2010 10th IEEE International Conference on Advanced Learning Technologies

A Social Network Analysis Perspective on Student Interaction Within the Twitter Microblogging Environment

Karen Stepanyan
Buckinghamshire New University
High Wycombe, United Kingdom
karen.stepanyan@bucks.ac.uk

Kerstin Borau, Carsten Ullrich
Shanghai Jiaotong University
Shanghai, China
kerstinborau@mail.onlinesjtu.com,
ullrich_c@sjtu.edu.cn

***Abstract* — This paper summarises the analyses of participant interaction within the Twitter microblogging environment. The study employs longitudinal probabilistic social network analysis (SNA) techniques to identify the patterns and trends of network dynamics. It explores the associations of student achievement records with the observed network. The results indicate tendencies towards: [i] reciprocal interaction; and [ii] adoption of a selective approach in communication over time, implying that students tend to communicate with fewer peers over time. The evaluations that examine achievement score attributes indicate [iii] network homogeneity and popularity effects associated to achievement scores – suggesting greater interaction among students of similar levels and more attention to higher achieving students.**

Keywords - microblogging; social network analysis; social networking; collaborative learning;

I. Introduction and Background

Twitter is a microblogging service that allows users to communicate with others by posting brief messages (so called updates) that describe their current status. Most recently, microblogging platforms acquired attention of educational practitioners and researchers [1]. The growing popularity and widening acceptance of the services compel educators to explore the possibilities of using the tools for educational purposes. This paper summarises an empirical study that evaluates the use of Twitter as part of a foreign language learning course. It particularly, analyses the interaction of participants (both learners and teachers) by using SNA techniques.

The recent learning literature acknowledges the central role of interaction and participation in a learning process. Interaction and participation is considered an essential and indivisible component of student engagement [2, 3]. Despite the growing body of e-learning research, the studies of online interaction are often incomprehensive due to limitation of the employed research methods and the complexity of the field in general. This study attempts to address this gap.

This paper extends the earlier conducted study [1] that analysed the benefits of using Twitter for second language learning. The rationale for this research is to identify interaction patterns that can inform educational designers, practitioners and technologists. This research looks at the dynamics of participant interaction within a microblogging environment and tries to inform educational practitioners about possible predictive measures that can be anticipated when adopting similar learning designs and technologies.

II. Description of the Study

The study was performed at the distant college of Shanghai Jiaotong University, China (Online-SJTU). The class sizes at the Online-SJTU vary from 80 to 120 students; thus not everybody has the chance to communicate with the instructor during class and practice his or her English skills. We therefore investigate using technology to provide additional communication opportunities.

In this study, we used the microblogging tool Twitter in an English course for native speakers of Chinese in the following way: the instructor created a new, personal Twitter account. The students of this class were prompted to create their own account and to become “friends” with the instructor’s account as well as with the accounts of the other students. Since each Twitter user receives the messages of his or her friends, each student who followed the instruction would receive the messages of his/her fellow students and of the instructor. The students were then told to post at least seven microblogging messages a week and to read the incoming messages of their fellow students. In order to increase the incentive to use Twitter, participation in Twitter contributed to the students’ final grade.

III. Methodology and Research Questions

SNA is a technique that allows analysis of human interaction and relationships between individuals, groups and communities. It can be employed for the studies of participant interaction, access as well as analysis of group and community development [4, 5]. The application of SNA can shed light on the hidden factors that may affect student participation, open collaboration and personal development. Thus, the use of SNA in educational research can become a valuable and a fundamental resource for understanding student interaction and participation, subsequently leading to improvement of teaching techniques and tools [6, 7]. The research methodology adopted in this study employs application of probabilistic longitudinal SNA research techniques that allow identification of network dynamics and trends, and permit reporting with statistical precision..

Analysing student interaction this study attempts to address the following questions: [1] What are the prevalent interaction patterns, measures of community development and measures of interaction dynamics within the studied

978-0-7695-4055-9/10 $26.00 © 2010 IEEE
DOI 10.1109/ICALT.2010.27

IEEE computer society

environment?; [2] Which network measures and trends are associated to participant scores?

