A New Approach to the Process of Identifying Lead Users in Open Source Software Communities

Mr. Geoff Breach *
geoff.breach@uts.edu.au
School of Management, University of Technology, Sydney

Dr. Roger Jenkins
roger.jenkins@uts.edu.au
School of Management, University of Technology, Sydney

ABSTRACT

This paper presents a methodology that is designed to improve upon established approaches to the identification of lead users amongst communities of end users.

While it is clear that organisations can benefit by applying the lead user method – an approach that recognises that certain end users of products are a valuable source and in some industries the only source of new product innovations, the process has been criticised for it's resource intensive nature. The individuals who author Open Source Software (OSS) – complex and high-quality computer software that is typically distributed free of charge and without restrictions on use – appear to share many common characteristics with Eric von Hippel's (2005) lead users. These skilled computer programmers develop their software collaboratively, and often form large communities that congregate in Internet 'chat rooms' to discuss and manage their software development projects. In response to a call for further research to improve the performance of the lead user method, this paper integrates concepts from the fields of innovation management, data visualisation and open source software to construct a process designed to improve the process of identifying lead users in OSS communities.

INTRODUCTION

Recent research in a range of areas has demonstrated that product manufacturers[1] are no longer the only developers of new innovations (von Hippel 1988). Product users also innovate, and in some fields almost all innovation is done by users (Shah 2000). Certain user innovators, called lead users,

1 The term 'manufacturer' in this area of the innovation literature refers to the producer of innovations in a broader sense than the literal definition that the operations management literature applies.

are of particular interest to product manufacturers for the insight that they can provide into new trends and techniques. These innovative individuals are “the unique few who experience the future in advance of the mass market” (Nortel Networks 2006:3) They expect to make such significant gains from solutions to their unique problems that they are likely to invent solutions where none are available in the marketplace (von Hippel 1986), and they are often happy to share their solutions (von Hippel 2005).

The lead user method is a process that identifies lead users and involves them in the development of new product concepts (Herstatt & von Hippel 1992). The lead user method has been criticised for its high resource cost – substantial commitments of time and effort are required (Luthje & Herstatt 2004)(Olson & Bakke 2001). The process of identifying lead users is one of the areas that has attracted this criticism. Two methods have been used for identifying lead users – the screening approach and the networking approach – but little empirical data exists regarding the relative merits of the two methods (Luthje *et al.* 2004). This study responds in part to Luthje & Herstatt's (2004) call for further research comparing the efficiency and effectiveness of the two approaches and proposes a new screening method that is designed to improve upon current practices for selecting lead users.

The new method relies on certain special characteristics of the individuals that participate in Open Source Software projects , and of the projects themselves, to derive the identities of lead users by observing the patterns of their Internet-based interactions. A form of social network graph is constructed from an analysis of the Internet Relay Chat conversations of group members (Mutton 2004b), and a rank ordered list of individuals who exhibit leading characteristics is produced. Because the new screening method is to be entirely automated, applications of this new approach will be inherently less resource intensive than the interview-based networking approach. The lead user method's networking approach (Herstatt, Lüthje & Lettl 2001) – referred to hereafter as the reference method – is used to produce a second list of leading users from the same population.

This paper begins with a review of the innovation literature focusing specifically on lead users and the lead user method. The communities of individuals who author open source computer software are introduced and a method is proposed to efficiently identify the lead users within those communities proposed. A review of the practice of combinatoric 'graph' analysis and it's application to Internet Relay Chat (IRC) conversations is made in support of the proposed new method. A methodology is proposed to test the new method by directly comparing its results to the established lead user method and, finally, implications for the research is discussed.

LITERATURE REVIEW

Perspectives on Innovation

Joseph Schumpeter's (1974) broad perspective considered innovation was borne of those with an entrepreneurial spirit, and that it was a key driver of economies . Where Schumpeter's view was broad, Everett Rogers and Eric von Hippel have taken far narrower and more specific perspectives. Rogers' defined innovation as an “idea, practice or object that is perceived as new” (Rogers 2003:12) and focused on the processes by which innovations are disseminated, or diffused (Rogers 1962).

