Social Network Learning Analytics: identification of students at risk of early school leaving

Social Network Learning Analytics: identificazione degli studenti a rischio di abbandono scolastico

Caterina Bembich
Dipartimento di Studi Umanistici, Università degli studi di Trieste, cbembich@units.it

Social Network Analysis (SNA) is gaining increasing attention in educational research as branch of Learning Analytics. This contribution offers an overview of Social Network Analysis, describing the origins of this methodology, the implications of its use for educational research, and its application in the study of early school leaving.
This paper presents a case study of the application of SNA in a group of students attending a vocational school. The study analyzes the external and internal relationships of the students in the school. The results highlight that students at risk of dropping out of school, tend to have less dense and less cohesive social networks, and exhibit a higher probability of establishing relationships with classmates with similar tendencies to leave school early. SNA draws attention to the relational structures that are established during school activities, and can help teachers and schools improve learning processes and learning environments more broadly. This study shows how a relational approach can be used to explore the phenomenon of early school leaving and highlights dysfunctional relational structures that accentuate the risk situation.

Keywords: Social Network Analysis; Learning Analytics; Early school leaving; Interdependence; Identification of risk

La Social Network Analysis (SNA) sta attualmente emergendo nella ricerca educativa, come un settore dei Learning Analytics. La prima parte del contributo descrive le origini della Social Network Analysis, le implicazioni nella ricerca educativa e la sua applicazione nello studio del fenomeno dell'abbandono scolastico.
Il lavoro presenta successivamente un caso studio di applicazione della SNA. Lo studio analizza le reti di relazioni, sia esterne che interne alla scuola, costruite da un gruppo di studenti frequentanti una scuola professionale. Dai risultati ottenuti si evidenzia come studenti a rischio di abbandono scolastico tendono ad avere reti sociali meno dense e meno coese, e a mostrare una probabilità più alta di stringere rapporti con compagni di classe con la loro stessa tendenza all'abbandono. Lo studio evidenzia come un approccio relazionale può essere utilizzato per esplorare il fenomeno dell'abbandono scolastico e mettere in luce strutture relazionali disfunzionali che accentuano la situazione di rischio. L'analisi delle reti sociali può aiutare gli insegnanti e la scuola a migliorare i processi di apprendimento attraverso la strutturazione di percorsi di apprendimento che possano favorire strutture di rete funzionali al raggiungimento del successo scolastico, attraverso il coinvolgimento attivo degli studenti.

Parole chiave: Social Network Analysis; Learning Analytics; Abbandono scolastico; Interdipendenza; Identificazione del rischio

© Pensa MultiMedia Editore srl – ISSN 2038-9744 (on line) - DOI 10.7346/SIRD-2S2019-P176

Introduction

"Learning Analytics" is understood as the sampling, measurement and examination of learning data, derived from environments and from learners, with the aim to improve learning settings and processes, using an evidence base approach (Buckingham Shum, 2012; Siemens, 2013). According to this definition, Learning Analytics (LA) can include different levels of analysis (macro, meso or micro levels). When the primary interest of a study is addressed to the learners and teachers, the analysis occurs at the micro level.

Micro-level analysis can aid in the identification of students at risk of failure, and subsequently, given the highly detailed nature of the analysis, can provide a strategy for improvement (Buckingham Shum, 2012). At this level of analysis, Social Network Analysis (SNA) can be adopted as a technique to collect and analyze interpersonal data. In educational research, SNA is currently emerging as branch of Learning Analytics, aiming to better understand and improve learning contexts (Siemens, 2005; Ferguson, Shum, 2012). SNA has a solid base in the learning sciences; however, from 2008 it begun to be increasingly applied within the literature related to Learning Analytics (Ferguson, 2014). Furthermore, pedagogical theories were increasingly referenced in the LA literature, due to the research of experts employing Social Network Analysis (Dawson, 2008; Dawson, McWilliam, 2008). These authors argue that learning processes are facilitated by the participation of individuals, adopting a socio-constructivist view, that considers learning as a process of social construction of knowledge and skills (Bruner, 1986).

SNA is a methodology that studies the relations and interactions between individuals in a social network (Domínguez, Hollstein, 2014). It originates in the field of sociology, and in network theory in particular, and was developed to study relationships and social structures among individuals in a group (Siemens, 2005). Moreno (1934) developed sociometric theory, from which sociomatrix and sociogram are derive as tools to undertake social analysis. Currently, new approaches to SNA have been developed, through the use of computerized statistical models, that have helped evolve the analysis of students' interactions in learning environments (Siemens, Long, 2011; Robins et al., 2007).

