

# **Leveraging the Louvain Algorithm for Enhanced Group Formation and Collaboration in Online Learning Environments**

**Minkyung Lee<sup>1</sup>, Priya Sharma<sup>2</sup>**

**<sup>1,2</sup> Department of Learning and Performance Systems, Pennsylvania State University**

Minkyung Lee

Ph.D. Candidate at the Department of Learning and Performance Systems  
Research Assistant at The Leonhard Center in the College of Engineering  
Penn State University

Email: [minkle010@gmail.com](mailto:minkle010@gmail.com)

Priya Sharma

Associate Professor at the Department of Learning and Performance Systems  
Penn State University

Email: [pus3@psu.edu](mailto:pus3@psu.edu)

We have no known conflict of interest to disclose.

Corresponding concerning this article should be addressed to Minkyung Lee,  
[minkle010@gmail.com](mailto:minkle010@gmail.com).

# **Leveraging the Louvain Algorithm for Enhanced Group Formation and Collaboration in Online Learning Environments**

## **Abstract**

This study explores the dynamics of student interaction networks within an online asynchronous discussion forum, focusing on both whole group discussions and subgroup interactions distinguished by the Louvain algorithm, a renowned community detection method. Analyzing 2481 posts from 154 undergraduate students across three sections of a communications course centered on discussions about movie clips or social phenomena to enhance media literacy, this research aims to interpret the interaction patterns in these virtual spaces.

Traditional methods of group formation, such as teacher intervention and self-selection, often fail to create balanced and effective groups, especially in large online courses. The Louvain algorithm, known for its efficiency in modularity optimization, identifies clusters based on actual student interaction patterns. By leveraging both global and local network analyses, the study provides a comprehensive understanding of interaction structures. The global network analysis offers a macro view of overall interaction structures, revealing diverse patterns despite identical course designs, suggesting the influence of unique group dynamics. The local analysis, focusing on the intricacies of node and edge connections, underscores that the Louvain algorithm's classifications correlate with heightened cohesiveness and collaborative potential.

The results indicate that algorithmically detected groups exhibit strong internal communication and cohesiveness, as evidenced by high clustering coefficients, density values, and weighted degrees. These findings underscore the potential of network analysis to optimize online student interactions, providing valuable insights for refining educational design to promote student engagement and collaborative problem-solving. This research highlights the transformative potential of integrating advanced data-driven techniques in educational technology to improve group formation and collaborative learning outcomes, offering empirical insights for educators to enhance online interactions and expand pedagogical understanding.

**Keywords:** Louvain algorithm, Group Formation, Asynchronous Discussion, Cohesiveness, Network Analysis, Online Learning, Community Detection, Collaborative Learning, Educational Technology, Media Literacy, Students Engagement

## Introduction

The rise of online education has introduced new challenges in effectively forming groups that encourage participation and enhance the learning experience. Unlike traditional classrooms, online environments lack physical presence and spontaneous interactions, making it difficult for educators to facilitate student engagement. This study investigates the use of the Louvain algorithm, a method from network analysis, to improve group formation and collaboration in online learning environments.

Group projects and peer interactions in traditional classrooms support deeper comprehension and critical thinking (Springer et al., 1999). However, translating these benefits to an online domain, with its geographical barriers, requires innovative approaches. Group discussions in online courses serve as crucial spaces for student interaction. Yet, forming effective groups in online courses, especially those with extensive enrollments, remains a challenge due to its impact on collaborative learning outcomes (Kozlowski & Bell, 2003). This challenge is a critical research agenda in Computer Supported Collaborative Learning (CSCL) (Ouyang & Scharber, 2017; Rodríguez et al., 2011; Weinberger et al., 2005).

The existing literature primarily explores traditional group formation methods, such as teacher intervention and self-selection, which have limitations in ensuring balanced and effective groups (Felder & Brent, 2001). While some studies have attempted algorithmic solutions for automatic group formation (Müller et al., 2022), these have largely been based on students' demographic and psychological characteristics. However, there is a lack of research exploring the utility of network science algorithms, like the Louvain algorithm, for group formation based on actual interaction patterns.

The Louvain algorithm, renowned for its efficiency in modularity optimization in networks (Blondel et al., 2008), is employed here to discern clusters among students based on their interaction patterns. This study hypothesizes that these algorithmically detected groups correspond to naturally occurring collaborative units. By analyzing interaction data from an undergraduate communications course, this study aims to assess the algorithm's efficacy in identifying cohesive and collaborative student groups.

In addition to facilitating group formation, the Louvain algorithm provides analytical insights into collaborative dynamics within digital educational spaces. These insights are essential for refining educational design to promote student engagement and collaborative problem-solving (Rodríguez et al., 2011). Accordingly, this paper outlines our methodology, presents findings from the Louvain algorithm's application, and discusses its potential to enhance the online educational experience. The goal is to demonstrate the value of network analysis in educational technology, offering a fresh perspective on optimizing online student interactions for improved educational outcomes.

## Background

### Online Learning Communities

The notion of "community" is integral to effectiveness and satisfaction in online learning environments. Rovai (2002) and Wenger (1998) have emphasized that a sense of community in online courses enhances the students' overall learning experience and positively impacts their academic outcomes. Within the expanding domain of online education, fostering a sense of community has gained considerable attention from researchers. Much research ascertained that a strong sense of community in online learning environments elevates the quality of educational experiences and positively influences academic outcomes (Shea et al., 2006; Singh et al., 2022).

This perspective gains extra significance in the context of Asynchronous Online Discussions (AOD). The asynchronous nature of these platforms such as discussion board can improve the role of community, as the absence of real-time interaction necessitates other forms of rich, meaningful discussion (Author\_1, 2021; Wise et al., 2012).

These discussions serve as more than only informational exchanges; they evolve into environments for developing skills such as critical thinking and problem-solving. The Community of Inquiry (CoI) framework (Garrison et al., 2000) highlights that such interactions are vital for fostering an online learning environment where knowledge is co-constructed. Within such communities, the collaborative engagement among peers serves as a medium for social, cognitive, and teaching presence, the core elements of the CoI framework. A study by Xie et al. (2014) indicated that well-facilitated online communities could lead to higher grades, increased engagement, and a deeper understanding of course content. This becomes increasingly significant in the situation where student retention is a concern. Rovai (2007) and, more recently, Peacock & Cowan (2019) suggest that a sense of belonging, often nurtured through strong community interactions, can be crucial in retaining students in online courses. However, the process of forming these communities, or discussion groups, remains an open question. Traditional methods like teacher intervention and self-selection have been critiqued for their limitations in ensuring balanced and effective groups (Felder & Brent, 2001). While some have attempted algorithmic solutions for automatic group formation (Müller et al., 2022), these have primarily been based on students' demographic and psychological characteristics.

Herein lies the research gap this study aims to fill. While the existing literature provides a foundation for the importance of online learning communities and suggests the necessity for effective group formation, there is few research found exploring the utility of network science

algorithms like the Louvain algorithm for this purpose (Author\_1& Author\_2). The Louvain algorithm, known for its efficiency in modularity optimization in networks (Blondel et al., 2008), offers a data-driven approach to cluster students based on their actual interactions, thereby promising more cohesive and naturally formed groups.

### *Group formation*

Group formation is vital for effective collaborative learning since collaborative learning is based on student interaction (Cohen 1994). This generally involves organizing the arrangement of group members, and the groups are composed of the attributes of members in a group, and these factors heavily impact the effectiveness and efficiency of group processes (Müller et al., 2022). According to Bell (2007), group composition variables can be categorized into surface-level and deep-level. Surface-level variables refer to demographic characteristics such as age, race, education level, and organizational tenure, while deep-level variables refer to psychological traits such as personality factors, values, and attitudes. Bell suggests that these team composition variables can be used to compose teams and improve team performance effectively with the consideration of contexts. Many small group studies focus on group composition, examining heterogeneous groups compared to homogeneous groups in collaborative situations. For example, Donovan et al. (2018) investigated how students with different levels of competence in biology in different group compositions and found low competence students had better learning outcomes in heterogeneous groups. In contrast, Jensen & Lawson (2011) looked at heterogeneous upper-level biotechnology lab groups and found that students paired with students of a different academic level (undergrad and grad) earned better grades than those at the same level. However, Miller et al. (2012) found no differences between

students working in homogeneous or heterogeneous groups in undergraduate physics courses. Likewise, different group compositions can influence students' learning.

