Chapter 4: Cacophony of Networks in Learning Analytics

Oleksandra Poquet,[1] Srećko Joksimović[1]

[1] Centre for Change and Complexity in Learning (C3L), University of South Australia, Adelaide, Australia

DOI: 10.18608/hla22.004

ABSTRACT

Network analysis, a suite of techniques to quantify relations, is among core methods in learning analytics (LA). However, insights derived from the application of network analysis in LA have been disjointed and difficult to synthesize. We suggest that such is due to the naïve adoption of network analysis method into the methodologies of measuring and modelling interpersonal activity in digital learning. This chapter describes the diversity of empirical research using network analysis as a cacophony of network approaches. Focusing on LA studies that evaluate social behavior of individuals or model networks, the chapter exemplifies aspects of the analytical process that require rigor, justification, and alignment to overcome the cacophony of empirical findings. The chapter argues that the clarity of network definitions, hypotheses about network formation, and examination of the validity of individual-level measures are essential for coherent empirical insights and indicators. Future work should also make effort to model the temporal nature, multiplex ties, and dynamic interaction between the levels where interpersonal interactions unfold.

Keywords: Online networks, digital traces, digital learning, network modelling

Learning analytics (LA) have been a part of the scholarly discourse now for almost a decade. LA scholarship continues to mature, and institutional adoption of LA is on the rise. Against this backdrop, researchers are urged to demonstrate how LA impacts the practices of teaching and learning [10]. Addressing such a call for impact today is feasible in some areas of LA, such as predictive modelling, writing analytics, and analytics of self-regulation processes. Their applications in LA have been used across diverse technologies, courses, and institutions, and can provide insights to inform teaching and learning practices. However, some areas of LA have not advanced to offer the trusted insights.

Among areas in need of refinement and rigor is that of social learning analytics, here defined as the analysis of interpersonal activity in digital learning. The interest in social learning analytics is driven by the premise that social learner activity contributes to the quality of learning and student experiences. Among mainstream approaches used to examine online interactions is network analysis, a suite of techniques for analyzing relations between objects. LA studies that have used network analysis to understand interpersonal activity offer limited insight, as their disjointed empirical findings are difficult to synthesize. This chapter argues that this lack of coherence is due to the complexity of analytical decisions that arise on the intersection of network analysis and LA. This chapter critically discusses extant LA studies that apply network analysis and highlights aspects of the analytical process that require rigor, justification, and alignment across diverse cases.

1 FOUNDATIONS OF NETWORK STUDIES IN LEARNING ANALYTICS

Analysis of learner networks and social network analysis (SNA) has been adopted in LA since the first LAK conference. In 2010, online teaching practices centered on learner-to-learner interactions via educational technology and web 2.0. Early LA studies built on student retention research in higher education, where social aspects of learning such as social integration, social capital, and the sense of belonging were emphasized [46, 54]. Within such a context, networks constructed from digital data in learning environments could capture social interactions between learners and potentially improve social aspects of university experience.

Analysis of social learning in digital settings was enabled by the social scientists whose tools, examples, and conceptualizations are widely used in LA. SNA as an approach for the analysis of social relations has a long-standing tradition in social sciences [14]. SNA differs from other approaches that analyze randomly sampled individual-level observations. Instead, SNA quantifies patterns within the sets of interdependent relations. Research on social networks, where network ties represent self-reported relationships between people, is widely used in LA, drawing on SNA insights about social structures, theories of how they evolve, and SNA techniques [57].

SNA has played a prominent role in learning sciences, offering tools to understand activities and social processes that students and teachers engage with in technology-mediated settings. For instance, Haythornthwaite [23, 22] analyzed types of exchanges and types of media that support collaboration, socializing, and emotional support in an e-learning environment. Haythornthwaite examined networks of online interactions, where ties represented interpersonal activity captured online, not the self-reported relations between individuals as common in SNA. Early LA work navigated between the insights and interpretations from SNA research towards social structures gleaned from digital relational data. Dawson [8] examined to what extent position of centrality in a network of learners was associated with beneficial learning outcomes, such as individual's sense of belonging. The hypothesis linking network position with benefits reflected the prevailing understanding from the SNA literature that centrality to the network, i.e. positioning within the network ties, can be associated with enhanced access to resources and information [13, 18].