1. SNA fundamental principles

The SNA methodology takes a relational perspective: it represents social networks through a graph structure, viewing actors as nodes in a network and ties representing links between pairs of actors (Wasserman, Faust, 1994). As a part of social learning analytics (e.g., Ferguson, Shum, 2012), network visualization is considered as a potentially helpful feedback for learners, that can stimulate them to reflect on their social interactions in a group, with the aim of building collective knowledge (e.g., Dawson, 2010).

The nature of the social network is defined by the type of pairwise connection, represented by different network structures. The data analysis of social network quantifies the importance of actors among the network and allows recognition of group(s) of actors connected more densely than others. Through SNA, we can understand if ties and relations between individuals are weak or strong, depending on the frequency, quality or importance of bonds (Domínguez, Hollstein, 2014).

The analysis of networks is founded on a mathematical approach based on graph theory (Van Steen, 2010), which represents the social networks as a binary matrix, called an adjacency matrix. The socio matrix is a method commonly used to conduct social network analysis, to analyze and understand in a quantitative way the interactions between individuals. It represents the presence or strength of ties between the members of a group. Moreover, there are statistical techniques and models to analyze and describe an observed network, which measure its properties (Skyler, Bruce, 2010).

The perspective of analysis can be directed towards the individual (egocentric approach) or to the entire network (complement network design). The first approach can be useful to identify contextual factors that influence learning and that support an individual's learning. In the second approach, a global view identifies elements that hold the network together and provides information's regarding a set of people (Haythornthwaite, de Laat, 2010); it is useful to identify communities and individuals' affinity groups within a network, that can or cannot support learning (Gee, 2004).

2. Application of Social network analysis and early school leaving

Through educational research we can investigate the factors that can interfere with the school participation of at-risk students (Boaler, Staples, 2008; O'Connor , Michaels, 1996), contrasting the tendency to explain early school leaving as a phenomenon reducible to individual traits, or to a "culture of poverty" (Gutierrez, Rogoff, 2003).

The factors which explain early school leaving are complex and include a variety of situations: failure to attend school and drop out before the conclusion of studies; repetition, irregular frequency in attendance; and delays with respect to school age. Moreover, there is also a "covert dispersion", that include frequent delays and absences from involvement during lessons. In Italy between the school years 2015/2016 and 2016/2017, 1.35% of students attending first grade of secondary school and 4.31% of students attending the second grade of secondary school, abandoned their schooling (see The National Student Registry, established by Legislative Act 76 on 15 April 2005).

Scholastic outcomes can be influenced by several factors; according to a Bioecological perspective (Brofrenbrenner, 1995), success or failure in school is never determined by a single factor. Rather the outcome in the students' learning pathways is determined by interaction between child, family, school and social context, and their characteristics. Potentially at risk students can nevertheless achieve academic success against the odds, not only thanks to their personal characteristics, but also thanks to the contributions of people around them (Siraj and Mayo, 2014). Students are embedded within social networks, inside and outside school, that can affect in positive or negative way their learning pathway (Borgatti, Halgin, 2011).

The relational context at school, assumes different forms that can be studied in order to identify students at potential risk of abandonment. To contrast learning failure, is important to stimulate the building of strongly connected interactional networks, that favor the construction of learning communities (Brown, Campione, 1990). Belonging to a community constitutes a part of an individual's identity: a positive sense of community emerges when individuals feel both important and needed in that group. It is enhanced by promoting interdependence between individuals, namely the collaboration between students with the specific aim of learning from each other (Sarazin, 2017). Researchers argue that interdependence is important in promoting social group cohesion and learning in educational settings (e.g., Osterman, 2000).

Through the application of Social Network Analysis (Dominguez, Hollstein, 2014), it is possible to highlight relational structures that are established during school activities, and identify possible situations of risk or isolated students. On the basis of SNA results, teachers and schools can design activities that promote cohesion among students and knowledge exchange to promote the implementation of interdisciplinary projects, and encourage cooperative connections and collaboration between students (Haythornthwaite, De Laat, 2010).

