Investigating recognition-based performance in an open content community: A social capital perspective

Chitu Okoli a,*, Wonseok Oh b

a John Molson School of Business, Concordia University, 1455 de Maisonneuve Blvd. West, Montréal, QC, Canada H3G 1M8
b Desautels Faculty of Management, McGill University, Montréal, Canada

Received 24 March 2005; received in revised form 10 May 2006; accepted 21 December 2006

Abstract

As the open source movement grows, it becomes important to understand the dynamics that affect the motivation of participants who contribute their time freely to such projects. One important motivation that has been identified is the desire for formal recognition in the open source community. We investigated the impact of social capital in participants’ social networks on their recognition-based performance; i.e., the formal status they are accorded in the community. We used a sample of 465 active participants in the Wikipedia open content encyclopedia community to investigate the effects of two types of social capital and found that network closure, measured by direct and indirect ties, had a significant positive effect on increasing participants’ recognition-based performance. Structural holes had mixed effects on participants’ status, but were generally a source of social capital.

Keywords: Social networks; Social capital; Open content; Open source software; Recognition-based performance; Social status; Virtual communities

1. Introduction

The open source movement is a puzzle to those who are more comfortable with the traditional economic model of software development, where a large fixed cost of development is offset by mass distribution under strictly controlled intellectual property licenses. Obviously, participants in open source projects are not economically motivated [52]. As the movement continues to grow, it becomes important to examine the social fabric of the communities to better understand how and why they “work”.

A number of studies have considered what incentives motivate participants in open source projects [23,24,32]. One highlighted the participants’ desire to be recognized in their virtual community of open source participants. This occurs when a participant receives informal praise and acknowledgment from their fellows, and also by tangible tokens such as recognition in the open source communities by the granting of administrative, or “insider”, rights that permit high-quality contributors to add their changes directly with minimal prior review, and to have an important role in deciding the direction of the project. Thus the action is equivalent to promoting an employee to manager status in a traditional organization and thus is a recognition of achievements.

Our study therefore investigated how social capital [8,15,21,35] in the community influenced an OSS participant’s recognition-based performance. Because

* Corresponding author. Tel.: +1 514 848 2424; fax: +1 514 848 2824.
E-mail address: Chitu.Okoli@concordia.ca (C. Okoli).

0378-7206/$ – see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.im.2006.12.007

open source communities operate under social structures [12], social capital theory provided a valuable perspective for understanding how participants lever-aged social resources to achieve their performance objectives. More specifically, we asked: Do the virtual connections through which open content participants operate influence their performance?

We addressed this question by examining the activities of Wikipedia (http://en.wikipedia.org, a general encyclopedia), currently the largest open source content community, because its members span every aspect of the human demography and anyone with Internet access may access and contribute to it.

2. Open source software and social capital

Though open source activities have been practiced since the mid-1980s, when the Free Software Foundation was founded, little research has been attempted to try to maximize its benefits.

2.1. Open source software (OSS)

The emergence of open source software has changed the technological landscape of the computing industry, affecting the strategic dynamics of some commercial enterprises, including interactions among proprietary software developers, hardware manufacturers, and makers of network products [51]. OSS has been defined as “software that is made freely available to all”; it is often developed and maintained by many contributors scattered around the world, but they interact through a virtual community. The majority of these public-good producers contribute freely to the project in return for only intrinsic rewards.

However, “freedom” has a particular meaning in open source contexts: “Free [open source] software is a matter of the users’ freedom to run, copy, distribute, study, change and improve the software” [48], as opposed to freeware, which is software provided at no price. An OSS license often permits sale of the product or a system developed from or with it, with the caveat that the buyer be informed of his or her right to full access to the source code at no charge, instead of, or in addition to, paying for compiled executable programs.

2.1.1. Motivations for participating in OSS projects

Research on OSS has mainly focused on dimensions such as the economic and sociological foundations, motivational issues of participants, and its managerial and legal aspects. Von Hippel and von Krogh consider OSS developed a “private-collective” innovative model, and illustrated how it deviated from private investment and collective action models. They believed that the “private-collective” model offered “the best of both worlds” by resolving problems such as revelation of innovations developed by private funds and the free-riding phenomenon. From the perspectives of labor economics and industrial organization, Lerner and Tirole examined OSS economics and argued that signaling incentives (career concern incentives, future job opportunities and ego gratification) were the main drivers of the volunteers’ participation. These authors also discussed the importance of leadership due to the unique management structure of open source organizations and the competitive dynamics between open and commercial software development.

