



The impact of study load on the dynamics of longitudinal email communications among students



Shahadat Uddin ^{a,*}, Kate Thompson ^{b,1}, Beat Schwendimann ^{b,2}, Mahendra Piraveenan ^{c,3}

^a Project Management Program & Centre for Complex Systems Research, The University of Sydney, Room 402, School of Civil Engineering, Redfern, NSW 2006, Australia

^b Faculty of Education and Social Work, The University of Sydney, Australia

^c Project Management Program & Centre for Complex Systems Research, The University of Sydney, Room 334, School of Civil Engineering, Redfern, NSW 2006, Australia

ARTICLE INFO

Article history:

Received 14 August 2013

Received in revised form

12 November 2013

Accepted 14 November 2013

Keywords:

Email communication

Social network analysis

Exponential random graph models

ABSTRACT

With the advent of information technology, emails have gained wide acceptability among students as an asynchronous communication tool. According to the current pedagogy literature the overall trend of the use of email communication by university students has been increasing significantly since its inception, despite the rapid growth of the popularity and acceptability of other social mediums (e.g. Mobile phone and Facebook). In this study, we explore a longitudinal email communication network, which evolved under an increasing study load among 38 students throughout a university semester, using measures of social network analysis (SNA) and exponential random graph (ERG) models. This longitudinal network was divided into three waves, where each wave represents the portion of the complete longitudinal network that evolves between two consecutive observations. An increased study load was imposed through the assessment components of the course. SNA measures of *degree centrality* (i.e. the activity of an actor or actor popularity), *betweenness centrality* (i.e. the capacity to control the flow of information in a network), *closeness centrality* (i.e. reachable to other nodes) and *reciprocity* (i.e. tendency to make reciprocal links) are considered to explore this longitudinal network. ERG models are probabilistic models that are presented by locally determined explanatory variables and can effectively identify structural properties of networks. From the analysis of this email communication network, we notice that students' network positions and behaviours change with the changes in their study load. In particular, we find that (i) students make an increased number of email communications with the increase of study load; (ii) the email communication network become sparse with the increase of study load; and (iii) the 2-star parameter (a subset of three nodes in which one node is connected to each of the other two nodes) and the triangle parameter (a subset of three nodes in which each node is connected to the other two nodes) can effectively explain the formation of network in wave3; whereas, the 3-star parameter (a subset of four nodes in which one node is connected to each of other three nodes) can effectively explain the formation of network in wave1 and wave2. Interpretations of these findings for the monitoring of student behaviour in online learning environments, as well as the implications for the design of assessment and the use of asynchronous tools are discussed in this paper.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Instructors of online learning environments often only see students' final products and miss the ongoing communication processes that occur between students during the course (Race, 2001). The ability to collect data that informs us about the processes of communication and

* Corresponding author. Tel.: +61 02 9351 2118; fax: +61 02 9662 6566.

E-mail addresses: shahadat.uddin@sydney.edu.au (S. Uddin), kate.thompson@sydney.edu.au (K. Thompson), beat.schwendimann@sydney.edu.au (B. Schwendimann), mahendra.piraveenan@sydney.edu.au (M. Piraveenan).

¹ Tel.: +61 2 9351 3773; fax: +61 2 9036 5205.

² Tel.: +61 2 9351 3773; fax: +61 2 9351 5027.

³ Fax: +61 02 9662 6566.

collaboration, as well as the analytic techniques to apply to such data have both expanded over the last five to ten years (Dillenbourg, Järvelä, & Fischer, 2009; Goodyear, Jones, & Thompson, 2014; Hakkarainen, Lipponen, & Järvelä, 2013; Reimann, 2009; Stahl, Koschmann, & Suthers, 2006). Suggestions for methods of analysing processes that show changes in the dynamics of groups have primarily focused on coded data (e.g. Kapur, 2011; Kennedy-Clark & Thompson, 2013; Reimann, 2009) and transferability to real-time feedback to teachers has been limited. In this paper, we present the results of a three-stage analysis of email communication between postgraduate students during a 13-week course. While the information we have about the *learning* is limited, the potential of the method to inform teachers of student progress is evident. The first section will provide an overview of methods of analysing collaborative learning, followed by a longer section outlining the methods used in this paper. The case study will then be outlined, followed by the discussion including an overview of the implications of this method for future work.

2. Background literature

Collaborative learning has been defined as those situations in which the group is expected to “complete the task together through dialogue and joint action” (Goodyear et al., 2014). Working in collaborative settings is increasingly important for both learning and preparation for future workspaces (Dillenbourg et al., 2011). Computer-supported collaborative learning (CSCL) is a situation in which computer technology plays a role (Goodyear et al., 2014), and may refer to a piece of hardware (such as a mobile device or interactive whiteboard), software running on a computer (such as a simulation model or virtual world) or tools accessible through a browser (such as online forums or email). A challenge for any instructor in an online or computer supported collaborative environment is to understand the activity undertaken by learners in response to a given task. Such knowledge can allow instructors to detect any problems early and develop interventions. Without it, instructors may find it challenging to evaluate collaborative processes, such as the symmetry of participation (Dillenbourg, 1999). Orchestration is a term that has recently been used to describe the processes involved in teaching in CSCL environments. Building on the analogy of the conductor of an orchestra, orchestration in CSCL is the process of productively coordinating supportive interventions across multiple learning activities. Moon (2001) uses ‘orchestration’ to refer to “the process of managing a whole learning group in such a way as to maintain progress towards the learning outcomes and improvement of practice for all” (p. 120). Orchestration can involve different forms of coordination, including activities at different social, contextual and media levels; scaffolds at different social levels; self-regulation and external regulation; individual motivation and social processes (Dillenbourg & Jermann, 2010; Dillenbourg et al., in press).

Research that examines the *processes* of learning in collaborative settings has developed in the past decade with advances in data collection techniques, and the availability of ‘big data’ related to the activities of learners (see Goodyear et al., 2014 for a detailed discussion about this). Work in this area has focused on the development of methods of analysis in subject areas such as decision making (Kapur, 2011; Reimann, 2009), knowledge building (van Aalst, 2009), or argumentation (Ding, 2009). Of most relevance to this paper is that which relates to the representation of time and order in analysing processes of learning. Reimann (2009) advocated for event-based analysis that takes into account a group’s history when analysing their behaviour. Thompson and Kelly (2012) make a case for a two-stage approach to this kind of analysis, the first using visualisations to determine phases of group work, and then the application of event-based analysis. These authors suggest that this better accounts for the changes in processes of learning over time.

Methods from social network analysis (Wasserman & Faust, 2003) can be applied to computer-supported collaborative learning environments to identify emerging groups (cliques), and (strong) links between participants, and (weak) links between groups – or the lack thereof (Granovetter, 1973, 1983). Other studies have used social network analysis to analyse email communication (e.g. Kazienko, Musial, & Zrzywa, 2009; Uddin & Jacobson, 2013), examining the node position to determine importance in the network.

