

What Influences the Decision to Participate in Audience-bounded Online Communities?

Claudia López, Rosta Farzan,
Shaghayegh (Sherry) Sahebi, Peter Brusilovsky
University of Pittsburgh
{cal95, rfarzan, shs106, peterb}@pitt.edu

Abstract

Building online communities to support small, audience-bounded offline social structures such as neighborhoods or organizations can be challenging. Due to the small size of their potential audience, the contribution volume is likely to be insufficient to maintain a sustainable community-driven system. In our research, we are interested in studying how different characteristics of the offline structure of these communities can influence their online behavior. Particularly, we analyzed participation of researchers in a social system for conferences. Our analysis shows that centrality in the academic social structure is a significant predictor of the likelihood to accept an invitation to participate in the system. These results suggest that an understanding of the users' offline context can increase the effectiveness of user engagement strategies in an online context.

Keywords: audience-bounded online communities, offline structure, participation

Introduction

Online communities have been widely known for enabling communication among large numbers of people distributed around the world. However, online communities can also greatly benefit smaller existing offline groups. Providing virtual spaces for communication to existing offline groups can increase community awareness and engagement (Schuler, 1994; Carroll, 2012). These kinds of online communities aim to reach much smaller cohorts of people who share an offline context and have opportunities to interact outside the online community. For example, there are private online social networks serving neighborhoods (e.g. Nextdoor.com); online communities to enhance the experiences of the event attendees (Farnham, Brown, & Schwartz, 2009; Farzan & Brusilovsky, 2008); educational online communities to support students¹; and organization-specific social network sites to share knowledge and to increase networking among employees (DiMicco et al., 2008).

One of the main challenges for online communities is to maintain a sufficient level of contribution. Even with a potentially global reach, online communities often fail to maintain a critical mass (Cummings, Butler, & Kraut, 2002; Raban, Moldovan, & Jones, 2010). Building self-sustainable social online systems for already existing offline communities is even more challenging because of the much smaller population of potential participants (Resnick, Janney, Buis, & Richardson, 2010). Therefore, special efforts focusing on user engagement are required.

The shared offline context has the potential to affect members' motivation to participate in the online community. For instance, it can influence the perceived usefulness of the online community. The fact that the potential users of an audience-bounded online community are likely to know each other or share some offline context adds to the challenge. The perceived benefits of using the systems need to surpass not only the costs of participating as in every online community (Butler, 2001), but it also needs to provide benefits that cannot be easily achieved through the offline interactions. Our research agenda aims to understand how

¹ <http://studentscircle.net>

offline social properties of potential users are associated with the patterns of participation in audience-bounded online communities. To achieve this goal, we are collecting quantitative and qualitative data about the usage of online communities for conferences, graduate schools and neighborhood-oriented online communities.

In this paper, we analyze how offline centrality in an academic social network influences the likelihood of participation in an online community for academic conferences. Our results show more central members in the offline social structure are significantly more likely than peripheral members to respond to invitations to participate in the online community. Our findings reveals that the offline context of potential users can be used to design a more efficient user engagement strategies.

Research Questions

Economic modeling provides a framework for understanding peoples' motivation in making decisions to take an action, including participation in online communities (Harper, Li, Chen, & Konstan, 2005). Often people compare the cost of performing an action with its benefits in order to decide to take the action. To participate in an online community, the perceived benefits of using the systems need to be more significant than the costs of participating (Butler, 2001).

A potential benefit of online communities is easier access to information (Schwabe & Prestipino, 2005; Cotten & Gupta, 2004). Online communities can provide more opportunities for individuals in terms of exposure to those who hold relevant information. Centrality in a social network has been found to be associated with access to more information (Brass, 1984). In this sense, we can argue that more central members of the offline community can get most of the information they need through their numerous offline connections while more peripheral members have less access to information because of fewer connections in the offline world. To compensate for this lack of information access, peripheral members can take advantage of online communities to acquire the information they cannot access offline. Thus, the perceived benefits of a social system can be higher for more peripheral members of the community. This leads us to our first hypothesis:

H1a: Peripheral offline community members are more likely to decide to participate in an audience-bounded online community than more central community members in order to compensate for their lack of connections offline.