Drawing from the extended Klandermans model of voluntary action [28], Hertel et al. explored the factors that motivated the participation of volunteers. Some of these factors, based on a Web-survey of 141 participants, included pragmatic (e.g., improving the Linux kernel for personal advantages), social (i.e., supporting the work with other volunteers), and hedonic motives (i.e., intrinsic motivations). Bonaccorsi and Rossi [5] discussed several economic issues, such as motivation, coordination and diffusion from the perspectives of collective action, governance structure, and network externalities, respectively.

In open source projects, it is necessary to consider the performance of the participating developers, since most are volunteers with individual motivations for participation. Aligning performance measures with motivations would be helpful to organizers and evaluators of OSS projects. According to the resource-based model of social structure sustainability, online social structures should provide members with positive benefits, such as the opportunity to be influential [55], to affiliate or champion [37], and the ability to disseminate ideas rapidly [27]. While not absolutely necessary for participation, it is very important for open source participants to be recognized for their contributions [48]; thus they gain status and respect in the community. To better encourage participation in open source projects, research is needed to understand the dynamics of recognition as an incentive, and how the attainment of this motivation serves as an indication of a participant’s performance, in the sense that they have achieved a desired individual goal.

2.1.2. OSS and open content encyclopedias

An important feature of OSS communities is that there is limited opportunity for participation for those who are not capable programmers: non-programmer
members typically do not have the same status (such as voting rights on important design issues). OSS projects form virtual communities in which people interact to achieve a common goal; the communities have power structures, norms, values, and traditions. However, most are somewhat narrow in their scope of contributors [30], being limited to skilled computer programmers.

We examined a particular type of OSS whose characteristics draw from an almost universal spectrum of people—the open content encyclopedia (OCE), which allows almost anyone to contribute to it, and that permits modification and distribution of its content. Quality is loosely maintained by applying Linus' Law [38]: “Given enough eyeballs, all bugs are shallow”. With thousands of contributors, erroneous and poor quality content is removed and only good quality content is retained. Wikipedia is by far the most extensive and best known, but there are others, such as ISPedia and the Enciclopedia Libre Universal en Español. OCEs are based on the Wiki concept, a Web technology that allows viewers to add content to Web pages using minimal technical skills. OCEs are released with a license that permits the free redistribution and modification of their content: Wikipedia uses the GNU Free Documentation License, the textual complement to the GNU General Public License for traditional OSS.

OCEs are structured to eliminate or minimize individual agendas and strive towards public welfare in setting policies for the development of their products. Their organizational structure is somewhat anarchistic, but there is some central administration. However, by employing Wiki technology, OCEs remove the technical barrier to access that exists in traditional OSS communities. Participants only need to be computer literate and have access to the Internet. A new mechanism related to the development of OCEs is being spearheaded by the creative commons (CC, http://www.creativecommons.org), a resource that creates licenses on demand for literary, audio, and video works; it allows creators of content to choose among several groups of intellectual property rights under which they can license their works.

3. Social capital and social networks

The principle idea of social capital theory is simple and straightforward: social relationships, norms, and values [36] attached to social capital determine the performance of individuals, groups and organizations who are part of a socially or economically connected network [31,56]. While “performance” is sometimes conceptualized in objective terms, such as salary and promotions, it is more often measured in terms of subjective recognition-based measures such as organizational influence, and peer acknowledgement; this corresponds with our focus on recognition-based performance in OCEs. Many studies have examined how individuals invest in social relations and how they access and exploit the resources [40]. But many others have placed a greater emphasis on group-level issues related to social capital. Bourdieu and Wacquant [7] and Putnam [42] studied the process by which groups generated and maintained social capital as a collective asset, illustrating how the collective asset mutually benefited the group members: see Lin for a detailed review of these issues. Despite variation in the level of analysis, researchers have viewed social capital as “institutionalized social relations with embedded resources” which can benefit both the collective and the individuals in the collective.

Tsai and Ghoshal [49] explored the structural, relational and cognitive dimensions of social capital, and found that the first two significantly influenced product innovation. Ahuja [1] examined three aspects (direct ties, indirect ties and structural holes), of a firm’s ego network (the personal network with the individual as its focal point), in order to determine how a firm’s inter-organizational network structure affected the degree of subsequent innovation: both direct and indirect ties were found to positively influence innovation. Interestingly, however structural holes were negatively associated with innovation. In a study of biotechnology start-ups, Shan et al. [45] and Powell et al. [41] examined the relationship between a firm’s network position and organizational performance. They found that the number of ties [47] and centrality were significantly related to innovation and faster growth, respectively.