Quantitative information about relationships between groups and participants could be helpful for instructors. However, if the quantitative information about students’ interactions is only available after the course, instructors may not use this information about students to change their strategies for subsequent cohorts (Tyler, 2011). Presenting information about students’ interactions to students after the interaction is finished may be too late to affect group processes. To be able to address emerging problems, instructors need access to carefully selected real-time information (Segedy, Sulcer, & Biswas, 2010). The provision of real-time feedback can be useful for students managing their group work, as well as instructors, needing to know whether and when to intervene. A natural marker of progress through a course is the assessments given, as used in this paper to divide the data collected. Kay, Yacef, and Reimann (2007) developed several representations of interactions between students. They were shown to students after the collaborative task, and students were asked to reflect on whether these represented their understanding of the collaborative processes in their group. One representation developed was a wattle diagram, which showed students’ activity in a wiki over time, in terms of work performed and management of group tasks (Kay, Maisonneuve, Yacef, & Reimann, 2006a, 2006b). The analogy of a wattle tree was used, with the main trunk of the tree representing the timeline of student activity (with each day marked), the round yellow flowers represented the wiki-related activity, and orange round flowers represented SVN-related activity. The green leaves represented tickets given to members of the groups, and closed, which indicated the allocation and completion of work. In their analysis, different patterns of activity were identified as the assessment approached. Well organised groups had several, small ‘leaves’ (tasks allocated and closed), between which were the yellow and orange flowers (activity), evenly spaced. Those groups who worked inefficiently, with most work completed during the week when the task was due to be submitted, had a different pattern observed in their wattle diagram. When examined in combination with other interaction diagrams, they were able to gain insight into roles that were adopted in groups, as well as identify social loafers (Kay et al., 2006a, 2006b).

Email has been used for more than a decade as the official means of communication between instructors and students, as well as between students, in university settings. There are few studies that examine student-to-student interactions (the focus of this paper), most studies concentrate on instructor–student interactions. Research into student satisfaction with instructor’s use of email found that students felt that it increased their access to instructors and improved these interactions (e.g. Atamian & DeMerville, 1998; Boles, 1999; Hassini, 2006). In a context analysis of 385 email messages from 42 students, Hassini (2006), found that 33% (the largest proportion) were related to the assignments, followed by the project in which students were expected to participate (26%), followed by grades (21%). Tolmie and Boyle (2000) suggested that asynchronous email exchange facilitated discussion between students, which was seen to be beneficial. Similar to Hassini’s (2006), they found that email was used mostly for exchanging drafts of assignments, as well as updating team members on group

Table 1

Duration of each wave, the distribution of different assessment components over these waves and the relative weight of each assessment component compared to the total mark.

| Wave ID | Week considered | Assessment component (relative weight in % to the final mark) |
|---------|---|--|
| Wave1 | Week1 to Week5 | Ongoing assessment of tutorial activities (except Week1) |
| Wave2 | Week6 to Week9 + 1 week Mid semester vacation | <ul style="list-style-type: none"> • Ongoing assessment of tutorial activities • Online quiz1, due in Week7 (10%) • Presentation of <i>Group Assignment</i>, due in Week9 (5%) • Separate discussions between each group and the lecturer regarding the preparation of the report of <i>Group Assignment</i>, which were also taken under consideration in determining the contribution of an individual to the <i>Group Assignment</i>. |
| Wave3 | Week10 to Week13 + 1 week vacation before the final examination | <ul style="list-style-type: none"> • Ongoing assessment of tutorial activities • Online quiz2, due in Week12 (10%) • Report submission of group assignment, due in Week13 (10%) • Final examination (60%) |

*Ongoing assessment of tutorial activities contributes 5% to the final grade.

progress. Face-to-face meetings were used for ongoing discussions. Email has also been found to be beneficial to student learning. [Absalom and Pais Marden \(2004\)](#) examined the use of email in language learning, including the content of the emails, and found that the expressions used were very similar to those used in a telephone conversation, as if the writer was speaking to the recipient in real-time. They also found that, due to the asynchronous nature of the email communication, learners were allowed to fully develop ideas, even though the language appeared more spontaneous. More recently, research has turned to other asynchronous methods of communication, such as wikis (e.g. [Dohn, 2010](#)) and discussion boards (e.g. [Wise & Chiu, 2011](#)); however, the findings are similar. [Zenios and Holmes \(2010\)](#) also found that these asynchronous collaboration and communication tools could be used for the development of written work.

Measures of communication within a network that can be used for analysis include centrality, direction of causality, levels of analysis, explanatory goals, and explanatory mechanisms ([Borgatti & Foster, 2003](#)) as well as identification of structural holes ([Burt, 1992](#)). Selected quantitative information about students' network structure can provide instructors with valuable information to design effective interventions.

3. Research methods

3.1. The dataset

This study utilises a dataset composed of the network of students' email communications. This dataset does not include the content of the emails or any personal information about the students. The communication network, involving 39 students, evolved over one semester, during which there was an increasing study load on students. The semester consisted of 13 semester-weeks (for lectures and tutorials), one week of mid-semester vacation and one week vacation before the final examination. The masters-degree course, entitled Quantitative Methods in Project Management, was delivered in *face-to-face* mode. In each semester-week students were expected to attend one 1.5 h lecture and one 1.5 h tutorial. The course was maintained by *Blackboard WebCT*⁴ which is a web-based tool for course management and is used by educational institutions for delivering courses through *online* and *face-to-face* modes ([Ngai, Poon, & Chan, 2007](#)). For all course-related communication, students were advised to communicate with other students, the tutor and the lecturer through the designated email communication system of *Blackboard WebCT*. Using this email communication system, a student can send an email to a single recipient or a group of recipients. An email sent to a specific recipient is not accessible by others. The course coordinator can access all emails sent by students, lecturer or tutor. In this study, the course coordinator shared the dataset after applying an encryption algorithm to de-identify sender and receiver email addresses. Informed consent was also given by each student before the start of the semester.

The email communication system of *Blackboard WebCT* was configured in such a way that it allowed students to send any of these four categories of emails – *socialisation*, *general discussion*, *tutorial* and *questions & inquiries*. For this purpose, four sub-folders were created within the email communication systems for the course. Emails that had been created by students for the purpose of socialisation with their peers and course coordinator belong to the *socialisation* category. The *general discussion* emails had been created for the purpose of online brainstorming and knowledge sharing for activities such as the solution of supplementary problems, knowledge sharing of the application of different course materials, discussion about group assignment and online quizzes. The *tutorial* sub-folder contained all emails that were related to the discussion about how to solve tutorial problems. Students had created *questions & inquiries* type of emails for the purpose of making questions or queries regarding online quizzes, group assignments or the final examination to the course coordinator and to their peers.

