A User Segmentation Model for Social Networking Websites (SNSs)

De-Kui, Li[*] *and Bhubate Samutachak*[**]

Abstract

Usage of the Internet has been largely limited to passive reception of available content until the advent of large social networking websites, e.g. Facebook, Twitter, MySpace, etc. Through these websites online users create contents and build relationships with each other, amplifying their dependence on and the power of the online network. While the online social network is gaining momentum in online world with magnifying user size, inadequate attention has been given to behavioral analysis of online social network members, particularly concerning their contributions to the networks. This paper uses Kozinets' model of Internet user segmentation to categorize users into 4 groups: collectors, onlookers, VIP users, and joiners. Two differentiating factors, i.e. contributions and social capital are introduced. Unlike most studies on online networking behaviors, this study uses the real behavioral data (as opposed to surveyed, subjective data) of Hao Kan Pu, a Chinese-based social network website. It was found that idle members, those having low contributions and low connections (social capital) dominate the network, accounting for 98.47% of the samples. Active members, those having high contributions and vast connections, accounted for only 0.21%. High-contribution members with low connections (collectors), and low-contribution members with high connections, accounted for 0.64% and 0.67%, respectively. These structural data regarding the behavior of online networks are crucial for online social network analysis, e-marketers, and website management.

Keywords: *Social Network Websites, Internet Users, Segmentation Model, Social Capital, Social Network Analysis, NodeXL*

[*] Doctoral student, Kwangwoon University, 447-1, Wolye-Dong, Nowon Gu, SOUTH KOREA 139-701

[**] Assistant professor, Dhurakij Pundit University, 110/1-4 Prachachuen Rd., Laksi, Bangkok 10210, THAILAND

ตัวแบบการแบ่งส่วนผู้ใช้งานเว็บไซต์สังคมออนไลน์

เตอ คุย ลี[*] และ ภูเบศร์ สมุทรจักร[**]

บทคัดย่อ

ในอดีต บทบาทของผู้ใช้งานอินเตอร์เนตมักจะเป็นผู้คอยบริโภคเนื้อหา โดยไม่ค่อยได้มีส่วนร่วมในการสร้างเนื้อหาด้วยตนเองจนกระทั่งเมื่อมีการสร้างเว็บไซต์สังคมออนไลน์ เช่น เฟสบุ๊ค ทวิตเตอร์ มายเสปซ ฯลฯ ซึ่งเว็บไซต์สังคมออนไลน์เหล่านี้เปิดโอกาสให้สมาชิกสามารถสร้างเนื้อหา และเป็นเวทีสำหรับการสร้างคามสัมพันธ์กับสมาชิกคนอื่นๆ และในขณะที่เว็บไซต์สังคมออนไลน์กำลังเป็นที่นิยมและมีบทบาทเพิ่มขึ้นเป็นอย่างมากในสังคม โดยดูจากจำนวนของสมาชิกที่เพิ่มขึ้นอย่างมาก และรวดเร็ว ความพยายามในการทำความเข้าใจกับพฤติกรรมในโลกออนไลน์ของสมาชิกในเว็บไซต์สังคมออนไลน์กลับยังไม่มากและลึกพอ ในงานวิจัยฉบับนี้ ผู้วิจัยใช้ตัวแบบการแบ่งส่วนผู้ใช้อินเตอร์เนตของ Kozinets ซึ่งแบ่งกลุ่มผู้ใช้ออกเป็น 4 กลุ่มได้แก่ collectors, onlookers, VIP users, และ joiners โดยใช้ปัจจัยสำคัญสองประการในการแบ่งกลุ่มได้แก่ ปริมาณเนื้อหาที่สมาชิกสร้างขึ้น (contributions) และทุนทางสังคม (social capital) ของสมาชิก ซึ่งดูจากระดับความกว้างขวางของสมาชิกคนหนึ่งๆ ในสังคมออนไลน์ โดยข้อมูลที่นำมาใช้ในการศึกษาครั้งนี้เป็นข้อมูลพฤติกรรมจริงในสังคมออนไลน์ของสมาชิกเว็บไซต์ เห่าค่านปู้ (Hao Kan Pu) ซึ่งเป็นเว็บไซต์สังคมออนไลน์สัญชาติจีน แตกต่างจากการศึกษาอื่นๆ ที่มักใช้วิธีการสำรวจและเก็บข้อมูลจิตวิสัย (Subjective) ผลการวิจัยพบว่าสมาชิกที่สร้างเนื้อหาน้อย และไม่ค่อยเชื่อมต่อกับสมาชิกคนอื่นมีสัดส่วนมากถึงร้อยละ 98.47 ของกลุ่มตัวอย่าง ในขณะที่สมาชิกที่สร้างเนื้อหาปริมาณมาก และมีการเชื่อมต่อกับสมาชิกคนอื่นมาก กลับมีเพียงร้อยละ 0.21 และสมาชิกที่สร้างเนื้อหาปริมาณมาก แต่เชื่อมต่อกับสมาชิกคนอื่นค่อนข้างน้อยมีสัดส่วนร้อยละ 0.64 ส่วนสมาชิกที่สร้างเนื้อหาปริมาณค่อนข้างน้อยแต่เชื่อมต่อกับสมาชิกคนอื่นค่อนข้างมาก มีสัดส่วนร้อยละ 0.67 ข้อค้นพบดังกล่าวเป็นประโยชน์อย่างยิ่งต่อการประเมินอำนาจและความสำคัญของเว็บไซต์สังคมออนไลน์ นักการตลาดออนไลน์ และผู้บริหารเว็บไซต์สังคมออนไลน์

