Inferring Leadership of Online Development Community using Topological Structure of its Social Network

Sho TUGAWA[1]・Hiroyuki OHSAKI[2]・Makoto IMASE[3]
[1,2,3]Graduate School of Information Science and Technology, Osaka University
[1]Graduate Student[i]・[2]Associate Professor・[3]Professor

［Abstract］

To facilitate an online community and lead it to success, community administrators, managers, and facilitators must have adequate knowledge of the current state of the community. For online development and socialization communities, in particular, recognizing the leadership levels is valuable for administrators, since the leadership level in an online community is known to strongly affect the success of the online community. Therefore, in this paper, we propose a method to infer the strength of leadership in an online community as a whole by using betweenness centrality, which is obtained from a social network. We also examine the effectiveness of the proposed method by utilizing log data of the activities in SourceForge, a major forum for online development communities. Consequently, we show that (1) the leadership level of a participant can be accurately inferred from his/her betweenness centrality; (2) the proposed index called LSI (Leadership Strength Index), which estimates the level of leadership in an online community, has sufficient positive correlation (correlation coefficient ~0.3) with the software maturity and productivity of a medium community (i.e., community where the number of participants is 50-150), and is applicable for comparative purposes; and (3) an eight month observation of a community's communication history is adequate for inferring the leadership of a medium community.

［Key word］
Social Network, Centrality, Leadership

1. Introduction

In recent years, social activities have rapidly shifted into networked environments. As a result, several types of online communities have been formed, such as socialization communities (e.g., Facebook for networking with others), knowledge-sharing communities (e.g., Wikipedia for Web-based knowledge sharing), and development communities (e.g., SourceForge for open-source software development) [1].

In this paper, we focus on the online development community.

To facilitate an online community and lead it to success, community administrators, managers, and facilitators must have adequate knowledge of the current state of the community [2, 3]. For an online community to be successful, community administrators are required to understand the participants' needs and to take appropriate action in order to satisfy those needs [3]. Furthermore, success is dependent on community administrators frequently monitoring the progress of a task toward completion [2].

Since *leadership*, *information quality*, *system quality*, and *pro-sharing norms* are reported to be major factors in the success of an online community [4–6], knowing the current state of these factors is useful for community administrators. Thus, online community administrators can take appropriate action if they know the leadership levels of participants, the quality of the information exchanged, the quality of the service system, and the norm of participants in the community.

For online development and socialization communities, in particular, recognizing the leadership levels is valuable for administrators, since the leadership level in an online community is known to strongly affect the success of the online community [4, 5]. Hence, as a first step in inferring the current state of an online community, we concentrate on estimating the level of leadership in an online development community. There exist several leadership roles and definitions of leadership. In this paper, as the leadership roles, we focus on leaders' activities such as monitoring progress and supporting participants for problem solving since such activities are found to be important for community success [5]. Our definition of leadership of a participant follows Osborn's [25]:"Leadership is not only incremental influence of a boss toward subordinates, but most important it is the collective incremental influence of leaders in and around the system." We therefore define the leadership level in an online community as the collective amount of influence from who play central role

[i] s-tugawa@ist.osaka-u.ac.jp

(i.e., leaders) to the other participants (i.e., followers).

In the literature, by studying the topology of social networks in both online and offline communities, several techniques have been proposed to discover the leaders in a group [7, 8], as well as to infer the influence of one participant on others [9–11]. For instance, Shetty *et al*. reported that leaders can be discovered based on graph entropy, which is obtained from a social network's topology [7]. In contrast, to infer the influence of one participant on others, Freeman proposed three indices called degree centrality, closeness centrality, and betweenness centrality that measure structural centrality in a social network [9].

Significant positive correlation has been reported to exist between the level of leadership of a participant and his/her betweenness centrality [12–14]. Thus, we expect that the leadership of participants in an online development community can be inferred from betweenness centrality. In those studies [12–14] the relation between leadership levels and centrality measures have been investigated under different environments by using a number of methods.

However, betweenness centrality and other conventional techniques cannot be applied directly to infer the state of an online community itself, since these techniques focus on the influence of an individual on the other participants in a community. Namely, they cannot be simply applied for comparing leadership of different online communities. The purpose of existing techniques is to discover participants who play a central role or to infer the influence of a participant on other participants in a community.