Students enrolled in this course (i.e. Quantitative Methods in Project Management) as part of a professional qualification degree, entitled *The Master of Project Management*. For this course, the weekly lecture and tutorial time were from 6:00 pm to 7:30 pm (Thursday) and from 7:30 pm to 9:00 pm (Thursday) respectively. These schedules were chosen deliberately since many of the students were engaged in full-time employment. During the same semester, students could also enrol in other course(s), which were required for this professional degree, either in *face-to-face* or *online* mode. Given this information about the students, it is expected that they depended on email communication for most of their course-related discussions such as seeking help, collaborating with others or sharing the answers of weekly tutorial

⁴ <http://www.blackboard.com/>.

questions. Emails that passed a common message to all (e.g. introductory emails sent to all by most of the students in order to introduce themselves to the whole class) were excluded. Those emails that had a single recipient were considered for research analysis purposes as this type of email reflects more intensive and directed communications (Uddin, Hossain, Murshed, & Crawford, 2011; Uddin, Murshed, & Hossain, 2011). After these refinements, the dataset contained 617 emails that were included in the analysis. For sociological analysis, the size of a network is very important, which is usually justified by the level of connectedness among the participating actors (Wasserman & Faust, 2003). A large social network that has a higher number of participating actors (e.g. 2000 actors) may not be suitable for sociological analysis if it has a low level of actors' connectivity (e.g. existence of 0.01% connections compared to all possible connections). On the other hand, sociological analysis can be applied to a small social network that has a lower number of actors (e.g. 10 actors) with a higher level of connectivity (e.g. existence of >10% of connections compared to all possible connections) among participating actors (Wasserman & Faust, 2003). In our research dataset, there are 207 distinct email communications (i.e. unique combinations of senders and receivers) among 39 students. There are 1482 (i.e. ${}^{39}P_2 = 39 \times 38$) possible unique combinations of senders and receivers. Therefore, there exists 13.97% (207/1482) of all possible connections in the research dataset of this study, which makes it suitable to be explored following a sociological analysis approach.

For this study, the email communication network collected over the semester was divided into three five-week periods, according to the increasing expectations regarding study load. The portion of the complete longitudinal network that evolved between two consecutive observations is called a *wave* (or a *window*) (Krings, Karsai, Bernhardsson, Blondel, & Saramäki, 2012; Uddin, Hamra, & Hossain, 2013; Uddin, Hossain, et al., 2011; Uddin, Murshed, et al., 2011; Uddin, Piraveenan, Chung, & Hossain, 2013). The study load of the course was distributed among these three *waves* in such a way that the study load in *wave2* was higher than *wave1* and the study load in *wave3* was higher than *wave1* and *wave2*. Table 1 outlines the duration of each *wave* and the distribution of assessment components of the course among these *waves*. The corresponding sociogram for each *wave* is illustrated in Fig. 1. These sociograms illustrate the connectivity (i.e. who is connected with whom) among students during each *wave*. An in-depth network analysis is therefore required to explore structural changes of networks during these three *waves*. The duration of five weeks for each *wave* was chosen because, in the research dataset, it was noticed that a minimum time of five weeks was required to evolve sufficient communication links among students so that a research analysis could be conducted using social network analysis and simulation models.

The analysis used in this study followed three different approaches, as illustrated in Table 2. A simple coding and counting approach was used to report on the basic statistics of the student communication, and will be further discussed in the results section. The second and third approaches, network measures and simulation models using ERG respectively, will be discussed in more detail in the next sections.

3.2. Network measures

Measures of social network analysis (SNA) have been found to be very useful in investigating informal networks (e.g. students' email communication) and their changes over time (Wasserman & Faust, 2003). SNA can be seen as the mapping and measuring of relationships among participating actors in a network (Carrington, Scott, & Wasserman, 2005) and can provide both a visual and mathematical analysis of network relations among actors. It plays an important role in identifying and quantifying informal networks which function at level beyond the formal and traditional organisational structure of actor relationships (Wasserman & Faust, 2003). This study utilises the following four basic level SNA measures: *degree centralisation*, *closeness centralisation*, *betweenness centralisation* and *average network distance*.

3.2.1. Degree centralisation

Before explaining *degree centralisation*, *degree centrality* should be defined. *Degree centrality* is one of the basic measures of network centrality. It is the proportion of nodes that are adjacent to a particular actor in a network. *Degree centrality* highlights the node with the most links to other actors in a network, and can be defined by the following equation for the actor (or node) i in a network having N actors (Wasserman & Faust, 2003):

$$C'_D(n_i) = \frac{d(n_i)}{N-1}$$

The subscript D for 'degree' and $d(n_i)$ indicates the number of actors with whom actor i is connected. The maximum value for $C'_D(n_i)$ is 1 when actor i is linked with all other actors in the network.

The set of *degree centralities*, which represents the collection of *degree* indices of N actors in a network, can be summarised by the following equation to measure network *degree centralisation* (Freeman, Roeder, & Mulholland, 1979):

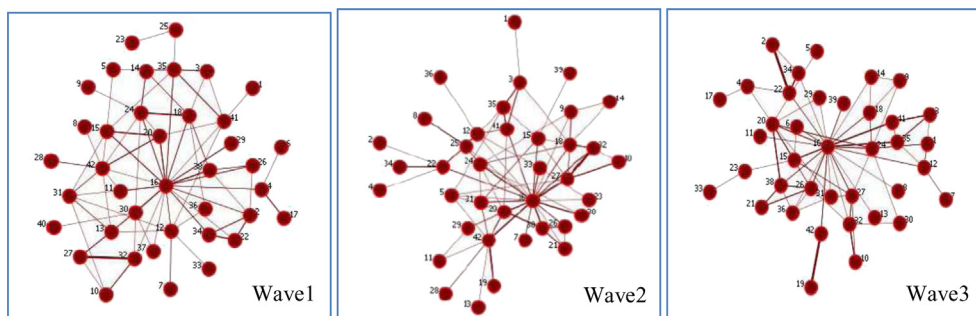


Fig. 1. Students' email communication network for wave1, wave2 and wave3.

Table 2

Measures and models used for different network analysis approaches.

| Approach to network analysis | Measures/Models used |
|---|--|
| Report statistics of basic network measures | Number of participants, number of emails, average emails per participant, number of links, number of reciprocal links and link weight. |
| Using network-level measures | Degree centralisation, closeness centralisation, betweenness centralisation, efficiency and hierarchy |
| Using simulation model | Exponential random graph models |

$$C_D = \frac{\sum_{i=1}^N [C_D(n^*) - C_D(n_i)]}{[(N-1)(N-2)]}$$

where, $\{C_D(n_i)\}$ are the *degree* indices of N actors and $C_D(n^*)$ is the largest observed value in the *degree* indices. For a network, *degree centralisation* (i.e. the index C_D) reaches its maximum value of 1 when one actor chooses all other $(N-1)$ actors and the other actors interact only with this one (i.e. the situation in a *star* graph as illustrated in Fig. 2). This index (i.e. C_D) attains its minimum value of 0 when all *degrees* are equal (i.e. the situation in a *circle* graph as illustrated in Fig. 2). Thus, C_D indicates varying amounts of centralisation of *degree* compared to both *star* and *circle* graphs.