Recent work by von Hippel has directed focus to the identity of the innovator and specifically, the sources of innovation (von Hippel 1988), and it is these sources that are the focus of this paper.

The concept of a functional source of innovation is a means to categorise companies and individuals by "the functional relationship through which they derive benefit from a given product, process, or service innovation" (von Hippel 1988:3). Manufacturer innovators derive benefit from producing the innovation for sale, user innovators benefit from direct use of the innovation. More than one classification may apply to any one innovator.

The term users is applied herein to describe individuals or companies that hope to gain from using a product or service.

Users as Innovators

von Hippel (1988) observes that while it has previously been assumed that the vast majority of new product innovations were developed by manufacturers, there are now many fields where this is not the case. User innovators are now known to be an important source of innovations (Lettl, Herstatt & Gemuenden 2006)(Lettl & Gemünden 2005). Commercialised product innovations were developed by users in 82% of cases of scientific instruments studied (Mantel & Meredith 1986 cited in Luthje *et al.* 2004) and 67% in the field of semiconductor assembly equipment (von Hippel 1976). Different numbers of users develop solutions for their own use in different fields (Luthje *et al.* 2004), and in some fields, most or even all new product innovations derive from user innovation (Shah 2000). Users, be they organisations or individuals, are motivated to innovate by their expectation that substantial benefit will flow to them as a result (von Hippel 1988).

Innovating users will, contrary to what might be expected, often freely reveal the technical details of their innovations (von Hippel 2005) and they stand to benefit significantly – for example by way of reputational gain – in doing so (Harhoff, Henkel & von Hippel 2003). The lead user method, discussed in section below, takes advantage of this preparedness to reveal details of innovations by inviting lead users to participate in new product development.

Lead Users

Lead Users are users who exhibit two key characteristics:

1. They are trend leaders who experience strong present needs that are pre-indicators of needs that will be experienced by the general market at a later point in time, and
2. they stand to profit substantially from innovations that resolve those needs, so much so that they will often invent their own solutions (von Hippel 1986).

These lead users are of special interest to new product developers because they can offer insights that cannot be gained via traditional market research techniques that use input from existing product users as guidance for developing new products. The process of canvassing existing users is flawed because those users have experientially based pre-existing conceptions about what a particular product should do (Luthje *et al.* 2004) and are unable to avoid using their pre-existing knowledge even when instructed to do so (Marsh, Ward & Landau 1999). Their experience interferes with their ability to conceive novel solutions (von Hippel 1986). Lead users on the other hand do not need to imagine that

they are in a new situation because they, being ahead of the field as they are, are already in the new situation, and will apply their advanced knowledge to the development of solutions for the problems they actually face (von Hippel 1986).

Having recognised the importance of the lead user, a process is required to identify and engage these innovative users. The Lead User Method is such a process.

The Lead User Method

As perhaps the only formal technique for identifying innovators (Intrachooto 2004), the Lead User Method is a technique designed to find lead users amongst communities of users, and to apply their skills to the development of new product ideas (von Hippel & Sonnack 1999a).

The lead user method has four main steps: First, work is done to define intended target markets and the types of innovations desired by stakeholders. Second, research is conducted to identify current trends in the field under study. Third, the research team must identify lead users, and Fourth, the lead users and company experts meet to develop ideas for innovative new products (von Hippel, Thomke & Sonnack 1999b).

The third stage may be approached in one of two ways, screening or networking.

The Screening and Networking Approaches

Two approaches to the third – lead user identification – stage have been used: A screening approach, applied in the HILTI study (Herstatt *et al.* 1992), tested a large number of existing product users for lead user qualities, and a networking approach, applied in the Johnson & Johnson study (Herstatt *et al.* 2001) begins by interviewing a small number of users in the target market and networks towards experts by asking the interviewees to refer the interviewer to other innovative individuals. Both approaches have been shown to be useful, but there is almost no empirical evidence of the relative merits of the two processes (Luthje *et al.* 2004). The following research question has been proposed:

> "*If both approaches (screening vs. networking) are explored in comparative studies: What approach has higher performance in terms of efficiency (search time, search cost) and effectiveness (identification of leading-edge users)?*" (Luthje & Herstatt 2004:564)

This paper proposes a new approach to screening for lead users that is designed to address some of the criticisms of the lead user method, and to validate it using the established networking method as the reference method.