Therefore, in this paper, we propose a method to infer the strength of leadership in an online community as a whole by using betweenness centrality in order to compare the leadership among different online communities. Specifically, we propose using the leadership strength index (LSI), which estimates the level of leadership in an online community and is applicable to performing a comparison of leadership in different online communities. LSI is obtained through three steps: (1) build a social network from the communication history among online community participants, (2) calculate betweenness centrality of each participant from the topological structure of the obtained network, and (3) derive graph centrality of the whole social network based on betweenness centrality values found for each participant.

We examine the effectiveness of the proposed method by utilizing log data of the activities in SourceForge [15], a major forum for online development communities. First, we investigate the relation between betweenness centrality of a participant obtained from the generated social network and the participant's role (i.e., administrator, developer, or user) in 507 SourceForge communities. This investigation then reveals the effectiveness of betweenness centrality for inferring a participant's leadership level in each development community. Second, to show the value of our method, we study the correlation between the strength of leadership in a community inferred by the proposed method and several standard indices for measuring a community's success. Consequently, we discuss the implications of using LSI to infer the current state of an online community.

The organization of the remainder of this paper is as follows. In Section 2, we present our method for inferring the strength of leadership in an online development community. In Section 3, we examine the effectiveness of our proposed method by utilizing log data of the activities in SourceForge. Finally, Section 4 concludes this paper and discusses future work.

2. Method for Inferring Leadership in an Online Community

To estimate the strength of leadership in an online community, we propose using LSI obtained through the following three steps: (1) build a social network from the communication history among online community participants, (2) calculate betweenness centrality of each participant from the topological structure of the obtained social network, and (3) derive graph centrality of the whole social network based on betweenness centrality values found for each participant.

(1) Build the Social Network

First, a social network is built from the communication history among online community participants. Since participants in an online community communicate with each other over a network using electronic tools, such as character-based or voice-based chat applications and bulletin boards, a social network representing the social ties among participants can be generated from the history of these communications.

A social network can be represented as a weighted undirected graph (see Fig. 1), where a vertex in the graph denotes a participant, an edge denotes a social tie between participants, and the weight on an edge denotes the strength of the social tie. The social network is thus a graphical representation of the social relations among participants in the online community. We deliberately use an undirected social network, since such networks have commonly been used in social network analysis and analytical tools for social network, such as centrality measures (e.g., betweenness, closeness, and degree centralities), are defined for undirected social networks [9, 16].

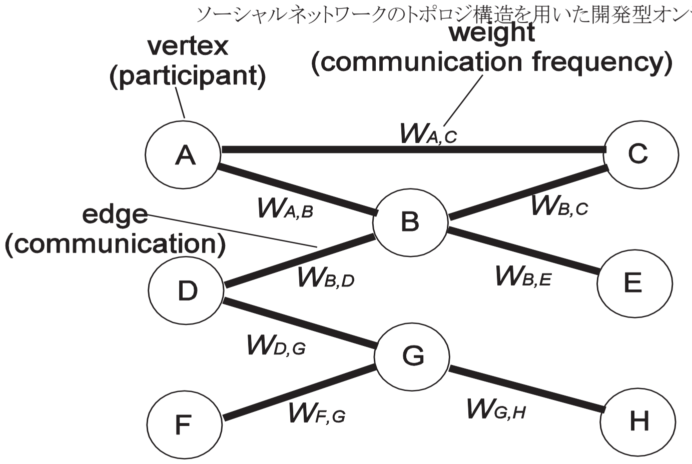


Figure 1: A social network (weighted undirected graph) built from the communication history among participants in an online community

An edge (i, j) is generated when communication exists between participants i and j in observation period T. The weight $w_{i,j}$ on the edge (i, j) is then defined as communication frequency. Hence, for a given observation period T, we can build a social network from the communication history in that period.

(2) Calculate Betweenness Centrality of Each Participant

Next, to infer the leadership level of each participant, each participant's betweenness centrality is calculated from the topological structure of the social network. Our method infers the strength of leadership of an entire online community based on the leadership of each participant.

Betweenness centrality measures the centrality of a vertex in a graph through an index that represents the proportion of shortest paths between all other vertices passing through the target vertex [9]. Specifically, betweenness centrality of vertex k in graph $G = (V, E)$ is defined as

$$C_B(k)=\frac{2\sum_i \sum_{i<j} b_{i,j}(k)}{n^2-3n+2}, \tag{1}$$

where $b_{i,j}(k)$ is the proportion of shortest paths from vertex i to vertex j that pass through k, and n is the number of vertices in G. In this paper, since a social network is described as a weighted undirected graph, the distance between i and j is defined as $d_{i,j} = 1/w_{i,j}$ [17].