There are different methods to form groups: random selection, self-selection, teacher selection, and automatic selection (Felder & Brent, 2001). Random selection is the simplest method, as it involves randomly selecting students from the pool. This method can be quick to implement but may result in unbalanced groups of members with different skills or interests. Thus, it may not be the best approach for tasks requiring collaboration and teamwork, as students may be unable to work effectively with their group mates. Self-selection, however, allows students to choose their group mates based on factors such as shared interests, skills, or prior relationships. This method can promote collaboration and engagement, as students have a voice in whom they work with. However, it may also result in groups that lack diversity or are dominated by a few individuals. Additionally, some students may feel excluded if they cannot find a group that fits their needs. Teacher selection involves the teacher manually assigning students to groups based on a range of criteria, such as skills, interests, or learning needs. This method can ensure that groups are balanced and well-suited for the task at hand, but it can be time-consuming and not feasible for large groups of students. Moreover, the teacher's knowledge of individual students may not be comprehensive, resulting in some students being placed in groups that are not optimal for their needs.

Finally, automatic selection uses algorithms to assign students to groups based on criteria such as skills, interests, or learning needs. This method can be both effective and economically superior, as optimal groupings can be quickly generated with a given set of criteria (Muller et al., 2022; Vallès-Català & Palau, 2023). Vallès-Català & Palau (2023) created an approach based on complex network theory to design an algorithm called Minimum Entropy Collaborative

Groupings (MECG) to form heterogeneous groups more effectively. They found that the groups created with MECG were more effective with less uncertainty in comparison to the randomized groups. The implementation of an automatic approach which would consider a range of learning needs and employ intelligent techniques for the structure of optimal student groups, is essential for promoting effective collaborative learning. Muller et al. (2022) highlight that automatic group formation based on deep level personality traits, such as extraversion and conscientiousness, can also lead to better satisfaction and group performance compared to other group formations. Krouska & Virvou (2020) emphasized genetic algorithms focused on three dimensions, academic, social, and cognitive dimensions for heterogeneous group formation and social learning.

Likewise, the criteria for algorithms varies by researchers and purposes and no general or evidence based recommendation or set of consistent criteria for feeding the algorithm for group formation exists. Additionally, it may not be possible to create an equally successful distribution for all groups simultaneously, as there may be contextual trade-offs between different criteria. Therefore, when using algorithms to group students, it is important to carefully consider the criteria used, as these algorithms will not be able to capture all relevant student characteristics. To address this challenge, Pai et al. (2014) and Muller et al. (2022) suggest that developing user-friendly online applications that facilitate the selection and weighting of relevant criteria and procedures for grouping could be a feasible solution. These applications could enable a more interactive and intuitive way of inputting and manipulating data, potentially resulting in more comprehensive and accurate groupings of students. Yassine et al. (2022) also highlighted the need to automate community detection techniques in online learning environments.



### *Louvain algorithm: Group formation by frequency of interaction*

Interactions serve as a multifaceted metric that offers invaluable insights into various aspects of the educational experience in online learning environment operating as a tangible measure for student engagement, allowing educators to distinguish between active and passive learners for targeted support (Moore & Kearsley, 2011). Interactions also reveal naturally forming learning clusters within an online course, which are crucial for peer support and collaborative learning (Rovai 2002; Dawson 2008). By analyzing the flow of interactions, educators can identify key nodes or students who significantly influence information dissemination and discussion facilitation (Haythornthwaite, 2001; Cela et al., 2015). This understanding of interaction patterns is helpful when forming groups for assignments or projects, as it enables strategic group formation to either leverage existing dynamics or encourage new interactions (Saqr et al., 2020). Moreover, the quality and quantity of these interactions have been shown to correlate with academic outcomes, making them a predictive tool for student success in numerous research (Arbaugh, 2008; Wise et al., 2014). Therefore, interactions are not just incidental exchanges but critical metrics that can shape, assess, and enhance the online learning experience. Likewise, SNA has been used to capture learner interaction and how the relationships affect learning outcomes. The purpose of using SNA is generally to assess whether and to what degree interaction and collaboration occur (Froehlich et al., 2020).

A few studies have focused on SNA to identify communities and examine learning. Jan & Vlachopoulos (2018) used the Integrated Methodological Framework using SNA to structurally identify communities in online learning. They suggest that community-based learning and structural similarity between networks and communities make SNA a natural choice for deeper understanding of group interaction. Their method emerged as an effective framework

for structural identification of a Community of Inquiry (CoI) and Community of practice (CoP). Yassine et al. (2021) used the label propagation algorithm and Louvain algorithm for community detection and found those different community detection algorithms can be implemented on learning networks to detect communities. In community detection algorithms, communities are defined as nodes with similar affiliations that are different from the rest of the network (Yang et al., 2010) or as cohesive network structures with the possibility of separation (Newman 2018). This concept of cohesiveness is not just theoretically rich but also empirically validated. Numerous studies have found that cohesion is positively related to various team effectiveness outcomes. Team cohesion is a complex, multidimensional emergent state extensively studied for its critical role in teamwork and team effectiveness (Beal et al., 2003). LePine et al. (2008) define it as the attraction and commitment team members feel towards their team and its tasks and Tekleab et al. (2009) highlight the group's united pursuit of goals as a key element of cohesion. Recently, Braun et al. (2020) defined it as “team members' shared commitment or attraction to their task/goal and to one another.” In the context of community detection, especially in SNA, cohesion typically refers to the degree to which members within a particular community or groups are more closely interconnected. A highly cohesive community is one where its members have many connections among themselves, but fewer connections to nodes outside the community. Thus, community can be separated when the internal links are larger than the external links, so a strong community has more internal links than external links (Wasserman & Faust, 1994). Though communities can be detected using a variety of methods, modularity-based algorithms are most frequently cited in relation to community detection.

The Louvain algorithm (Blondel et al., 2008) one of the widely used modularity optimization methods (Menczer et al., 2018), is interactive for detecting communities in

networks through modularity optimization. The algorithm consists of two main steps that are repeated until modularity can no longer be improved, ultimately determining the partition with the highest modularity. Initially, the algorithm assigns each node to its own community, and examines neighboring communities for potential modularity gains if a node were to be reassigned. Nodes are then moved to the neighboring community that results in the maximum modularity gain or they remain in their current community if no gain is possible. Following this, the network is simplified by replacing communities with single super-nodes. Links between distinct communities are transformed into weighted links between corresponding super-nodes, with weights based on the sum of original link weights. Additionally, internal community links are converted into self-loops for super-nodes, where self-loop weights are equal to the sum of the original internal link weights. Upon completing these two steps, the algorithm iterates again with the new, smaller weighted network generated in the second step. This process of reassigning nodes to communities and aggregating communities into super-nodes continues until no further modularity improvement is possible. When the algorithm converges, the final partition of nodes into communities represents the network's structure with the highest modularity. Our study incorporates the notion of "algorithmically formed communities," which are based on metrics like frequency and type of interaction, derived through the Louvain algorithm with a social network generation tool, Gephi.

### Purpose of the study

This study began with the primary objective of examining student interaction patterns across different online courses. Originally focused on identifying and comparing interaction patterns (RQ1), the study's scope expanded as initial findings suggested the potential utility of

the Louvain algorithm for classifying these interactions. The algorithm uses actual data to classify students into groups ensuring that the classification is based on actual interaction patterns as a data driven classification, so it identifies natural clusters within the data, revealing how students group organically themselves based on their interactions. This led to a deeper exploration of how the algorithm could effectively categorize interactions and influence our understanding of collaborative dynamics within student groups. This adaptive, exploratory approach reflects the emergent nature of the research (Maxwell 2012), leading to the development of additional research questions that address both the initial insights and the subsequent findings. The research questions are structured to reflect this progression:

1. Initial exploration of interaction patterns: How do interaction patterns among students vary across different online courses?
  - a. Are there discernible patterns that differentiate courses based on interaction metrics?
  - b. What global metrics (e.g., average degree, diameter, density) highlight the nature and quality of student interactions in these courses?
2. Emergent focus on classification effectiveness: Following the detection of significant patterns in initial analyses, how effectively does the Louvain algorithm classify students based on their interaction patterns?
  - a. How do local metrics (such as clustering coefficient, density, weighted degree, and the presence of triangles) vary across the groups formed by the Louvain algorithm?
  - b. What patterns within these metrics does the Louvain algorithm identify to distinguish one group from another?

