Evolutionary Social Network Analysis of Web-based Educational Environments

Andrew Laghos
Cyprus University of Technology
30 Arch. Kyprianos Str. Limassol, Cyprus
andrew.laghos@cut.ac.cy

ABSTRACT

This research provides a different perspective to analyzing the social networks that are formed in online learning environments. The novelty of this study lies in the investigation of the evolution of the online social networks as opposed to most current research which focuses only on the end state of a course when analyzing its computer-mediated communication. These educational social networks continually evolve and change over time, thus a method which can be used to assess their evolving nature is necessary. The contribution of this paper to the HCI community is the recommendation of Evolutionary Social Network Analysis as a method for studying students' communication interactions with each other during the duration of an online course. In the approach followed, the proposed method was applied to a case study and the results show that it can be a useful tool for researchers and educators in studying evolving social networks.

Keywords

Social Networks, Evolution, Analysis, Web-based Education.

ACM Classification

K.4 Computers and Society; H.5.3 Web-based Interaction

INTRODUCTION

One of the most important characteristics of the Internet today is the opportunities it offers for human communication and social networking by making use of Computer-Mediated Communication (CMC). These communication activities happen mainly through written communication and through audio and video [13]. By engaging in CMC, the students become part of evolving social networks, and an analysis or their communication activities becomes important in order to understand their behavior. Approaches that have been used for this purpose include Vrasidas's framework for studying human-human interaction in computer-mediated online environments [22], techniques like questionnaires, interviews, personas and log analysis, and a method called Social Network Analysis which is the basis of this study [12]. "Social Network Analysis (SNA) is the mapping and measuring of relationships and flows between people, groups, organizations, computers or other information/knowledge processing entities" [10].

Most studies that use SNA to analyze communication in e-Learning environments only focus on the final state of a network, but do not look into detail at how it changed over time [1]. Investigating the evolution of student connections in a course has several advantages. It enables the mapping out of the changes the network goes through and in addition it becomes possible to investigate how specific course amendments, participation in CMC, interface designs, and/or conversation topics positively or negatively influence the dynamics of the online community.

Although there have been some studies that used social network analysis to study the evolution of social networks [7, 15], they were not concerned with web-based educational environments. The purpose of these environments is learning and this plays a role in the evolution of their social networks.

This paper proposes the use of Evolutionary Social Network Analysis (ESNA) as a method to analyze the evolution of students' communication in online learning environments. The method was applied to a case study and the results showed that it can be used successfully for such a purpose.

BACKGROUND

In online educational environments, students engage in CMC when they are stuck, confused, excited, have questions, or just want to meet other peer students and have off topic discussions with them [18, 20].

Using CMC in these environments has its benefits as well as its limitations. For instance, a benefit of CMC is that the discussions are potentially richer than in face-to-face classrooms, but on the other hand users with poor writing skills may be at a disadvantage when using text-based CMC [18].

Studies show that students would prefer to contact their peer students (rather than their tutor) when they have difficulty with coursework, difficulty understanding lectures and difficulty assessing facilities [14]. In addition, knowledge is constructed in communities of practice through social interaction [3].

This makes it important to have tools that allow easy and straightforward ways for community members to interact with and support each other in a peer-to-peer fashion in web-based educational environments [11]. Major factors which influence interaction are course structure, class size, feedback, and prior CMC experience [21]. Research also shows that people who interact more in an online course tend to achieve higher marks at exams as opposed to lurking which is not as successful [3].

It is also necessary to have methods which can analyze the evolving nature of the students' CMC usage and this paper is addressing this by proposing ESNA.

METHODS

SNA has been used by analysts to determine if a network is tightly bounded, diversified or constricted, to find its density and clustering, and to study how the behavior of network members is affected by their positions and connections [6, 9, 19, 23].

SNA provides a philosophy and set of techniques for understanding how people and groups relate to each other [16]. It is concerned about dyadic attributes between pairs of actors (like kinship, roles, and actions), and has been used extensively by sociologists, communication researchers, and others.

The case study used was Learn Greek Online (LGO), a web-based course for learning the Modern Greek Language. In LGO, there were 15 lessons - each with its own discussion board. To follow the evolution of the students communications, instead of carrying out SNA on the end state of the network, it was instead applied to each individual lesson's discussion board, starting from the discussion board of Lesson 1 up until the student communications in the discussion board of Lesson 15. Using this form of Evolutionary SNA allowed for identifications of the students' patterns of communications along with the points of major changes in their communication behaviors.