(3) Infer the Leadership in the Community Using Graph Centrality

Finally, LSI is obtained as the graph centrality of the whole social network, which is calculated based on each participant's betweenness centrality. Whereas betweenness centrality is a measure of the centrality of a *vertex*, graph centrality measures the centralization of a *network*. We therefore infer the strength of leadership in a community from an index for measuring the centralization of the community's social network.

Several definitions of *graph centrality* exist, such as those based on the distance between vertices [18, 19] and those based on each individual's centrality (i.e., degree, closeness centrality, or betweenness centrality) [9]. Distance-based graph centrality attains a large value when the distances between vertices are short. In contrast, graph centrality defined based on each vertex's centrality attains a large value when the relative differences between the centrality of the most central vertex and that of all other vertices are large.

In this paper, we use graph centrality based on betweenness centrality [20]. It is defined as the sum of the relative differences between the betweenness centrality of the vertex having the largest betweenness centrality and those of other vertices normalized by the total number of vertices in the network. Hence, graph centrality based on betweenness centrality of a graph $G = (V, E)$ is

$$C_B=\frac{\sum_i C_B(p^*)-C_B(i)}{n-1}, \tag{2}$$

where p^* is the vertex with largest betweenness centrality, and n is the number of vertices in G. Note that graph centrality based on betweenness centrality is applicable for comparing online communities with different numbers of participants, since this measure of graph centrality is independent of the size (number of vertices) of a graph [20].

Our proposed method is thus based on the notion that if a participant who plays leadership role (i.e., he/she has high betweenness centrality) is at the center of a social network, then the leadership in that community is strong. Namely, we assume that the level of leadership in a community is high when a participant who influences others plays a central role in the community.

3. Experiment

3.1. Method

We examine the effectiveness of the proposed method for inferring the strength of leadership in an online community by utilizing log data of the activities in SourceForge [15], a major forum for online development communities. Since a number of statistics on the activities in SourceForge communities (known as "projects") are publicly available, we can examine the efficacy of our method using these data. For our experiments, we use data from the SourceForge Research Data Archive [21], which is a public archive of *SourceForge.net* maintained by the University of Notre Dame.

We first examine the effectiveness of using betweenness centrality to infer the leadership levels of participants in SourceForge communities by investigating the relation between a participant's betweenness centrality obtained from a generated social network and his/her position (i.e., administrator, developer, or user) in the community. In a SourceForge community, participants are either registered or unregistered to the community, and three types of participant positions exist [22]. A limited number of registered participants will be classified as *administrators* with administrative privileges, who have the authority to release software, recruit developers, and create/delete bulletin boards, as well act as developers. Other registered participants are *developers* who contribute to software development and can update software repositories. Finally, unregistered participants are *users* who can merely exchange messages on bulletin boards. SourceForge has no explicit rules or guidelines for the functions of administrators, developers, and users [15], and so the style of community management typically differs by community. In general, however, administrators are most likely to play a leadership role in a community. Conversely, developers will only sometimes play a leadership role, whereas users are unlikely to have such a role. We therefore examine the effectiveness of betweenness centrality for inferring leadership levels of participants through analyzing the relation between their positions in the community and their betweenness centrality scores obtained from a generated social network.

Secondly, we examine the applicability of the proposed LSI for comparing the strength of leadership in different SourceForge communities. To this end, we explore the correlation between the leadership levels of SourceForge communities inferred by LSI and several existing indices for measuring the success of online communities. To the best of our knowledge, an objective index does not currently exist for measuring the strength of leadership in online development communities; however, strong leadership has been reported to be one of major factors of a successful community [5]. Although measuring the leadership in an online development community is possible through a subjective evaluation determined by participant questionnaires or interviews, performing reproducible and large-scale experiments using subjective evaluation is non-trivial. We therefore compare the leadership levels of different SourceForge communities inferred using LSI, and investigate the correlation between these levels and the success of the communities.

Table 1: Means and standard deviations of the numbers of administrators, developers, users, and messages, and the success indices two years after each community's creation

	average	standard deviation
number of administrators	1.5	0.9
number of developers	5.6	7.7
number of users	59	106.6
number of messages	568.5	1264
software status	4.2	1.1
number of releases	19	30.1
number of downloads	84458.2	399366.8
number of bug reports	77.4	363.8
mean bug fixing time [days]	71.2	139.4

To evaluate their success, we use a number of indices that measure the activity, quality, and quantity of a community's open source software: the number of *developers*, software status, the number of software releases (i.e., initial and version upgrades), the number of software downloads, the number of bug reports posted in a bug tracker system, and the mean time taken to fix a bug after a report has been posted [23]. Software status is expressed in six stages— planning, pre-alpha, alpha, beta, stable, and mature—and in this experiment, we use an integer score of one to six, respectively, to denote each stage. If multiple software statuses are assigned to an open source software, we take the average of those scores.