- c. how can the correlations between local clustering coefficient, density, weighted degree, and triangles deepen our understanding of group collaboration and cohesiveness within student networks?
- d. Implications for Collaborative Dynamics: Given the insights from the Louvain algorithm's application, how does the algorithm's classification reflect the collaborative nature and cohesiveness of these groups?

This structure explicitly acknowledges the exploratory and iterative nature of the research, starting from a broad inquiry into interaction patterns and progressively focusing on specific analytical tools and their implications for understanding group dynamics in online learning environments. This phased approach underlines the flexibility and responsiveness of the research process to emerging data and insights, characteristic of exploratory studies.

## Methods

### Course Selection

The selection criteria were to include courses with sufficient interaction data to provide robust analytical outcomes while ensuring the representativeness of various interaction dynamics within the online learning environment. Three sections of the same course were selected for this study based on minimum course duration, the amount of interaction and the number of participants.

*Table 1. Course Information*

	Students	Posts	Topics
C4	53	808	9

C5	50	764	9
C6	51	839	9

---

The course in online undergraduate communications provided a rich and structured dataset that is relevant for exploring the research theme of group composition and interaction patterns. The course emphasizes media literacy, critical thinking, and analytical viewing skills, incorporating lectures, readings, and film screenings. It focuses on how films tell stories and make arguments through visual language, allowing students to interpret and analyze films to understand the cultural conditions that produced them. The Lesson Discussion Forums consisted of general question prompts to initiate discussion. There were eight odd-numbered lessons in total. Students were encouraged to engage in the discussion by drawing upon what they learned from the lesson. They were required to respond to the question thoughtfully and engage the ideas of their classmates by responding to at least two other students' posts. Respectful and constructive interaction was emphasized, and the total discussion participation average accounted for 20% of the final grade. This context is ideal for examining student interactions as it involves active engagement, critical discussions, and collaborative learning. By analyzing how students interact and form groups in the context of this course, the research can contribute to a broader understanding of collaborative learning in online environments.

### Data Collection and Processing

The research utilized anonymized discussion data from an undergraduate communications course focused on films, designed to enhance media literacy. The initial phase involved a rigorous data cleaning process to ensure participant anonymity by removing personally identifiable information. Additionally, irrelevant or redundant data were

systematically excluded to streamline the subsequent analysis. After cleaning, the data underwent a transformation process, where raw discussion posts were restructured into nodes, and replies were formatted as edges. This effectively converted the conversations into a graph-based representation amenable to network analysis.

For each section of the course, adjacency lists were compiled, detailing the interactions between students. These lists were crucial for both the visualization and detailed analysis of the network's structure and dynamics. The cleaned and formatted data was then imported into Gephi, an advanced tool for network visualization and analysis. Key network metrics such as weighted degree, diameter, and density were calculated, which are pivotal for assessing the interaction patterns and connectivity among students. To ensure the reliability and accuracy of the findings, the data was also validated using software packages R and SPSS.

### Analytical Approach

The analysis was conducted in two main stages to thoroughly explore student interaction patterns and the effectiveness of community detection algorithms in an educational setting. Initially planned as an analysis of student interaction patterns, the research evolved into an exploratory investigation into the utility of the Louvain algorithm for group composition. This shift occurred as initial data analysis revealed the algorithm's potential in uncovering hidden patterns in student interactions, providing a practical application of network analysis in collaborative learning. The study's methodology emphasizes the potential of network analysis to enhance the educational experience by optimizing collaborative learning dynamics and identifying natural group formations.

- **Global Network Analysis:** The study first examined the networks at a global level to understand their overall structure and interaction patterns. Metrics such as graph density, centralization, and clustering coefficient were calculated to provide insights into the topology and connectivity of the network. This global perspective helps to set the stage for a more detailed local analysis by identifying overarching patterns in student interactions.
- **Local Network Analysis:** This stage focused more deeply on the communities identified by the Louvain algorithm, a key component in our methodological approach. Here's how we conducted this analysis:
  - **Community Detection:** Using Gephi's built-in modularity function, which employs the Louvain algorithm, the study segmented the network based on patterns of connectivity among nodes. This process optimizes the modularity score to gauge the strength of divisions into distinct communities. The study kept the modularity resolution setting in Gephi at its default to ensure consistency across all datasets, based on preliminary tests that confirmed its effectiveness in accurately reflecting community structures within our data.
  - **Analysis of Local Metrics:** the study examined several local metrics within the identified communities, such as the local clustering coefficient, within-community density, and the roles of central nodes. These metrics helped us to understand the cohesiveness and collaborative potential of the groups formed within the network. Specifically, the study explored:
    - How local metrics vary across groups formed by the Louvain algorithm.
    - Patterns within these metrics that the Louvain algorithm identifies to distinguish one group from another.



- How correlations between local clustering coefficient, density, weighted degree, and triangles contribute to our understanding of group collaboration and cohesiveness within student networks.

This detailed examination of local network dynamics enriches our understanding of how students interact within smaller, more defined groups and provides insights into the structural and functional aspects of these interactions. It highlights the utility of the Louvain algorithm in identifying meaningful interaction patterns and naturally occurring student communities.

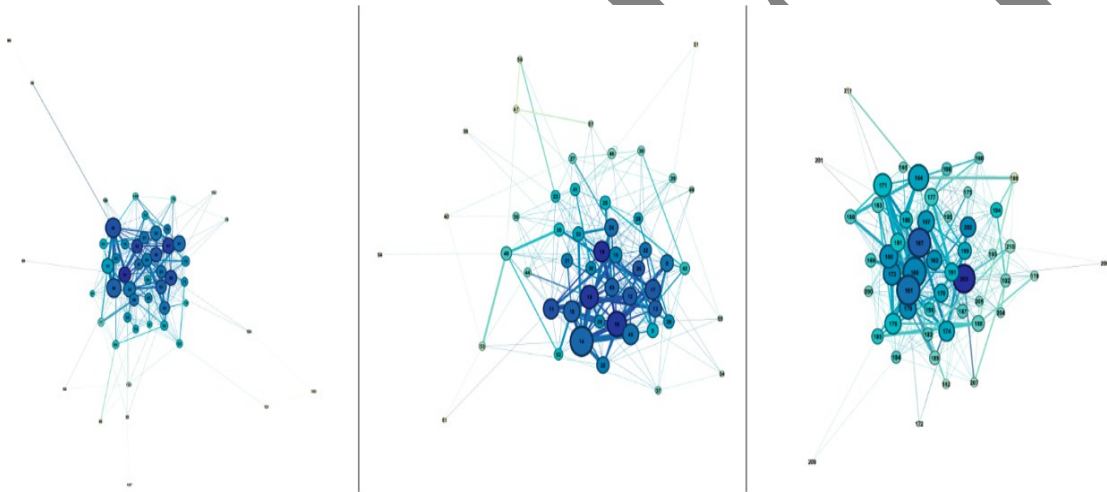
## Results

Networks can be analyzed from two perspectives: local and global. The local level examines specific nodes or edges, while the global level provides an overview of the entire network (Miele et al., 2019). Morrison et al. (2022) noted the importance of using both types of analysis for a thorough understanding. Following this idea, study leverages both types of analysis to gain a comprehensive understanding of interaction dynamics. The study began by assessing the networks globally to understand the overarching structure and subsequently focused on the group cohesions at the local level identified by the Louvain algorithm.

### Global metrics (RQ1)

Global metrics offer understanding into the overall characteristics, topology, and structure of networks (Morrison et al., 2022). Metrics such as density, centralization, and clustering coefficient provide a holistic view of a network's structure, and particularly, density and diameter are often used to compare the structures and extents of different networks.

For the research question 1 as an initial exploration of interaction patterns, the study focuses on how interaction patterns among students vary across different online courses. The first sub-question asking if there are discernible patterns that differentiate courses based on interaction metrics, the study found there are clear visual differences in the connectivity patterns across courses from the network diagrams even though those were with the same course content and instructor but the differences were only with students participating the course discussion. Three networks of the courses C4, C5, and C6 are shown in Figure1.