Starting from these considerations, in the following we draw on an SNA exploratory study to analyze the phenomenon of early school leaving, developed inside the FAMI-IMPACT FVG 2018-2020 project. The exploratory inquiry (Lumbelli, 1989) is addressed to examine the existing network configuration of a class members, in order to define variables or categories of observation, to control and use for future investigations.

We investigated the structure of the social networks in a group of students attending the first year of a vocational school: we analyzed the network of their external relationships (in order to explore the support received by the students outside of school), and the internal network in the class; moreover, we wondered if being at risk of school dropout affects individual school members' likelihood of having relationships in networks.

Therefore, the following research questions are:

RQ1: How is structured the supportive network of the risk students outside the school?

RQ2: How is structured the social network of the risk students, inside the class?

3. Exploratory inquiry

3.1. *Sample*

The sample consisted of 17 students attending their first year of a vocational school in Trieste (15 males, 2 females; age M=17; SD= 1.28). Students came from a variety of countries of origin: 9 students came from extra UE countries (Foreign students); 7 had Italian citizenship (Italian students); and 1 had Italian citizenship with one of their parents coming from extra UE countries (Migration background). Only 4 students were not repeating the year; the majority failed during the previous school years.

Teachers provided information on the school context; in general, the students enrolled came from previous negative school experiences, and presented lacking motivation. However, there were also some students which showed motivation and interest for studying, although their previous school difficulties; these students were well engaged in their educational path, considered as a possibility for their future job placement. The teachers also reported the difficulty to positively involve the families of the students, which in many cases were scarcely engaged in the school pathway of their sons.

3.2. *Measures*

Through a survey distributed to the students, we collected general information and relational data. The questions explored the following aspects:

1. General information: age, citizenship, sex
2. Risk of leaving school: "Have you ever thought of leaving school?"
3. Sociometric questions:
   3.1 Egocentric network-external support received for schooling (from family, or other reference adults):
      (a) "Who help you with your homework?"
   3.2 Complement network design within the class group:
      (a) Network existing before beginning school: "Did you know your classmates before starting this school?"
      (b) Network after beginning school:
      B1 "With which of your classmates do you hang out with during break or talk about personal things?"
      B2 "Which of your classmates do you hang out with even outside school?".

To answer these socio-metric questions, respondents were provided with a list of their classmates. In order to contribute to ensuring anonymous analysis of the data, the list contained a letter code for each actor.

Students were asked to indicate this letter code by completing the survey; there was no limitation to the number of classmates a respondent could indicate.

3.3. *Data analysis*

Social network properties were calculated using the software package UCINET 6.0 (Borgatti et al., 2002, 2009). Regarding question 3.2(b), the dependent variable is defined as the existence or absence of a relationship between two students (a dyad). For every pair of schoolmates, a value of 1 represents a relationship between them; a value of 0 indicates the absence of a tie between the two members.

We identified the following SNA indexes: network density, network degree centralization and homophily. These indexes provide basic information about the global structure of the network and the activity of the group members.

The density of a network describes the general level of the links between the points in a graph. The more nodes that are directly connected to each other, the denser a graph is. The density index assumes a value that varies from 0 to 1 (= density of a complete graph when all the nodes are adjacent to each other). Centrality helps to describe the power relations that are established within it. A high degree of centrality can mean that the node occupies a position of prestige compared to the others, and can be interpreted as the percentage of relationships that school team members maintain within the whole network.

Homophily is a measure of the tendency of the individuals to be more likely to have ties with others who are similar to themselves on specific attributes (for example, age, gender, education). Similar backgrounds increase the likelihood that students possess shared experiences and knowledge (Reagans, McEvily, 2003). The E-I index is used as the measure of homophily; the value of the index ranges from 1 to -1 (1 being totally heterophilous and -1 totally homophilous).

Moreover, we used a graphical visualization of the networks to represent a major feature of SNA: trough sociograms. These are useful to show graphically relevant information in the network (such as subgroups, relevant positions, more and less prominent members).

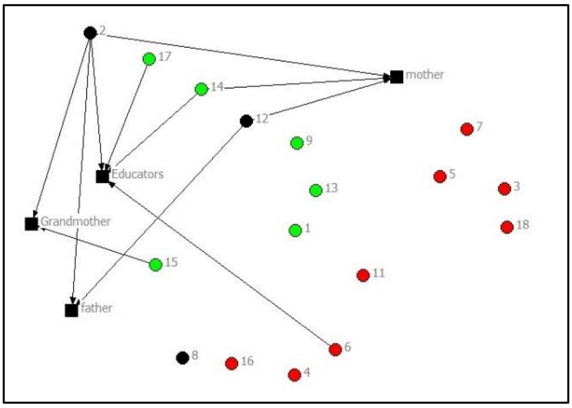
3.4. *Results*

Risk of school leaving: "Have you ever thought of leaving school?