3.2.2. Closeness centralisation

Closeness centrality needs to be defined before explaining *closeness centralisation*. *Closeness centrality*, another view of actor centrality based on closeness or distance, focuses on how ‘close’ an actor is to all the other actors in a network (Freeman et al., 1979). The idea is that an actor is central if it can quickly interact with all other actors in a network. In the context of a communication relation, such actors need not rely on other actors for the relaying of information. For an individual actor, it can be represented as a function of shortest distances between that actor and all other remaining actors in the network. The following equation represents the closeness centrality for a node i in a network having N actors (Freeman et al., 1979; Wasserman & Faust, 2003):

$$C'_C(n_i) = \frac{N-1}{\sum_{j=1}^N d(n_i, n_j)}$$

where, the subscript C for ‘closeness’, $d(n_i, n_j)$ is the number of lines in the shortest path between actor i and actor j , and the sum is taken over all $i \neq j$. A higher value of $C'_C(n_i)$ indicates that actor i is closer to other actors of the network, and $C'_C(n_i)$ will be one when actor i has direct links with all other actors of the network.

The set of *closeness centralities*, which represents the collection of *closeness* indices of N actors in a network, can be summarised by the following equation to measure network *closeness centralisation* (Freeman et al., 1979):

$$C_C = \frac{\sum_{i=1}^N [C'_C(n^*) - C'_C(n_i)]}{[(N-1)(N-2)/(2N-3)]}$$

where, $\{C'_C(n_i)\}$ are the *closeness* indices of N actors and $C'_C(n^*)$ is the largest observed value in *closeness* indices. For a network, *closeness centralisation* (i.e. the index C_C) reaches its maximum value of unity when one actor chooses all other $(N-1)$ actors and the other actors have *shortest distances* (i.e. geodesics) of length 2 to the remaining $(N-2)$ actors (i.e. the situation in a *star* graph as illustrated in Fig. 2). This index (i.e. C_C) can attain its minimum value of 0 when the lengths of *shortest distances* (i.e. geodesics) are all equal (i.e. the situation in a *complete* graph and *circle* graph as illustrated in Fig. 2). Thus, C_C indicates varying amounts of centralisation of *closeness* compared to *star*, *circle* and *complete* graph.

3.2.3. Betweenness centralisation

Betweenness centrality will be defined first before explaining *betweenness centralisation*. *Betweenness centrality* is obtained by determining how often a particular node is found on the shortest path between any pair of actors (or nodes) in the network. It views an actor as being in a favoured position to the extent that the actor falls on the shortest paths between other pairs of actors in the network. That is, nodes that occur on many shortest paths between other pairs of nodes have higher *betweenness centrality* than those that do not (Freeman, 1978). Therefore, it can be regarded as a measure of strategic advantage and information control. In a network of size N , the *betweenness centrality* for an actor (or node) i can be represented by the following equation (Wasserman & Faust, 2003):

$$C'_B(n_i) = \frac{\sum_{j < k} \frac{g_{jk}(n_i)}{g_{jk}}}{[(N-1)(N-2)]/2}$$

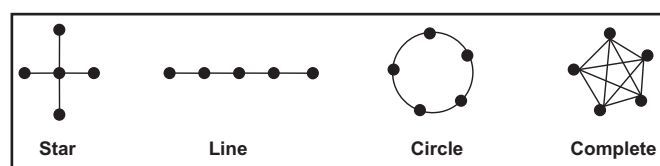


Fig. 2. Illustration of a *star*, a *line*, a *circle* and a *complete* graph.

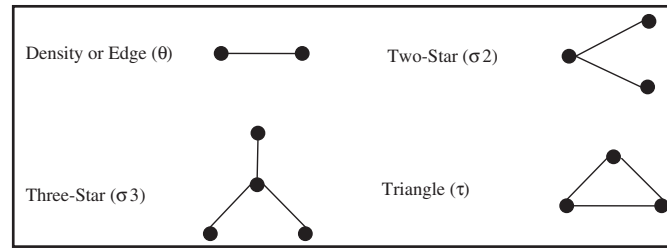


Fig. 3. Configurations of low-order Markov graph (Robins, Pattison, et al., 2007; Robins, Snijders, et al., 2007). This study considers only these four configurations for ERG modelling.

where, $i \neq j \neq k$; $g_{jk}(n_i)$ represents the number of shortest paths linking the two actors that contain actor i ; and g_{jk} is the number of shortest paths linking actor j and k .

From the set of *betweenness centralities* of N actors in a network *betweenness centralisation* can be defined by the following equation:

$$C_B = \frac{\sum_{i=1}^N [C_B(n^*) - C_B(n_i)]}{(N-1)}$$

where, $\{C_B(n_i)\}$ are the *betweenness* indices of N actors and $C_B(n^*)$ is the largest observed value in *betweenness* indices. For a network, *betweenness centralisation* (i.e. the index C_B) reaches its maximum value of unity when one actor chooses all other $(N-1)$ actors and the other actors have *shortest distances* (i.e. geodesics) of length 2 to the remaining $(N-2)$ actors (i.e. the situation in a *star* graph as illustrated in Fig. 2). This index (i.e. C_B) can attain its minimum value of 0 when all actors have exactly the same actor *betweenness* index (i.e. the situation in a *line* graph as illustrated in Fig. 2). Thus, C_B indicates varying amounts of centralisation of *betweenness* compared to both *star* and *line* graph.

3.2.4. Average network distance

Average network distance is an effective measure for capturing the aspect of how individuals are embedded in networks. The *average network distance* for a network is the average value of the shortest distances between all pairs of actors (Chung & Lu, 2004), which is calculated by the following equation :

$$\text{Average Network Distance} = \frac{2^* \sum_{ij} d(i,j)}{N(N-1)}$$

where, $d(i,j)$ is the shortest distance between actor i and j , and N is the number of actors of the network. In a given network, this measure computes the average amount of connections between any two members. The information about how far an actor is from all other actors of a network is very important for understanding its positional constraints and opportunities. Scott (2005) argued that the distances among actors in a network may be an important macro-characteristic of that network. It may take longer time for information to diffuse across a group of actors when distances among them are high.

3.3. Network simulation: exponential random graph models

For the third network analysis approach of Table 2 (i.e. network simulation), this study considers exponential random graph (ERG) models. ERG models can effectively identify structural properties in social networks (Wasserman & Pattison, 1996). This theory-driven modelling approach also allows the significance of structural parameters in the process of the formation of a given network to be tested (Hamra, Uddin, & Hossain, 2011; Uddin et al., 2013). For instance, a given social network may be explored using an ERG model to examine which micro-structures play statistically significant roles in the development process of that network. It simplifies a complex structure down to a combination of basic parameters. The advantage of this approach is that it is very general and scalable as the architecture of the graph is represented by locally determined explanatory variables, and the choice of explanatory variables is quite flexible and can be easily revised.