Criticisms of the Lead User Method

Despite significant successes in many fields (von Hippel *et al.* 1999b)(Nortel Networks 2006), the lead user method is not without its problems. The process places onerous demands of time and commitment on participants, to the extent that it distracts from routine duties (Luthje *et al.* 2004).

> "*In particular, the effort required to find, qualify, and recruit experts for trend analysis and lead users for concept generation were seen as the most burdensome task*" (Olson & Bakke 2001:391)

The process is so difficult that organisations may revert to their old methods of new product development, despite initial successes (Olson *et al.* 2001). The extensive interview processes have also been criticised for their intrusiveness: a trade-off must be made between interviewing a wide range of individuals and stressing the good will of the community, or focusing on a smaller group and risking a failure to locate the most appropriate experts (Kautz, Selman & Shah 1997).

A New Approach to Screening for Lead Users

Social network analysis and graph theory are applied to a programmatic analysis of the Internet based conversations of communities where Open Source Software is developed, supported and discussed.

In the following sections, the fields of Social Network Analysis, Open Source Software, and Internet Relay Chat are discussed as discrete fields before being assembled in Section to construct a new screening method, called the IRC method.

The IRC method is designed to improve the efficiency of the lead user identification process, and to reduce the degree of intrusion into communities that would be caused if interviews were used as in the case of the reference method.

Social Network Analysis

By studying the structure of social networks – in context of this study, networks of individuals – a researcher can expect to develop a clearer view of important linkages and key players in the networks (Angot & Josserand 2001). Network analysis techniques have been applied with success to a broad range of disciplines that include human resource management (Josserand & Teo 2004), infectious disease control (Rothenberg *et al.* 1998) (Rothenberg & Narramore 1996), control of unsolicited bulk email[2] (Golbeck & Hendler 2004), and identification of experts in Internet-based groups (Kautz *et al.* 1997).

Networks of referral (Brin & Page 1998) and trust (Golbeck 2005)(Dasgupta 1986) are strong indicators of reputation (Sabater & Sierra 2002). Newton's shoulders of giants, the academic tradition of citation (Garfield 1979)(Goffman 1971) and the PageRank (Page, Brin, Motwani & Winograd 1998) algorithm at the core of the popular search engine Google all exemplify the same principle: in the absence of the possibility of an individual evaluating all sources of information himself, the recommendations of his peers are an excellent (and possibly the only practical) substitute. It should also be noted that not all networks are reliable – evidence exists of apparent corporate abuse of Internet-based trust networks (David & Pinch 2005).

Open Source Software

Open Source Software (OSS) is computer software that is distributed under a license endorsed by the Open Source Initiative (Open Source Initiative 2006). OSS is distinguished by the terms of the licenses under which it is released. OSS licenses grant a series of rights, called freedoms, to the users of the software. Those rights typically include the freedom to observe the inner workings of the software, the freedom to make changes to the software, and the freedom to distribute copies of the

2 Colloquially known as 'SPAM'

software (Stallman 2001). OSS is not Freeware or Shareware, and it is not in the public domain – it is copyrighted, and strictly licensed (Perens 2001).

OSS Communities

OSS projects are embedded within strong communities (Lakhani & Wolf 2005). Large numbers of programmers work together to produce complex pieces of software that are often of very high quality (Raymond 2001).

Lakhani and Wolf's (2005) study revealed that a substantial proportion (58%) of the developers of OSS are experienced professional programmers and system administrators. Students (19.5%) and academic researchers (7%) also make up a significant proportion. At least half of all contributors have formal university-level training in information technology or computer science, and many more are vocationally or commercially trained computer programmers (9%). These findings are consistent with Raymond's (2001) claim that OSS culture only accepts the most talented programmers.

Raymond (2001) also observes that individual programmers begin by authoring personal solutions to a particular technical problem: “every good work of software starts by scratching a developer's personal itch” (p23). Empirical evidence supports this claim, 58% of OSS developers write their software to meet their needs – some at work and some outside of their jobs (Lakhani *et al.* 2005). The largest single gain for individuals who write OSS is intellectual stimulation (44.9%), and human capital (personal skill) development follows closely (41.8%) (Lakhani *et al.* 2005). Raymond acknowledges that programmers who submit their work to OSS projects benefit from a process of peer review of that work, and claims that their programming skills improve as a result (2001). Their technical solutions spread to the community at large “because the problem turns out to be typical for a large class of users” (Raymond 2001:49).

