Learning with Digital Media: Investigating the Relationship Between Student Citation Networks, Assignment Structures, and Learning Outcomes

Regina Collins
New Jersey Institute of Technology
rsb24@njit.edu

Anatoliy Gruzd
Ryerson University
gruzd@ryerson.ca

Abstract

Students are comfortable sharing digital content with others, yet the effect of sharing of digital media for learning remains largely unexplored. Building on research in social network analysis and learning analytics, this research explores the use and sharing of digital media in learning activities, analyzing the effects of the design of the learning activities on the resulting networks of students and their cited resources, and exploring relationships between attributes of these citation networks and students' perceptions of the learning outcomes. Results suggest that the extent to which an assignment is well-structured and converges towards a single solution positively influences the density and clustering coefficient of the resulting citation network, and that these network measures in turn have a positive influence on students' perceptions of learning from the assignment.

1. Introduction

The Internet places a wealth of digital content at our fingertips, and research has shown that students have a preference for using digital media to resolve their information needs [22], turning to Google as their first source of information [8]. Once defined as digitally-encoded content, digital media is now viewed as a rich medium that enables exploration and self-expression through creation and collaboration [7]. Using new information and communication technologies (ICTs) such as social curation systems (e.g. Pinterest and Pearltrees), social networking sites (e.g. Facebook), and social messaging sites (e.g. SnapChat, Instagram, and WhatsApp), students are sharing digital media for purposes ranging from entertainment to education [2, 11].

However, the sharing of digital media, particularly for the purposes of knowledge sharing and learning, remains largely unexplored due to several factors.

First, the informal exchange of digital media, even for learning, frequently occurs through social ICTs, making it difficult to capture and evaluate [2]. Additionally, even if students use digital media for learning, they rarely cite such sources in formal learning assignments due to instructors' concerns about the quality and reliability of digital media [6].

On the other hand, the increasing reliance on digital media in the teaching and learning context is beginning to fuel innovations in the emerging field of learning analytics. Learning analytics investigates student data (everything from actual assignments and grades to trace data such as number of logins or time spent on an activity), seeking to identify patterns that can improve the educational experience or learning outcomes [15, 29].

This exploratory study applies the concepts of learning analytics to the use of digital media in education by exploring student citation networks (networks connecting students and hyperlinks to the digital media they cite in an assignment). These student citation networks are created when students use and share digital media resources in a two-part assignment that requires them to 1) conduct individual research on the Internet and cite any relevant digital content for an individual research assignment; and 2) subsequently share their digital media with group members to create a group response to the same topic. Implemented as a mixed methods study [25], this research applies quantitative and qualitative methods and social network analysis to these student citation networks and to students' self-reported data about the learning activities to explore the following research questions:

RQ1: How does the design of the learning activity influence the structure of the resulting student citation network?

RQ2: What relationship exists between social network analysis measures and students' perceptions of the outcomes of the learning activity?

The remainder of this paper is organized as follows. Related literature that informed this research and its derived hypotheses are discussed in the next section.

HICSS

This is followed by a description of the study methodologies and results. The significance of the results, as well as limitations of the research, conclude the paper.

2. Related Literature

This study builds on prior research in the field of social network analysis, learning analytics, and computer-mediated collaborative learning. These topics are discussed in detail in the following sections, along with related research sub-questions for exploratory areas and related hypotheses for topics in which a directional relationship could be predicted.

2.1. Social Networks

All learners are, at least to some extent, members of some social network. Social networks are defined as "links from people to other people, groups, or information objects" [28]. Social networks can exist in physical and virtual groups, such as a classroom of students or a community of practice, and can be used to indicate friendship, competition, access to resources, or even shared knowledge.

The methodology referred to as social network analysis enables researchers to explore the characteristics of these often hidden networks. Social network analysis has been applied in the learning domain to explore the exchange of knowledge or information [14], the effects of teacher presence on the density of a network [19], the cohesion of small groups (in order to identify active groups and isolated nodes) [23], and participation in collaborative learning activities [20].

Characteristics of social networks can include network-level measures such as the overall density and transitivity of a network, as well as node-level measures such as local clustering coefficient. Network density indicates the extent to which the potential connections between nodes in a network are actual connections [5, 27]. The clustering coefficient represents the degree to which edges in a network tend to cluster together [30]. In this study, clustering coefficient is calculated by averaging the clustering coefficient of each node in the network, referred to as a local clustering coefficient [37]. Previous studies have found that both density and clustering coefficient can reflect the efficiency of a network in diffusing knowledge for learning [31, 32]. They are used in this study to assess shared knowledge (through citations to the same resources) amongst students during a learning activity.

