Social Network Structure as a Critical Success Condition for Virtual Communities

David Hinds and Ronald M. Lee
Department of Decision Sciences and Information Systems, Florida International University
david.hinds@fiu.edu, r.lee@fiu.edu

Abstract

Virtual communities have become an important new organizational form and yet relatively little is known about the conditions which lead to their success. In an attempt to address this knowledge gap, a particular subset of virtual communities - open source software project communities - is investigated and four hypotheses are asserted which relate social network structure to community success. The hypotheses, which are based on social network theory and related research, suggest that success is supported by high levels of affiliation with other communities, moderate levels of density within the network of community conversations, moderate levels of density in the communications between peripheral members and core members, and low levels of density in the communications between administrators and the rest of the community. Empirical research is underway to test these hypotheses based on a sample of over 200 open source software project communities.

1. Introduction

Virtual communities have become the subject of considerable interest in both research and practice. These communities encompass a broad spectrum of activities, ranging from social networking sites such as MySpace to online auction sites such as eBay to open source software project communities such as Linux.

In assessing the performance of virtual communities, there are various possible ways of defining and measuring "success" depending upon the perspective of the evaluator as well as the type of community that is involved [7, 9, 42]. Regardless of the definition of success, it has been observed that some virtual communities are more successful than others. Using open source software project communities as an example, only a small fraction of the 20,000 or more project communities that are registered each year have been observed to produce useful computer software.

Little is known about why some virtual communities are successful while others are not. Some general guidance is provided from the results of case studies and the observations of a few prominent practitioners, but only a few multi-community studies of success have been conducted with incomplete and fragmented results. Furthermore, the studies that do exist have tended to focus on a particular type of virtual community (or even an individual community as in a case study) and have not attempted to identify more fundamental theory-based conditions which may be applicable across a range of community types.

The general problem to be addressed in this research involves virtual communities and the conditions which influence their success. The specific objective of this study is to identify social network structures of a particular type of virtual community - the open source software project (OSSP) community - that might impact the success of the community (and associated project) and to study the relationship between these structures and community success. In this context, "success" is defined as the level of community activity and quantity of community work output.

Social network constructs are used for the independent variables because prior research shows that social network measures are useful predictors of various group outcomes [4, 6, 8, 48]. Social network constructs are quite flexible and can be applied across different types of virtual communities. In addition, these constructs can be used to reflect a wealth of qualitative knowledge regarding effective community management in a concise and mathematical form, and derived social network measures can then be used as a practical diagnostic and monitoring tool.

In section 2 of this paper, we propose a needs-based virtual community typology, followed by a discussion of OSSP communities and concepts of success in virtual communities. In section 3, we briefly review social network concepts and in section 4, we discuss social network research of virtual communities and then present our overall conceptual framework. We derive and discuss four hypotheses in section 5 and then discuss ongoing empirical testing in section 6. In

1530-1605/08 $25.00 © 2008 IEEE

section 7, we discuss the implications of this research for OSSP communities and virtual communities in general.

2. Virtual communities

While various definitions have been offered in the literature, a simple definition is adopted for the purposes of this research:

"A virtual community is a community in which the primary mode of interaction is electronic (online/virtual) and not face-to-face."

In this context, the notion of "community" is broadly interpreted and can involve either direct interactions among members (e.g. as in threaded conversations) or indirect interactions through the member creation, modification, and usage of persistent digital goods (e.g. as in interactions through a digital repository such as the Wikipedia).

2.1 Typology of virtual communities

Virtual communities are sometimes viewed in the literature as a homogeneous set of activities. In reality, there are various types of virtual communities that have become prominent and it is plausible that the characteristics of these different types are such that different theories apply to each. Various typologies have been suggested in the literature. For example, one typology offered by Dube et. al. [13] is applied to virtual communities of practice and covers the dimensions of demographics, organizational context, membership characteristics, and technology environment. The authors specifically exclude "casual" online communities and focus on organizationally-created entities. Hummel and Lechner [19] study 50 business-related virtual communities and propose a set of five "genres of business-relevant communities" including 1) gaming, 2) interest, 3) business-to-business, 4) business-to-consumer, and 5) consumer-to-consumer.

Many different typologies could be defined. For the purposes of this research, we have devised a typology in which each community type is defined by a particular set of needs that members expect to be satisfied through their participation in the community. This needs-based typology is chosen because it is expected to have relevance for virtual community success (as further discussed in section 4.3).

Each virtual community type is listed in the following subsections. For each type, the key member subgroups (stakeholder groups) are defined, along with examples of the needs that these members expect to satisfy through their community involvement. These needs may be instrumental (e.g. building software development skills, acquiring video content) or they may be psychological (e.g. feeling support or acceptance). Of course, we recognize that other needs-based typology classifications could be devised and this is just one possible set. Also, some communities may be hybrids in that their member needs may be represented by more than one type.