We infer the strength of leadership in SourceForge communities created between February 2003 and June 2008 from their communication history during the first T months after community creation, and obtain their success indices after two years. We examine 507 communities in which a minimum number of communications had been performed (i.e., at least 100 messages had been exchanged) in SourceForge. In the following experiments, which monitor long-term communication, the observation period T is 24 months unless explicitly stated. Namely, we build a social network from communication history for the 24 months after community creation. We also obtain the aforementioned success indices relating to the end of this 24 month period.

Since participants in SourceForge communities usually communicate with each other by exchanging messages on bulletin boards, we determine the relations among participants from the structure of message threads on those boards [16, 24]. We identify that communication from participant i to j occurred when participant i replied to a message posted by j. We ignore the original posts, since ascertaining to whom the original posts were sent is generally difficult. To exclude participants who did not directly contribute to development, we removed bulletin boards named "Help" and anonymous bulletin boards.

Means and standard deviations of the numbers of administrators, developers, users, and messages, and the success indices two years after each of the 507 communities' creations are shown in Tab. 1.

3.2. Result: Relation between Betweenness Centrality and Participant's Position

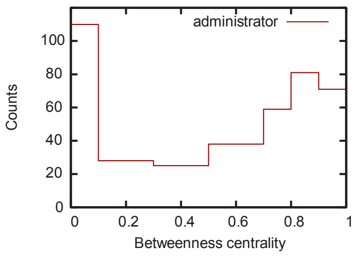
First, we classified 29,889 participants in the 507 SourceForge communities into three categories based on their positions in the communities (i.e., administrator, developer, or user) [22]. Histograms showing betweenness centrality distributions for the participants in each category are given in Fig. 2, and betweenness centrality means and standard deviations for each category are listed in Tab. 2. The numbers of participants (samples) placed into each category are also shown in the table.

From these results, the leadership level of a participant can clearly be inferred from his/her betweenness centrality. The betweenness centralities of *administrators*, who generally play leadership roles, are large, whereas those of *user*, who generally do not play leadership roles, are small in comparison. While the standard deviation of the *administrators'* betweenness centrality is relatively large at 0.34, that of *users* is much smaller at only 0.038. Although the number of participants in each category is considerably different, the betweenness centrality scores of most users are low (typically less than 0.1). This result suggests that we can distinguish users who have low leadership from others positions in development communities by using betweenness centrality.

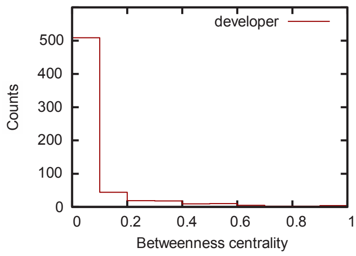
3.3. Result: Relation between LSI and Success Indices of Communities

Next, we calculate the rank correlation coefficients between the leadership levels in the 507 SourceForge communities inferred by using our proposed method and the indices for measuring the success of these communities (Fig. 3). Here, nonparametric rank correlation coefficients are employed since several indices, such as the inferred leadership of a community and the number of downloads, have biased distributions. For comparative purposes, the rank correlation coefficients between each success index and the number of participants who have exchanged any message on bulletin boards, the ratio of *administrators* (i.e., number of *administrators*/number of participants), the ratio of *developers* (i.e., number of *developers*/number of participants), the number of messages on bulletin boards, and the number of message per participant (i.e., number of messages/number of participant) are also shown.

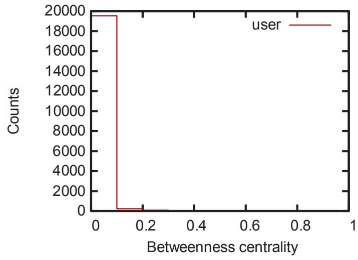
From the results in Fig. 3, we see that LSI has a significant positive correlation with the maturity of the software (i.e., the software status) and the productivity (i.e., the number of releases). Positive correlations with statistical significance of less than 1% exist between LSI and the software maturity and LSI and the productivity. In contrast, LSI is negatively correlated or has a statistically insignificant correlation with the activity (i.e., number of developers, number of bug reports, and mean bug fixing time) and popularity (i.e., number of downloads) measures. One of the causes of the low correlation between LSI and popularity and activity measures is that these measures are greatly dependent on the number of users of the software. Hence, leadership is considered to have little influence on the popularity and activity.