3.1. Social capital

Social capital has been defined in many ways, such as “the sum of the resources, actual or virtual, that accrue to an individual or group by virtue of possessing a durable network of more or less institutionalized relationships of mutual acquaintance and recognition” but despite differences in the contextual application, they coalesce around a central orchestrating theme that reflects the criticality of the relations. Social capital theory places emphasis on social relationships.
3.2. The two types of social capital

Although social capital is considered an important element of a person’s success, there are two schools of thought about the mechanism by which social capital should be produced and mobilized. Fig. 1 graphically illustrates the difference. By bridging various network segments, Robert plays an important “linking-pin” role, whereas James’s role is limited to heavy interaction with other people in the same sub-network. Advocates of the closed, dense or cohesive network view, such as Coleman and Walker et al. [54], have argued that closure or density of social relations is the primary ingredient for the generation of social capital. They have given empirical evidence to support their perspective (e.g., a successful operation in the labor market, and the formation of informal credit associations where pooled savings are allocated on a rotating basis [33]). They argue that such network formations provide members with solid grounds that allow for the maintenance of cohesion and preservation of trust, authority, and norms; that a cohesive network can provide an infrastructure that facilitates smooth coordination and cooperation from member to member in the pursuit of their interests while reducing uncertainties [43] and increasing trust.

In contrast, Granovetter and Burt who advocate the structural hole view, assert that “true” social capital can be efficiently produced and maintained under open or loosely coupled networks, in which members can access the resources available in heterogeneous sub-networks. They have stated that loosely developed networks rich in structural holes allow the individuals to access and mobilize social resources and that bridges, structural holes or weaker ties are the building blocks necessary to construct or configure the network and thus produce “fresh” information while “social redundancies” are minimized. Thus the diversity and uniqueness of information become the crucial aspects of social capital, and they require a network rich in structural holes [18]. Structural hole theory posits that closed networks hinder, rather than promote, organizational coordination [9]. Portes and Sensenbrenner claimed that instead of an asset, the social capital embedded in closed networks could become a social liability that hinders the organization’s performance.

The two views have some similarities, however. Both presume that reciprocity is the primary driving force in creating and reproducing social capital. They also agree that cohesive relations promote and amplify reciprocity. Nonetheless, these perspectives maintain a parallel stance with respect to the effect of reciprocity on a member’s coordination. The closed network theory posits that such reciprocity is a necessary component for assuring group cohesion, while the structural hole framework considers it to be structural arthritis [9], which deters coordination among group members.

Several sociologists [2] have attempted to resolve the conflict by stating that an explanation for the conceptual gap is truly contingent on what outcomes are expected by its members, and under what conditions they mobilize their network resources. Lin suggested, for example, that cohesive networks were more effective for “preserving or maintaining resources,” while “searching for and obtaining resources” can be better managed under the open network formation rich in structural holes. In short, the effectiveness of a network structure, in terms of creating social capital would depend on either promoting the use of current resources or by acquiring new resources.

We predicted that better performance and recognition in an open content community required individual volunteers to “occupy” a network location that could leverage both types of social capital to their as much as possible. Consistent with Ahuja, we consider these types of social capital as complementary resources. Members of open content communities can only perform well and move forward if solidarity exists among the participants.

3.2.1. Effects of network closure

A closed network structure in which many connections link the ego to its alters (directly or indirectly) provided a solid basis on which members ensure the trustworthiness and authority of other members: the direct ties that have been created and maintained by a participant provide four benefits that help improve overall performance: knowledge sharing, complementarity, quality control, and conflict resolution. A dense

![Fig. 1. The structural hole (adopted from Burt [11]).](image-url)
network facilitates knowledge sharing between the participants [4], a critical asset for their performance in the open content project. In addition, intense interactions and collaboration among individuals promotes norms and trust and facilitates complementary skills [44]. Open content initiatives often require a great deal of complementarity because of their complexity.

In addition to these benefits, closed networks may serve as an effective mechanism to control the quality of member contribution and resolve potential conflicts of interests. Such a network formation structurally discourages opportunistic behavior and shirking among its members. Those densely tied tend to improve their performance, which facilitates quality control. Moreover, they are aware that their performance will likely be monitored by the alters, and thus they are more likely to feel responsible and make frequent contributions.

Finally, a cohesive network rich in direct and indirect ties may facilitate the resolution of any interpersonal conflicts. Individuals who interact more frequently tend to agree on issues [29]. However, conflicts do arise in an open content project because the participants may have different background, specialty, and interests.

In the case of Wikipedia, one of the important criteria that members consider in voting for an administrator is his or her ability to work well with others in face of conflict. Thus, we expected that network density was important in encouraging members to amicably resolve disagreements with long-term colleagues.

Hypothesis 1. The greater the closure of a participant’s network in an open content project, the higher the performance of the participant.

3.2.2. Effects of structural holes

Although a closed network is a necessary component for members, this may not be sufficient to elicit their optimal performance. Participants who forge ties over an extended time without being subject to “external” review are likely to suffer from knowledge deficiency and experience a cognitive lock-in [20,50]. Consequently, performance would be limited by a lack of new ideas. Moreover, relational inertia, which prevents members from forming new ties, could occur in the community with strong bonds.