The data used in this study constitutes the messages posted by the participants within the Twitter microblogging environment. Students were able to communicate with one another by using the communication conventions widely used within the selected microblogging environment. The messages posted by one participant (actor i) to another (actor j) are defined as a directed tie (from i to j; i→j) only when there are more than three messages posted in total. The observed interaction was then combined into three equally timed groups (also referred as waves) of weighted and directed social network data. The interaction was observed throughout a period of 56 days. Out of total 5256 messages, posted by 108 participants, 1266 directed messages and 87 interacting participants were considered in the study.

IV. PROBABILISTIC ANALYSIS

This study employs stochastic models for capturing regularities with statistical accuracy. This section conjecture and test a set of hypothesis and reports the results to address the research questions. The study was conducted on four different levels: actor, dyadic, triadic and global.

TABLE I. RESEARCH HYPOTHESES, PARAMETERS AND CONDITIONS

Question	Hypotheses	Null Hypotheses and Conditions
Research Question 1	H1: The rate of network development stabilizes over time. (Global Level)	$H1_0$: *Rate1<=Rate2* at $\alpha<0.05$[a]
	H2: There is a tendency of exhibiting selective approach in choosing an interaction partner. (Actor Level)	$H2_0$: *Outdegree-Density>0;* at $\alpha<0.05$
	H3: There is a tendency of reciprocation in the network (*i→j and j→i*). (Dyadic Level)	$H3_0$: *Reciprocity <= 0;* at $\alpha<0.05$
	H4: There is a tendency of network closure (i.e. increasing transitivity and reducing distance between actors). (*i→j, j→k and i→k*). (Triadic Level)	$H4_0$: [a] *Transitivity <= 0;* [b] *Distance Two <= 0;* at $\alpha<0.05$
Research Question 2	H5: (Homophily effect) Participant with similar achievement records (score) tend to interact among themselves. (Actor Level)	$H5_0$: [a] *Same Score <= 0;*[b]*Similar Score<=0;*at $\alpha<0.05$
	H6: (Indegree popularity effect) Participant achievement is not related to his/her acquired attention. (Actor Level)	$H6_0$: *Score Covariate-alter = 0;* at $\alpha<0.05$
	H7: (Outdegree popularity effect) Participant achievement is not related to his/her outreach. (Actor Level)	$H7_0$: *Score Covariate-ego = 0;* at $\alpha<0.05$

To evaluate the dynamics of the developed network, a set of concepts for addressing issues related to the formation and evolution of social networks have been selected, namely *homophily (actor level), reciprocity (dyadic level) and transitivity (triadic level).* Hence a set of hypotheses (TABLE I.) were conjectured and tested as part of this study. The probabilistic analysis was performed by employing dynamic actor-driven models defined and evaluated with SIENA (v. 3.17) software jointly with the StOCNET graphical interface package [8, 9]. Based on the conjectured hypotheses a number of models were defined.

A. *Structural Effects of Network Dynamics*

Models 1, 2 and 3 are compiled to show the structural network dynamics effects, such as outdegree, reciprocity, transitivity and distance at two - indicatives of network closure effect [8] (i.e. tendency for a friend-of-a-friend to become a friend). TABLE II. summarises the models and values of the estimated parameters.

TABLE II. ESTIMATION RESULTS: TESTING FOR STRUCTURAL EFFECTS

Structural Effects	Network Dynamics Parameter	Model 1	Model 2	Model 3
H2&H3: Outdegree and Reciprocity effects	Outdegree (density)	-2.30 (0.05)*	-5.72 (0.25)*	-2.6 (0.07)*
	Reciprocity	2.11 (0.1)*	5.10 (0.45)*	2.16 (0.11)*
H4: Network Closure	Transitivity	-	0.59 (0.29)	-
	Distance at Two	-	-	0.29 (0.05)*
H1: Network Development Rate	Rate 1	28.34 (6.01)*	15.12 (4.97)*	33.8 (9.84)*
	Rate 2	9.68 (0.93)*	2.84 (0.73)*	9.39 (1.01)*

*- indicates statistically significant effect at $\alpha < 0.05$.