2.2. Citation Networks and Shared Resources

Social networks are frequently based on communication patterns or interactions between learners [9, 32]. For example, connections (edges) between nodes may represent responses in discussion forums or self-reported ties such as who a student asks for help with homework.

In this study, the ties in the networks represent citations to the same resources. These resources were identified independently by students conducting individual research on the Internet. Citation networks are a particular type of social network that have been most commonly explored in relation to academic or research publications. Citation networks are typically bipartite, bimodal networks that connect authors or researchers with the papers they cite in their publications. Prior studies have focused on scholars' positions in citation networks, finding that central nodes in such networks indicate influential researchers whose papers are cited most frequently [21] and that highly cited authors tend to cite one another [10].

Learning analytics researchers explore a variety of trace data available from students' learning activities as inputs to improve learning through formative feedback, summative assessment, and even to drive educational policymaking [24]. Examples of trace data previously explored include time spent on learning activities, frequency of logins to learning management systems, and participation in online discussion activities [29]. In this research, citations to the same resources serve as a proxy for shared knowledge that is expected to lead to improved perceptions of learning.

2.3. Complexity and Problem Definition of Assignments

The complexity and amount of structure integrated into a learning problem or activity has been shown to have a significant effect on the learning process. Specifically, well-structured problems have convergent solutions that require specific processes and knowledge [13, 17], while ill-structured problems are more vaguely defined and may offer multiple paths or solutions [17, 37]. Although there is a distinction between well-structured and ill-structured assignment tasks, in actuality they represent the opposite ends of a spectrum along which assignment characteristics can be designed or evaluated. In the domain of collaborative learning, previous studies have found that learning activities using structured, asynchronous networks resulted in high critical thinking and the development of cohesive cliques in learning networks, while learning activities implementing informal,

unstructured, open discussion forums resulted in low critical thinking and few cohesive cliques [3].

Similarly, studies exploring the structures and knowledge sharing efficiencies of problem-solving social networks found that, contrary to results from modeled networks, real-world activities requiring participants to work together to solve complex problems performed better when the problem-solving networks were more efficient as evidenced by high density and clustering coefficient [31].

Together, these findings suggest that the connectivity between nodes in a network, as measured by network density and clustering coefficient, may be affected by the level to which learning activities are complex and well-structured, suggesting the following research subquestions:

RQ1a: Does network density increase as the structure and complexity of the learning assignment increases?

RQ1b: Does network clustering coefficient increase as the structure and complexity of the learning assignment increases?

2.4. Network Visualizations

Visualizations of networks have previously been studied as an effective method of displaying the relationships between entities including people, organizations, and even data. Prior research has shown the importance of social network visualizations not only for researchers seeking to understand the interactions between actors but also for those immersed in the network to gain insights from their positions and the positions of others [26]. In this paper, visualizations of student citation networks are explored as a supplemental method for understanding the impact of assignment structures of the resulting networks.

2.5. Perceived Learning

In this study, perceived learning [4] is used to capture students' perceptions about how much they learned from an assignment that required the storage, use, and sharing of digital media. Although self-reported perceptions can be problematic, analysis of the actual grades students received for the assignments in this research revealed significant inconsistencies between instructors. For example, doctoral students who served as instructors showed little variance in their grades; in one extreme example, a doctoral student gave scores of either 90 or 100 to all groups for their group assignment submissions.

Because of these issues, the self-reported variable of perceived learning was instead used in this study to explore students perceptions of learning from such an activity, suggesting the following two hypotheses:

H2a: Network density will be positively correlated with students' perceptions of learning from the assignment.

H2b: Network clustering coefficient will be positively correlated with students' perceptions of learning from the assignment.

3. Research Methodology

This research is a mixed-methods study [25] that begins with the application of social network analysis to the student citation networks to calculate network-level measures that are compared to the level of structure in the assignments to identify emerging relationships. To extract additional information from the network structures, network visualizations are subsequently inspected to identify visual cues that may provide meaningful insights. Finally, network density and clustering coefficient are examined in relation to self-reported data capturing students' perceptions of learning during the assignment.

