Literature Review Assignment
MACS 30200, Dr.Evans
Jingyuan Zhou
April.24th

Research question: Do critical users on online social Q&A communities have broader interests in topics compared to others?

This specific research is one of a series that is meant to understand the difference between influential people and the others in social communities. In this series, we do not try to conclude a causal relationship but to observe patterns that are able to differentiate two groups. We attempt to approach this question by analyzing history of activities of users on zhihu.com, an online social Q&A platform. Due to the nature of this research, this literature review is split into two parts: works on diffusion of influence and distinguishing influential users and works based on social Q&A communities.

1. Related works on diffusion of influence

Influence has been an important topic for people to understand how and why some innovations or ideas are adopted by larger population faster than it others. Thus, it is a critical topic for fields including sociology, communication, marketing, and political science (Rogers 1962; Katz and Lazarsfeld 1955). Many empirical studies have focused on the general diffusion of influence and the roles of influencers.

In the early stages, researchers have used activities on online blogs and e-commerce websites to address these questions. Some have used time-stamped observations of posts to infer a transmission network between bloggers under the assumption that transmission is an independent cascade model (Gruhl et al 2004), others attempt to infer how people are influenced by the number of contacts who recommend them a certain product by utilizing the referral on an e-commerce website (Leskovec et al 2007). These earlier researches have suffered from the lack of network structure of their data; however, due to the prevalence of social media sites including Twitter and Facebook where network structure is explicitly imposed for their users, later researches were able to take advantage of this feature to better understand diffusion of influence and even measure the difference of influence between users. For example, Sun et al. used Facebook to analyze diffusion trees of fan pages.

A series of researches based on Twitter data were managed to make significant contributions in this topic. Kwak et al. show that the rankings of most influential users based on number of followers, number of retweets or page-rank are different. Cha et al. find that the most followed users are not necessarily the most influential ones according to their measurement by comparing three statistics — number of followers, number of mentions and number of retweets. Bakshy et al. show efficacy of ordinary influencers as oppose to word-of-mouth strategies that depend on triggering "social epidemics" by targeting special individuals.

This abundance of literature on influence provides us with empirical results on the unreliability of determining influential users solely based on their global network measurements. In fact, these results correspond to a modern view of information flow that emphasizes the importance of prevailing culture instead of the role of influentials (Domingos and Richardson 2001). Thus, we arrive at the idea of determining influential users of a network based on their ranking in their own communities instead of their network measurements in the entire network.

2. Related studies based on online Q&A sites

Soical question answering (SQA) sites are online communities for information seeking by asking natural language questions to other users in a network (Shah, 2008). Social question answering may occur by a user posting a question in designated SQA services or systems such as Yahoo! Answers, Quora, or Zhihu (Harper, Moy, & Konstan, 2009; Kim, 2010)

2.1 Collaboration

SQA sites contain both informational questions to solicit specific facts and conversational questions to carry on discussions (Harper et al., 2009).

Previous research has reported social collaborations in SQA and collaborative information seeking (Hansen and Rvelin, 2005; Gazan, 2010; Wang, Gill & Mohanlal, 2013). For example, two thirds of social collaboration in collaborative information retrieval is document-related, while one third is human-related (Hansen and Rvelin, 2005). Research has found that in SQA sites, collaboration takes place in brief, informal episodes, and users with higher ranking are found to contribute more content (Gazan, 2010).

Factors influencing users' collaborative behaviors include willingness to share information, altruism and morality, perceived pleasure, social capital and resources, and affective factors (Hertzum, 2008; Gazan, 2010; Zhang, 2012).

2.2 Answer quality

Answer quality is an important part of SQA research. Researchers attempt to learn the criteria that users use to evaluate the quality of the answer in a social Q&A community. On the basis of SQA service of Yahoo! Answers, Shah and Pomerantz summarized a small set of questions, with at least five answers for each, then asked Amazon Mechanical Turk workers to assess the quality of each answer for a given question based on 13 different criteria. Zhu et al. developed a multi-dimensional model which includes another 13 indicators for users to evaluate the answer quality of an SQA site. Soojung and Sanghee used the criteria of selecting the best answers in Yahoo! Answers and analyzed 2,140 comments with the content analysis and identified that 23 individual relevance criteria could be divided into six classes, which are content, cognition, utility, information sources, extrinsic state, and socio-emotion. In addition, they also find that the importance degree of individual criteria varies according to topic categories, and socio-emotion is a popular criterion in discussion-oriented categories of SQA sites.

