Author: *Han, Sangman, C. L. Magee, Y. Kim.*
Article to be submitted to *Management Science*

Hyperactive Agents in Social Network Sites: the evidence for life-cycles and determinants of lifecycle variability

Sangman Han
School of Business Administration, SungKyunKwan University, Seoul, 110-745, smhan@skku.edu

Christopher L. Magee
Center for Innovation in Product Development, Massachusetts Institute of Technology, Cambridge, MA 02139, cmagee@mit.edu

Yunsik Kim
Center for Innovation in Product Development, Massachusetts Institute of Technology, Cambridge, MA 02139, yunskim@mit.edu

This paper examines a highly used social network site (SNS) by studying the behavior of more than 11 million members over a 20 month period. The importance of the most highly active members to the overall network is demonstrated by the significant fraction of total visits by extremely active members. By examining the time dependence, we find that such members do not have extensive life spans as hyperactive agents but instead have average lifespan of only 2.5 months. We form and test a number of hypotheses concerning these hyperactive agents and the determinants of their lifespan. It is confirmed that persons aged 18-22 are much over-represented in the population of hyperactive agents whereas contrary to expectations women are statistically more likely to be hyperactive agents than men. We also find that speed of achieving hyperactive agent status increases the lifespan of a hyperactive agent. The norm of reciprocity is strongly confirmed to be present in the overall population and in the hyperactive agent population. However, contrary to expectations, local clustering in the vicinity of hyperactive agents is smaller (rather than larger) than overall clustering. These results show some important ways that the SNS is similar to real social networks but also important ways that they are different. In particular the fact that agents practice the "norm of reciprocity" is homologous with real social networks but the strong dominance behavior and lack of strong local clusters are very different for the virtual social network relative to what is known about real social networks. These findings have sociological as well as managerial implications which are discussed in the paper.

Key words: social networks, hub, lifecycle, reciprocity, network structural properties

History:

---

1. Introduction

The early 1990s development of the world-wide web as an application layer utilizing the power of the Internet was largely seen by developers from an essentially sociological perspective (Berners-Lee 1999).

Indeed, it was not long after the world wide web initiation that academic studies began to explicitly recognize that "when computer networks link people as well as machines, they become social networks" (Wellman et al 1996). The early work on Computer Supported Social Networks focused upon work groups using computer communication to achieve tasks in a geographically dispersed manner (Wellman et al 1996). A few years later, a broader role for Computer Supported Social Networks was recognized (Wellman 2001) including impact on non-work communities and suggestions for classifying types of communication in virtual communities (Burnett 2000). More recent work along the sociological dimension has focused upon learning and relational perspectives on learning (Borgatti and Cross 2003) as well as the application of behavior setting theory to virtual communities (Blanchard 2004).

The economic and business aspects of the world-wide web were much less considered in its development (Berners-Lee 1999, Clark 1998, Castells 2001). Thus, the early efforts to explicate the business and economic aspects of virtual communities were not surprisingly searching for the correct framework (for example, Balasubramanian and Mahajan, 2001, Williams and Cothrel, 2000). The recent successes of search and auction firms have now demonstrated two viable business models. However, for our purposes the rapid recent evolution of SNS (Social Network Sites) that appear economically viable is particularly noteworthy and such virtual social communities are the principle focus of this work. Such SNS (Social Network Sites) offer personal profiles, blogs, clubs, photos, music, video streams and allow users to submit friends as possibly of interest for their virtual social network. In these sites, young people use digital images, music and postings to express themselves and share experiences with others (Business week 2005). They support personal homepages to strengthen relationships with each other and enable them to establish an online community. Through maintenance of a personal homepage, users can optimize their self presentation and identity with photos, music and other uploaded information. Recently, many SNS (Social Network Sites) have been launched –for example in Korea, Cyworld.com[1] was launched in 1999, in the U.S., myspace.com and facebook.com were launched in 2004, in Japan, Mixi was launched in 2004,

[1] This SNS is the subject of the research and is described in more detail in the following section

in U.K. Bebo was launched in 2005, and in China, Sina blog and Qzone were launched in 2005. Myspace.com and facebook.com are the most popular social networking sites among the young generation in US, especially, myspace.com which gained the top ranking of US website in 2006(Reuters 2006). Similarly, the earlier launched cyworld.com is most popular in Korea with 22 million memberships in Oct 2006. Some basic facts about significant (in addition to Cyworld) SNS (social network sites) are shown in table 1.

Table 1: The types and market size of selected social network sites[2].

	Facebook	**Myspace**	**Bebo**	**Mixi**	**Sina Blog**	**Qzone**
Language	English	English	English	Japanese	Chinese	Chinese
Type	SNS based on campus life	Music community & friend-making Community	SNS based on school and college community	Japanese Local SNS	blog (portal supported) based SNS	Blog (portal supported) based SNS
Market Share	- 7.5M+ users - Alexa ranking : 33rd of the World	- 127M+ users - 77% of users from 14 to 35 - Alexa Ranking: 6th of the World	- 22M+ users - The 1st of UK SNS (12.91%) - Alexa Ranking: 174th of the World	- No. 1 of Japanese SNS (about 70% Market Share): over 8 million members - 52nd of the World, 6th in Japan	- 32.86% of market share in Chinese internet service - Sina Portal : No. 10 worldwide, No. 3 in china	- 19.28% of market share in Chinese internet service - QQ portal : No. 9 worldwide, No. 2 in China

Simultaneously with the recognition of the World Wide Web as a social (as well as technical) phenomenon, there has arisen a strong use of "Network Analysis" in the study of such large-scale socio-technical systems (Watts 2004, Newman 2003). Some of this work has emphasized the existence of power laws in degree distribution (Barabasi 2002, Barabasi and Bonabeau 2003, Price 1965, 1976) and have called attention to highly connected nodes in networks (they are often called hubs- we will call them "hyperactive agents" in this study). Much of this work has simply emphasized connections or links but some recent

[2] The data in table 1 comes from 'the social networking faceoff', 'Internet Guide' and 'Miki' that can be accessed at http://www.readwriteweb.com/archives/social_network_faceoff.php, http://tech.sina.com.cn/socus/2006_GUIDE2007/index.shtml, and http://en.wikipedia.org/wiki/Mixi, respectively.

dynamic work (Braha and Bar-Yam 2004) calls attention to the importance of actual use of the links rather than just existence of links. In our study of the SNS, we particularly emphasize the activity of these highly connected agents over time. They are found to be vital in information flow, disease propagation, or word-of-mouth spread in the network. Previous research found the mechanism called preferential attachment as the process of network evolution. However, there has been no study on what influences the life-span of the highly connected agents. Such agents seem to be very important in keeping the network active and appealing, and thus their longevity and patterns of activity over time have potentially interesting social network and business implications. Therefore, the major foci of this paper are 1) the relative preponderance of such agents relative to agent and network characteristics and 2) the time dependence of such agents to examine the nature and determinants of their life-cycle. In particular, we explore the demographic and network properties these hyper-active agents possess, the time dependence or life-cycle of these hyper-active agents, and which factors influence the life-span of the hyper-active agents.

2. Overall Characteristics of the network

2.1 Characteristics of data studied for Cyworld.com

Cyworld.com was launched in 1999 and is currently the most popular social network site in Korea with 22 million memberships in Oct 2006. With formal launch in Korea in 1999, Cyworld was merged into nate.com which is a popular portal service in Korea in 2004. It has been reported that as much as 90% of the Korean population in their 20's and a third of the total population of Korea are registered user of Cyworld. In their homepages, people can accommodate a lot of documents, photos, and appealing items for free but many choose to decorate his/her "minihompy" (Mini homepage) with paid for items. Many people in Korea consider Cyworld as part of every day life with regard to building relationships with each other and publish his/her daily life on their minihompy to share with others. The number of monthly unique visitors is about 20 million in Cyworld. Cyworld generates revenue from the sale of cyber money which is called dotori and is worth about 0.3 million dollars a day. The revenue of Cyworld comes from the pay-to-decorate model and the paid advertising model. Cyworld makes an estimated more than $7 per person a year from the pay-to-decorate model.