2.1.1 Socialization communities. Virtual socialization communities are commonly associated with social networking sites and personal blogs. The primary member group consists of individuals who wish to socialize with others, especially those having shared interests. Member needs are those associated with socialization such as the need for approval, support or friendship. Examples of this type include communities formed around MySpace, Facebook, and SecondLife.

2.1.2 Gaming communities. Virtual gaming communities are often associated with massively multiplayer online role playing games (MMORPGs) such as EverQuest, although they may also include non-role playing games such as Full Tilt Poker. Members of these communities have an instrumental need to play a game with others, involving a need for competition and/or fantasy.

2.1.3 Content sharing communities. A virtual content sharing community involves the sharing of digital content such as music or videos. Some members are content providers and some are content users, although the same individual may perform both roles. Members may have instrumental needs to acquire a particular song or they may have psychological needs to express themselves artistically. Examples include communities supported by video sharing sites like YouTube and peer-to-peer file sharing sites such as Kazaa.

2.1.4 Knowledge sharing communities. One of the broadest of the virtual community types, knowledge sharing communities are formed to support learning, opinion expression, and information dissemination, and are organized around sites such as Wikipedia or Slashdot. Primary member groups include knowledge providers and knowledge seekers. Some members have an instrumental need for knowledge on a particular topic while others have a need to build their reputation.

2.1.5 Activism communities . Virtual activism communities are formed for the purpose of organizing the members for some type of offline action, such as with an online political action group or a consumer group. The primary member group includes those who have an interest in the action, such as the skilled immigrants who formed Immigration Voice. The member needs are instrumental and relate to the desired action.

2.1.6 Development communities. Virtual development communities are organized for technological development and innovation purposes. The most prominent example includes open source software project communities, such as Linux developers and users, as well as the broader open source communities such as SourceForge. Also included within this type are innovation communities such as ThinkCycle. Some members have task-related needs to create and/or use designs or artifacts (e.g. software) and may also have psychological needs such as feeling a sense of accomplishment. Other members (corporate sponsors) may have an instrumental need to make profits (e.g. by providing support services).

2.1.7 Exchange communities. Virtual exchange communities are organized to support the economic exchange of goods or services. Primary member groups include buyers and sellers, although the platform provider is also a key stakeholder. Most members have an instrumental need to engage in a buy-sell transaction, while the platform provider has a need for profit. Examples include eBay and Craigslist.

2.2 Open source software project communities

OSSP communities are typically initiated by an individual (or group of individuals) who provides (or provides access to) systems and development components, as well as communication infrastructure. Participants are usually volunteers and contributors are not normally motivated by traditional economic incentives, but rather by instrumental factors associated with fulfilling a need, and by intrinsic factors such as enhanced reputation, expertise development (learning), self-fulfillment, as well as basic fun and enjoyment [22, 24, 33, 34, 46].

The individuals that participate in open source software projects are often described as comprising a community. These communities have been described as having an onion-like structure, with a central core of highly active individuals, surrounded by other layers of progressively less active individuals. One example of this is presented by Ye et. al. [52] in which the central core is composed of the project leaders and core members, with five outer layers containing active developers, peripheral developers, bug reporters, passive users, and stakeholders, respectively.

OSSP communities have been referred to as virtual communities [12] and are depicted as part of the "development" type in section 2.1.6. However, open source research is often viewed as its own track and is not frequently associated with the study of virtual communities. This may be due to the fact that much of this open source research addresses software/technical/work process issues [10] and not social/community issues. This particular research project is focused on the social and community aspects, and we consider OSSP communities to be a subset of virtual communities.

2.3 Success of virtual communities

For virtual communities in general, only a few studies have addressed the issue of success. Lin et. al. [25] study "web-based knowledge communities" and validate a research model which relates usability and sociability factors to community success. This work uses Preece's community success framework [30]. In a study by Leimeister, et. al. [23], the authors review the literature for success factors in virtual communities and note the different definitions of success that are associated with the different stakeholder groups. They identify 32 success factors and then assess their importance using a stakeholder survey. In a study related to community success, Sangwan [35] investigates the motivational factors expressed by community participants.