The Rate parameter indicates a frequency at which network changes are estimated to occur; i.e. the participant may obtain, maintain or cease communication with participants affecting the change of network structure between waves [9]. Hence, the considerably greater value of Rate 1 in relation to the value of Rate 2 suggests a tendency of stabilising network dynamics that leads to observation of less sporadic communication. This pattern is consistent and statistically significant in all the three models – leading to the rejection of $H1_0$ in support of the H1.

Model 1 was introduced for testing the effects of outdegree density and reciprocity in the network. The results of estimation indicate negative and statistically significant results for the outdegree (-2.30) effect. This pattern indicates a tendency for exhibiting a selective approach in reaching out other participants rather than a tendency of decreasing outdegree density [10]. Similar to the earlier discussed parameter of rate, the negative effect of outdegree density indicates the stabilizing tendency within the network. Based on the consistent (Models 1, 2 and 3) observation of negative outdegree density parameter values the $H2_0$ can be rejected.

The estimation values for the parameter of reciprocity are positive (2.11) and statistically significant in Model 1. The results are consistent in Models2&3 and, therefore, $H3_0$ can be rejected. This result can be interpreted as a tendency towards reciprocating the incoming ties (i.e. the participants tend to respond to the initiated communication).

The test of H4 involves consideration of two mutually inverse parameters: transitivity and distance at two. Both of the parameters are indicative of network closure effect [8]; an effect that suggest a tendency of link formation between three participants of which two are already connected (for instance, friend of a friend tends to become a friend). The transitivity parameter (Model 2) estimation result (0.59) is not statistically significant. While, the estimation of the distance at two (Model 3), indicates a positive (0.29) and statistically significant value. Therefore, $H4_0$ can be rejected, only on the basis of the second [b] condition. Hence, the results do not reveal a network closure effect.

B. Homophily and Popularity Effects

The concept of homophily associates certain network structures with the similar actor attributes within the specified network. In line with the proverb - 'birds of a feather flock together' - homophily effect is present when contact between similar actors occurs more frequently than among dissimilar actors [11]. This paper conjectures and tests the homophily effects based on participant score.

In addition to the homophily effect, this study explores the effect of participant score on the 'popularity' of the actor. The popularity, is defined by the in/out degree centrality values, in other words the number of incoming or outgoing ties to a specified actor [12]. The rationale of these tests is to identify whether the number of incoming or outgoing ties are in any way related to participant score. The Models 4, 5 and 6 are compiled to test the possible homophily and in/outdegree popularity effects.

TABLE III. ESTIMATION RESULTS: TESTING FOR SCORE EFFECTS

Net. Effects	Network Dynamics	Model 4	Model 5	Model 6
Outdegree and Reciprocity effects	Outdegree (density)	-5.29 (0.23)*	-5.66 (0.31)*	-5.7 (0.28)*
	Reciprocity	4.85 (0.39)*	4.39 (0.55)*	4.4 (0.41)*
H5: Homophily	Score Similarity	2.70 (0.92)*	-	-0.51 (0.85)
	Same Score	-0.56 (0.32)	-	-
H6: Poularity	Score Ego (Outdegree)	-	0.61 (0.15)*	0.51 (0.18)*
H7: Poularity	Score Alter (Indegree)	-	0.46 (0.16)*	0.66 (0.16)*

*- indicates statistically significant effect at $\alpha < 0.05$.

The estimation results (Model 4) indicate positive and statistically significant homophily effect (2.70) for the Score Similarity covariate. Hence, $H5_0$ can be rejected on the first [a] condition, which conjectured the lack of homophily effect among participants with similar scores. On the other hand, estimation with the Same Score covariate does not reveal statistically significant results, which restrain the rejection of $H5_0$ based on the second [b] condition.