3.1. Assignments

The assignments in this study were two-part assignments in which students had to first conduct Internet research for an individual assignment and subsequently share their cited Internet resources with their group members to arrive at a group submission. In total, this study included 10 assignments completed in seven different courses in the Information Systems discipline at a major northeastern polytechnic university. Four of the courses were graduate-level courses and three were undergraduate courses. All courses were taught face-to-face. Assignments were developed with input from the instructors and only after approval of the study from the university's Institutional Review Board.

Assignments ranged from very well-structured assignments that converged on a single solution to very ill-structured assignments in which solutions depended heavily on the information students found. Because students were aware of the group portion of the assignment, even in assignments where students were free to choose a case or topic, group members had to agree on a case or topic in advance to facilitate their knowledge sharing during the second part of the assignment.

Assignments were typically based on prior semesters' assignments; the topics were similar, but the details of the assignments were altered to 1) require Internet-based research to cite digital resources and 2)

include both an individual and a group submission. Table 1 lists each assignment along with a general rating of the structure of the assignment. Assignment structure was evaluated by the first author with input from participating instructors.

Table 1. Assignment identifiers and descriptions

ID	Assignment
A1	**Moderately well-structured**: students had to explain the concept of object-oriented modeling but could select to describe one of several benefits to software development.
A2	**Ill-structured**: students could select any one of three topics (object-oriented databases, database security, or data warehouses); because the topics were broad, students could focus on definitions, examples, or implementations of the topics.
A3	**Ill-structured**: students had to research the concept of nonrelational databases (NoSQL, MongoDB) but could focus on either of the systems mentioned
A4	**Moderately well-structured**: students had to research the 2013 Target data breach and Target's response and relate it to information about TJX's business model and IT strategy.
A5	**Ill-structured**: students had to select three IS failures from a list provided and apply the four components of an information system to analyze the failure.
A6	**Ill-structured**: students had to conduct a case study and could choose from a number of provided cases.
A7	**Well-structured**: students had to find the values for Hofstede's cultural dimensions as calculated for India and apply these dimensions to the structures of two pre-selected Indian consulting firms.
A8	**Ill-structured**: students had to conduct a case study and could choose from a number of provided cases.
A9	**Moderately well-structured**: students had to explain the four types of organization structures as defined in the text and then apply them to the FBI and explain the failure of their virtual case file project.
A10	**Well-structured**: students had to review the Facebook emotional contagion study and identify resources discussing its ethical concerns and justifications. As this debate was emerging at the time of the assignment, limited resources were available, driving students to a converged solution.

Each assignment was a required part of the course activity. Students willing to participate in the research portion of the assignment for extra credit also completed a pre-assignment survey and a post-assignment survey. The procedures, systems, and instruments were pilot tested in one course. Slight modifications to the survey instruments and assignment instructions were made based on the findings from the pilot prior to beginning the larger, quasi-experimental field study that was conducted over two semesters. The research variable perceived learning was captured in the post-assignment survey.

3.2. Student Citation Networks and Network Measures

For each individual assignment, the hyperlinks of the digital media students cited were captured and entered into a matrix (along with the students' identifiers) to create bipartite, bimodal citation networks. (Networks of group hyperlinks were not informative because each group member would have access to the same shared resources.) Bimodal networks include two different types of nodes (in this case, students and hyperlinks to the digital media they cited and stored). Bipartite networks have connections only between different types of nodes – in other words, students can only be connected to digital resources and not directly to other students. The student citation networks were created in Carnegie Mellon's CASOS Organizational Risk Analysis (ORA) software Version 2.2.9. This application facilitates the creation, exploration, manipulation, and visualization of social networks.

Although bipartite, bimodal citation networks are useful for visual analysis because they show both the students and their cited resources, for the purpose of generating network measures, one-mode networks were created using ORA's Matrix Algebra Dot Fold function. The resulting one-mode student networks connect students who cited the same digital media. These one-mode networks were used to calculate the network-level measures of density and clustering coefficient. (In fact, clustering coefficient must be computed on one-mode networks because, in a bipartite network, the clustering coefficient would have a zero value.) These values were subsequently imported into IBM SPSS V22.0.0.1. for statistical analysis.

ORA was also used to generate network visualizations. Visualization is a powerful tool for understanding and investigating data and communicating the meaning of that data to others [12]. Therefore, to extract additional information from the

student citation networks, extreme examples are visually investigated to identify patterns and clues regarding student learning.