*Figure 1. Networks of C4, C5, and C6 (from left)*

These networks represent the structures of communication networks within different course environments. Each node represents a student, with its size representing its weighted degree, indicating the intensity of interactions. Nodes with darker colors indicate hubs or students with higher connectivity, emphasizing relative connectivity rather than just the number of connections. These different networks depict varied student participation and interaction

behaviors. For instance, C6 appears denser with more connections than C4 and C5. This observation is substantiated when considering the following global metrics in Table 2.

*Table 2. Global Network Metrics for three sections of an online course*

Metric	Description	C4	C5	C6
<b>Number of Nodes</b>	Total number of students participating in the forum	53	50	51
<b>Number of Edges</b>	Total number of interactions between students	314	261	340
<b>Graph Density</b>	Ratio of actual connections to possible connections, higher values indicate more interconnectivity	0.228	0.213	0.267
<b>Weighted Degree</b>	Average level of connectivity per student, indicating interaction intensity	16.679	14.960	17.294
<b>Path Length</b>	Average shortest path between any two nodes, lower values indicate quicker information spread	2.111	1.996	1.807
<b>Diameter</b>	Longest of all the shortest paths in the network, smaller values suggest closer connectivity	5	4	4
<b>Global Clustering Coefficient</b>	Degree to which nodes cluster together, higher values suggest a tighter community	0.345	0.282	0.372

The second sub-question, which focuses on identifying global metrics that can highlight the nature and quality of student interactions across the sections, evaluated several key indicators.

The network for C4, depicted in the left diagram, reveals a relatively sparse structure with fewer connections per node. This observation aligns with the quantitative data, which shows a graph density of 0.228. This indicates that interactions among students are not as frequent. The average weighted degree of 16.679 suggests moderate interaction levels, but these interactions are not well-distributed throughout the network. The network diameter of 5 and path length of 2.111 indicate that information or interactions need to traverse more nodes to reach others, implying less direct communication pathways. Additionally, the global clustering coefficient (CC) of 0.345 points to a moderate tendency for students to form clusters, suggesting some level of group collaboration but not as cohesive as might be ideal.

The middle diagram for C5 shows a network with a moderate level of connectivity. The graph density is slightly lower than C4, at 0.213, suggesting fewer connections relative to the number of possible interactions. The average weighted degree is 14.960 further supports this observation, indicating that a relatively moderate frequency of interactions among students when compared to the higher interaction levels in C4. However, despite the lower overall density and weighted degree, the network structure of C5 reveals a more efficient configuration of connections. This is evidenced by a diameter of 4 and a shorter path length of 1.996, which are indicative of a more compact network. That is, although C4 has a slightly higher graph density (0.228) compared to C5 (0.213), it also has a higher diameter of 5 compared to 4 in C5. This contrast of higher density but larger diameter can seem counterintuitive initially, as typically, a higher density would suggest a more compact network, potentially leading to a smaller diameter.

This phenomenon can be explained with a few aspects in network. First, although a higher density indicates students in C4 on average students in C4 have more connections than C5, these connections might not be optimally distributed to minimize the longest path between

nodes. If these connections in C4 are more clustered within smaller groups, the overall network might still have longer paths between distant nodes, resulting in a higher diameter. Thus, higher clustering within certain subgroups can lead to localized high density but still maintain a high diameter if these clusters are not well interconnected. In other words, students within a cluster might be very well connected but there could be a few connections between clusters, necessitating longer paths to travel between students in different clusters. Also, this could mean the lack of central hubs as seen from the figure 1. Central hubs play a crucial role in reducing the diameter of a network by providing shortcuts that reduce the number of steps needed to connect any two nodes. C5 might have more effective central hubs that reduce the overall diameter, even with a lower density, whereas C4 had a higher density but lacks these central hubs, resulting in a less efficient overall network structure. The global clustering coefficient of 0.282, the lowest among the three courses, further explains that there are fewer tightly-knit groups, and students are less likely to form cohesive collaborative clusters in C5.

C6, with the highest density of 0.267, the highest weighted degree of 17.294 and a diameter of 4, demonstrates the most robust as well as frequent interaction patterns. The high density and lower diameter reflect a well-integrated network where students are not only connected but are part of a closely-knit community. The shortest path length of 1.807 further supports this, indicating efficient communication and interaction pathways throughout the course. The high global clustering coefficient of 0.372 signifies strong group cohesiveness, suggesting that students in C6 engage in frequent and meaningful interactions, enhancing collaborative learning.

These observations highlight the nuanced nature of network dynamics in online courses, where simply increasing the number of connections (density) does not necessarily lead to a more

efficiently connected network (lower diameter). Strategic interventions to enhance inter-cluster connectivity and the presence of central hubs can significantly impact the overall structure and effectiveness of student interactions.

### Local Metrics (RQ2)

Local Network Analysis focused more deeply on the communities identified by the Louvain algorithm, a key component in our methodological approach. Using Gephi's built-in modularity function, which employs the Louvain algorithm, the study segmented the network based on patterns of connectivity among nodes. This process optimizes the modularity score to gauge the strength of divisions into distinct communities. The study kept the modularity resolution setting in Gephi at its default to ensure consistency across all datasets, based on preliminary tests that confirmed its effectiveness in accurately reflecting community structures within our data. The study examined several local metrics within the identified communities, such as the local clustering coefficient, within-community density, and the roles of central nodes. These metrics helped us to understand the cohesiveness and collaborative potential of the groups formed within the network. Specifically, the study explored how local metrics vary across groups, patterns within these metrics, how correlations between them contribute to our understanding and finally reflecting collaborative nature and cohesiveness by the algorithm's classification.

These examination of local network dynamics enriched our understanding of how students interact within smaller, more defined groups and provides insights into the structural and functional aspects of these interactions. The study also highlights the utility of the Louvain algorithm in identifying meaningful interaction patterns and naturally occurring student communities. The Louvain algorithm helps to classify groups based on the actual interaction

patterns as a communication detection method to uncover high modularity groups within large networks. It basically optimizes the modularity of a network by aggregating nodes into communities in a way that maximizes the density of links inside communities compared to links between communities. For this research question, the study analyzed the groups after the Louvain algorithm had identified groups as a post-classification analysis to understand their structural characteristics and here is how each metric can be used post-classification to interpret the nature of the groups formed in Table 4:

*Table 3. Role of local metrics in post-Louvain classification analysis*

<b>Metric</b>	<b>Purpose</b>	<b>Post-classification use</b>	<b>Interpretation of collaborative nature and cohesiveness</b>
<b>Clustering Coefficient (LCC)</b>	Indicates how close nodes in a network tend to cluster together.	Higher values suggest tightly knit members within a community.	High LCC indicates strong internal communication and cohesion.
<b>Density</b>	Measures the proportion of actual connections to possible connections within a group.	Higher density suggests a well-connected group.	High density implies effective communication and collaboration.
<b>Triangles</b>	Counts the number of closed triplets within the network.	Measures local cohesiveness and strong subgroup relationships.	Many triangles indicate robust mutual connections within subgroups.
<b>Weighted Degree</b>	Reflects the sum of the weights of the edges connected to a node.	Identifies influential or central nodes within each group.	High weighted degree indicates potential key individuals crucial for information flow.

---

DRAFT



This table provides a clear framework for understanding the role and interpretation of local metrics in analyzing the structure of groups identified by Louvain algorithm.

#### *Variation in local metrics (RQ2\_a)*

To evaluate the Louvain algorithm's classification effectiveness, the study analyzed the local metrics within the groups it identified. Table 3 presents these metrics for groups in courses C4, C5, and C6, highlighting variations in clustering coefficient, number of triangles, density, and weighted degree. These selection as local metrics was strategic for understanding group cohesion in online learning environments. To be specific, these metrics can offer a comprehensive analysis of how students interact within their groups, the strength of their connections, and the roles of central nodes in facilitating collaborative learning.