On the other hand, research in organizational behavior suggests that more central members have a stronger sense of attachment to the organization (Burt, 2001). Their network position within the organization facilitates their understanding of what it means to be part of the organization and strengthens their sense of identification with the organization (Burt, 2001). In an online site for a community of practice, centrality in the online social structure was found to be correlated to higher contribution rates and more helpful contributions (Wasko & Faraj, 2005). In an offline context, people who define themselves as more knowledgeable and more invested in their local neighborhood wrote more reviews on social media sites than those who did not identify themselves as local experts (Antin, Sa, & Churchill, 2012). As a result, even though the personal benefit of contribution to the online community can be smaller for a central member, because of their stronger sense of attachment to the community, the perception of community benefit can be higher for them. This line of argument leads us to a competing hypothesis that central people are more likely to contribute because they have stronger attachment to the community and they care more about the community. Thus, an alternative hypothesis is:

H1b: Central offline community members are more likely to decide to participate in an audience-bounded online community than peripheral community members because of their stronger attachment to the community.

Data Collection

To test our hypotheses, we analyzed participation patterns of users in an online community for academic conferences. More specifically, we conducted an analysis of users' behavior online in the Conference Navigator (CN) system during two conferences: the User Modeling, Adaptation and Personalization Conference (UMAP) 2011 and 2012.

Conference Navigator (Farzan & Brusilovsky, 2008) is an online community designed to help conference attendees make decisions about which talks to attend. CN users can schedule talks they want to

attend at the conference, and they can tag their own papers with relevant keywords to increase the visibility of their papers within the system. Based on user-generated content, CN guides conference attendees by augmenting the conference schedule with meta-information about talks such as popularity and by providing personalized recommendations of relevant talks to each user. Additionally, the users can connect to other researchers in the field through the system. Logging into the system is necessary to schedule talks, receive recommendations and connect with other researchers.

The system was initially introduced to the authors solely. The conference chair invited all the papers' authors to preview the system before the conference. Two invitation emails were sent to each author. The first invitation letter invited all authors to schedule their five most interesting papers in the conference. The authors were told that this information will help the conference chairs to plan the schedule of the conference by identifying clusters of relevant papers and trying to schedule them for a single session during the conference. This first message was sent ten weeks before the conference. Six weeks later, after the conference schedule was released, the conference chair sent another email to the same list of authors. The authors were invited to check their papers' information and inform the system administrator if they found any error in the paper and presentation details. We collected the data of the authors' behavior in the system during the interval of time from the first email until a week before the conference, when other conference attendees were invited to use Conference Navigator.

To model the offline social structure among the conference attendees, a network of co-authorship was created by using the data about the papers published in the UMAP Conference series from 2001 through 2011. Co-authorship relationships are defined when a pair of authors has written a paper together for UMAP series of conferences. These relationships are weighted based on the number of papers two authors have written together, number of papers by each of the authors, and number of co-authors they had in common. We employed the method described by (Liu, Bollen, Nelson, & Sompel, 2005) to calculate these weights. We used Gephi tool² to calculate network measures including clustering coefficient and degree centrality. The network of authors includes 491 authors. The average degree centrality is 4.74, the average clustering coefficient is 8.81, and the average path length is 2.46.

Data Analysis

We analyzed how the UMAP authors' network properties were associated with their decision to participate in Conference Navigator. We used logistic regression to model these relationships. The dependent variable is a binary variable: "1" for the authors who participated in the CN system at least once and "0" for those who never logged into the system. The independent variable is the centrality of the author in the academic network.

It is possible that first authors perceive the system to be more beneficial as they may view the system as a medium to propagate information about their paper. This is especially important if our first hypothesis is true and personal gain is the main motivational factor for participation in the system. To account for that, we controlled for the fact that the authors were a first or a secondary author of a paper in the conference.

Beyond network centrality, degree of connection in users' neighborhood can influence their perceived benefits of the system. If community benefit is what most motivates users, users whose neighborhoods are less connected may perceive that the system can help to increase neighbors' awareness about each other and help them to connect to each other. To account for that, we also controlled for the level of connection among the author's neighborhood by including the clustering coefficient.

Therefore, the covariates included in the model are:

- *Author Role* indicates if the user is a first author or secondary author of a paper in the conference.
- *Weighted Degree* measures the users' centrality in the academic network.
- *Clustering Coefficient* measures the connectedness the users' neighborhood. A fully-connected neighborhood has a value of "1" and a completely disconnected neighborhood has a value of "0".