Computer-supported collaborative learning and networked learning also influenced LA network research. De Laat et al. [35] suggested to integrate network analysis that reflects who talks to whom with content analysis that reflects what they are talking about. De Laat et al. [35] utilized SNA to reveal the most influential participants in learning discussions and to explain patterns of connections between the peers. The authors further applied a qualitative coding scheme for analyzing negotiation of meaning and social construction of knowledge. Haythornthwaite and De Laat [23, 22] explored the intersection of learning and social structures, discussing various possibilities for what could constitute a tie in a learning network. They also proposed analytical questions that SNA can explore in learning settings, such as 'who learns from whom', 'what learners learn from each other', 'what kinds of interactions happen between people who learn together', 'which directions do resources flow', 'how frequently do learning interactions happen', and 'how important are they for people involved' (p.354).

To summarize, from the early studies in LA, to interpret patterns in digital interaction networks, researchers borrowed the constructs derived from SNA and learning sciences. To this end, they often contextualized observed digital data by complementing it with other information, such as types of media used for interaction [21], self-reported instruments [8], and content of what was exchanged online with interaction trace data [35].

This link between digitally mediated interactions and their interpretations borrowed from SNA remained in LA network studies. To maintain the distinction between social relations and digitally mediated interactions, we will use SNA to refer to the studies of social networks, i.e. where ties are operationalized as self-reported relational states between people, such as 'trust' or 'friendship'. We will use network analysis to refer to the studies of other networks. Since networks, also known as graphs, can include any objects, or nodes, linked by any relations, or edges [61], LA has adopted network analysis to analyze diverse data sources. Analytical techniques and method-related principles that quantify patterns in a graph are the same, regardless of the network type. Studies of social digital environments in LA that analyze relational data are not limited to social networks, and include networks of learner interaction, text networks, networks of individual clickstream activity, or networks of curriculum modules, among others.

2 NETWORK STUDIES IN LEARNING ANALYTICS

Today, a large portion of network analysis in LA is geared towards a better understanding of the social aspects of the student experience and their relevance for learning and student success. Digital traces of interactions in socio-technical systems have been collected in a vast variety of settings. Some studies have examined university online courses [16] where groups of learners are bounded by similar motivation, similar curriculum trajectories, and likely higher homogeneity in prior knowledge. Other studies focused on MOOCs [28] where learners heterogeneous in their motivation and prior knowledge are bounded by course enrolment, but their patterns of social participation and commitment are fluid [45]. Network analysis has also been applied to open-ended social contexts where group boundaries are ill-defined, to inquire into informal learning in open Internet communities [20, 34]. Finally, network analysis has gained prominence in social text- and video-annotation contexts [24, 36] where artefacts that mediate student learning are explicit and have affordances of their own. Artefact-driven social contexts have often been analyzed using two-mode networks where artefacts and learners are equal actors shaping the structure of interactions [26].

In a digital learning setting, network analysis makes use of the patterns of relations between individuals and artefacts. For instance, network analysis can derive node-level metrics, such as describing the position and a relative importance of a node (a person, a word, a web page in the course, or other) in a network. Alternatively, network analysis can reveal closely interconnected groups of nodes, or provide network-level metrics that describe the entirety of the network structure. Research questions that network analysis can address can be broadly classified, though not limited to: (1) What is the relationship between node characteristics, node positioning, and the outcomes of such a position; (2) Why ties form, i.e. what mechanisms generate observed network structure; and (3) How node attributes influence network formation, as well as how network structure impacts node attributes.

LA studies have addressed the entire spectrum of such network analytical questions. For example, node-level analyses in LA examined how individual positioning captured through network centralities relates to performance and learning-related outcomes in a co-enrolment network [15]; or how a position in a communication network relates to learner linguistic properties [11]. Sub-graph analyses have been prominent in bipartite networks (i.e. where nodes are of two types). In such studies, researchers can detect learner communities based on engagement patterns [26] and identify clusters of learners based on similarity in learning and social activities [24]. Network-level studies have provided metrics to describe structures that represent interactions in different technological and pedagogical contexts [5, 6]. In addition, network-level analyses are applied in epistemic network analysis (ENA, see [48]), a particular methodology that represents epistemic views of individuals and groups as network structures to demonstrate similarities and differences between them. Using network-level studies in LA, researchers also have statistically modelled online learner networks to describe the mechanisms that can explain what drives the formation of network ties [29, 45].