คำสำคัญ : *เว็บไซต์สังคมออนไลน์ ผู้ใช้งานอินเตอร์เนต ตัวแบบการแบ่งส่วนตลาด ทุนทางสังคม การวิเคราะห์เครือข่ายทางสังคม NodeXL*

[*] นักศึกษาระดับปริญญาเอก คณะบริหารธุรกิจ มหาวิทยาลัยควางอูน สาธารณรัฐเกาหลี
เลขที่ 447-1 เขต วอลเย ถนน โนวอน สาธารณรัฐเกาหลี (เกาหลีใต้) 139-701

[**] ผู้ช่วยศาสตราจารย์ สาขาวิชาธุรกิจระหว่างประเทศ คณะบริหารธุรกิจ มหาวิทยาลัยธุรกิจบัณฑิตย์
เลขที่ 110/1-4 ถนนประชาชื่น เขตหลักสี่ กรุงเทพฯ 10210

Introduction

In today's competitive online social networking websites (SNSs) environment, the ability to identify main users in order to build their long-term loyalty and steadily expand existing relationships is a key competitive factor for social networking website (Greenberg, 2010). The advent and prevalence of social networking websites, e.g. Facebook, Twitter, MySpace, has given Internet users a virtual platform to demonstrate their influence and interdependence on others, thus giving rise to word-of-mouth (or viral) marketing as a crucial strategy in marketing domains. Through these SNSs, word-of-mouth can spread more easily, quickly, and vastly. The essence of word-of-mouth marketing is to reach out to a broad set of potential customers and attract considerable attention via social interactions. Unlike direct and mass marketing, which only recognize the intrinsic value of a customer, word-of-mouth marketing additionally exploits the network effect of the customer by taking network value into consideration to measure the customer's real value (Richardson & Domingos, 2002).

A typical social networking website combines text (basic content), images or video (multimedia) and a variety of links (network-based linkage). In this paper, these factors are categorized into two dimensions: contribution-based and social capital-based factors. Both contribution-based and social capital-based sources should be considered in order to develop a more comprehensive and robust model of influence and to estimate the precise value of marketing influence. It is therefore sensible that marketing analyses may benefit from these properties of an individual as a criteria of market segmentation in addition to the traditional demographic, psychological, behavioral properties, among others.

As a result, the purpose of this paper is to develop a segmentation model of users that contribute content and contact other users frequently on social networking websites. In this paper, users are modeled as a graph-based representation, which is formed by nodes. Nodes stand for social networking websites with some social characteristics or behaviors and have social influence value. The most important result of this model is to identify nodes that have core value, growing value, and general value. While most models of this kind may rest their analysis on surveyed and subjective data, this model uses real users' data of a network derived from the NodeXL software.

The paper is structured as follows: First, we examine existing user segmentations for online communities. Second, we proceed to identify the most useful segmentation models and update them to applications for uses with social networking websites. Last, we use the users' data from a real social networking website to validate our segmentation model.