*Table 4. Local level metrics in groups*

Course	Group	LCC	Triangles	Density	W. Degree
<b>C4</b>	Group0	0.485	2.455	0.291	5
	Group1	0.429	0.857	0.381	4.25
	Group2	0.818	5.625	0.607	7.75
	Group3	0.771	7.286	0.81	15.7
	Group4	0.658	4.500	0.607	8
	Group5	0.470	4.500	0.364	5
<b>C5</b>	Group6	0.619	4.714	0.714	9.143
	Group7	0.463	3.700	0.467	5.8
	Group8	0.528	2.500	0.667	9

	Group9	0.595	3.000	0.619	7.714
	Group10	0.385	1.000	0.417	4
	Group11	0.015	0.091	0.236	3.6
<b>C6</b>	Group12	0.571	3.000	0.619	6.286
	Group13	0.562	3.429	0.619	6
	Group14	0.562	4.000	0.5	4.667
	Group15	0.381	2.143	0.524	4.286
	Group16	0.622	3.222	0.417	4.667
	Group17	0.576	3.857	0.667	6.571
	Group18	0.900	4.200	0.9	9.6

The clustering coefficient indicates the likelihood that a node's neighbors are also connected. High LCC values suggest tight-knit groups where members are likely to form triangles. For example, in the section C6, Group 18 exhibits an exceptionally high LCC of 0.900, indicating that nearly every student in this group forms a triangle with their peers, reflecting very high cohesiveness. High LCC values indicate that students are likely to engage in mutual interactions, forming strong triadic relationships. This cohesiveness is beneficial for collaborative learning as it fosters a supportive environment where students can easily share ideas and resources. On the other hands, Groups in C4 and C5 show more variability. To be specific, Group 2 in C4 has a high LCC of 0.818, while Group 11 in C5 has an extremely low LCC of 0.015, indicating a lack of cohesive clusters in that group. This variability can highlight areas where intervention is needed to foster better group cohesion and collaboration.

Density measures how many of the possible connections in a group are actually present. Higher density values suggest more interconnected and potentially more collaborative groups. Group 18 in C6 also shows the highest density at 0.9, indicating that almost all potential connections are present, suggesting an environment where students are highly interconnected. High-density groups are likely to have efficient communication channels, enhancing collaborative efforts. Density values vary significantly in other courses. For instance, Group 3 in C4 has a high density of 0.81, suggesting strong internal connectivity, while Group 11 in C5 has a much lower density of 0.236, indicating sparse connections.

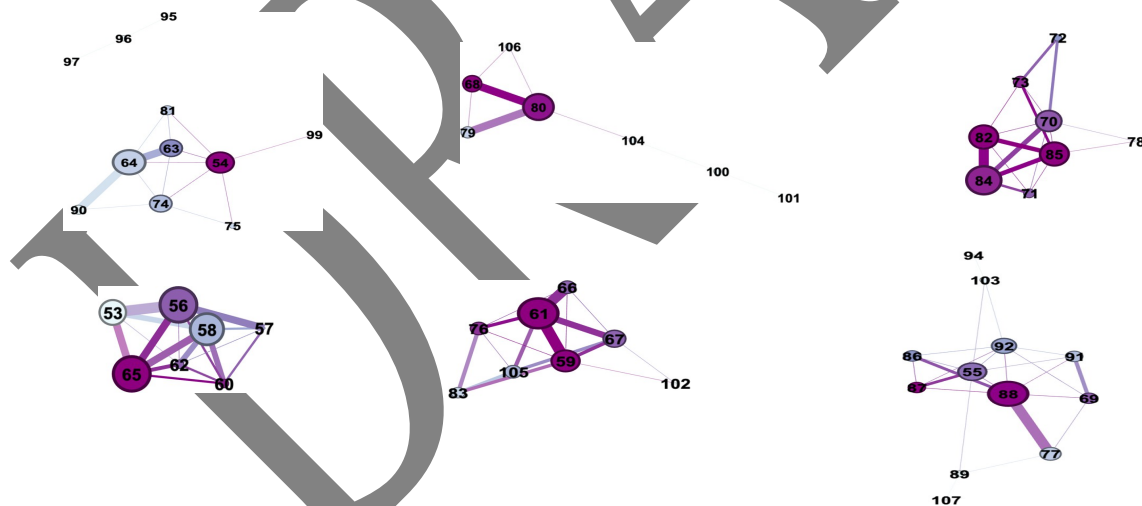
The weighted degree reflects the total interaction volume a node (student) has. Higher values suggest that some students are particularly active and central within their groups. Group 3 in C4 shows a high weighted degree of 15.7, indicating that students in this group are very active and central to the group's interactions. In C6, Group 18 also exhibits a high weighted degree of 9.6, further demonstrating the high level of interaction within this group.

The number of triangles in a network indicates local clustering and the presence of cohesive subgroups. Groups in C6, such as Group 18 with 4.200 triangles, indicate robust mutual connections and strong subgroup formation. C4's Group 3 shows a high number of triangles (7.286), suggesting strong cohesiveness within the group. Conversely, C5's Group 11 has very few triangles (0.091), indicating weak subgroup formation.

*Patterns identified by the Louvain Algorithm (RQ 2\_b)*

*Table 5. Descriptive statistics for the group cohesiveness*

Metric	LCC	Density	Triangles	W. Degree
N (Valid)	19	19	19	19
Mean	0.548	0.549	3.373	6.686
Median	0.562	0.607	3.429	6
St. Deviation	0.189	0.173	1.704	2.875
Skewness	-0.796	0.077	0.12	1.79
Kurtosis	2.853	-0.315	0.645	4.389



*Figure 2. C4 Group Networks (0-2 left to right, top row; 3-5 left to right, bottom row)*

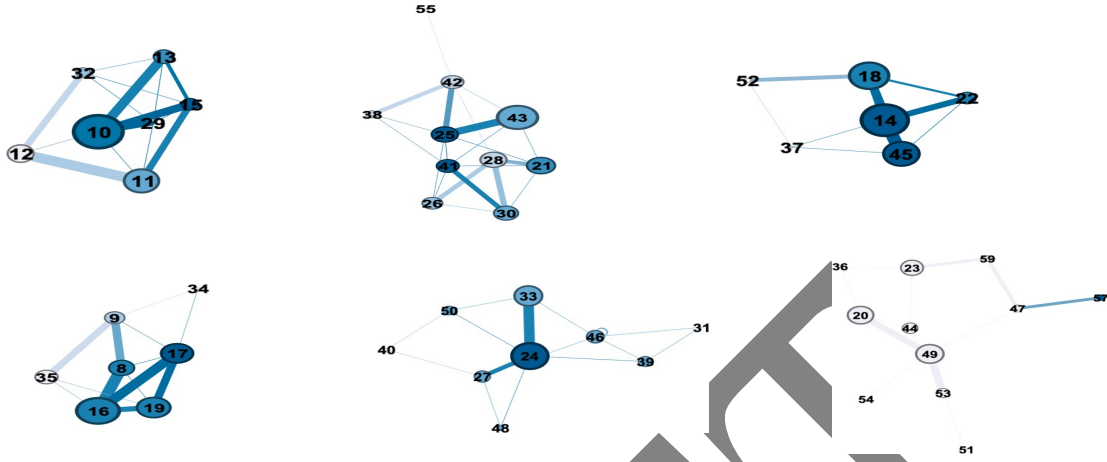


Figure 3. C5 Group Networks (6-8 left to right, top row; 9-11 left to right, bottom row)

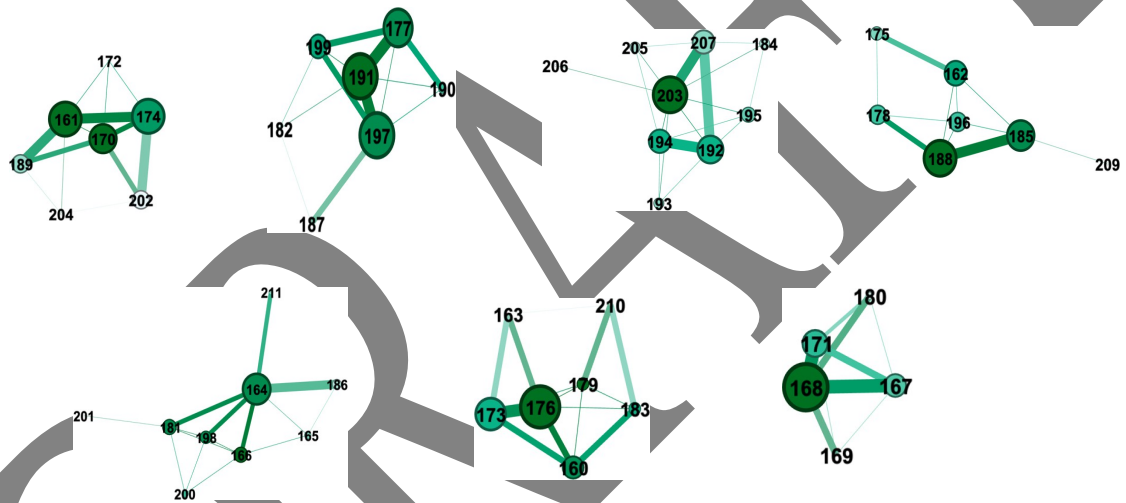


Figure 4. C6 Group Networks (12-15 left to right, top row; 16-18 left to right, bottom)

The average **Clustering Coefficient (LCC)** across all groups is 0.548, with a median of 0.562. This indicates a general tendency for students to form moderately cohesive clusters. Groups with high LCC values, such as Group 18 in C6 (0.900) (Figure 4) and Group 3 in C4 (0.771) (Figure 2), are characterized by strong local clustering. These groups are likely to have

robust internal interactions where students frequently form triadic relationships, enhancing collaborative learning, whereas groups with low LCC values, such as Group 11 in C5 (0.015) (Figure 3), show minimal clustering, indicating less cohesive and more fragmented interactions.