We obtained anonymous records for 11,163,690 members on cyworld.com for a 20 month period from Dec 2003 to July 2005. In the data, we excluded 10,074 members, which are outliers, having extreme value of either Kin or Kout. These outliers are special sites whose purpose is to promote their websites for commercials such as the brand minihompy or are sites for other business purposes. We studied the network for a series of 1 month periods and define members as nodes and a visit to another node during each 1 month period as a link in our analysis. Thus our links metric is an indicator of activity for the 1 month period.

Figure 1 shows the undirected cumulative degree distribution of the entire membership over 20 months and the dotted line is a power law fit of form P(x)≈ Cx-α with α = 4.07 (since we used a cumulative degree distribution, the observed exponent is α – 1 or 3.07)

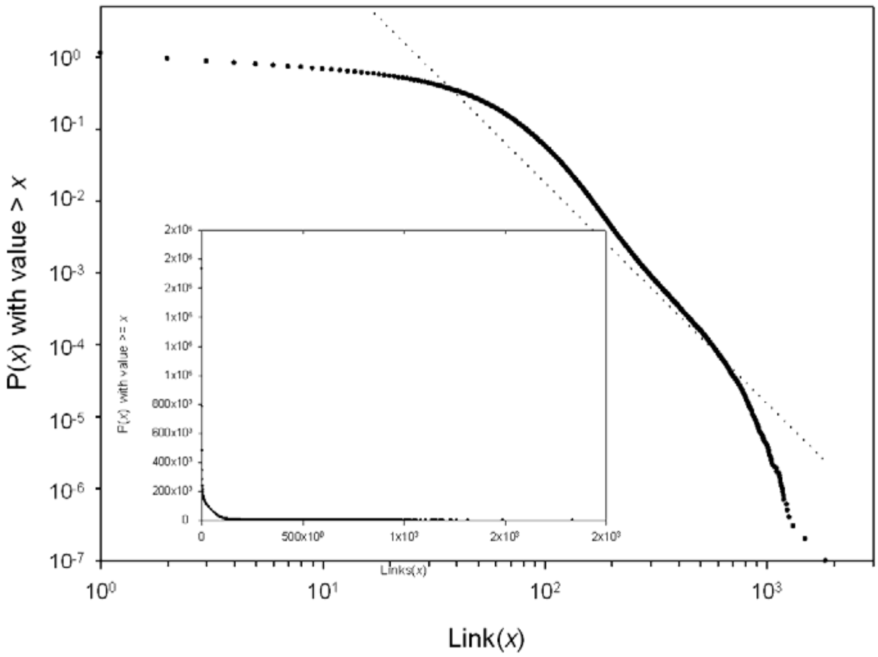

Figure 1: Cumulative degree distribution of total agent population.

Figure 1 is a plot of the cumulative distribution function P(*x*), which is the fraction of nodes with links greater than or equal to *x*, where x is the number of links connecting to given nodes (Newman 2005). It shows that some agents are particularly active (>1000 visited sites in a period) but many other agents are not very active (> 5 million have < 10 sites visited in a period)

Our regression showed a reasonable fit (R^2 = 0.91) with significant parameter values at significant level but clearly only fits for k between ~ 30 and 800. The result shows that the degree distribution (P(*x*)) follows a power low distribution over the range of ~30 to 800 links per agent with exponent 4.07. The exponent α is found to be within the range defined by other network research (Dorogovtsev and Mendes 2002, 2003, Albert and Barabasi 2002, Newman 2003, 2005).

2.2 Definition of hyperactive agents

We characterize activity in 2 distinct dimensions: Kin and Kout. Kin refers to the total number of incoming visits and Kout refers to the total number of visiting others in the network. We classify agents having greater than 1 standard deviation above the mean as active and those with higher than 3 standard deviations above the mean as hyperactive and thus arrive at 9 categories as shown in Figure 2.

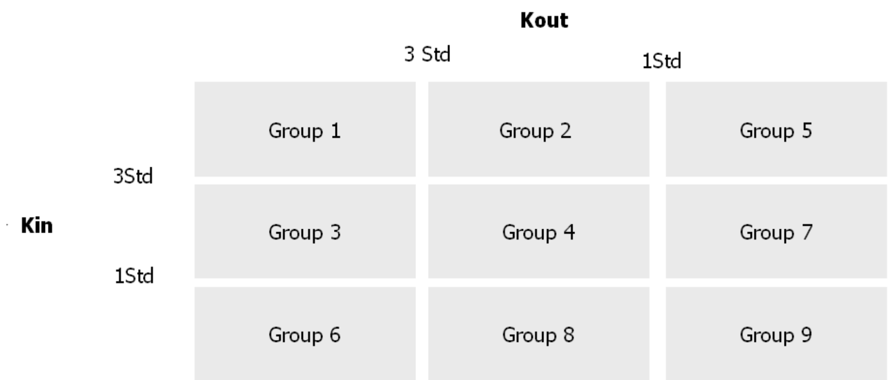

Figure 2: Characterization of nodes.

Group 1 consists of agents where both Kin and Kout are greater than the mean activity plus 3 standard deviations. Group 5 refers to agents with average Kout and hyperactive Kin while group 6 refers to agents with average Kin and hyperactive Kout. We conceptualize group 1, 2, 3, 5 and 6 as hyperactive agents in their network and define group 1 as balanced hyperactive agents, group 2 and 3 as unbalanced hyperactive agents and group 5 and 6 as maximum unbalanced hyperactive agents. Group 9 contains agents having average or below visits and average or below visiting others in their network.

Since one focus of this research is the birth and death of hyperactive agents, the overall time dependence of the number and concentration of such agents is important background information. Consisistent with the "fat tail" power law seen in Section 2.1, there are far more of such agents than would be expected from a normal distribution., Figure 3 shows that the change of the number of balanced hyperactive agents closely follows the total number of agents. The average ratio of balanced hyperactive agents to the entire membership is about 1.25% ± 0.15% over the entire time period.

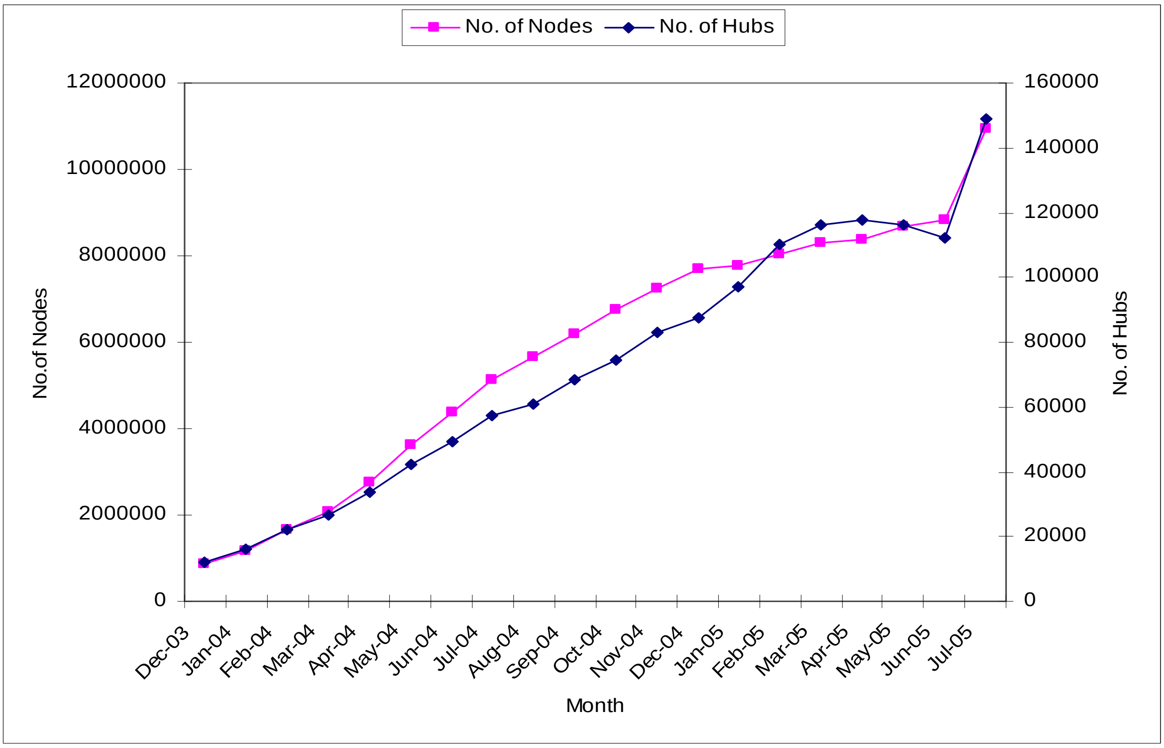

Figure 2: The number of balanced hyperactive agents (group 1) over time.

