Measuring Participation in Distance Education Online Discussion Forums Using Social Network Analysis

Luiz Fernando Cal da Silva
Department of Software Engineering and Information Systems, Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, MG, 30640-070, Brazil. E-mail: luiz.cal@sga.pucminas.br

Marcelo Werneck Barbosa
Department of Software Engineering and Information Systems, Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, MG, 30535-901, Brazil. E-mail: mwerneck@pucminas.br

Rodrigo Richard Gomes
Department of Software Engineering and Information Systems, Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, MG, 30640-070, Brazil. E-mail: richard@pucminas.br

Distance Education professionals have been constantly coming up with methods and techniques to increase student participation in an environment where learning happens continuously and asynchronously. An online discussion forum (ODF) is one of these mechanisms, but it will only be successful if students are willing to participate. Stimulating students is a challenge many institutions currently face. The objective of this study was to analyze the social interaction among participants in ODFs using Social Network Analysis. Knowing the characteristics of these networks and its participants is important to design actions to improve the use of ODFs. As a case study, data were collected from ODF logs of the majors in Business Administration and Accounting in a Brazilian private university. This study found out that these interaction networks are sparse, which shows that students could be more engaged in interacting and collaborating with others. Students, in general, tend to interact more in the first semester and interaction diminishes as time passes. The number of active ODF participants has been around 45–50%, which shows that students currently do not participate very often in ODFs. Their main incentive seems to exist when they are graded. Popular ODFs were also analyzed.

Additional Supporting Information may be found in the online version of this article.

Received July 17, 2017; revised May 6, 2018; accepted May 27, 2018

 DOI: 10.1002/asi.24080

Introduction

Distance Education (DE) has usually been defined as education using one or more technologies to deliver instruction to students separated from instructors and other students who do not need to be present or accessible simultaneously (Caruth & Caruth, 2013). To diminish the isolation of self-paced learning and promote a learning community in a collaborative environment, students are encouraged to learn together while retaining individual control over their time, space, presence, activity, identity, and relationship (Minocha, 2009). In collaborative learning, knowledge is not transmitted to students but emerges from active dialog and group work among participants who seek to understand and apply concepts and techniques. In this context, the teacher is, therefore, an agent who facilitates the construction of knowledge through various means. Collaborative learning techniques stimulate the creation of support networks for students and promote deeper learning than more conventional approaches to teaching (Camarero, Rodríguez, & San José, 2014; Lay & Balakrishnan, 2017). The use of a collaboration strategy is highly recommended to minimize disadvantages inherent in distance learning (Anaya & Boticario, 2011).

Because the collaborative approach seeks to create enriching learning processes through dialog, debate, and even confrontation and argument, some e-learning tools are able to facilitate these activities enormously. Among such tools, online discussion forums (ODFs) are one of the most widely used asynchronous communication tools in online courses (Durairaj & Umar, 2015a) that can be used to complement learning and teaching. Forums constitute a

virtual learning environment in which students asynchronously post their responses and opinions concerning a topic, based where necessary on links, materials, and reading, thus enabling students to learn from one another (Camarero et al., 2014).

ODFs can be used in several ways. A traditional approach for creating interaction in ODFs is to use questions that require students to critically reflect on each other's responses. These forums can be mandatory or optional. Students interact through posting replies to each other's messages. ODFs can also be used for assignments—the teacher posts an initial question and students are required to post an answer to that initial question as well as reply to another student's posting (Strang, 2015). Tutors can also participate in these forums if he or she is in charge of analyzing the collaboration process regularly and frequently, especially if the course involves over 100 students (Anaya & Boticario, 2011).

The wide application of online and asynchronous discussions as a learning tool makes understanding and characterizing learning in these discussions essential, but the rapid evolution and expansion of this medium has made it difficult for education research methodologies to keep pace and provide tools for developing this understanding (Weltzer-ward, 2011). Besides, nowadays students use other popular communication tools such as Skype and social networking communities such as Facebook (Minocha, 2009) or WhatsApp (Lay & Balakrishnan, 2017).

Durairaj and Umar (2015b) state that it is important to understand the level of interaction and students' engagement in an ODF, which can be analyzed using Social Network Analysis (SNA) (Wasserman & Faust, 1994). SNA has become one of the most popular interdisciplinary analysis techniques (Ye, Li, & Law, 2011). When participants interact in ODFs, they create a social network, in which nodes are these participants and a connection exists between them if they have replied to a post from another participant in the forum. In this way, participants who present active behavior in forums might be considered to be more central or well-positioned in this network.

SNA has been used in DE research in various ways. da Silva et al. (2014) used SNA to assess the participation of members of monitors of telecentres in a digital inclusion context. Similar uses of SNA could be observed in the studies of Dorner (2012) and Sharma and Tietjen (2016). The former identified patterns of relationships among participants in a learning network and quantified interaction patterns, whereas the latter examined and mapped interactions among participants in two sections of a specific course. Cowan and Menchaca (2014) used SNA to assess communities of practice. Zawacki-Richter and Von Prummer (2010) performed an analysis of the impact of gender and collaboration patterns among researchers in research methods, topics, and productivity. Zawacki-Richter and Anderson (2011) mapped the structure of DE journals using cross-citation data to investigate the interactions and information flow among high-ranked journals. Kinshuk, Sampson, and Chen (2013) analyzed the research type, topics, and methods, international collaborations, participant levels, and frequently appearing keywords in the top-20 highly cited articles in the *Educational Technology and Society* journal in the period between 2003 and 2010. Gomes and Barbosa (2018) characterized the research community of DE with respect to coauthorship considering articles published in 11 relevant DE journals. The DE community has been characterized as under development, sparse, and poorly connected. Influential authors and the topological properties of the networks were also identified.