To represent ERG models, this paper follows the notation and terminology described in Robins, Pattison, Kalish, and Lusher, (2007) and Robins, Snijders, Wang, Handcock, and Pattison, (2007). For each pair (e.g. i and j) of a set of N actors, X_{ij} is a network tie variable with $X_{ij} = 1$ if there is a network tie from i to j , and otherwise $X_{ij} = 0$. This paper specifies x_{ij} as the observed value of X_{ij} where X is the matrix of all variables and x is the matrix of observed ties of the network. A configuration is a set of nodes and a subset of ties among them. For example, an *edge* is a subset of two nodes in which one node is connected by a tie to another, and a *3-star* is a subset of four nodes in which one node is connected by a tie to each of the other three nodes. Configurations are defined hierarchically, so that a *triangle* also includes three *2-stars*. The general form of the class of (homogeneous) ERG models is as follows (Robins, Pattison, et al., 2007; Robins, Snijders, et al., 2007):

$$\Pr(X = x) = \frac{1}{k} \exp \left\{ \sum_A \eta_A g_A(x) \right\}$$

where, (i) the summation is over configuration types A ; different sets of configuration types represent different models (e.g. dyadic independence or Markov random graph); (ii) η_A is the parameter corresponding to configuration of type A ; (iii) $g_A(x)$ is the network statistic corresponding to configuration A (for homogeneous Markov graph models this is the number of configurations of type A observed in the network: for example, the number of triangles); and (iv) k is a normalising quantity to ensure a proper probability distribution. A commonly

used sub-class of ERG model is the Markov random graph in which a possible tie from i to j is assumed conditionally dependent only on other possible ties involving i and/or j (Frank & Strauss, 1986). This sub-class of ERG models is also known as low-order model. For ERG models, this study considers only low-order Markov graphs. Different configurations for low-order Markov graphs are illustrated in Fig. 3. For the purposes of estimation, this study utilises a maximum Monte Carlo Markov chain maximum (MCMC) estimator.

4. Findings of this study

4.1. Findings from the statistics of basic network measures and email categorisation

The statistics of basic network measures at different waves are reported in Table 3. The number of emails, sent by the students, increased consistently from *wave1* (164 emails) to *wave3* (242 emails) although the number of network participants remained almost unchanged. The average number of emails sent by individual students also increased (from 4.21 at *wave1* to 6.37 at *wave3*). When a student sent an email to another student then a link was established between them, which was given a weight of one. If these two students had five email communications between them, then the link between them had a weight of five. The total number of links changed slightly: no change was observed between *wave1* and *wave2* and there was a slight decrease between *wave2* and *wave3*. When considered with the increase in the total number of emails, this meant that students communicated more frequently, but only with those students with whom they had email communications before. This is reflected in the results of the percentage of reciprocal links which increased (40% in *wave1* to 46% in *wave3*) over time. The average link weight also increased during that time (1.55 at *wave1* to 2.43 at *wave3*), as did the maximum link weight (from 8 at *wave1* to 13 at *wave3*) and the standard deviation of link weight (from 1.01 at *wave1* to 2.39 at *wave3*). The minimum link weight remained unchanged (always 1). In summary, as study load increased during the semester, students increased the amount of email communication in which they were engaged. While some of the measures that were related to links (i.e. percentage of reciprocal links, average link weight, maximum link weight and standard deviation of link weight) in their email communication network increased, examination of the total number of network links did not reveal any pattern of change.

The distribution of different categories of emails throughout the three waves is presented in Table 4. There are increasing patterns for emails that belong to the *general discussion* and *questions & inquiries* categories throughout three waves. The opposite pattern (i.e. decreasing pattern) was observed for the *socialisation* and *tutorial* categories. Compared to the second and third waves, students sent a very higher percentage of email communications for the purpose of socialisation in the first wave, probably indicating an initial phase in which students became familiar with their classmates.

4.2. Findings from the network-level measures

The results of the analysis of network-level measures are presented in Table 5. The change in *degree centralisation* increased between *wave1* (0.62) and *wave2* (0.91) and then decreased between *wave2* and *wave3* (0.88). Between *wave1* and *wave2* the *closeness centralisation* decreased slightly (from 0.13 to 0.11), and then sharply between *wave2* and *wave3* (0.02). This indicates that the varying amount of the centralisation of *closeness* for the students' email communication network moved towards either a *circle* or a *complete graph* structure with the increased amount of study load. In a *complete graph*, there is a maximum number of links among member nodes. Since the total number of links in different waves did not change significantly (see Table 3), the students' communication network moved towards a *circle* structure as study load increased and the semester progressed. *Betweenness centralisation* of students' email communication network decreased (0.47 at *wave1* to 0.44 at *wave3*) slightly with the increase of study load. This meant that the varying amount of the centralisation of *betweenness* for students' email communication network moved towards a *line* structure with the increase in study load. Unlike *closeness* and *betweenness* centralisation, *average network distance* of students' email communication network increased (from 3.67 at *wave1* to 5.66 at *wave3*) with the increase in study load. Since the *distance* between two nodes in a network quantifies the number of intermediaries that are needed so that one node can reach the other node (Chung & Lu, 2004), this finding indicates that students separated in their email communication network with the increase of their study load.

We explore the impact of network positions of individual actors on their achieved marks in the course. For that, we divided students into two groups- students who got above the average mark (which is 78.8) belong to the *group1* and other students belong to the *group2*. For calculating network position of each student, we consider the three basic SNA centrality measures (i.e. *degree centrality*, *closeness centrality* and *betweenness centrality*) and a single network that consists of all communication links of *wave1*, *wave2* and *wave3*. We do not consider *network distance* measure since this is a network-level measure and cannot be utilised to quantify the position of an individual actor. The result is presented in Table 6. Students who belong to the *group1* have higher values for the SNA measures compared to the students of the *group2*.

Table 3
Statistics of basic network measures of different waves.

| Network items | Wave1 | Wave2 | Wave3 |
|-------------------------------|-------|-------|-------|
| Number of participants | 39 | 38 | 38 |
| Number of emails | 164 | 211 | 242 |
| Average email per participant | 4.21 | 5.55 | 6.37 |
| Number of links | 105 | 105 | 99 |
| Percentage of reciprocal link | 40% | 44% | 46% |
| <i>Link weight</i> | | | |
| Minimum | 1 | 1 | 1 |
| Maximum | 8 | 9 | 13 |
| Average | 1.55 | 2.01 | 2.43 |
| Standard deviation | 1.01 | 1.43 | 2.39 |

Table 4
Distribution of different categories of emails throughout three waves.

| Wave ID | Percentage of emails in different categories | | | | Total |
|---------|--|--------------------|----------|-----------------------|-------|
| | Socialisation | General discussion | Tutorial | Questions & inquiries | |
| Wave1 | 31% | 8% | 51% | 10% | 100% |
| Wave2 | 4% | 34% | 43% | 19% | 100% |
| Wave3 | 2% | 38% | 31% | 29% | 100% |

Table 5
Results of network-level network measures for different waves.

| Network measures | Wave1 | Wave2 | Wave3 |
|----------------------------|-------|-------|-------|
| Degree centralisation | 0.62 | 0.91 | 0.88 |
| Closeness centralisation | 0.13 | 0.11 | 0.02 |
| Betweenness centralisation | 0.47 | 0.46 | 0.44 |
| Average network distance | 3.67 | 4.41 | 5.66 |

4.3. Findings from exponential random graph models

Pnet⁵ was used to fit ERG models with students' email communication networks for *wave1*, *wave2* and *wave3*. Following several iterative processes, the model (i.e. *edge*, *2-star*, *3-star* and *triangle*) had been found to fit with students' email communication network for *wave1*, *wave2* and *wave3*. The result for this model is shown in Table 7. The weight of different micro-structures (e.g. *edge* and *2-star*) of this model can be tested using a *t-value* (also known as *t-statistics*) which is defined by dividing the *estimate* by its *standard error*. Thus, the *t-value* measures how many *standard errors* the *estimate* is away from zero. Generally, any *t-value* greater than +2 or less than -2 (i.e. absolute *t-value* is greater than 2) is acceptable. The higher the *t-value*, the greater the confidence showed by the parameter under consideration as a predictor.