3.4. Research Variables

The perceived learning measurement scale included in this research is a validated scale from prior research [4] and was captured through the post-assignment survey. Responses were based on a five-point, Likert-type scale ranging from Strongly Disagree (1) to Strongly Agree (5) with a Neutral option (3). Survey responses were first screened individually for unengaged responses. All data analysis was conducted using IBM SPSS Version 22.0.0.1. A Spearman's rank correlation was used to compare network-level measures (density and clustering coefficient) to perceived learning. Spearman's rho (ρ) was used because the data violated the assumption of normality and has been used in other social network analysis research exploring relationships between performance and network measures [1].

4. Results

Demographic information about participants is provided in Table 2. The final sample contained 210 complete survey responses.

Table 2. Participant demographics

Gender	147 Male (70%)	63 Female (30%)
Educational Level	74 Undergraduate (35.2%)	136 Graduate (64.8%)
Degree Program		
Information Systems	76 (36.2%)	
Other (Information Technology, Computer Science, Business Information Systems)	134 (63.8%)	

4.1. Assignment Structure and Network Characteristics

Research question 1 asks, "How does the design of the learning activity influence the network measures of density and clustering coefficient of the resulting student citation network?" A review of the assignments listed in Table 1 suggests that five of the 10 assignments were ill-structured and more "open" in the sense that the solutions to the problems were open to interpretation. For example, A5 allowed students to identify three IS failures to analyze, where the causes of each failure are open to students' interpretation.

Three of the 10 assignments were moderately well-structured. The solutions to these assignments were more likely to have similar elements and therefore require similar digital resources. For example, A2 asked students to select from three distinct topics about databases. The topics were broad (e.g., data warehouses), suggesting that students' solutions should have similar elements but there was still a lack of convergence on a specific solution.

Conversely, two of the 10 assignment topics were well-structured, providing clear knowledge needs and convergent solutions. These assignments identified very specific topics that left little room for students to branch off in other directions. For example, A7 required that students identify Hofstede's cultural dimensions as they relate to India [16]. This was a very specific topic that drove many students to identify and cite, among others, the same one or two digital resources. The ethical analysis of the Facebook emotional contagion study (A10) [18] was even more well-structured in the sense that it required exploration of a particular study conducted using Facebook, rather than an analysis of Facebook itself. Additionally, A10 was assigned while the debate over the ethics of the emotional contagion study was still occurring, limiting the amount of digital media that was available for students to identify and cite. Table 3 lists the 10 assignments in ascending order of network density and suggests an alignment between the structure of the assignment, network density, and clustering coefficient.

Table 3. Network densities and clustering coefficients

Assign	Density	Clustering Coefficient	
			Ill-Structured
A5	0.032	.000	
A3	0.032	.190	
A2	0.048	.316	
A6	0.108	.322	
A9	0.123	.564	
A4	0.138	.434	
A8	0.250	.484	
A1	0.282	.461	
A10	0.504	.692	
A7	0.692	.814	
			Well-Structured

These results suggest that there is a relationship between the extent to which a research topic is structured and well-defined and the resulting densities of the student citation networks, suggesting that RQ1a is supported. Specifically, it appears that complex and

more highly structured assignments in which there are only single solutions tend to produce denser networks. (This pattern is examined and explained in more detail in the following section.)

Results also suggest that network density and clustering coefficient are sensitive to whether an assignment provides a single topic or multiple topics from which students could choose. The emerging pattern of values suggests that, in assignments with multiple topics, network density may be low (because groups of students are working on different topics), but there is a significant amount of clustering among group members, leading to an increased clustering coefficient; A2 is an illustrative example. Students could choose from one of three topics about databases (group members coordinated before the individual assignment to determine which topic they would work on). Although overall network density was very low (0.048), the clustering coefficient was higher (0.316) because of multiple shared resources within each topic.

Researchers seeking to further explore the relationship between assignment structure, student citation networks, and learning outcomes can take advantage of these relationships to drive resulting network patterns through careful instructional design to provide a single or multiple topics that require convergent solutions (to produce dense networks) or divergent solutions (for sparse network structures). Highly connected yet separate network components can be encouraged through multiple topics, each of which converges towards a single solution.

4.2. Network Visualizations

Because network visualizations can often be helpful in identifying patterns, all networks were visualized, but only extreme cases were examined in detail. In these visualizations of the bipartite, bimodal networks, the red circles represent students, and the green triangles represent their cited digital resources.