Although asynchronous online discussions can afford certain benefits, such benefits can only be achieved if students are willing to participate in the discussions in the first place (Hew & Cheung, 2008). Stimulating students is a challenge many institutions currently face. Therefore, the objective of this study was to analyze the social interaction among participants in ODFs to describe this social network and identify who the most central participants are. Centrality in a social network is a concept that illustrates the most important and prominent actors in the network. Actors with high centrality possess a strategic location within the network (Giannakis, 2012). Actors who are the most prominent in the community are often located in strategic locations that may allow them to communicate directly and be close to many other actors and to serve as intermediary nodes in the interactions of many other pairs of actors (Barbosa, Ladeira, & de la Calle Vicente, 2017). To achieve this objective, SNA is used. Data were collected from ODF logs of the majors in Business Administration and Accounting in a Brazilian private university. More specifically, this work is interested in answering the following research questions (RQs):

- RQ1: What are the macro and micro topological properties of the social network formed by participants in online discussion forums?
- RQ2: Who are the most influential participants in a course's online discussion forums in terms of active participation?
- RQ3: Are these topological properties and interaction patterns different in distinct courses?
- RQ4: What are the types of the online discussions with more interactions?
- RQ5: How has interaction among participants evolved in the last years?

This article contributes to scholars because it applies an important social analysis technique in a relevant research stream. It also contributes to practitioners and educational institutions because the most central participants can be used as hubs to publish information and convey ideas and behaviors in the network. It shows that SNA is a technique suitable to derive useful information on interactions in ODFs and aid conceiving new guidelines and actions to better manage virtual courses. Besides, course administrators can benefit from a discussion on how students interact with each other and how to stimulate them to do so.

DOI: 10.1002/asi

The remaining sections of this article are organized as follows. The section Participation and Interaction in Online Discussion Forums provides the theoretical background of this work, focusing more specifically on the use of ODFs to promote interaction. Research Methodology describes the research methodology used and the Results and Discussion presents the results found by our analysis. Finally, the section Conclusion and Future Work finishes the article presenting our conclusions, the limitations to this study, and suggestions for further studies.

Participation and Interaction in Online Discussion Forums

ODFs are usually organized by threads. A discussion thread is a hierarchically organized collection of messages in which all messages, except the one that started the discussion, are written as replies to earlier messages (Hew & Cheung, 2008). ODFs tend to be accessible, simple to use, compatible with teaching and learning practices, exist within virtual learning environments, allow discussions to extend beyond the classroom environment, and allow students to freely participate in discussions, offer suggestions, and give their opinions. Students can ask questions about concepts, give examples of how they encounter class material in their everyday lives, or discuss particularly difficult topics to gain greater understanding (Salter & Conneely, 2015). In addition, ODFs give both learners and teachers the opportunity to reflect before posting a response or a message by providing a nonintimidating environment where learners feel able to give opinions, offer suggestions, and ask questions (Leflay & Groves, 2013). ODFs can attract students' attention; in some courses students spend as much or more time on using the discussion forums as they do viewing lectures or doing homework, which suggests that discussing with others is particularly important (Gillani & Eynon, 2014).

In an ODF, a structure can be provided to students to respond to by posting starter questions designed to engage students and guide the discussion. On the other hand, the instructor can choose to provide less structure by not posting any questions, which requires students to create their own discussion (Salter & Conneely, 2015). The instructor's presence can heavily influence the types of interactions observed during group work, particularly during the earlier stages, due to its supportive, facilitator role, guiding the groups toward the patterns of work identified as more productive. Truly, the instructor's presence is especially important in groups with work patterns that may indicate difficulties hindering their development of a successful knowledge convergence process (Oliveira, Tinoca, & Pereira, 2011). Other instructor facilitator roles involve helping students overcome technical difficulties or concerns on how to access the online discussions and setting explicit expectations for student participation in them (Hew & Cheung, 2008). Interaction with other students and the instructor is one powerful way to provide learning support to students. Interaction is also a key element of quality of online learning. ODFs are valued by instructors as an important instrument for understanding and intervening in learning activities (Wise, Cui, Jin, & Vytasek, 2017). Instructor facilitators are advised to keep the online discussions on track because they can lose their original focus and deviate for a long period of time as compared with face-to-face discussions. Some of the ways to keep the discussions on track include designing good questions, providing guidelines for students to use when preparing their responses, rewording the question when discussions go off topic, and by providing discussion summaries. Other researchers suggest that an instructor facilitator should play the role of an encourager (Hew & Cheung, 2008).

Participants in ODFs form a virtual social community that provides them with the opportunity to take ownership over the content and direction of their learning; to be responsible for managing their learning and cooperate; and to "focus on their own learning and development from a critical, reflective perspective, combined with an understanding of relevant academic ideas and concepts" (Oliveira et al., 2011, p. 1). In such a community, the connectivity among students embraces social activity, including the conversational practices of online interaction and the learning of argumentation and discourse community norms. It goes beyond the simple interaction, including interconnected ideas, actors, and authors (Haythornthwaite, Andrews, Fransman, & Meyers, 2007). By participating in collaborative activities in a community, students acquire individual knowledge and create new knowledge that is usable for the community at large. These outcomes are evident through the engagement in the codevelopment of artifacts and practices.

Participation is an important factor in achieving a desirable outcome in learning. Much research has been conducted to show the positive influence of students' participation on their e-learning experience. Participation improves the quality of discussions and, when measured by interaction, has been shown to have a positive relationship with perceived learning achievement, grades, and students' attitude, depending on the manner and degree of interaction (Kim, 2013). The relationship between the use of ODFs and academic performance has been studied previously and some have suggested that forums could also increase student engagement (Salter & Conneely, 2015).