Theoretical Background

Contributions in Social Media

Previous studies of participation in online communities have focused on two types of social system: information commons, where many individuals contribute to the construction of a small number of shared artifacts, and online discussion groups, where individuals exchange messages on a given topic (Moira, Cameron & Thomas, 2009). Content contribution can be described through the interaction between the following elements: the people that are involved, the content and the artifacts they produce and share or the engagement between people around content, and distribution, i.e. the way in which people discover and consume content.

Previous contributions of knowledge to electronic networks of practice seem to be paradoxical. Research has argued, for example, that giving away knowledge eventually causes the possessor to lose his or her unique value relative to what others know and benefits all others except the contributor (Thorn & Connolly, 1987).

For a social network, the success of the system is tied to the amount of contributions any member's social contacts produce, and on outcome, which is dependent on the eventual participation of a large portion of the user base.

Social Capital Theory

Social capital is typically defined as "resources embedded in a social structure that are accessed and/or mobilized in purposive action" (Lin, 2001). However, what exactly is social capital? A thorough review of the literature on the subject shows that researchers have defined the construct in terms of social networks, trust, civic engagement, life satisfaction, and a variety of other concepts (Lin, 2001; Putnam & Tuning, 1995). The core idea of social capital, however, is straightforward: it is the resources available to people through their social interactions (Putnam, 2004). Individuals with a large and diverse network of contacts are thought to have more social capital than individuals with small, less diverse networks. Although people often accumulate social capital as a result of their daily interactions with friends, coworkers, and strangers, it is also possible to make conscious investments in social interaction (Resnick, 2002; Wusch-Vincent, 2007). This is what transpires when people report that their main reason for joining Facebook, for example, and spending time on the site is to keep in touch with old friends and to strengthen bonds with colleagues. Using SNSs, individuals seek to maintain and expand their social networks (Ellison, Steinfield, and Lampe, 2007; Joinson, 2008).

The key difference between social capital and other forms of capital is that social capital is embedded in the social realm. While other forms of capital are based on assets or individuals, social capital resides in the fabric of relationships between individuals and their connections to their communities (Putnam, 2004).

The literature suggests that social capital is difficult to develop or to be transferred to electronic networks of practice because social capital is more likely to emerge in collectives characterized by a shared history, high interdependence, frequent interaction, and closed structures. It has also been argued that electronic networks cannot support significant knowledge outcomes because knowledge is often tacit and highly embedded, requiring high-band-width communication that is difficult to sustain through technology (Brown & Duguid, 2001).

Extending this rationale to SNSs, we could conclude that their impact on social capital should be contingent upon the specific uses and gratifications sought by users. Using Putnam's (1995) concepts of bridging, weak-tie social capital (i.e., across diverse social groups) versus bonding, strong-tie social capital (i.e., across homogeneous groups), Williams (2006) noted that the type of relationships within social networks can predict different kinds of social capital. Weak-tie networks produce bridging social capital because they connect people from different life situations. These networks broaden the set of information and opportunities for users in the network. However, individuals in weak-tie relationships do not gain the benefits of bonding social capital, such as the emotional support that occurs based on the interdependence and commonalities of strong-tie networks. As we shall see, the features of Facebook allow for the production and maintenance of both strong ties and weak ties and, by extension, can influence users' lives positively.

User Segmentation

The idea of segmenting a homogeneous market into sub-segments is intertwined with the idea of tailoring products to fit each sub-segment, i.e. product differentiation. The ideas originated from the classical work of Wendell R. Smith (1956). Thereafter, researchers developed a variety of segmentation approaches such as benefit segmentation, life style segmentation, psychographic segmentation, etc.

A primary concern of all marketers has been the identification of their users. Likewise, the flourish and fall of social networking websites hinge on the clear knowledge of their users' identity which immensely affects the interconnectedness of the members thus commanding the marketing power of the network. Word-of-mouth, which can be transmitted only through a social network, has become a powerful marketing instrument. People are likely to be affected by the decisions of their friends and colleagues (Kempe, Kleinberg, and Tardos, 2003), and some researches have investigated the diffusion process of word-of-mouth and viral marketing effects in the success of new products (Domingos & Richardson, 2001; Richardson & Domingos, 2002). Zhan et al. (2009) emphasize the important role of writing and referring product reviews on the Internet (such as blogosphere or online communities). In the case of the methodologies to implement opinion mining, many scholars focus on the identification of the author's attitude, for example whether it is positive or negative (Putnam & Tuning, 1995). Motivated by applications to marketing, several probabilistic models have been proposed for choosing customers with a large overall effect on the social network (Domingos & Richardson, 2001; Richardson & Domingos, 2002). Notably, the problem is addressed that we hope to market a new product/service via the adoption of the power of word-of-mouth in the network. Or should we start by initially targeting a few influential members as the virus of the network to diffuse and propagate the information even recommendations to their friends? (Kempe, Kleinberg, and Tardos, 2003). But how should we choose the influential seeds with the strongest virulence? Or which characteristics should be taken into consideration to identify the influential nodes?