RESULTS

618 students took part in the course. Their communication was characterized by direction (who sent a message to whom) and strength (i.e. how many times student x sent a message to student y) [4]. The data was collected directly from the discussion boards of the course and consisted of the sender and the receivers of each message. Once this information was obtained, the students' communication interactions were tabulated in the form of network matrices allowing for the analysis to take place. Below are the results from the analysis of inclusiveness, reachability, geodesic distance, centrality, and cliques.

Inclusiveness

Inclusiveness is the number of connected points expressed as a proportion of the total number of points [17]. In other words it is the number of connected students over the total number of students in the course. Isolates are students that are part of the network but have not made any communication interactions with any of their peers. They might be observing the discussions silently, however they have not contributed to any of the discussions and no other students have made contact with them. Figure 1 shows the evolution of Participants vs. Isolates in the LGO network.

Before the course began all the students were isolates since there were no connections with their peers. The major shift was at lesson 1 were over 80% of the students participated in the discussion forums. A content analysis of lesson 1's discussion board revealed that this was mainly because at the beginning of the course the students were more excited, they wanted to make sure they are in the right course, and they wanted to meet their fellow class mates. As the course evolved more students engaged in the discussions and finally every student ended the course with at least one connection with another student. This means that by the end of the course there were no isolates and inclusiveness was 100%.

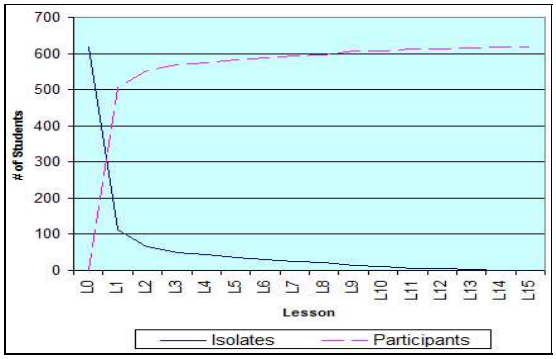

Figure 1. Participates vs. Isolates

Reachability and Geodesic Distance

If a path exists between two nodes they are said to be reachable. The length of the shortest path between two nodes (often referred to as the optimal connection between two actors) is called Geodesic distance [8].

As can be seen in Figure 2, the mean geodesic distance is on a decreasing trend from 2.63 to 2.57 indicating that the network becomes more connected since the mean shortest path between two nodes decreases.

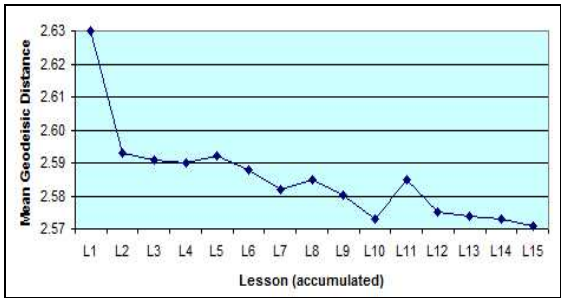

Figure 2. Geodesic Distance

At the same time, the average number of reachable students increased from 381.46 to 585.48 (Figure 3).

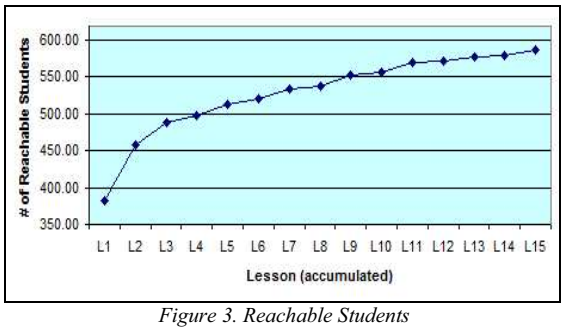

Figure 3. Reachable Students

This increase in reachable nodes by around 53% (from Lesson 1 to Lesson 15) tells us that as the course progressed the students had interacted with more of their peers that they hadn't interacted with before. This made

the network members closer and more connected with each other, since each student was indirectly connected to most of his/her peers through other students of the course. Even though as identified earlier there are no isolates in LGO, the number of reachable nodes is not 618 (as the total number of participating students). This is because some of the students have only made single connections with students who themselves have no other connections. Thus although these students have made connections with each other, they are still outside of the overall network connection and a shortest path to them does not exist.

Centrality

Degree centrality is measured by the portion of nodes that are adjacent to each node. The nodes with the highest degree scores are the ones who are more central (powerful) in the network.