Raymond (2001) goes on to argue that the open source movement is a “very strict meritocracy” (p89) and observes that the programmers who make the greatest contributions to the community are the ones that are most respected by the community and who rise to take on leadership roles. The IRC method recognises the meritocratic nature of the community in highlighting lead users.

Internet Relay Chat

Of further interest is the means by which many of the more geographically diverse OSS communities communicate. Many communities hold formal project meetings, social gatherings and provide technical support to end users via an Internet based mechanism called Internet Relay Chat, or simply IRC.

IRC was created by Jarkko Oikarinen in 1988 to add a multi-participant capability to the then ubiquitous 'talk' command that was installed on most Unix computer systems of the day. While talk lacks the functionality to convene large scale communication, it is usual for a network of five or ten or more IRC servers to host hundreds of thousands of individual participants, each participating in conversations that cross multiple servers.

IRC is a text-based means of exchange. Individual participants use a piece of software, called an IRC client or 'chat client', to connect to an IRC server. In most cases, access to IRC is free of charge,

unrestricted, and largely anonymous. Participants join an IRC channel – a term analogous to the concept of a citizens band radio channel - and immediately begin to receive a rolling textual conversation on their screen. Individuals make their contributions to the conversation by way of a typed message that appears at the bottom of all participant's screens.

IRC was initially recognised by academics as a useful tool for collaboration amongst groups of geographically disconnected individuals, but a common use of the tool has been entertainment (Rheingold 1995).

Nature of Analysis of IRC Communications

The usual mode of operation for IRC client software is interactive – an individual uses the tool to join channels, and participate in conversations. This study proposes to use an automated IRC client program that records conversations for later analysis without actively participating in the conversations. An automated process of monitoring conversations and inference of relationships is proposed that will present data in a form suitable for combinatoric analysis.

IRC conversations are particularly suitable for computerised analysis (Tuulos & Tirri 2004), and many examples exist of analysis that have been made for the purpose of entertaining and informing the participants themselves (Donath 2002). Mutton's (2004) process identifies relationships between participants by detecting when they address each other by name, and infers relationships when two or more individuals speak with close temporal proximity. Despite its apparent simplicity, the process has been shown to produce social network graphs that actual participants agreed were reliable (Mutton 2004b).

All participants in IRC communications have a self-chosen name, typically a pseudonym, called a 'nick' (a contraction of 'nickname'). Nicknames are typically persistent over extended periods of time and while the identity of an individual can not usually be inferred from the name they choose, it is relatively safe to assume that the person using a given nickname today is the same person who used it on previous days (Rheingold 1995). This characteristic is relied upon to ensure that the same individuals are observed from day to day, and that those identified by the reference method and by the new screening method are able to be compared.

Common Characteristics of Actors

We have seen that lead users are individuals who experience needs that are later experienced by a community of users. They are compelled to develop their own solutions because no suitable solutions are already available and because they stand to gain substantially by doing so. The Lead User Method's networking based approach searches within communities with a view to identifying these most innovative individuals.

OSS is produced by communities of skilled individuals who recognise and are led by those who contribute solutions to needs experienced by the community at large. These individuals are innovators who set out to “scratch a personal itch” and develop their solutions into computer software that can resolve problems for themselves and for other community members. They are rewarded for their innovative contributions with intellectual stimulation, skill development and the respect of their peers.

It is clear that the case of the OSS developer shares many common characteristics with that of the lead user. These common characteristics are central to the IRC method and to the claim that the individuals identified by the new method are in fact lead users.

The IRC Method

The IRC method relies upon the characteristics of OSS developers that are common to those of lead users to identify individuals within OSS communities who are likely to be lead users. A form of social network graph is constructed by observing the IRC interactions of individuals in the OSS communities, and lead users are identified by applying established methods of social network analysis. The implementation of the IRC method is discussed in detail in the Methodology section, .