One of the networks graphically explored is the network resulting from assignment A5, shown in Figure 1. This assignment instructed students to read about the four components of an information system, review a mini-case, and subsequently research three notable information system failures from a list provided in the textbook. The resulting one-mode network had one of the lowest densities of all the citation networks, as well as a clustering coefficient of zero. An inspection of the resulting bipartite network suggests that students did not conduct much research about the mini-case but focused on researching system failures of various companies. Because students could select any three companies from the list provided, the resulting network is visually fragmented and very

sparse (density=0.032), as shown in Figure 1. It also has extremely low clustering in the folded network, with only one triplet of students sharing a citation.

In general, the characteristics of the networks resulting from ill-structured assignments are visually similar: they have low density, they are visually fragmented, and the nodes in the network are not highly clustered or connected to each other.
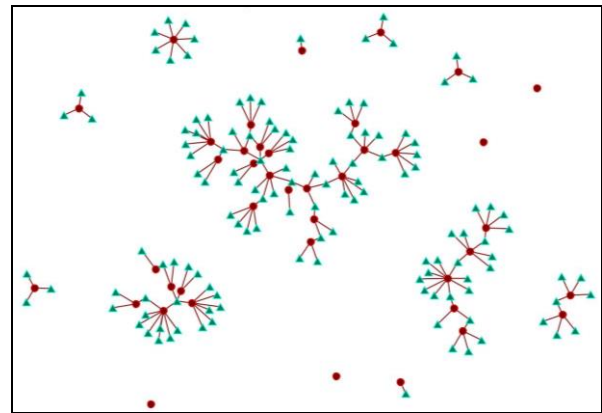

Figure 1. A sparse network with very low density

Alternatively, a visualization of the bipartite, bimodal network of one of the well-structured assignments is shown in Figure 2. Assignment A10 asked students to research the ethical considerations of the Facebook emotional contagion study [18]. The students were first asked to read the study paper and subsequently to find resources that would assist them in analyzing the ethics of the study and how it was conducted. The resulting network is shown in Figure 2.
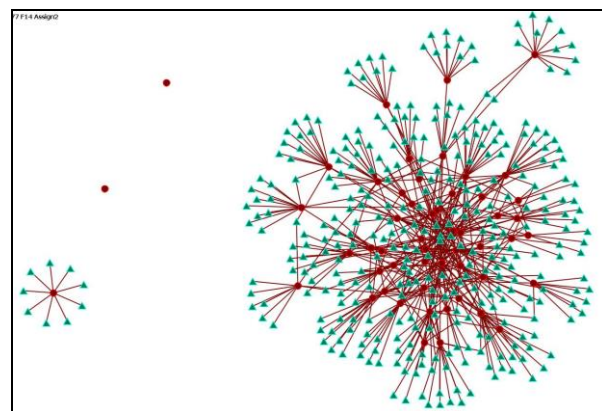

Figure 2. A dense network resulting from a highly structured assignment

In this network, almost all students who participated in the study cited at least one resource in

common with other students, resulting in a very dense network (density=0.504) with high clustering coefficient (0.692). Because not many resources were available about this topic, many students cited the same resources, resulting in the highest network density of all the assignments. Only one student did not share any resources with other students (visualized at lower left). Upon further investigation, this student (an international student) misunderstood the assignment topic and found resources unrelated to the Facebook study. Such outliers may indicate a problem: isolated nodes in otherwise dense networks can serve as a visual cue indicating students who may be off topic and require intervention.

Another interesting pattern emerges in the visualization of the A2 network. This assignment allowed students to select from one of three topics: object-oriented databases, database security, or data warehouses, reflected in the three distinct components shown in Figure 3. This assignment structure resulted in a network with low density (0.048) due to students finding many unique resources, but a higher clustering coefficient (0.316) due to students being highly connected within each topic component.
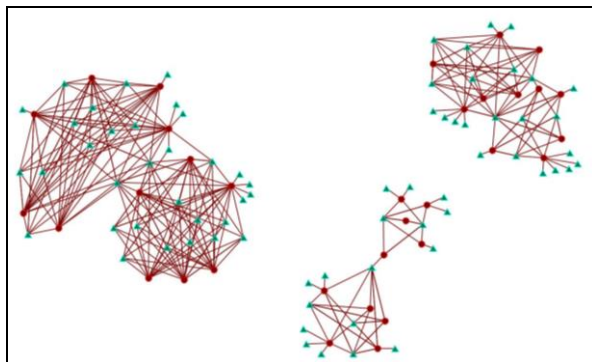

Figure 3. A sparse network with multiple clusters

The component on the left contains the students and resources relating to database security. The sparse component at the lower center contains the students and resources relating to object-oriented databases. The component in the upper right relates to data warehouses. Even though there are no connections between components, the students within each component are well-connected to each other.