In an open source community, cognitive lock-in and relational inertia may pose a great threat to the development of open content initiatives; contributors who establish direct links and make frequent interactions are likely to be “trapped in their own network” and pay less attention to the issues of those outside their network, preventing the entry of new members with fresh ideas. Consequently, an open-network formation, which allows brokerage opportunities, provides the constituents with the platform to interact dynamically with many others, and to gain new knowledge and insights. A Wikipedian interested in adminship rights can become more “influential” in the community by playing a bridging role between otherwise separate subnetworks. By occupying such a linking positions, he or she can increase control over the information.

Hypothesis 2. The more structural hole capital a participant controls in an open project, the higher the performance of the participant.

4. Methodology

4.1. Data source

As the database for our study, we downloaded a full replication of the English-language version of Wikipedia as of 16 June 2004. This 30 GB MySQL database included the edits of over 80,000 Wikipedians (not counting the tens of thousands of anonymous contributors), representing over 300,000 articles and 2,000,000 edits. However, a network of 80,000 people is much too large for practical analysis, and the vast majority of the material would not have been relevant for our research questions. Therefore, we decided to reduce the participant list to less than 500 Wikipedians.

The components of a social network are the nodes and the ties between them. In our work, the nodes were the participants, and the ties were completed when two Wikipedians interacted when creating an article. In the Wikipedia process, different participants create articles by generating or making edits to an article page through the Wiki interface. Each article has an associated “talk page” on which participants can discuss the article, ask questions of other contributors, and thus resolve conflicts. These “talk pages” are not real-time chat rooms, but living, documented discussions. They are the primary way of communication for the virtual community of Wikipedia participants, and are used to forge a feeling of community among members [39]. In our extensive observation of the use these pages, we found that those who write comments generally read what has already been written in order to understand the state of the discussion before they make any comments themselves. Unlike newsgroups or forums, each article has a single talk page that records all the discussion; thus it is easy for contributors to access the full discussion history. However, for extremely lengthy discussions, older comments can be archived.

Discussions on talk pages are open interactions, since contributors read and respond to each other by name. Thus, we noted a tie when two Wikipedians had written comments on the same talk page. In our study, we examined the interactions of nearly 500 Wikipedians and used their mutual editing of the main article talk pages as ways to determine the ties between them. We used queries in MySQL to compile the links, and restructured them appropriately using Microsoft Access and Excel. We weighted all ties equally: 1 if it existed, and 0 if not.

4.2. The dependent variable: administrative rights in Wikipedia

We wished to investigate the effects of social network structure on the performance of open source participants and conceptualize participant performance as the recognition of their efforts in the open source community: in our study this was measured by the increased status that attained by achieving a formal designated title: in the Wikipedia organizational unit, this is accomplished when participants are promoted to increasing levels of administratorship (adminship).

Wikipedia permits anyone, to edit any article and to create new articles. Participants who decided to “sign-in” can also move articles. Any Wikipedian can request administrative rights, which, according to policy, is normally accorded liberally to anyone who demonstrates that they are responsible, contributing members of the community. Adminship gives the additional rights to delete articles permanently, and to undo the revisions of the vandals who deliberately efface articles. The administrators can also protect pages from editing when a serious problem arises, and even discipline unruly Wikipedians by temporarily blocking them from access to Wikipedia. Thus promotion gives participants both senior-editor and police status.

Many contributors, however, do not seek out such privileges they: just want to contribute. Their reward is apparently the pleasure gained from contributing their knowledge and sharing in the creation of a valuable resource. Thus, we had to try to distinguish between contributors interested in administrative powers and the adminship process, and those who merely contributed. Wikipedians become administrators by a vote open to all community members (though few actually participate), after all have had the opportunity to discuss and review a nominee’s contribution history.

Beyond adminship, there are a number of super-administrator ranks. A bureaucrat has all administrator privileges and can appoint or depose administrators, The highest rank is that of steward, which has full rights to access the entire Wikipedia database and make changes to any Wikipedians’ rights.

In our study, our dependent variable was the Wikipedians’ performance in terms of formal rights and responsibilities accorded in recognition of their contribution to the community. However, we observed that many Wikipedians were not interested in adminship. Thus, we selected only those Wikipedians who had shown an interest in the adminship process—indicated by their contribution to the page “Requests for adminship,” which included all those who had ever been nominated (including those self-nominated), had commented on nominees, had contributed to the formulation of adminship policy, or had made any other comment on either the “Requests for adminship” project page or on its associate talk pages. Through this process, 512 Wikipedians were identified. However, of these, only 465 participated in talk pages of English Wikipedia articles, so this was our population. Of course, this sample does not represent the whole set of Wikipedia contributors.