The next two models (Model 5&6) test the existence of popularity effects within the network. The results indicate positive and statistically significant values for Score Ego (Outdegree popularity) and Score Alter (Indegree popularity) parameters. Hence, both $H6_0$ and $H7_0$ can be rejected in support of H6 and H7. The positive values of estimation results suggest that higher scoring participants both attract more incoming and initiate more outgoing communication ties over time. Furthermore, estimation for the Score Ego and Alter effects jointly with the Score Similarity effect (Model 6) indicates a different (-0.51) and not statistically significant result for the Score Similarity parameter. This indicates the dominance of the Score Ego and Alter popularity effects over the Score Similarity. In other words, while participants with similar scores tend to have greater interaction among each other, the predominant pattern within the studied network is the popularity of participants with higher scores.

V. CONCLUSIONS

Employing SNA techniques, this study identified interaction patterns that may not otherwise be immediately evident. The results indicate tendencies of: [a] stabilising interaction network; [b] selective approach in choosing interaction partners; and [c] reciprocating initiated contacts. Furthermore, [d] no tendencies of network closure were identified. The probabilistic approach to the network analysis allows reporting the results with statistical precision. The identified tendencies may suggest that learners, over time, may prefer to narrow their interaction to selected participants. Yet, the learners consistently exhibit a tendency to respond to initiated communication. Hence, the practitioners who intend to expose learners to the wider community may want to integrate designs that encourage learners to initiate communication with others. Furthermore, this study extends to identifying possible interrelation between network dynamics and learner achievement. The results indicate: [e] a homophily effect – a preference of participants to interact with others who were awarded similar scores; [f] a popularity effect – an effect of learner score on attracting or being attracted by other participants; and [g] a dominance of the popularity effect over the homophily effect. In other words, the higher scoring participants both attract more incoming and initiate more outgoing communication ties. Hence, alternative pedagogical designs may be considered to compensate for the predominant interaction patterns related to participant score.

REFERENCES

[1] K. Borau, C. Ullrich, J. Feng, and R. Shen, "Microblogging for Language Learning: Using Twitter to Train Communicative and Cultural Competence," 2009, p. 87.

[2] P. Chen, R. Gonyea, and G. Kuh, "Learning at a distance: Engaged or not," Innovate, 2008.

[3] J. Taylor, "Teaching and learning online: The workers, the lurkers and the shirkers," in *CRIDALA Conference*, Hong Kong, 2002.

[4] P. R. Monge and N. S. Contractor, *Theories of Communication Networks*: Oxford University Press, USA, 2003.

[5] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*: Cambridge University Press, 1994.

[6] A. Martínez, Y. Dimitriadis, B. Rubia, E. Gómez, and P. De la Fuente, "Combining qualitative evaluation and social network analysis for the study of classroom social interactions," *Computers & Education,* vol. 41, pp. 353-368, 2003.

[7] C. Haythornthwaite, "Social Network Methods and Measures for Examining E-learning," *Social Networks,* 2005.

[8] T. A. B. Snijders, "The Statistical Evaluation of Social Network Dynamics," *Sociological Methodology,* vol. 31, pp. 361-395, 2001.

[9] C. Steglich, T. A. B. Snijders, and P. West, "Applying SIENA: An Illustrative Analysis of the Coevolution of Adolescents' Friendship Networks, Taste in Music, and Alcohol Consumption," *Methodology,* vol. 2, pp. 48-56, 2006.

[10] T. A. B. Snijders, C. E. G. Steglich, M. Schweinberger, and M. Huisman, "Manual for SIENA, version 3," Groningen, The Netherlands: University of Groningen, 2006.

[11] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Reviews in Sociology,* vol. 27, pp. 415-444, 2001.

[12] T. Snijders, C. Steglich, and M. Schweinberger, "Modeling the co-evolution of networks and behavior," *Longitudinal models in the behavioral and related sciences,* pp. 41–71, 2007.