The mean **Density** of 0.549 and median of 0.607 suggest that, on average, groups maintain about half of their potential connections. High-density groups, like Group 18 (Figure 4) in C6 (0.9) and Group 3 in C4 (0.81) (Figure 2), exhibit strong interconnectedness among members, fostering efficient communication and collaboration whereas groups with lower density, such as Group 11 in C5 (0.236) (Figure 3), are less interconnected, which may impede effective group interactions and collaboration.

The average number of **Triangles** is 3.373, with a median of 3.429. Triangles indicate the presence of closed triplets, which are crucial for local cohesion. Groups like Group 3 in C4 (7.286) (Figure 2), and Group 6 in C5 (4.714) (Figure 3), have a high number of triangles, suggesting strong subgroup formations and robust mutual connections while Groups with fewer triangles, such as Group 11 in C5 (0.091) (Figure 4), lack these strong mutual connections, reflecting weaker subgroup cohesion.

The mean **Weighted Degree** is 6.686, with a median of 6. This indicates a moderate level of interaction volume per node. Groups with high weighted degrees, like Group 3 in C4 (15.7) (Figure 2), and Group 6 in C5 (9.143) (Figure 4), have central nodes with significant interaction volumes, suggesting that certain students play key roles in maintaining group dynamics while groups with lower weighted degrees, such as Group 11 in C5 (3.6) (Figure 3), may have less active participation, indicating potential areas for intervention to boost engagement.

To summarize, the mean values for LCC (0.548), density (0.549), and triangles (3.373) across all groups suggest moderate overall cohesiveness and connectivity. These metrics indicate

that, on average, groups are fairly cohesive, but there's variability. When considering standard deviation: Variability in LCC (0.189), density (0.173), and triangles (1.704) confirms that some groups are more cohesive and interconnected than others, as visually represented in the diagrams. Finally, negative skewness in LCC (-0.796) suggests a tail towards lower values, indicating some groups with very low cohesiveness. High kurtosis in weighted degree (4.389) points to the presence of outliers, likely groups with exceptionally high central node activity.

#### *Correlational analysis of group dynamics (RQ 2\_c)*

The study further utilized Spearman's rank-order correlation coefficients to analyze the relationships between various structural metrics within student interaction networks, as identified by the Louvain algorithm. These metrics—Local Clustering Coefficient (LCC), Triangles, Density, and Weighted Degree—offer insights into the collaborative nature and cohesiveness of the hypothetical groups formed within the online courses.

*Table 6. Spearman's Rank-Order Correlations among Group Metrics*

<b>Metric</b>	<b>Triangles</b>	<b>Density</b>	<b>W. Degree</b>
<b>Local CC</b>	.762**	.677**	.801**
<b>Triangles</b>	-	.531*	.719**
<b>Density</b>	.531*	-	.860**
<b>Weighted Degree</b>	.719**	.860**	-

Note: \*\*p < .01, \*p < .05.

A strong positive correlation (0.762) was observed between the Local Clustering Coefficient (LCC) and the number of Triangles. This correlation suggests that groups with higher

LCC, indicative of more cohesive units, also tend to have a greater number of closed triangles. Such a structure implies that these groups not only have tightly-knit connections among members but also exhibit a high level of mutual interconnectedness, facilitating robust subgroup formations within the network. Similarly, the LCC showed a significant positive correlation with Density (0.677), indicating that more cohesive groups are generally denser. This relationship underscores that as groups become more cohesive, the proportion of actual connections relative to possible connections increases, reflecting a more interconnected network where collaborative interactions are likely more frequent and substantial. The correlation between LCC and Weighted Degree (0.801) further highlights that groups characterized by higher cohesiveness also feature nodes with higher levels of interaction volume. This finding suggests that in groups where members are closely connected, there are typically one or more individuals who are central to maintaining and facilitating these interactions, thus enhancing the group's overall communicative activity.

Examining the relationships involving Triangles, a moderate positive correlation (0.531) with Density was found, supporting the notion that networks with more closed triangles tend to be denser. This correlation aligns with the idea that increased group interconnectivity, as evidenced by numerous triadic relationships, contributes to a denser overall network structure, potentially enhancing the collaborative environment. Furthermore, the strong positive correlation between Triangles and Weighted Degree (0.719) indicates that groups with numerous triangles not only have dense connections but also contain highly active nodes. This suggests that these groups not only form tight-knit communities but are also dynamic, with substantial engagement across members.



The most noticeable correlation was between Density and Weighted Degree (0.860), illustrating that denser groups typically feature nodes with higher weighted degrees. This robust relationship implies that in networks where connections are plentiful, individual nodes tend to engage more actively, contributing significantly to the network's liveliness and the efficacy of collaborative efforts.

These correlations collectively found a comprehensive picture of the group dynamics within the studied networks. The strong interdependencies between cohesiveness (as measured by LCC and Triangles) and both the quantity (Density) and quality (Weighted Degree) of interactions suggest that more cohesive groups are not only well-connected but also highly active. This detailed understanding of network structures provides valuable insights into how different aspects of group interactions contribute to the collaborative potential and educational outcomes within online learning environments.

#### *Reflection of Collaborative Nature and Cohesiveness by the Algorithm's Classification (RQ2\_d)*

The Louvain algorithm's classification provides insights into the collaborative nature and cohesiveness of the groups within the network. Although these groups are hypothetical, the analysis of their interaction patterns reveals important findings about potential group dynamics. By using the algorithm to identify these interaction patterns, we can systematically uncover and validate the structural properties that signify effective collaboration and group cohesion.

Groups with high clustering coefficients and density values reflect strong potential for collaboration and mutual support. For instance, Group 18 in C6 exhibits a high LCC of 0.900 and a density of 0.9. The visual representation of this group (Figure 4) shows dense interconnections, suggesting that, if this group were real, students would likely form a tightly-

knit network with frequent interactions. This high level of cohesion indicates a strong potential for effective collaboration and mutual support among group members. Such metrics and visual representations provide empirical evidence of cohesive group structures which are essential for collaborative learning environments. Conversely, groups with low clustering coefficients and density values indicate weaker connections and less cohesive interactions. Group 11 in C5, for example, has an LCC of 0.015 and a density of 0.236. The corresponding network diagram (Figure 3) shows sparse connections and isolated nodes. These metrics suggest that, in a real-world scenario, this group might struggle with collaboration due to fragmented interactions. The lack of cohesion highlighted by both the metrics and the visual representation points to the need for targeted interventions to enhance group connectivity and engagement. The ability to systematically identify such groups using the Louvain algorithm validates its utility in detecting areas that require pedagogical intervention to improve group dynamics.

Groups with high weighted degrees often feature central nodes that play a significant role in maintaining communication and interaction within the group. For instance, Group 3 in C4 has a high weighted degree of 15.7, indicating a substantial interaction volume per node. The network diagram for this group (Figure 2) shows several nodes with multiple connections, suggesting that key participants are driving the group's interactions. These central figures are crucial for facilitating communication and fostering a collaborative environment within the group. Identifying these key participants is vital for educators to support and utilize them in enhancing overall group effectiveness.

