The Efficacy of Knowledge Sharing: Centralized Vs. Self-Organizing Online Communities

Jaideep Godara

Thesis submitted to the faculty of the
Virginia Polytechnic Institute and State University
in
partial fulfillment of the requirements for the degree of

Master of Science
in
Industrial and Systems Engineering

Dr. Andrea Kavanaugh, Chair

Dr. Brian M. Kleiner, Co-Chair

Dr. Tonya L. Smith-Jackson, Member

May 4th 2007

Blacksburg, Virginia

Keywords: Online communities, Communities of Practice, Weblogs, Discussion forums, Public Deliberation, Legitimate Peripheral Participation, Weak-ties, Knowledge Sharing.

The Efficacy of Knowledge Sharing: Centralized Vs. Self-Organizing Online Communities

Jaideep Godara

ABSTRACT

This study investigates the impact of an online community's control structure on the knowledge sharing process in that community. Using a framework comprised of legitimate peripheral participation theory and the weak-ties phenomenon, the study focuses on a comparative analysis of self-organizing online communities (e.g., weblog networks) and centralized online communities (e.g., discussion forums communities) with respect to the efficacy of knowledge sharing in these communities. The findings of this study indicate that self-organizing communities of practice have more weak-ties among their members compared to centralized communities. As per weak-ties theory of Granovetter (1973, 1983), these findings suggest that self-organizing communities facilitate greater dissemination of knowledge and flow of information among their members than centralized communities. The abundance of weak-ties in their community structure also makes self-organizing communities better environments for the discovery of new information compared to centralized community environments.

This study did not find any evidence of community structure impact on peripheral participation and the interaction activity level among peripheral participants of a given online community. These observations may have stemmed from the limitations of research design, however, it is safe to say as of now that verdict on peripheral participation differences in different community structures is inconclusive at best.

AKNOWLEDGEMENTS

Looking back at the time when I started working on this study, and even before that when I first started working with Digital Government research group at Center for HCI, Virginia Tech, I feel really amazed how far I have come and how much I have learned during this period. This journey has had its due shares of difficulties, but in the end, it has proven to be one of the most intellectually stimulating experiences of my life. Without a doubt, this study and my own quest for learning would not have been possible without support of my professors, mentors, and friends.

First, I want to acknowledge the support and assistance of my advisory committee, Dr. Kavanaugh, Dr. Kleiner, and Dr. Smith-Jackson. I express my heartiest gratitude to my mentor and the chair of my committee, Dr. Andrea Kavanaugh. She has been a driving force for this study from the moment I started working with her; it was her guidance, undeterred support and constant encouragement that helped me wade through the difficult times when I was almost down and out. Thank you, Dr. Kavanaugh.

My special thanks to Dr. Brian Kleiner, Co-chair of my advisory committee, for his patience, guidance, and helping me with designing and refining this study. I would also like to thank Dr. Smith-Jackson for her guidance, invaluable feedback, and suggestions for improving this study, as well for her services as a committee member on my advisory committee. I feel extremely fortunate to be a student of Dr. Kleiner and Dr. Smith-Jackson; their knowledgeable teaching and guidance have given a new life and vigor to my interest in the discipline of Human factor. Thank you, Dr. Kleiner and Dr. Smith-Jackson.

I owe special thanks to Dr. Manuel Perez and Philip Isenhour from Digital Government research group, who have helped me throughout this research to keep my thinking sharp and review my work. Their ideas and early involvement with this study have steered my research throughout this study. Thank you, Dr. Perez and Philip. I am similarly indebted to my fellow research student, Joon Kim, for his untiring and selfless help with Social network analysis and in refining data analysis for this study. My special thanks to Brian Cochran, who took the time to read and refine my early thesis drafts.

I owe the most to my family–my parents, my brother, Manish, and my sister, Namita. Their unwavering faith and support have been a constant motivation for me throughout my life. They have encouraged me to pursue my dream of higher studies and learning, amidst all of the demands of family life. These years of learning away from home have not been easy for them in terms of time and absence from home, and I am forever grateful for their love and support. I owe you more than I can say in words.

I owe special thanks to my close friend George, who drove me to places and did errands for me when I had little time to focus on anything else but this study, and above everything else, who kept me high in my spirits. Thank you, George. My sincere thanks to Phil and Val, who lent me their place for long writing sprees and faithfully reminded me of time to sleep! I hope to return the favor someday. Finally, I extend my sincere appreciation to my friends Anu, Juan, Marshall, Ryan, Steph, Virginia, Vishal, and Xiaoyi for being my support system in Blacksburg.

To all of you, Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

1.1 Background Information

In the spring of 2006, a report based on a detailed survey of 431 human resources officials by the Conference Board, the world's preeminent business research organization, highlighted a fear that many business analysts had already expressed: "The future workforce is here, and it is ill-prepared."(Conference Board, 2006). Apparently, the rapidly aging workforce – caused mainly by the number of the baby boomers slowly retiring – has led to a widening of the chasm between much needed workplace skills of the existing workforce and the incoming worker generation. Furthermore, the priorities of U.S. companies have shifted from knowledge creation to real-time knowledge sharing and collaboration.

This shift of focus also reflects an important change away from the mantra that "knowledge is power, so hoard it in order to stay ahead of the competition. The new mantra of the organizational and corporate world is that knowledge is power, so share it and it will multiply". This paradigm shift is also evident in the recent survey findings by the Conference Board and the American Management Association that shows that more than 78% of U.S. companies have undertaken knowledge management initiatives (Conference Board, 2005). The same report gives a more in-depth analysis about the causes behind this concern: one-half of companies surveyed felt that the departure of mature workers presents potential knowledge vulnerabilities, a crucial factor in the competitive global market. More than one-half of these companies have some form of mentoring program in place to share and transfer knowledge. Knowledge sharing has become strategically crucial for business entities and organizations.

In the modern knowledge-intensive economy, knowledge is a recognized strategic asset and its management is crucial for any organization to have a sustainable competitive advantage (O'Dell & Grayson, 1998; Osterloh & Frey, 2000). The obstacle in achieving smooth and effective knowledge exchange, however, is that knowledge is a distributed and situated cognitive entity (Brown, Collins, & Duguid, 1989; Lave & Wenger, 1991; Suchman, 1987). That is, it resides in and is distributed among different individuals, tasks, and sub-networks or groups within any organization (Argote & Ingram, 2000). This phenomenon explains the limited success of various knowledge management approaches, which were only concerned with the knowledge that could be captured, codified and stored - in other words, with explicit knowledge only (Hildreth & Kimble, 2002).

To address this problem, which arises because of the situated nature of knowledge, it is essential to transform distributed individual knowledge into group or organizational knowledge. Knowledge sharing provides a viable solution to this problem, as knowledge sharing has been a natural practice for humans to learn new ideas and information. Story telling, apprenticeship, and face-to-face meetings are some of the most common ways of learning new information within this practice.

Knowledge sharing is a process in which individuals or groups mutually exchange implicit and/or explicit knowledge. It helps translating individual knowledge into group or organizational knowledge; therefore, it assists organizations in managing their knowledge resources more efficiently (Hooff & Ridder, 2004). In addition, knowledge sharing helps build a shared knowledge base among participating individuals and units, which is helpful for group decision making or action-taking processes. It has, therefore,

crucial applications in domains, such as public deliberation and organizational management, which require a shared knowledge base for well-informed group-decision making.

To understand the knowledge sharing process better, Lave and Wenger (1991) conducted case studies of how newcomers become part of various occupational groups that do not have formal training systems for newcomers. Based on the findings of these case studies, they proposed the view that knowledge is a situated cognitive dimension in these groups or activity systems. Various scholars have now come to accept this concept of knowledge as situated and distributed cognition (Boland & Tenkasi 1995; Lave & Wenger, 1991; Suchman, 1987). Their viewpoint is that knowledge is distributed among people and artifacts and cannot be completely captured or codified.

The situated nature of knowledge makes interaction among the participating individuals or groups essential for all knowledge sharing purposes. In the fields of knowledge sharing and knowledge management, the individuals and groups that participate in knowledge sharing processes are known as communities of practice (CoP). These communities of practice evolve around shared common goals or themes, and breed interaction among their participating members. In particular, these communities are environments for the sharing of *implicit* knowledge, which is very difficult to capture or record using traditional knowledge management practices. This is the reason that current research concerning knowledge management and knowledge sharing uses a ‘community based’ approach (Scarbrough & Swan, 2001), in which shared themes and practices among the members of such a community are the basis of knowledge sharing, and not their formal organizational roles.

To explain the evolution of these communities and the knowledge sharing process that takes place in their environments, Lave and Wenger (1991) proposed an alternative form of learning called situated learning through legitimate peripheral participation (LPP). The legitimate peripheral participation (LPP) framework is a theoretical description of how knowledge sharing occurs among the experts, participants and newcomers in a domain. It explains how newcomers become experienced members and eventually old timers in a community of practice or collaborative work.

Based upon their LPP framework, Wenger, McDermott, and Snyder (2002) defined a community of practice (CoP) as "a group of people who share a concern, a set of problems, or a passion about a topic, and who deepen their understanding and knowledge of this area by interacting on an ongoing basis" (Wenger, McDermott, & Snyder, 2002). Since knowledge is a situated cognitive dimension, communities of practice function as knowledge repositories, and are crucial for knowledge sharing, knowledge maintenance, reproduction, and extension (Brown & Duguid, 2001).

Traditionally, these communities of practice have adopted face-to-face communication as a knowledge sharing mechanism in situations, such as in town-hall meetings and brainstorming sessions. Information and communication technologies (ICT) have extended traditional ways of knowledge sharing, as these technologies seem to bridge the space and time dispersion of communities of practice (Huysman & Van Baalen, 2001; Roberts, 2000). ICT applications, such as email and chat rooms, have already paved the way for a new wave of knowledge sharing; innovative web applications are playing an ever-increasing role in bringing dispersed individuals and

communities together and facilitating the exchange of ideas, experiences and other types of information (Scarbrough & Swan, 2001, Huysman & De Wit, 2002).

Fueled by ICTs, knowledge sharing is experiencing an unprecedented growth. Individuals and groups are increasingly able to create, contribute, collaborate and share knowledge with ease. This means that corporations are better equipped technologically to address their knowledge sharing needs, as there are myriad collaborative tools available including portals, weblogs, discussion forums, and wikis to facilitate greater knowledge sharing. There is one catch, however, in this seemingly promising scenario: that is, the complex nature of knowledge sharing and the influence of many factors that contribute to its overall success (Hooff, Elving, Meeuwsen & Dumoulin, 2003; Scarbrough & Swan, 2001). It is very crucial for the overall success of any knowledge sharing initiative that the chosen collaboration technology is effectively suited to that particular context of knowledge sharing.

The current research work in knowledge sharing domain, however, has so far only focused on the broader impact of ICT on knowledge sharing in communities. To further the existing research, we need to examine the unique characteristics of collaborative software, and to distinguish its effects on the knowledge sharing process (Hooff, Elving, Meeuwsen & Dumoulin, 2003). Since communities of practice are the breeding grounds for knowledge sharing, it is only fitting to focus on a ‘community-based’ approach to get a broader understanding of the influence that ICTs exert on knowledge sharing. This may help us understand the suitability of collaborative technologies for knowledge sharing in various contexts.

This study examines the effects of collaborative software on the structure of online communities of practice, and on the knowledge sharing process in these communities. This research also addresses, therefore, the need for explicit comparison of different social software technologies, and in particular, their respective efficacies for the knowledge sharing process. An insight into the existing knowledge sharing process in the communities aided by these technologies would not only enable us to further improve these technologies but also help us in finding new applications of them in domains like learning, organizational communication and public deliberation.

1.2 Problem Statement

The existing research demonstrates that knowledge sharing is critical for the performance and sustainability of organizations (O'Dell & Grayson, 1998; Osterloh & Frey, 2000). Collaborative software, such as discussion forums, weblogs, and wikis, help to bridge the time and space constraints of knowledge sharing. The rapid growth of collaborative technologies, however, offers both a challenge and an opportunity for organizations, which have a vested interest in finding innovative and effective ways to facilitate knowledge sharing in the workplace.

The biggest challenge presented by the rapid growth of collaborative technologies is the need to evaluate and measure the effectiveness of various software tools for knowledge sharing. There are a variety of tools available to facilitate knowledge sharing including portals, discussion forums, weblogs, and wikis; however, each of these tools has their unique characteristics that may lead to advantages or disadvantages depending upon the context of use and specific requirements of knowledge sharing in that context.

The existing research, however, has yet to probe the relationship between the attributes of web-based collaboration technologies and knowledge sharing (Hooff, Elving, Meeuwsen & Dumoulin, 2003). Since knowledge sharing is a complex process with a multitude of variables (Davenport and Prusak, 1998; Hooff & Ridder, 2003), understanding and mapping the influence of existing collaborative technologies on the knowledge sharing process is a challenge. Before attempting to understand the relationship, there is a need to understand the knowledge sharing process thoroughly and the environment in which it takes place.

Online knowledge sharing is taking place in two types of communities of practice, what Hendriks (2002) calls "micro" and the "macro" online communities. Both of these community types comprise theme-oriented online communities. The micro type of community comprises structured, self-contained, centralized communities, such as discussion forums and mailing lists. In this study, I have defined centralized communities as the communities where designated person(s) or group(s) manage and/or moderate the discussion themes and membership issues for the community. Discussion themes could be either very focused (such as, human computer interaction or usability engineering) or very generic (such as, life in the city of St. Paul). The participants have a vested interest in the discussion themes which transforms members' interactions into shared practices of the community. Guided and regular discussion brings a sense of community among the participants. Consequently, a common discussion theme and shared sense of community are the variables that define the virtual boundary and structure of the community.

The macro type of online community, on the other hand comprises a wider online discussion sphere and the online flow of public discussion in formal and informal

settings. These macro communities differ from micro ones in that individual members define their own discussion themes or shared practices. In other words, every participating member of these communities defines and controls the theme for their own knowledge contribution. Finally, these macro communities obtain their structure and sense of community when there is an external trigger that connects these separated, individualistic discussion themes via a shared goal or practice. Similar to micro communities, the shared practices in macro communities can vary greatly from specialized topics and interests, such as politics, engineering, art, and sciences, to very generic ones, such as a common geographical region or daily-life dairies. In this study, I consider this macro type of community as a self-organizing community since these communities comprise scattered individuals or groups that come together to form a community when connected by an external trigger.

This study examines different community structures of centralized and self-organizing communities of practice, and the effects that their unique structural attributes have on the knowledge sharing process. I selected discussion forums and weblog networks as the representative collaborative software systems for the centralized and self-organizing communities, respectively. Online discussion forums are one of the earliest collaborative tools designed for knowledge sharing. Most of these forums are centralized networks in nature in the sense that a governing individual (s) or group defines and enforces the discussion themes, or shared practices, along with participation rules, policies and regulations. Typically, these regulations ensure that discussions in the forum adhere to the stated themes or goals of the forum community. Typically, online discussion systems aim to aggregate public deliberation within a centralized site or forum

(for example, Minnesota E-Democracy, UK E-Democracy, and MeetUp.com, among others). While these centralized online discussion forums have been successful in stimulating deliberation, they have several limitations. These limitations include the tendency to attract the usual activists, difficulties in scaling up beyond their core group and limited breadth of information exchanged (Kavanaugh et al., 2005). The recent advancement of web technologies and applications has, however, expanded the available options and presented an array of easily accessible tools for online communities.