The parameter interpretation of the fitted ERG model is summarised as follows. The *edge* parameter showed significant confidence as a predictor for *wave1*, *wave2* and *wave3*. However, the *2-star* parameter revealed significant confidence as a predictor only for *wave3*. The positive *2-star* parameter indicates that there is a tendency for multiple network partners. Thus, students tended to have multiple network partnerships with their peers during the period of having the highest amount of study load (i.e. *wave3*). During this time (i.e. *wave3*) their email communication network became well connected but there was a low probability of having a network-hub (i.e. highly connected actor), which indicates that the *wave3* network is decentralised.

5. Discussion and conclusion

In the discussion, we aim to explain the findings in terms of the implications for the management of collaborative work, rather than make any conclusions about student learning. In addition, we will reflect on the use of this method for future work, aiming to examine the implications for learning.

The three waves examined in this paper correspond to increasing study load for the students involved. While we cannot make any significant claims as to the content of the emails, previous studies (Hassini, 2006; Tolmie & Boyle, 2000) have shown that students do use email to discuss assignments, projects and grades. However, given the classifications illustrated in Table 4, we can assume that at least part of the email communication between students, as study load related to assessment increased, was related to the assessment (for the *general discussion* and *questions & inquiries* categories). The findings indicate that students changed from a circular pattern of interaction to one in which more members were connected to each other (with a *complete* or *star* pattern). It seems that students came to interact with more members of the class as the semester carried on. As a method that could be used by teachers to inform about the class progress, this shows promise. In this case, this can be regarded as a positive indicator for the class as a whole; however this does not explain the behaviour of students in smaller groups.

Another indicator that could be used by teachers is the total number of emails, both overall and within particular waves. The total number of emails increased over the course, indicated by the increase in emails sent on average. There was a large jump between the first and second waves, in terms of the *degree centralisation*, and between *wave2* and *wave3*, there was little difference. This could indicate to the teacher that something happened between *wave1* and *wave2* that moved the group from a more circular pattern to one closer to a star. The closeness centralisation decreased, this could indicate movement to a complete pattern. *Betweenness centralisation* was similar; however, there was a slight increase. They were marginally closer to the *line* than to the *star* pattern. In addition, the average *network distance* increased, which meant it took longer for any information to diffuse in the network. Given that we did not have access to the content of emails, we cannot make any conclusion about what may have happened. But this could be useful tool for teachers to know when to intervene, or for students to monitor their own collaborative processes around assessments. In terms of the basic SNA centrality measures (i.e. *degree centrality*, *closeness centrality* and *betweenness centrality*), students who got higher marks in the course had stronger positions within the email communication network. A higher value of the basic SNA centrality measures for a student indicates that student has higher-level of network participation within the network. That means that student is more active, has more collaboration with her peers and is being easily accessible for knowledge sharing by her peers within the network (Wasserman & Faust, 2003). This could be a further indicator for teachers to intervene 'slow-progressing' students of their classes.

⁵ <http://www.sna.unimelb.edu.au/pnet/pnet.html>.

Table 6
Impact of network positions of individual students on their perceived total marks in the course.

| Network measures | Group ID | Statistics | | | Significance (using <i>t</i> -test) |
|------------------------|----------|------------|-------|-------|-------------------------------------|
| | | <i>N</i> | Avg. | Std. | |
| Degree centrality | Group1 | 19 | 0.146 | 0.110 | 0.035 |
| | Group2 | 13 | 0.091 | 0.097 | |
| Closeness centrality | Group1 | 19 | 0.142 | 0.036 | 0.271 |
| | Group2 | 13 | 0.127 | 0.044 | |
| Betweenness centrality | Group1 | 19 | 0.052 | 0.046 | 0.015 |
| | Group2 | 13 | 0.020 | 0.023 | |

For the purpose of empirical validation, we followed a similar protocol in order to repeat the experiment of this study for different cohorts of students who (i) enrolled the same course (i.e. Quantitative Methods in Project Management) in another semester; and (ii) enrolled different course (i.e. Statistical Methods in Project Management) in another semester. In both cases, we noticed similar outcomes as observed in this study.

Worthy of further examination is the interaction of measures and methods of analysis, and the implications for the further use of this method. A summary of these findings can be seen in Table 8. These findings indicate that the combination of methods could provide insight into the collaborative practices of students. A high value for each of three centralisation measures represents the *star*-like structure. For the configurations of ERG models, the *edge* is present on any of the four network structures. *2-star* and *3-star* are present on *star* and *complete graph*. The *triangle* is present only in the *complete graph*. Since a *complete graph* represents higher level of connectivity, the significance of the presence/absence of any of the three configurations (i.e. *2-star*, *3-star* and *triangle*) would indicate the collaborative behaviour of students.

The changes in the distribution of different categories of emails have been represented through the underlying changes of different SNA measures utilised in this study. With the increase in study load, distributions of the four categories of emails changed significantly. The *general discussion* and *questions & inquiries* categories were observed to show their highest proportion in the third wave; whereas, Table 4 showed that the *socialisation* and *tutorial* categories had higher percentages in the first wave. The *general discussion* emails were mainly created for discussion regarding group assignments, online quizzes and other course materials with peers. Therefore, as study load increased, students started to become more involved in online knowledge sharing and collaboration at a higher level. The discussion within the *questions & inquiries* emails also increased at the same time when the majority of the assessment components were due.

The potential implications for learning and education are important to consider at this stage. Generally, an examination of the processes of learning indicates that students need to pass through different phases of group work before reaching an ideal model (Fleming, Glass, Fujisaki, & Toner, 2010). Scaffolds may help learners to interact in such a way that a more integrated approach is reached in a more efficient manner. Further examination of the relationship of such patterns in complex learning environments would also be useful. Other studies have found that complex combinations of tool use were found concurrently with complex patterns of decision-making (Thompson & Kelly, 2012), and so it may be that the methods discussed here show similar alignment. In complex learning environments, the roles of students can also be more specialised in group work. Research that examines the development of expertise in group work (Thompson et al., 2013) suggests that expertise can be recognised in tool use as well as in the focus of attention. Patterns of interaction could be another way to recognise the establishment of different roles in groups. In some groups, a centralised pattern may be important, whereas in others, a decentralised pattern may indicate effective collaboration. Measures of effective collaboration need to relate back to the original design for learning. Further research is required to explore if the model of collaboration that assumes that all students should communicate equally and should be equally connected within the network corresponds best with effective designs for learning and learning outcomes. Such measures could be of use for teachers in managing online collaborative groups, and also for students in taking steps to manage their own communication and collaboration during group activities.