Recently, there have been several attempts to segment the market by employing the online networking properties of the individual. Tobey (2010) has conduced a study on Cyworld, a famous Korean social networking websites (Tobey, 2010). The data were collected from 50,000 sample users and analyzed using ABC segmentation methodology (Active users, buyers, and Connectors). The study's results of the study suggest a new business model called Cyworld 2.0. Another research was conducted on a blogosphere (Li, 2009), which created a marketing influence value model in order to identify influential users of the network for marketing use. In addition, for the specific purpose of online communities of consumption Kozinets (1999) has defined a typology of user segmentation. Kozinets describes two central measures that affect the "formation of lasting identification" with the community: the intensity of the relationship within the community and the centrality of the consumption of the user. Figure 1 shows four segments of users in a matrix, with the intensity of the relationship and the centrality of consumption being the two axes.

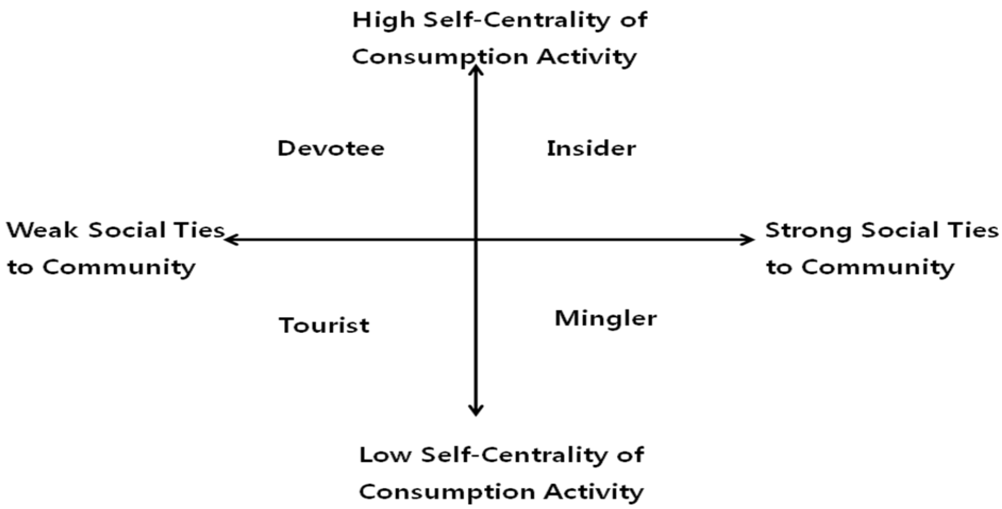

Figure 1: **User typology for communities of consumption**

SNSs User Segmentation Model

This paper uses Kozinets' model [Figure 2] of market segmentation to differentiate the Internet users according to their behaviors regarding SNSs. The model categorizes users' behaviors, based on their contributions and social capital, into 4 groups, i.e. collectors, VIP users, onlookers, and joiners. According to Kozinets, contribution represents the contents of various kinds produced by consumers. Users are different in terms of the quantity of the content they produce, which make each of them different in commercial and social value. These contents include the quantity of topics, comments, photos, as well as messages. With respect to social capital, this research uses the concept of social network analysis which categorizes social members according to their social properties, i.e. centrality degree, centrality betweenness, and closeness centrality. The centrality degree of a member is a count of the number of other members that they are connected to. Centrality betweenness is the number of shortest paths from all members from all others that pass through that member. Therefore, betweenness centrality is a useful measurement of member's connectivity. Closeness centrality is the distance of a member to all other members in the network measured by the average shortest path length. In a digital network it indicates how fast or how efficiently a member can access the network and how likely it is that information will reach that member (Trier, 2008).