Table 1 and Table 2 illustrate the descriptive statistics of our data. Among the 491 authors, 16% participated in the system (78 authors: 51 first authors and 27 secondary authors, see Table 1).

² <http://gephi.org/>

Table 1
Descriptive Statistics of Covariates

Type of author	Participation		Total
	No (ratio)	Yes (ratio)	
First	107 (.68)	51(.32)	158
Secondary	306 (.92)	27(.08)	333
Total	413 (.84)	78(.16)	491

As shown in Table 2, the mean weighted degree among first authors is 1.67 and among secondary authors is 1.41. The data suggests that first authors and secondary authors are not different in terms of centrality in the community. Regarding the clustering coefficient, the mean value among first authors is .72; for secondary authors, the mean clustering coefficient is .87.

Table 2
Authors' Participation in Conference Navigator

Covariate	Kind of author	Mean	Std.Dev.	Min	Max
Weighted degree	First	1.67	1.32	.99	11.25
	Secondary	1.41	1.15	.99	7.79
Clustering Coefficient	First	.72	.41	0	1
	Secondar	.87	.29	.0	1

Table 3 describes the results of the logistic regression. Members with higher centrality value in the offline network are significantly more likely to respond to our invitation to participate in CN after controlling for the other covariates ($p < .0001$). One unit increase in the weighted degree increases the chance of participation 1.4 times ($e^{.348}=1.4$). This result supports our hypothesis H1b that community benefit for central members is stronger motivating factor than personal benefit for peripheral members.

Keeping the network features constant, the first authors are 4.9 times ($e^{1.587}=4.9$) more likely to participate than secondary authors ($p < .0001$). This follows the general idea of cost-benefit analysis of participation that, independent of network centrality, the benefit of participation can be higher for first authors. They can use online networks as a medium to learn more about a community they care about and to promote their position in that community.

Members whose neighborhood is less connected are more likely to participate. However, this effect is only marginally significant.

Table 3
Regression Results - Authorship- Activity

	B	S.E.	z	Sig.
First Author	1.587	.275	5.76	.000
Weighted Degree	.348	.091	3.80	.001
Clustering Coefficient	-.595	.355	-1.67	.094
Constant	-4.099	.603	-6.80	0.000

n=491, non-zero=78, $p < .0001$

Overall and individual fit analyses were conducted to assess the robustness of these results. The Hosmer-Lemeshow goodness-of-fit test shows that there is no evidence of lack of fit of the model ($p = 0.43$). Some potential outliers were identified using the standardized Pearson residuals scores, Pregibon leverage and delta-beta statistics analyses. After removing these potential outliers, the interval confidence of the effect of first author fell below the positive range (lower than 1). However, centrality remained a significant predictor of the likelihood of participation in the online community. This suggests that the model is robust to outliers.

We observed the same result in a logistic regression to predict the likelihood to contribute (i.e. schedule a talk, tag a talk, or connect with others). More central members and first authors were more likely to contribute. We also analyzed the relationship between network centrality and contribution or participation levels in the system for those who logged in at least once. We conducted a robust Poisson regression with the number of contributions and number of actions as outcome measures with the same predictors. None of the models was significant, which may be due to the fact that a small percentage of users logged into the system.

Discussion and Limitations

This paper reports a study that aims to understand the relationship between centrality in an existing offline community and the likelihood of participating in a related online community. Our results support the hypothesis that higher centrality in the offline community significantly increases the odds to decide to participate in the online space for the community. This finding proposes that designers deploying an audience-based online community would benefit from understanding the social structure of the targeted offline community. Identifying central people and strategically engaging them earlier in the online conversation could be an effective user engagement strategy. Engaging more central people early can help to create enough initial content that can be later used to engage more peripheral members. Members that are less central to the community may need more reasons to join the system. Having valuable content contributed by central members and an active stream of activity can make the system more valuable for them.

We are interested in understanding why more central community members are more likely to join the online system. We argue that central community members are more committed to the community; therefore, they are more likely to embrace a new initiative that can provide benefits to the community as a whole. It is also possible that more central members are more comfortable publicly contributing to the system because they know the audience better. Although these are plausible arguments, further investigation is necessary to confirm these speculations. Our future work will focus on collecting qualitative data to examine what motivates offline central members in the UMAP community to participate in an online community.