The collaborative learning that happens through interactions in ODFs and the learning community that is formed by its participants present characteristics in common with Communities of Practice (CoP). The concept of a CoP was introduced by Lave and Wenger (1991) as part of a discussion of how informal learning takes place in organizations. As learners participate in such communities, they develop knowledge and skills (Rosenbaum & Shachaf, 2010). Although there is little agreement on the definition of CoPs (Dessne & Byström, 2015), they can be understood as groups of individuals who share a common profession or trade and engage in collaboration to achieve a shared set of objectives. The CoP model has been promoted for its

DOI: 10.1002/asi

TABLE 1. Summary of the algorithm

FIRST PROCESSING
- Read the log to obtain all unique user ids;
- Create a list with all user ids (its size is equal to the number of unique user ids);
- Create a square matrix with dimensions (number of user ids X number of users ids)
SECOND PROCESSING
- For each line in the log;
- Check if the line corresponds to a forum interaction ("has created the post" OR "has posted content in the forum post")
- If so, obtain the user id and discussion id;
- Associate the user id with the discussion id;
- At the end of the processing, all discussions will be associated with every user who participated in them
USER INTERACTION MATRIX GENERATION
- Construct a square matrix with dimensions equal to the number of unique user ids. The head of each row and column contains the user ids;
- For each discussion identified previously, increment the weight of the relation of the users involved, that is, for each user participating in the discussion, a relationship with each and every user in the discussion is established

potential to bring together diverse groups of people, inspire cross-disciplinary learning, and to enable changes in an institution. The use of the CoP model has extended across classroom teaching, research groups, organizational change initiatives, and professional learning (Pedersen, 2016). The concept of CoP in recent years has been broadly interpreted, adapted, and applied to better understand how people learn with each other online (Gillani & Eynon, 2014). CoPs are informal and self-organized networks of peers with diverse skills and experience in an area of practice or a profession. Members of these informal networks are bound by a desire to share and develop knowledge together (Bardon & Borzillo, 2016). The term CoP suggests that knowledge construction is relational and dynamic. In a CoP, there is a close relationship between the people and the context of their activities. The learners also may become more involved with and engaged in social activities, and this contributes to the evolution and the form of membership in the community. The knowledge sharing theme reflected in CoPs has become increasingly popular among its practitioners (Cheng & Lee, 2014). To develop student learning and gain real value from group work, the process of belonging to a team (as opposed to a random group) creating boundaries and aligning with organizational goals become significant (Fearon, McLaughlin, & Eng, 2012). Wenger (2000) also advocated that as a community works more closely together, it gets engaged in shared activities and develops a greater chance of achieving a joint purpose. CoPs are also self-initiating and based on voluntary membership, so they emerge through the participation of people instead of being controlled by others. As identities develop, their perception may influence potential engagement in a CoP (Dessne & Byström, 2015). As so, knowing who the most active and central participants are in this community may influence how others participate and engage themselves in this community. These central participants are identified in this study by using SNA, whose use is detailed in the next section.

Research Methodology

This investigation is both exploratory and descriptive. A case study (Yin, 2017) strategy was used to investigate the studied phenomena. It is considered a suitable research approach when exploring emerging complex phenomena, such as the ones in this study, in order to induce theories. The case study was carried out in a Brazilian private university, more specifically in the majors of Business Administration and Accounting, which are offered completely online. The university has around 45,000 students in undergraduate majors and around 15,000 students in online undergraduate majors and postgraduate courses. The study employs SNA to obtain insights into the structural attributes of interaction among participants in ODFs. The research methodology used in this study was based on De Stefano, Giordano, and Vitale (2011), who define the following stages for research in SNA: data collection, setting network boundaries, definition of the interaction matrix, and network data analysis and interpretation of results. These stages are described as follows.

Data Collection

The goal of this study was to analyze and characterize interactions in ODFs. To do so, a network of interactions among participants needs to be constructed. In this network, vertices are participants and a relationship between them exists if they have participated in the same ODF. To identify the participants that form the ODF community, Moodle logs were extracted and analyzed. A log was generated for each course of the majors in Accounting and Business Administration in different semesters. The logs were generated with no selected filters, that is, they contained each and every interaction Moodle users had within the environment. The lines that represent actions that concern the goals of this research are the ones with the following format:

- The user with id <user_id> created the post with id <post_id> in the discussion with id <discussion_id> in the forum with course module id <module_id>.
- The user with id <user_id> posted content in the forum post with id <post_id> in the discussion <discussion_id> located in the forum with course module id <module_id>;

DOI: 10.1002/asi

TABLE 2. General characteristics of the ODFs (2016/02)

Variables	Accounting	Business administration
Number of users	2,041	1,941
Number of participants in ODFs	1,058	904
% of participation	51.84%	46.57%
Number of discussions	2,489	1,984
Number of discussion posts	6,145	5,155
Average number of posts per discussion	2.47	2.60
Number of courses	39	45

To identify and extract the information corresponding to interactions among participants in ODFs, a computer program was constructed. The program reads each log twice. At first, it captures each user_id registered in the log to identify every user that is registered in that course because the interaction network must contain every registered user, even those who have never posted any content in ODFs. In the second processing round, the program captures only the lines that represent interactions in ODFs as previously shown. The processing algorithm is summarized in Table 1.

At first, all log files of the courses of the second semester of 2016 in the Accounting and Business Administration majors were analyzed. Table 2 summarizes their main characteristics regarding ODFs. First of all, it is possible to observe that only near half of the students created a post in a discussion in both majors (51.84% for Accounting and 46.57% for Business Administration). In general, a discussion does not have many posts, only 2.47 for the Accounting major and 2.60 for Business Administration. These first results show that students could be more stimulated to participate in discussions and that, in general, a discussion proposal is presented but not many students decide to engage in it.