The increasingly popular weblogs (blogs) is one such web technology that affords users easy content generation, publishing, and information sharing. The ease of use and simplicity of collaborative tools (such as weblogs, wikis and aggregator tools like RSS and ATOM) have attracted users who are interested in communicating with others in their social networks, their geographic communities and/or the greater public (Herring et al., 2005). Online users together represent a kind of self-organizing social system that allows them to interact and learn from each other through the exchange of ideas and information (Kavanaugh et al., 2005). The absence of any empirical study on the knowledge sharing patterns in the weblog networks, however, has barred us from conclusive insights about the implications of their network characteristics on knowledge sharing. This study investigates if blogs are different from traditional discussion forums in terms of knowledge sharing patterns and the type of community structure they promote.

To understand the knowledge sharing process in communities of practice the two most important influencing factors are the community's structure (Davenport and Prusak, 1998; Smith and McKeen, 2002) and its interaction modes (i.e., the flow of information)

(Hooff & Ridder, 2003). This study began with a premise, therefore, that exploring the influence of certain collaborative technologies, online discussion forums or weblogs on the structure of communities would lead us to a better understanding of the relationships among the technology, CoP and the knowledge sharing process that takes place in these environments. To investigate these two most critical variables – CoP structure and interaction mode – this study has used a combined framework of legitimate peripheral participation theory (Lave & Wenger, 1991) and weak-ties phenomenon (Granovetter, 1973).

Lave and Wenger (Lave & Wenger, 1991) proposed peripheral participation theory as an alternative form of learning called situated learning to explain the knowledge sharing process and the evolution of communities of practice. Legitimate peripheral participation (LPP) framework is a theoretical description of how knowledge sharing occurs among the core members of a community, participants and newcomers in a given domain. It explains the evolution of communities of practice (CoP) through which newcomers become experienced members and eventually veterans of a community of practice. The theory of 'weak-ties' refers to the notions of close-knit social bonding and indirect, less personal social contacts. The interesting twist in this theory is that in spite of more social trust and frequent interaction, strong-ties communities are less conducive for knowledge exchange compared to weak-ties communities. There is an argument that social systems lacking in weak-ties are isolated, and knowledge would have difficulty reaching masses in those systems (Granovetter, 1973; Granovetter, 1983). Granovetter (1973, 1983) also claims that smaller, tighter networks can be less useful to their members' knowledge sharing needs than networks with lots of loose connections (weak

ties) among participating individuals outside the main network. The dense and bounded networks tend to be isolated from other social networks, which results in an obstruction of the flow of information, and the reach of knowledge to the masses is limited. On the other hand, open networks, with many weak ties and social connections, are more likely to introduce new ideas and opportunities to their members than closed networks with many redundant ties.

The combination of an LPP framework and the weak-ties theory has provided a framework to investigate the relationship between the community structure and knowledge exchange. Figure 1 illustrates LPP theoretical framework:

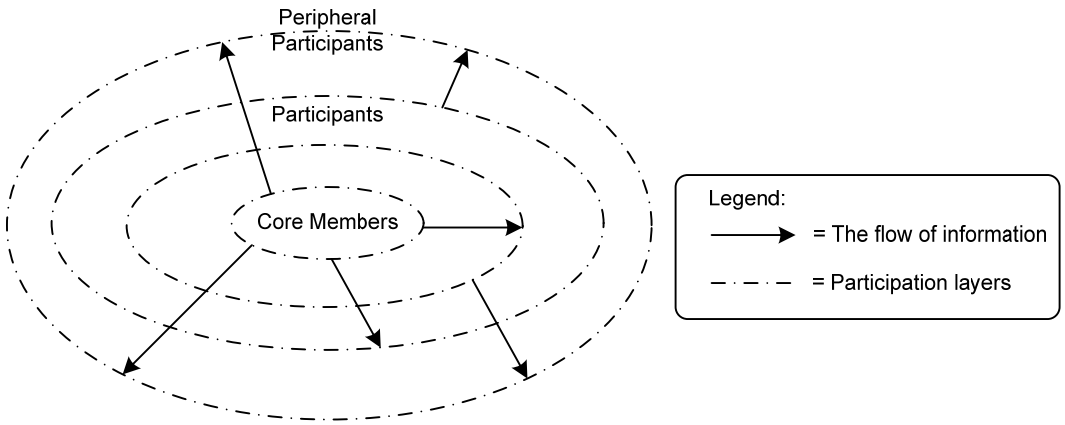

Figure 1: The theoretical framework illustrating a typical CoP structure

The illustration in figure-1 depicts the structure of a community and information exchange that takes place in a community of practice. In this environment, there are three hypothetical layers of participants based upon the participants' perceived authority and activity levels. Dotted lines depict the fluid boundaries of these different participation layers and arrows show the direction of the flow of information during knowledge exchange. In a broader sense, this figure illustrates the evolution of a typical community

of practice in which, through frequent interaction with the core members, newcomers become experienced members. This study used this model in conjunction with weak-ties theory to test for possible differences in CoP structure and patterns of knowledge exchange in different online communities.

1.3 Research purpose

The goal of this study is to investigate knowledge sharing in online communities of practice. To achieve this goal, I examined the structural differences of two types of CoP – centralized and self-organizing communities – and the implications of these structural differences on the knowledge sharing process.

The fundamental differences between these two community structures are the shared practices and theme-orientation; centralized communities have more control over shared practices and designated themes than self-organizing CoPs do. Self-organizing communities obtain a sense of community only when scattered individualistic practices and themes connect to each other via an external trigger or event resulting in a shared practice and goal oriented network. These fundamental differences of shared practices and theme designation between these two CoP types explains why they are centralized or self-organizing communities, and why one community has more control over the type of shared practice and theme compared to the other.

The main research objective of this study, therefore, is to understand the implications of the structure of CoP on knowledge sharing. In addition, this study also attempted to provide a more independent and objective means for classifying online communities of practice, particularly in terms of their structural properties and

information flow, which are the critical parameters for the knowledge sharing process that takes place in these communities.

1.4 Research questions

This study has attempted to answer the following questions:

1. What roles does the emerging collaborative software (namely weblogs and online discussion forums) play in the knowledge sharing process? Are there underlying differences in the knowledge sharing process facilitated by blogs and centralized online discussion forums?
2. How does the structure of a community of practice affect the knowledge sharing process that takes place in that environment? Do self-organizing online communities facilitate knowledge sharing better than centralized communities or vice-versa?
3. How does the flow of information differ in weblog networks and discussion forum communities, and what are the implications of these differences?
4. Compare the effectiveness of centralized and structured community structure versus decentralized, self-organizing community structure for the knowledge sharing process.

1.5 Hypotheses

H-1: Compared to centralized discussion forums communities, blog communities are more conducive to weak-ties phenomenon.

Wellman and Wortley (1990) have shown that some online communities exhibit standard characteristics for strong ties, such as demonstrating frequent, reciprocal, companionable and often supportive and multiplex contact. Moreover, the absence of

physical space constraint in computer-mediated interactions facilitates long-term contact without the loss of relationships that often accompanies residential mobility and schedule conflict. Community members come to regard each other as their closest friends even though they seldom or never met in-person (Hiltz & Turoff 1993).

Although, online ties are strong, it is nevertheless true that online social networks and communities are excellent supporters of weak-ties. There are low logistical and social costs to participating in online social networks; people can participate within the comfort and safety of their own homes or offices, at any time and at their own convenience. Limited social cues available online also encourage contact between weak ties as they take away the social hierarchical blocks in interaction (Wellman, 1996). Many online ties are between persons who have never met face to face, are socially and physically distant, and not bound into work or community structures (Nardi, 2004a).

Granovetter (1973, 1983) has mentioned the need to encourage weak ties bonding to facilitate greater knowledge sharing. Where dense, tight networks with small range are good for conserving existing resources, sparse, loosely bounded networks with a large range are good for obtaining additional resources. Sparsely knit, loosely bounded networks with a large range provide the network members better prospects of obtaining a wide range of information and resources (Granovetter, 1973). The existing research, however, lacks scholarly studies that have tried to compare or analyze if certain online social networking tools work better than others in facilitating weak-ties structures do.

Gruhl et al. (2004) have studied that blogs lead to a self-organizing online behavior which results in sudden bursts of connectivity and interaction around various real-life events and common topics. On the other hand, centralized discussion forums

have traditionally facilitated structured and centralized online communities (Preece et al., 2003). Based on these findings, I hypothesize that an online self-organizing community would have more weak-ties compared to a structured and centralized community.

H-2: Self-organizing communities provide a more conducive environment for peripheral participation compared to centralized, structured communities.

The first hypothesis of this study is that blog CoPs would exhibit weak-ties phenomenon to a larger extent than online centralized discussion forums. If this were a valid hypothesis then it would mean that blog networks are sparsely knit, loosely bounded networks with a large range. This type of environment would have a greater likelihood of attracting new connections, regardless of the life span of such connections. This leads to my second hypothesis that communities with more weak-ties are better suited to attract newcomers or, in terms of the LPP framework, legitimate peripheral participation.

H-3: Self-organizing communities have more multilateral interaction flow among their peripheral members than the peripheral members of centralized communities do.

It is generally accepted that strong ties contribute to intensive resource exchange and result in close communities, whereas weak ties integrate relatively separated social groups into larger social networks (Granovetter, 1973; Wellman & Wortley, 1990). Traditionally, according to the LPP framework, the knowledge or information flow in a community of practice is unilateral, and is directed from core (or expert) members to less knowledgeable participants or peripheral participants. The LPP framework, however, also suggests that the boundaries of CoP are fluid and that once a shared practice or goal changes, peripheral participants of a certain community could become the core members

of another community. Since self-organizing communities achieve a sense of community through common interests and shared practices of their members, therefore, when members of two such communities also have common interests and they connect to each other, they form a larger community. For examples, a blog networks that start with a focused common interest, such as trail conservation, and eventually grow larger as it connects to other blog networks through common members with similar shared practices, such as trail biking or jogging. It is more likely, therefore, that evolution of shared practices would be easier in such environments compared to the centralized communities where shared practices are agreed upon in the beginning of community creation.

1.6 Research Justification

1.6.1 Diffusion of innovation

Diffusion of innovation theory attempts to explain how, why, and at what rate new ideas, practices or technology spread through cultures (Roger, 1995). Further, diffusion of innovation theory claims that the structure of societies and the relations among different societies are shaped largely by the flow of new ideas and information among members (Rogers, 1983).

Since communities of practice are proliferation spaces for the exchange of new information and knowledge sharing, many scholars have studied diffusion of innovation through social networks (Young, 2002). Kumar, Novak, Raghavan, and Tomkins (2003) have observed and modeled sudden bursts of connectivity within the blogosphere based on an analysis of an evolving link structure. Other scholars have examine how new technologies like weblogs are changing the way people use their social networks to discover and spread new information among members (Preece et al., 2003).

Gruhl et al. (2004) have observed that topics discussed in blogs are mostly composed of chatter (that is, ongoing discussion whose subtopic flow is determined by decisions of the authors) and spikes (that is, short-term, high-intensity discussion of real-world events that are relevant to the topic). Here the spikes can be seen as a large scale propagation of information, and the writers and readers are the population exposed to them. For example, when a blogger reads a blog about a topic she doesn't know about or finds an external news source she did not know existed, it could be seen as discovery of information. Once she writes about this new piece of information on her blog and others read it, this takes the form of information propagation.

This study investigated whether there are substantial differences among collaborative software, blogs and centralized discussion forums, in terms of information diffusion. The findings of this study may provide insight into the underlying structures of a social network that influence the way information propagates among its members and to other social networks.

1.6.2 Collaborative learning

Lave and Wenger (1991) originally developed the notion of communities of practice in order to compare apprenticeship or collaborative learning to school based learning. Other scholars used this notion of alternative forms of learning and there have been many studies on collaborative learning and the knowledge sharing process (Connell, 2005; Godwin-Jones, 2003; Phoha, 2001; Soller, 2004). Since blogs and other online publishing tools are relatively new media, there are few studies that evaluate the effectiveness of using them for collaborative learning vis-à-vis older applications, such as online discussion forums and chat rooms (Cohen & Clemens, 2005). This study may help

in filling some gaps in these areas as well broadening the base of collaborative learning and knowledge sharing research.

1.6.3 Public deliberation and Online discourse

Online deliberation is a term associated with an emerging body of practice and research dedicated to fostering purposeful discourse over the Internet. This process of online deliberation includes both knowledge acquisition and knowledge transfer from one participating unit to another. Evidently, knowledge sharing seems to be a factor that may have an influence on the success or failure of the online deliberation process. Knowledge sharing, therefore, may provide a new perspective from which to evaluate the online deliberation and may help ongoing research efforts to measure online deliberation

A public deliberation whether online or offline is essential to a democratic society. However, civic engagement in America has been in decline since at least the early 1960s according to Putnam (2000). By promoting basic norms of a society, such as social trust, deliberation can help foster civic and public engagement (Price & Cappella, 2002).

The capability of ICT to facilitate deliberation is not well understood or established. Various studies have shown that ICTs facilitate storing and sharing information and accessing shared knowledge base (Hooff, Elving, Meeuwsen & Dumoulin, 2003; Scarbrough & Swan, 2001). These capabilities are crucial for the success of any deliberation process. An investigation of particular ICT applications and their contribution to knowledge sharing, therefore, might help us evaluate the effectiveness of deliberation in different situations. It has already been established in various studies that the Internet can facilitate participation by alleviating constraints of

time (i.e., providing anytime/anywhere information access); therefore, collaborative software that foster knowledge sharing may help to stimulate discussions or deliberation and greater civic participation (Patterson and Kavanaugh, 1994; Schuler, 1996; Davis and Owen, 1998; Wilhelm, 2000; Norris 2001).

1.6.4 Self-organizing networks and collaboration

There have been a few recent attempts to support collaboration and drive innovation within organizations by using different social networking ICT applications. The design process and features of these applications have been inspired by innovative applications, such as discussion forums and blogs (Cohen & Clemens, 2005). These applications can distribute responsibility for content creation, commentary and quality control across a community of users (i.e., writers, commentators, and readers). This trend portrays discussion forums and blogs as collaborative technologies and provides users with the tools they need to self-organize for knowledge sharing purposes.

Self-organization models for human behavior have already been used in economics (Krugman, 1996) and in studies of computer supported cooperative work (Eriksson and Wulf, 1990; Wulf, 1999). There is, however, a lack of research on the effectiveness of collaborative technologies for self-organization.

CHAPTER 2. LITERATURE REVIEW

2.1 Knowledge Sharing

Information is the data that reduces uncertainty about a proposition whereas knowledge is information transformed into a capability for effective action. This means knowledge produces competence leading to an effective action. This precise definition is the motivation behind great emphasis on the development of methods for turning information into knowledge (Marshall, 1997). While information as a growing part of organizations' asset base is often efficiently captured and stored, capturing knowledge and sharing it across the organization still requires a more insightful strategy to be effective (Hislop, 2002).

The obstacles in capturing knowledge are due to the dual nature of knowledge (Hildreth & Kimble, 2002). Knowledge can be divided in two categories: explicit knowledge and tacit knowledge. Explicit knowledge is easy to capture, codify, and store in physical objects, however, tacit knowledge or implicit knowledge resides in people's minds and situations, and is, therefore, difficult to access or capture. The concept of the tacit dimension of knowledge has come from the work of Michael Polanyi (Polanyi, 1983), who exemplified tacit knowledge through a famous aphorism: "we know more than we can tell".

Nonaka and Takeuchi (1995) highlighted the criticality of tacit knowledge for organizational innovation using the case studies of Japanese companies. They observed that Japanese companies were more innovative because they were able to collectivize successfully individual tacit knowledge to the organizational level of knowledge. This transformation of individual tacit knowledge to organizational tacit knowledge is the most critical and helpful aspect of knowledge sharing.