Sixty-five percent (65%) of students declared that they had thought about abandoning school. Students justified their response as being due to relational difficulties in the classroom, demotivation and low consideration of the usefulness of schooling for their life or to find a job.

1. Egocentric network: External support received for school:

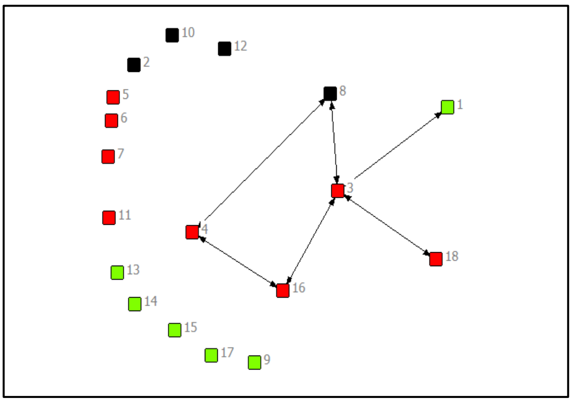
In general, students received low support at home from their family during homework, and was particularly lacking in students at risk of school dropout. Students most frequently received support from mothers and educators (graph 1).



Graph 1: External support received by the students: red circle = student at risk; green = not at-risk student; black = student who didn't give a response to question 3.1

2. Complement network design within the class group:

3.1 Network before beginning school: Most of the students did not know each other from before beginning school (Graph 2).

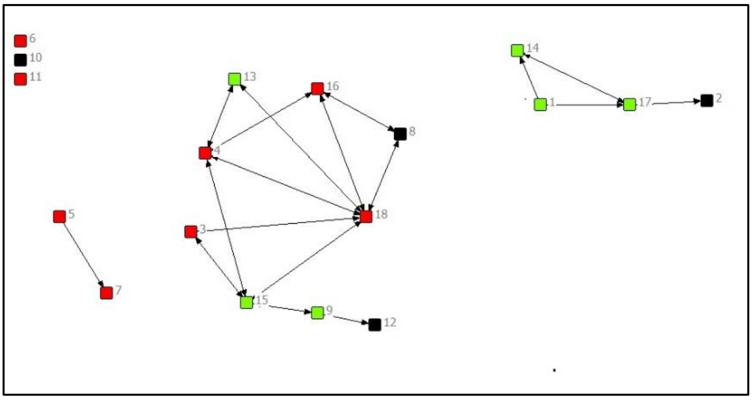


Graph 2: Network before beginning school: red circle = student at risk; green = student not at risk; black = student who didn't give an answer to question 3.2 (a)

3.2 Network during school: results show low density of the network: question B1= 0.06; SD=0.24; question B2= 0.06; sd=0.25. The degree of centrality is also low: question B1=0.14; question B2= 0.2.

We calculated the likelihood of engaging in a relationship based on dyadic similarity (cf. Homophily), considering the attribute "risk of drop out". Results show slightly a tendency to homophily (E-I index= -0.28) in answers to question B2 ("Who do you attend even outside the school?").

The network constituents after beginning school are represented in Graph 3.



Graph 3: Network after beginning school: red circle = student at risk; green = student not at risk; black= student who didn't give a response to question 3.2 (b)

3.5. *Discussion of results*

RQ1: How is structured the supportive network of the risk students outside the school?

Our analysis was conducted in a class where the majority of students were not on track with their school attendance and most had experienced school failure in previous years; looking at the results shown in Graph 1, it is possible to see how the structure of the support network outside the school is very dispersed in general, and how it is almost entirely absent for students who consider leaving school. This result is in line with previous research, which highlight that students at risk of school failure have usually poor social support (Richman, Bowen, 1997). Family has a very strong impact on the possible scholastic success or failure of the children; Richman et al. (1998), have analyzed the social support provided to high school students, identified by school as at risk of school failure. Researchers found that when students at risk of failure received from adult caretakers a positive support, they were able to reach good school performance, with positive consequences for their learning outcome. The risk factor was therefore mitigated by the presence of functional and supportive external social network.