3. Agent characteristics influencing the life span of hyperactive agents

The vitality of a virtual social network is determined by the membership size and the activity of the agents in the network. The most active agents are much more active than the average agent and thus may play a particularly important role in network vitality. Indeed, only 4.19% of agents are hyperactive agents.

However, they account for on average 9.01% of the total visits by agents. Our research is based upon the assumption that the life span or permanence of the most active agents is of particular importance to the long term health of any virtual social network. In the previous section, we have defined the agents according to their visiting and visited activities. We define the life-span of "hyperactive" agents as the time interval during which these agents keep their activity level above the defined thresholds (this and other definitions are discussed in Section 4.1). In this section, we develop hypotheses on the factors which influence the life-span of hyperactive agents. The factors are classified into three categories—demographic factors, agent-related factors and network-related factors.

Demographic factors

Rogers (1995) studied the characteristics of initial adopters of a wide variety of new technologies and found that the initial adopters are more likely to be young, male and better educated. This finding supports the view of a digital divide between males and females. There have also been numerous studies on the gender differences in Internet usage. These studies consistently found that men tend to be more interested in computers than woman, which may carry over to gender differences in Internet use (Shashaani 1997, Herring 2000). Ono and Zavodny (2003) examine the extent of gender differences in several measures of internet usage during the period 1997-2001. The results show that women are significantly less frequent and less intense users of the Internet.

According to Korgaonkar and Wolin (1999), even though the Internet gender gap is disappearing and women and men are equally likely to use the Internet, men are still more likely to purchase products from the internet. Herring (2000) shows that a gender difference exists in online communication. In mixed-sex public discussion groups, females post fewer messages and are less likely to continue their posting activity when their messages receive no response. When women use the Internet, they tend to participate more in exchanging private e-mails rather than participating in public discussion groups and chat rooms, whereas men tend to participate more in public discussion groups and chat rooms (Hoffman et al. 1996).

Based upon this prior work, there should be gender differences in terms of both the number and the life-span of female and male hyperactive agents. Specifically the following hypotheses are indicated:

H1a: *Females will be less represented in the hyperactive agent group than they are in overall social network community.*

H1b: *Female hyperactive agents will have shorter life-span than male hyperactive agents.*

Age differences have been found in people's attitudes toward the Internet. Younger people express more enjoyment and self-efficacy than older people (Zhang 2002, Patricia, Perotti and Widrick 2005). There is an interesting study by Mikami (2002) that people of age 18-24 spend the most time socializing with friends in both the U.S. and Japan. Since those in late teens and early twenties are also voracious users of new technologies—in particular, Internet related technologies such as SMS, Instant messenger service, on-line community, etc, they are more frequent and intense users of on-line social networks. Both of the two effects described above—1) more propensities for socializing with friends and 2) more frequent and intense use of Internet-related technologies make those in their late teens and early twenties likely to have higher activity levels in specific on-line social network sites.

Thus, we develop our second hypothesis as follows:

H2a: *People in their late teens and early twenties will be more strongly represented in the hyperactive agent group than they are in the overall population of the social network.*

H2b: *Hyperactive agents in their late teens and early twenties will have a longer lifespan than other hyperactive agents.*

Agent related factors

Previous studies found that individuals in real social networks pay close attention to their peers, constantly sending out signals and adjusting their behaviors based on feedback from those with whom they are interacting (Resnick 2004). Positive feedback response from peers is a sign of recognition, approval and even respect by other members of the social network. Members of social networks spend their time and energy in ways that increase the recognition and trust by their peers. This feedback response from peers can strengthen the engagement of users and deepen their commitment to the social network. Thus, the more feedback response the users gets from their peers, the more time and energy the users spend in the social network, which results in a positive relationship between the level of feedback response and the level of internet usage in the on-line community site. This phenomenon of feedback and response is called the norm of reciprocity. The norm of reciprocity can lead to people matching behaviors experienced from others with actions performed for others, giving in proportion to what they receive (Carr 2006). The norm of reciprocity has been found in many real communities (Thorn and Connolly 1987, Constant, Sproull and Kiesler 1996).

As in a real community, the norm of reciprocity prevails in the on-line community. It is a general norm that whatever is given ought to be repaid (Wellman and Gulia 1999). Users in the on-line community reciprocate visits to each other, even to those strangers who visit them for the first time. Constant, Sproull and Kiesler (1996) suggest two explanations for this norm of reciprocity. The first explanation is that the process of reciprocating visits to other members is a means of expressing self identity. By reciprocating the visits to other members, the user can strengthen one's self-identity and attain a certain status. The second explanation is that the norm of reciprocity is a means of showing a strong attachment to the community and a strong bond to the members of the same community.

We can think of those agents with hyperactive status as the ones who have a stronger desire to increase self-esteem, gain respect from others and attain certain status in the social network. This high desire for self-esteem make those hyperactive agents try harder to reciprocate visits they receive from their neighbors. The reciprocity of relationship between the hyperactive agents and their peers is assessed by the balance of incoming visits and outgoing visits of the agents. Thus, we develop the following hypotheses:

H3a: The *ratio of incoming visits and outgoing visits are highly correlated.*

H3b: *balanced hyperactive agents are more abundant than would be expected if outgoing and ingoing visits were uncorrelated.*

H3c: *The more balanced hyperactive agents are in terms of incoming and outgoing visits, the longer life-span they tend to have.*

Network analysts have shown that many real-world networks exhibit power laws in degree distribution. Social networks typically exhibit this characteristic as well (Price 1965). Price (1976) and later Barabasi and Albert (1999) argued that when new agents decide where to establish a link, they prefer to attach to an existing agent that already has many other connections. This basic mechanism was called cumulative advantage by Price (1976) and preferential attachment by Barabasi and Albert (1999). This mechanism can lead to a social network dominated by hyperactive agents.

Hyperactive agents, because of their large numbers of connections within the social network, know where the most interesting information is located. Because they enjoy the trust of other members of the network, they are also able to disseminate it to other members more easily (Kleinberg 1999). This role of hyperactive agents can be strengthened if they dominate their local social network. In other words, the more dominating hyperactive agents are in the relationship within its local network, the more important their role becomes and the more active they become in collecting and disseminating the information. If a hyperactive agent is sharing the same local network with other hyperactive agents, his/her dominance in the local network will be weakened and he/she becomes less active in information collection and dissemination. This implies that the more an agent dominates the local social network, the more that agent is engaged in the role of collecting and disseminating information in the social network.