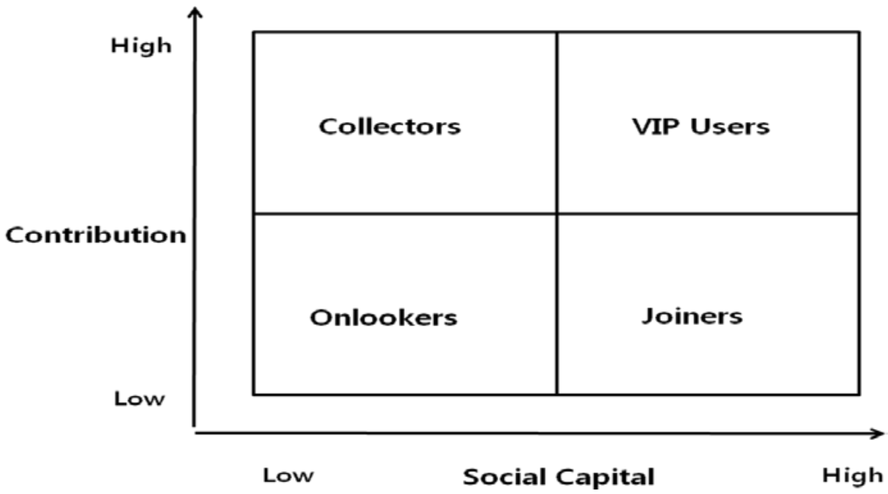

Figure 2: User segmentation model for SNSs

Contribution Factor

It is common in social networking websites that provide functions for members to post photos and comments, or send messages to friends. These activities contributes contents that can be measured according to the number of subjects, the number of comments, and the size of the subject, comment or message. In this paper, subject refers to the issue or topic of conversation that members initiate and post on the SNSs. Members can creates topics using scripts or photos to declare the state of their minds (diary), whereas, comments are responses of members to those subjects. In addition, members may personally send messages to other members in order to avoid public acknowledgement.

In this research, quantification of contribution was performed by counting the number and size of subjects, comments, and messages produced by each member. The greater the number and size of these items, the greater is the contributions of the member to the network. Moreover, mutual response among members through posting comments and sending messages implies the strength of their relationships.

Social Capital Factor

With respect to social network theory, their popularity and importance to the society of members comprise their social capital (Hatala, 2006). Freeman (1997) suggests, in his research on the network's communication activity, degree-based and betweenness-based measurements for social capital operationalization. Although social network theory suggests another item of measurement, closeness centrality, it is not used in this paper as it tends to be close to zero in the large network. Centrality degree is the number of links attached to a node. In SNSs, centrality degree is measured by counting the number of friends of each member. While centrality degree signifies the popularity and importance of a member, centrality betweenness measure its bridging power. The member that has high centrality betweenness links a great number of members together. Without high betweenness-centrality members, the network is not vastly connected.

Data Collection and Analysis

While most of the analysis of the network of this kind rely on surveyed data, this research was able to access, with the permission of the founder, the data structure of the Hao Kan Pu, a China-based online social network website found in 2007 with over 178,000 registered members. Similar to other SNSs, this website allows member to conduct social networking activities such as photos and diaries (status of topics) with posting as well as messaging. Details of these activities, such as the size of the information posted, frequency, the ID numbers of members that initiate any kind of social networking activities were recorded in the database of the website.

The researchers contacted the management of Hao Kan Pu in order to request the details on the members' social networking activities. The data were collected from May 1 to November 29, 2010.

Figure 3 depicts the process of user segmentation. First, the data on the members' contributions, i.e. number and size of (subject) diary, comment, and photos and messages, were extracted from the website's database. Information on the source and destination of the comment and messages were also collected. Then the researchers use NodeXL, a social network analysis software, to calculate centrality degree and centrality betweenness. NodeXL is a free, open-source template for Excel 2007 and 2010 that provides a sociometry analysis of a social network with graphic presentation. The software includes powerful automated features, while allowing for manual control of individual vertex placement, labeling, color properties, and the like. Calculation of the contribution and social capital value of each member as performed as follows;

Contribution value = (diary length/number of diaries) + (message length/number of messages) + (comment length/number of comments)

Capital value = Centrality betweenness + Centrality degree

Note: the length of each activity was measured in kilobyte

It should be noted that no previous literature provides methods of contribution value or social capital value calculations. The researchers therefore invented the above calculating methods. After deriving the necessary data on the contribution and social capital, the data were organized and analyzed using Tableau 6.0.