Definition of the Interaction Matrices

Different interaction networks were assembled to answer the research questions. For every network, a square adjacency matrix representation was used in which participants' names were both columns and rows and the intersection between these columns and rows was filled in with the number of interactions among participants involved. For research questions RQ2 and RQ4, one network, comprising all interactions made in the second semester of 2016 of all courses in the Accounting and Business Administration majors, was formed. To answer RQ1, RQ3, and RQ5 and to study the evolution of the interactions in ODFs, three networks were formed, each comprising the second semester of the last 3 years: (2014/02), (2015/02), and (2016/02). The second semesters of each year were chosen because there could be variations in the number and profile of students in the first and second semesters that could influence the results of the study. In Brazil, younger students who have just finished high school typically prefer to start their undergraduate majors in the beginning of the year (first semester), whereas in the second semester, a greater number of adult learners is observed. The first semester comprises the period from February to June, and the second semester starts in August and finishes in December. To answer RQ4, the types of the discussions with more interactions (posts) were analyzed in both majors.

In a network, edges can be directed or undirected. Directed edges are appropriate when mapping friendship, for example, where A likes B but the friendship is not reciprocated (Henneberg, Swart, Naudé, Jiang, & Mouzas, 2009). We understand that the relationship among participants in an ODF is undirected, meaning that they have a mutual relationship; one participant posts a message and the other one reads it and may reply to it or not. If some participant has posted a message, he/she will probably continue to follow updates on that discussion. So, we think of a discussion thread as a space where participants interact with all the others by either posting or reading messages (after they have posted at least one message). The linkages among participants are thus nondirectional and valued (the number of participations increases the weight of the edge). Besides, the centrality measures used in this study to identify central participants are common in nondirectional relationships. Centrality should not be confused with the notion of prestige, which is relevant only in directional networks where the actors explicitly indicate such choices in their relationships with each other (Behara, Sunil, & Smart, 2014).

Network Data Analysis and Interpretation of Results

The relationships among participants in each network and the role each one plays in these networks have been identified by evaluating several SNA metrics. The simplest one is the degree centrality, which represents the total number of nodes a specific node is connected to. Another centrality measure is node-betweenness centrality, which refers to the extent to which a particular point lies between the various other nodes in the graph. Betweenness refers to the number of paths that pass through an actor on the shortest paths connecting two other actors. The concept behind this metric is based on information flow: a node has high betweenness if it lies on many shortest paths connecting two other nodes. This is an important metric because a node with high betweenness centrality has better access to information, is better positioned to act as an intermediary in exchanges, and may be able to control flows of

DOI: 10.1002/asi

TABLE 3. General network measures (2016/02)

	Accounting				Business administration			
	Density	Avg. distance	Diameter	Avg. degree	Density	Avg. distance	Diameter	Avg. degree
1st	0.054	1.939	5	27.117	0.018	2.257	5	10.916
2 nd	0.020	2.101	7	9.976	0.059	1.880	4	10.330
3 th	0.010	2.543	6	3.979	0.015	2.951	7	4.141
4 th	0.022	2.146	6	12.904	0.097	1.598	4	18.611
5 th	0.047	2.234	7	10.640	0.003	2.946	5	1.075
6 th	0.028	2.123	5	13.706	0.004	2.895	5	2.181
7 th	0.006	3.295	8	3.394	0.024	1.957	5	6.520
8 th	0.016	2.209	5	3.847	0.126	1.801	5	57.194
All	0.009	2.732	8	19.040	0.011	2.771	7	21.785
Avg	0.025	2.324	6.125	10.695	0.043	2.286	5.000	13.871
StD	0.016	0.400	1.053	7.327	0.043	0.528	0.866	17.211

information or exchange (Sloane & O'Reilly, 2013). Besides, it can control communication flows and can potentially serve as a liaison between isolated areas of the network (Carter, Ellram, & Tate, 2007).

Network analyses were carried out using the Ucinet software, which features a large number of metrics that can be used to characterize whole networks and positions of nodes within networks (Borgatti, Everett, & Freeman, 2014). To aid our analysis, network diagrams were drawn using Gephi (https://gephi.org/). The results are shown and analyzed in the next section, considering each of the research questions and networks described previously.

Results and Discussion

This section describes the interaction networks as described previously. First, the general characteristics of the ODF interaction networks of the Accounting and Business Administration courses are presented using data from the second semester of 2016 (2016/02). Then, for each network, the most central participants are identified using SNA metrics such as centrality degree and betweenness. Then the main components of each network are analyzed. The evolution of the interaction network is evaluated considering relationships in three semesters 2014/02, 2015/02, and 2016/02. Finally, the discussion topics with more interactions are analyzed.

Online Discussion Forums Network: General Characteristics

In this section, the results for the interaction networks in ODFs in the majors of Accounting and Business Administration are presented, considering data obtained in the second semester of 2016, giving support to answer RQ1. Nine different networks were formed, considering each major semester, from the 1st to 8th and also a network with data from all courses of each major. Table 3 shows the network measures calculated for each of these networks. Considering the whole majors, they present similar metrics. In a semester, an ODF participant is connected to 10.69 other participants in the Accounting major and is related to 13.87 other different people in the Business Administration major. Although the Business Administration major presents greater average degree, the standard deviation for this metric is also higher, which shows that there are people who participate a lot in ODFs but also people who do not participate very often. The average distance between nodes is around 2.3 for both majors, which indicates that with few steps it is possible to reach every node in each network. Networks are sparse with low density, which shows that only a few of the possible connections among participants exist. Considering course semesters, an irregular interaction pattern is observed when the average degree is analyzed. In the Accounting major, participants interact more in the first semester and interaction diminishes in the last semesters. In the Business Administration major, there is also more interaction in the first semesters but also a very intense collaboration when students reach the end of their major. The reasons why students interact more at the beginning of their university lives could be related to being more enthusiastic at first, to a tendency of following guidelines provided by tutors and coordinators more closely, and also to a deeper concern of coordinators or even teachers with students who are entering the university, especially in a fully online major. Anyway, such reasons should be deeply investigated in future work. With this analysis, major coordinators and professors could come up with new actions to keep students engaged as well as identify semesters in which students have not been participating in discussions adequately.