(a) administrator



(b) developer



(c) user

Figure 2: Betweenness centrality distributions for all participants in each position category

Table 2: Means and standard deviations of betweenness centrality for participants in each position category

	number of samples	average	standard deviation
administrator	737	0.51	0.34
developer	868	0.07	0.15
user	28,284	0.008	0.038

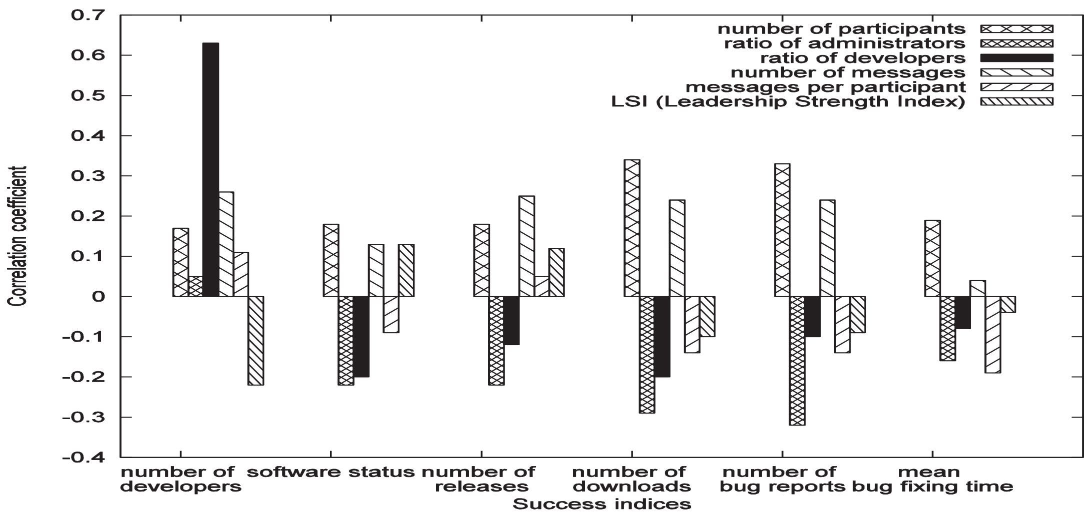


Figure 3: Rank correlation coefficients between indices measuring the states and success of the 507 communities

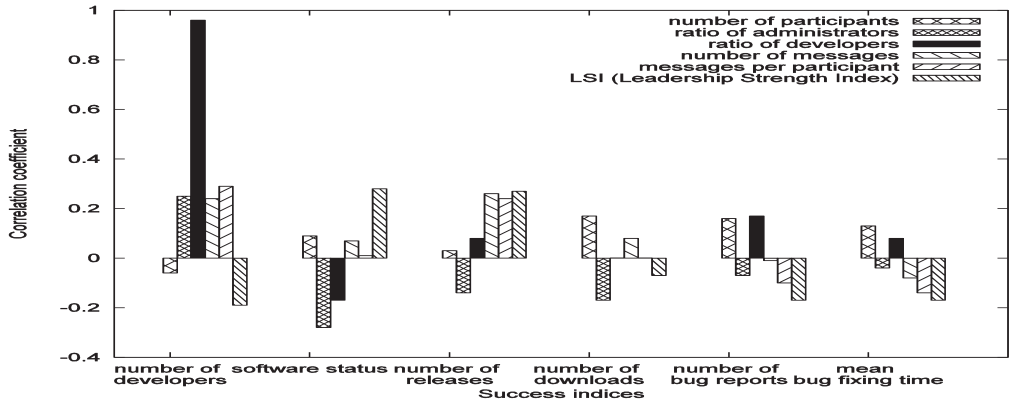


Figure 4: Rank correlation coefficients between indices measuring the states and success of 119 communities having 50–150 participants

Note that the number of participants and the number of messages have a stronger positive correlation with the success indices than LSI. This high correlation is primarily due to the size of a community strongly affecting its success. Namely, communities with a large number of participants or with participants that actively communicate with each other have a natural tendency to be successful compared with other communities. Therefore, effects resulting from the size of a community must be eliminated when analyzing the correlations between LSI and the success indices.