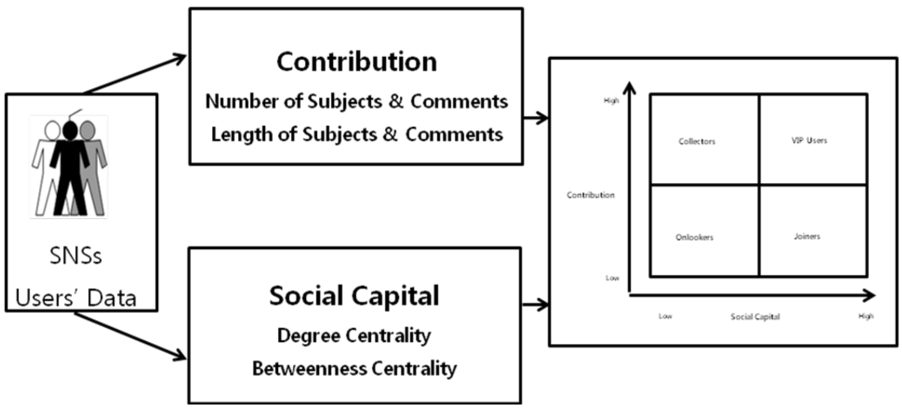

Figure 3: User segmentation process

Results and Discussion

Collaborating with the management of the website, the researchers were able to extract data on the members' activities from the network's database. Figure 4 shows a sociogram of the comments that members make on photos posted. Loops simply mean that a member commented on his/her photos. During the observed period, there were 613,566 photos posted on Haw Kan Pu. These photos attracted 165,228 comments from other members. On the other hand, figure 5 is a sociogram of the comments on diaries, where 2,535 comments were made on 9,021 diaries posted. The obvious difference on the thickness and density of the two sociograms implies that posting and comment on photos were more favored than on diaries.

Figure 4: Network map for comments on photos

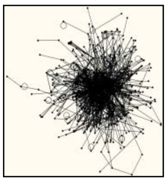

Figure 5: Network map for comments on diaries

The number of messaging activities, however, was by far greater than the previous two activities. During the same period of time, users sent 340,437 messages and received 45,131,401 replies. Their magnitude made it impossible for the software to produce a sociogram.

Assigning the level of contribution on the vertical axis and the level of social capital on the horizontal axis, a panel of user segmentation was formed. A horizontal line at the average contribution of the sample and a vertical line at the average social capital were drawn to demarcate 4 quadrants of user behaviors, i.e. collectors, VIP users, onlookers, and joiners. Figure 6 exhibits how the behavioral data of users fits into each quadrant. From an interview with the management of the network it was found that the network's staff input a certain amount of contributions in order to promote the liveliness of the network. The researchers hence deleted the contributions from the network' staff, resulting in Figure 7. It is interesting to note that most of the members of the network are onlookers. These members, accounting for 98.47% of the samples, produce relatively low contributions and have relatively low social capital. Their role in the network is quite passive. Many members of this category do not produce any content and have no or very few friends. On the other hand, those that contribute relatively actively and possess high social capital, the VIP users, were very few. Most of the members in this quadrant flock were along the high social-capital zone and made an average number of contributions. It should be mentioned that members of this kind make the network lively and sustainable. As for the upper left quadrant, i.e. collector, members of this group are content producers with relatively limited connections. The network is a performing stage for them where they can express themselves. However, the limited number of connections may signify that their content is irrelevant, thus not being followed by a great number of other members. Users of this sort accounted for 0.64% of the samples. The lower right quadrant shows the members that had relatively high social capital but produced relatively low contributions. These types of members, accounting for 0.67% of the samples, have a great amount of connections and are very important for information dissemination. However, they are not very active in producing content. It is this type of member about whom the network should acquire knowledge. The essential task of the network is to craft a strategy to encourage this type of member to make greater contributions and to become a VIP user.

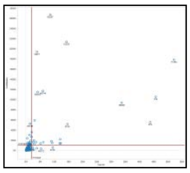

Figure 6: Full view of the H social networking website

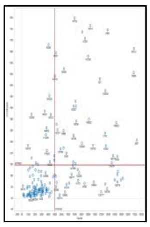

Figure 7: The corrected view of the H social networking website

Conclusion

This study offer a user segmentation model for the online network. Content contribution and social capital, borrowed from the social capital theory, were used as joint differentiating factors in order to place network members into 4 categories, i.e. collectors, VIP users, onlookers, and joiners. While other researches investigating the networks' members behaviors have relied mainly on surveyed data, this paper used the real online behavioral data of an SNS and fit it into a segmentation model, revealing the actual situation of the online network.