Online Discussion Forums Network: Central Participants

RQ2 inquires about who the most central participants are in the interaction network comprising all ODF interactions retrieved in 2016/02. The most basic metric is the centrality degree, which indicates the number of connections, in this case ODF interactions, each participant in the network has. Thus, the degree centrality only reflects how many interactions a participant has had with others. Table 4 presents the 15 participants who have had the highest centrality degrees in the 2016/02 network for the Accounting and Business Administration majors. For confidentiality reasons,

DOI: 10.1002/asi

TABLE 4. Top 15 participants for centrality degree (2016/02)

Rank	Accounting			Business administration		
	Participant	Degree value	Participant type	Participant	Degree value	Participant type
1	acc_stu_01	200	Student	ba_tut_01	253	Tutor
2	acc_stu_02	198	Student	ba_stu_01	252	Student
3	acc_stu_03	197	Student	ba_stu_02	242	Student
4	acc_stu_04	196	Student	ba_stu_03	220	Student
5	acc_stu_05	193	Student	ba_stu_04	216	Student
6	acc_stu_06	193	Student	ba_stu_05	214	Student
7	acc_stu_07	192	Student	ba_stu_06	214	Student
8	acc_stu_08	185	Student	ba_stu_07	199	Student
9	acc_stu_09	164	Student	ba_stu_08	199	Student
10	acc_stu_10	163	Student	ba_stu_09	198	Student
11	acc_stu_11	160	Student	ba_stu_10	185	Student
12	acc_stu_12	157	Student	ba_stu_11	182	Student
13	acc_coord	156	Coordinator	ba_stu_12	176	Student
14	acc_stu_13	156	Student	ba_stu_13	169	Student
15	acc_stu_14	155	Student	ba_stu_14	166	Student

participants' names were substituted with unique identifiers. The participant type is also presented. It can be observed that for the Accounting major, students participate intensely. The student with more participations is connected to 200 other participants of the ODFs. The coordinator of the major ranks 13th, showing that she is constantly in contact with her students in the ODFs. In the Business Administration major, there are also students who are greatly related to other participants, but the participant with more connections is a tutor. These results show that these students who are greatly connected to others can be used as stimulators for increasing students' engagement or conveying information and expected behavior of the coordination of the major.

Betweenness centrality is another SNA measure that represents the importance of a node in the network. As explained before, this index measures if a node is on the shortest path of many pairs of nodes, and consequently if it is in a critical position to act as an information hub. In an interaction network, a participant presenting high betweenness centrality can be considered an actor who connects different groups who would not be linked to one another if it were not for this participant in particular. Table 5 displays the 15 most central participants, considering the betweenness centrality index, in the interaction networks of the second semester of 2016 of the Accounting and Business Administration majors. By comparing data presented in Table 4, it is possible to infer that more collaborations with others do not necessarily imply that a participant is more central in the network, for just a few participants appear in both tables.

Figures S1 and S2 (available in Supporting Information Appendix, online) depict the ODF interaction network for the majors in Accounting and Business Administration in the second semester of 2016. In these graphs, each circle is an ODF active participant and its size is related to the participant's degree in the network. Red circles represent students, green vertices are professors, and blue nodes correspond to tutors. With such visualization, it is possible to see how tutors and professors interact with students and how they are positioned in the network. It is also possible

TABLE 5. Top 15 participants for centrality betweenness (2016/02)

Rank	Accounting			Business administration		
	Participant	Degree value	Participant type	Participant	Degree value	Participant type
1	acc_coord	23,184.814	Coordinator	ba_tut_01	31,391.15	Tutor
2	acc_stu_09	15,101.268	Student	ba_tut_02	25,971.73	Tutor
3	acc_prof_01	12,888.054	Professor	ba_stu_15	21,424.53	Student
4	acc_stu_08	12,613.968	Student	ba_tut_03	20,924.34	Tutor
5	acc_prof_02	11,460.319	Professor	ba_stu_16	18,832.84	Student
6	acc_stu_02	10,513.168	Student	ba_stu_03	17,436.76	Student
7	acc_stu_15	9,739.545	Student	ba_stu_17	16,756.93	Student
8	acc_stu_11	9,172.422	Student	ba_stu_18	13,650.07	Student
9	acc_stu_14	8,286.586	Student	ba_stu_01	11,322.79	Student
10	acc_stu_16	8,009.446	Student	ba_prof_01	10,831.61	Professor
11	acc_stu_17	7,806.622	Student	ba_tut_04	10,062.77	Tutor
12	acc_tut_01	7,571.280	Tutor	ba_tut_05	9,069.961	Tutor
13	acc_prof_03	7,388.169	Professor	ba_stu_19	8,769.652	Student
14	acc_stu_18	7,352.679	Student	ba_stu_20	7,378.655	Student
15	acc_stu_19	6,845.214	Student	ba_stu_21	6,689.392	Student

DOI: 10.1002/asi

TABLE 6. Largest components in the networks

	Accounting			Business administration		
Semester	Number of participants in the network	Number of participants in the largest component	% of participants in the largest component	Number of participants in the network	Number of participants in the largest component	% of participants in the largest component
2014/02	872	753	86.35%	1,004	950	94.62%
2015/02	858	764	89.04%	836	815	97.49%
2016/02	1,058	852	80.53%	904	767	84.07%

to see the formation of groups, as it is clearer in the Business Administration major, where there is a large set of active students in ODFs related to a tutor who was also very active. Other smaller groups cluster themselves around a professor.

