirical research into informational support and retention

Many studies on informational support have examined the contents of posts to forums. To identify the characteristics of informational support, for example, Zhang, He, and Sang (2013) qualitatively analyzed the content of 1352 posts sampled from various online health communities. Hwang et al. (2010) examined the informational support provided via an online weight-loss community. They conducted semi-structured interviews with 13 patients and characterized the support these patients received by qualitatively analyzing the transcripts of their interviews. Many other studies have investigated informational support from a relationship perspective, examining how informational support relates to other problems and/or entities (e.g. coping with diseases and a sense of community). For instance, Welbourne, Blanchard, and Boughton (2009) administered surveys to 122 members of an online infertility community. They found that informational support improves a user's sense of community and serves as a buffer between their physical health symptoms and their stress. Ginossar (2008) statistically analyzed 1424 posts to two online cancer communities to identify differences in how men and women use informational support. Høybye, Johansen, and Tjørnhøj-Thomsen