Following the above experiment, we thus calculate the rank correlation coefficients between LSI and success indices of only those communities (119 in total out of 507) having 50–150 participants (Fig. 4).

The number of participants, the ratio of *administrators*, the ratio of *developers*, the number of messages, and messages per participant are not significantly correlated to software status. Only LSI has a sufficient positive correlation with the software maturity. In addition to LSI, the number of messages and the messages per participant have significant positive correlation with the productivity. However, as a result of eliminating community size effects, other than LSI, the correlations among status and success indices are weakened. Conversely, the correlations between LSI and software maturity and between LSI and productivity are strengthened. Note that the values of correlation coefficients related to the activities of leaders in an online community and the community's success are approximately 0.3–0.6 [5]. We therefore conclude that LSI has sufficient positive correlations with software maturity and productivity.

These observations suggest that the proposed LSI is applicable for comparing the strength of leadership in different online communities. Our experiments reveal that LSI is positively correlated to two of the success indices, suggesting that we can analyze the success of online communities through a comparison of their leadership levels by using LSI.

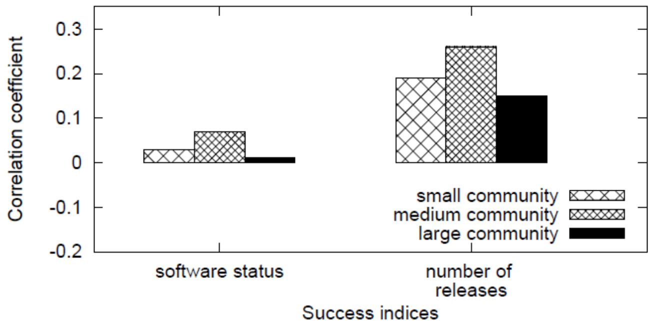
3.4. Result: Effect of the Size of Community

We now examine precisely how the correlations between LSI and the success indices change with the size of the community. In Section 3.3, we found that the correlations when all 507 communities were considered were different from those when only 119 communities in a certain size range were included in the analysis. From this result, we expect that correlations between LSI and the success indices change with the size of the community.

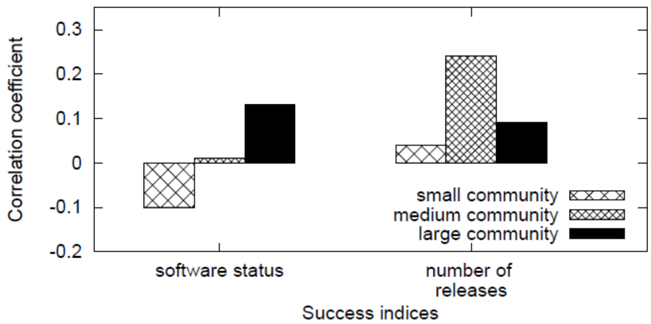
We classify the 507 communities into three classes based on their number of participants: small communities (with $\leq$ 49 participants), medium communities (with 50–150 participants), and large communities (with $\leq$151 participants).

The rank correlation coefficients for these communities calculated between each of the status and the success indices are shown in Fig. 5, where only those combinations with significant positive correlations in Section 3.3 are presented. Namely, we use the number of messages, messages per participant, and LSI as indices representing the state of a community, and the software maturity and productivity as indices measuring the community's success.

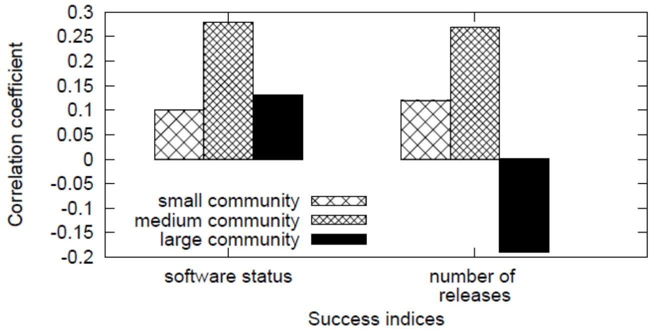
The positive correlations between LSI and each of the software maturity and productivity in small communities are weaker than those of medium communities. Moreover, for large communities, LSI is negatively correlated to productivity, whereas LSI is positively correlated to software maturity. Thus, the correlations between LSI and the success indices change with the size of the community, with the strongest positive correlations found for medium communities. Hence, LSI is considered to be particularly effective for inferring the state of medium-sized communities. The styles of the software development are generally different in different size communities. We therefore expect that the development style affects the effectiveness of the LSI to infer the current state of the online development communities. For instance, in small-sized communities, software development could be completed without any leaders. In large-sized communities, the leadership styles could be different from those in medium-sized communities.