3 CACOPHONY OF NETWORKS IN LEARNING ANALYTICS

These diverse examples show how flexible network analysis can be. The intuition for network analysis is, in part, responsible for its naive applications. That is, any set of relations can be viewed as a graph, and network tools will provide metrics describing them. The problems may begin when the metrics from network analysis are used to interpret indicators, constructs, or processes related to learning. In these instances, network analysis is no longer just a tool, but becomes a methodology with its own theoretical assumptions. Such assumptions include an understanding of what networks represent, but these assumptions are often implicit within the research choices.

Insufficient attention to the assumptions underlying research design can result in the naïve adoption of network analysis [37]. In our view, LA studies often take up network analysis without reflecting on the methodological decisions associated with it. The danger of naïve adoption is that the results are then interpreted through eclectic claims potentially incompatible with the design of the study [59]. Put simply, as methodologies of applying network analysis are not explicit, it is difficult to draw any conclusions as to the meaning of the metrics, even before metrics can be compared across different studies. We refer to this problem as the cacophony of network approaches in LA. We use cacophony to contrast this development with the notion of productive multivocality [53] where diverse disciplines with divergent views build upon one another to produce complementary insights.

Cacophony of findings in network studies results from the misalignment between network construction, analysis choices, and interpretations, impacting generalizability. To highlight areas of misalignment, we distinguish between (1) using network analysis as a method to reduce high-dimensional data and (2) using network analytical methodologies to understand socially shared communication and interpersonal activity in learning settings. When network analysis is a methodology, network construction, metrics and ways of modelling, as well as metric interpretations are at risk of misalignment. By discussing how LA studies evaluate social behavior of individuals and model networks in their entirety using network analysis methodologies, we outline areas where caution is needed and suggest potential ways forward.

4 NETWORK ANALYSIS AS A METHOD

Network analysis as a method summarizes relational data, without particular theoretical meaning assigned to the metrics. The method quantifies relational patterns and identifies clusters based on the relations between the observations of interest. These relations are, at least in part, interdependent, and node-level metrics quantifying them are often non-normally distributed. In LA, nodes linked by relationships can be people, words, learning resources, types of learning behavior captured through clickstream data, topics in the course, and similar. Applying network analysis techniques to these data can reduce its dimensionality and classify nodes. For instance, Joksimovic et al. [30] utilized community detection in networks of words to identify topics discussed in the course. Sirbu et al. [50] deployed ‘coherence network analysis’ to group learners based on the similarity in the linguistic properties of their discourse. Van Labeke and colleagues [56] used network techniques applied to text networks to help identify text quality for automated essay analysis. Besides applying graph analytical techniques to text networks, graph analytical techniques have been shown useful in analyzing relations between clickstream data. For example, Matcha et al. [39] demonstrate that learning strategies can be detected from networks of learner-level clickstream data, where ties between events represent co-occurrence of learning actions.

5 NETWORK ANALYSIS AS A METHODOLOGY

The challenges associated with network studies in LA come through when networks are used to represent socio-technical systems in learning environments. As we argue throughout this section, this shift from representing relational data as a network to representing a theorized construct as a network transforms network analysis from a method to a methodology. The way ties, and therefore, the entire network, are defined, may not work well with the metrics selected by the researchers. Chosen statistical models, i.e. hypothesized generative mechanisms that underpin statistical network analysis, may also be at odds conceptually with the chosen representation. Finally, the theory used to interpret the metrics may also be only in part relevant to the analyzed network.

5.1 Network Construction Issues

Naive adoption of network analysis in LA starts with naive network construction. When network ties, nodes, and boundaries are arbitrary, so are the selected data points, networks metrics derived from them, and their interpretations. Wise and colleagues [58] and Fincham et al. [12] show the variation that results from identical analyses of differently constructed online learner networks. Decisions about network constructions should be theory-based and systematic, and '... a network model should be viewed explicitly as yielding a network representation of something' [2, p. 2]. A close relationship between theoretical definition and interpretation "commits one to assumptions about what is interacting, the nature of that interaction, and the time scale on which that interaction takes place' [3, p. 416]. To align parts of the network analysis methodology, analyzed phenomenon needs to be theorized through literature, abstracted to the network concept, and represented in the network data through theorized and systematic definition of ties and non-ties, nodes, and network boundaries (for guidelines, see [27]).