Knowledge sharing is a process in which individuals or groups mutually exchange implicit and/or explicit knowledge. Since, it helps in translating individual level into group or organizational level, knowledge sharing assists organizations in managing their resources more efficiently (Hooff & Ridder, 2004; Nonaka & Takeuchi, 1995). Knowledge sharing also helps build a shared knowledge base among participating units that is helpful for group decision making or action processes. It has, therefore, a role in other phenomena, too, such as organizational decision making, collaborative work, and public deliberation that requires a shared knowledge base for well-informed group-decision making.

Knowledge sharing also plays a role in capturing tacit information. For example, the World Wide Web has large amounts of information available and it is increasing every day. This information base is transformed in to knowledge with the help of mainstream knowledge sharing technologies, such as instant messaging, e-mails, discussion forums and more recently blogs (Rosenbloom, 2004). To understand the knowledge sharing process and the environment in which it takes place, we need to understand its role in communities of practice.

2.1.1 Communities of Practice

Knowledge sharing requires a continuous interaction among the participating individuals or groups called *communities of practice* (CoP) in the field of knowledge sharing and knowledge management. CoPs crucial for knowledge sharing because they are the repositories for knowledge, and its maintenance, reproduction, and extension (Brown & Duguid, 2001).

A CoP defines itself along three characteristics (Wenger, 1998):

Objective – The common goal as understood and continually renegotiated by its members.

Sense of community - Mutual engagement that binds members together into a social entity.

Shared practices - The shared knowledge base and repertoire of communal recourses (routines, artifacts, vocabulary, and styles) that community members have developed over time.

A community, however, does not necessarily imply co-presence, a well-defined, identifiable group, or socially visible boundaries (Lave & Wenger, 1991). There are online communities that have emerged around the usage of discussion forums, weblogs and news groups, therefore, that seem to fulfill the criteria for being communities of practice (Wellman, 2001). These communities have the following characteristics (Whittaker, Issacs, & O'Day, 1997):

• Members have a shared goal, interest, need, or activity that provides the primary reason for belonging to the community.

• Members engage in repeated, active participation and there are often intense interactions and shared activities occurring between participants.

• Members have access to shared resources and there are policies for determining access to those resources.

• Reciprocity of information, support and services between members.

These characteristics are an extension of the communities of practice characteristics that Wenger (1998) formalized. These types of online communities, therefore, can be fittingly treated as communities of practice.

Online communities that act as CoPs provide shared space for their members and effuse relationships among members. The fact that online interactions are capable of fostering traditional community interactions is established in various case studies with the Blacksburg Electronic Village (BEV) being among the (Cohill & Kavanaugh, 2000). best known. In Blacksburg, research indicates that 'computer networks are not just reinforcing but even expanding existing social networks within an existing geographic community' (Kavanaugh, 1999).

Interactions in a community of practice occur in the context of built environment or structure of community. The characteristics of this environment or community structure facilitate certain modes and forms of interaction, while discouraging others. For a given context of knowledge sharing and interactions taking place in a community of practice, it is important to understand the underlying structure of a community as well the implications of this structure on the way knowledge sharing occurs.

2.1.2 Forms of Community: Strong Ties and Weak Ties

Among other characteristics that define communities are social relationships and the overall community structure that emerges out of knowledge exchange. These two characteristics have implications for the structure of a community as well as on the community's shared practices and knowledge base (Lave & Wenger, 1991).

Social network theory views social relationships in terms of *nodes* and *ties*. Nodes are the individual actors within the networks, and ties are the relationships between the

actors (Scott, 1991; Wasserman & Faust, 1994). There can be many kinds of ties among the nodes. In its most simple form, a social network is a map of all of the relevant ties between the nodes being studied. The strength of ties is, therefore, a relational, network concept that influences the shape of the social network.

Granovetter (1973) introduced the concept of weak ties and its importance in the domain of knowledge sharing. Strength of ties mostly involves closeness of bond; therefore, it can be characterized by a combination of interaction frequency, reciprocity, companionship and supportive contact. Wellman (1992) characterizes strong ties as intimate relationships through voluntary investment and companionship with the tie partner, frequent interactions over a prolonged period, and supportive, multiplex contact.

Granovetter (1973, 1983) claims that smaller and tighter networks are less useful to their members' knowledge sharing needs than the networks with lots of loose connections (weak ties) to individuals outside the main network. The dense and bounded networks tend to be isolated from other social network, which results in an obstruction of flow of information, and reach of knowledge to the masses is limited. On the other hand, open networks, with many weak ties and social connections, are more likely to introduce new ideas and opportunities to their members than closed networks with many redundant ties.

Dense, bounded groups have considerable and frequent communication among members. By contrast, few members of sparsely knit, loosely bounded networks communicate directly and frequently with each other. The theory of weak-ties proposes that a group of individuals with connections to other social worlds is likely to have access to a wider range of information.

ICT application and online social network tools support both dense, bounded groups and open, loosely bounded networks. They provide a space and time bridging platform to maintain strong ties, such as those among family members, and they are especially good for forming and maintaining weak-ties (Wellman & Gulia, 1999). For example, special interest discussion forums, blogger communities (blogrolls) and list-serves all connect users to numerous individuals through weak-ties. Weak-ties networks are larger and more diverse than dense and bounded strong-ties networks, therefore, weak ties networks bring their members more opportunities for obtaining and disseminating information. This argument can lead us to the proposition that while dense and strong-ties networks are better for collective action, open and self-organizing networks facilitate the exploration and dissemination of new information.

2.1.3 Self-organizing social systems

Online self-organizing social systems are online communities or networks that allow large numbers of individuals to self-organize in a highly decentralized manner in order to solve problems and accomplish common objectives (Wiley & Edwards, 2002). Their distributed and highly decentralized nature could prove disruptive to traditional notions of online collaboration and knowledge sharing. Self-organizing systems at times behave like nonlinear dynamical systems that under certain conditions exhibit a phenomenon known as chaos though over time some order typically emerges (Ott, 2002).

Maeterlinck (1927) studied self-organization in white ants and found out that white ants forage for resources, store resources, provide needed resources to others at the proper place and time, discriminate between optimal sources of food, build nests, hives,

or domes, and solve a variety of other complex problems without any visible centralized controlling mechanism.

ICTs facilitate self-organization of social networks or communities as it fosters open communication across space and time barriers while the lack of status cues fosters connections across hierarchical or other forms of status barriers (Eveland & Bikson 1988). The relative lack of social presence online fosters relationships with network members who have more diverse social characteristics than is normally encountered in person. This allows relationships to develop because of shared interests rather than be stunted at the outset by differences in social status (Hiltz & Turoff 1993). The phenomenon is a technologically supported continuation of a long-term shift to social networks organized by shared interests rather than by shared locality (Wellman 1992). When a computerized conference focuses on concerns that are held dear, physically distant members may have more in common than they do with their neighbors.

The weblog is one such ICT application that has shown the characteristics of self-organizing social networks. Kumar et al.(2003) have studied and modeled sudden bursts of connectivity within the blogosphere based on an analysis of the evolving link structure. Their study shows that dispersed individual bloggers come together to discuss the events and this collective conversation gets bigger to take the form of a spike.

Gruhl et al. (2004) have observed that generally discussion in Weblog communities is composed of chatter (ongoing discussion whose subtopic flow is determined by decisions of the authors), although, at times spikes (short-term, high-intensity discussion of real-world events that are relevant to the topic) appear and a large number of bloggers are exposed to this spread. Here the spikes can be seen as large-scale

propagation of information and the writers and readers are population exposed to them. This is another indication of self-organization among weblog communities when the micro-level entities change their behavior accordingly and self-organize themselves to participate on a greater scale. The figures below show the remarkable correlation between different related queries from their study.

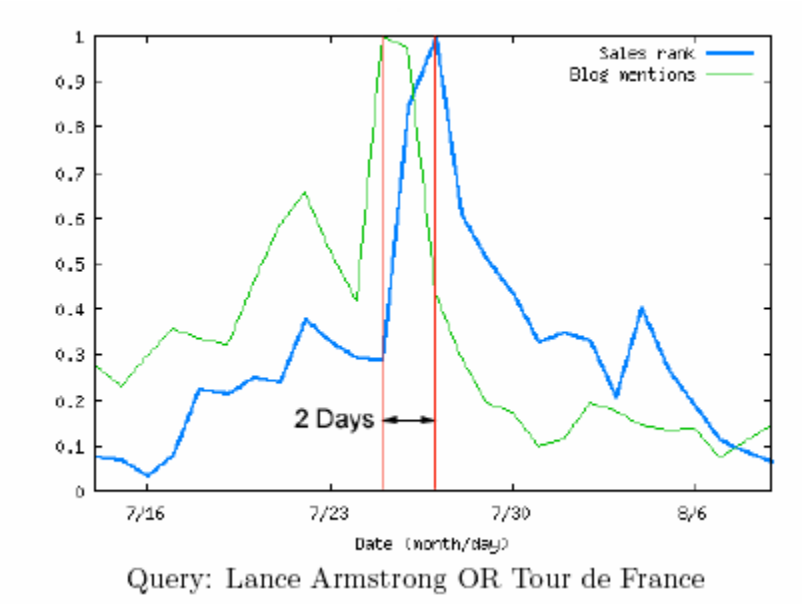

Figure 2: The Correlation between two topics on the blogosphere (Gruhl et al. 2004)

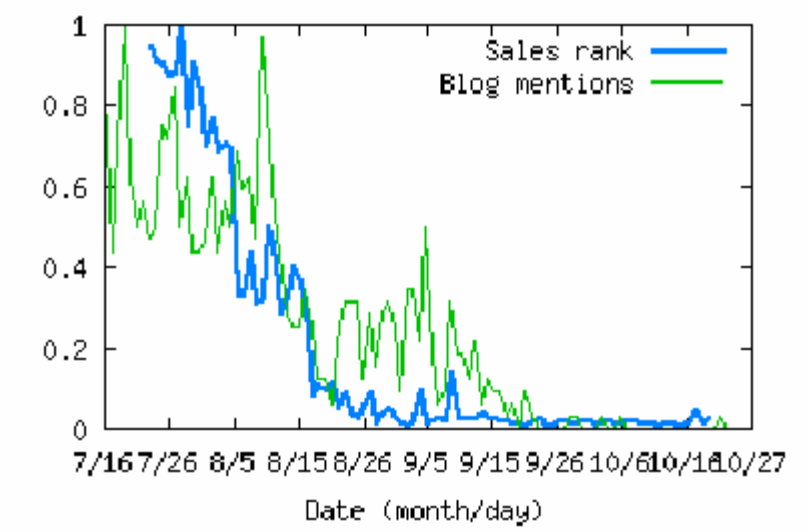

Figure 3: Falling sales and the blogs mention of "The Notebook" (Gruhl et al., 2004)

More recently, Eriksson and Wulf (1999) have begun exploring the relationships between self-organizing systems and the notion of computer-supported collaborative work.

Current research brings us to the point where the self-organizational potential of human social systems has been recognized and documented, and investigations are beginning into the ability of networked technology to facilitate this self-organizing activity for individuals who are geographically distributed.

2.2 Weblogs (Blogs)

2.2.1 Introduction

A weblog (usually shortened to blog, but occasionally spelled web log or weblog) is a web-based publication technology consisting primarily of periodic articles, most often in reverse chronological order. Weblog is the most recent application in the domain of computer mediated communication technologies that enables user to publish content online with ease. Blogs can be hosted by dedicated blog hosting services, or they can be run using blog software on regular web hosting services.

The entry (also called a post or a message) is the basic unit of a blog. The entries are where the content of the blog is located, and the most recent entry, displayed first on the page, is what a visitor is most likely to see at the first glance and read first. The writer of these blog entries is normally called a blogger. Besides the entry text itself, entries usually have both a header and footer with additional pieces of information (Herring et al., 2003). A blog entry typically consists of the following:

- *Title* - main title, or headline, of the post.
- *Body* - main content of the post.
- *Permalink* – the URL of the full, individual article.
- *Post Date* - date and time the post was published.

A blog entry optionally includes the following:

- *Comments* - comments added by readers
- *Categories* (or tags)- subject categories that the entries discuss
- *Trackback (Occasionally Pingback)* - links to other sites that refer to the entry

Only the newest entries are displayed on the main blog page while older entries are usually arranged in archives where they can be accessed on a later date. These archives are usually categorized on basis of dates/months/years, various topics and subject tags. Many blogs nowadays also allow readers to post comments to individual entries in a similar fashion as they would do in replying to threads on a discussion forum. These blogs differ in styles and ways to present content, navigation and functionality.

2.2.2 The basics of blogs

Every time a blogger puts an update on an blog, a post is created. Posts are also sometimes called entries. A blog entry typically consists of the following:

- Title: main title of the post or entry
- Body: main content of the post or entry
- Comments: Blogs are often referred to as conversations, and it's the ability of your readers to leave comments on each post you make to your blog that creates the feel of a conversation. Comments are usually time-stamped and identified by the author's name and perhaps a link to their Web site or blog. On some blogs, comments are threaded so that readers can comment on other comments, but on most blogs comments are simply displayed reverse chronologically.
- Category (or tags): Categories permit a blogger to subdivide content, putting posts about politics into one basket and posts about celebrities in another. Categorization helps readers read only what they are most interested in and is a good tool for those scanning a blog's archives.
- Post scheduling: time and date when the entry was published
- Permalink: a link pointing to the archived, permanent URL of the entry, done for the convenience of people who want to link to a specific entry from their blogs or other places.
- Trackback: a function that allows the tracking of entries in other blogs that have linked back to each entry and the ability to search on similar entries based on keywords. In Figure below, functioning of trackback feature is explained.

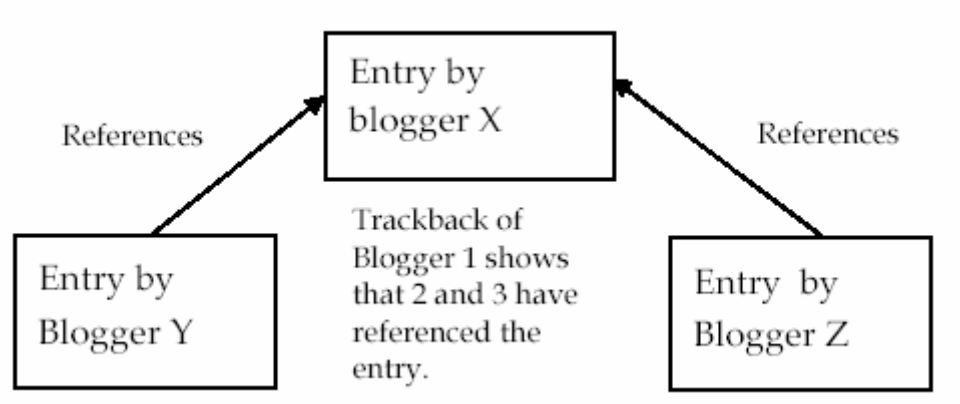

Figure 4: The concept of trackback in weblogs

Other basic characteristic features and terms associated with a weblog include:

Blogroll/Lists: The long list of other blogs alongside the posts in a blog is called blogroll, a list of the blogs read by the blogger. Sometimes lists are also kept to recommend books and other media, as well.

News Aggregation: Many blog software packages allow users to pull in and display the RSS or Atom feed of another blog. This is useful if a user wants to create a site with constantly updated content fed by other blogs. For example, a blogger who posts about politics could pull in the feeds of other political blogs without visiting other blogs.

RSS / Atom: In the blogosphere, syndication is a important. With millions of blogs available, many consumers use news aggregators, or readers, to pull in posts and read them, rather than visiting 150 blogs every day. RSS and Atom are two features of blog syndication.

Ping: There are several blogging tracking Web sites where users can search for other blogs and look for recent posts. If blog software 'pings' those sites when a user posts, that post gets included in the ping site's index, potentially increasing traffic of user's weblog.