(a) number of messages



(b) messages per participant



(c) LSI

Figure 5: Rank correlation coefficients between state and success indices for each class of community (small, medium, and large)

3.5. Result: Effect of Observation Period

Finally, we investigate the relation between LSI scores and the observation period of communication history utilized for building the social networks.

As we discussed in Section 3.1, an observation period T of 24 months was used in the above experiments. Specifically, we built each social network using the communication history of the 24 months following a community's creation, and analyzed the correlations between LSI and the success indices at the end of this period.

In Section 2, we stated that LSI is one of the indices for measuring the state of a community. In contrast, success indices such as software maturity and productivity are determined by the accumulation of the activities of a community. Namely, whereas the level of leadership in a community represents the goodness of a short-term state of the community, the success indices represent the goodness of long-term activities of the community.

However, studies have shown that the communication structure in a community does not change frequently [16]. Therefore, we expect that the strength of leadership in a community can be inferred to an extent, even if the observation period of the communication history is not that long.

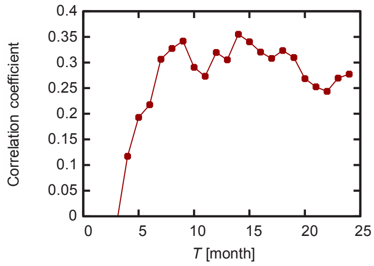


Figure 6: Rank correlation coefficients between LSI and software maturity for 119 medium-sized communities (with 50–150 participants) when changing the period T of observing communication history

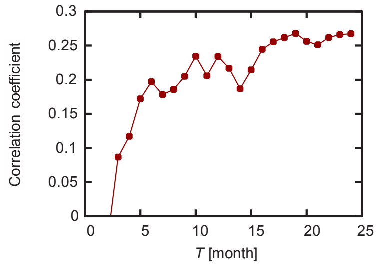


Figure 7: Rank correlation coefficients between LSI and productivity for 119 medium-sized communities (with 50–150 participants) when changing the period T of observing communication history

Rank correlation coefficients between LSI and each of the software maturity and productivity for changing T between one and 24 months are shown in Figs. 6 and 7, respectively. To eliminate community size effects, medium communities (with 50–150 participants) from Section. 3.4 were again selected as sample communities.

We see that LSI has significant positive correlations with both software maturity and productivity when T is greater than eight months. These results suggest that eight months need to pass from a community's creation in order to infer the strength of leadership in the community. Conversely, the results also suggest that an observation period exceeding one year contributes little to improving the accuracy of inferred leadership levels.

Although we cannot generalize based on only our experiments, the results indicate that the state of a community (in particular, the strength of its leadership) can be accurately estimated eight months after its creation.

Hence, we conclude that (1) the leadership level of a participant can be accurately inferred from his/her betweenness centrality; (2) the proposed LSI has sufficient positive correlation (correlation coefficient ~ 0.3) with the software maturity and productivity of a medium community, and is applicable for comparative purposes; and (3) an eight month observation of a community's communication history is adequate for inferring the leadership of a medium community.

4. Conclusion and Future Works

In this paper, we have proposed a method that uses LSI to infer *the strength of leadership in an online community as a whole* in order to compare the leadership among different online communities. We examined the effectiveness of our proposed method by utilizing log data of the activities in SourceForge, a major forum for online development communities. Consequently, we have shown that (1) the leadership level of a participant is accurately inferred from his/her betweenness centrality; (2) LSI has sufficient positive correlation (correlation coefficient ~0.3) with indices measuring the software maturity and the productivity of the medium community, and LSI is applicable for comparative purposes; and (3) an eight month observation of an online community's communication history is adequate for inferring the leadership of the medium community.

As future work, we plan to investigate the properties of LSI when applied as an index for not only development but also other types of online communities. As we have discussed in Section 1, leadership is also a major factor in the success of online socialization communities. Hence, the leadership of such communities will be explored using LSI.

Acknowledgement

The authors would like to thank Prof. Masayuki Murata for his kind support and valuable discussions. This research was partly supported by "Global COE (Centers of Excellence) Program" of the Ministry of Education, Culture, Sports, Science and Technology, Japan and Grant-in-Aid for Scientific Research (B) (21300022).