If an existing agent becomes hyperactive much faster, it monopolizes the cumulative advantage mechanism for itself. In other words, if there is only one agent with very high activity in a local social network, the preferential attachment or cumulative advantage will favor only that agent. As a result, the time for this agent to become hyperactive becomes shorter. This hyperactive agent, in turn, dominates the local social network, and its role as collector and distributor of information is strengthened. Therefore, the shorter the time to become hyperactive, the more dominating the agents are in their role of collecting and disseminating information, which leads to longer life-span. We develop the following hypothesis:

H4a: *The shorter is the time for an agent to become hyperactive, the longer becomes its life-span.*

H4b: *The more active the hyperactive agent is, the longer will be its lifespan.*

Network related factors

We argue in the previous hypothesis development that as the more dominating hyperactive agents are in the relationship within its local network, the more important their role becomes and the more active they become in collecting and disseminating the information (Kleinberg 1999). This means that if the dominance of the hyperactive agents diminishes, their role of information collection and dissemination may also become weaker. In other words, the more a hyperactive agent is sharing the role of information collection and dissemination with other neighbor members in the same local network, his/her dominance in the local network will be weakened and he/she becomes less active in information collection and dissemination. This implies that the more active the neighbor agents become, the hyperactive agents become less dominant and therefore less active in their local network.

Thus, we develop the following hypothesis:

H5: *The more active a hyperactive agent's neighbors are, the shorter its life-span.*

Putnam (1995) defined social capital as the features of social organization, such as trust, norms and networks that can improve the efficiency of society by facilitating coordinated actions. Social capital refers to the amount and quality of communication about a community that takes place among its members within the social networks (Kavanaugh and Patterson 2001). The increase in social capital leads to a strong feeling of companionship, emotional bond, a sense of belonging, which in turn increases the participation of members in community-related organizations and activities. Thus, communities with vibrant communication networks are likely to achieve common social goals and to sustain their longevity. Since the longevity of a community is expected to be correlated with the longevity of its central node (Dorogovtsev and Mendes 2003), we expect that higher social capital can also lead to longer life-span of hyperactive agents in the community.

As Putnam (1995) suggested, social capital can increase companionship and a sense of belonging, and it can motivate members to participate in voluntary activities for the community. In order to create the social capital, it is vital to have sub-communities within which coherent members are densely clustered and vibrantly communicate with each other. Oh and Jeon (2007) found that social capital plays a significant role in sustaining a community, especially one that is founded on member's volunteerism. They showed that key factors affecting the longevity of an OSS(Open Source Software) community is the dynamic interaction of its members. Koku and Wellman (2002) also found from the study of scholarly networks that the high density of the network predicts the network's success in fostering contact. In particular, small sets of clusters were found, within which coherent sets of researchers work together, and these blocks of coherent clusters increase the collaboration between researchers.

The clustering coefficients of hyperactive agents indicate the level of dense connection among their neighbors. Thus, high clustering coefficients of the hyperactive agents imply a higher level of connections within their local networks, which enables vibrant communication within the local network. This leads to the higher social capital of the hyperactive agents' local communities and the longevity of the local communities, which leads to the longevity of the hyperactive agents. From these arguments, we develop the following hypothesis:

H6: *The higher is the clustering coefficient of a hyperactive agent, the longer is its life-span.*

4. Methods and results

4.1 Definition of variables

We define life span for hyperactive agents as the time a given agent/node remains at a hyperactive agent status and TBH as the time to become a hyperactive agent for a given agent/node.

We define Imbalance for an agent in its network as the extent of imbalance of outgoing visits to incoming visits specifically defined as Imbalance for the *i*th agent as

$$B_i = \frac{\sum_{t=1}^{20} \frac{|kin_{it} - kout_{it}|}{kin_{it} + kout_{it}}}{T},$$

Where T is the periods and *i* is an agent. B is 1 for agents who only visit others (or who visit no-one but just receive visitors) and is 0 for agents where Kin = Kout.

A definition of a local clustering coefficient which was introduced by Watts and Strogatz is useful for us because we are interested in how the clustering coefficient around the hyperactive agents has any influence on their life-span. Following Watts and Strogatz, we define that the clustering coefficient for node *i* in terms of the interconnection among the neighbors of node *i*. We define the k_i as the number of neighboring nodes node *i*. $k_i(k_i - 1)/2$ links can exist among these neighbors for node *i* and n_i is the actual number of links existing among the neighboring nodes connected with node *i*.

The clustering coefficient is defined as the ratio of actual to possible links among neighbors of nodes *i*,

$$C_i = \frac{n_i}{k_i(k_i - 1)/2},$$

If every neighbor connected to node *i* is also connected to all other neighbors of node *i*, then C becomes 1. If none of the nodes connected to node *i* is connected to each other, then C becomes 0.

We take the weighted average to calculate the average of the variables when some nodes belong to more than 2 groups. If all the weights are equal, it is the same as an arithmetic mean. Table 2 gives a summary of the variables, and the variables are defined by the following equations.

The weight w_{ij} is measured as the proportion of each group *j* in Life-span of node *i*.

$$w_{ij} = \frac{Life\,span_{ij}}{\sum_{j=1}^{n} Life\,span_{ij}},$$

$$Imbalance_i = \sum_{j=1}^{n} w_{ij} B_{ij},$$

$$Activity_i = \sum_{j=1}^{n} w_{ij} activity_{ij},$$

$$CC_i = \sum_{j=1}^{n} w_{ij} CC_{ij},$$

$$Act\,Neigh_i = \sum_{j=1}^{n} w_{ij} Act\,Neigh_{ij},$$

Where *j* stands for groups, and *i* stands for a node.

Table 2: summary of variables on analysis

Variable	Variable description
Life span$_i$	The time to remain a hyperactive agent status for agent *i*.
TBH$_i$	The time to become a hyperactive agent for agent *i*.
Kin$_{it}$	A total number of incoming visits for agent *i* in period *t*.
Kout$_{it}$	A total number of outgoing visits for agent *i* in period *t*.
Activity$_{it}$	The sum of Kin and Kout for agent *i* in period *t*.
Imbalance$_i$	The extent of balance of outgoing visits to incoming visits, 0 means very balanced and 1 very unbalanced.
CC$_{it}$	Clustering coefficient for agent *i* in period *t*.
Act Neigh$_{it}$	The activity of agents *i*' neighbors in period *t*.
Gender	The categorical variable (male 0, female 1).
Age	From 13 to 39 categorized into 13-17, 18-24, 25-29 and 30-39.

4.2 Demographic properties of hyperactive agents

This data are obtained from a database extracted from Cyworld from Dec 2003 to July 2005 for 20 months and the total number of agents used for our analysis is 11,163,690. The number of hyperactive agents is 468,278; among them 230,491 are balanced hyperactive agents. Table 3 describes the demographic properties of hyperactive agents, balanced hyperactive agents, and the entire membership. As we predicted in hypothesis 2a, members in the age 18-24 age group are more strongly represented in the hyperactive agent group than they are in the overall population of the social network. In the overall population, the proportion of 18-24 age group is 31.87%, while the same age group represents 67.56% in the hyperactive agents. This figure becomes even higher for balanced hyperactive agents (71.55%). Chi-square test for the difference of the 18-24 age group participation between hyperactive agents and entire membership shows that there is a significant difference ($\chi 2$=257,717, p<.001) between the two groups.

This result supports our hypothesis (H2a) that people of age 18-24 are more likely to be highly active in virtual social networks.

While we expected to observe higher proportion of males as hyperactive agents, table 3 shows that the composition of females is slightly larger than that of males. Even though the proportion of females is only slightly larger (48.49% for male and 51.51% for female in hyperactive agents), this is a statistically meaningful finding that contradicts the view of a digital divide between males and females. For the entire population, we cannot reject the prevailing view that men are more intense users of internet-related activities (the proportion of female is 49.89% while that of male is 50.11%). Chi-square test shows that there is a significant difference in the proportion of females between hyperactive agents and entire membership ($\chi 2$=472.4541, p<.001).