The Blogosphere: The social network of weblogs (also known as blogspace or blog ecosystem)

2.2.3 Previous research

The tradition in weblog research is new and the methods used for the research have mostly been adapted from the field of research. This has led to a wide variety of approaches to researching blogs. In the field of HCI, most of the research has been either quantitative research on characteristics of blogs, such as propagation of information or studies of various characteristics of links (Herring et al., 2004) or qualitative research on blog entries (Krishnamurthy, 2002).

Interview methods and ethnography have been employed in studying the habits and viewpoints of the bloggers (Henning, 2003; Nardi et al., 2004). Nardi et al. (2004) have done ethnographic studies on bloggers' blogging practices and come up with different reasons for blogging. Their methods consisted of doing both quantitative and qualitative analysis of blogs. They found five main blogging motivations: documenting one's life; providing commentary and opinions; working out emotional issues; thinking by writing; and promoting conversation and community.

Based on their study of blog entries related to September 11 tragedy, Krishnamurthy and colleagues (2002) proposed a classification for blogs into four types along two dimensions: personal–topical and individual–community. These classifications are still useful, though the blogging community has expanded since 2001 and the lines

between personal and topical, for example, are often blurred. Krishnamurthy also claims that the most insightful and controversial posts attract the most comments hence the most readership.

Of the approximately 4.12 million weblogs hosted on the eight leading blog hosting services in 2003, as many as 66 % had been either permanently or temporarily abandoned. Since creating blogs is easy, it seems that many start them without actually continuing to commit to updating them regularly. The size and demographics of the blogosphere show that a majority of blogs were started by teenagers or people in their twenties with a slight majority being female bloggers (60 %) (Henning, 2003). Of the approximately 4.12 million blogs, Henning also makes a distinction between the "blogging iceberg" – popular blogs with tens of thousands of daily readers that are updated frequently and millions of blogs with "nano-audiences" consisting of friends and family members.

A random sample of 203 weblogs by Herring and colleagues (2004) found that blogs shared characteristics with offline genres, such as diaries and newspaper editorials. This study also compared weblogs with other online and offline types of publishing and discussion, placing them somewhere between asynchronous discussion forums and standard WebPages. According to them, blogs share features with both types of web publishing and discussion and "bridge the technological gap" between the two. For example, blogs offer the multimedia elements of web pages but also give readers some communication rights in form of comments. They found out that commenting and inter-blog linking was not as common as their perceptions of blogging had led them to expect.

Kumar et al.(2003) have studied and modeled sudden bursts of connectivity within the blogosphere based on an analysis of the evolving link structure. Their study shows that dispersed individual bloggers come together to discuss about the events and this collective conversation gets bigger to take the form of a spike. Gruhl et al. (2004) furthered the work of Kumar et al. (2003) and found out that generally discussion in Weblog communities is composed of chatter (ongoing discussion whose subtopic flow is determined by decisions of the authors), although, at times spikes (short-term, high-intensity discussion of real-world events that are relevant to the topic) appear and a large number of bloggers are exposed to this spread.

2.3 Online Discussion Forums

A discussion forum is a facility on the World Wide Web for holding discussions, or the software used to provide the facility. A sense of community often develops around forums that have regular users. Technology, computer games, and politics are popular areas for forum topics, but there are forums for a huge number of different topics. Discussion forums are also commonly referred to as web forums, message boards, discussion boards, Internet forums, discussion groups, bulletin boards or simply forums**.**

Discussion forums support connections with large numbers of people, providing possibilities to form a loosely connected social network around common interests or objectives. Because all members of newsgroups and discussion groups can read all messages -- just as in a cafe conversation -- groups of people can talk to each other casually and get to know other individuals.

The emergence of discussion forum Communities of Practice can be traced back to the emergence of Bulletin boards (Preece et al., 2003b). Bulletin boards are designed

based on the metaphor of a physical bulletin board. People post messages to the board and they are displayed in various ways. Usually the messages are threaded which means that messages on the same topic are associated with each other. The first message forms the beginning of the thread and later responses are stacked beneath it. During the last few years, these systems have offered many fine enhancements: search engines enable users to search on topics, user name, date; emoticons; private conversation spaces; links to email, user profiles and web pages; and graphical two-dimensional pictures and avatars.

2.3.1 Bulletin board systems

Bulletin Board systems started out as simple posting centers that ran on a single home computer that people could be connect to one person at a time through a simple modem connection. The post and reply prototype was the simple beginnings of the asynchronous threaded discussion. The system had to be asynchronous because only one person at a time could connect through a modem to the home computer running the BBS.

The Citadel BBS software was born in December of 1981, a new format of BBS software based on the concept of "rooms". Each room is essentially a message area, with the name of the room suggesting the topic of discussion within the room. One of the unique features of Citadel is that users can create their own rooms; thus, the Citadel grows larger because the users themselves become part of the construction crew. This is the birth of users spawning their own micro communities inside a macro structure.

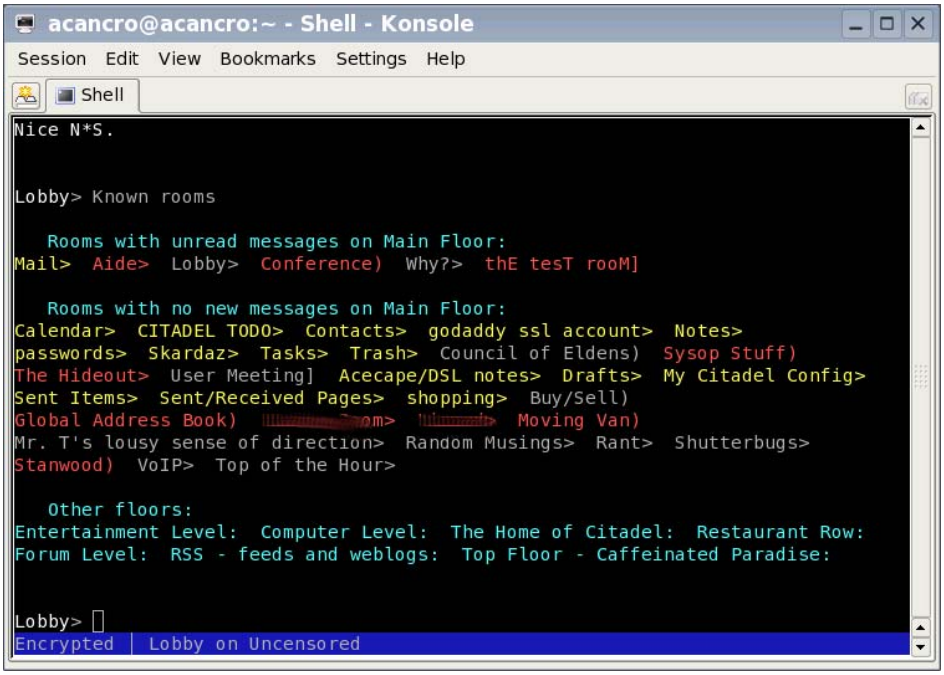

Figure 5: The screenshot of BBS running citadel

Wenger (1998) stresses the importance of communities being able to fracture and grow on their own to meet the changing needs of community members. While simplistic in nature the Citadel software of 1981already had the makings of a mature online CoP. It could also be shown that BBSes were the true starting points for community networks.

2.3.2 News Groups

The emergence of news groups were overlapped with the decline of local BBSes. This emergence of news groups is an important factor in tracing the development of today's CoPs. News Groups were a key player in the early developmental stages of the Internet. These early Internet News Groups were asynchronous thread discussions that were organized by subject of interest. They were essentially local BBSs moved to the global environment of the Internet. These News Groups could be access via client software referred to as News Group Readers. The news group readers allowed you to

access the multitude of news groups and then organize, read and reply to postings. One of the more popular client readers was “Agent” (see figure below).

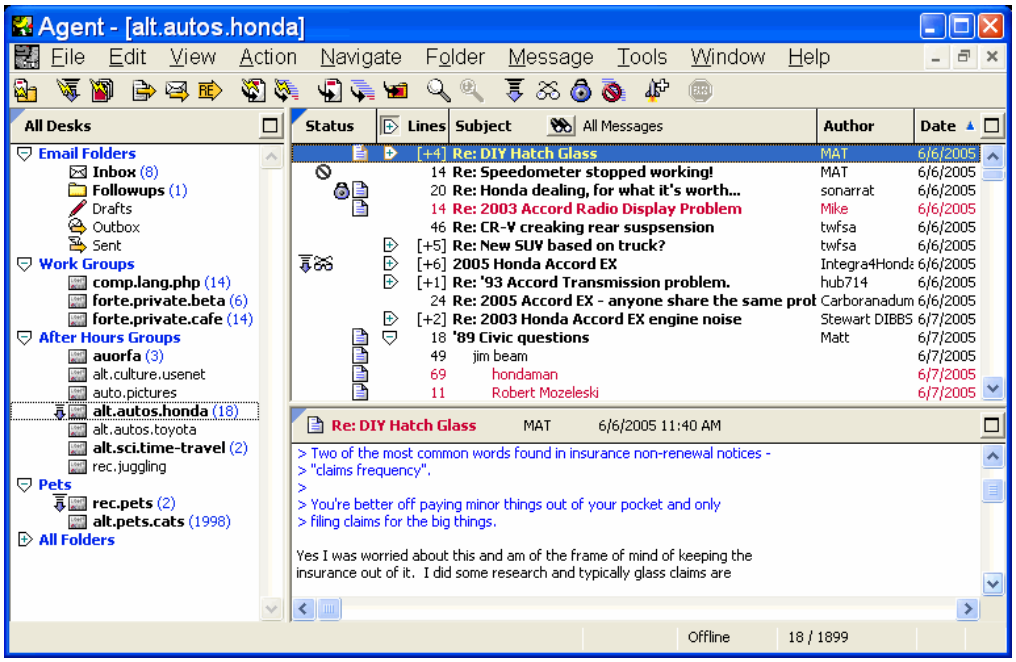

Figure 6: The screenshot of Agent user interface

2.3.3 Differences between weblogs and discussion forums

Table 1: Differences between weblogs and online discussion forums

	Online Discussion Forums	**Weblogs**
Information flow	Discussion flows linearly, with participants replying to each other within a thread.	Discussion is not linear as it moves from one blog to another. The participants reply to or reference each other in their blogs using hyperlinks.
Organization	Entries are organized under different boards and thread topics.	Entries in individual blogs only, organized in different categories or in archives.

Ownership	Community	Personal / Community
Moderation	Centralized moderation	Moderated by the author of blog only
Chronology	Depends on the forum technology, the initial entry in a topic or thread is usually displayed on top.	Entries are displayed reverse-chronologically.
Referencing	Entries are connected to each other by being arranged under a thread or topic, and referenced with quotes.	Entries are distributed throughout blogs, connecting and referencing is done with help of links, quotes and trackback
Participation	Anyone can usually read the forum topics, though posting and replying usually requires registration to the forum	In order to post the entries one must have a personal blog or be a part of a collaborative blog. Visitors can post comments to the entries without having to register.

2.4 Social Network Analysis

A social network is a social structure made of nodes, which are generally individuals. It indicates the ways in which nodes are connected through various social relationships ranging from casual acquaintance to close familial bonds. The presence of relational information is a critical and defining feature of social networks (Scott, 1991; Wasserman & Faust, 1994). Social network analysis is a set of methods for mapping and analyzing relations among people and groups. The fundamentals of social network analysis are grounded in graph and system theories.

The social network analysis provides ways to think about social relationships and a set of structural variables for a given social network, such as the density of the network, how tightly it is bounded, and whether it is diversified or constricted in its size and heterogeneity, how narrowly specialized or broadly multiplex are its relationships, and how indirect connections and positions in social networks affect behavior (Wellman, 1996). Moreover, social network analysis allows measurement of structure and systems, which would be almost impossible to describe without relational concepts, and provides tests of hypotheses about these structural properties.

The internet has posed new challenges about the nature of social networks and opened new perspectives for social network analysis (Garton, Haythornthwaite, & Wellman, 1999; Wellman, 2001). In particular, the hypertextual structure of the web makes linking explicit. These links function as a relationship between different online places or artifacts or simply nodes. As a consequence, as Jackson (1997) points out, social network analysis is ideally suited to the web environment. In addition to being directly observable, hyperlinks are easy to map and analyze with specialized software programs.

Most recently, social network analysis methods have begun to be applied to weblogs. Kumar, Novak, Raghavan, and Tomkins (2003) observed and modeled temporally concentrated bursts of connectivity within blog communities over time, concluding that the blogosphere has been expanding rapidly since the end of 2001, “not just in metrics of scale, but also in metrics of community structure and connectedness”. Gruhl and colleagues (2004) have studied the propagation of information in the blogosphere. Marlow (2004) uses social network analysis to identify "authoritative" blog

authors, and compares them with measures of opinion leadership and authority in the popular press. Delwiche's (2004) study found that the most authoritative authors on the blogosphere are actually members of the A-list.

2.4.1 Fundamental concepts in Social network analysis

Actor: Social network analysis attempts to explore and learn about the linkages among social entities and underlying inferences of these linkages. The social entities in here are referred to as *actors.* Actors could be discrete individual, corporate, or collective social units. For example, people in a group, departments within an organization, or different nations in world could be treated as actors in a social network analysis (Wasserman & Faust, 1994).

Affiliation Networks: Affiliation networks are a special type of social networks that demonstrate collection of actors rather than just linkages between pairs of actors. They consist of a set of actors and a set of events (Wasserman & Faust, 1994). These are sometimes also referred as membership networks (Breiger, 1974) or hypernetworks (McPherson, 1982). Discussion forums and weblogs both are the examples of the affiliation network types.

The motivation behind studying affiliation networks is that affiliations of actors with events provides an opportunity for direct linkage, either between the actors through membership in different events, or between events through common members (Wasserman & Faust, 1994).

Relational tie: In social networks, social relations or ties connect actors to each other. The tie, therefore, is a linkage between a pair of actors (Scott, 1991; Wasserman & Faust, 1994). Since Social network analysis (SNA) focuses on relations and ties in studying the

behavior of actors. Thus, the positions of actors within a network and strength of ties between them are of critical importance for the analysis of a given network.

Strength of ties mostly involves closeness of bonding between actors. There is general agreement that strong ties contribute to intensive resource exchange and close communities, whereas weak ties provide integration of relatively separated social groups into larger social networks (Granovetter, 1973; Wellman & Wortley, 1990).

2.4.2 Fundamental properties of Social networks data

Betweenness centrality: Degree an individual lies between other individuals in the network; the extent to which a node is directly connected only to those other nodes that are not directly connected to each other; an intermediary; liaisons; bridges. Therefore, it is the number of people who a person is connected to indirectly through their direct links

Boundedness: Boundedness refers to the proportion of network members' ties that stay within the boundaries of the social networks (Laumann, Marsden, & Prensky 1983). Networks can be bounded groups or permeable ramifying networks in which people can reach out widely to connect with others.

Centrality: Centrality is a measure of the importance of a node in a network. It assigns relative scores to all nodes in the network based on the principle that connections to nodes having a high score contribute more to the score of the node in question (Wasserman & Faust, 1994).

Closeness: The degree an individual is near all other individuals in a network (directly or indirectly). It reflects the ability to access information through the "grapevine" of network members. Thus, closeness is the inverse of the sum of the shortest distances between each individual and every other person in the network.

Density: Individual-level density is the degree a respondents know one another. Network or global-level density is the proportion of ties in a network relative to the total number possible (sparse versus dense networks) (Scott, 1991; Wasserman & Faust, 1994).

Dense, bounded networks usually have a small range because a large network becomes unbounded relatively quickly. As the number of network members increases the population basis for more diversity within the network increases. Moreover, as the number of network members increases arithmetically, the number of connections required to sustain full connectivity increases geometrically (Scott, 1991; Wasserman & Faust, 1994).