Networks where ties represent students responding to one another may only to some extent overlap with social networks between interacting students. Therefore, a large degree of caution is required when networks of student communication are interpreted using SNA theories. More complex tie operationalizations, such as aggregating interactions across different types of exchanges, across longer periods of time, or as validated by self-reported measures of affect, may be a better fit to provide insights about social networks from digital data. For instance, Gruzd & Haythornthwaite [19] only include ties between the learners who address one another by names or nicknames. Poquet et al. [44] includes interactions only between learners who sustain participation over a longer period of time. Goggins et al. [17] and Suthers [52] combine information about where, when, or why interpersonal interactions took place, using diverse clickstream information, with semantic similarity between the text, to derive the presence of a tie between learners.

5.2 Choosing and Interpreting Centrality Measures

Learner centrality metrics, i.e. node-level metrics derived from ties in the network describe learner position in relation to others within a network. In LA, measures of learner centrality (e.g. degree, betweenness, closeness centralities, among others) are often contrasted with other process indicators or final assessment results [7]. Researchers also have investigated the relationship between learner centrality in communication and co-enrolment networks with measures of perceived belonging [8], creativity [9]), social capital [28], and discourse features [11].

These studies, however, often are conducted on networks where ties are operationalized differently. Beyond these issues of validity, the misalignment in research design can occur when network measures and their interpretation embed SNA assumptions, but the specific network representation does not afford those assumptions. To explain, we can look at measures of degree, betweenness and closeness centrality. The premise that learner network position, captured through the centrality, is associated with particular benefits stems from SNA. In social networks, an individual's position represents access to resources, such as information flow or support [1]. In SNA, degree centrality, a local measure of centrality that takes into account the number of connections an individual has, is equivalent to the number of social relationships an individual has. LA studies use degree as a measure of popularity, influence or capital, transferring interpretations of centrality that assume that ties represent relationships. But the interpretation for centrality in online settings can be different from that in social networks. Based on an empirical experiment, Poquet et al. (2020) modelled online interaction networks to demonstrate that degree centrality in online learner interactions is associated with in-course individual-level activity, rather than social choices made by learners. The authors use empirical simulations to claim that centrality is merely a proxy of individual learner characteristics and not of social dynamics, as is in SNA.

Interpretation of betweenness and closeness centrality measures in online settings is even further away from their use in SNA. Their use in learner interaction networks can be controversial not only in interpretation, but the metric itself may be inapplicable. Centrality measures such as betweenness and closeness are distance-based, i.e. the formula takes into account the entire network structure. For instance, betweenness centrality is derived from the number of shortest paths that go through each node. In SNA interpretation of this measure presumes that the absence of ties is equivalent to the absence of access. Hence, in SNA nodes with high betweenness can be interpreted as having privileged access to resources. Online interaction networks are constructed from event data where ties are transient events (e.g. A replied at time X) not relational states (e.g. A is friends with B). The absence of a tie in the context of ties as events does not imply limits of access. Therefore, distance between the individuals in the network and uniqueness of positioning (embedded in the measure) in communication environments is not at all equivalent to its SNA counterpart, or its interpretation.

5.3 Comparing, Interpreting, and Modelling Networks

Further challenges arise when network-level studies are conducted. Research questions asked at a network level can describe network structures and mechanisms generating them (e.g. [4, 29, 60]). This becomes useful because a network structure can serve as group-level indicator caused by a specific pedagogical and technological setting [5, 42] or as a signal of desired outcomes, such as team's performance [43]. In such network-level studies, methodological flaws can easily occur (1) when researchers directly compare descriptive networks metrics from different settings, and (2) when they use hypothesis from SNA theory to model how socio-technical networks in learning environments form.

For instance, researchers conduct descriptive analysis of several courses in the same study, and then descriptively compare their network metrics, such as density (overall interconnectedness of the graph), transitivity (presence of triads in the graph) or centralization (reliance of graph connectivity on one or several nodes). Such studies then commonly report that as the course progresses interaction networks increase in the number of connections between the individuals (density), in reciprocation between pairs (reciprocity), and in triad formation (transitivity) [33, 41, 55, 62]. The challenge arises when researchers start explaining forces behind these metrics. A network in course A may have evolved from a different generative mechanism than in course B. This implies that network density observed in course A may have been random, whereas network density observed in course B may have been beyond chance. Descriptive cross-network comparisons do not provide this information.