In a directed network (where the direction of the communication is important) like the LGO case, the in-degree centrality is the portion of nodes that are adjacent to each node, and out-degree centrality is the portion of nodes that are adjacent from each node [5]. Through the evolution of the course the mean centrality in-degree (Figure 4) and out-degree values increased only slightly from lesson 1 (0.041) to lesson 15 (0.047). This indicated that the more central students had gained and maintained their powerful status early on in the course.

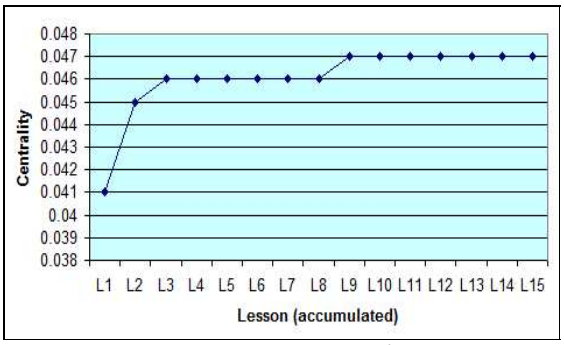

Figure 4. Degree Centrality

The students' individual centrality scores revealed that the students who participated and posted the highest number of messages in the discussion boards, were the same ones who received the highest number of incoming messages.

Cliques

A clique is a maximal complete subgraph of three or more nodes which are adjacent to each other, and there are no other nodes in the network that are also adjacent to all of the members of the clique. Cliques may overlap, meaning a node can be a member of more than one clique [2]. In the LGO case, I have carried out the evolutionary clique analysis on cliques with a minimum number of 3, 5, 10, 20, 50 and 100 members. Figure 5 shows the results. There was only 1 clique that had 50+ members (it was the same clique that had 100+ members) and it had developed from as early as lesson 1. This was because at the beginning of the course, the students were interested to get to know one another and make friends with their fellow classmates.

The cliques with 10+ and 20+ students increased gradually from lesson to lesson. The biggest increase however was in the number of cliques that had 3+ and 5+ members. These results show that communication in the social network remained active, while more and more students would interact with peers they hadn't exchanged messages with before.

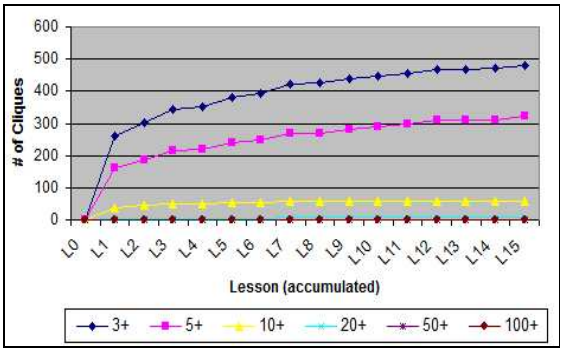

Figure 5. Cliques

DISCUSSION

The main advantage of using ESNA is that it assesses the social networks over the duration of a course. Current methods analyze the end state of a social network omitting the changes and growth that these people networks have gone through before reaching their final state. Through the use of ESNA in this case study it has been found that the students seem to be more excited and communicate more at the start of the course, but their overall participation rates are on a decrease during the duration of the course. In addition, the centrality in-degree scores were approximately equal to the centrality out-degree scores. This outcome suggests that if you participate and post messages in the discussion boards, you are more likely to get replies and in-coming messages from other students yourself. Finally, it was identified that the central students in the course took their powerful status from the beginning of the course. Although less powerful students gained more status during the duration of the course, the most central ones had already gained and maintained their status from the beginning.

ESNA is useful for keeping track of the network changes, while investigating how specific conversation topics or course amendments positively or negatively influence the dynamics of an online community. This way, people who use ESNA to assess their online communities benefit in the ability to predict how certain actions will affect their network, and to incorporate various methodologies to alter the state of their network.

HCI is an important discipline which is quite often omitted when planning and designing online learning environments. However, HCI issues must be addressed in order to be able to implement more user-friendly systems and interfaces as these ultimately play a role in the usage of the system by the students, and can affect their communication, motivation and learning outcomes.

Future research directions could include applications of ESNA to different e-Learning domains. It would be beneficial to use ESNA to study the evolution of social

networks in subject areas different to this one (like for example history, math, economics, music and so on) to see the differences and similarities of the students' communication in these courses.

CONCLUSION

This paper has provided insights into the use of Computer Mediated Communication by students in web-based educational environments. More specifically, a case study was used whereby the students' interactions and connections with their peers in the discussion boards of an online course were analyzed using the method of Evolutionary Social Network Analysis. In online courses, these people networks continually evolve and change over time thus following their evolution is significant.