However, a closer examination of the structure of each component suggests that students focused on different aspects of the topic they were researching. For example, in the database security component on the left side of Figure 3, students on the upper left-hand side stored and cited digital media that explored instances of data theft and issues relating to database security, while students on the lower right-hand side

found digital resources focusing on database vulnerabilities and prevention methods. Students in the center of the component, bridging these two groups, found resources that were common to both of these topics, suggesting that they developed a broader understanding than students focusing on one aspect of database security or another.

This bridging role is more apparent in the component of students who researched object-oriented databases (shown in the lower center). Here, one student clearly bridges the different aspects of the research topic, with students in the lower area focusing on providing a conceptual explanation of object-oriented databases and students in the upper area storing and citing digital media that discuss actual object-oriented database systems. The data warehouse component similarly has several students who act as bridges connecting different perspectives of the research topic.

Together, these visualizations indicate that there is a value in looking beyond network measures to explore the visual patterns in the networks.

4.3. Network Densities and Perceptions of the Assignment

To explore RQ2, Spearman's rho (ρ) was calculated to evaluate the relationship between each assignment's network density, clustering coefficient, and perceived learning (H2). Table 4 shows the resulting correlations, which are significant at the 95% confidence level.

Table 4. Spearman's correlations between perceived learning and network measures

Variable	ρ	Sign.
H2a: Network density with perceived learning	0.657	0.04
H2b: Clustering coefficient with perceived learning	0.745	0.01

While both correlations were significant at the 95% confidence level, clustering coefficient was more highly correlated with perceived learning than network density. This is understandable considering that density can be affected by assignments in which students could select from several topics or cases, while clustering coefficient more accurately reflects the amount of similar resources shared by students. Together, these findings suggest that students feel they learn better when they individually identify at least some of the same resources as their groupmates.

5. Discussion

Using student citation networks generated through Internet-based research assignments, this study explores the relationships between assignment structure and the characteristics of the resulting networks as well as the relationships between network-level measures and students' perceptions of learning from the assignment.

In response to RQ1, this study finds that network densities and clustering coefficients do increase as the amount of structure and definition of the research topics also increase. More specifically, network density increases as the structure and convergence of an assignment also increases, but is sensitive to assignments allowing students to choose from a set of topics. Clustering coefficient, on the other hand, is more closely tied to the amount of convergence in the assignment solutions, even when multiple topics are available to students.

These findings suggest strategies for instructors and researchers seeking to create assignments that would result in either sparse or dense citation networks with single or multiple components. Well-structured topics that are clearly defined and that converge towards a single solution will result in more similar resources being cited, leading to an increased clustering coefficient. Ill-structured topics that are broad and do not offer a single solution will result in lower network density and lower clustering coefficient. Assignments offering multiple topics are more likely to result in individual components that may result in lower network density while clustering coefficient will be more sensitive to the amount of solution convergence within each topic.

Visually representing the resulting citation networks can provide important feedback to researchers and instructors alike, helping to understand the patterns in network density and clustering coefficient. Moreover, visualizations of citation networks can be used to explore the importance of students' positions in such networks. For example, isolated nodes in otherwise dense networks may indicate students requiring intervention, while students in bridging positions may have a broader perspective than students at the periphery of the network.

The opportunity to extract richer contextual information from citation network visualizations may also be of interest for system designers or researchers developing formative feedback learning dashboards. Students who find themselves isolated in otherwise dense networks (such as A10) may question why they are disconnected from all other students in the course, prompting them to realize that they are off-topic. Similarly, instructors seeing students isolated in this way can intervene to ensure that the student understands the learning activity. Students who see themselves at one end of a cluster may expend additional effort to explore some of the digital content stored by students on the opposite end of the cluster to gain a broader understanding of the research topic. Instructors may wish to identify students who are in bridging positions in these networks and have them mentor other students or share their digital resources with others.

Evaluating the relationships between perceived learning, network density, and clustering coefficient, results of RQ2 suggest that there is a positive relationship between network density and students' perceptions of learning from the assignment, but that the relationship between clustering coefficient and perceived learning is even stronger. Together, these relationships provide a foundation upon which future research can build in evaluating relationships between network measures and learning. Instructors and instructional designers can improve perceptions of learning through careful design of research assignments featuring one or multiple topics with solutions that converge on a single solution.