Degree: The count of the number of ties to other actors in the network. It has two types: Indegree & Outdegree. For the study of weblogs and online discussion forums, in-degree could be measured as number of inbound links from other blogs and discussion forums while out-degree could be measured as the outbound links to other online places (Herring et al., 2005).

Network range or size: The range of a network describes how large and diverse is the population within its boundaries (Wasserman & Faust, 1994).

Reach: The degree any member of a network can reach other members of the network. Graph theory states that two nodes in a graph are adjacent if there is a line between them, and they are reachable if there is a path between them.

Reciprocity: Reciprocity arises when pairs of actors have a bidirectional response link; that is, they respond to each other. In SNA, reciprocity is also known as mutuality or mutual dyads (Wasserman & Faust, 1994). Wellman and Gulia (1999) defined reciprocity

as one of the defining attributes of any community. Herring et al. (2005) also included reciprocity as a behavioral indicator for the emergence of a community in their study of conversations in the blogosphere.

CHAPTER 3. METHODS

3.1 Research Design

Wenger (1998) observed during observational case studies that communities of practice evolve over time, and are difficult to build in a short time. In addition, online communities, such as weblogs and discussion forum communities, are dispersed in space and time, which reduces the feasibility of a controlled experimental design for the study. These observations indicate that any external control would seriously weaken the context in which these communities exist, therefore, this study is an analytical field study that has used Data mining (Archival research) for data collection and Social Network Analysis for data analysis.

Compared to the research in traditional communities of practice field, the research in online communities, such as weblog networks and discussion forum communities, is relatively young. The methods used for the research of online communities, however, draw heavily from the existing methods used in the traditional communities of practice research. Although there are a wide variety of approaches to research CoP, most of the existing research has used the social network analysis technique for the analysis of various communities including the weblog networks and discussion forum communities (Herring et al., 2004; Herring et al., 2005; Kumar et al., 2003; Wellman, 1996).

In addition, the internet and hypertextual environments are merely the extension of traditional relationships and communities of practice environments into the virtual world, which makes them ideally suited for the social network analysis (Garton, Haythornthwaite, & Wellman, 1999; Wellman, 2001). To elaborate, the hypertextual structure of the web makes linking explicit and these links function as a relationship between different online places or artifacts or simply nodes. As a consequence, as

Jackson (1997) points out, social network analysis is ideally suited to the web environment. In addition to being directly observable, hyperlinks are easy to map and analyze with specialized software programs.

The main objectives of this research were to study the impact of community structure on peripheral participation, on information flow, and knowledge sharing in that community. For the earlier mentioned reasons, the current study uses social network analysis methodology to investigate the structural and interaction modal characteristics of the weblog networks and discussion forum communities. Of course, as communities are informal groups that do not have clear and fixed boundaries, it was difficult to be sure before investigating whether the studied network was a community. To overcome this limitation, the study has distinguished characteristics that encourage the emergence of communities. The selected characteristics included the size of network, location, and overall theme or mission of the network. Since this is an exploratory study, I chose to use four existing online communities selected based on the above mentioned community characteristics. This study has attempted to examine the structural and information flow properties of the selected online communities using three commonly used social network measures: density (centrality), and connectedness (Outdegree and Indegree).

3.2 Independent variables:

The independent variable chosen for this study is the community control structure, because the objective of this study was to investigate the impact of community structure on knowledge sharing. For the purpose of this study, I chose two common types of community structures in existence: centralized communities and self-organizing communities. For ease and better control of research design, I define the centralized

communities as the communities, in which an individual or governing body defined the participation rules, policies and regulations for the community. Decentralized communities are the communities in which every individual member or group(s) defines their own discussion themes and goals, and in which the shared practices of community emerge out of a common objective or goal connected by external events or triggers. As explained earlier, discussion forums have shown the above-mentioned characteristics of centralized communities while weblog networks exhibit the characteristics of self-organizing communities. The two levels of independent variable, community structure, are:

- Centralized community structure: represented by discussion forum communities.
- Self-organizing community structure: represented by weblog networks.

3.3 Dependent variables:

The LPP framework, which is a theoretical framework for this study, suggests three roles of participants in any community of practice. These roles are based on an individual's activity and perceived authority levels. For example, a core member in a community would have high authority while a newcomer would have a comparatively low level of authority. This is important to note, however, that the authority here is not formal authority but perceived authority of that individual by other community members.

Based upon the above interpretation of the LPP framework, the current study has three dependent variables strength of ties, peripheral participation levels, and levels of peripheral multi-lateral interaction.

3.3.1 Strength of ties

A social network is a social structure made of nodes, which are generally individuals. In its most simplest form, a social network is a map of all relevant ties between the nodes of that network. The strength of ties or relations is, therefore, a relational, network concept that influences the shape of the social network. For a given network, the network's Degree Centrality is a measure of strength of ties since strength of ties mostly involves closeness of bonding between actors or in other words, overall density of network.

Degree centrality is the count of the number of ties that a given community member (actor or node) has with other members in the network (Faust, 1997). In undirected networks, e.g., relationship networks, nodes differ from one another only in number of connections they have. With directed networks, (e.g., Information networks) however, it is important to distinguish the degree centrality based on in-degree from the degree centrality based on out-degree. This is because high in-degree centrality shows the prominence or high-authority of that particular node, while high out-degree centrality indicates relatively high connectedness or activity levels.

3.3.2 Peripheral Participation Levels

As per the LPP framework, a community of practice structure has three different yet fluid layers of participation. The innermost layer comprises core members, the members who enjoy high authority in that community. Similarly, the outermost layer consists of peripheral participants. These are newcomers in that community and have comparatively low authority among the community members.

In social network analysis, authority is conceptualized as a particular pattern of social ties, in which the authority of a node is the number of unique in-degrees it attracts. Since the node in this context does not have any control on what kind of and how many in-degrees it would receive, this type of authority is called structural authority. Structural authority is different from the social authority that emerges from the hierarchical order of a node in a society or organization.

The peripheral participation level of a node, therefore, depends upon two variables: authority of that node in the community and its activity level. In a social network, a node's structural authority determines its position among three participation layers. Node with relatively high authority would occupy core layer while nodes with lower authority would be on periphery of the network. To calculate the participation levels of these peripheral nodes, I need out-degrees values since out-degree of a node in directed networks dictates its comparative 'Activity level'.

3.3.3 Levels of Peripheral Multi-lateral Interaction

The LPP framework states that in traditional communities of practice flow of information or knowledge is from inner layers to outer layers (Lave & Wenger, 1991). In other words, core members of a community disseminate information and knowledge while newcomers or peripheral participants work as absorbents. For the objectives of this study concerning information flow, I needed to first identify peripheral participants based on relative authority levels of nodes followed by the identifying peripheral nodes that possess high activity level.

Table 2: Summary of Independent and Dependent Variables

Independent Variables	Dependent Variables
Community Control Structure	Strength of Ties
	Peripheral Participation Level
	Peripheral Multi-lateral Interaction Level

3.4 Procedure

3.4.1 Sampling and Data collection

As discussed earlier, communities are informal groups that do not have clear and fixed boundaries, and communities of practice cannot be created in a lab setting but evolve over time. In addition, the ICT applications used in online communities, weblogs and discussion forums, may have differences rooted in their granular features. These constraints of community and technology characteristics compelled me to define and distinguish the characteristics of communities. To balance the confounding variables that may emerge because of disparate characteristics of chosen communities, I selected three attributes – the size of network, location, and overall theme or mission of the community – to select participating communities for the study.

Based on the above-mentioned characteristics, four independent online communities (two Discussion groups and two Weblog networks) were selected for this study. These online communities were:

- Brattleboro Community Forum: A discussion forum community for Brattleboro, VT at (www.ibrattleboro.com).

- St. Paul chapter - Minneapolis E-democracy: A discussion forum community for St. Paul (MN) at (www.e-democracy.org/stpaul).
- Greensboro Community Weblogs: A weblog network for Greensboro community at (www.greensboro101.com).
- Richmond Weblogs: A weblog network for Richmond (VA) community at (www.rvablogs.com).

All of these selected communities had common themes or shared practices in politics, local events & news, and civic issues for their respective geographical locations. To collect measurements on the social relations among network members, this study used the 'Full Network Method' (Hanneman & Riddle, 2005; Wasserman & Faust, 1994) of sampling social ties. In full network method, I collected information about each node's ties with all other nodes. In essence, this approach takes a census of ties in a population of nodes -- rather than a sample. The data collection duration for this study was one month (May 1 to May 30, 2006). Since discussion forums do not have hypertextual linking features of weblog networks, such as trackbacks and permalinks, this study used a qualitative content analysis approach (Gunawardena, 1997) along with archival research.

Gunawardena's (1997) approach describes a node as engaged in a knowledge sharing process or having network characteristics if an author of entry refers to other entry by means of textual reference or by linking or replying to it. Following these data collection approaches – quantitative and qualitative data collection – I counted the number of messages posted and received by each node (individual member) of selected online communities along with the unique references that node received in the entries of

other nodes. Table 3 and Table 4 show the samples of data collected from a given discussion forum community for this study:

Table 3: A Sample of Data Collection for DF_Greensboro Online Community

Blog Name	Trackbacks	Comments	In-degree	Posts Written	Replies to Others	Out-Degree
Automatic Writing	0	23	23	2	14	16
Back to the Garden	0	6	6	1	0	1
Black Box Voting	0	0	0	0	1	1
BloggingPoet.com	0	0	0	0	2	2
Chewie World Order	0	1	1	1	8	9
Chosen Fast	0	6	6	2	7	9
connecting*the*dots	0	1	1	3	2	5
David Boyd	0	1	1	1	2	3
Ed Cone	0	26	26	6	8	14

Table 4: A Sample of Network Measures for DF_Greensboro Online Community

Node	Destination Node	Out Degree	In Degree
Automatic Writing	Ed Cone	3	4
Automatic Writing	Greensboro101		
Automatic Writing	Chosen Fast		
Black Box Voting	Keeping Reality in Sight	1	0
BloggingPoet.com	Ed Cone	2	0
BloggingPoet.com	Chosen Fast		
Chewie World Order	Ed Cone	6	1
Chewie World Order	Back to the Garden		
Chewie World Order	Editor's Log		
Chewie World Order	Greensboro101		
Chewie World Order	Greensboro Peer Pressure		
Chewie World Order	Joe Guarino		
Chosen Fast	Greensboro Peer Pressure	2	3
Chosen Fast	Joe Guarino		

3.4.2 Visualization

To create the visualizations for data analysis, the current study used the Pajek visualization tool (Batagelj & Vrvar, 2004). This tool uses the graph theory and social network analysis approaches to produce visual representations of various network characteristics including degree centrality, in-degrees, and out-degrees. The visualization

immensely helps in the identification of subgroups or sub-clusters within clusters resulting in better analysis of network properties.

3.5 Statistical Analysis

For the statistical analysis of collected data, this study used a combination of social network analysis and one-way analysis of variance (ANOVA) for each online community. This study included three dependent variables strength of ties, peripheral participation levels, and levels of peripheral multi-lateral interaction. All of these variables depend upon social network measures, which were calculated using social network analysis techniques.

The following table shows various social network measures for each dependent variable:

Table 5: Summary of Social Network Measures for Dependent Variables

Dependent variable	Corresponding Social Network Measures
Strength of Ties	Network Degree Centrality
Peripheral Participation Levels	Nodal In-Degree
	Nodal Out-Degree for peripheral nodes
Levels of Peripheral Multi-lateral Interaction	Nodal In-Degree
	Nodal out-degree for peripheral sub-network

3.5.1 Network Degree centrality:

In social network analysis, various measures of node centrality determine the relative importance of a node within the network. For example, how important a person is within a social network, or how important a city is in a transport business network.

This study has used degree centrality measure using Freeman's graph centralization measures (Freeman, 1979).

Freeman's graph centralization measures define degree centrality of a node as the number of edges attached to the node. To know the standardized degree centrality score to compare the degree centralities of different network, every nodal degree centrality score is divided by (N-1), where N is total number of nodes in the network.

3.5.2 Nodal In-Degree and Out-Degree:

Nodal degree is the number of lines that are incident with a particular node. For a node n_i, nodal degree is denoted by $d(n_i)$. Nodal degree has two different types based on the nature of ties:

- Nodal In-degree: Number of ties that originated from the outside to a node, such as links to a particular weblog on other weblogs.
- Nodal Out-degree: Number of ties that originated from source node to outside nodes, such as links on a given weblog to other weblogs.

In other words, in-degree of a node is the number of ties it receives, while out-degree is the number of ties a node sends out. To standardize the calculation of social measure, this study has considered the distinct ties only, which mean that I treated multiple ties between two nodes as one tie only. This is because multiple ties are a measure of social trust and strength of tie between the participating nodes; however, since the subject communities for this study are information networks, I was only interested in finding out the relative importance of nodes in obtaining or disseminating information in these networks.

3.5.3 Quantitative Analysis

The resultant data obtained from the afore-mentioned social network measures was further tested for statistical significance using one-way ANOVA. The post-hoc evaluation of Tukey HSD was also used to analyze the data. This allowed for a more detailed display of where the significance specifically occurred.

3.6 Research Design: Advantages

Wenger (1998) observed during observational case studies that communities of practice could not be developed overnight, rather they evolve over time. This observation projects that any external control would seriously weaken the context in which these communities exist because online communities in the weblog networks are self-organizing in nature. These communities emerge when individuals, who share a common interest or goal, connect together to share their experiences or knowledge. The fact that these weblogs and discussion forum communities are dispersed in space and time discourages the feasibility of a controlled experimental design for the study.

These observations and facts steered the current study to follow existing communities of practice within the blogosphere and discussion forums communities instead of creating prototype community spaces. Since the study is exploratory in nature, these internal validity limitations of research design do not indicate a significant flaw for the purpose of the study.

The advantages of using an analytical field study approach are that firstly, it does not lose context in which these communities emerge and exist, and secondly, information retrieved in this manner is in form of relational data and can be stored in matrix to analyze interactions and knowledge sharing patterns. For these precise reasons, various

scholars (Herring et al., 2005; Preece, 2003a; Preece, 2003b; Wellman, 1996) involved in CoP and research on online communities have recommended and used exclusively social network analysis (SNA) to analyze the structural properties of online communities.

Moreover, the Social network analysis approach is a very useful technique in an analytical study, as it does not need a strict controlled environment to investigate all the network measures. This fact is supported by social network research in computer-mediated communication for the last decade, which has shown that it is possible to explore characteristics of online communities and interaction through various field research techniques, such as interviews, archival research or ethnography (Scott, 1991; Wasserman & Faust, 1997).

CHAPTER 4. RESULTS

4.1 Participant Demographics

This study investigated four online communities, of which two were discussion forum communities and two were weblog networks communities. All of these online communities had shared practices of politics, local events & news, and civic issues for their respective geographical locations. The following table summarizes the total number of members and active members in these communities:

Table 6: Summary of the participant Online Communities

Online Community	Group Identifier	Total Number of Members	Total Number of Active Members*
eDemocracy Discussion Forum-St. Paul Chapter	DF_StPaul	410	65
IBrattleboro Discussion Forum	DF_Brattleboro	92	65
Greensboro Aggregator Blog	Blog_Greensboro	107	34
RVA Aggregator Blog	Blog_RVA	90	59

For the scope of this study, I have defined active members as the members of these online communities who either posted or received an entry during the data collection period of one month (May 1st, 2006 to May 30th, 2006). With limited data set of the current study, it was not possible to measure the frequency with which a particular node is linked to within the World Wide Web as a whole—such a measure would require that we have the entire internet for analysis. However, I measured the in-degree (number of inbound links) within data sample as a rough substitute for global popularity of node in the network. I also measured out-degree (number of outbound ties) to determine whether

a particular node was an active or a passive participant in the larger community, and analyzed the relationship between the two measures.