[References]

[1] D. Hinds and R. M. Lee, "Social network structure as a critical success condition for virtual communities," in *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS'08)*, Jan. 2008, pp. 323–323.

[2] B. S. Bell and S. W. J. Kozlowski, "A typology of virtual teams: Implications for effective leadership," *Group & Organization Management*, vol. 27, no. 1, pp. 14–49, Mar. 2002.

[3] R. L. Williams and J. Cothrel, "Four smart ways to run online communities," *SLOAN MANAGEMENT REVIEW*, vol. 41, no. 4, pp. 81–91, Jul. 2000.

[4] J. Koh and Y.-G. Kim, "Sense of virtual community: A conceptual framework and empirical validation," *International Journal of Electronic Commerce*, vol. 8, no. 2, pp. 75–94, 2003.

[5] A. Bourhis, L. Dubé, and R. Jacob, "The success of virtual communities of practice: The leadership factor," *Electronic Journal of Knowledge Management*, vol. 3, no. 1, pp. 23–34, Jul. 2005.

[6] H. Lin, W. Fan, L. Wallace, and Z. Zhang, "An empirical study of web-based knowledge community success," in *Proceedings of the 40th Annual Hawaii International Conference on System Sciences(HICSS'07)*, Jan. 2007, p. 178c.

[7] J. Shetty and J. Adibi, "Discovering important nodes through graph entropy the case of Enron email database," in *Proceedings of the 3rd international workshop on Link discovery*, Aug. 2005, pp. 74–81.

[8] A. Goyal, F. Bonchi, and L. V. S. Lakshmanan, "Discovering leaders from community actions," in *Proceeding of the 17th ACM Conference on Information and knowledge management (CIKM'08)*, Oct. 2008, pp. 499–508.

[9] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social networks*, vol. 1, no. 3, pp. 215–239, 1979.

[10] P. Bonacich, "Power and centrality: A family of measures," *The American Journal of Sociology*, vol. 92, no. 5, pp. 1170–1182, Mar. 1987.

[11] V. Latora and M. Marchiori, "A measure of centrality based on network efficiency," *New Journal of Physics*, vol. 9, no. 6, p. 188, Jun. 2007.

[12] L. C. Freeman, D. Roeder, and R. R. Mulholland, "Centrality in Social Networks:II. Experimental Results," *Social Networks*, vol. 2, no. 2, pp. 119–141, 1979.

[13] B. Mullen, C. Johnson, and E. Salas, "Effects of communication network structure: Components of positional centrality," *Social Networks*, vol. 13, no. 2, pp. 169–185, Jun. 1991.

[14] D. J. Brass, "Being in the right place: A structural analysis of individual influence in an organization," *Administrative Science Quarterly*, vol. 29, no. 4, pp. 518–539, Dec. 1984.

[15] "SourceForge," http://www.sourceforge.net/.

[16] K. Goh, Y. Eom, H. Jeong, B. Kahng, and D. Kim, "Structure and evolution of online social relationships: Heterogeneity in unrestricted discussions," *Physical Review E*, vol. 73, no. 6, p. 66123, Jun. 2006.

[17] M. E. J. Newman, "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality," *Physical review E*, vol. 64, no. 1, p. 16132, Jun. 2001.

[18] A. Bavelas, "Communication patterns in task-oriented groups," *Acoustical Society of America Journal*, vol. 22, pp. 725–730, Nov. 1950.

[19] C. Flament, *Applications of graph theory to group structure*. Englewood Cliffs, New Jersey:PrenticeHall, 1963.

[20] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, Mar. 1977.

[21] G. Madey and Others, "The SourceForge Research Data Archive (SRDA)," http://zerlot.cse.nd.edu/, 2010.

[22] S. Krishnamurthy, "Cave or community ? an empirical examination of 100 mature open source projects," *First Monday*, vol. 7, no. 6, pp. 20–56, 2002.

[23] K. Crowston, H. Annabi, and J. Howison, "Defining open source software project success," in *Proceedings of the 24th International Conference on Information Systems (ICIS 2003)*, Dec. 2003, pp. 327–340.

[24] V. Gómez, A. Kaltenbrunner, and V. López, "Statistical analysis of the social network and discussion threads in Slashdot," in *Proceedings of the 17th international conference on World Wide Web*, Apr. 2008, pp. 645–654.

[25] R . Osborn, "Toward a contextual theory of leadership," *The Leadership Quarterly*, Vol.13, No.6, pp. 797–837, 2002.

（2012年4月29日受理）