In terms of specific studies of OSSP community success, again only a few papers are available. A statistical analysis by Krishnamurthy [21] was conducted on SourceForge projects which were categorized as being in a "mature" development status. Stewart and Ammeter [41] conducted an analysis of 240 open source software projects to investigate factors which lead to attracting user attention ("popularity") and developer activity ("vitality"). Their preliminary results indicate that vitality significantly affects popularity, and that sponsored projects are more popular than non-sponsored projects. Crowston and Scozzi [12] conducted a multiple regression analysis of SourceForge data and found support for Katzy and Crowston's [20] theory of competency rallying, which suggests that project members must have necessary competencies, understand market opportunities, apply competencies to meet opportunities, and manage cooperative processes. In a more recent effort to address open source software project success, Crowston et. al. [10] outlined an approach for studying the work practices of open source software project groups and relating these practices to team effectiveness.

In a paper devoted to the subject of success measures, Crowston et. al. [9] present a range of measures that could be used to assess the success of open source software projects. Their measures include system and information quality, user satisfaction, use, individual or organizational impacts, project output, process, and outcomes for project members, as well as the opinions of participants with respect to users, products, processes, developers, uses, recognition, and influence.

In general, these research works regarding virtual community success refer to one particular type of virtual community, and it is plausible that success conditions will be different depending upon the community type. For example, (referring to the

typology presented in Section 2.1) the success conditions for a socialization community may not be the same as for a gaming community or for an exchange community.

3. Social networks

Social networks are graph theoretical representations of the relationships between social entities, in which the social entities (such as people or organizations) are represented as a set of nodes and the relationships (such as advice-giving or trade) are represented as a set of ties which connect the nodes. The structure of a social network is defined by certain mathematical properties of the associated graph such as density or centralization.

The term "social network analysis" refers to a broad set of methods and tools for coding and analyzing social network representations. In contrast, the domain of social network theory involves the application of network concepts and perspectives to various aspects of social psychology, sociology, and organizational science. In the next section a brief review of social network analysis is provided followed by a discussion of social capital theory, which is one of the most prominent social network theories.

3.1 Social network analysis

First seen in 1934 in the "sociograms" of Moreno [26], social network analysis has grown into a large collection of methodologies, measurements, and tools that can be used for the description and analysis of social networks and social structure [6, 48]. The primary mathematical foundation for social network analysis is provided by graph theory, and social network analytical methods draw heavily on matrix algebra for coding and manipulating network data.

Social network analysis has a number of positive features with respect to its use as an analytical tool. Its use can reveal patterns that are not discernable with other methods. These patterns may be reflected in quantitative social network analysis measurements or they may be observed qualitatively in two- or three-dimensional graphical network representations. Further, the use of social network analysis provides a quantitative method for studying complex social phenomena such as kinship, community structure, corporate interlocks, and elite power, whose investigation would otherwise be limited to the use of qualitative tools.

3.2 Social capital theory

One of the primary theoretical foundations for this study is social capital theory which has been used extensively in research regarding the effect of social network structures on group outcomes. Social capital theory considers both the social network itself and the knowledge and other resources that may be mobilized through the network. The theory uses an information processing paradigm [38] to explain how social network structure affects social outcomes at the individual level and at the group level. Social ties are viewed as conduits for the flow of information, knowledge or other resources. In this study, we primarily utilize a group level perspective for social capital [31] as opposed to an individual level perspective [4].

One assertion of social capital theory is that a social network structure with high closure, or a tightly knit set of ties within the network, will facilitate the utilization of resources. Closure is seen to create cohesive groups, and this type of structure is expected to support shared norms and trust [8]. However, high closure can also have negative impacts on the network, for example as seen in the "groupthink effect".

Social capital theory further suggests that a social network structure with extensive bridging ties, which extend outside of the group, will facilitate the access to resources. At the group level, this is an extension of Granovetter's notion of weak ties [15] and their positive impact on information transfer. Bridging also involves Burt's [4, 5] concept of "brokerage" in which brokering individuals can provide the network (group) with access to non-redundant information and knowledge. However, this notion of brokerage has two interpretations based upon the intention of the actor in the broker position. If the actor intends to keep the other actors isolated in order to appropriate value from them, this is referred to as the "tertius gaudens" orientation, or "the third who benefits". This compares with the "tertius iungens" orientation, or "the third who joins" [27], in which the individual utilizes the broker position to help connect the other actors to their benefit and/or the benefit of others in the network. The tertius iungens orientation is often referred to as "bridging".

4. Social network structure in virtual communities

In this section, we review the limited existing research regarding social network structures in open source software project communities, and we make reference to other social network oriented studies involving other kinds of virtual communities. This is followed by a discussion of community subgroups research which can be related to social network concepts. We then suggest a conceptual model which posits the relationship between social network structure and virtual community success, as mediated by virtual community type.