4.2 Measures of Strength of Ties

This section discusses the tests for the first hypothesis on strength of ties in online communities. The hypothesis that related to the dependent measure of strength of ties was:

H-1 Compared to centralized discussion forums communities, blog communities are more conducive to the weak-ties phenomenon.

The strength of ties directly correlates with the cohesiveness of a community's social structure. A community structure with many strong ties would be dense and bounded structure; on the other hand, a community structure with lots of weak-ties would be comparatively open and porous.

The strength of ties was calculated using the social network measure of network centrality. Since the participating communities were information networks, i.e., the social ties in these networks were for information and knowledge exchange, and not due to formal organizational and personal relations, I have used the degree centrality measure for network centrality calculations. The degree centrality is an important social network measure for information networks, because it helps to identify nodes that are important for knowledge exchange. The degree centrality measure is based on the premise that a node with high nodal degree would be connected to many unique nodes in the network, therefore, is able to exchange information or knowledge with all those nodes directly.

For the calculation of network degree centrality, this study used the Freeman's graph centralization formula (Freeman, 1979). Freeman's graph centralization measure

expresses the degree of inequality or variance in a given network as a percentage of that of a perfect star network of a same size. For a network with N unique nodes, the following formula calculates the Freeman's degree centrality:

$$\text{Freeman's Degree Centrality} = \frac{\sum (D\max - di)}{(N-1)(N-2)} \times 100$$

Where :

Dmax = Maximum degree value in the network

di = Nodal degree

N = Total number of nodes in the network

Figure 7: Formula to compute Freeman's Degree Centrality

For every node in the network, nodal degrees were calculated using the unique ties with other nodes in the network. Every entry posted or reply to other nodes were considered as one out-degree, while every reply or reference in other entries were considered as one in-degree for that node. Since the study focuses on comparison across networks of different sizes and densities, I 'standardized' all the in-degree and out-degree using the following formula:

For a Network of N nodes

$$\text{Normalized Degree} = \frac{d_i}{(N-1)} \times 100$$

Where :

di = degree of a given node in the network

Figure 8: Normalized nodal degree for a given node

The following table summarizes the mean scores for the degree centrality index of the four online communities:

Table 7: Summary of Degree Centrality scores

Groups	Count	Sum	Average	Variance	Degree Centrality
DF_Brattleboro	65	835.94	12.86	186.98	0.77
DF_StPaul	65	1704.69	26.23	429.80	0.60
Blog_Greensboro	34	496.97	14.62	108.23	0.30
Blog_RVA	59	1168.97	19.81	287.46	0.65

The social network measure responses were summarized and a one-way analysis of variance (ANOVA) was performed to test degree centrality differences among the four online communities. The ANOVA results for network measure of degree centrality are summarized in the following table:

Table 8: ANOVA results for Degree Centrality

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	6532.03	3	2177.34	7.98	0.000045	2.65
Within Groups	59718.33	219	272.69			
Total	66250.36	222				

The results of the ANOVA indicate that a significant effect of community structure exists. The $F(3,219)=7.98$ and the $p\text{-value}<0.0001$. This indicates that there exists a significant difference in the mean scores of degree centrality measures. Since degree centrality scores are indicators of strength of ties in target networks, the ANOVA results indicate that there is statistically significant difference among the strength of ties in subject communities. The upper extreme in the strength of ties values indicate a dense, tighter community structure with high strong-ties, while lower extreme indicates an open community structure with high weak-ties.

The pair wise Tukey HSD (honestly significant difference) test was used as the post-hoc test for the ANOVA. With this test, each unique experimental condition can be compared to another to determine where significant difference occurs or does not occur. The results of Tukey HSD (95% CI) post-hoc test are summarized in the following table:

Table 9: Summary of post-hoc Tukey HSD (95% CI) test for hypothesis-1

Contrast	Difference	Tukey 95% CI		
DF_Brattleboro v Blog_RVA	-6.952	-14.640	to 0.735	
DF_Brattleboro v DF_StPaul	-13.365	-20.864	to -5.866	**(significant)**
DF_Brattleboro v Blog_Greensboro	-1.756	-10.804	to 7.292	
Blog_RVA v DF_StPaul	-6.413	-14.100	to 1.274	
Blog_RVA v Blog_Greensboro	5.196	-4.009	to 14.401	
DF_StPaul v Blog_Greensboro	11.609	2.561	to 20.657	**(significant)**

From the post-hoc test of Tukey HSD (95% CI), it was found that there was significant difference between degree centrality scores of DF_StPaul and Blog_Greensboro. Table 9 shows that difference between the mean scores of DF_StPaul and Blog_Greensboro was 11.61. As a consequence, DF_StPaul online community (0.60) emerges as comparatively more centralized than Blog_Greensboro online community (0.30).

In addition, there was a statistically difference found between degree centrality scores of DF_Brattleboro and DF_StPaul. Comparing the mean scores (Network Degree Centrality) showed that DF_Brattleboro online community (0.77) was more centralized network compared to DF_StPaul online community (0.60). To demonstrate the comparative network centrality, Figure 8 below summarizes the mean scores of three online communities among which there was a statistical significant difference:

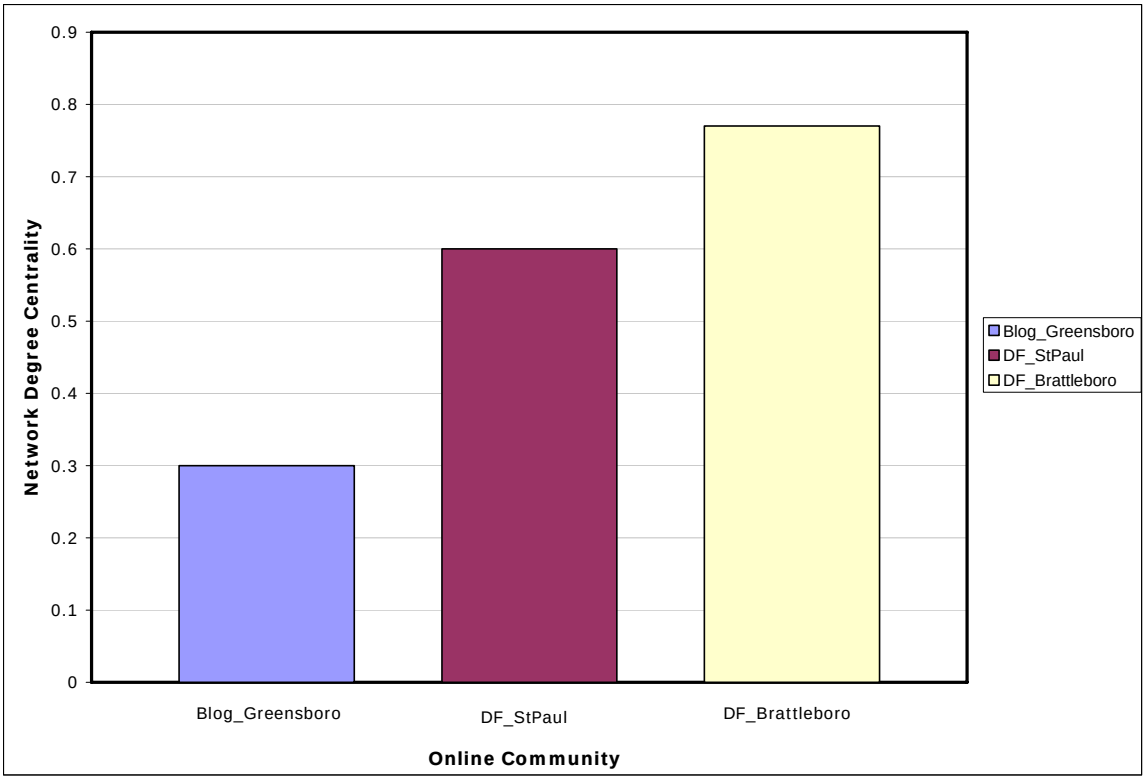

Figure 9: Summary of Network Degree Centrality Scores

4.3 Measures of Peripheral Participation Levels

This section tests for the second hypothesis on peripheral participation levels in online communities. The hypothesis that related to the social network analysis measure of peripheral participation levels was:

H-2 Self-organizing communities provide a more conducive environment for peripheral participation compared to centralized, structured communities.

The peripheral participation levels were calculated using the social network measure of nodal in-degree and nodal out-degree. Nodal degree is an important social network measure to identify nodes that are important for knowledge exchange. Since the larger the number of sources accessible to a node, the easier it is to obtain information; in this sense, social ties constitute a social capital that may be utilized to gain different

social resources. However, in directed networks, such as in the case of participating online communities, it is important to distinguish nodal degree based on in-degree and out-degree types of nodal degree.

A node with high out-degree would be comparatively well connected with other, and therefore, would have a comparatively high activity level of exchanging information with many others. A node with high in-degree, on the other hand, is considered prestigious because nominations or incoming ties on a positive social relation (e.g., acknowledgement of one's knowledge) express popularity of that node.

As per the LPP theoretical framework for this study, peripheral members of a network have the lowest authority, theoretically close to zero. However, since it is very difficult to identify such participants who have zero level of authority or in other words, zero in-degree, I decided to choose the nodes with lowest 10% authority levels as peripheral participants for this study. Firstly, in-degrees for all the nodes of each online community were computed, and then these in-degrees were converted into normalized in-degree values using the following formula:

For a Network of N nodes

$$\text{Normalized In - Degree} = \frac{d_i}{(N-1)} \times 100$$

Where :

di = in - degree of a given node in the network

Figure 10: Formula to compute normalized in-degree

Once all the normalized in-degrees were calculated, I identified nodes with the lowest 10% normalized in-degrees as the peripheral participants of that network. To find out the activity levels of these selected nodes, I calculated their out-degrees and to

compare across the networks, computed their normalized out-degrees using the following formula:

For a Network of N nodes

$$\text{Normalized Out - Degree} = \frac{d_i}{(N-1)} \times 100$$

Where :

di = out - degree of a given node in the network

Figure 11: Formula to compute the normalized out-degree

The following table summarizes the mean scores for the normalized out-degree index of peripheral participants in subject online communities:

Table 10: Summary of Normalized out-degree scores for peripheral participants

Groups	Count	Sum	Average	Variance
DF_Brattleboro	50	228.13	4.56	18.92
DF_StPaul	33	296.89	9.00	50.66
Blog_Greensboro	24	160.61	6.69	29.53
Blog_RVA	25	87.93	3.52	5.57

The responses for network measure 'normalized out-degree' of every peripheral participant node were summarized and a one-way analysis of variance (ANOVA) was performed to test activity level differences among the four online communities. The ANOVA results for network measure of normalized out-degrees are summarized in the following table:

Table 11: ANOVA results for normalized out-degree scores of peripheral participants

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	562.58	3	187.53	7.14	0.00018	2.68
Within Groups	3360.99	128	26.26			
Total	3923.57	131				

The results of the ANOVA indicate that a significant difference in normalized out-degree scores of peripheral participants exists. The $F(3,128)=7.14$ and the $p\text{-value}<0.0002$. This indicates that there exists a significant difference in the mean scores of normalized out-degree measures for the peripheral participants. Since normalized out-degree values are indicators of activity level for given nodes, these ANOVA results indicate that there exists a statistically significant difference in the activity levels of the peripheral participants in four subject communities.

The pair wise Tukey HSD (honestly significant difference) test was used as the post-hoc test for the ANOVA. With this test, each unique experimental condition can be compared to another to determine where significant difference occurs or does not occur. The results of Tukey HSD (95% CI) post-hoc test are summarized in the following table:

Table 12: Summary of post-hoc Tukey HSD (95% CI) test for hypothesis-2

Contrast	Difference	Tukey 95% CI		
DF_Brattleboro v DF_StPaul	-4.434	-7.425	to -1.442	**(significant)**
DF_Brattleboro v Blog_Greensboro	-2.129	-5.442	to 1.183	
DF_Brattleboro v Blog_RVA	1.045	-2.222	to 4.313	
DF_StPaul v Blog_Greensboro	2.304	-1.274	to 5.883	
Blog_RVA v DF_StPaul	5.479	1.942	to 9.016	**(significant)**
Blog_RVA v Blog_Greensboro	3.175	-0.637	to 6.987	

From the post-hoc test of Tukey HSD (95% CI), it was found that there was significant difference between normalized out-degree scores of groups DF_StPaul and Blog_RVA. In addition, there was statistically difference found between normalized out-degree scores of groups of DF_Brattleboro and DF_StPaul. Figure 12 summarizes the mean scores for normalized out-degrees in online communities among which there was a significant difference:

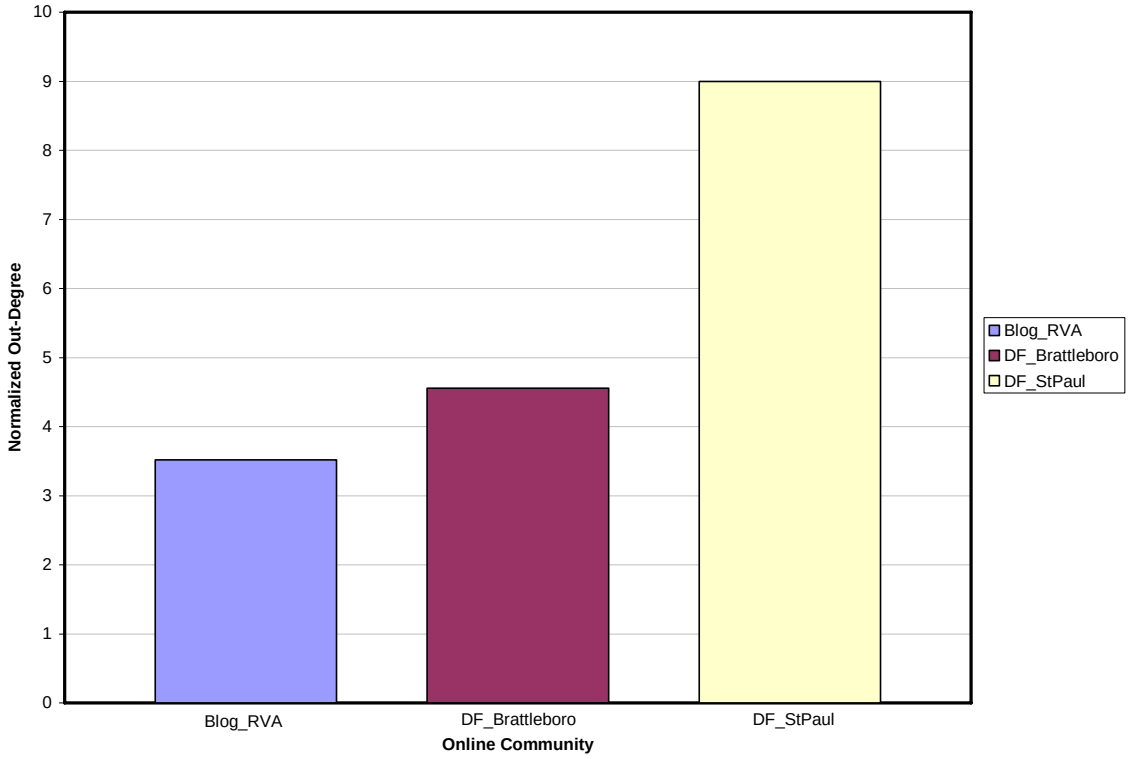

Figure 12: Summary of Normalized Out-Degree mean scores

4.4 Measures of Peripheral Multi-lateral Interaction Levels

This section tests for the third hypothesis on levels of multi-lateral interaction among peripheral participants in subject online communities. The hypothesis that related to the dependent measure of peripheral multi-lateral interaction levels was:

H-3 Self-organizing communities have more multilateral interaction flow among their peripheral members than the peripheral members of centralized communities.