The paper has reported on the results of an initial analysis of a dataset using these novel methods. There are, of course, limitations to this work. The main limitation is that the content of the emails were not captured, and plans for further studies certainly include the addition of this source of data. A number of studies have expressed the advantages of multimodal data collection (Kennedy-Clark & Thompson, 2013; Thompson & Kelly, 2012; Worsley & Blikstein, 2013). Other means of communication between students (e.g. face-to-face, third party email service, social media, phone and texting) could also be added to the dataset. Students' email communications within the online learning environment might have been affected by the students' awareness that their emails were recorded for this study (despite the fact that all emails would be de-identified). It would also be interesting to examine whether patterns could be identified in different discipline areas, and in preparation for different types of assessment tasks.

To extend this research even further, an analysis of the content, in line with other studies of email (e.g. Hassini, 2006), could reveal whether dividing the data by assessment is an appropriate division to make. Methods for analysing the content, for example by using algorithms for word frequency analysis, would ensure that any potential for close-to-real-time feedback could still be met. Automated

Table 7
Estimation result for ERG model with (*edge*, *2-star*, *3-star* and *triangle*) configuration.

| | Wave1 | | | Wave2 | | | Wave3 | | |
|----------|-----------|--------|-------------|-----------|--------|-------------|-----------|--------|-------------|
| | Estimates | Stderr | Est./Stderr | Estimates | Stderr | Est./Stderr | Estimates | Stderr | Est./Stderr |
| Edge | -3.04 | 0.82 | -3.71 | -3.81 | 0.76 | -5.01 | -4.25 | 0.57 | -7.46 |
| 2-star | 0.17 | 0.19 | 0.89 | 0.37 | 0.22 | 1.68 | 0.44 | 0.21 | 2.10 |
| 3-star | -0.03 | 0.05 | -0.60 | -0.07 | 0.06 | -1.17 | -0.07 | 0.08 | -0.88 |
| Triangle | -0.05 | 0.35 | -0.14 | -0.31 | 0.43 | -0.72 | -0.77 | 0.72 | -1.07 |

Table 8
Presence/absence and maximum/minimum values of different SNA measures and ERG configurations in various network structures (i.e. Star, Line, Circle and Complete graph).

| Network structure | Degree centralisation | Closeness centralisation | Betweenness centralisation | Edge | 2-star | 3-star | Triangle |
|-------------------|-----------------------|--------------------------|----------------------------|------|--------|--------|----------|
| Star | ~1 | ~1 | ~1 | ✓ | ✓ | ✓ | × |
| Line | × | × | ~0 | ✓ | × | × | × |
| Circle | ~0 | ~0 | × | ✓ | × | × | × |
| Complete graph | × | ~0 | × | ✓ | ✓ | ✓ | ✓ |

content analysis identifying selected keywords could reveal the purpose of emails (Thompson et al., 2013), for example to distinguish work-related email from personal emails or the emergence of new terms introduced in the course. Using social network analysis, frequent senders and recipients of emails could be determined to identify which students are more central and which are less integrated (not part of a clique). This could provide important real-time information for instructors to better integrate students outside of cliques or facilitate communications between cliques.

Making selected information effectively and efficiently accessible to instructors to support real-time monitoring of student progress as well as to serve as a resource for designing or revising courses is an important consideration. It remains to be determined which information will be most useful to instructors and what formats can make information efficiently accessible. A dashboard approach could be used to provide selected information useful to instructors in real-time to inspect the collaborative processes and recognise patterns of behaviour. Few (2006) defines dashboards as “a visual display of the most important information needed to achieve one or more objectives; consolidated on a single screen so the information can be monitored at a glance”. Martinez et al. (2012) developed visual representations of students' collaboration patterns for instructors' dashboards. Carefully selected summaries of real-time data presented in a visual dashboard (for example of isolated students or cliques) have the potential to support instructors' detection of problems in terms of collaboration earlier and develop interventions.

In addition to instructors, real-time information about email communication could be made available to students to improve their awareness of their collaborative interactions. Visual representations of interactions between learners can reveal patterns and serve as indicators for the quality of collaborative group work (Erickson, 1999). For example, Donath (2002) visualised participation in online group tasks using a Bloom visualisation. Kay et al. (2006a, 2006b) developed a set of visualisations to identify anomalies in online teamwork that indicate the amount of participation, interactions between group members, and leadership roles.

This paper has initiated an investigation of the potential of new methods for identifying effective collaboration in complex learning environments. By combining SNA and ERG in the analysis of student interactions, changes in patterns of collaboration could be observed, as the course progressed. This has inspired numerous ideas for future research, as there are many implications for analysing the relationships between the behaviour of learners, the design of complex learning environments, and the processes of learning.