4.1 Social network research of virtual communities

Much of the social network research regarding virtual communities is exploratory and descriptive. Substantive work in this area has been published by Wellman [17, 18, 49, 50, 51], Hampton [17], Haythornthwaite [18, 49, 51], Boyd [3], and Smith [39]. Schenkel et. al. [37] define five structural properties which can be used to characterize communities of practice. These include: 1) connectedness, 2) graph-theoretic distance, 3) density, 4) core/periphery structure, and 5) coreness. In comparing and synthesizing her prior studies, Teigland [44] notes that there are significant differences in the social structures of different community forms. In a study which mainly focuses on the "invisible" participants in communities (the "lurkers"), Rafaeli et. al. [32] use a social network perspective to measure the activities of individuals as they move from lurker role to active participant role.

In the area of OSSP communities, Crowston and Howison [11] examine 120 project teams (communities) from SourceForge, and analyze interactions associated with the bug reporting archives. In particular, they measure and compare the centralization measures of the different projects. They conclude that it is wrong to assume that all open source software projects are associated with a particular social structure, and that the examination of social structure offers an interesting avenue for future research. In a practice sense, they suggest that open source software project teams should spend more effort on creating social structures which are considered to be favorable.

4.2 Community subgroups research

A good deal of community research makes reference to (or implies the existence of) certain subgroups within the community. For example, three subgroups that are frequently mentioned in studies of communities include 1) leaders, 2) core members, and 3) peripheral members. The results of much of the research regarding these three kinds of subgroups can be reflected in a social network theory involving the connectedness of subgroups.

In this context, "subgroup connectedness", based on the notion of "cohesion" [2], refers to the extent to which a focal subgroup is central to or connected with the remaining members of the project group. In our study, this concept is measured using "two-mode density", whereby the mode-1 actors include members of the focal subgroup and the mode-2 actors include the remaining members of the community. For a focal subgroup, then, the two-mode density is defined as the total number of ties between the mode-1 members (the subgroup) and the mode-2 members (the rest of the community) divided by the maximum possible number of such ties.

Using this social network concept of subgroup connectedness, we can interpret some of Leimeister et. al.'s [23] supported hypotheses for peripheral members and leaders. For example, with respect to limiting leader connectedness, the study concludes that community managers "should intervene in community life as little as possible" and that "it is more important to operators of virtual communities to sustain neutrality than to constantly extend their offerings for community members". With respect to increasing peripheral connectedness, Leimeister, et. al. [23] concludes that "before changing lay-out or functionalities of a community site, it is important to give members the possibility to take part in the modification of design/functionality of the offerings first".

With respect to OSSP communities, three similar subgroups are frequently defined including 1) administrators, 2) core developers, and 3) peripheral developers. Administrators are developers who lead the project. They take responsibility for monitoring and guiding the progress of the project, and their special role is recognized by most group members [1, 43, 52]. Core developers are developers who are actively involved with the project and who contribute the majority of design concepts and source code for the project software. (Administrators are typically also core developers.) Peripheral developers are developers who are somewhat involved with the project and who have either contributed source code or have posted requests or comments to the public project communication records. These individuals are often beta-testers and/or they submit bug reports and feature requests.

Some of the research results for OSSP communities suggest that project members are motivated by a sense of ownership in the project, and that heavy-handed control by administrators can reduce the motivation of both core developers and peripheral developers. Interpreted through the lens of subgroup connectedness, this argues for lower levels of administrator connectedness. The research also suggests the importance of involving the peripheral developers in order to fulfill their need for challenge and skills development, and to increase their identification with the project which tends to increase their level of contribution. This research conclusion argues that higher levels of peripheral connectedness are more favorable to success.

4.3 Conceptual social network model of virtual community success

In this section, we propose a conceptual model which suggests that social network structure is a critical success condition for virtual community success (see Figure 1). The model provides a framework for the hypotheses which are described in the next section, and it is also intended to provide a starting point for future

research which may lead towards a contingency theory of virtual community success.

In presenting this model, it is suggested that certain social network structures will positively or negatively affect the perceived satisfaction of member needs associated with community participation. For example, higher levels of peripheral subgroup connectedness may support the satisfaction of peripheral members' need to feel a sense of legitimacy and acceptance, while higher levels of administrator subgroup connectedness may inhibit members' need to feel a sense of ownership and control in community matters. The perceived satisfaction of member needs will in turn lead to increased member motivation and community participation levels, which will ultimately result in a greater level of community success, especially considering the voluntary nature of these communities. In effect, virtual communities tend to be more successful if they are able to attract and motivate more participants, and therefore, the success of the community is likely to be closely tied with (and driven by) the extent to which participant needs are perceived to be satisfied by the community.

With respect to virtual community types, it is suggested that different kinds of social network structures may be supportive of different kinds of member needs, the satisfaction of which will ultimately lead to community success, as described above. For example, the member needs for approval, support, or friendship associated with socialization communities may require higher levels of communication network density, while the task-related needs of open source software developers in development communities may require low to moderate levels of communication network density. Thus, the needs-based virtual community type is shown as a mediating factor in the model. While the typology presented in section 2.1 may be useful for understanding this mediation effect, it is recognized that this typology might not be complete or optimal and that other kinds of typologies might be more useful.