Our dependent variable consisted of four consecutive categories that covered the 465 subjects: there were 181 regular Wikipedians without administrative rights, 257 with simple administrative rights, 18 bureaucrats, and 9 stewards. Because of the distribution of this variable, we could group the sample into either two categories (181 regular Wikipedians and 284 administrators) or three categories (181 regular Wikipedians, 257 simple administrators, and 27 super-administrators). We used both categorizations in our analysis.

4.3. Independent variables: measures of social network structure

We used UCINet [6] to compute social network variables for structural holes and network closure, obtained from the social network data generated from the MySQL queries. Building on Ahuja’s research framework, we used direct and indirect ties to represent the network closure of each ego’s network. This counted the number of direct partners of the focal firm and the number of other firms to which it was tied (at a path distance of two or greater) in order to operationalize the constructs of direct ties and indirect ties, respectively. To compute a count of direct ties, we assessed the number of direct collaborators of each participant in the sample. The number of indirect ties was based on the number of other collaborators in a particular Wikipedia network to which it was tied at path distances of two or greater.

Network size, density, and hierarchy, on the other hand, represented the structural holes in each ego’s network, along with the constraint as a composite index. To derive the numerical indices for each component, we used the procedures specified in Burt’s article [10] (Table 1): he had used these three components as separate independent variables in an effort to explore the gender differences in social capital.

Network size referred to the number of contacts in a network. An inverse relationship existed between the degree of network constraint and the size of a network because any participant’s relational intensity in terms of time and energy with one particular collaborator diminished as the number of contacts increased Burt argued that network size was positively associated with performance because individuals attached to a network can obtain diverse information as the number of contacts increases. Similarly, a Wikipedian’s performance (in terms of receiving peer recognition) increased as the size of the network increased. Density indicated the average strength of connection in an ego network. If a Wikipedian formed a dense network by working closely with only a limited number of collaborators, he or she would not be “globally” recognized.

Finally, hierarchy denoted the extent to which node connections were concentrated in one central contact. Burt operationalized hierarchy through the use of the Coleman-Theil disorder index [14], which was measured by the sum of the average level of constraint multiplied by its natural logarithm, quantity divided by the maximum sum possible:

$$\sum_j (c_{ij} / (C/N)) \ln(c_{ij} / (C/N)) / N \ln(N)$$

where $C_{ij}$ is the degree of constraint for $i$ posed by contact $j$, $C$ the sum of constraint across all $N$ relationships, $N$ the number of contacts in an ego network and $C/N$ is the mean level of constraint per contact.

In our study, $C_{ij}$ represented the degree of constraint to Wikipedian $i$ posed by Wikipedian $j$ due to their hierarchical relationship. For example, a participant’s opportunity to receive a wide peer-recognition decreased when

1. he or she committed a large investment of time and energy in reaching Wikipedian $j$ and
2. $j$ had few structural holes. Indeed, Wikipedians $i$ and $j$ may have interacted with a similar pool of “local” Wikipedians, thereby severely constraining $i$’s capacity to expand globally.

### 4.4. Control variables

The two major considerations given when voting for adminship are the quality of a Wikipedian’s contribution, and his or her ability to judiciously and calmly resolve conflicts with other Wikipedians, especially

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Social capital</strong></td>
</tr>
<tr>
<td>Component</td>
</tr>
<tr>
<td>Network closure</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Structural holes (components)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

when they are themselves involved in the conflict. Unfortunately for our analysis, these are highly qualitative measures, and it was impractical to analyze the hundreds of thousands of edits that the selected Wikipedians had made and then to try to code these values. In lieu of qualitative control variables, we tested the effects of a number of quantitative proxies in association with our social network variables to test their effect on the Wikipedians’ performance; specifically we tested the effects of the:

- total number of edits that a Wikipedia had made to articles;
- length of membership as measured from the date of his or her first article contribution;
- average number of edits made, as was the ratio of the preceding two measures;
- number of distinct articles to which the Wikipedia had contributed.

We used MySQL queries to obtain all these values from the Wikipedia database. Although not a proxy of quality, we also calculated the number of other Wikipedians among our 465 with whom the focal Wikipedian had interacted on an article talk page.

We also adapted a proxy quality measure from Lih, who used the ratio of the number of edits to an article to the number of unique contributors as a measure he termed the “rigor” of an article, indicating the number of individuals who worked on (and peer-reviewed) it. As an analogous measure, we computed the ratio of a Wikipedia’s total number of edits to the number of distinct articles to which the Wikipedia had contributed as the “degree of focus” of the Wikipedia in contributing to articles.