In summary, this paper makes several contributions to the domains of educational research and learning analytics. First, it shows the importance of students finding and citing similar digital resources when learning collaboratively, particularly when those resources must be shared as part of a group assignment. Secondly, it begins to untangle the relationship between network characteristics, assignment structure, and perceptions of learning.

6. Limitations and Future Work

This study has several factors that limit its generalizability. First, it was conducted at a single northeastern polytechnic university; results may vary if conducted at other universities, particularly non-technical universities. Additionally, all of the courses included in this study were in the Information Systems discipline, although students in the courses were enrolled in a variety of majors. Students from other programs, and courses in other disciplines, may influence the findings. Finally, all of the courses were taught in a traditional, face-to-face environment. Results may differ when this study is replicated in online or distance education courses.

Methodologically, this study only examined two prominent SNA measures: network density and clustering coefficient. Future work will examine other network-level measures such as diameter, centralization and modularity as well as node-level

centrality measures that have also proven useful in the learning [33] and community building [34, 35] contexts. Also, from the data collection perspective, this study would be difficult (i.e. tedious) to replicate. All hyperlink citations were extracted manually from submitted assignments to create the network matrices. In the future, add-ons could be developed for tools such as Turnitin so that links to cited resources could be automatically extracted.

Finally, this research captures student self-reported data. Although grades were captured for those students who gave permission through the Family and Educational Rights to Privacy Act (FERPA), actual assignment grades proved problematic due to instructor inconsistencies and/or lack of variance in grades. Future research could introduce scoring rubrics and trained raters to provide a more objective assessment of learning and the structure of the assignment as evidenced by the student's assignment submission.

7. Conclusion

Researchers in the field of learning analytics are exploring various types of learning trace data to predict or improve learning. Still in its infancy, this field frequently takes an exploratory approach, identifying available data sources and examining their relationship to other data sources or to reported learning. This study uses this approach to conduct an exploratory study of students' citation networks that result from assignments in which students are encouraged to conduct research using digital media found on the Internet. Results suggest that student citation networks have the potential to provide meaningful learning feedback, both statistically and visually, to learners and instructors alike. By understanding how to drive network density and clustering coefficient through appropriate assignment design, researchers can build on these results to refine a framework linking assignments, cited digital media, network characteristics, and learning.

8. Acknowledgements

9. References

[1] Abbasi, A., J. Altmann, and L. Hossain, "Identifying the effects of coauthorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures," *Journal of Infometrics* 5, 2011, pp. 594-607.

[2] Alelyani, T., R. Collins, S.R. Hiltz, and Y. Xiong, "Not just silly cat videos: Exploring student knowledge sharing via social media," *Proceedings of Americas Conference on Information Systems*, 2015, n.p.

[3] Aviv, R., Z. Erlich, G. Ravid, and A. Geva, "Network analysis of knowledge construction in asynchronous learning networks," *Journal of Asynchronous Learning Networks* 7(3), 2003, pp. 1-23.

[4] Benbunan-Fich, R., and J.B. Arbaugh, "Separating the effects of knowledge construction and group collaboration in learning outcomes of web-based courses," *Information & Management* 43(6), 2006, pp. 778-793.

[5] Borgatti, S.P., M.G. Everett, and J.C. Johnson, *Analyzing Social Networks*, SAGE Publications, Thousand Oaks, CA, 2013.

[6] Bruce, C., "Information literacy as a catalyst for educational change. A background paper," 2004, pp. 8-19.

[7] Buckingham, D., "Digital media literacies: Rethinking media education in the age of the Internet," *Research in Comparative and International Education* 2(1), 2007, pp. 43-55.

[8] Connaway, L.S., M.L. Radford, T.J. Dickey, J. De Angelis Williams, and P. Confer, "Sense-making and synchronicity: Information-seeking behaviors of Millennials and Baby Boomers," *Libri* 58(2), 2008, pp. 123-135.

[9] De Laat, M., V. Lally, L. Lipponen, and R.J. Simons, "Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for Social Network Analysis," *International Journal of Computer-Supported Collaborative Learning* 2(1), 2007, pp. 87-103.

[10] Ding, Y., "Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks," *Journal of Infometrics* 5(1), 2011, pp. 187-203.

[11] Eid, M.I.M., and M. Al-Jabri, "Social networking, knowledge sharing, and student learning: The case of university students," *Computers & Education* 99, 2016, pp. 14-27.

[12] Few, S., and PERCEPTUAL EDGE, "Data visualization: past, present, and future," *IBM Cognos Innovation Center*, 2007, n.p.