The first step towards finding out the interaction levels of peripheral participants in subject communities was to find out the peripheral participants in these communities. As described in Section 4.3, I calculated the normalized in-degree values for all the nodes in target communities and, as per the LPP theoretical framework, chose the nodes with lowest 10% authority levels as peripheral participants in those communities. To investigate the multi-lateral peripheral interaction, I was interested in the interaction among only the chosen peripheral participants; therefore, I formed sub-networks of all the selected peripheral participants and re-calculated the normalized out-degrees for nodes in these sub-networks.

The following table summarizes the mean scores for the normalized out-degree index of nodes in the configured sub-networks:

Table 13: Summary of normalized out-degree scores in the sub-networks

Groups	Count	Sum	Average	Variance
DF_Brattleboro	50	49	0.98	1.78
DF_StPaul	33	40	1.21	2.92
Blog_Greensboro	24	15	0.63	0.77
Blog_RVA	25	17	0.68	0.73

The responses for measure ‘normalized out-degree’ for nodes in the sub-networks (peripheral participant nodes of original networks) were summarized and a one-way analysis of variance (ANOVA) was performed to test interaction level differences among the four online communities. The ANOVA results for network measure of normalized out-degrees are summarized in the following table:

Table 14: ANOVA results for normalized out-degree scores of nodes in sub-networks

Source of Variation	SS	df	MS	F	P-value	F crit.
Between Groups	6.52	3	2.17	1.29	0.28	2.68
Within Groups	215.56	128	1.68			
Total	222.08	131				

The results of the ANOVA indicate that there is a non-significant effect of community structure on the normalized out-degree scores of nodes in the configured sub-networks. The *F(3,1289)=1.29* and the *p-value=0.28.* Since normalized out-degree scores are indicators of activity levels of nodes in the reconfigured networks, the ANOVA results indicate that there is no statistically significant difference present among the interaction levels or activity levels of these participating nodes.

CHAPTER 5. DISCUSSION AND CONCLUSIONS

5.1 Validity of Hypotheses

The study was performed to test three hypotheses about knowledge sharing process in online communities of different community structures. Data gathered and analyzed in this study do not answer specific research questions; rather they are used cumulatively to examine the hypotheses and address the research questions. The study was performed using an analytical field study method of data mining (archival research) to explore the relationship between online communities of weblog networks & discussion forums and the knowledge sharing process that takes place in them.

For quick reference, the examined hypotheses are restated as follows:

H-1: Compared to centralized discussion forums communities, weblog communities are more conducive to weak-ties phenomenon.

H-2: Self-organizing communities provide a more conducive environment for peripheral participation compared to centralized, structured communities.

H-3: Self-organizing communities have more multilateral interaction flow among their peripheral members than the peripheral members of centralized communities.

In addition, it will be important to reference the research questions that formed the hypotheses. The following are research questions to help provide additional insight into the discussion of hypotheses:

1. What roles do the emerging ICT applications (namely weblogs and online discussion forums) play in the knowledge sharing process? What are the underlying differences in the knowledge sharing process facilitated by weblogs and online discussion forums respectively?

2. How does the structure of a community of practice affect the knowledge sharing process that takes place in that environment? Do self-organizing online communities facilitate knowledge sharing better than centralized and structured communities or vice-versa?
3. How does the flow of information differ in weblog networks and discussion forum communities, and what are the implications of these differences?
4. Compare the effectiveness of centralized and structured community structure versus decentralized, self-organizing community structure for the knowledge sharing process.

5.1.2 Strength of ties in a community structure

The study results fully support the first hypothesis, which stated that self-organizing communities, e.g., weblog communities, are more supportive of weak-ties within their structure. The study results showed that there was a statistically significant difference between degree centrality scores of a weblog community, Blog_Greensboro, and a discussion forum community, DF_StPaul. It was found that weblog community, Blog_Greensboro had network degree centrality value of only 0.30 compared to network degree centrality value of 0.60 for discussion forum community, DF_StPaul. When these two results combined, it showed that weblog network community, Blog_Greensboro, had relatively high number of weak-ties, and the difference is significant when compared to discussion forum community of DF_StPaul.

However, the study results also showed that there was significant difference between degree centrality scores of two discussion forum communities, DF_Brattleboro and DF_StPaul. Since the absolute network degree centrality scores of DF_Brattleboro

and DF_StPaul are 0.77 and 0.60 respectively. It can be safely interpreted that DF_Brattleboro has comparatively stronger ties in its community structure than DF_StPaul does. This interpretation indicates that order of the extent of weak-ties in a community structure is as follows: DF_Brattleboro < DF_StPaul < Blog_Greensboro.

These observations suggest that self-organizing communities of practice, such as the weblog networks, do exhibit the weak-ties phenomenon. When combined with Granovetter's weak-ties theory, this argument furthers the debate that self-organizing communities possess the characteristics of an excellent knowledge-sharing environment. The findings of this study led us to the conclusion that while dense and strong-ties communities may be better for collective action, self-organizing communities of practice would facilitate the exploration and dissemination of new information. In addition, comparatively better knowledge dissemination abilities of self-organizing communities make them more suitable environment to offer solutions to knowledge-transfer problems (Hansen, 1999) compared to the centralized communities.

5.1.3 Peripheral participation levels

The second hypothesis, which states that self-organizing communities, such as blog networks, facilitate peripheral participation better than the centralized and bounded communities do, was not completely supported by the results of this study. The study results showed that there was statistically significant difference between activity level scores (Normalized out-degree) of a weblog community, Blog_RVA, and a discussion forum community, DF_StPaul. However, weblog community had relatively lower mean scores for normalized out-degrees of their nodes compared to discussion forum

community implying that discussion forum community had higher levels of peripheral participation.

The study results, therefore, did not support the second hypothesis and rather exhibited opposite of second hypothesis that centralized discussion forum communities had more peripheral participation among their members. When these observations are analyzed along with the results of first hypothesis, they seem to contradict the weak-ties theory that a community with lot of weak-ties should be more conducive to newcomers in that community as there are lower barriers of social trust. The explanation of these seemingly contradictory observations may be that the peripheral participants in this study had relatively higher perceived authority levels than suggested in the LPP theoretical framework.

To find out the peripheral participants as per the theoretical framework of LPP (Lave & Wenger, 1991), it was necessary to identify the participants with zero connections (zero out-degrees and in-degrees). Such participants are referred as lurkers in a community and are very difficult to spot without extensive administrative control over community participation. Due to these practical reasons of spotting peripheral participants with zero connections, I had selected peripheral participants with higher perceived authority levels (0-10% normalized in-degree scores). This limitation of research design may have been confounding and, consequently, led to observation contradictory to the implications of weak-ties theory (Granovetter, 1973) that sparse, loosely bound networks allow relatively more newcomers to participate as the social trust levels required to participate in shared practices of a community are low compared to the centralized communities.

5.1.4 Levels of Peripheral multi-lateral interaction

The study results support rejecting the third hypothesis, which stated that self-organizing communities, such as weblog networks, would have more multi-lateral interaction among their peripheral participants than the centralized and bounded communities would do. No indication was found that self-organizing communities, e.g., weblog networks, have higher activities among their peripheral participants even though these communities exhibit relatively higher extent of weak-ties among their members.

Along with the second hypothesis, the third hypothesis was dependent upon the variable of peripheral participation. The LPP theory (Lave & Wenger, 1991) defines peripheral participants as participants with zero authority, that is, newcomers in a community who has zero perceived authority among the existing members of that community. However, as described in previous section, due to practical limitations of identifying true peripheral participants, participants with higher activities yet relatively lower scores (0-10% normalized in-degree scores) than rest of the participants in that community were chosen as peripheral participants to examine this hypothesis.

The 'weak-ties' phenomenon (Granovetter, 1973 & 1983) suggests that sparse, loosely bounded networks with a large range are good for peripheral participation as they require lower social trust and less frequent interaction to participate in shared practices of a community. LPP framework (Lave & Wenger, 1991) states that peripheral participants of a community of practice could be core (or expert) members of another community of practice that makes the boundaries of these networks fluid and porous boundaries. Since these characteristics facilitate overlap of online communities with different shared practices through common members, it seemed logical to hypothesize that self-organizing

communities – relatively sparse and loosely bound networks as per observations of first hypothesis – would have more multilateral interaction through their relatively greater number of common peripheral participants with other overlapping communities of practice. The shortcomings of research design in finding out the participants with zero perceived authority levels and including all the participants in case of self-organizing communities may provide an explanation why the findings of this study are not coherent with the propositions of weak-ties theory about multilateral peripheral participation as stated above.

5.2 Limitations of the Study

There were several limitations of this study, which may have been confounding factors and consequently may have skewed the results. First limitation of this study was its inability to identify participants who are truly newcomer to a given community. As per the LPP framework, a true newcomer (peripheral participant) of the community would only absorb the knowledge disseminated by core members of the community but would not contribute to the community in terms of information and knowledge about the shared practices of that community. When this theory is used for online communities, true peripheral participants are the ones who read the discussions on a discussion forum or weblog, but rarely or never actively participates. Such passive members of a community make up a sizeable proportion of many online communities; however, without explicit administrative controls it is not possible to find out about their activities. Various measures defined by system administrators, such as unique number of readers for a post, may enrich the future studies investigating peripheral participation.

This study employed a data-mining (archival research) approach for the data collection, which provides the data for social network analysis measures like nodal degree and degree centrality in terms of number of posts or entries made or received by a particular member of community. In addition, a content analysis scheme was used to find out if there was any mention of a particular member in posts or entries by other members of communities without explicit linking. However, both of the above data collection measures fail to capture instances of social ties and linking among members in the absence of explicit linking. In addition, most of the community members develop trust and knowledge of each other after a period of interaction that allows them to use reference names for other members, which may be different from the screen names of those members.

Finally, a critical limitation of study was its limited success in capturing true self-organization nature of weblog networks. For the purpose of this study, I had selected two aggregator weblog sites that captures the feeds from their member weblogs and posts them on a central place. However, all of these member weblogs may be connected to many other weblogs, which are not part of chosen weblog aggregator site, therefore, a true representation of their self-organizing weblog network would be to include all of these weblogs connected to each other. As found in the study of Herring et al. (2004), the number of connected weblogs grows exponentially and a true self-organized weblog network may comprise of hundreds of thousands weblog, which would make comparison of bounded communities like discussion forums against such open and vast communities very difficult and full of confounding variables.

5.3 Design Implications and Recommendations

This study has attempted to investigate knowledge sharing in online communities within a broad and important set of concerns that are applicable to a range of contexts, and which are not dependent upon a particular technology or application. The findings of this study could have important strategic implications for designers of online communities in the extent to which they use tools and approaches that are more centralized or decentralized. In addition, the findings of this study could help organizations, individuals, and other entities select online community tools and applications for their unique needs.

This study has tackled the question of knowledge sharing in online communities from a fresh perspective using web links and an adaptation of social network analysis. The study has adopted its central premise from the LPP framework and weak-ties theory that knowledge sharing capabilities will function differently in different types of online communities of practice.

The primary finding of this study that has important design implications for the online communities is that self-organizing communities, such as blog networks, exhibit weak-ties phenomenon to higher extent than centralized communities. As per weak-ties theory of Granovetter (1973, 1983), prevalence of weak-ties in a community suggests that community structures of these self-organized communities would facilitate comparatively greater dissemination of knowledge and flow of information among their members than centralized communities. The findings of this study, therefore, demonstrate that self-organizing online communities, such as blog networks, promote a sparse and loosely

bound community structure with comparatively greater extent of weak-ties among their members than centralized online communities.

Granovetter (1973) put forth the idea that while strong ties are good for taking collective action, it is weak-ties that are crucial for finding and disseminating new information. He explained this idea using concepts of ‘friends’ and ‘acquaintances’, friends have more frequent interaction among themselves and enjoy strong-ties while acquaintances have comparatively less-frequent interaction, therefore, ties among them are comparatively weak. Granovetter (1983) argued that since friends usually belong to same social circles, therefore, the information they receive from each other would overlap considerably with what they already know. Acquaintances, by contrast, have relatively lower chances of belonging to same social circle; therefore, the information they receive from each other would turn out to be relatively novel information. The abundance of weak-ties in self-organizing community structure, therefore, makes these communities better environments for the discovery of new information compared to centralized community environments, which would make them suitable environments for dissemination of new information, therefore, these self-organizing communities may offer solutions to what Hansen (1999) calls as knowledge transfer problems, that is, situations in which individual(s) or groups are not aware of existence of information that may be useful to them.

The findings of this study, however, did not provide any evidence of community structure’s impact on peripheral participation and the interaction activity level among peripheral participants of a given online community. These findings mean that there is inconclusive evidence if a certain community structure would encourage newcomers to

participate more in the shared practices of a community. In addition, the findings did not demonstrate if there is any strong correlation between structure of a community and interaction among its peripheral participants. It is, therefore, inconclusive based on the findings to state what type of community structure would be more successful in inciting interest and higher participation among newcomers of a given community.

Apart from the direct design implications for future designs of online communities, this study also paves the way for mapping of various organizational and social networks to measure knowledge sharing in these communities. The combined theoretical framework of LPP and weak-ties theory provides a backbone tool to compare different networks for community structure, flow of information, and overall knowledge sharing. Since LPP framework uses the concept of 'perceived' authority instead of formal or explicit authority, the theoretical framework used in this study could be used in shadow or 'implicit' that lie behind the organizational hierarchy.

During the data collection and data analysis phases of this study, I learned about various issues that are relevant to future studies on online communities. These observations are presented below as design recommendations.

For the comparative studies of online communities containing two or more, it is crucial to control for sample sizes. The theory of LPP, however, suggests that communities of practice evolve with little or no control; therefore, it is difficult to have lab control or restrict a community to a particular size. Since varying sample sizes of subject communities may have confounding impacts on research designs, different sample sizes could make the communities vulnerable to logical inconsistencies and anomalies. To control this difference without affecting the attributes of communities of

practice, I recommend the technique of 'Normalization' as it transforms almost all of social network measures in a standard measure not dependent on network sample size. The formula used for 'Normalization' in this study is stated in Figure 8 and is reiterated below for quick reference:

For a Network of N nodes

$$\text{Normalized Degree} = \frac{d_i}{(N-1)} \times 100$$

Where :

di = degree of a given node in the network

Figure 13: Formula to Compute Normalized Nodal Degree of a Given Node

Here N denotes the sample size of network. To obtain standardized social network measures for a particular node, network measure is divided by overall network size minus ego (i.e., the subject node) and multiplied by 100 to obtain percentage form.

The following design recommendations are regarding approaches to capture peripheral participation in online communities. The theory of LPP (Lave & Wenger, 1991) states that peripheral participants or newcomers in a community are the ones who have zero perceived authority among existing members of a community of practice. In online communities of practice, peripheral participants are commonly referred as 'lurkers'. A lurker is a member who passively participates in the shared practices of a community by reading discussions, but rarely participates actively by writing or replying. By definition, lurkers do not connect explicitly with other members of a community, therefore, it becomes very difficult to observe their activities and obtain network measures.

Various administrative controls over online community may help in obtaining information about lurkers' activities. Administrative controls may provide statistics about the participation levels of lurkers with help of session logs. To measure the network measures, such as in-degree and out-degree, number of unique hits for a particular post may prove crucial, however, since a particular user may use different machines or IP addresses, it is immensely difficult to obtain this data without a margin of error.

Another big challenge that researchers may face while studying online communities is to observe and investigate self-organizing communities. These self-organizing communities comprise scattered individual(s) and groups who do not have explicit connections among themselves, however, these communities form when triggered by an external event these individuals and groups start connecting to each other. By definition, there is no single place to capture all the members of a self-organizing community making it difficult to investigate a community using a particular website, such as an aggregator blog website. In addition, the changes in the external triggers, connections among members, and therefore shape and size of overall network, may change as well.