The presence of triangles within the network highlights the formation of cohesive subgroups. Group 3 in C4, with a high number of triangles (7.286), demonstrates robust mutual connections. The visual diagram (Figure 2) shows multiple interconnected triangles, indicating

strong subgroup formations within the group. These subgroups reflect a high level of local cohesion, which is essential for effective collaboration and support within the larger group. This detailed analysis using the Louvain algorithm helps in understanding the underlying microstructures that contribute to the overall cohesiveness of the group.

Balanced metrics across different groups provide a comprehensive view of their potential for effective collaboration. Group 6 in C5, with a moderate to high LCC of 0.619, density of 0.714, and weighted degree of 9.143, represents a healthy hypothetical group. The network diagram (Figure 3) shows a well-connected network, suggesting that students in this group would likely engage and support each other effectively. Such balanced metrics indicate strong collaborative potential and robust group dynamics. The algorithm's ability to detect these patterns validates its use in educational settings for fostering optimal group formations.

The correlation analysis further reinforces the understanding of these group dynamics. The strong positive correlation between LCC and Triangles (0.762) indicates that groups with higher clustering coefficients tend to have more closed triangles, suggesting robust mutual connections and cohesive subgroup formations. The correlation between LCC and Density (0.677) shows that more cohesive groups are also denser, reflecting a higher proportion of actual connections relative to possible connections. The very strong correlation between LCC and Weighted Degree (0.801) suggests that highly cohesive groups also have members with high interaction volumes, highlighting the importance of active participants in maintaining group connectivity.

Additionally, the moderate correlation between Triangles and Density (0.531) and the strong correlation between Triangles and Weighted Degree (0.719) indicate that groups with more interconnected triads tend to be denser and have members with higher levels of

engagement. The very strong correlation between Density and Weighted Degree (0.860) underscores that denser groups feature more active nodes, contributing significantly to the network's vibrancy and collaborative potential.

The classification of student groups by the Louvain algorithm provides a systematic reflection of the collaborative nature and cohesiveness of these hypothetical groups. High values in these metrics correspond to more cohesive and potentially collaborative groups, as evidenced by dense, interconnected network diagrams. Conversely, low values indicate fragmented groups with weak interactions, highlighting areas needing targeted interventions. This analysis enhances our understanding of group dynamics, offering insights that can inform the design of more effective collaborative learning environments.

### Implication

The application of the Louvain algorithm within online learning environments offers valuable insights and practical applications in two primary areas. First, as a group formation strategy, the algorithm leverages naturally occurring interaction patterns among students to organize them into groups. This data-driven approach can lead to the creation of learning clusters that are inherently more cohesive due to shared communication dynamics. Such groupings are inclined to active participation and collaboration, which are critical components of successful online learning. By identifying key participants and forming groups with high clustering coefficients and density values, educators can foster an environment where effective collaboration and mutual support are more likely to thrive.

Secondly, the algorithm acts as a powerful analytic tool, dissecting the web of student interactions to provide educators with a clearer understanding of how collaborative dynamics

unfold in digital spaces. The strong correlations identified between metrics such as LCC, triangles, density, and weighted degree offer empirical evidence of the structural properties that signify effective collaboration and group cohesion. This knowledge aids in the refinement of group formation and enhances the overall instructional design, potentially leading to educational spaces that are more conducive to student engagement and collective problem-solving. The study emphasizes the potential of the Louvain algorithm, traditionally employed across various domains to identify communities with robust connections. This method is suitable for online platforms and seamlessly integrates into face-to-face and hybrid settings.

However, this study's application of the Louvain algorithm also reveals some limitations. The efficacy of the algorithm is predicated on the presence of rich interaction data, which may not be available in all learning contexts. Additionally, the success of this strategy in an undergraduate communications course may not directly translate to different subjects or educational settings where interaction patterns could vary significantly. Thus, further research is needed to explore the versatility and effectiveness of the Louvain algorithm in diverse online learning scenarios. Another potential limitation is the assumption that algorithmically generated groups will inherently function effectively. The dynamic nature of group interactions and the importance of individual student contributions mean that the Louvain algorithm should be used as one of several tools in the instructional design process. It is important to recognize that the quality and depth of interactions can also be influenced by course content, instructional strategies, and individual student motivations. Also, the study's predominant reliance on quantitative network metrics and the Louvain algorithm might not fully encapsulate the nuanced qualitative aspects of student interactions.

## Conclusion

This study serves as an initial exploration into the use of the Louvain algorithm in online learning settings. Our findings reveal the algorithm's dual potential as a tool for group formation and for analyzing interaction patterns to improve collaborative learning environments. The results offer compelling evidence that supports the utility of the Louvain algorithm in both creating cohesive learning groups and providing insights into the nature of student interaction.

Nevertheless, it is important to note the limitations in this exploratory research. The scope of our study is limited to specific courses, and the generalizability of the findings may be restricted. The impact of the algorithm-informed group structures on actual learning outcomes remains to be empirically validated in future studies.

Therefore, future research should aim to replicate these findings across different contexts and explore the longitudinal effects of algorithm-assisted group formation on student engagement and achievement. Additionally, there is a need to compare the Louvain algorithm with other group formation strategies to better understand its relative strengths and weaknesses. Supplementing our quantitative analysis with qualitative research methods would also offer a more holistic view of the student experience. This study lays the groundwork for subsequent research that could refine our understanding of how technology can be employed to improve the quality of interaction in online learning communities. The insights gained here are a step towards creating online educational spaces where collaboration and interaction are actively enhanced through data-driven strategies.

## References

Authors\_1. (2021)

Author\_1 & Author\_2. (2023)

Arbaugh, J. B. (2008). Does the Community of Inquiry Framework Predict Outcomes in Online MBA Courses? *International Review of Research in Open and Distributed Learning*, 9(2).

Bates, A. W. (2015). *Teaching in a digital age: Guidelines for designing teaching and learning*. BCcampus.

Beal, D. J., Cohen, R. R., Burke, M. J., & McLendon, C. L. (2003). Cohesion and performance in groups: A meta-analytic clarification of construct relations. *Journal of Applied Psychology*, 88, 989–1004. <https://doi.org/10.1037/0021-9010.88.6.989>

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. *Journal of Applied Psychology*, 92(3), 595–615.  
<https://doi.org/10.1037/0021-9010.92.3.595>

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10), P10008.

Braun, M. T., Kozlowski, S. W. J., Brown, T. A. R., & DeShon, R. P. (2020). Exploring the dynamic team cohesion-performance and coordination-performance relationships of newly formed teams. *Small Group Research*, 1-30. <https://DOI:10.1177/1046496420907157>

Carolan, B. V. (2013). *Social network analysis and education: Theory, methods & applications*. Sage Publications.

- Cela, K. L., Sicilia, M. Á., & Sánchez, S. (2015). Social Network Analysis in E-Learning Environments: A Preliminary Systematic Review. *Educational Psychology Review*, 27(1), 219–246. <https://doi.org/10.1007/s10648-014-9276-0>
- Cohen, E. G. (2023). *Restructuring the Classroom: Conditions for Productive Small Groups*.
- Dawson, S. (2008). A study of the relationship between student social networks and sense of community. *Journal of educational technology & society*, 11(3), 224-238.
- De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology*, 88(4), 741–749. <https://doi.org/10.1037/0021-9010.88.4.741>
- Dillenbourg P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed), *Collaborative-learning: Cognitive and computational approaches* (pp. 1-19). Oxford: Elsevier.
- Felder, R. M., & Brent, R. (2001). Effective strategies for cooperative learning. *Journal of Cooperation & Collaboration in College Teaching*, 10(2), 69-75.
- Froehlich, D. E., Van Waes, S., & Schäfer, H. (2020). Linking quantitative and qualitative network approaches: A review of mixed methods social network analysis in education research. *Review of Research in Education*, 44(1), 244-268.
- Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education*, 2(2-3), 87-105.
- Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12), 7821-7826.
- Grunspan, D. Z., Wiggins, B. L., & Goodreau, S. M. (2014). Understanding classrooms through