[13] Greeno, J.G., "Natures of problem-solving abilities," *Handbook of learning and cognitive processes* 5, 1978, pp. 239-270.

[14] Haythornthwaite, C., "Social network analysis: an approach and technique for the study of information," *Library & Information Science Research* 18(4), 1996, pp. 323-342.

[15] Haythornthwaite, C., M. De Laat, S. Dawson, and D. Suthers, "Introduction to learning analytics and networked learning minitrack," *Proceedings of Hawaii International Conference on System Sciences*, 2013, pp. 3077.

[16] Hofstede, G., *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.) Sage Publications, Thousand Oaks, CA, 2001.

[17] Jonassen, D., "Toward a design theory of problem solving," *Educational technology research and development* 48(4), 2000, pp. 63-85.

[18] Kramer, A.D. "The spread of emotion via Facebook," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012, pp. 767-770.

[19] Martinez, A., Y. Dimitriadis, B. Rubio, E. Gomez, and P. De la Fuente, "Combining qualitative evaluation and social network analysis for the study of classroom social interactions," *Computers & Education* 41(4), 2003, pp. 353-368.

[20] Nurmela, K., E. Lehtinen, and T. Palonen, "Evaluating CSCL log files by social network analysis," *Proceedings of the 1999 conference on Computer Support for Collaborative Learning*, 1999, pp. 54.

[21] Otte, E., and R. Rousseau, "Social network analysis: a powerful strategy, also for the information sciences," *Journal of Information Science* 28(6), 2002, pp. 441-453.

[22] Prensky, M., "Listen to the natives," *Educational Leadership* 63(4), 2005, pp. 8-13.

[23] Reffay, C., and T. Chanier, How social network analysis can help to measures cohesion in collaborative distance-learning," *Designing for Change in Networked Learning Environments*, Springer, Netherlands, 2003, pp. 343-352.

[24] Siemens, G., D. Gasevic, C. Haythorthwaite, S. Dawson, S.B. Shum, R. Ferguson, E Duval, K. Verbert, and R.S. J.D. Baker, *Open Learning Analytics: an integrated & modularized platform*, Doctoral dissertation, Open University Press, 2011.

[25] Venkatesh, V., S.A. Brown, and H. Bala, "Bridging the Qualitative-Quantitative Divide: Guidelines for conducting mixed methods research in Information Systems," *MIS Quarterly* 37(1), 2013, pp. 21-54.

[26] Viegas, F.B., and J. Donath, "Social network visualization: Can we go beyond the graph," *Workshop on social networks, CSCW*, 2004, pp. 6-10.

[27] Wasserman, S., and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, New York, NY, 1994.

[28] White, C.M., *Social media, crisis communication, and emergency management: Leveraging Web 2.0 technologies*, CRC Press, Boca Raton, FL, 2011.

[29] Wise, A., Y. Zhao, and S. Hausknecht, "Learning analytics for online discussions: Embedded and extracted approaches," *Journal of Learning Analytics* 1(2), 2014, pp. 48-71.

[30] Watts, D.J., and S.H. Strogatz, "Collective dynamics of 'small-world' networks", *Nature* 393(6684), 1998, pp. 440 – 442.

[31] Mason, W, and D.J. Watts, "Collaborative learning in networks," *Proceedings of the National Academy of Sciences* 109(3), 2012, pp. 764-769.

[32] Hamulic, I., and N. Bijedic, "Social network analysis in virtual learning community at faculty of information technology (fit), Mostar," *Procedia Social and Behavioral Sciences* 1, 2009, pp. 2269-2273.

[33] Gruzd, A., D. Paulin and C. Haythornthwaite, "Analyzing Social Media and Learning Through Content and Social Network Analysis: A Faceted Methodological Approach," *Journal of Learning Analytics* 3(3), 2016.

[34] Gruzd, A. and K. Tsyganova, "Information wars and online activism during the 2013/2014 crisis in Ukraine: Examining the social structures of Pro- and Anti-Maidan groups," *Policy & Internet* 7(2), 2015, 121–158.

[35] Gruzd, A. and C. Haythornthwaite, "Enabling community through social media," *Journal of Medical Internet Research* 15(10):e248, 2013.

[36] Collins, R., "Assessing learning outcomes and social capital formation resulting from the use and sharing of Internet knowledge resources," Doctoral Dissertation, 2015, NJIT.

[37] Newman, M. *Networks: an introduction*. Oxford University Press, Oxford, 2010.