Components

Wasserman and Faust (1994) define a component as a maximal connected subgraph, that is, a path exists between all authors in the subgraph (all nodes are reachable) and there is no path between a node in the component and any node outside the component. The main component is the one with the largest number of actors. In a network, there are usually components of varying sizes and researchers often investigate whether or not a giant component is present. A giant component is the component having the largest number of connected vertices. It is usually present in established and mature networks and it should comprise at least 50% of the nodes of the network (Newman, 2004). Studies have found large components of varying sizes, from 57% to almost 90% of the network size (De Stefano et al., 2011). As the global and local metrics are measured for the giant component, detecting the largest component and determining its size is crucial to understanding the topological features of the network (Kumar & Jan, 2013). So, the analysis of components in all networks gives support to answers to RQ1, RQ3, and RQ5.

Very active participants may be responsible for the formation of the giant component because active participants are, on average, more centrally located in the network when compared with all other participants with less interaction. Moreover, participants who are linked by chains of interactions are more likely to be of a similar interest group. Table 6 displays information on the largest component of the interaction network of both majors in the last 3 second semesters of the past years. The giant component phenomenon was observed in all six networks constructed. The size of the largest component obtained varied from 80.53% to 97.49%.

Evolution of the ODF Community

To evaluate how collaborations in the ODF networks have evolved in the past years and to answer RQ5, interactions from ODFs in these majors were analyzed in the second semester of the past 3 years (2014/02, 2015/02, 2016/02). Table 7 shows the general characteristics of these networks. The network properties present, in general, constant and similar characteristics. The results do not seem to exhibit a clear tendency of improving interaction in ODFs. In the last 3 years, the number of active participants has been around 45% to 50% in both majors, which represents the need to develop incentives so that students can interact more and engage themselves in collaborative learning. The Accounting major seems to be slowly developing a strategy to improve interaction because participants present a higher degree (are connected to more participants) and lower average distance (are more connected and reachable with fewer steps) in 2016/02. The Accounting network has presented an increase in density, also demonstrating that more connections have been established among participants.

Most Popular Discussions

Finally, to better comprehend interactions among participants in the network, discussion types were also analyzed. Table 8 displays information on the most popular discussions of the Accounting and Business Administration majors in 2016/02, giving support to answer RQ4. It is possible to

TABLE 7. Evolution of the ODF networks

	Accounting			Business administration		
Network property	2014/02	2015/02	2016/02	2014/02	2015/02	2016/02
Number of participants	1,828	1,920	2,041	2,037	1,945	1,941
Number of active participants	872	858	1,058	1,004	836	904
% of active participants	47.70%	44.69%	51.84%	49.29%	42.98%	46.57%
Average degree	9.615	13.875	19.045	20.288	16.537	21.785
Density	0.005	0.007	0.009	0.010	0.009	0.011
Number of components	1,067	1,155	1,186	1,085	1,129	1,174
Average distance	3.095	2.988	2.731	2.762	2.832	2.771
Diameter	7	7	8	7	6	7

DOI: 10.1002/asi

TABLE 8. Discussion types and posts (2016/02)

Rank	Accounting		Business administration	
	Number of posts	Type	Number of posts	Type
1	183	Assignment	187	Assignment
2	168	Assignment	174	Assignment
3	147	Assignment	108	Assignment
4	141	Assignment	101	Assignment
5	115	Assignment	92	Assignment
6	70	Assignment	83	Assignment
7	56	WhatsApp Group	81	Assignment
8	55	WhatsApp Group	79	Assignment
9	50	Assignment	73	Assignment
10	41	Assignment	70	Assignment
11	39	WhatsApp Group	64	Assignment
12	34	Assignment	53	Assignment
13	34	WhatsApp Group	51	Assignment
14	33	Course Feedback	49	Assignment
15	29	WhatsApp Group	35	Assignment

see that the number of posts rapidly decline, showing that only few discussions are able to make participants really interact. Among the most popular discussions, three different types were found: Assignment (a discussion created in which student participation is graded by the professor), WhatsApp Group (a discussion in which students request participation in an external WhatsApp group), and Course Feedback (a discussion in which students give feedback on the course to the professor). One can observe in Table 8 that in the Business Administration major, all top 15 discussions are of the Assignment type. In the Accounting major, Assignment discussions also frequently occur, but there are also discussions of the WhatsApp Group and Course Feedback types. These results show that students currently do not actively participate in ODFs, except when they are graded by the professor. Students also seem to feel less confident to engage themselves in the ODFs in the university virtual environment and might prefer to create an external environment, the WhatsApp groups, where they may feel more secure to give their opinions and express themselves because professors are not invited to these groups.

Tang and Hew (2017) argue that there is currently a growing obsession with smartphone applications, in particular instant messaging, which is probably the most popular, especially among young adults. Among such applications, WhatsApp is now leading the worldwide market by owning more than 1 billion customers in over 180 countries (https://www.whatsapp.com/about/). By carrying out a literature review, the authors observed that instant messaging applications are used in different ways with pedagogical purposes, such as journaling (posting self-reflection comments), dialogic (emphasizing the discourse among participants), communication (broadcasting or transmitting information to students), constructionist with peer feedback (construction of a meaningful product), helpline (solving homework problems), and assessment (conducting summative assessments). In the courses where this study took place, WhatsApp was not used by lecturers, only by students. They decided to form their own group outside the learning virtual environment. Investigating the reasons students decided to do so is important and considered in future work.

WhatsApp instant messaging makes learning easy, favors problem solving, and resolves learning difficulties related to the learning process or content. The mobile learning technology helps students to create a learning community, to easily construct knowledge, and to share it with other members of a WhatsApp group through instant messaging. The use of such instant messaging applications might be understood as the adoption of a more learner-centered approach. Students feel free to post their own questions and comments, to reply to other peers' posts, and to share knowledge and experiences. Lay and Balakrishnan (2017) state that the shift from a traditional teacher-centered to a student-centered approach emphasizes the importance of fostering students' engagement and interactions with their peers and lecturers. Moreover, the use of mobile technology in the classroom improves students' engagement by providing a platform to gather their feedback and to respond to their queries.