Comparing descriptive indicators across networks requires statistical analyses that rely on the so-called null models that explain how socio-technical networks form. Null models are random networks simulated using hypothesized generative rules, such as 'learners are likely to respond to those who interacted with them earlier' or 'learners interact on a given day when the assignment was posted'. These generative principles should explain why networks form in digital settings, derived from the theories about digital learning and social processes. By comparing observed network to the distribution of random networks generated from the null model, a researcher can interpret if density, transitivity, or any other network measure appears in the observed network by chance or resulted from some particular influences. Many different approaches exist to how null models can be generated, such as tie permutation [40], exponential random graph modelling (ERGM) [38], stochastic actor-oriented modelling [51], among others approaches to network reconstruction [25].

LA largely lacks validated null models that explain how networks form in digital learning environments. Thus far, statistical modelling of networks in digital settings had predominantly used hypotheses derived from why ties form in social networks [29, 12, 60]. For instance, SNA hypothesizes that 'the tie will form between A and C, if A and B as well as B and C are already connected' – based on the principle 'a friend of a friend becomes a friend' observed in social networks. LA researchers can adopt this principle and model online communication network to observe if it describes the random structure, i.e. can explain observed patterns. To demonstrate that these theorized principles can explain formation of ties, researchers need to show that random networks generated by the same principles are similar to the observed network through the goodness of fit plots. Creating network models supplemented with goodness of fit plots would demonstrate where the generative models fail to explain the data. LA studies rarely include such plots for statistical modelling of networks that uses SNA hypotheses. By implication, there is little ground to evaluate how well the models reflect the data.

This highlights the need for formulating and testing generative principles that suit digital learning. Theoretical considerations currently omitted from statistical modelling of digital learning networks include diversity of contexts, as well as lack of attention to time and learner activity level. First, social contexts where LA examines technology-mediated interactions between learners, instructors, online platforms, and course artefacts are markedly different. Given the diversity of social contexts examined in LA, it is likely that the processes generating ties between individuals in them are also different, and theories as to how they form are yet to be put forward. Second, statistical modelling in LA has only recently started to explicitly include temporal aspects of learner activity in socio-technical networks and overall participation levels at the node level (e.g. [4]). Otherwise, researchers used ERGMs to model forum communication as a network of binary ties between the learners, not as a network of valued ties (e.g. where a tie has a value equal to the sum of posts shared between two learners). Excluding information about the weight of ties from a communication network removes some dyadic observations from the modelled data, and therefore, requires a conceptual justification. In light of these shortcomings, current evidence derived from statistical modelling that validates network-level indicators to evaluate socially shared learning and communication can be perceived as limited.

6 FUTURE RESEARCH

The chapter reviewed empirical studies in LA that utilize network approaches. The chapter highlighted the aims of network studies and major caveats associated with them. We emphasize that the researchers who use network analysis as a methodology need to be more explicit about the assumptions they bring from the literature. We call for explicit and rigorous operationalization of networks as phenomena they represent. At minimum, a clear description of network models is needed, to enable further synthesis of insights and prevent naive transfer of interpretations from self-reported network research into the network measures of online learner networks.

Addressing the issues presented throughout the chapter can help constrain LA to better model and understand socially shared learning, with diverse ties and actors at different levels and scales interacting dynamically. That is, learning in socio-technical systems unfolds through temporal interactions between socio-material agents, linked through diverse interactions, and at different levels. A socio-technical view of learning emphasizes that these networks form through mutually interdependent interactions between the artefacts, technology, people, and ideas [31, 32, 49]. Socio-cognitive processes underpinning the diverse interactions drive community development and knowledge building [47]. Knowledge building processes unfold through the interaction of words, topics, themes, social norms stated through discourse, linguistic markers of identity, and similar.

Despite these rich theorizations, current network modelling approaches in LA do not reflect this theoretical richness. A new generation of network studies is needed to use the potential of complex network modelling to integrate dynamic, relational, spatial, multi-level, and multiplex nature of models of social learning with technology. For network analysis methodologies to deliver on the promise for rich insights and indicators to inform about learning, explicit modelling of socio-technical learning processes and better alignment of theory with the methodologies is needed.