2.3 User roles

Related researches have recognized three groups of user roles: administrators, content contributors, and marginal roles. Administrators could be split into two groups: mediators/ moderators that maintain the order of online communities by preventing flames, filtering spams and facilitating ongoing discussions (Preece, 2000; Gazan, 2010); Vandal fighters/flame warriors that sanction norm violators (Geiger & Ribes, 2010). Content contributors could be split into questioners (Gazan, 2010), answer people (Gleave, 2009; Haythornthwaite, 2005; Turner, 2005), discussion people (Gleave, 2009; Haythornthwaite, 2005; Turner, 2005) and technical editor (Gleave, 2009; Geiger & Ribes, 2010). Marginal roles include fans (Haythornthwaite 2005; Turner, 2005) and lurkers who do not actively contribute contents or connect with others (Preece, 2000; Gleave, 2009).

Contribution

Existing literature based on social Q&A websites have mainly focused on collaboration behavior and answer evaluation. In the discussion of roles, previous researches have defined specific roles according to their behavior in certain part of the network. Under their framework, a user could be questioner in some topics but discussion people or technical editor in others. We are proposing to view users as different combinations of these specific roles and define their role based on their influence within their own communities under the network structure provided by zhihu.com. Thus, by traversing activity history of users, we can understand if there is a difference between influential users and others under this new definition of roles. We've shown that this approach is both new on defining influential users in social network and analyzing social Q&A sites.

Reference

Rogers, E. M. 1962. Diffusion of Innovations. Free Press.

Katz, E., and Lazarsfeld, P. 1955. Personal Influence: The Part Played by People in the Flow of Mass Communications. New York: The Free Press.

D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins. Information diffusion through blogspace. pages 491–501. ACM New York, NY, USA, 2004.

J. Leskovec, A. Adamic, Lada, and A. Huberman, Bernardo. The dynamics of viral marketing. ACM Trans. Web, 1(1):5, 2007.

E. S. Sun, I. Rosenn, C. A. Marlow, and T. M. Lento. Gesundheit! modeling contagion through facebook news feed. In International Conference on Weblogs and Social Media, San Jose, CA, 2009. AAAI.

H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? pages 591–600. ACM, 2010.

M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummad. Measuring user influence on twitter: The million follower fallacy. In 4th Int'l AAAI Conference on Weblogs and Social Media, Washington, DC, 2010.

E. Bakshy, B. Karrer, and A. Adamic, Lada. Social influence and the diffusion of user-created content. In 10th ACM Conference on Electronic Commerce, Stanford, California, 2009. Association of Computing Machinery.

Domingos, P., and Richardson, M. 2001. Mining the Network Value of Customers. In ACM SIGKDD.

Shah, C., Oh, J. S., & Oh, S. (2008). Exploring characteristics and effects of user participation in online social Q&A sites. First Monday, 13 (9).

Harper, F. M., Moy, D., & Konstan, J. A. (2009, April). Facts or friends: distinguishing informational and conversational questions in social Q&A sites. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 759–768. ACM.

Kim S. (2010). Questioners' credibility judgments of answers in a social question and answer site. Information Research, 15 (2), 5.

Hansen, P., & Järvelin, K. (2005). Collaborative information retrieval in an information-intensive domain. Information Processing & Management, 41 (5), 1101–1119.

Gazan, R. (2010). Microcollaborations in a social Q&A community. Information processing & management, 46 (6), 693–702.

Wang, G., Gill, K., Mohanlal, M., Zheng, H., & Zhao, B. Y. (2013, May). Wisdom in the social crowd: an analysis of quora. In Proceedings of the 22nd international conference on World Wide Web, 1341–1352. ACM.

Hertzum, M. (2008). Collaborative information seeking: The combined activity of information seeking and collaborative grounding. Information Processing & Management, 44 (2), 957–962.

Zhang, P. (2012). Information seeking through microblog questions: The impact of social capital and relationships. Proceedings of the American Society for Information Science and Technology, 49 (1), 1–9.

Zhu, Z.M., Bernhard, D., & Gurevych, I. A multi-dimensional model for assessing the quality of answers in social Q&A.

Soojung, K., & Sanghee, O. Uses' relevance criteria for evaluating answers in a social Q&A site. Journal of the American Society for Information Science and Technology, 2009, 60(4): 716-727.

Preece, J. (2000). Online Communities: Designing Usability, Supporting Sociability. Jonh Wiley & Sons. Inc., New York, NY, USA.

Gleave, E., Welser, H. T., Lento, T. M., & Smith, M. A. (2009, January). A conceptual and operational definition of 'social role' in online community. System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference, 1–11. IEEE.

Haythornthwaite, C., & Hagar, C. (2005). The social worlds of the Web. Annual review of information science and technology, 39 (1), 311–346.

Turner, T. C., Smith, M. A., Fisher, D., & Welser, H. T. (2005). Picturing Usenet: Mapping computer-mediated collective action. Journal of Computer-Mediated Communication, 10 (4).

Shah, C., & Pomerantz, J. Evaluating and predicting answer quality in community QA. Proceedings of the 33rd international ACM SIGIR conference on research and development in information retrieval. ACM, 2010: 411-418.