AIMS AND HYPOTHESIS

This paper establishes a methodology to construct and evaluate a new screening method – the IRC method – that is designed to be less resource intensive than the existing approaches to lead user identification that are used in stage 3 of the lead user method.

Individuals within OSS communities share common characteristics with the individuals sought by the lead user method, and it is in the nature of those communities to recognise and privilege the individuals who exhibit those characteristics. The IRC method seeks to identify these individuals.

Therefore, it is hypothesised that the IRC method will validly identify lead users in OSS communities.

METHODOLOGY

This proposal is in part a response to Luthje & Herstatt's (2004) call for empirical comparison of the different approaches to identifying lead users and it proposes to do so by applying both the networking approach and the IRC method to the same user community and comparing the outcomes.

Subsequent to selection of a population for analysis, the methodology for this study has three key components: First, the IRC method was used to programmatically generate a social network graph from the IRC interactions of OSS project members and to identify central users. Second, the interview-based networking approach to lead user identification (the reference method) is conducted amongst the same community. Third, the two lists of identified individuals are compared to assess the degree of correlation between the two lists.

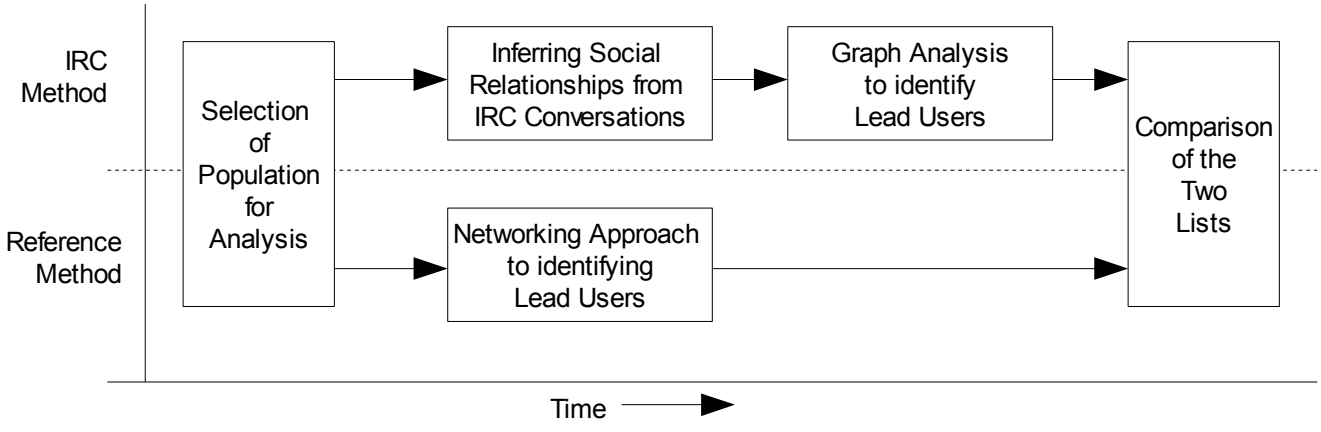

Figure 1: Steps in data collection and analysis

Four key steps occur in the collection and analysis of data. The time-based ordering of those steps is shown in Figure 1. All data collection steps, including the interview process, should be completed before data analysis begins to ensure that the researcher's knowledge of the outcome of the IRC method does not contaminate the interview process. Details of each step are discussed in the following sections.

Selection of Population for Analysis

An appropriate population for this study are the members of a large OSS project that specifies IRC as a primary means for community interaction and that has a visibly active IRC channel. By was of example, and appropriate target for a study may be the Gentoo Linux project and its corresponding #gentoo channel on the Freenode IRC network because it fits the abovedescribed criteria.

The IRC Method

The IRC method constructs a social network graph by inferring relationships between participants on a suitable IRC channel and applies basic combinatoric technique to identify the most significant participants. The following sections detail the process of deriving a list of the most significant participants from the IRC data.

Inferring Social Relationships from IRC Conversations

Mutton's (2006) PieSky software proceeded to a destination of producing visualisations of the social network graphs (see Figure 2). Whilst the generated images are useful visual indicators they are not believed to be sufficient on their own to sustain rigorous analysis. A modified version of the software to be used in this study will generate graph data in a form suitable for formal combinatoric analysis.