References

- van Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourses. *International Journal of Computer-supported Collaborative Learning*, 4(3), 259–287.
- Absalom, M., & Pais Marden, M. (2004). Email communication and language learning at university—an Australian case study. *Computer Assisted Language Learning*, 17(3–4), 403–440.
- Ataman, R., & DeMolive, W. (1998). Office hours—none: an e-mail experiment. *College Teaching*, 46(1), 31–35.
- Boles, W. (1999). Classroom assessment for improved learning: a case study in using e-mail and involving students in preparing assignments. *Higher Education Research & Development*, 18(1), 145–159.
- Borgatti, S. P., & Foster, P. C. (2003). The network paradigm in organizational research: a review and typology. *Journal of Management*, 29(6), 991–1013.
- Burt, R. (1992). *Structural holes: The social structure of competition*. Harvard Univ Pr.
- Carrington, P., Scott, J., & Wasserman, S. (2005). *Models and methods in social network analysis*. New York: Cambridge Univ Pr.
- Chung, F., & Lu, L. (2004). The average distance in a random graph with given expected degrees. *Internet Mathematics*, 1(1), 91–113.
- Dillenbourg, P. (1999). What do you mean by collaborative learning?. In *Collaborative-learning: Cognitive and computational approaches* (pp. 1–19). Oxford: Elsevier.
- Dillenbourg, P., Dimitriadis, Y., Nussbaum, M., Roschelle, J., Looi, C. K., Asensio, J. I., et al. (2012). Design for classroom orchestration. *Computers & Education* (in press).
- Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In *Technology-enhanced learning* (pp. 3–19). Springer.
- Dillenbourg, P., & Jermann, P. (2010). Technology for classroom orchestration. In *New science of learning* (pp. 525–552). Springer.
- Dillenbourg, P., Zufferey, G., Alavi, H., Jermann, P., Do-Lenh, S., Bonnard, Q., et al. (2011). Classroom orchestration: the third circle of usability. In , Vol. 1. *CSC2011 Proceedings* (pp. 510–517).
- Ding, N. (2009). Visualizing the sequential process of knowledge elaboration in computer-supported collaborative problem solving. *Computers & Education*, 52(2), 509–519.
- Dohn, N. B. (2010). Teaching with wikis and blogs: potentials and pitfalls. In *Proceedings of the 7th International conference on networked learning*.
- Donath, J. (2002). A semantic approach to visualizing online conversations. *Communications of the ACM*, 45(4), 45–49.
- Erickson, T. (1999). Persistent conversation: an introduction. *Journal of Computer-Mediated Communication*, 4(4), 0–0.
- Few, S. (2006). *Information dashboard design*. O'Reilly.
- Fleming, L. M., Glass, J. A., Fujisaki, S., & Toner, S. L. (2010). Group process and learning: a grounded theory model of group supervision. *Training and Education in Professional Psychology*, 4(3), 194.
- Frank, O., & Strauss, D. (1986). Markov graphs. *Journal of the American Statistical Association*, 81(395), 832–842.
- Freeman, L. (1978). Centrality in social networks: conceptual clarification. *Social Networks*, 1(3), 215–239.
- Freeman, L., Roeder, D., & Mulholland, R. (1979). Centrality in social networks: II. Experimental results. *Social Networks*, 2(80), 119–141.
- Goodyear, P., Jones, C., & Thompson, K. (2014). Computer-supported collaborative learning: instructional approaches, group processes and educational designs. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th ed.) (pp. 439–452). New York: Springer.
- Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380.
- Granovetter, M. (1983). The strength of weak ties: a network theory revisited. *Sociological Theory*, 1(1), 201–233.
- Hakkarainen, K., Lipponen, L., & Järvelä, S. (2013). *Epistemology of inquiry and computer-supported collaborative learning*. Computer supported collaborative learning.
- Hamra, J., Uddin, S., & Hossain, L. (2011). Exponential random graph modeling of communication networks to understand organizational crisis. In *SIGMIS annual conference on computer personnel research*. New York: ACM.
- Hassini, E. (2006). Student–instructor communication: the role of email. *Computers & Education*, 47(1), 29–40.
- Kapur, M. (2011). A further study of productive failure in mathematical problem solving: unpacking the design components. *Instructional Science*, 39(4), 561–579.
- Kay, J., Maisonneuve, N., Yacef, K., & Reimann, P. (2006a). The big five and visualisations of team work activity. *Intelligent Tutoring Systems*, 4053, 197–206.
- Kay, J., Maisonneuve, N., Yacef, K., & Reimann, P. (2006b). *The big five and visualisations of team work activity*. In *Intelligent tutoring systems*. Springer.

- Kay, J., Yacef, K., & Reimann, P. (2007). Visualisations for team learning: small teams working on long-term projects. In *Proceedings of the 8th international conference on computer supported collaborative learning*. International Society of the Learning Sciences.
- Kazienko, P., Musial, K., & Zgrzywa, A. (2009). Evaluation of node position based on email communication. *Control and Cybernetics*, 38(1), 67–86.
- Kennedy-Clark, S., & Thompson, K. (2013). Using multimodal discourse analysis to identify patterns of problem solving processes in a computer-supported collaborative environment. In F. Sudweeks, & H. Lim (Eds.), *Innovative methods and technologies for electronic discourse analysis*.
- Krings, G., Karsai, M., Bernhardtsson, S., Blondel, V. D., & Saramäki, J. (2012). Effects of time window size and placement on the structure of an aggregated communication network. *EPJ Data Science*, 1(1), 1–16.
- Martinez, J. E. (2012). *High performance network monitoring*. Los Alamos National Laboratory (LANL).
- Moon, J. A. (2001). *Short courses & workshops: Improving the impact of learning, training & professional development*. Routledge.
- Ngai, E. W. T., Poon, J. K. L., & Chan, Y. H. C. (2007). Empirical examination of the adoption of WebCT using TAM. *Computers & Education*, 48(2), 250–267.
- Race, P. (2001). *A briefing on self, peer & group assessment*. Learning and Teaching Support Network York.
- Reimann, P. (2009). Time is precious: variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-supported Collaborative Learning*, 4(3), 239–257.
- Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p^*) models for social networks. *Social Networks*, 29(2), 173–191.
- Robins, G., Snijders, T., Wang, P., Handcock, M., & Pattison, P. (2007). Recent developments in exponential random graph (p^*) models for social networks. *Social Networks*, 29(2), 192–215.
- Scott, J. (2005). *Social network analysis: A handbook*. London: Sage Publications Ltd.
- Segedy, J., Sulcer, B., & Biswas, G. (2010). Are ILEs ready for the classroom? Bringing teachers into the feedback loop. In *Intelligent tutoring systems*. Springer.
- Stahl, G., Koschmann, T., & Suthers, D. (2006). *Computer-supported collaborative learning: An historical perspective*. Cambridge handbook of the learning sciences.
- Thompson, K., Ashe, D., Carvalho, L., Goodyear, P., Kelly, N., & Parisio, M. (2013). Processing and visualizing data in complex learning environments. *American Behavioral Scientist*. <http://dx.doi.org/10.1177/0002764213479368>.
- Thompson, K., & Kelly, N. (2012). Processes of decision-making with adaptive combinations of wiki and chat tools. In *The future of learning: Proceedings of the 10th international conference of the learning sciences (ICLS 2012)*.
- Tolmie, A., & Boyle, J. (2000). Factors influencing the success of computer mediated communication (CMC) environments in university teaching: a review and case study. *Computers & Education*, 34(2), 119–140.
- Tyler, W. (2011). *The sociology of educational inequality*. Routledge.
- Uddin, S., Hamra, J., & Hossain, L. (2013). Exploring communication networks to understand organizational crisis using exponential random graph models. *Computational & Mathematical Organization Theory*, 19(1), 25–41.
- Uddin, S., Hossain, L., Murshed, S. T., & Crawford, J. W. (2011). Static versus dynamic topology of complex communications network during organizational crisis. *Complexity*, 16(5), 27–36.
- Uddin, S., & Jacobson, M. J. (2013). Dynamics of email communications among university students throughout a semester. *Computers & Education*, 64, 95–103.
- Uddin, S., Murshed, S. T. H., & Hossain, L. (2011). Power-law behaviour in complex organizational communication network during crisis. *Physica A: Statistical Mechanics and Its Applications*, 390(15), 2845–2853.
- Uddin, S., Piraveenan, M., Chung, K., & Hossain, L. (2013). Topological analysis of longitudinal networks. In *Hawaii international conference on system sciences (HICSS-46)*, Hawaii, USA.
- Wasserman, S., & Faust, K. (2003). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.
- Wasserman, S., & Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and P. *Psychometrika*, 61(3), 401–425.
- Wise, A. F., & Chiu, M. M. (2011). Analyzing temporal patterns of knowledge construction in a role-based online discussion. *International Journal of Computer-supported Collaborative Learning*, 6(3), 445–470.
- Worsley, M., & Blikstein, P. (2013). Towards the development of multimodal action based assessment. In *Proceedings of the third international conference on learning analytics and knowledge*. ACM.
- Zenios, M., & Holmes, B. (2010). Knowledge creation in networked learning: combined tools and affordances. In *Proceedings of the 7th international conference on networked learning*.