Using Social Network Analysis in the classroom: a case study applying NodeXL

Antonieta Kuz[1], Mariana Falco[1], Enzo de Renzis[1], Leopoldo Nahuel[1], Roxana Giandini[1,2]

[1] LINSI, Universidad Tecnológica Nacional, Facultad Regional La Plata, La Plata, Buenos Aires, Argentina.
{akuz, mfalco, ederenzis, lnahuel} @frlp.utn.edu.ar
[2] LIFIA, Facultad de Informática, Universidad Nacional de La Plata. La Plata, Buenos Aires, Argentina.
giandini@info.unlp.edu.ar

Abstract. The interest in social networks has extended to different disciplines, such as Computer Science. This approach brings Social Network Analysis (SNA) as the study of social structure in different environments, like companies, establishment, and schools, among others. For this reason, this article highlights basic network information like graphs using actors and relations and important concepts related to classroom like structure which constitutes it. Therefore, the main aim of this research is to examine the educational use of SNA. Furthermore, schools in society are a system of actors joined by relationships. Accordingly, the current paper presents a qualitative analysis through a practical approach of SNA, by describing the classroom as a sociometric experiment using NodeXL, to verify the contrast keeping attention in show how students make informal contact and the knowledge that this brings.

Keywords: SNA, Sociometry, Graphs, Classroom, NodeXL

1 Introduction

Information and Communication Technologies (ICTs) can contribute to universal access to education, equity in education, delivery of quality learning and teaching, teachers' professional development and more efficient management in education, governance and administration. As such, the ICTs are becoming an ubiquitous component of classroom learning. They are able to provide additional opportunities to support the learning process and it may be able through the future growing, to transform educational practices [1].

The interest in Social Networks has been increasing and evolving across a wide variety of fields and researches, such as Physics, Psychology and Computer Science [2]. Social Network Analysis (the acronym SNA) was developed in a relatively non-technical manner from the structural concerns of the anthropologist Radcliffe-Brown.

He started to develop a concept of social structure and a web of social life. Social networks have also been studied by Milgram's small world research [3] [4].

Social interactions between students are a major and underexplored part of undergraduate education. Understanding how learning relationships form in undergraduate classrooms, as well as the impacts these relationships have on learning outcomes, can inform educators in unique ways and improve educational reform. Social Network Analysis provides the necessary tool kit for investigating questions involving relational data. We introduce basic concepts within SNA, along with methods for data collection, data processing, and data analysis. In order to make this proposal practical, we present a case study developed in the NodeXL [5] tool to create visualizations of the social networks studied which assist us in our analysis.

We consider the classroom as the main target of this investigation, because this is the environment where professors and students interact every day, involving relational data set to obtain the patterning of relationships among students. Through SNA, it will be possible to deepen the knowledge of social phenomena [6]. This perspective will allow us to point out the intensity of the relationships within the group to study; the degree of cohesion, the structure of a group and each of the positions occupied by each member in the classroom.

This paper is organized as follows: in section 2 we introduce the main Social Network concepts, while section 3 defines SNA. Section 4 presents a Case Study with NodeXL in a school environment and the description of some results obtained through different metrics about the case study. Finally, conclusions and future work are presented in Section 5.

2 Social Networks: Structure and Basic Concepts

With the growing of online social media, everything is connected, people, students, employees, information, events, places, among others. A practical way of making sense of the tangle of connections is through analyzing them as networks. Social networks can be named as a well-defined set of actors such as individuals, groups, organizations, communities, etc. linked to each other, through a relationship or a set of social relations. Another field which formally studies Social Networks is Graph Theory, which is a branch of Mathematics [7].

In social networks, nodes and vertices in a graph represent the actors and relations respectively. Most of the bounds between actors have a purpose. The interpersonal bounds in a network are characterized by the roles and the context of those where they have been developed. The dynamic character of the network appears because sometimes the relations are more or less persistent or also it might be for the existence or not between actors.

A social network is formed by actors. It is very important to try to identify the central individuals in the network. On the one hand, its attributes refer to different aspects, characteristics, and intrinsic properties of the individuals such as opinions, comments, suggestions, and so on. Actors do not act independently; they are influenced in their behavior and attitudes by other actors to whom they are tied. On the other hand, relational data are the contacts, ties and connections, which relate one

actor to another actor. The relations connect pairs and express linkages of actors, they are specific to the context, and the context depends on the interactions among them. Examples of relations are friendship, job relations, flow of information, among others topics [3].