Table 3: Demographics of agents

	Total Hyperactive Agents	**Balanced hyperactive agents**	**Entire membership**
No of nodes	468,278	230,491	11,163,690
Age	22.79	22.69	25.74
13 to 17	3.38%	2.06%	10.76%
18 to 24	67.56%	71.55%	31.87%
25 to 29	25.13%	23.37%	28.82%
30 to 39	3.93%	3.02%	28.56%
Sex			
Male	48.49%	49.01%	50.11%
Female	51.51%	50.99%	49.89%

4.3 Life-Cycle Analysis of Hyperactive Agents

Hyperactive agents play a role in indirectly linking members in 2 steps who otherwise would only be connected together in far more steps. Hyperactive agents are in many cases the sources of information flow in the network by collecting information and disseminating it to their peers. Previous research has identified the process of how a certain agent becomes a hyperactive agent over time. “Growth” and “Preferential Attachment” are found to be the governing principle of the birth of hyperactive agents in some networks. However, no previous research has investigated the life-cycle of hyperactive agents over time because of the lack of the relevant dataset. In this section, we try to answer the following questions: 1) is there a life-cycle for hyperactive agents? 2) How long is the cycle? 3) How do the hyperactive agents become inactive over time?

Table 4 shows that hyperactive agents do show some form of life-cycle. It takes an average of about 7.14 months after they become a member until they reach hyperactive agent status. On average, they stay for

2.45 months as hyperactive agents. The average activity level (sum of Kin and Kout) is 122 for hyperactive agents and 141 for balanced hyperactive agents. Hyperactive agents are in general very balanced (the imbalance score is only 0.15) and this implies that the norm of reciprocity prevails. The life-span of hyperactive agents is far shorter than we expected. One of the reasons could be the extreme definition of hyperactive agents. We define hyperactive agents as those whose Kin or Kout is at least 3 standard deviation away from its mean, and we measure the life-span during which they retain the status of hyperactive agents. Another reason could be the ceiling effects of the reciprocal behaviors of hyperactive agents. Ceiling effects imply that there exists some kind of an upper limit of time and efforts for hyperactive agents to spend in order to maintain the norm of reciprocity. The results suggest that hyperactive agents cannot maintain their highly active and reciprocal behavior for more than 2.5 month on average.

Table 4: Descriptive analysis of hyperactive agents and balanced hyperactive agents

	Total Hyperactive agents		**Balanced hyperactive agents**	
Variable	**Mean**	**SD**	**Mean**	**SD**
Life Span	2.4546	2.0994	2.3932	2.0014
TBH	7.1473	4.2889	7.0309	4.2684
Activity	122.0430	29.4810	141.7390	33.2348
Imbalance	0.1514	0.2005	0.0669	0.0639
CC	0.0783	0.0494	0.0799	0.0430
Act Neigh	70.4612	18.5920	74.7426	17.5805

Table 5 shows migration patterns among groups. It provides an interesting finding regarding the life-cycle path of hyperactive agents in the network. If we examine the column percent of hyperactive agents (groups 1, 2, 3, 5, and 6), the largest proportion of new hyperactive agents arises from the new entries (31.4%-47.4%). More than 30% of hyperactive agents are new entries in the period of examination. The next largest proportion of new hyperactive agents is from more balanced groups such as group 4 (3.1%-12.3%) and group 9 (7.3%-26.9%). This shows that more balanced agents are more likely to evolve into hyperactive agents over time.

The row percent of hyperactive agents (group 1, 2, 3, 5, and 6) also shows an interesting finding that hyperactive agents drop out of the network completely at a much higher ratio than we expected (31.5%-50.2%). This result with the pervious finding that new entries supply the largest proportion of hyperactive agents implies that the composition of hyperactive agents is very unstable. The turn-over of hyperactive agents is very high (more than 30% of them in any period are new and more than 30% of current hyperactive agents lose their status in every period.

The row percent of balanced hyperactive agents (group 1) also shows an interesting finding about the path of decline in their life-cycle. The migration rate from group 1 to group 2 is larger than that from group 1 to group 3. The migration rate from group 1 to group 5 is also larger than that to group 6. Hyperactive agents in group 2 and 5 have larger Kin than Kout, while those in group 3 and 6 have larger Kout than Kin. Thus, when hyperactive agents decline in activity, they apparently first reduce their visis to others. Even though they reduce their visiting activities, the visits from others do not decrease immediately. This results in larger proportion of migration from group 1 to groups 2 and 5. This implies that in order to maintain the network vitality, managers may benefit if they can incentivize the visiting activity level of hyperactive agents. In summary, we found a life-cycle of birth, growth, mature, and decline for the hyperactive agents, and it appears important to find ways to sustain the visiting activities of hyperactive agents for the vitality of social networks.

Table 5: Migration patterns among groups (from June 2005 to July 2005)

			Period t+1										
			Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Drop-out	Total (t)
Period t	Group1	No of Nodes	13963	3698	1315	3207	150	85	1390	333	1736	12205	38082
		%	0.17	0.04	0.02	0.04	0	0	0.02	0	0.02	0.15	0.46
		Row %	36.67	9.71	3.45	8.42	0.39	0.22	3.65	0.87	4.56	32.05	
		Column %	30.55	12.07	4.78	1.34	11.68	1.88	0.82	0.3	0.03	1.34	
	Group2	No of Nodes	2271	3922	347	2317	106	12	1140	138	932	11271	22456
		%	0.03	0.05	0	0.03	0	0	0.01	0	0.01	0.13	0.27
		Row %	10.11	17.47	1.55	10.32	0.47	0.05	5.08	0.61	4.15	50.19	
		Column %	4.97	12.8	1.26	0.97	8.26	0.27	0.67	0.13	0.01	1.24	
	Group3	No of Nodes	2133	863	2861	2663	13	108	1221	406	2171	11480	23919
		%	0.03	0.01	0.03	0.03	0	0	0.01	0	0.03	0.14	0.29
		Row %	8.92	3.61	11.96	11.13	0.05	0.45	5.1	1.7	9.08	48	
		Column %	4.67	2.82	10.39	1.11	1.01	2.39	0.72	0.37	0.03	1.26	
	Group4	No of Nodes	4280	3551	3406	81358	105	139	19330	5393	54085	34691	206338
		%	0.05	0.04	0.04	0.97	0	0	0.23	0.06	0.65	0.41	2.47
		Row %	2.07	1.72	1.65	39.43	0.05	0.07	9.37	2.61	26.21	16.81	
		Column %	9.37	11.59	12.37	33.89	8.18	3.08	11.41	4.89	0.79	3.8	
	Group5	No of Nodes	41	72	5	43	135	0	102	2	86	395	881
		%	0	0	0	0	0	0	0	0	0	0	0.01
		Row %	4.65	8.17	0.57	4.88	15.32	0	11.58	0.23	9.76	44.84	
		Column %	0.09	0.23	0.02	0.02	10.51	0	0.06	0	0	0.04	
	Group6	No of Nodes	83	17	144	116	0	1043	39	466	1097	1379	4384
		%	0	0	0	0	0	0.01	0	0.01	0.01	0.02	0.05
		Row %	1.89	0.39	3.28	2.65	0	23.79	0.89	10.63	25.02	31.46	
		Column %	0.18	0.06	0.52	0.05	0	23.11	0.02	0.42	0.02	0.15	
	Group7	No of Nodes	1550	1473	1003	12564	122	34	26825	2096	46521	33387	125575
		%	0.02	0.02	0.01	0.15	0	0	0.32	0.03	0.56	0.4	1.5
		Row %	1.23	1.17	0.8	10.01	0.1	0.03	21.36	1.67	37.05	26.59	
		Column %	3.39	4.81	3.64	5.23	9.5	0.75	15.84	1.9	0.68	3.66	
	Group8	No of Nodes	781	287	1529	10915	2	458	5186	15677	46091	28380	109306
		%	0.01	0	0.02	0.13	0	0.01	0.06	0.19	0.55	0.34	1.31
		Row %	0.71	0.26	1.4	9.99	0	0.42	4.74	14.34	42.17	25.96	
		Column %	1.71	0.94	5.55	4.55	0.16	10.15	3.06	14.2	0.68	3.11	
	Group9	No of Nodes	3763	2233	5959	101253	128	1216	87059	75020	4728150	778985	5783766
		%	0.05	0.03	0.07	1.21	0	0.01	1.04	0.9	56.56	9.32	69.19
		Row %	0.07	0.04	0.1	1.75	0	0.02	1.51	1.3	81.75	13.47	
		Column %	8.23	7.29	21.65	42.18	9.97	26.94	51.4	67.97	69.35	85.4	
	New Entry	No of Nodes	16836	14532	10960	25602	523	1419	27092	10837	1937241	0	2045042
		%	0.2	0.17	0.13	0.31	0.01	0.02	0.32	0.13	23.17	0	24.46
		Row %	0.82	0.71	0.54	1.25	0.03	0.07	1.32	0.53	94.73	0	
		Column %	36.84	47.42	39.81	10.67	40.73	31.44	15.99	9.82	28.41	0	
Total(t+1)			45701	30648	27529	240038	1284	4514	169384	110368	6818110	912173	8359749
%			0.55	0.37	0.33	2.87	0.02	0.05	2.03	1.32	81.56	10.91	100