RQ2: How is structured the social network of the risk students, inside the class?

The results show that most of the boys did not know each other before starting school, and therefore the social network currently formed in the class was built after the start of school. The network shows that relationships between the students are not close, connections are not high, there are subgroups and some students are isolated. There are some more popular students who have more numerous connections, but in general there are no positions of prestige or leadership. The results therefore show a group where members have low interdependence and lack cohesion within a group. Prior research who examined the social relationship of the students at risk inside high school, highlights that the social integration was a protective factor for the drop out: students well integrated with peers, with dense and more centralized social network, were less likely to drop out; by contrast students isolated were more likely to early school leaving (Staff, Kreager, 2008; South et al., 2007).

Furthermore, our results show that students who have considered leaving school show a slightly tendency to associate outside school with classmates who have the same idea. This result highlights just a trend that should be deeper investigated. Making a comparison with previous research who have analyzed the homophily in student social networks, our result is in line with the evidences that at-risk students tend to bond with other students at risk, increasing significantly the like hood of drop out of high school (Ellenbogen, Chamberland, 1997; Ream, Rumberger, 2008).

Conclusion

Social Network Learning Analytics has been used to study the networks inside groups of students, detecting relevant influences on at-risk students, analyzing how they are connected into the network. The combination of data collection on a social network, with statistics and visualizations of connections through graphs, can increase the understanding of the learning networks, highlighting what constitutes a learning bond and how learning links are activated and supported (Dawson, 2008).

Social Network Analysis (Dominguez, Hollstein, 2014) makes it possible to identify structures of affiliation and reciprocity that are established during school activities, in order to develop action research projects that elaborate activities that promote cohesion, and the exchange of knowledge among students. Through social network analysis, it is possible to investigate how students are affected by or use their social networks in activities related to education.

Our preliminary results highlight as the "structure of the network" is a variable of analysis that can provide important information for the study of the phenomenon of early school leaving, and address educational intervention. This variable should be deeper in the future research, in order to collect comparative data from students with different background and students coming from different educational institutes.

Social networks can be an important source for students to support themselves, affecting individuals' behaviors and attitudes. However, networks do not always have a positive function. When networks are poorly connected, and a strong relational structure is missing, the sense of belonging to a learning community is lost; if the external networks in the social context are lacking, there is no drive and containment for children in difficulty who run the risk of losing their way. If the relationships within the network are dysfunctional, there is the risk that nonfunctional learning behaviors will be enhanced and strengthened.

Practitioners and teachers should include social network information to plan intervention strategies to support learning pathway of students at risk. It is important to understand the relational structure of the student social network, and work subsequently to build cohesive group, by stimulating a positive interdependence (Brown, Campione, 1990). This suggests a rethinking of teaching practices and the design of learning paths that can foster functional network structures to achieve academic success, through the active involvement of students. Moreover, educational intervention should develop strategies to increase the effectiveness of resources presented in the external environment system, individuating the source available and the reference caretakers, even if this process is often difficult.

References

Boaler J., Staples M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record*, *110*(3), 608-645.

Borgatti S.P., Mehra A., Brass D.J., Labianca G. (2009). Network analysis in the social sciences. *Science,* 323(5916), 892–895.

Borgatti S.P., Halgin D.S. (2011). On network theory'. *Organization Science*, 22 (5), 1168–81.

Borgatti S.P., Everett M.G., Freeman L.C. (2002). *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.

Bronfenbrenner U. (1995). The bioecological model from a life course perspective: Reflections of a participant observer. In P. Moen, G. H. Elder, Jr., K. Lüscher (Eds.), *Examining lives in context: Perspectives on the ecology of human development* (pp. 599-618). Washington, DC, US: American Psychological Association.

Brown A.L., Campione J.C. (1990). Communities of learning and thinking, or a context by any other name. *Human Development*, 21, 108-125.

Bruner J. (1986). *Actual minds. Possible worlds*. Cambridge: Harvard University Press

Buckingham Shum S. (2012). *Learning analytics. UNESCO policy brief*. Retrieved from http://iite.unesco.org/pics/publications/en/files/3214711.pdf

Dawson S. (2008). A study of the relationship between student social networks and sense of community. *Journal of educational technology & society*, *11*(3), 224-238.