Mutton's (2006) software uses three techniques - direct addressing of exchanges, temporal proximity of exchanges, and temporal density of exchanges – to infer relationships between individuals. Those techniques are well documented by Mutton (2006, 2004) and so not repeated here. Because the primary goal for Mutton's implementation was to draw visual graphs and not to provide data for analysis, an implementation of the proposed methodology should test each of the three techniques to gauge their appropriateness in the data gathering context.

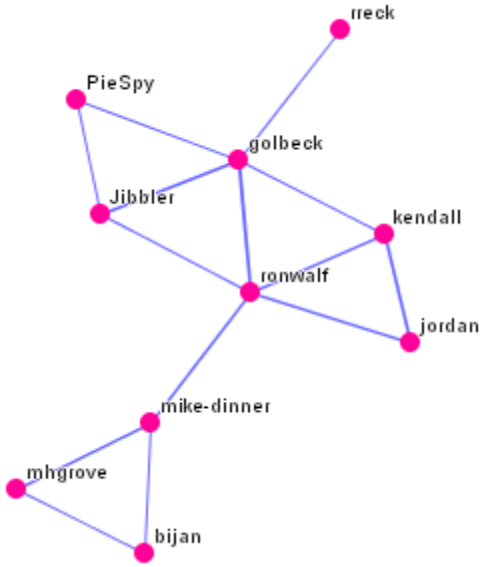

Figure 2: A simple social network visualisation.
Image source: (Mutton 2004a)

Conduct of Interviews

An approach derived from that discussed by von Hippel, Thomke & Sonnack (1999), and also used at Johnson & Johnson (Herstatt *et al.* 2001), is to be used to generate interview data to correlate with that produced by the IRC method. Individuals will be asked questions designed to assess their level of innovativeness. Additionally, are also to be asked to identify other individuals who they consider to be experts. The individuals who are identified by their peers are to be in turn invited to participate in

a similar interview. By then asking the same or similar questions of the peer-identified experts, the interviewer works his way up a "pyramid of expertise" (von Hippel *et al.* 1999b).

According to the Lead User Method, the interview process ceases when repeated identifications of the same experts indicates that the most expert individuals have all been identified. A list of lead-users is produced by identifying those that are most often referred by their peers and who exhibit the characteristics of lead users when interviewed.

Comparison of Results

The primary output from both processes of analysis is an ordered list of participant's names. A concordance will be established to link all names present in the results of both methods. The two lists are to be compared using a test such as the Kruskal-Wallis test for non-parametric comparison of independently drawn samples (Keller & Warrack 1994). Prior to comparison the results must be tested for evidence of periodicity in the data the result of the international nature of the communities under study – individuals can reasonably be expected to be present or not present at certain times of the day and week. Tests must also confirm that the results are repeatable with consistent results over an extended period of time.

DISCUSSION AND CONCLUSION

Expected Outcomes

The first author has observed in the course of his long participation amongst OSS communities that, consistent with Raymond's (2001) observations, certain individuals develop strong reputations as they make valuable contributions to their communities. In the IRC context it is usual for these individuals to become central participants in the technical discussions that take place. Applications of the Mutton (2004) software to these discussions produced visualisations of social network graphs that placed specific individuals centrally within communities that were wholly consistent with the first author's own subjective interpretations of the conversations observed over an extended period of time. Other individuals who were informally asked "*Who do you think are the technical leaders in this group?*" gave answers that were consistent with the visual data and with the author's own perceptions.

The automated nature of the proposed IRC method is that it will require far fewer resources than the established interview-based networking method. Given the nature of the preliminary observations made above, it is expected that the IRC method will produce results that very much reduce the level of face-to-face interviewing required to identify lead users.

Contributions of this Research

If the proposed IRC method is successful then, for contexts where there is a need to identify lead users amongst groups of OSS programmers, substantial resource savings stand to be achieved. Because, as we have seen, the programmers who author OSS are highly skilled individuals their skills are eminently transferable – and so too the benefits of the IRC method are likely to be of benefit in commercialised applications where their skills are required. OSS communities stand to benefit when their lead users are identified and their contributions to industry funded.