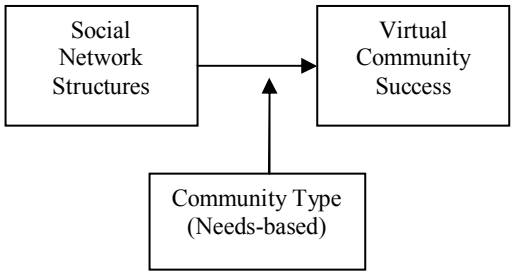

Figure 1. Conceptual social network model of virtual community success

5. Hypotheses

Using the above conceptual model as a framework, four hypotheses are devised which relate social network structures to successful outcomes in OSSP communities (see Figure 2). The first two hypotheses are derived from social capital theory while the last two involve the application of subgroup connectedness concepts to research regarding the activities of the administrators and peripheral developers subgroups. It is recognized that success conditions may differ for mature communities versus early-stage communities, and these hypotheses are more suited for application to early-stage communities.

Four dependent variables are defined to measure the "community success" construct which is referred to in each hypothesis (resulting in four versions of each hypothesis to be tested). These measures consist of "number of downloads" and "number of page views" (reflecting community activity levels), as well as "number of software commits" and "number of software releases" (reflecting community work output).

The hypotheses further refer to two different kinds of networks: 1) a communication network and 2) a membership network. The communication network involves the communications among group members on public discussion facilities (e.g. public forums, mailing lists, etc.). Crowston and Howison [11] use a similar type of communication network to study the social structural patterns of open source software projects by extracting conversational data from bug report trackers. The membership network is defined by the set of projects for which an individual is a member. Gao, et. al. [14] define such a membership network in studying the connections between open source software projects hosted at SourceForge.

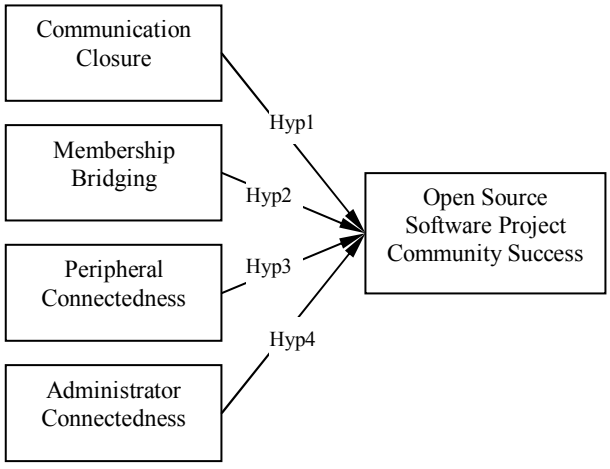

Figure 2. Social network model of OSSP community success

5.1 Communication closure

As defined in this study, closure is the extent to which the members of the community are connected to each other through the communication network. In effect, closure represents the extent to which all members talk with all other members. This construct is measured by "group communication density", which is the number of ties within the community communication network divided by the maximum possible number of such ties, where a tie between two members is recognized if the two members are co-participants in at least one discussion thread on the project public forums.

In social capital theory, closure is generally portrayed as leading to positive social outcomes involving utilization of resources and group health and viability. However, some negative effects are noted regarding groupthink and a reduced tendency to associate with outsiders. In group effectiveness studies, closure has been generally associated with a positive effect, although at least one study suggested that the relationship is an inverted-U shape [28].

For the community as a whole, it is anticipated that a certain level of closure is required in order to carry out the activities of the community. However, after a certain point, additional closure becomes a burden, it is distracting, and therefore it reduces the effectiveness of the group in carrying out community activities. This is a "cost-of-ties effect" in which the time and effort associated with creating and maintaining a tie can act to reduce the benefits associated with the tie.

In the case of communication ties within OSSP communities, the effort required to post a comment to a discussion forum would seem to be rather small. However, consideration should be given to the time necessary to read and understand the content of previous forum posts and to start and maintain a dialogue with other actors. Further, open source software projects, as defined in this research, involve volunteers who typically have a limited amount of time to contribute to the project. Thus, each additional tie that is established represents a cost to the actors involved and the group as a whole.