Burt [8] specifies that all social actors involved in a social system that incorporate other actors are significant landmarks in each other's decisions. The relationships that an actor has with others can affect their actions, perceptions and behaviors. So, SNA is focused on uncovering the patterning of how people's interactions will result on data sets. The analysis of the relations between actors allows delimiting the dynamics of flow circulation between actors located in different places in the network.

3 Social Network Analysis: Definition and Tools

Social Network Analysis aims to understand the determinants, structure, and consequences of relationships between actors. SNA is the mapping and measuring of relationships and flows between people, groups, organizations, computers or other information/knowledge processing entities. Social network analysis is the study of structure [9]. It involves relational datasets. That is, structure is derived from the regularities in the patterning of relationships among social entities, which might be people, groups, or organizations.

SNA is a method for visualizing the people and connection power, leading us to identify how we can best interacts to share knowledge. Thus, SNA brings the explanation of behavior of relations that requires an analysis of how the actors are connected to one another considered in a particular environment with contextual factors. In the next section, we provide an analysis of the structure relations between social actors in the school environment [3].

The importance of relationships and emergent structures formed by relationships makes SNA different from other research paradigms, which often focus solely on the attributes of actors. For example, traditional analyses may separate students into groups based on their attributes and search for disproportional outcomes based on those attributes. A social network perspective would focus instead on how individuals may have similar network positions due to shared attributes. These similar network positions may present the same social influences on both individuals, and these social influences may be an important part of the causal chain to the shared outcome. In situations in which a presence or absence of social support is suspected to be important to outcomes of interest, such as formal learning within a classroom, the SNA paradigm is appealing.

NodeXL is a tool for interactive network visualization that leverages the widely available MS Excel application as the platform for representing generic graph data, performing advanced network analysis and visual exploration of networks (see Figure 1) [5]. Likewise, NodeXL is a practical tool, because it uses a highly structured workbook template that includes multiple worksheets to store all the information needed to represent a network graph. NodeXL provides several visualization options availables.

We selected this software tool, because it allowed us to use certain metrics from SNA. Furthermore, we have used this case study to show how the combination of social network and qualitative analysis work together, characterized by communication network defined as a set of "interconnected individuals who are linked by patterned communication flows" [3].

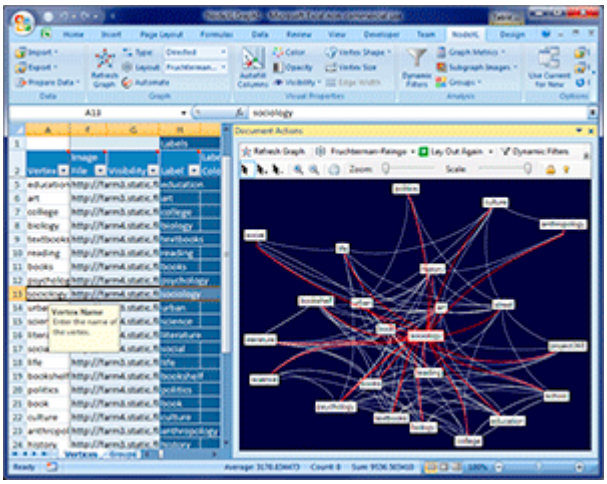

Fig. 1. NodeXL tool

4 Case Study

Burt [8] specifies that all social actors involved in a social system that incorporate other actors are significant landmarks in each other's decisions. The relationships that an actor has with others can affect their actions, perceptions and behaviors. So, SNA is focused on uncovering the patterning of how people's interaction will result on data sets. The analysis of the relations between actors allows delimiting the dynamics of flow circulation between actors located in different places in the network.

Most of the bounds between actors have a purpose. The interpersonal bounds in a network are characterized by the roles and the context of the roles they have developed. The dynamic character of the network appears because sometimes the relations are more or less persistent or also it might be for the existence or not of his actors.

As Hoffman [10] points out, "Sociometry is based on the fact that people make choices in interpersonal relationships. Whenever people gather, they make choices about where to sit or stand; choices about who is perceived as friendly and who is not, who is central to the group, who is rejected, who is isolated". Moreno [11] explained that all criteria have this in common: that the respondents have some actual experience in reference to them, whether ex post facto or present; in sociometric language, they are still "warmed up" to them otherwise the questions would not arouse any significant response.

4.1 Applying Techniques of Collection and Delivery of Social Networks

This section allows us to describe the set of methods and techniques used to characterize a school group specifying the methodological approach of the research, the techniques used to collect the data and the procedures performed for the analysis and interpretation of results.