Additionally, we will attempt to replicate the study within larger conferences. The current study focuses on a small-size community featuring a highly connected offline structure. Compared with well-known average of six-degrees of separation in random communities, the UMAP community has an average path length of 2.46; i.e. this community is more strongly connected than a random community. Therefore, our results may not be directly generalizable to other social systems that aim to support bigger and more loosely-connected social networks. Therefore, we believe that replication of this study in larger conferences or other offline communities are necessary to make our results more robust.

Despite the limitations, we believe that our research provides a gateway to enhancing users' engagement strategies in audience-based online communities.

References

- Antin, J., Sa, M. de, & Churchill, E. F. (2012). Local experts and online review sites. In *Proceedings of the acm 2012 conference on computer supported cooperative work companion* (pp. 55–58). New York, NY, USA: ACM. Available from <http://doi.acm.org/10.1145/2141512.2141541>
- Brass, D. J. (1984). Being in the right place: A structural analysis of individual influence in an organization. *Administrative Science Quarterly*, 29(4), 518–539.
- Burt, R. S. (2001). Attachment, decay, and social network. *Journal of Organizational Behavior*, 22(6), 619–643. Available from <http://dx.doi.org/10.1002/job.106>
- Butler, B. (2001). Membership size, communication activity, and sustainability: A resource-based model of online social structures. *INFORMATION SYSTEMS RESEARCH*, 12(4).
- Carroll, J. M. (2012). The neighborhood in the internet. In (pp. 23–47). Routledge.
- Cotten, S. R., & Gupta, S. S. (2004). Characteristics of online and offline health information seekers and factors that discriminate between them. *Social Science & Medicine*, 59(9), 1795 - 1806. Available from <http://www.sciencedirect.com/science/article/pii/S0277953604000796>
- Cummings, J. N., Butler, B., & Kraut, R. (2002, July). The quality of online social relationships. *Commun. ACM*, 45, 103–108. Available from <http://doi.acm.org/10.1145/514236.514242>

-
- DiMicco, J., Millen, D. R., Geyer, W., Dugan, C., Brownholtz, B., & Muller, M. (2008). Motivations for social networking at work. In *Proceedings of the 2008 acm conference on computer supported cooperative work* (pp. 711–720). New York, NY, USA: ACM. Available from <http://doi.acm.org/10.1145/1460563.1460674>
- Farnham, S. D., Brown, P. T., & Schwartz, J. L. K. (2009). Leveraging social software for social networking and unity development at events. In *Conference on communities and technologies* (p. 235-244).
- Farzan, R., & Brusilovsky, P. (2008). Where did the researchers go?: supporting social navigation at a large academic. In *Proceedings of the nineteenth acm conference on hypertext and hypermedia* (pp. 203–212). New York, NY, USA: ACM. Available from <http://doi.acm.org/10.1145/1379092.1379131>
- Harper, F. M., Li, X., Chen, Y., & Konstan, J. A. (2005). An economic model of user rating in an online recommender system. In *Proceedings of the 10th international conference on user modeling* (pp. 307–316). Berlin, Heidelberg: Springer-Verlag. Available from http://dx.doi.org/10.1007/11527886_40
- Liu, X., Bollen, J., Nelson, M. L., & Sompel, H. Van de. (2005, December). Co-authorship networks in the digital library research community. *Inf. Process. Manage.*, 41(6), 1462–1480. Available from <http://dx.doi.org/10.1016/j.ipm.2005.03.012>
- Raban, D. R., Moldovan, M., & Jones, Q. (2010). An empirical study of critical mass and online community survival. In *Proceedings of the 2010 acm conference on computer supported cooperative work* (pp. 71–80). Available from <http://doi.acm.org/10.1145/1718918.1718932>
- Resnick, P., Janney, A., Buis, L. R., & Richardson, C. R. (2010). Starting an online community on demand: A case study of adding forums to a physical activity promotion program. *Journal of Medical Internet Research*.
- Schuler, D. (1994, January). Community networks: building a new participatory medium. *Commun. ACM*, 37(1), 38–51. Available from <http://doi.acm.org/10.1145/175222.175225>
- Schwabe, G., & Prestipino, M. (2005). How tourism communities can change travel information quality. In *Ecis* (p. 1685-1696).
- Wasko, M. M., & Faraj, S. (2005). Why should I share? examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, 29.