Thus, closure has a positive influence on community success, but only to a point, after which the cost-of-ties effect begins to reduce group effectiveness. Therefore, the relationship between communication density and community success is posited as follows:

Hypothesis 1

OSSP community success is maximized at a moderate level of group communication density (i.e. group communication density has an inverted-U relationship with community success)

5.2 Membership bridging

Bridging is the extent to which community members are connected (via "bridging ties") to other OSSP communities through the membership network. The more other communities that an individual is tied to, the higher the value of the bridging construct. This construct is measured by "average core affiliation degree", where "affiliation degree" is the total number of other OSSP communities in which the focal actor is affiliated as a core member, averaged over all core developer members of the focal community.

In the social capital literature, bridging is generally associated with improved access to resources and an associated increase in performance. There are various positive effects associated with bridging ties. The bridging actors will have access to new ideas such as production and design methods. In addition, these actors may be able to bring in members from other communities on a one-time basis to solve particular problems and/or provide other special kinds of support. Bridging ties may also increase the likelihood of recruiting new members from other communities, as the focal actor utilizes his or her bridging ties to communicate the features of the focal project to potential members from other groups. These effects result in additional resources which should help to improve community performance.

In the case of bridging ties, the cost-of-ties effect is only a burden on the individual actor, because the ties are between that actor and the members of other communities (not the focal community). However it is possible that too many bridging ties would result in a lack of time and attention given to the focal project by the bridging actor. Again, this is only one actor and the net negative effect of this on the overall project is expected to be minor. Therefore, the general relationship between the bridging measure and community success is expected to be positive:

Hypothesis 2

OSSP community success increases as average core affiliation degree increases (i.e. average core affiliation degree has a positive relationship with community success)

5.3 Peripheral connectedness

For peripheral developers, a greater level of subgroup connectedness should lead to a greater sense of identification with the project, as well as feelings of satisfaction and challenge. In this study, peripheral communication connectedness is measured by "peripheral two-mode density", which is the total number of ties between the peripheral developers (mode-1) and the rest of the community (mode-2) divided by the maximum possible number of such ties, where a tie between two members is recognized if the

two members are co-participants in at least one discussion thread on the project public forums.

As Raymond [34] notes, it is important to "listen to the beta testers". Greater connectedness should translate into increased feelings of obligation and commitment to make contributions and to remain with the project. More connected peripheral developers are more likely to contribute code, report bugs, and assist with the production of the project software. These peripheral developers may be the source of new ideas and methods of development that could improve the group processes. Further, one or more may decide, at some point, to become core developers.

On the negative side, the cost-of-ties effect may become a significant factor as peripheral developer connectedness increases. Higher levels of connectedness with the core developers may become a distraction for these more active individuals. As the two-mode density increases, the effect of increasing cost-of-ties is expected to offset the benefits of having more motivated peripheral developers. Therefore, the relation between peripheral two-mode density and community success is expected to be positive, but only to a point:

Hypothesis 3

OSSP community success is maximized at a moderate level of peripheral two-mode density (i.e. peripheral two-mode density has an inverted-U relationship with community success)

5.4 Administrator connectedness

In this study, administrator communication connectedness is measured by "administrator two-mode density", which is the total number of ties between the administrators (mode-1) and the rest of the community (mode-2) divided by the maximum possible number of such ties, where a tie between two members is recognized if the two members are co-participants in at least one discussion thread on the project public forums. For administrators, a certain level of subgroup connectedness is necessary in order for them to coordinate and integrate the work of the other developers. However, as the level of connectedness increases, the administrators face the possibility of becoming overburdened and subject to "burn-out", which would have significant negative effects on community success. This is essentially the cost-of-ties effect observed at the individual level of the administrator. Pavlicek [29] suggests that administrators should delegate as much as possible.

With regard to effects on the other (non-administrator) community members, again, a certain level of connectedness is valuable in that these members need to feel welcomed and accepted into the group. At a point, however, too much connectedness of the administrator subgroup can lead to a loss in the "feeling of ownership" that is so important for open source software contributors [47]. Although some small level of administrator involvement is necessary, the cost-of-ties effect becomes significant and results in a compounded depressive effect on member motivation and administrator overload. Therefore, the net effect of administrator two-mode density on community success is expected to be negative:

Hypothesis 4

OSSP community success decreases as administrator two-mode density increases (i.e. administrator two-mode density has a negative relationship with community success)

6. Empirical testing of the hypotheses

Research is currently underway to empirically test a set of 10 hypotheses which include the above 4 hypotheses and 6 other variations. A sample of over 200 OSSP communities is selected and an unobtrusive research method is used to extract and analyze existing statistics from the public archives of Sourceforge.com. In addition, data is extracted from the University of Notre Dame SourceForge database [45] and the CVSAnalY database [16], both of which are created from data dumps which are periodically produced by the SourceForge organization.