Conclusion and Future Work

The objective of this study was to analyze through SNA the social interaction among participants in ODFs. Data were collected from ODF logs of the majors in Business Administration and Accounting in a Brazilian private university. This study identified who the most central participants in the community are, the topological properties of the networks, the interaction patterns, and analyzed the evolution of the interactions in the last 3 years. This study found that these interaction networks are sparse, with low density, which shows that only a few of all the possible connections among participants exist and students could be more engaged in participating, interacting, and collaborating with others. An irregular interaction pattern is observed as far as major's semesters are concerned. In the Accounting major, participants interact more in the first semester and interaction diminishes in the last semesters. In the Business

DOI: 10.1002/asi

Administration major, there is also more interaction in the first semester, but also a very intense collaboration when students reach the end of their major. The giant component phenomenon was observed in all networks constructed. The size of the largest component obtained varied from 80.53% to 97.49%. In the last 3 years, the number of active participants has been around 45% to 50%. The results have also shown that the main incentive to participate in an ODF seems to exist when students are graded by the professor. Students also seem to feel less confident to engage themselves in the ODFs in the university environment and seem willing to create a second environment via WhatsApp groups, where they might feel more secure to give their opinions and express themselves.

In a DE environment, collaboration among students is a key factor for promoting and developing learning and engagement. We believe that students participating in ODFs form a Community of Practice. To develop student learning and gain real value from group work, the process of belonging to a team becomes significant. Thus, it is important to identify key participants in such communities because learners will also become more involved with and engaged in activities if they are stimulated by others' behavior.

This work contributes to scholars and practitioners by presenting a case study that adopted SNA to investigate interaction among participants in ODFs. The results of this study show that, by using SNA, major coordinators and professors could identify and characterize interactions as well as develop new actions to keep students engaged by identifying courses in which students have not been participating in discussions adequately. Also, SNA allowed the identification of those students who are greatly connected to others and can be used to stimulate student participation or convey information and expected behavior from the coordinators.

Limitations and Future Research

There are some limitations to this study. First of all, this study assumes that the relationship among participants in an ODF is undirected, meaning that they have a mutual relationship; one participant posts a message and the others read it and may reply to it or not. In this context, participants in an ODF interact with all the others by either posting or reading messages. Thus, the relationships among participants are nondirectional. Assembling networks with directed edges would be interesting if individuals' prestige was to be investigated. In this case, we would analyze the choices participants make on whom to reply to and who the participants with more replies are. Because this study concerns participation patterns and the discussion thread is seen as a community in which all members who have posted at least one message interact, networks were formed with undirected links. Consequently, no distinction as to whether a post was a response to a specific post or a reply to the entire thread was made. In conclusion, we assumed that by posting a comment or a reply in a discussion thread, the participant interacts with all the other participants of the thread. Considering these facts, we understand that investigating different forms of assembling the interaction network could also yield interesting results and is planned as future work.

This study analyzed the evolution of the ODF networks in the past 3 years considering each year's second semester. Although no substantial difference among semesters was observed, investigating students' participation in ODF in the first semester of the past years should be done in the short term to understand if there are variations in the conclusions made so far.

This study opens several possibilities for future work. More case studies are needed to investigate different interaction patterns among distinct majors and careers. In the university where the study took place, there are only two totally online majors. Comparing the patterns observed in this study with other Brazilian universities or even international universities is considered future research. In addition, different actions to stimulate students' participation might be proposed and its effect on ODFs could be measured to assess these actions' efficacy.

References

Anaya, A.R., & Boticario, J.G. (2011). Application of machine learning techniques to analyse student interactions and improve the collaboration process. Expert Systems with Applications, 38, 1171–1181.

Barbosa, M.W., Ladeira, M.B., & de la Calle Vicente, A. (2017). An analysis of international coauthorship networks in the supply chain analytics research area. Scientometrics, 111(3), 1703–1731.

Bardon, T., & Borzillo, S. (2016). Communities of practice: Control or autonomy? Journal of Business Strategy, 37(1), 11–18.

Behara, R.S., Sunil, B., & Smart, P.A. (2014). Leadership in OM research: A social network analysis of European researchers. International Journal of Operations & Production Management, 34(12), 1537–1563.

Borgatti, S.P., Everett, M.G., & Freeman, L.C. (2014). UCINET. In Encyclopedia of Social Network Analysis and Mining (pp. 2261–2267). New York: Springer.

Camarero, C., Rodríguez, J., & San José, R. (2014). An exploratory study of online forums as a collaborative learning tool. Online Information Review, 36(4), 568–586.

Carter, C.R., Ellram, L.M., & Tate, W.L. (2007). The use of social network analysis in logistics research. Journal of Business Logistics, 28(1), 137–168.

Caruth, G.D., & Caruth, D.L. (2013). The impact of distance education on higher education: A case study of the United States. Turkish Online Journal of Distance Education, 14(4), 121–131.

Cheng, E.C.K., & Lee, J.C.K. (2014). Developing strategies for communities of practice. International Journal of Educational Management, 28(6), 751–764.

Cowan, J.E., & Menchaca, M.P. (2014). Investigating value creation in a community of practice with social network analysis in a hybrid online graduate education program. Distance Education, 35(1), 43–74.

da Silva, A.S., Rossy de Brito, S., Martins, D.L., Vijaykumar, N.L., Jorge da Rocha, C.A., Costa, J.C.W.A., & Francês, C.R.L. (2014). Social networks analysis and participation in learning environments to digital inclusion based on large-scale distance education. International Journal of Distance Education Technologies, 12(2), 1–25.

Dessne, K., & Byström, K. (2015). Imitating CoPs: Imposing formality on informality. Journal of the American Society for Information Science and Technology, 61(11), 2277–2284.