REFERENCES

[1] Stephen Borgatti. “Centrality and network flow”. In: *Social Networks* 27.1 (2005), pp. 55–71.

[2] Ulrik Brandes et al. “What is network science?” In: *Network Science* 1.01 (2013), pp. 1–15.

[3] Carter Butts. “Revisiting the foundations of network analysis”. In: *science* 325.5939 (2009). Publisher: American Association for the Advancement of Science, pp. 414–416.

[4] Bodong Chen and Oleksandra Poquet. “Socio-Temporal Dynamics in Peer Interaction Events”. In: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (2020), pp. 203–208.

[5] Bodong Chen et al. “Fostering student engagement in online discussion through social learning analytics”. In: *The Internet and Higher Education* 37 (2018), pp. 21–30.

[6] Regina Collins and Anatoliy Gruzd. “Learning within digital media: Investigating the relationships between student citation networks, assignment structures, and learning outcomes”. In: *Proceedings of Hawaii International Conference on System Sciences*. 2017. DOI: 10.24251/HICSS.2017.253.

[7] Marielle Dado and Daniel Bodemer. “A review of methodological applications of social network analysis in computer-supported collaborative learning”. In: *Educational Research Review* 22 (2017), pp. 159–180.

[8] Shane Dawson. “A study of the relationship between student social networks and sense of community”. In: *Journal of educational technology & society* 11.3 (2008), pp. 224–238.

[9] Shane Dawson, Jennifer Pei Ling Tan, and Erica McWilliam. “Measuring creative potential: Using social network analysis to monitor a learners’ creative capacity”. In: *Australasian Journal of Educational Technology* 27.6 (2011), pp. 924–942.

[10] Shane Dawson et al. “Increasing the impact of learning analytics”. In: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*. 2019, pp. 446–455.

[11] Nia Dowell et al. “Modeling Learners’ Social Centrality and Performance through Language and Discourse”. en. In: *International Educational Data Mining Society* (2015), pp. 250–257.

[12] Ed Fincham, Dragan Gaševic, and Abelardo Pardo. “From Social Ties to Network Processes: Do Tie Definitions Matter?” In: *Journal of Learning Analytics* 5.2 (2018), pp. 9–28.

[13] Linton Freeman. “Centrality in social networks conceptual clarification”. In: *Social networks* 1.3 (1978). Publisher: North-Holland, pp. 215–239.

[14] Linton Freeman. “The development of social network analysis”. In: *A Study in the Sociology of Science* 1 (2004). Publisher: Empirical press, p. 687.

[15] Josh Gardner and Christopher Brooks. “Coenrollment networks and their relationship to grades in undergraduate education”. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. ACM, 2018, pp. 295–304.

[16] Dragan Gasevic, Amal Zouaq, and Robert Janzen. “"Choose Your Classmates, Your GPA Is at Stake!": The Association of Cross-Class Social Ties and Academic Performance”. In: *American Behavioral Scientist* 57.10 (2013), pp. 1460–1479. DOI: 10.1177/0002764213479362.

[17] Sean Goggins, Christopher Mascaro, and Giuseppe Valetto. “Group informatics: A methodological approach and ontology for sociotechnical group research”. In: *Journal of the American Society for Information Science and Technology* 64.3 (2013), pp. 516–539.

[18] Mark Granovetter. “The strength of weak ties: A network theory revisited”. In: *Sociological theory* 1.1 (1973), pp. 201–233. (Visited on 06/03/2015).

[19] Anatoliy Gruzd and Caroline Haythornthwaite. “Automated discovery and analysis of social networks from threaded discussions”. In: *The International Network of Social Network Analysis Conference*. St Pete Beach, FL, USA, 2008. URL: https://www.ideals.illinois.edu/handle/2142/11528.

[20] Anatoliy Gruzd et al. “Learning analytics for the social media age”. In: *Proceedins of the Fourth International Conference on Learning Analytics And Knowledge - LAK ’14* (2014), pp. 254–256. DOI: 10.1145/2567574.2576773.

[21] Caroline Haythornthwaite. “Exploring multiplexity: Social network structures in a computer-supported distance learning class”. In: *The information society* 17.3 (2001). Publisher: Taylor & Francis, pp. 211–226.