Sociometric methods became part of SNA and have been developed like a tool that generates an excellent material from the group from specific questionnaires [11]. We take into account and explored a classroom environment composed of a set of eleven (11) students from twelve (12) and thirteen (13) years with a GPA intermediate with individuals highlighted in the positive aspect, i.e. "students with good grades". In order to preserve the identity of the students, we used labels (English, labels), and all the group has been numbered from A1 to A11, while the teacher is A12.

In our study case, we simulate a set of questions to students to obtain a dataset. The sociogram is considered a character sociometric technique, i.e., a method for measuring social relations among members of a group, where its elements are known, have common goals and influence each other. Graphically, a sociogram represents relationships by dots (individuals) appearing together by one or more lines (inter-relationships).

With this information, it is plausible for us to illustrate the most appropriate way in which teachers might make decisions in the classroom. Sociometric data were collected by questions that are formulated online to the students listed in the following list:

1. Who would you choose as a class leader?
2. Who doesn't get on with whom?
3. Who would choose for project teams?
4. Who would you trust in knowing what is really taking place?

4.2 Discovering Group Dynamics through Graphs Topologies

Group dynamics is a system of behaviors and psychological processes occurring within a social group (intragroup dynamics), or between social groups (intergroup dynamics). Kurt Lewin [12] coined the term "group dynamics" to describe the way groups and individuals act and react to changing circumstances. Therefore, the study of group dynamics can be useful in understanding decision-making behavior, tracking the spread of diseases in society, creating effective therapy techniques, and following the emergence and popularity of new ideas and technologies [13].

We use NodeXL to represent graphs from data sets previously obtained from the precedent sociometric questionnaire [5] [14]. Graphs are visual representations of networks, displaying actors as nodes and the relational ties connecting actors as lines. In this case, each node represents a student, considering A1 to A11. Each actor was provided with a list of all actors in the network and asked to indicate those with whom he or she has a particular relation according to the questions previously mentioned. According to the questions set above, it is correct to note that the connections might involve identifying people with whom the student expresses he or she frequently socializes. After that, we display the graph by clicking on the Read Workbook button

directly above the graph pane; a network showing the ties is displayed, as we show in Figure 2.

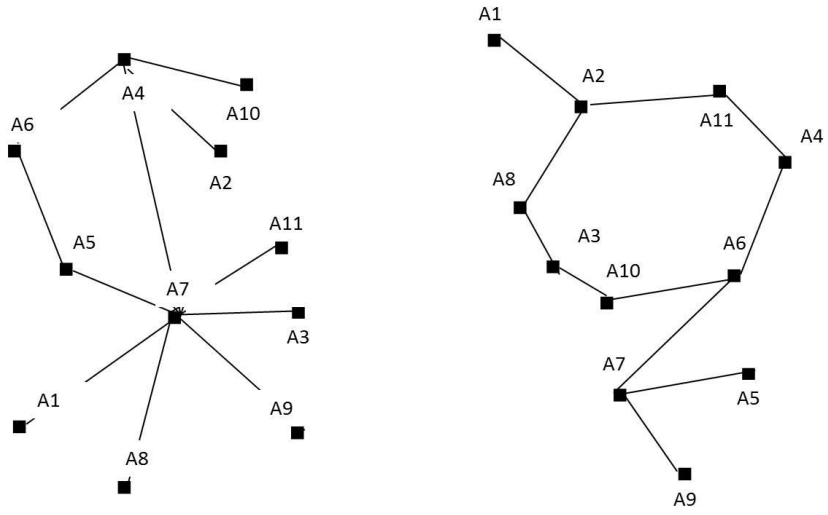

Fig. 2. Social network with a flow of information in the classroom (questions 1 and 3)

Regarding to question 1, we highlight several issues about relationships in the classroom. First, there are no mutual choices because relationships require bilateral content of two involved actors; therefore similarity between members partly increases the likelihood of forming the links within the groups. Then, we emphasize that 5 people elect A7, which makes it a great informal leader (students chosen by both current and by peers).

After that, there are a number of people within the classroom we described as isolates, such as a group of individuals prevented by social barriers from interbreeding with others of their kind and social exclusion. They are: A9, A1, A8, A11, A10 and A2. As in the previous case, there are no mutual choices in response to question 3. It happens that A7 and A2 have been chosen by two people, both present the most choices. Then, we point that between A6, A10, A3, A8, A2, A11 and A4 occurs what graph theory calls a cycle graph which consists of a single cycle with some vertices connected in a closed chain. In this case, we have a directed cycle graph with all the edges being oriented in the same direction. We could say that these relationships suggest a certain bond where friends choose each other (without prior agreement).