Further Research

Scope for expansion of this research exists in two key directions: First, Internet-based communities exist in support of a far broader range of activities than OSS and so it is likely that the IRC method will be transferable to other fields where similar communities exist. Second, IRC is one of a great many means by which Internet-based communities interact, and scope exists for the application of similar approaches to the one proposed to other mediums.

Conclusion

The IRC method was devised of the observation that Raymond's (2001) anthropological discussions of OSS communities and von Hippel's (2004; 1985) work on lead users discussed very similar characteristics in the individuals under study. This application of Mutton's (2004) method to IRC based OSS communities is designed to produced a screening method that identifies leading users without the high resource costs and intrusion into the community that the networking approach suffers.

REFERENCES

Angot, J. & Josserand, E. (2001). Analyzing social networks. In *Doing Management Research*. Sage, London. pp. 312-331

Brin, S. & Page, L. (1998) The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems*, *30*, 1-7, pp107-117

Dasgupta, P. (1986). Trust as a Commodity. In *Trust: Making and Breaking Cooperative Relations*. University of Cambridge Department of Applied Economics. pp. 49-72

David, S. & Pinch, T. (2005). Six Degrees of Reputation: The Use and Abuse of Online Review and Recommendation Systems. In , *Economic Sociology and Technology Conference*

Donath, J. (2002) A Semantic Approach to Visualizing Online Conversations. *Communications of the ACM*, *45*, 4, p45-49

Garfield, E. (1979) Citation Indexing: Its Theory and Application in Science, Technology, and Humanities, Wiley New York

Goffman, W. (1971) A Mathematical Method for Analyzing the Growth of a Scientific Discipline. *Journal of the ACM*, *18*, 2, p173-185

Golbeck, J. (2005). *Computing and Applying Trust in Web-based Social Networks*. Unpublished doctorial thesis, University of Maryland

Golbeck, J. & Hendler, J. (2004) Reputation Network Analysis for Email Filtering. *Proceedings of the First Conference on Email and Anti-Spam*, pp30-31

Harhoff, D., Henkel, J. & von Hippel, E. (2003) Profiting from Voluntary Information Spillovers: How Users Benefit by Freely Revealing their Innovations. *Research Policy*, *32*, 10, pp1753-1769

Herstatt, C. & von Hippel, E. (1992) Developing New Product Concepts via the Lead User Method: A Case Study in a Low Tech Field. *Journal of Product Innovation Management*, *9*, 3, pp213-221

Herstatt, C., Lüthje, C. & Lettl, C. (2001) Innovationsfelder mit Lead Usern erschliessen. *Workingpaper des Lehr-und Forschungsbereichs Technologie-und Innovationsmanagement der Technischen Universität Hamburg-Harburg (forthcoming in Harvard-Manager)*

Intrachooto, S. (2004) Research and Concepts: Lead Users Concept in Building Design: Its Applicability to Member Selection in Technologically Innovative Projects. *The TQM Magazine*, *16*, 5, p359

Josserand, E. & Teo, S. (2004). Role Identity and Social Capital Development During HRM Role Change. In , *64th Academy of Management Meetings, New Orleans, USA*

Kautz, H., Selman, B. & Shah, M. (1997) Referral Web: Combining Social Networks and Collaborative Filtering. *Communications of the ACM*, *40*, 3, pp63-65

Keller, G. & Warrack, B. (1994) Statistics for Management and Economics, Fourth Edition, Duxbury Press Belmont, Calif

Lakhani, K. & Wolf, R. (2005). Why Hackers Do What They Do: Understanding Effort and Motivation in Open Source Projects. In *Perspectives on Free and Open Source Software*. MIT Press, Cambridge, MA

Lettl, C. & Gemünden, H. (2005) The Entrepreneurial Role of Innovative Users. *Journal of Business & Industrial Marketing*, *20*, 7, pp339-346

Lettl, C., Herstatt, C. & Gemuenden, H. (2006) Users' Contributions to Radical Innovation: Evidence from Four Cases in the Field of Medical Equipment Technology. *R&D Management*, *36*, 3, pp251-272

Luthje, C. & Herstatt, C. (2004) The Lead User Method: An Outline of Empirical Findings and Issues for Future Research. *R & D Management*, *34*, 5, p553