Using the above data sources, social network measures and community success measures are derived from statistics of forum discussion thread archives, project records, and membership records. This study involves a cross-sectional design with a time-adjusted observation window, whereby variables are tracked for the two year period following the date of first software release, regardless of the actual start date of the project. This design results in a more homogeneous set of projects (all being at a similar stage of development), compared with other designs which select the sample at a given point in time and therefore include a more heterogeneous group of projects (in various stages of development).

Because it is plausible that group size will be positively related to community success (especially with regard to community activity levels), the effect of the independent variable for each hypothesis will be isolated from the effect of group size (as measured by "number of community members") using a hierarchical regression method.

7. Implications

The results of this work will provide OSSP community leaders and sponsors with guidance regarding the kinds of social network structures that are supportive of community success. Based on these results, actions can be devised regarding appropriate management / leadership approaches, community rules

and norms, and also the technical design of the web platform for the OSSP community. For example, administrators may choose to limit their involvement in forum discussions and encourage others to communicate with peripheral developers. In a pragmatic sense, the results will provide practical measurement tools which can be efficiently applied to pre-existing digital archives, such as email, instant messaging and online forums. Project administrators can use such measures to assess their own communities and to determine if they have the right kinds of structures or if changes might be necessary.

With regard to virtual communities in general, further research is needed to validate the results for other virtual community types. It is expected that the conceptual model and typology presented in this paper will provide a starting point for this work.

8. References

[1] Almarzouk, M., L. Zheng, et al. (2005). "Open source: concepts, benefits and challenges." Communications of the AIS 16: 756-784.

[2] Borgatti, S.P. and Everett, M.G. (2006) "A graph-theoretic perspective on centrality." Social Networks 28(4): 466-482.

[3] Boyd, D., H.-Y. Lee, et al. (2002). Developing legible visualizations for online social spaces. 35th Annual Hawaii International Conference on System Sciences, Hawaii.

[4] Burt, R. S. (1992). Structural holes: The social structure of competition. Cambridge, Mass., Harvard University Press.

[5] Burt, R. S. (2001). Structural holes versus network closure as social capital. Social capital: theory and research. N. Lin, K. Cook and R. S. Burt. New York, Aldine De Gruyter: 31-56.

[6] Carrington, P. J., J. Scott, et al., Eds. (2005). Models and methods in social network analysis. Cambridge, Mass., Cambridge University Press.

[7] Colazo, J. A., Y. Fang, et al. (2005). Development success in open source software projects: exploring the impact of copylefted licenses. The Americas Conference on Information Systems, Omaha, NE.

[8] Coleman, J. S. (1988). "Social capital in the creation of human capital." American Journal of Sociology 94: S95-S120.

[9] Crowston, K., H. Annabi, et al. (2004). Towards a portfolio of FLOSS project success measures. The 4th Workshop on Open Source Software Engineering, Edinburgh, Scotland.

[10] Crowston, K., H. Annabi, et al. (2005). "Effective work practices for FLOSS development: A model and propositions." 38th Hawaii International Conference on System Sciences - 2005, Hawaii.

[11] Crowston, K. and J. Howison (2004). The social structure of free and open source software development. Syracuse FLOSS Working Paper.

[12] Crowston, K. and B. Scozzi (2002). "Open source software projects as virtual organizations: competency rallying for software development." IEE Proceedings Software 149(1): 3-17.

[13] Dube, L., Bourhis, A., and Jacob, R. (2006). "Towards a Typology of Virtual Communities of Practice." Interdisciplinary Journal of Information, Knowledge, and Management, 1, 69-93.

[14] Gao, Y., V. Freeh, et al. (2003). Analysis and modeling of open source software community. North American Association for Computational Social and Organizational Science (NAACSOS) Conference 2003.

[15] Granovetter, M. (1973). "The strength of weak ties." American Journal of Sociology **78**: 1360-1380.

[16] GSYC/LibreSoft, "CVSAnalY-SF database" release of June 2006, <http://libresoft.urjc.es/Data/CVSAnalY_SF>

[17] Hampton, K. N. and B. Wellman (1999). "Netville online and offline: observing and surveying a wired suburb." American Behavioural Scientist **43**(3): 475-492.

[18] Haythornthwaite, C. and B. Wellman (1998). "Work, friendship, and media use for information exchange in a networked organization." Journal of the American Society for Information Science **49**(12): 1101-1114.

[19] Hummel, J. and Lechner, U. (2002) "Social profiles of Virtual Communities." In: Sprague, E., (Ed.) Proc. of the Int. Hawaiian Conf. on System Sciences (HICSS 2002), IEEE Press]

[20] Katzy, B. R. and K. Crowston (2000). "A process theory of competency rallying in engineering projects." Submitted to IEEE Transactions on Engineering Management.

[21] Krishnamurthy, S. (2002). "Cave or community? An empirical examination of 100 mature open source projects." First Monday 7(6).