Since we obtain the graph of the classroom conferring to the questions 1 and 3, we will continuous to analyze some metrics. It is important to understand that the metrics ultimately became the principal aspect of analysis, and they are obtained with the same software tool.

4.3 Discovering Group Dynamics through Graphs Topologies

Part of the SNA studies all the interaction between individuals and organizations, and flows of information. The analysis of the relations between actors allows delimiting the dynamics of flow circulation between actors located in different places in the network. Most of the bounds between actors have a purpose or may have it because there exist some interest. Therefore, the interpersonal bounds in a network are characterized by the roles and the context of the roles they have developed.

One key direction for education researchers is to study network formation within classrooms, in order to elucidate how the realized networks affect learning outcomes. Network analysis can give a baseline understanding of classroom network norms and illuminate major aspects of students learning. Therefore, network data are collected at the individual level, but the analyses occur at the structural level, with the use of different measures. This allows us to obtain the classroom structural characteristics.

This property refers to the overall pattern of relationships of relationships between the system's actors, by capturing the size and internal connectivity of a network as well as attributes of each node. NodeXL supports a minimal set of the most crucial network measures for individual nodes, such as: density, distance, reachability, degree and betweenness.

4.3.1 Cohesion-density

Perhaps the most basic measurement in network analysis is network density, which points out how many links are observed in a whole network divided by the total number of links that could exist if every actor were connected to every other actor. In Figure 2 it is possible to discriminate that there are 11 ties out of a possible 121 for the organizational network, giving a density of 0.091.

4.3.2 Cohesion-distance

Consider two persons, call them A and B, which each might have five friends. But suppose that none of A's friends have any friends except A. B's friends, in contrast, have five friends each. The information available to B is that B's friends have potential for influence which is far greater than A's; usually known as being a "friend of a friend" may be quite consequential. To capture this aspect of how individuals are embedded in networks, one main approach is to examine the distance that an actor is from others.

If two actors are adjacent, the distance between them is one (that is, it takes one step to go from the source to the receiver). As shown in Figure 2, for question 1 A7 is a distance of 2 from A6, A2, A10, and A9; and for question 3 A7 is a distance of 7 from A4 instead of A11 who can reach A4 within an only tie. This is the notion of ‘‘degrees of separation’’ made familiar to many by a popular play [15].

4.3.3 Cohesion – Reachability

An actor is "reachable" by another if there exists any set of connections by which we can trace from the source to the target actor, regardless of how many others fall between them. In other words, reachability measures whether actors within a network are related, either directly or indirectly, to all other actors [16]. If the data are asymmetric or directed, it is possible that actor A can reach actor B, but that actor B cannot reach actor A.

Regarding the question 1, with the exception of the five isolates (A8, A11, A1, A10, and A9) all of the remaining actors in figure 2 can reach one another. For question 3 with the exception of the three isolates (A5, A9, and A1) all of the remaining actors in figure 2 can reach one another.

4.3.4 Degree Centrality

Centrality measures identify the most prominent actors, those who are extensively involved in relationships with other network members. [17] Centrality indicates one type of ‘‘importance’’ of actors in a network: in lay terms, these are the ‘‘key’’ players. Degree centrality is the sum of all other actors who are directly connected to ego. It signifies activity or popularity. Lots of ties coming in and lots of ties coming out of an actor would increase degree centrality.

For question 1, A7 has the highest degree centrality with five direct ties and A4 is the next most central with three direct ties. For question 3, A7 and A2 have the highest degree centrality with two direct ties each.

4.3.5 Betweenness Centrality

Betweenness centrality is the number of times an actor connects pairs of other actors, who otherwise would not be able to reach one another. It is a measure of the potential for control as an actor who is high in ‘‘betweenness’’ is able to act as a gatekeeper controlling the flow of resources between the alters that he or she connects [17]. A7 is by far the most powerful actor depicted in Figure 2.

All actors in the network must go through A7 to reach A1, A11, A8, A3 and A9. The next most powerful is A4, because all actors must get through him to reach A2 and A10. Also, in question 3, all actors in the network must go through A6 to reach A10, A3, A8, A2, A11 and A4. By going through A2, we can reach A1. In the other way, through A7 we can reach A5 and A9.