5. Data analysis and results

Our dependent variable was categorized as being split into three consecutive classes, or in two. We tested the two-category form (regular Wikipedians, administrators) using logistic regression, which was an appropriate analytical technique with continuous independent variables and a binary dependent one [22]. (The use of multiple discriminant analysis would have been inappropriate for this analysis, as it requires continuous independent variables and a categorical, non-ordinal dependent variable—the inverse of ANOVA.) For the three-category form (regular Wikipedians, simple administrators, and super-administrators), we tried to use both multinomial ordinal logit and multinomial ordinal probit regression procedures; these are appropriate for data that is multinomial (more than two values) or ordinal (in successive rank order of magnitude) but not interval (values between data points are unevenly spaced). However, the results with such analyses were virtually identical to those achieved when we used multiple linear regression (which assumes interval data—that the distance between a regular Wikipedia and a simple administrator is the same as the distance between a simple administrator and a super-administrator). Because linear multiple regression was simpler to report and interpret, we will only discuss the results of the linear multiple regression and logistic regression here. However, we have shown the results of the ordinal multinomial logit regression in a table for comparison.

For our analysis, we began by creating a base model that consisted of the control variables that had significant effect on the dependent variable: the rights. After regressing total edits, length of membership, average edits, number of distinct articles, and the “degree of focus” measure, we found that only the length of membership and the total number of edits together provided a significant model. We only show the results of this final base model in Table 2.

The negative effect of the first edit was expected, since the earlier the date of a Wikipedia’s first edit, the longer he or she has been a member, and his or her status would be expected to be higher. Also, the greater the total number of edits, the higher would be the participant’s status. The log likelihood (−2LL) is an estimation of the goodness of logistic model fit, with lower values being better. Conversely, higher $R^2$ values for linear regression indicate better model fit. −2LL is $\chi^2$-distributed with the degrees of freedom equal to the number of independent variables in the model. The $\chi^2$ of the base model (103.9 for $-2LL = 517.7$ at 2 degrees of freedom) gave us a benchmark against which to compare our social network models.

A regression of the network closure variables on admanship with the control variables indicated a high degree of multicollinearity among the two (variance inflation factor [VIF] for direct ties was 34.34 and indirect ties was 34.06, both much larger than the recommended maximum of 4.0–5.0), resulting from their high (0.985) correlation. Thus, we regressed each of them separately. In Table 3, we only show the regression for direct ties, as the results for indirect ties are almost identical. Both linear and logistic regressions showed that network closure as expressed by ties was statistically significantly related to admanship ($p < 0.001$ for direct ties). The effect was positive, indicating that the greater the closure of a Wikipedia’s personal network, as measured by direct (or indirect) ties, the more likely it...
The linear model had an adjusted $R^2$ of 0.242, and the logistic model had a $-2LL$ of 490.6. We tested whether this model added any explanatory power beyond our base model by comparing the $\chi^2$ of the two. The difference in $\chi^2$ was $131.0 - 103.9 = 27.1$, which is statistically significant ($p < 0.001$) at 3 – 2 = 1 degree of freedom (the difference in degrees of freedom of the two models). This confirmed that the network closure model did indeed explain Wikipedians’ adminship status beyond the base model of control variables.

6. Discussion and implications

We found that, both direct and indirect ties were positively related to the administrative rights accorded in the OSS community. The number of direct connections of a participant with others indicated those with whom the Wikipedia had discussed. Discussion indicated that the contributors to the talk page actually associated comments with user names, which enabled them to know and respect each other. Among the Wikipedians, direct and indirect ties were highly correlated ($r = 0.985$).

Table 3

The impact of network closure on performance

<table>
<thead>
<tr>
<th>Adminship</th>
<th>Linear regression</th>
<th>Logistic regression</th>
<th>Ordinal logit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef. $p &gt;</td>
<td>r</td>
<td>$</td>
</tr>
<tr>
<td>First edit</td>
<td>0.000 0.000***</td>
<td>-0.002 0.000***</td>
<td>-0.002 0.000***</td>
</tr>
<tr>
<td>Total edits</td>
<td>0.000 0.000***</td>
<td>0.000 0.000***</td>
<td>0.000 0.000***</td>
</tr>
<tr>
<td>Direct ties</td>
<td>0.002 0.000***</td>
<td>0.006 0.000***</td>
<td>0.007 0.000***</td>
</tr>
<tr>
<td>Constant</td>
<td>14.8 0.000***</td>
<td>59.4 0.000***</td>
<td>665.0 0.000***</td>
</tr>
</tbody>
</table>

$R^2 = 0.25$; adjusted $R^2 = 0.24$; $p > F = 0.000***$

$-2LL = 490.6; \chi^2 (3 \text{ d.f.}) = 131.0$;