[22] Lakhani, K. R. and R. G. Wolf (2005). Why hackers do what they do: understanding motivation and effort in free/open source software projects. Perspectives on free and open source software. J. Feller, B. Fitzgerald, S. Hissam and K. R. Lakhani. Cambridge, MA, MIT Press.

[23] Leimeister, J. M., Sidiras, P. and Krcmar, H. "Success factors of virtual communities from the perspective of members and operators: an empirical study". Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS-37), Big Island, Hawaii, 05–08 January, 2004.

[24] Lerner, J. and J. Tirole (2002). "The simple economics of open source." Journal of Industrial Economics 52: 197-234.

[25] Lin, H., W. Fan, et al. (2007). "An empirical study of web-based knowledge community success." Proceedings of the 40th Hawaii International Conference on System Sciences - 2007.

[26] Moreno, J. (1934). Who shall survive? New York, Beacon Press.

[27] Obstfeld, D. (2005). "Social networks, the tertius iungens orientation, and involvement in innovation." Administrative Science Quarterly 50: 100-130.

[28] Oh, H., M.-H. Chung, et al. (2004). "Group social capital and group effectiveness: The role of informal socializing ties." Academy of Management Journal 47: 860-875.

[29] Pavlicek, R. C. (2000). Embracing insanity: open source software development. Indianapolis, IN, Sams.

[30] Preece, J., Online Communities – Designing Usability, Supporting Sociability. 2000, Chichester, New York, Weinheim, Brisbane, Singapore, Tokio: John Wiley and Sons. 439.

[31] Putnam, R. D. (2000). Bowling alone: the collapse and revival of American community. New York, Simon and Schuster.

[32] Rafaeli, S., G. Ravid, V. Soroka (2004). "De-lurking in virtual communities: A social communication network approach to measuring the effects of social and cultural capital." Proceedings of the 37th Hawaii International Conference on System Sciences – 2004.

[33] Raymond, E. S. (1998). "Homesteading the noosphere." First Monday 3(10): <<http://www.firstmonday.dk/issues/issue3_10/raymond/>>.

[34] Raymond, E. S. (1999). The cathedral and the bazaar: musings on Linux and open source by an accidental revolutionary. Sebastopol, CA, O'Reilly & Associates, Inc.

[35] Sangwan, S. (2005). "Virtual community success: A uses and gratifications perspective." Proceedings of the 38th Hawaii International Conference on System Sciences - 2005.

[36] Savannah (2005). <http://savannah.gnu.org/.>

[37] Schenkel, A., R. Teigland, et al. (2000). Theorizing communities of practice: a social network approach. Academy of Management: Organization and Management Theory Division.

[38] Simon, H. A. (1976). Administrative behavior (3rd edition). New York, Free Press.

[39] Smith, M. A. and P. Kollock, Eds. (1999). Communities in cyberspace. London, Routledge.

[40] SourceForge (2005). <http://www.sourceforge.net.>

[41] Stewart, K. J. and T. Ammeter (2002). An exploratory study of factors influencing the level of vitality and popularity of open source projects. Twenty-Third International Conference on Information Systems.

[42] Stewart, K. J. and S. Gosain (2006). "Impacts of ideology, trust, and communication on effectiveness in open source software development teams." MIS Quarterly **30**(2): 291-314.

[43] Sturmer, M. (2005). Open source community building. MIT Open Source collection.

[44] Teigland, R. (2003). Knowledge networking: structure and performance in networks of practice. Institute of International Business. Stockholm, Stockholm School of Economics.

[45] University Notre Dame, SourceForge Research Archive, 2007, <http://zerlot.cse.nd.edu/>

[46] von Hippel, E. and G. von Krogh (2003). "Open source software and the "private-collective" innovation model: issues for organization science." Organization Science 14(2): 209-223.

[47] von Krogh, G. (2003). "Open-source software development." Sloan Management Review: 14-18.

[48] Wasserman, S. and K. Faust (1994). Social network analysis: methods and applications. Cambridge, UK, Cambridge University Press.

[49] Wellman, B. and C. Haythornthwaite, Eds. (2002). The internet in everyday life. Oxford, Blackwell.

[50] Wellman, B. and B. Leighton (1979). "Networks, neighborhoods and communities." Urban Affairs Quarterly **14**: 363-390.

[51] Wellman, B., J. Salaff, et al. (1996). "Computer networks as social networks: collaborative work, telework, and virtual community." Annual Review of Sociology **22**: 213-238.

[52] Ye, Y., K. Nakakoji, et al. (2005). The co-evolution of systems and communities in free and open source software development. Free/open source software development. S. Koch. Hershey, PA, Idea Group Inc (IGI): 59-82.