These measures of centrality are purely structural measures of popularity, efficiency, and power in a network, namely that the more connected or central an actor is the more popular, efficient, or powerful.

5 Conclusions and future work

Beyond considering a group of students and determining the dynamics, this research focused on deepening the search for a simple way in which it is plausible to help not only the teacher, but mainly the students, paying attention to their views and needs [11]. We have used NodeXL, a software tool through which we obtained different results on relations in the classroom, and by the metrics we appreciate and support the textual analysis of the graphs. Therefore, this study not only was a description of the graph, it was also an objective assessment of the hidden reality on students, their feelings, their relationships with classmates and teachers, as well as their perception of the organizational hierarchy.

As we mentioned at the beginning, SNA not only can be applied to an organization like schools, but also to different environments. In the school, we see that SNA is very important, because in this environment children and teenagers improve their character through emotions, feelings but especially, by fulfilling the meaning of relationships. They learn how to interact with the equals, how to build what we know as social consciousness. The connection and exchange between students are the most important sources of information and knowledge, and this is it because students trust more those they know than those they don't.

As regards future work, we will try to analyze the network formed by students, and after that, we are going to study their choices, which will determine what others will see and how they will be connected to others. So we will intend to consider the concept of trust and the different concepts related to it, like confidence and their differences. In this way, we will introduce trust to Computer Science, and we will also place it in the center of our attention in SNA. Moreover, we are going to use NodeXL to represent that. In this manner, we will take into account the properties and the most important components: the trustor and the trustee, and how they interact.

Acknowledgments. The authors wish to thank the authorities of the National Technological University - Regional La Plata (UTN - FRLP) for their support, as well as the Innovation Laboratory Information Systems (LINSI).

References

1. Greenhow, C. and Robelia, B.: Old Communication, New Literacies: Social Network Sites as Social Learning Resources. Journal of Computer-Mediated Communication, 14:130–1161. doi: 10.1111/j.1083-6101.2009.01484.x (2009)
2. Hanneman, R. A. & Riddle, M.: Introduction to social network methods, University of California, Riverside, Riverside, CA, (2005)
3. Scott, J.: Social Network Analysis: A Handbook, Sage Publications (2000)
4. Watts, D.: Six Degrees: The Science of a Connected Age, W. W. Norton (2004)
5. Matei, S. A.: 'Analyzing Social Media Networks with NodeXL: Insights from a Connected World by Derek Hansen, Ben Shneiderman, and Marc A. Smith', Int. J. Hum. Comput. Interaction 27(4), 405–408 (2011)
6. Kuz, A., Falco, M.: Herramientas sociométricas aplicadas al ambiente áulico, in

Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2013) (En línea) Córdoba, Argentina. ISSN: 2347-0372 (2013)
7. Tang, L., Liu, H.: Graph mining applications to social network analysis. In C. Aggarwal and H. Wang, editors, Managing and Mining Graph Data, chapter 16, pages 487-513. Springer (2010)
8. Burt, R.S.: Models of Network Structure. Annual Review of Sociology, vol. 6, pp. 79-114. JSTOR publishers (1980)
9. Wellman, B., Berkowitz, S.D.: Social structures: a network approach. Greenwich: JAI Press, (1997)
10. Hoffman, C., Wilcox, L., Gomez, E., & Hollander, C.: Sociometric applications in a corporate environment. Journal of Group Psychotherapy, Psychodrama & Sociometry, 45(1), Spring, 3-16.(ISSN: 0731-1273 (1992)
11. Moreno, J.L.: Who Shall Survive? 2nd edition, New York. Beacon House (1953)
12. Lewin, K.:Dinámica de la personalidad: selección de artículos, Filosofía, Psicología, Pedagogía, Morata (1973)
13. Backstrom, L.; Huttenlocher, D.; Kleinberg, J.; Lan, X.: "Group formation in large social networks". Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '06. p. 44. doi:10.1145/1150402.1150412. ISBN 1595933395 (2006)
14. Dunne, C., Shneiderman, B.: Improving graph drawing readability by incorporating readability metrics: A software tool for network analysts , University of Maryland, Human-Computer Interaction Lab Tech Report HCIL-2009-13 (2009)
15. Guare, J.: Six degrees of separation: a play. New York: Vintage Books (1990)
16. Doreian, P.: On the connectivity of social networks. Journal of Mathematical Sociology 1974;3:245–58.
17. Freeman, L. C.: 'Centrality in social networks conceptual clarification', Social Networks p. 215 (1978)