Dawson S. (2010). "Seeing" the learning community: An exploration of the development of a resource for monitoring online student networking. *British Journal of Educational Technology, 41*(5), 736-752.

Dawson S., McWilliam E. (2008). Investigating the application of IT generated data as an indicator of learning and teaching performance. *Queensland University of Technology and the University of British Columbia*, *41*.

Domínguez S., Hollstein B. (2014). *Mixed methods social networks research: Design and applications* (Vol. 36). Cambridge University Press.

Ellenbogen S., Chamberland C. (1997). The peer relations of dropouts: a comparative study of at-risk and not at-risk youths. *Journal of Adolescence*, 20 (4), 355-367.

Ferguson R. (2014). Learning Analytics: fattori trainanti, sviluppi e storie. *Italian Journal of Educational Technology*, *22*(3), 138-147.

Ferguson R., Shum S.B. (2012). Social learning analytics: Five approaches. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (LAK '12), 29 April–2 May 2012, Vancouver, BC, Canada (pp. 23-33). New York: ACM.

Gee J. (2004). *Situated Language and Learning: A Critique of Traditional Schooling*. Routledge.

Gutierrez K.D., Rogoff B. (2003). Cultural ways of learning: Individual traits or repertoires of practice, *Educational Researcher* 32 (5), 19-25.

Haythornthwaite C., De Laat M. (2010). Social networks and learning networks: Using social network perspectives to understand social learning. In *Proceedings of the 7th international conference on networked learning* (pp. 183-190). Aalborg, Denmark: Lancaster University.

Legislative Act 76 of 15 April 2005 "Definizione delle norme generali sul diritto-dovere all'istruzione e alla formazione ai sensi dell'articolo 2, comma 1, lettera c, della legge 28 marzo 2003, n. 53".

Lumbelli L. (1989). Pedagogia sperimentale e ricerca esplorativa, *Scuola e città*, 3, 97-104.

Moreno J.L. (1934). *Who shall survive: A new approach to the problem of human interrelations.* Washington, DC, US: Nervous and Mental Disease Publishing Co.

O'Connor M.C., Michaels S. (1996). Shifting participant frameworks: Orches-

trating thinking practices in group discussion. In Hicks (Ed.), C*hild Discourse & Society Learning: An Interdisciplinary Perspective* (pp. 63-103). Cambridge: Cambridge University Press.
Osterman K.F. (2000). "Students' need for belonging in the school community'. *Review of Educational Research*, 70 (3), 323-67.
Reagans R., McEvily B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative science quarterly*, *48*(2), 240-267.
Ream R., Rumberger R. (2008). Student engagement, peer social capital, and school dropout among Mexican American and non-Latino white students. *Sociology of Education* 81, 109-139.
Richman I.M., Bowen G.L. (1997). School failure: An ecological-interactional-developmental perspective. In M.W. Fraser (Ed.), *Risk and resilience in childhood: An ecological perspective* (pp. 95-116). Washington, DC: Press.
Richman J.M., Rosenfeld L.B., Bowen G.L. (1998). Social support for adolescents at risk of school failure. *Social work*, 43(4), 309-323.
Robins G., Pattison P., Kalish Y., Lusher D. (2007). An introduction to exponential random graph models for social networks. *Social Networks* 29, 173-191.
Sarazin M. (2017). Can student interdependence be experienced negatively in collective music education programmes? A contextual approach. *London Review of Education*, 15, 488-504.
Siemens G. (2005). Connectivism: A learning theory for the digital age. *International journal of instructional technology and distance learning 2*, 1, 3-10.
Siemens G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, *57*(10), 1380-1400.
Siemens G., Long P. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review 46*, 5, 30-32.
Siraj I., Mayo A. (2014). *Social class and educational inequality: The impact of parents and schools*. Cambridge University Press.
Skyler J., Bruce A. (2010). Inferential Network Analysis with Exponential Random Graph Models. *Society*, 11(3), 2008, 224-238.
South S., Haynie D., Bose S. (2007). Student mobility and school dropout. *Social Science Research*, 36, 68-94.
Staff J., Kreager D. (2008). Too cool for school? Peer status and high school dropout. *Social Forces*, 87, 445-471.
Van Steen, M. (2010). An Introduction to Graph Theory and Complex Networks. *Copyrighted material.*
Wasserman S., Faust K. (1994). *Social network analysis: Methods and applications*. New York: Cambridge University Press.