4.4 Test of the Norm of Reciprocity Assumption

In Hypothesis 3a, we expect the ratio of incoming visits and outgoing visits to be highly correlated. This hypothesis is based on the notion of the norm of reciprocity. The norm of reciprocity is regarded as a general principle of behavior both in a real community and in a virtual social network (Wellman and Gulia 1999, Constant, Sproull and Kiesler 1996). In order to test whether this is true in our dataset, we run a simple regression of Kin on Kout. If the general principle holds in the virtual social network we study, we expect the coefficient of Kout to be close to 1. We run this simple regression on all 11,163,690 members. The regression is as follows:

$Kin_{it} = a_t + \beta_t Kout_{it}$

Kin_{it} (and Kou_{it}) in this equation means the number of other agents who visited the agent *i* in period t (and number of other agents whom agent *i* visited in period t). The regression is for all agents over the entire 20 month period. "Perfect Reciprocity" should yield $\alpha_t = 0$, $\beta_t = 1$ and $R^2 = 1$; "Perfect Non-reciprocity" should yield α_t = average of Kin, $\beta_t = 0$ and R2= 0.

In table 6, we list the coefficients of Kout, R^2, and average of Kin and Kout over 20 months. For the entire period, α is 1.13(vs. 8.17 for average of Kin), β is .853 and R^2 is .853 showing very strong support for the norm of reciprocity in the network. The monthly results similarly show that the norm of reciprocity holds in general. The coefficients are in the range of 0.73-0.88, and R^2 is also quite high (0.71-0.88). Thus, the long-known "norm of reciprocity" is found to be very well followed in this social network.

Next, we tested the hypothesis 3b to see whether balanced hyperactive agents are more abundant than would be expected if outgoing and ingoing visits were uncorrelated. Table 7 contains average balance, actual and expected number (and proportion) of members over 20 months for each group of hyperactive agents. Groups 1, 2, and 3--those groups with low scores of imbalance (which means they are more balanced) show a big difference between actual and expected proportion of members. In particular, the expected proportion of balanced hyperactive agents (group 1) is 0.00% while in reality the proportion is 1.56% consistent with the power law. Group 5 and 6 show large scores of balance (0.5726 and 0.8456).and their actual proportion is relatively close to the expected random proportion and much lower than the power law expectation. This results support the hypothesis that balanced hyperactive agents are far more abundant than would be expected, and consequently confirm that the norm of reciprocity prevails in the group of hyperactive agents. We can also compare the imbalance score of hyperactive agents with that of other agents. The imbalance score of hyperactive agents (imbalance=0.15) turned out to be much lower than the other agents (imbalance=0.53), which supports our contention that hyperactive agents try harder to reciprocate the visits from other agents.

Table 6: Regression analysis of Kin on Kout for all agents

Period (t)	a_t	β_t	R^2	**Average of Kin**	**Average of Kout**
Month1	0.4727	0.7972	0.8165	3.0107	3.1834
Month2	0.6100	0.8375	0.8648	4.5050	4.6507
Month3	0.7097	0.8858	0.8833	6.2491	6.2532
Month4	0.9067	0.8711	0.8687	7.4498	7.5114
Month5	1.1826	0.8573	0.8371	9.1508	9.2949
Month6	1.4881	0.8535	0.8367	10.7635	10.8670
Month7	1.8319	0.8394	0.8267	11.9765	12.0862
Month8	2.1528	0.8201	0.8112	12.0996	12.1291
Month9	2.1525	0.8190	0.8087	12.1917	12.2579
Month10	2.3995	0.8059	0.7965	12.4087	12.4197
Month11	2.4357	0.7967	0.7860	11.7430	11.6823
Month12	2.4879	0.7898	0.7748	11.4830	11.3893
Month13	3.0301	0.7475	0.7382	11.6520	11.5340
Month14	2.7658	0.7466	0.7291	10.5854	10.4731
Month15	2.5907	0.7596	0.7428	10.3687	10.2400
Month16	2.3361	0.7795	0.7525	10.0429	9.8866
Month17	2.5627	0.7507	0.7287	9.8469	9.7029
Month18	2.6632	0.7411	0.7174	9.7864	9.6118
Month19	2.7053	0.7320	0.7108	9.5929	9.4097
Month20	2.4218	0.7401	0.7208	8.0170	7.5596
Entire Period	1.1261	0.8528	0.8526	8.1748	8.1073

Table 7: the imbalance, and actual vs. expected number of membership for each agent group

	Group1	Group2	Group3	Group5	Group6
Balance	0.0669	0.0683	0.1744	0.5726	0.8456
Actual Number and proportion	230,491 1.56%	165,762 1.12%	196,332 1.33%	8,368 0.06%	30,113 0.20%
Expected Number and proportion	0.00001% 0.5	0.00002% 2.0	0.00002% 2.3	0.00814% 909.3	0.01437% 1603.8

4.5 Determinants of Life-span of Hyperactive Agents

In this section, we examine the factors which influence the life-span of hyperactive agents. In the hypothesis development –see Section 3, we classified the factors into three categories—demographic factors, agent related factors and network related factors. Age and gender are included as demographic factors. Balance, activity level, and time to become hyperactive agent are used as agent related factors. Two network related factors are included—activity level of neighbor nodes and clustering coefficient for hyperactive agents.

Regression on life span of hyperactive agents and balanced hyperactive agents

We performed a regression analysis on 468,278 hyperactive and on 230,491 balanced hyperactive agents respectively. As table 8 shows, all the variables included in the analysis came out to be significant even though R^2 is not high (R^2=0.144 for hyperactive agents and R^2=0.143 for balanced hyperactive agents). The influence on the life-span of hyperactive agents is in the order of TBH (standardized β= -0.251, p<0.01) and activity level (standardized β=0.197, p<0.01), followed by clustering coefficient (standardized β= -0.066, p<0.01). For the life-span of balanced hyperactive agents, activity level (standardized β=0.278, p<0.01) and TBH (standardized β= -0.225, p<0.01), followed by balance (standardized β= -0.196, p<0.01) are the top three variables in terms of the influence. Among the factors, gender and age seem to have the least influence on the life-span of both hyperactive agents and balanced hyperactive agents.