The findings from the case study highlighted important communication patterns of the students taking part in the course that could not be otherwise investigated with existing methods. The approach provided in this paper can be a useful methodology for developers and maintainers of online learning communities as it provides characteristics about the nature and dynamics of their community enabling them to develop strategies for altering the state of their student social networks.

REFERENCES

1. Aviv, R. (2004). Concept Network Analysis of Students' quotes on Asynchronous Learning Networks. Available at http://telem-pub.openu.ac.il/users/aviv/papers/ConceptNetAnalysisofQuotes.pdf
2. Bock, R.D., and Husain, S.Z. (1950). An adaptation of Holzinger's B-coefficients for the analysis of sociometric data. Sociometry. 13, 146-153.
3. De Angeli, A, Sue, K (2005). Learning conversations: A case study into e-learning communities. Proceedings of the Interact 2005 eLearning and Human-Computer Interaction Workshop. 12-16 Septemeber, 2005.
4. De Nooy, W., Mrvar, A., & Batagelj, V. (2005). Exploratory social network analysis with Pajek. New York: Cambridge University Press.
5. Freeman L C. (1979). Centrality in Social Networks: Conceptual clarification. Social Networks. 1, 215-239.
6. Garton, L., Haythorthwaite, C., & Wellman, B. (1997). Studying On-line Social Networks. In Jones, S. (Eds.), Doing Internet Research. Thousand Oaks CA: Sage.
7. Gisela, B., and Aili, M. (2008). A Social Network Analysis of the evolution of the environmental criminology and crime analysis (ECCA) Symposiums. Crime Patterns and Analysis, Vol 1(1), 5-22.
8. Hanneman, R. A. (2001). Introduction to Social Network Methods. Available at http://faculty.ucr.edu/~hanneman/SOC157/TEXT/TextIndex.html
9. Knoke, D., & Kuklinski, J.H. (1982). Network Analysis. Sage University Paper Series on Quantitative Applications in Social Sciences. Serious no. 07-001. Beverly Hills and London: Sage Publications.
10. Krebs, V. (2011). An Introduction to Social Network Analysis. Available at http://www.orgnet.com/sna.html
11. Kurhila, J., Miettinen, M., Nokelainen, P., & Tirri, H. (2004). The Role of the Learning Platform in Student-Centered E-Learning. Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies.
12. Laghos, A. (2012). Multimedia Social Networks & e-Learning. In Kanellopoulos, D. (Eds.), Intelligent Multimedia Technologies for Networking Applications: Techniques and Tools. IGI Global. Hershey, Pennsylvania, USA.
13. Laghos, A., Masoura, S., and Skordi, A. (2012). Linguistics in Social Networks. American International Journal of Contemporary Reasearch. Vol 2(1), pp1-5.
14. Lockley, E., Pritchard, C., & Foster, E. (2004). Professional Evaluation Students supporting students - lessons learnt from an environmental health peer support scheme. Journal of Environmental Health Research, Volume 3, Issue 2.
15. Michelle C. Kegler, Jessica Rigler, Maya K. Ravani (2010). Using network analysis to assess the evolution of organizational collaboration in response to a major environmental health threat. Health Education Research, Vol. 25, No. 3., pp. 413-424.
16. Preece, J. (2002). Online Communities: Designing Usability, Supporting Sociability. John Wiley and Sons: Chichester, UK.
17. Roberts Jr, J.M. (2000). Simple methods for simulating sociomatrices with given marginal totals. Social Networks. 22, 273-283.
18. Scotcit. (2003). Enabling large-scale institutional implementation of communications and information technology (ELICIT). Using Computer Mediated Conferencing. Available at http://www.elicit.scotcit.ac.uk/modules/cmc1/welcome.htm
19. Scott, J. (2000). Social Network Analysis: A handbook. Second edition. London: Sage.
20. Sumner, J., & Dewar, K. (2002). Peer-to-Peer eLearning and the Team Effect on Course Completion. ICCE: 369-370.
21. Vrasidas, C., & McIsaac, S. M. (1999). Factors influencing interaction in an online course. The American Journal of Distance Education, 13(3), 22-36.
22. Vrasidas, C. (2001). Studying human-human interaction in computer mediated online environments. In Manopoulos, Y. & Evripidou, S. (Eds). Proceedings of the 8th Panhellenic Conference on Informatics, Nicosia, Cyprus.
23. Wellman, B. (1997). An Electronic Group is Virtually a Social Network. In Kiesler, S. (Ed.), Culture of the Internet. Hillside, New Jersey: Lawrence Erlbaum Assoc. 179-205.