The segmentation of SNS users reveals an alarming finding that the majority of network members are idle and poorly connected. This large number of users use up the resources of the network and make less than a moderate contribution. The smallest group of users was the group whose members were well connected and actively contributed to other members. The results of this study give a valid warning to both e-marketers and network management. Most of the SNSs boast about their networking power, referring to the size of members. However, one should keep in mind the large proportion of idle users. As a matter of fact, according to social capital theory, what dynamizes a network are the resources embedded in a social structure. If the social structure is not well connected, in other words, most of its members stay idle, these resources will not be transmitted or shared, thus making the network less popular. Efficient SNSs should be able to report the size of each user category so that the networking power of the network can be correctly understood. In addition, SNS management should craft strategies specifically for users of each category in order to promote interconnectedness and activeness, which will help sustain and expand the network.

References

Brown, J. S., Duguid, P. (2001). Knowledge and Organization: A Soda I-Practice Perspective, *Organization Science*, pp. 198-213

David Tobey, Cyworld: User Segmentation and Targeting, 2010.

Domingos, P. and Richardson, M. (2001). Mining the network value of customers, *ACM SIGKDD international conference on Knowledge discovery and data mining*

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The Benefits of Facebook "friends:" Social capital and college students use of online social network sites, *Journal of Computer-Mediated Communication*, 12, 1143—1168

Freeman, Linton C. (1977). A Set of Measures of Centrality Based on Betweenness, *Sociometry*, pp. 35-41,.

Hatala, J.P. (2006). "Social Network Analysis in Human Resource Development: A New Methodology." *Human Resource Development Review* 5(1): 27.

Joinson, A. N. (2008). Looking at, looking up or keeping up with people?: Motives and use of Facebook, *SIGCHI conference on Human Factors in Computing Systems, Florence, Italy*

Kempe, D., Kleinberg, J. and Tardos, E. (2003). Maximizing the spread of influence through a social network, A*CM SIGKDD international conference on Knowledge discovery and data mining*

Kozinets, R.V. (1999). E-tribalized marketing?: the strategic implications of virtual communities of consumption, *European Management Journal*, pp. 252-264

Lin, N. (2001). *Social Capital*, Cambridge University Press, Cambridge, U.K.

Moira, B., Cameron, M., Thomas, L. (2009). *Feed me: motivating newcomer contribution in social network sites*, pp.945—954,

Paul Greenberg (2010). *CRM at the Speed of Light*, Fourth Edition, Mc Graw Hill

Putnam, R, Tuning (1995). in. *Tuning Out: The Strange Disappearance of Social Capital in America, Political Science and Politics*, pp. 664-683,.

Putnam, R.D. (2004). *Bowling together*. OECD Observer

Resnick, P. (2002). Beyond bowling together: Socio-technical capital, In J. M. Carroll (Ed.)., *Human-computer interaction in the new millennium*, pp. 247—272

Richardson, M. and Domingos, P. (2002). Mining knowledge-sharing sites for viral marketing, *ACM SIGKDD international conference on Knowledge discovery and data mining*

Thorn, B.K., d Connolly, T. (1987). Discretionary Data Bases. A Theory and Some Experimental Findings, Communication Research, pp.512-528

Trier, M. (2008). Towards Dynamic Visualization for Understanding Evolution of Digital Communication Networks, *Information Systems Research* 19(3): 17.

Wendell R. Smith, (1956). Product Differentiation and Market Segmentation as Alternative Marketing Strategies, *The Journal of Marketing*

Williams, D. (2006). On and off the net: Scales for social capital in an online era, Journal of Computer-Mediated Communication, 11, pp.593—628

Wunsch-Vincent, S. (2007) Participative web and user-created content: Web 2.0, OECD

Li, Yung-Ming (2009). Identifying Bloggers with Marketing Influence in the Blogosphere, *ICE*

Zhan, J., Loh, H.T. and Liu, Y. (2009). Gather customer concerns from online product reviews - A text summarization approach, Expert System and Applications 36, 2, 2107