- social network analysis: A primer for social network analysis in education research. *CBE—Life Sciences Education*, 13(2), 167-178.
- Haythornthwaite, C. (2001). Social Network Analysis: An Approach and Technique for the Study of Information Exchange. *Library & Information Science Research*, 18(4), 323-342.
- Hilliger, I., Ortiz-Rojas, M., Pesántez-Cabrera, P., Scheihing, E., Tsai, Y. S., Muñoz-Merino, P. J., ... & Pérez-Sanagustín, M. (2020). Identifying needs for learning analytics adoption in Latin American universities: A mixed-methods approach. *The Internet and Higher Education*, 45, 100726.
- Jalan, S., Sarkar, C., Madhusudanan, A., & Dwivedi, S. K. (2014). Uncovering randomness and success in society. *PloS one*, 9(2), e88249.
- Jan, S.K., Vlachopoulos, P. (2019). Social Network Analysis: A Framework for Identifying Communities in Higher Education Online Learning. *Tech Know Learn* **24**, 621–639. <https://doi.org/10.1007/s10758-018-9375-y>
- Jensen, J. L., & Lawson, A. (2011). Effects of collaborative group composition and inquiry instruction on reasoning gains and achievement in undergraduate biology. *CBE—Life Sciences Education*, 10(1), 64-73.
- LePine, J. A., Piccolo, R. F., Jackson, C. L., Mathieu, J. E., & Saul, J. R. (2008). A meta-analysis of teamwork processes: Tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel Psychology*, 61, 273–307. <https://doi.org/10.1111/j.1744-6570.2008.00114.x>
- Luisa Sanz-Martínez, Erkan Er, Alejandra Martínez-Monés, Yannis Dimitriadis & Miguel L. Bote-Lorenzo (2019): Creating collaborative groups in a MOOC: a homogeneous

- engagement grouping approach, Behaviour & Information Technology, DOI: 10.1080/0144929X.2019.1571109
- Kourtellis, N., Alahakoon, T., Simha, R., Iamniti, A., & Tripathi, R. (2013). Identifying high betweenness centrality nodes in large social networks. *Social Network Analysis and Mining*, 3(4), 899-914.
- Kozlowski, S. W., & Bell, B. S. (2003). Work groups and teams in organizations.
- Krouska, A., & Virvou, M. (2019). An enhanced genetic algorithm for heterogeneous group formation based on multi-characteristics in social-networking-based learning. *IEEE Transactions on Learning Technologies*, 13(3), 465-476.
- Magpili, N. C., & Pazos, P. (2018). Self-managing team performance: A systematic review of multilevel input factors. *Small Group Research*, 49(1), 3–33. <https://doi.org/10.1177/1046496417710500>
- Marsden, Peter V. (2004). "[Network Analysis](#)." Encyclopedia of Social Measurement, edited by Kimberly Kempf-Leonard, 819-825. San Diego: Academic Press.
- Maxwell, J. A. (2012). *A realist approach for qualitative research*. Sage.
- Menczer, F., Fortunato, S., & Davis, C. (2020). *A First Course in Network Science*. Cambridge: Cambridge University Press. doi:10.1017/9781108653947
- Miele, V., Matias, C., Robin, S., & Dray, S. (2019). Nine quick tips for analyzing network data. *PLoS computational biology*, 15(12), e1007434.
- Miller, H. B., Witherow, D. S., & Carson, S. (2012). Student learning outcomes and attitudes when Biotechnology lab partners are of different academic levels. *CBE-Life Sciences Education*, 11(3), 323-332.

- Moore, M. G., & Kearsley, G. (2011). *Distance Education: A Systems View of Online Learning*. Cengage Learning.
- Morrison, D., Bedinger, M., Beevers, L. *et al.* Exploring the *raison d'être* behind metric selection in network analysis: a systematic review. *Appl Netw Sci* **7**, 50 (2022).
- Müller, A., Bellhäuser, H., Konert, J., & Röpke, R. (2022). Effects of group formation on student satisfaction and performance: A field experiment. *Small Group Research*, *53*(2), 244-273.
- Newman, M. (2018). *Networks*. Oxford university press.
- Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, *69*(2), 026113.
- Opsahl, T., 2013. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Social Networks* *35*, doi: 10.1016/j.socnet.2011.07.001.
- Ouyang, F., & Scharber, C. (2017). The influences of an experienced instructor's discussion design and facilitation on an online learning community development: A social network analysis study. *The Internet and Higher Education*, *35*, 3  
<https://doi.org/10.1016/j.iheduc.2017.07.002>
- Pai, H. H., Sears, D. A., & Maeda, Y. (2015). Effects of small-group learning on transfer: A meta-analysis. *Educational psychology review*, *27*, 79-102.
- Peacock, S. & Cowan, J. (2019). Promoting sense of belonging in online learning communities of inquiry at accredited courses. *Online Learning*, *23*(2), 67-81.  
doi:10.24059/olj.v23i2.1488
- Rovai, A. P. (2002). Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. *The Internet and Higher Education*, *5*(4), 319-332.

- Rovai, A. P. (2007). Facilitating online discussions effectively. *The Internet and Higher Education*, 10(1), 77–88. <https://doi.org/10.1016/j.iheduc.2006.10.001>
- Rodríguez, D., Sicilia, M. Á., Sánchez-Alonso, S., Lezcano, L., & García-Barriocanal, E. (2011). Exploring affiliation network models as a collaborative filtering mechanism in e-learning. *Interactive Learning Environments*, 19(4), 317-331.
- Samarasinghe S, Strickert G (2013) Mixed-method integration and advances in fuzzy cognitive maps for computational policy simulations for natural hazard mitigation. *Environ Model Softw* 39:188–200. <https://doi.org/10.1016/j.envsoft.2012.06.008>
- Saqr M., Nouri J., Vartiainen H., Tedre M. (2020). Robustness and rich clubs in collaborative learning groups: a learning analytics study using network science. *Scientific Reports*, vol. 10(1), art. no. 14445. doi: 10.1038/s41598-020-71483-z.
- Singh, J., Singh, L., & Matthees, B. (2022). Establishing social, cognitive, and teaching presence in online learning—A panacea in COVID-19 pandemic, post vaccine and post pandemic times. *Journal of Educational Technology Systems*, 51(1), 28-45.
- Shea, P., Li, C. S., & Pickett, A. (2006). A study of teaching presence and student sense of learning community in fully online and web-enhanced college courses. *The Internet and higher education*, 9(3), 175-190.
- Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, 69(1), 21–51. <https://doi.org/10.3102/00346543069001021>.

- Tekleab, A. G., Quigley, N. R., & Tesluk, P. E. (2009). A longitudinal study of team conflict, conflict management, cohesion, and team effectiveness. *Group & Organization Management*, 34, 170–205. <https://doi.org/10.1177/1059601108331218>
- Thomas, L., Herbert, J., & Teras, M. (2014). A sense of belonging to enhance, participation, success and retention in online programs. *The International Journal of the First Year in Higher Education*, 5(2), 69–80.
- Vallès-Català, T., & Palau, R. (2023). Minimum entropy collaborative groupings: A tool for an automatic heterogeneous learning group formation. *Plos one*, 18(3), e0280604.
- Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre (2008) Fast unfolding of communities in large networks, in *Journal of Statistical Mechanics: Theory and Experiment*.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*.
- Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer supported collaborative learning. *Instructional Science*, 33(1), 1–30. <https://doi.org/10.1007/s11251-004-2322-4>
- Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems thinker*, 9(5), 2-3.
- Wise, A. F., Saghaian, M., & Padmanabhan, P. (2012). Towards more precise design guidance: Specifying and testing the functions of assigned student roles in online discussions. *Educational Technology Research and Development*, 60, 55-82.
- Wise, A. F., Hausknecht, S. N., & Zhao, Y. (2014). Attending to others' posts in asynchronous discussions: Learners' online "listening" and its relationship to speaking. *International*

*Journal of Computer-Supported Collaborative Learning*, 9(2), 185–209.

<https://doi.org/10.1007/s11412-014-9192-9>

Yang, D. Liu, J. Liu (2010). Discovering communities from social networks: methodologies and applications. *Handbook of Social Network Technologies and Applications*, Springer, pp. 331-346

Yassine, S., Kadry, S., & Sicilia, M. A. (2021). Application of community detection algorithms on learning networks. The case of Khan Academy repository. *Computer Applications in Engineering Education*, 29(2), 411-424.

Yassine, S., Kadry, S., & Sicilia, M. A. (2022). Detecting communities using social network analysis in online learning environments: Systematic literature review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 12(1), e1431.

Xie, K., Yu, C., & Bradshaw, A. C. (2014). Impacts of role assignment and participation in asynchronous discussions in college-level online classes. *The Internet and Higher Education*, 20, 10-19.