Table 8: Regression Analysis on Life-span

	Total Hyperactive agents			**Balanced hyperactive agents**		
Variables	**Coefficients**	**T-Value**	**Standardized *β***	**Coefficients**	**T-Value**	**Standardized *β***
Intercept	1.174**	49.7	0	1.320**	38.53	0
TBH	-0.123**	-179.26	-0.251	-0.106**	-112	-0.225
Activity	0.014**	134.13	0.197	0.017**	107.29	0.278
Imbalance	-0.516**	-33.83	-0.049	-6.154**	-78.04	-0.196
CC	-2.797**	-42.42	-0.066	-5.261**	-49.73	-0.113
Act neigh	0.004**	20.23	0.032	-0.004**	-14.73	-0.035
Gender	0.214**	36.89	0.051	0.110**	13.85	0.027
Age						

13-17	-0.472**	-22.03	-0.041	-0.074*	-2.08	-0.005
18-24	0.447**	29.18	0.100	0.626**	26.8	0.141
25-29	0.379**	24.49	0.078	0.301**	12.71	0.064
R^2	0.144			0.143		
Adjusted R^2	0.144			0.143		
F	8751.47**			4282.55**		

* p<0.05, * * p<0.01

Demographic factors

In hypothesis 2b, we expected that those hyperactive agents in their late teens and early twenties would have a longer lifespan than other hyperactive agents. Our hypothesis was based on the findings that people in late teens and early twenties (ages ranged from 18 to 24) spend the most time socializing with friends, and they are also intense users of on-line social networks. The results support our hypothesis. For both hyperactive and balanced hyperactive agents, the 18-24 age group shows the largest coefficients, which means that those hyperactive agents in the age of 18-24 tend to have longer life-span.

In hypothesis 1b, we expected that female hyperactive agents will have shorter life-span than male hyperactive agents. This hypothesis is rejected because contrary to what we expected, gender (0 for male and 1 for female) shows a positive coefficient for both hyperactive agents (standardized β= 0.051, p<0.01) and balanced hyperactive agents (standardized β= 0.027, p<0.01). According to previous research (Hoffman et al. 1996, Herring 2000), even though males are more intense users of online communication such as discussion groups or chat rooms, females tend to participate more in exchanging private e-mails than male. We believe that participating in virtual social networks may share some common properties with exchanging private e-mail communication. Females seem to be more active than males in building relationship in virtual social networks.

Agent related factors

For agent related factors, the results support our hypotheses. Both TBH and imbalance are significant and are negatively related with life-span (standardized β= -0.251, p<0.01 for TBH, standardized β= -0.049, p<0.01 for balance). Activity is also significant and has positive effects on life span (standardized

β=0.197, $p<0.01$). Thus, the faster an agent becomes a hyperactive agent, the more balanced a hyperactive agent is, and the more active in visits a hyperactive agent is, s/he tends to stay longer as a hyperactive agent. The correlation matrix analysis also shows that TBH is negatively correlated with activity. The results imply that managers need to pay more attention to those hyperactive agents who are more balanced, have shorter TBH and are more active in order to keep the vitality of social networks.

The hypotheses regarding agent related factors are also supported for the balanced hyperactive agents. TBH, activity, and balance have significant effects on life-span of balanced hyperactive agents (standardized β= -0.225, $p<0.01$ for TBH, standardized β= 0.278, $p<0.01$ for activity, standardized β= -0.196, $p<0.01$ for balance). One thing to note is that balance becomes a much more influential factor on life-span in the case of balanced hyperactive agents compared to overall hyperactive agents. This means that among balanced hyperactive agents it is more important to keep the balance between visiting others and being visited by others in order to retain the hyperactive agent status longer. As we discussed in the previous section, the norm of reciprocity is very important in the life-span of the balanced hyperactive agents.

Network related factors

The results in table 8 show that hypothesis 5 is only partially supported. Even though decreasing activity of the neighbor nodes increases the life-span of balanced hyperactive agents (standardized β= -0.035, $p<0.01$), its influence on the life-span of entire hyperactive agents is positive (standardized β= 0.032, $p<0.01$). Thus, the argument made in developing hypothesis 5 that hyperactivity is nurtured by dominance is only supported for the balanced hyperactive agents. For the unbalanced hyperactive agents, it is important to increase the activity level of neighbor agents in order to keep the unbalanced hyperactive agents active.

In hypothesis 6, we predicted that the larger the clustering coefficients are, the longer the life-span. However, this hypothesis is not supported. On the contrary to what we expected, the clustering coefficient is negatively related with life-span (standardized β= -0.066, $p<0.01$). We expect that higher clustering coefficients would generate more dense local networks, and these dense networks would, in turn, nurture the social capital within the local networks around the hyperactive agents. However, it turns out that the neighbors of long-lived hyperactive agents do not interact strongly with each other, so clustering coefficients are low around long-lived hyperactive agents. As a result, low clustering coefficients are associated with long life-span of hyperactive agents.

5. Discussion

This study shows that agents have a broad range of activity described by power laws in activity. Although this is similar to the findings of Price, Barabasi and many others relating to the distribution of the number of links or connections between nodes (or agents for social networks), it is also different because the power law now relates to the activity of the agents in visiting and being visited. We have shown that hyperactive agents are of particular interest in the social network we have studied. We have studied the hyperactive agents further finding characteristics of the agents and factors that affect their longevity as hyperactive agents.

The strongest feature of the hyperactive agents is their balance relative to visiting and being visited by other agents. The number of outgoing visits is highly correlated with the number of incoming visits. Thus, balanced hyperactive agents are far more abundant than would be expected if outgoing and ingoing visits were uncorrelated. Thus, the long-known "norm of reciprocity" is found to be very well followed in this computer-assisted young people social network.

We find in agreement with prior work that younger agents are more prevalent in the hyperactive population than in the overall network (peak hyperactivity is in the 18-24 age range). We find in mild disagreement with prior indications that females are (slightly) more prevalent than males in the hyperactive agent population than they are in the overall network. These results support the concept that the virtual network is quite real to the participants as the age group activity is what is expected in real social interaction. Moreover, women rather than being less social (as computer use might have indicated) are –as in the real world- equally or more active socially.

Our finding that hyperactive agents follow a life-cycle pattern with short life-span (average 2.5 months) is a potentially important observation. This indicates that in this virtual social network, the role of the "hubs" (we call them "hyperactive agents" in this study) is not fixed to specific members. We observe that different members take the role of "hyperactive agents" in various periods of time. This finding also suggests the possibility that virtual social networks which depend heavily on extreme users are unstable and may disappear regardless of their current success. We identify two exceptions to the death of virtual social networks. The first exception occurs if new members can be continuously injected into the networks which our data support for this social network site over our study period. The new members may be younger generations or new members from an untapped market. The second exception occurs if the specific virtual social network itself can evolve in a periodic interval into a new and upgraded version. There are various possibilities for such evolutionary transformation beyond anything we have studied but some of our results are suggestive of some aspects that could be important in such transformation.

We have found that the hyperactive agents have a lifespan and this lifespan is positively increased by lower age, female gender, and by increased balance as expected from the prevalence results. Moreover, we have found that the most prevalent hyperactive agents are balanced and their exit from hyperactivity begins with decreasing visiting activity. These results suggest that management attention and strategies for maintaining vigor might well emphasize younger, more active and more balanced agents and find ways to encourage continued visiting activities.

We also find at least partially for the balanced hyperactive agents that long lifespan is associated with hyperactive agent's dominance of their local network so that long-lived agents have neighbors that tend to be less active than average agents. In addition, the neighbors of long-lived agents do not interact strongly with their other neighbors so clustering coefficients are low around longer-lived hyperactive agents. These results indicate that "social capital" in these young people's virtual social networks is of a clearly different structure than what has usually been found (communities of dense interaction without dominant individuals) for real social networks. This is a dilemma for the managers of virtual social networks since the coherence of local network may work positively for the longevity of the local communities, but it may have negative influence on the longevity of hyperactive agents who are centers of the local communities. We note that low clustering (low local coherence) has recently been found to have what the authors considered a surprising positive influence on the evolution of cooperation in networks (Hanaki et al, 2007).

This finding is intriguingly similar to ours in that we expected clustering around hyperactive agents to increase their lifespan but we found the opposite effect. While the two networks have different "goals", we believe that future research needs to be done on the interaction between the coherence of local networks and the dominance of hyperactive agents in the local networks. Moreover, the different structure of virtual social networks is also indicated by our finding that hyperactive agents often migrate out of the network in a short period. Such results support the concept that the information and communication revolution is changing the behavior as well as the mode of communication.