[22] Caroline Haythornthwaite and Maarten de Laat. “Social Network Informed Design for Learning with Educational Technology”. In: *Informed Design of Educational Technologies in Higher Education : Enhanced Learning and Teaching*. Ed. by Anders D Olofsson et al. IGI Global, 2012, pp. 352–374.

[23] Caroline Haythornthwaite and Maarten de Laat. "Social Networks and Learning Networks: Using Social Network Perspectives to Understand Social Learning". In: *Proceedings of the 7th international conference on networked learning*. Aalborg, Denmark: Lancaster University, 2010, pp. 183–190.

[24] Tobias Hecking et al. "Using network-text analysis to characterise learner engagement in active video watching". In: *ICCE 2017 Main Conference Proceedings*. Asia-Pacific Society for Computers in Education, 2017, pp. 326–335.

[25] Elizabeth Hobson et al. "A guide to choosing and implementing reference models for social network analysis". In: *arXiv preprint arXiv:2012.04720* (2020).

[26] Urlich Hoppe. "Computational Methods for the Analysis of Learning and Knowledge Building Communities". In: *Handbook of Learning Analytics*. Ed. by Charles Lang et al. First. Society for Learning Analytics Research (SoLAR), May 2017, pp. 23–33.

[27] James Howison, Andrea Wiggins, and Kevin Crowston. "Validity issues in the use of social network analysis with digital trace data". In: *Journal of the Association for Information Systems* 12.12 (2011), p. 2.

[28] Srećko Joksimović et al. "Exploring development of social capital in a CMOOC through language and discourse". In: *The Internet and Higher Education* 36 (2018), pp. 54–64.

[29] Srećko Joksimović et al. "Translating network position into performance: Importance of centrality in different network configurations". In: *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. LAK '16. Edinburgh, United Kingdom. ACM, 2016, pp. 314–323.

[30] Srećko Joksimović et al. "What do cMOOC participants talk about in social media? A topic analysis of discourse in a cMOOC". In: *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. 2015, pp. 156–165.

[31] Christopher Jones. *Networked Learning: An Educational Paradigm for the Age of Digital Networks*. Springer, 2015.

[32] Victor Kaptelinin and Bonnie Nardi. "Affordances in HCI: toward a mediated action perspective". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 967–976.

[33] Carmel Kent, Amit Rechavi, and Sheizaf Rafaeli. "Networked Learning Analytics: A Theoretically Informed Methodology for Analytics of Collaborative Learning". In: *Learning In a Networked Society*. Springer, 2019, pp. 145–175.

[34] Kirsty Kitto et al. "Learning Analytics Beyond the LMS: The Connected Learning Analytics Toolkit". In: *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. LAK '15. Poughkeepsie, New York: ACM, 2015, pp. 11–15. DOI: 10.1145/2723576.2723627.

[35] Maarten de Laat et al. "Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis". In: *International Journal of Computer-Supported Collaborative Learning* 2.1 (2007), pp. 87–103. DOI: 10.1007/s11412-007-9006-4.

[36] Alwyn Vwen Yen Lee and Seng Chee Tan. "Promising ideas for collective advancement of communal knowledge using temporal analytics and cluster analysis". In: *Journal of Learning Analytics* 4.3 (2017), pp. 76–101.

[37] Kristine Lund and Dan Suthers. "Le déterminisme méthodologique et le chercheur agissant". In: *Éducation & didactique* 10.1 (2016), pp. 27–37.

[38] Dean Lusher, Johan Koskinen, and Garry Robins. *Exponential random graph models for social networks: Theory, methods, and applications*. Cambridge University Press, 2012.

[39] Wannisa Matcha et al. "Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches". In: *European Conference on Technology Enhanced Learning*. Springer, 2019, pp. 525–540.

[40] Mark Newman. *Networks*. Oxford university press, 2018.

[41] Fan Ouyang and Yu-Hui Chang. "The relationships between social participatory roles and cognitive engagement levels in online discussions: Relationships on social and cognitive engagements". In: *British Journal of Educational Technology* 52.3 (July 2018), pp. 1396–1414.

[42] Fan Ouyang and Cassandra Scharber. "The influences of an experienced instructor's discussion design and facilitation on an online learning community development: A social network analysis study". In: *The Internet and Higher Education* 35 (2017), pp. 34–47.