I recommend, therefore, studying self-organizing communities using a single external trigger only. For example, a study would be less vulnerable to inconsistencies if it studied all the blogs and their inter-connections that are discussing a particular event, such as Katrina disaster or Presidential election, rather studying all the blogs from a particular geographical area. In former case, blogs are influenced by a single external trigger (i.e., Katrina tragedy or Presidential election) to form a self-organizing community. The later case, however, may have multiple external triggers (i.e., both

Katrina tragedy and Presidential election) to form multiple self-organizing communities. While studies of these stand-alone self-organizing communities could be useful, their comparison to each other may have confounding inconsistencies because of multitude of shared practices.

5.4 Future Research

This study has investigated the impact of an online community's community structure on the knowledge sharing process in that community. To examine the knowledge sharing process in a community as a function of a community member's social ties and network measures, I used a theoretical framework comprised of legitimate peripheral participation theory and the weak-ties phenomenon. Further, I performed a comparative analysis of two different types of community structures – self-organizing online communities (e.g., weblog networks) and centralized online communities (e.g., discussion forums communities)–with respect to different social network measures, such as nodal degrees and degree centrality. These network measures helped us to find out and compare community attributes, such as strength of ties and peripheral participation, which have implications on overall knowledge sharing in that community.

There are certain areas of the current study which could be further explored in future studies to enhance their scope in order to examine various attributes of knowledge sharing that were not covered in the current study. One such area is the method to delineate the community structure; for the purposes of current study, the method was used to categorize different community structures in centralized and self-organizing communities. However, there could be potentially better ways to delve into and explore the relationship of community structure with knowledge sharing. Hoof et al. (2003) have

highlighted certain other independent variables of community structure that could probably be of interest for future studies.

The other potential area for future exploration is to find out the impacts of the granular feature of ICT applications, such as weblogs and discussion forums. There is a need to investigate if the online communities using ICT applications, such as weblogs and discussion forums, also encompass other characteristics of communities of practice, and what could be set of rules to determine which category a given community fits into? How would one define their boundaries and overall size? Are theses community sizes potential factors to consider?

In this study, I have explored knowledge sharing in a given community using social network measures, such as explicit links and ties only. There is definitely a potential scope to explore knowledge sharing among individual participants of community along with looking at knowledge sharing as a community phenomenon. There are interesting attributes and variables related to knowledge creation and accumulation that can be further explored among the individual participants themselves, through surveys, interviews and content analysis of postings.

REFERENCES

Argote, L., & P. Ingram (2000). "Knowledge transfer A Basis for Competitive Advantage in Firms." Organizational Behavior and Human Decision Processes, 82 (1): 150-169.

Bart van den, H., & Jan A. de Ridder (2004) "Knowledge sharing in context: the influence of organizational commitment, communication climate and CMC use on knowledge sharing". Journal of Knowledge Management, 8 (6): 117-130.

Bart van den Hooff, Wim Elving, Jan Michiel Meeuwsen, & Claudette Dumoulin (2003). "Knowledge sharing in knowledge communities". M. Huysman et al. (eds.), Communities and Technologies, 119-141.

Batagelj, V. & Vrvar, A. (2004). "Pajek: package for large networks: version 1.00." Ljubljana: University of Ljubljana. http://vlado.fmf.uni-lj.si/pub/networks/pajek/.

Breiger, R.L. (1974). "The duality of persons and groups." Social forces. 53, 181-190.

Brown, J. S., & Duguid, P. (2001). "Knowledge and Organization: A Social-Practice Perspective." Organization Science, 12 (2): 198-213.

Conference Board, the, user survey (Oct 2, 2006). Retrieved Dec. 12, 2006, from http://www.conference-board.org/utilities/pressDetail.cfm?press_ID=2971

Davenport, T., & Prusak, L. (1998). "Working knowledge. How organizations manage what they know." Cambridge, MA: Harvard Business School Press.

Delwiche, A. (2004). "Agenda-setting, opinion leadership, and the world of web logs". Paper presented at the International Communication Association Conference, May 27-June 1, New Orleans, LA.

Eriksson, D. M. & Wulf, V. (1999). Self-organising social systems: A challenge to computer supported cooperative work. Cybernetics and Human Knowing, 6(2), 1-4.

Eveland, J. D., & Bikson, T. (1988). “Work Group Structures and Computer Support”. ACM Transactions on Office Information Systems, 6, 354-79.

Faust, K. (1997). “Centrality in affiliation networks.” Social Networks, 19, 157-191.

Freeman, L. C. (1979). “Centrality in social networks: Conceptual clarification.” Social Networks, 1(3), 215-239.

Garton, L., Haythornthwaite, C., & Wellman, B. (1999). “Studying on-line social networks.” In S. G. Jones (Ed.), Doing Internet research (pp. 75-106). Thousand Oaks, CA: Sage Publications.

Gibson, D., Kleinberg, J., & Raghavan, P. (1998). “Inferring Web communities from link topology.” Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia.

Granovetter, Mark S. (1973). “The Strength of Weak Ties”. American Journal of Sociology, 78 (6): 1360- 1380.

Granovetter, Mark S. (1983). “The strength of weak ties: a network theory revisited”. Sociological Theory, 1: 201-233.

Gruhl, D., Guha, R., Liben-Nowell, D., & Tomkins, A. (2004). “Information diffusion through blogspace.” WWW2004, May 17-22, New York.

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). “Analysis of global online debate and the development of an interaction analysis model for examining social

construction of knowledge in computer conferencing". Journal of Educational Computing Research, 17(4), 397-431.

Hanneman, R., & Riddle, M. (2005). "Introduction to social network methods." Riverside, CA: University of California, Riverside (published in digital form at http://faculty.ucr.edu/~hanneman/).

Hendriks, C. (2002). "Institutions of deliberative democratic processes and interest groups: roles, tensions and incentives." Australian Journal of Public Administration, 61(1): 64-75.

Henning, J. (2003). "The Blogging Iceberg of 4.12 Million Hosted Weblogs, Most Little Seen, Quickly Abandoned." Survey, Perseus Development Corp. (http://www.perseus.com/blogsurvey/) Last accessed Sep. 9, 2005.

Herring, S. C., Scheidt, L. A., Bonus, S., & Wright, E. (2004). "Bridging the gap: A genre analysis of weblogs." Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS-37). Los Alamitos: IEEE Press.

Herring, S. C., Kouper, I., Paolillo, J. C., Scheidt, L. A., Tyworth, M., Welsch, P., Wright, E., and Yu, N. (2005). "Conversations in the blogosphere: An analysis "from the bottom up." Proceedings of the Thirty-Eighth Hawaii International Conference on System Sciences (HICSS-38). Los Alamitos: IEEE Press.

Hiltz, S. R., & Turoff, M. (1993). "The Network Nation". (second ed.). Cambridge, MA: MIT Press.

Hinds, P., & Pfeffer, J. (2003), "Why organizations don't 'know what they know': cognitive and motivational factors affecting the transfer of expertise," in

Ackerman, M., Pipek, V., & Wulf, V. (Eds),Beyond Knowledge Management: Sharing Expertise, MIT Press, Cambridge, MA.

Hislop, D., (2002). “Mission impossible? Communicating and sharing knowledge via information technology.” Journal of information technology, 17:165-177.

Huysman, M.H., & De Wit, D. (2002), Knowledge Sharing in Practice, Kluwer Academics, Dordrecht.

Ito, J. “Emergent Democracy.” http://extremedemocracy.com/chapters.

Jackson, M. H. (1997). “Assessing the structure of communication on the World Wide Web”. Journal of Computer-Mediated Communication, 3 (1). http://www.ascusc.org/jcmc/vol3/issue1/jackson.html Last accessed: Dec. 12, 2005.

Jarvenpaa, S.L., & Staples, D.S. (2001), “Exploring perceptions of organizational ownership of information and expertise,” Journal of Management Information Systems, 18 (1): 151-83.

Katz, J., Rice, R., & Aspden, P. (2001). “The Internet, 1995-2000: Access, civic involvement, and social interaction”. American Behavioral Scientist, 45 (3): 405-419.

Kavanaugh, A. (1999). “The impact of computer networking on community: A social network analysis approach.” Paper presented at the telecommunications policy research conference, 27-29 September, 1999. http://www.bev.net/about/research/reports/docs/TPRC.UserStudy.Kavanaugh.pdf Last accessed on 2/18/2006.

Kavanaugh, A., Cohill, A., & Patterson, S. (2000). “Use and Impact of Community Networking in Blacksburg.” Boston: Artech House (2nd edition).

Kavanaugh, A., & Patterson, S. (2001). “The Impact of Community Computer Networks on Social Capital and Community Involvement”. American Behavioral Scientist, 45 (3): 496-509.

Kavanaugh, A., Carroll, J.M., Rosson, M.B., Reese D.D., & Zin, T.T. (2005). “Participating in Civil Society: The case of networked communities”. Interacting with Computers, 17: 9-33.

Kavanaugh, A., Isenhour, P., Cooper, M., Carroll, J.M., Rosson, M.B., & Schmitz, J. (2005). “Information Technology in Support of Public Deliberation”. In P. van den Besselaar and J. Preece (Eds.). Communities and Technologies Proceedings, The Netherlands: Kluwer Academic Publishers.

Kavanaugh, A., Isenhour, P., Godara, J., Cooper, M., Midha, A., & Randolph, W. (2005). “Detecting and Facilitating Deliberation at the Local Level”. second conference on online deliberation/DIAC 2005, Stanford University, May 19-21, 2005.

Kirn, K. (2002). “Building Social Capital on the Web: The case of Minnesota E-Democracy”. In Turow, J (Ed.), Energizing Voters Online: Best Practices from Election 2000. Report No.- 39, Annenberg Public Policy Center, University of Pennsylvania.

Kleinberg, J. (1999). “Authoritative sources in a hyperlinked environment.” Journal of ACM (JASM), 46(5), 604-632.

Kraut, R., & Attewell, P. (1993). Electronic Mail and Organizational Knowledge. Working paper, Department of Social and Decision Sciences Carnegie Mellon University.

Krishnamurthy, S. (2002). “The multidimensionality of Blog conversations: the virtual enactment of September 11.” In Maastricht, The Netherlands: Internet research 3.0.

Kumar, R., Novak, P., Raghavan, S., & Tomkins, A. (2003). “On the bursty evolution of Blogspace.” Proceedings of the Twelfth International World Wide Web Conference, Budapest, Hungary.

Laumann, E., Marsden, P., & Prensky, D. (1983). R. Burt & M. Minor (Eds.), in “Applied Network Analysis”. Beverly Hills, CA: Sage.

Lave, J., & Wenger, E. (1991). “Situated learning: Legitimate peripheral participation”. Cambridge: Cambridge University Press.

Maeterlinck, M. (1927). The life of the white ant. London: George Allen and Unwin.

Marlow, C. (2004). “Audience, structure and authority in the weblog community.” Paper presented at the International communication association conference, May 27-June 1, New Orleans, LA.

Marshall, L., (1997). “facilitating knowledge management and knowledge sharing: new opportunities for information professionals”. Online, 21(5): 92-98.

McPherson, J.M. (1982). “Hypernetwork sampling: Duality and differentiation among voluntary organizations.” Social networks. 3, 225-249.

Nardi, B.A., Schiano, D.J., Gumbrecht, M., Schwartz, L. (2004a). “I’m blogging this” A Closer Look at Why People Blog.[On line] Available at: http://www.ics.uci.edu/~jpd/classes/ics234cw04/nardi.pdf

Nardi, B., Schiano, D., & Gumbrecht, M. (2004b). “Blogging as social activity, or, would you let 900 million people read your diary?” CSCW’04, Nov. 6-10, 2004, Chicago-IL.

Nonaka, I., & Takeuchi, H. (1995). “The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation.” Oxford University Press, New York.

O'Dell, C., & Jackson Grayson, C. (1998). “If we only knew what we know: identification and transfer of internal best practices.” California Management Review, 40 (3): 154-174.

Osterloh, M., & Frey, B.S. (2000), “Motivation, knowledge transfer, and organizational forms,” Organization Science, 11 (5): 538-50.

Ott, E. (2002). “Chaos in Dynamical Systems.” Cambridge University Press New, York.

Pan, S.L., & Scarbrough, H. (1998). “A socio-technical view of knowledge sharing at Buckman Laboratories.” Journal of Knowledge management, 2 (1): 55-66.

Polanyi, M. (1966). “The Tacit Dimension”. Mass.: Peter Smith (Reprint 1983).

Preece, J., & Diane Maloney-Krichmar (2003a). “Online Communities.” In J. Jacko and A. Sears, A. (Eds.) Handbook of Human-Computer Interaction, Lawrence Erlbaum Associates Inc. Publishers. Mahwah: NJ. 596-620.

Preece, J., Maloney-Krichmar, D., & Abras, C. (2003b) “History of Emergence of Online Communities.” In B. Wellman (Ed.), Encyclopedia of Community. Berkshire Publishing Group, Sage.

Price, V., & Cappella J. N. (2002). “online deliberation and its influence: the electronic dialogue project in campaign 2000”. IT & Society, 1 (1): 303-329.

Putnam, R. (2000). Bowling Alone: The collapse and revival of American community. NY: Simon & Schuster.

Rogers, E. (1983). “Diffusion of Innovations.” (3rd ed.) NY: free press.

Scott, J. (1991). “Social network analysis: A handbook”. London: Sage.

Smith, H.A., & McKeen, J.D. (2002), “Instilling a knowledge-sharing culture,” Proceedings of the third European Conference on Organizational Knowledge, Learning and Capabilities, ALBA, Athens.

Suchman, L. (1987). “Plans and Situated Actions: The Problem of Human/Machine Communication”. New York: Cambridge University Press.

Van den Hooff, B., & De Ridder, J. (2003), “The context of knowledge sharing,” Organizational Communication Division, San Diego, CA, May.

Wasserman, S., & Faust, K. (1994). “Social network analysis: Methods and applications”. Cambridge: Cambridge university press.

Wellman, B. (1982). “Studying personal communities.” In P. M. N. Lin (Ed.), Social Structure and Network Analysis . Beverly Hills: CA: Sage.

Wellman, B., & Wortley, S. (1990). “Different Strokes From Different Folks: Community Ties and Social Support.” American Journal of Sociology, 96, 558-88.

Wellman, B. (1992). “Which types of ties and network gives what kinds of social support.” Advances in group processes, 9: 207-235.

Wellman, B., & Gulia, M. (1999). “Virtual Communities as Communities: Net Surfers Don't Ride Alone.” In P. Kollock & M. Smith (Eds.), Communities in Cyberspace . Berkeley, CA: Routledge.

Wellman, B. (2001). “Computer networks as social networks”. Science, 293 (Sep.): 2031-2034.

Wenger, E. (1998). “Communities of practice. Learning, meaning and identity.” Cambridge: Cambridge University Press.

Wenger, E., McDermott, R., & Snyder, W. (2002), “Cultivating Communities of practice”, HBS.

Whittaker, S., Issacs, E., & O'Day, V. (1997). “Widening the net. Workshop report on the theory and practice of physical and network communities.” SIGCHI Bulletin, 29(3): 27-30.

Wiley, D., & Edwards, E. (2002). “Online self-organizing social systems: The Decentralized Future of Education.” http://wiley.ed.usu.edu/docs/ososs.pdf. Last accessed: May 18th, 2005

Young, P. (2002). “Diffusion in social networks”, Sante Fe Institute Working Paper 02-04-018, 2002.

Zárraga, C., & García-Falcón, J.M. (2003), “Factors favoring knowledge management in work teams,” Journal of Knowledge Management, 7 (2), 81-96.