$p > \chi^2 = 0.000***$

Table 4
The impact of structural holes on performance

<table>
<thead>
<tr>
<th>Adminship</th>
<th>Linear regression</th>
<th>Logistic regression</th>
<th>Ordinal logit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>p &gt;</td>
<td>t</td>
</tr>
<tr>
<td>First edit</td>
<td>0.000</td>
<td>0.000***</td>
<td>-0.002</td>
</tr>
<tr>
<td>Total edits</td>
<td>0.000</td>
<td>0.039**</td>
<td>0.000</td>
</tr>
<tr>
<td>Size</td>
<td>0.002</td>
<td>0.000***</td>
<td>0.006</td>
</tr>
<tr>
<td>Density</td>
<td>-0.018</td>
<td>0.021**</td>
<td>-0.057</td>
</tr>
<tr>
<td>Hierarchy</td>
<td>-0.500</td>
<td>0.137</td>
<td>-2.314</td>
</tr>
<tr>
<td>Constant</td>
<td>17.6</td>
<td>0.000***</td>
<td>72.7</td>
</tr>
</tbody>
</table>

\[ R^2 = 0.26; \text{adjusted } R^2 = 0.26; \]
\[ p > F = 0.000*** \]

<table>
<thead>
<tr>
<th>Adminship</th>
<th>Linear regression</th>
<th>Logistic regression</th>
<th>Ordinal logit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>p &gt;</td>
<td>t</td>
</tr>
<tr>
<td>First edit</td>
<td>-0.001</td>
<td>0.000***</td>
<td>-0.002</td>
</tr>
<tr>
<td>Total edits</td>
<td>0.000</td>
<td>0.000***</td>
<td>0.000</td>
</tr>
<tr>
<td>Constraint</td>
<td>-0.439</td>
<td>0.105</td>
<td>-3.948</td>
</tr>
<tr>
<td>Constant</td>
<td>22.6</td>
<td>0.000***</td>
<td>78.0</td>
</tr>
</tbody>
</table>

\[ R^2 = 0.18; \text{adjusted } R^2 = 0.18; p > F = 0.000*** \]

\[ 2LL = 483.7; \chi^2 (5 \text{ d.f.}) = 133.1; \]
\[ p > \chi^2 = 0.000*** \]

\[ 2LL = 646.9; \chi^2 (5 \text{ d.f.}) = 143.6; \]
\[ p > \chi^2 = 0.000*** \]

\[ 2LL = 699.2; \chi^2 (3 \text{ d.f.}) = 100.8; \]
\[ p > \chi^2 = 0.000*** \]

\* p < 0.1.
\** p < 0.05.
\*** p < 0.01.

Thus, the number of indirect ties significantly and positively related to their administrative status. Because Wikipedia is a very active community with over 500 participants who contribute more than 100 edits each month, it is no surprise that those who have many direct ties would also have many indirect ties.

The structural holes argument posits that social capital, leading to better performance, is due to a person’s ability to exploit gaps in the social network. We found that the effective size component of structural holes was significantly and positively related to admship, as we hypothesized. Density was also negatively related, but this effect was significant only in the linear regression model that differentiated regular Wikipedians, regular administrators, and super administrators; it was not significant when we grouped both kinds of administrators. However, the hierarchy of the network was not significantly related to admship. Nor did the composite constraint measure have a significant effect.

Volunteers who formed a dense and closed network with many direct ties linking alters in the Wikipedia community appeared to be separated by others who did not belong to the network. We suspected that the participants who did not form a dense network were reluctant to create a new relationship with someone who had already established a dense network. When a participant interacted extensively only with the other participants in their network, the people who did not belong did not appreciate their contribution or even just ignore their postings. This suggested that people should broaden their network in order to be recognized by others.

Our results also indicated that the size of a person’s network significantly affected their admship. Burt argued that networks with larger effective sizes tended to provide more information that could be used to exploit structural holes. We further suggest that with a network of larger size there are greater chances that holes exist.

We found that social network hierarchy did not affect admship. In the case of Wikipedia, admship is granted through a communal vote. Even when there is a central Wikipedian who is well connected, he or she cannot help others gain admship in any way other than giving their support and casting their single vote. The Wikipedia community has a very low power distance culture, thus hierarchy has little effect.

The overall structural hole index (constraint) had a negative effect, as hypothesized, but was only marginally significant, with an average p value of 0.10 across both models. This is not surprising, considering that hierarchy is not significant, and that density is not consistently so.