Mantel, S. & Meredith, J. (1986). The Role of Customer Cooperation in the Development, Marketing, and Implementation of Innovations. In *The Art and Science of Innovation Management*. Elsevier, Amsterdam. pp. 27-36

Marsh, R., Ward, T. & Landau, J. (1999) The Inadvertent Use of Prior knowledge in a Generative Cognitive Task. *Memory & Cognition*, *27*, 1, pp94-105

Mutton, P. (2004a). *The PieSky Social Network Bot*. Retrieved from http://www.jibble.org/piespy/. Accessed 10th July, 2007

Mutton, P. (2004b) Inferring and Visualizing Social Networks on Internet Relay Chat. *Eighth International Conference on Information Visualisation, 2004. IV 2004. Proceedings*, pp35-43

Nortel Networks (2006) Lead Users and Dynamic Information Transfer, Nortel Networks, USA

Olson, E. L. & Bakke, G. (2001) Implementing the Lead User Method in a High Technology Firm: A Longitudinal Study of Intentions versus Actions. *The Journal of Product Innovation Management*, *18*, 6, pp388-395

Open Source Initiative (2006). *Open Source Initiative OSI - Certification Mark and Process*. Retrieved from http://opensource.org/docs/certification_mark.php. 26th November 2006

Page, L., Brin, S., Motwani, R. & Winograd, T. (1998). *The PageRank Citation Ranking: Bringing Order to the Web*. Unpublished technical report, Stanford University

Perens, B. (2001). Open Source Software. In *Revolution OS, 35mm film*. Wonderview Productions, California

Raymond, E. (2001) The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary, O'Rielly and Associates, Sebastapol, CA

Rheingold, H. (1995) The Virtual Community: Finding Connection in a Computerized World, Martin Secker & Warburg Ltd, England

Rogers, E. (1962) Diffusion of Innovations, Free Press New York

Rogers, E. (2003) Diffusion of Innovations, 5th Ed, Free Press, New York

Rothenberg, R. & Narramore, J. (1996) Commentary: The Relevance of Social Network Concepts to Sexually Transmitted Disease Control. *Sexually Transmitted Diseases - Journal of the American STD Association*, *23*, , pp24-29

Rothenberg, R., Potterat, J., Woodhouse, D., Muth, S., Darrow, W. & Klovdahl, A. (1998) Social Network Dynamics and HIV Transmission. *AIDS - Official Journal of the International AIDS Society*, *12*, 12, pp1529-1536

Sabater, J. & Sierra, C. (2002) Reputation and social network analysis in multi-agent systems. *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*, pp475-482

Schumpeter, J. (1974) Capitalism, Socialism and Democracy, Fourth Edition, Routledge New York

Shah, S. (2000) Sources and Patterns of Innovation in a Consumer Products Field: Innovations in Sporting Equipment. *WP-4105, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Mass*

Stallman, R. (2001). *Transcript of Richard M. Stallman's speech, "Free Software: Freedom and Cooperation" New York University in New York, New York on 29 May 2001*. Retrieved from http://www.gnu.org/events/rms-nyu-2001-transcript.txt. Accessed: 26th November 2006

Tuulos, V. & Tirri, H. (2004) Combining Topic Models and Social Networks for Chat Data Mining. *WI'04: Proceedings of the Web Intelligence, IEEE/WIC/ACM International Conference on (WI'04)*, pp206-213

von Hippel, E. (1976) The Dominant Role of users in the Scientific Instrument Innovation Process. *Research Policy*, *5,* pp212-239

von Hippel, E. (1986) Lead Users: A Source of Novel Product Concepts. *Management Science*, *32*, 7, p691

von Hippel, E. (1988) The Sources of Innovation, Oxford University Press, USA

von Hippel, E. (2005) Democratizing Innovation, MIT Press, Cambridge, MA

von Hippel, E. & Sonnack, M. (1999a) Breakthroughs to Order at 3M. *MIT Sloan School of Management Working Papers*

von Hippel, E., Thomke, S. & Sonnack, M. (1999b) Creating breakthroughs at 3M. *Harvard Business Review*, *77*, 5, p47