[43] Andrew Pilny et al. "An illustration of the relational event model to analyze group interaction processes." In: *Group Dynamics: Theory, Research, and Practice* 20.3 (2016), p. 181.

[44] Oleksandra Poquet, Shane Dawson, and Nia Dowell. "How effective is your facilitation?: Group-level analytics of MOOC forums". In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. LAK '17. ACM, 2017, pp. 208–217. DOI: 10.1145/3027385.3027404.

[45] Oleksandra Poquet, Liubov Tupikina, and Marc Santolini. "Are forum networks social networks? A methodological perspective". In: *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK '20)*, Frankfurt, Germany, 2020. DOI: https://doi.org/10.1145/3375462.3375531sharma.

[46] Alfred Rovai. “Building Sense of Community at a Distance”. In: *The International Review of Research in Open and Distributed Learning* 3.1 (2002), pp. 1–16. DOI: 10.19173/irrodl.v3i1.79.

[47] Marlene Scardamalia and Carl Bereiter. “Computer Support for Knowledge Building Communities”. In: *CSCL: Theory and Practice of an Emerging Paradigm*. Ed. by Timothy Koschmann. Malwah, New Jersey: Lawrence Erlbaum Associates Inc. Publishers, 1996.

[48] David Williamson Shaffer, Wesley Collier, and Andrew R. Ruis. “A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data”. In: *Journal of Learning Analytics* 3.3 (2016), pp. 9–45.

[49] George Siemens. “Connectivism”. In: *A Learning Theory for the Digital Age. In eLearnSpace* (2004).

[50] Maria-Dorinela Sirbu et al. “Exploring online course sociograms using cohesion network analysis”. In: *International Conference on Artificial Intelligence in Education*. Springer, 2018, pp. 337–342.

[51] Tom Snijders, Gerhard van de Bunt, and Christian Steglich. *Introduction to stochastic actor-based models for network dynamics*. Vol. 32. 1. 2010, pp. 44–60.

[52] Dan Suthers. “From contingencies to network-level phenomena: Multilevel analysis of activity and actors in heterogeneous networked learning environments”. In: *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*. 2015, pp. 368–377.

[53] Dan Suthers. *Productive Multivocality in the Analysis of Group Interactions*. Ed. by Dan Suthers et al. Boston, MA: Springer US, 2013. ISBN: 978-1-4614-8959-7. URL: http://link.springer.com/10.1007/978-1-4614-8960-3.

[54] Vincent Tinto. “Dropout from higher education: A theoretical synthesis of recent research”. In: *Review of educational research* 45.1 (1975), pp. 89–125.

[55] Adrienne Traxler, Andrew Gavrin, and Rebecca Lindell. “Networks identify productive forum discussions”. In: *Physical Review Physics Education Research* 14.2 (2018), p. 020107.

[56] Nicolas Van Labeke et al. “OpenEssayist: extractive summarisation and formative assessment of free-text essays”. In: *International Workshop on Discourse-Centric Learning Analytics*. 2013.

[57] Stanley Wasserman and Katherine Faust. *Social network analysis: Methods and applications*. Vol. 8. Cambridge university press, 1994.

[58] Alyssa Wise, Yi Cui, and Wan Qi Jin. “Honing in on social learning networks in MOOC forums: examining critical network definition decisions”. In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (2017), pp. 383–392.

[59] Stephen Yanchar and David Williams. “Reconsidering the compatibility thesis and eclecticism: Five proposed guidelines for method use”. In: *Educational researcher* 35.9 (2006), pp. 3–12.

[60] Jingjing Zhang, Maxim Skryabin, and Xiongwei Song. “Understanding the dynamics of MOOC discussion forums with simulation investigation for empirical network analysis (SIENA)”. In: *Distance Education* 37.3 (2016), pp. 270–286. DOI: 10.1080/01587919.2016.1226230.

[61] Katharina Zweig. *Network Analysis Literacy: A Practical Approach to the Analysis of Networks*. en. Lecture Notes in Social Networks. Wien: Springer-Verlag, 2016. DOI: 10.1007/978-3-7091-0741-6.

[62] Justyna Zwolak et al. “Students’ network integration as a predictor of persistence in introductory physics courses”. In: *Physical Review Physics Education Research* 13.1 (2017), p. 010113.