6.1. Implications for open source contributors and communities

Although open source contributors are generally altruistic, they nonetheless often have a desire to be...
A large number of the direct ties came from participating in many diverse articles and thus being able to interact on talk pages with a wider spectrum of contributors. It is not sufficient to simply participate in many different articles in the same category, as this might expose the Wikipedian only to the same set of coauthors. More generally, in an open source community, contributors who want higher status must try to participate in a wide diversity of projects within the community. Of course, genuine skill is hard to fake, and only those with multifaceted abilities can do this. But this indicated that open source communities tend to reward such persons.

In accordance with Burt’s findings, our results showed that a degree of density, one of the elements of network constraints, had a generally negative impact on performance as measured by obtaining administrative rights. This result suggested that failure to diversify gave collaborators little opportunity to be known and commended by promotion, but this could also be perceived as unfriendly or insular and harm a participant’s opportunities for promotion.

In an open source community, high density indicates that several people are working on the same project. In the model that only distinguished between those with some administrative status and those with none, density was not found to add social capital significantly. This indicated that regular open source participants had no need to fear participating in heavily supported projects—in fact, this could expose them to more people, and increase the social capital. However, when we distinguished between higher-level and lower-level administrators, dense networks apparently reduced participants’ chances for promotion, as the structural hole theory hypothesizes. This would suggest that when open source participants have already established their status in the community and want to differentiate themselves further, they should seek projects or articles that are not widely frequented, in order to extend their skill base and exploit the structural holes that accrue from diversifying interests.

6.2. Implications for other open source communities

Considering the importance of formal community recognition and status in motivating the performance of the highest performers, our findings could be helpful in guiding open source community leaders in improving their volunteer force. In communities where there is no limit on the number of administrators, it would be helpful to promote as many participants as possible to ranks of higher privilege and responsibility. This study has shown that the direct ties, indirect ties, and effective size of a person’s network are important sources of social capital. These show that the more the number of interactions with a participant, the better his or her performance in the community. Contrary to what might be expected in a traditional organization, there is little effect of hierarchy in this case, indicating that the power of other members – and the power of the members to whom the ego is tied – is not relevant. We can only apply this conclusion to communities like Wikipedia, where recognition and promotion are conferred by democratic processes, which is not usually the case in open source communities. This would suggest that open source communities that want its members to maximally benefit from their ties should implement such open methods of according ranks.

Despite many differences (i.e., nature of the product, member characteristics, goals, etc.), Wikipedia follows the same principles by which other OSS communities operate. Therefore the results of our study could be generalized to other active OSS networks. It is worth noting that the Wikipedia contributors who are most interested in recognition rewards happen to be the most prolific. It is therefore in the interests of community leaders to understand the value of social networks and re-configure their structure in ways that best mobilize the embedded social resource.

6.3. Implications for corporations

Recently, corporations have become interested in implementing the Wiki as their new knowledge management mechanism. In fact, it has already helped several companies transform the way they brainstorm, track projects and coordinate marketing, while allowing them to preserve their physical resources (eliminate meetings, reduce conference calls, etc.) [25].

However, managers should be aware of a potential negative consequence as a result of adopting the Wiki as their coordination channel. Social capital is a fundamental resource that facilitates smooth knowledge
sharing. The Wiki as a coordination mechanism is likely to succeed when a company can construct or reconfigure its network in which both categories of social capital are efficiently produced and reproduced. Companies may find it extremely difficult to ensure “pure” voluntary participation; norms or trust must be present or opportunity and shirking are likely to occur. Therefore, to avoid failures, managers may have to provide financial incentives or use their authority to force participation.

6.4. Implications for academia

Wiki technology has the potential to promote more research collaboration both within and across academic fields, providing a platform through which researchers who have never met can collaborate and produce intellectual capital. An open source encyclopedia has already been established for the IS research community, using Wiki technology (http://ispedia.terry.uga.edu). In fact, each academic field could develop its own OCE.

7. Conclusion

Despite much interest in the development of open source software, little research has been performed in this area. In particular, due to the difficulty of data collection, few empirical studies have been attempted. Based on the perspective of social capital, we examined empirically the extent to which network closures and structural holes in the network influenced participants’ performance. Our approach may help address one issue that is often cited as a key challenge to open source community stakeholders: how to keep the interest of participants and motivate them to continue to contribute. Economic incentives might dilute the original spirit over which OSS communities are established and may have negative consequences. One alternative for maintaining members’ interests is to construct or adjust the network structure to attract participants and improve its social capital resources. Furthermore, OSS leaders should fine-tune or reconfigure their network to be efficient with equilibrium between the network closures and structural holes.

Uncited references

[3, 13, 16, 17, 19, 26, 34, 46, 53].

Acknowledgements

We gratefully acknowledge the invaluable assistance of Claudio Vejar. This research was partially funded by Fonds de recherché sur la societe et la culture [FQRSC 2006-NP-105602], and by the Faculty Research Development Program of Concordia University.

References