DOI: 10.1002/asi

De Stefano, D., Giordano, G., & Vitale, M.P. (2011). Issues in the analysis of co-authorship networks. Quality & Quantity, 45(5), 1091–1107.

Dorner, H. (2012). Effects of online mentoring in computer-supported collaborative learning environments: Mentor presence and cognitive engagement. American Journal of Distance Education, 26(3), 157–171.

Durairaj, K., & Umar, I.N. (2015a). A proposed conceptual framework in measuring social interaction and knowledge construction level in asynchronous forum among university students. Procedia - Social and Behavioral Sciences, 176, 451–457.

Durairaj, K., & Umar, I.N. (2015b). Analysis of students' listening behavior patterns in an asynchronous discussion forum. Procedia - Social and Behavioral Sciences, 176, 27–34.

Fearon, C., McLaughlin, H., & Eng, T.Y. (2012). Using student group work in higher education to emulate professional communities of practice. Education + Training, 54(2/3), 114–125.

Giannakis, M. (2012). The intellectual structure of the supply chain management discipline. Journal of Enterprise Information Management, 25(2), 136–169.

Gillani, N., & Eynon, R. (2014). Communication patterns in massively open online courses. Internet and Higher Education, 23, 18–26.

Gomes, R.R., & Barbosa, M.W. (2018). An analysis of the structure and evolution of the distance education research area community in terms of coauthorships. International Journal of Distance Education Technologies, 16(2), 65–79.

Haythornthwaite, C., Andrews, R., Fransman, J., & Meyers, E.M. (2007). Handbook of e-learning research. New York: Sage.

Henneberg, S.C., Swart, J., Naudé, P., Jiang, Z., & Mouzas, S. (2009). Mobilizing ideas in knowledge networks: A social network analysis of the human resource management community 1990-2005. The Learning Organization, 16(6), 443–459.

Hew, K.F., & Cheung, W.S. (2008). Attracting student participation in asynchronous online discussions: A case study of peer facilitation. Computers and Education, 51(3), 1111–1124.

Kim, J. (2013). Influence of group size on students' participation in online discussion forums. Computers & Education, 62, 123–129.

Kinshuk, H.H., Sampson, D., & Chen, N.S. (2013). Trends in educational technology through the lens of the highly cited articles published in the Journal of Educational Technology and Society. Educational Technology & Society, 16(2), 3–20.

Kumar, S., & Jan, J.M. (2013). On giant components in research collaboration networks: Case of engineering disciplines in Malaysia. Malaysian Journal of Library & Information Science, 18(2), 65–78.

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge, UK: Cambridge University Press.

Lay, C., & Balakrishnan, V. (2017). Enhancing classroom interaction via IMMAP – An interactive mobile messaging app. Telematics and Informatics, 34(1), 230–243.

Leflay, K., & Groves, M. (2013). Using online forums for encouraging higher order thinking and "deep" learning in an undergraduate Sports Sociology module. Journal of Hospitality, Leisure, Sport & Tourism Education, 13, 226–232.

Minocha, S. (2009). Role of social software tools in education: A literature review. Education + Training, 51(5/6), 353–369.

Newman, M.E.J. (2004). Coauthorship networks and patterns of scientific collaboration. Proceedings of the National Academy of Sciences of the United States of America, 101(1), 5200–5205.

Oliveira, I., Tinoca, L., & Pereira, A. (2011). Online group work patterns: How to promote a successful collaboration. Computers and Education, 57(1), 1348–1357.

Pedersen, K.W. (2016). Supporting collaborative and continuing professional development in education for sustainability through a communities of practice approach. International Journal of Sustainability in Higher Education, 18(5), 681–696. https://doi.org/10.1108/IJSHE-02-2016-0033.

Rosenbaum, H., & Shachaf, P. (2010). A structuration approach to online communities of practice: The case of Q&A communities. Journal of the American Society for Information Science and Technology, 61(9), 1933–1944.

Salter, N.P., & Conneely, M.R. (2015). Structured and unstructured discussion forums as tools for student engagement. Computers in Human Behavior, 46, 18–25.

Sharma, P., & Tietjen, P. (2016). Examining patterns of participation and meaning making in student blogs: A case study in higher education. American Journal of Distance Education, 30(1), 2–13.

Sloane, A., & O'Reilly, S. (2013). The emergence of supply network ecosystems: A social network analysis perspective. Production Planning & Control: The Management of Operations, 24(7), 621–639.

Strang, K.D. (2015). How can discussion forum questions be effective in online MBA courses? Campus-Wide Information Systems, 28(2), 80–92.

Tang, Y., & Hew, K.F. (2017). Is mobile instant messaging (MIM) useful in education? Examining its technological, pedagogical, and social affordances. Educational Research Review, 21, 85–104.

Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications. Cambridge, UK: Cambridge University Press.

Weltzer-Ward, L. (2011). Content analysis coding schemes for online asynchronous discussion. Campus-Wide Information Systems, 28(1), 56–74.

Wenger, E. (2000). Communities of practice and social learning systems. Organization, 7(2), 225–246.

Wise, A.F., Cui, Y., Jin, W.Q., & Vytasek, J. (2017). Mining for gold: Identifying content-related MOOC discussion threads across domains through linguistic modeling. Internet and Higher Education, 32, 11–28.

Ye, Q., Li, T., & Law, R. (2011). A coauthorship network analysis of tourism and hospitality research collaboration. Journal of Hospitality & Tourism Research, 37(1), 51–76.

Yin, R.K. (2017). Case study research: Design and methods (6th ed.). Newbury Park, CA: Sage.

Zawacki-Richter, O., & Anderson, T. (2011). The geography of distance education - Bibliographic characteristics of a journal network. Distance Education, 32(3), 441–456.

Zawacki-Richter, O., & Von Prummer, C. (2010). Gender and collaboration patterns in distance education research. Open Learning: The Journal of Open and Distance Learning, 25(2), 95–114.