In conclusion, the contribution of this study is that we show the life-cycle pattern of hyperactive agents (or "hubs") and we identify the factors which influence the life-span of the hyperactive agents. We also found that the balance of visiting and visits from peers is the key ingredient for the growth of the hyperactive agents. However, the overall R2 of the regression analysis with all of the factors studied is fairly low (R2 <.15). Thus, it appears that the large part of the life-span of hyperactive agents may depend on their own needs and personality factors rather than the factors identified here.

References

Albert, R., Barabasi, A. L. 2002. Statistical mechanics of complex networks. *Reviews of modern Physics* **74** 49-97.

Barabasi, A. 2002. *Linked: The New Science of Networks*, Perseus Publishing.

Barabasi, A., E. Bonabeau. 2003. Scale-Free Networks. *Scientific American* **288**(5) 50-59.

Barabasi, A., R Albert. 1996. Emergence of Scaling in Random Networks. *Science* **286** 509-512.

Balasubramanian, S., V. Mahajan. 2001. The Economic Leverage of the Virtual Community. *International Journal of Electronic Commerce* **5**(3) 103-138.

Bebo.com, Facebook.com, Myspace.com. 2006. The Social networking Faceoff. Retrieved Sep 21, 2006, http://www.readwriteweb.com/archives/social_network_faceoff.php

Berners-Lee, T., M. Fischetti, 1999. *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web*. Harper Audio.

Blanchard, A. 2004. Virtual behavior settings: An application of behavior setting theories to virtual communities. *Journal of Computer Mediated Communication* **92**, http://jcmc.indiana.edu/vol9/issue2/blanchard.html.

Blog.sina.com, Qzone.com. 2007. Internet Guide 2007, Retrieved Jan 2007, http://tech.sina.com.cn/socus/2006_GUIDE2007/index.shtml

Borgatti, S.P., R. Cross. 2003. A Relational View of Information Seeking and Learning in Social Networks. *Management Science* **49**(4) 432-445.

Braha, D., Y. Bar-Yam. 2004. Information flow structure in large-scale product development organizational networks. *Journal of Information Technology* **19**(4) 234-244.

Burnett, G. 2000. Information exchange in virtual communities: a typology. *Information Research* **5**(4), http://informationr.net/ir/5-4/paper82.html

BusinessWeek.com. 2005. E-Society: My World Is Cyworld. Retrieved Sep 26, 2005, http://www.businessweek.com/magazine/content/05_39/b3952405.htm

Carr, C. L. 2006. Reciprocity: The Golden Rule of IS-User Service Relationship Quality and Cooperation. *Communications of the ACM* **49**(6) 77-83.

Castells, M. 2001. *The Internet Galaxy: reflections on the internet, business, and society*. Oxford University Press.

Clark, D. 1998. *A Taxonomy of Internet Telephony Applications. Telephony, the Internet, and the Media*. J. MacKie-Mason, D. Waterman, eds. Lawrence Erlbaum Associates.

Constant, D., L. Sproull, S. Kiesler 1996. The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice. *Organizational Science* **7**(2) 119-135.

Dorogovtsev, S. N., J. F. F. Mendes 2002. Evolution of Networks. *Advances in Physics* **51**(4) 1079-1187.

Dorogovtsev, S. N., J. F. F. Mendes 2003. *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford.

Hanaki, N., A. Peterhansl, P. S. Dodds, D. J. Watts. 2007. Cooperation in Evolving Social Networks. *Management Science* **53**(7) 1036-1050.

Herring, S. 2000. Gender differences in CMC: Findings and implications. *Computer Professionals for Social Responsibility Journal* **18** (1). Retrieved January 5, 2005,

http://archive.cpsr.net/publications/newsletters/issues/2000/winter2000/herring.html

Hoffman, D. L. Wm. D. Kalsbeek, T. P. Novak. 1996. Internet and web use in the U.S. *Communications of ACM* **39**(12) 36-46.

Koku, E. F., B. Wellman 2002. *Scholarly Networks as Learning Communities: The Case of TechNet. Designing Virtual Communities in the Service of Learning*. Edited by Barab, Kling, Cambridge University Press.

Kavanaugh, A., S. Patterson 2001. The impact of Community Computer Networks on Social Capital and Community Involvement. *American Behavioral Scientist* **45**(3) 496-509.

Kleinberg, J. M. 1999. Authoritative Sources in a Hyperlinked Environment. *Journal of the ACM* **46**(5) 604-632.

Korgaonkar, P., L. A. Wolin 1999. A multivariate analysis of web usage. *Journal of Advertising research* **39**(2) 53-68.

Mikami, S. 2002. Internet Use, Sociability in Japan. *IT & Society* **1**(1) 242-250.

Miki.com. 2007. Miki. Retrieved July 29, 2007, http://en.wikipedia.org/wiki/Mixi.

Newman, M. E. 2003. The structure and function of complex networks. *SIAM Review* **45** 167-256.

Newman, M. E. 2003. The structure of scientific collaboration networks. *PNAS : Proceedings of the National Academy of Sciences of the United States of America* **98** 404-409.

Newman, M. E. 2005. Power laws, Pareto distributions and Zipf's law. *Comtemporary Physics* **46**(5) 323-351.

Oh, W., S. Jeon. 2007 Memership Herding and Network Stability in the Open Source Community: The Ising Perspective. *Management Science* **53**(7) 1086-1101.

Ono, H., M. Zavodny. 2003. Gender and the Internet. *Social Science Quarterly* 84(1) p111-121.

Patricia, S. V., S. W. Perotti. 2005. Attitude and age differences in online buying. *International Journal of Retail & Distribution Management* **33**(2) 122-132.

Price, D. J. de Solla. 1965. Networks of Scientific Papers. *Science* **149** 510-515

Price, D. J. de Solla. 1976. A general Theory of Bibliometric and other Cumulative Advantage Processes. *Journal of the American Society of Information Science* **27** 292.

Putnam, R. 1995. Turning in, turning out: The strange disappearance of social capital in America. *PS: Political Science & Politics* **28**(4) 664-684.

Resnick, M. 2004. New Styles of Thinking for the Era of the Organic Network. *BT Technology Journal* **22**(4) 113-120.

Reuters.com. 2006. MySpace gains top ranking of US Web sites, Reuters, Retrieved Jul 11, 2006, www.reuters.com

Rogers, E. M. 1995. *Diffusion of Innovations*. New York, Free Press.

Shashaani, L. 1997. Gender Differences in Computer Attitudes and Use among College Students, *Journal of Educational Computing Research* **16** 37-51.

Thorn, B. K., T. Connolly 1987. Discretionary Data Bases: A Theory, Some Experimental Findings. *Communication Research* **14** 512-528.

Watts, D. J. 2004. The new science of networks. *Annual Review of Sociology* **30** 243-270.

Watts, D. J., Steven H. Strogatz 1998. Collective dynamics of ‘small-world’ networks. *Nature* **393** 440-442.

Wellman, B. 2001. Computer Networks as Social Networks. *Science* **293** 2031-2034.

Wellman, B., J. Salaff, D. Dimitrova, L.Garton, M. Gulia, C. Haythornthwaite 1996. Computer Networks as Social Networks: Collaborative Work, Telework, and Virtual Community. *Annual Review of Sociology* **22** 213-238.

Wellman, B., M. Gulia 1999. *Net Surfers Don’t Ride Alone: Virtual Communities as Communities*. Networks in the Global Village, edited by Barry Wellman, Boulder, CO, Westview.

Willams, R. L., J. Cothrel. 2000. Four Smart Ways to Run Online Communities. *Sloan Management Review* Summer **41**(4) 81-91.

Zhang, Y. 2002. Comparison of Internet Attitudes between Industrial Employees and College Students. *Cyber Psychology & Behavior* **5**(2) 143-149.