Friendship networks and educational networks and their connection

Hanie Hatami[1]
[1]*Department of Physics, Sharif University of Technology, Tehran, Iran*

INTRODUCTION

background

In the ever-evolving campus environment, the interconnectedness of students forms a complex network that is as fascinating as it is influential. This article explores two networks of communication between students and the network of common courses taken by students in the current semester, in the faculty environment of the Faculty of Physics at Sharif University of Technology with the perspective of network science.
Network science is a multidisciplinary field that has seen significant growth in recent years due to advances in computational methods and increased access to large-scale data. Key literature in this field has examined the structure and dynamics of various networks, from social networks to biological networks. Related concepts include centrality measures, community detection algorithms, and network motifs that can be used to understand complex communications in academic settings.

The importance of the subject

The communication network can reveal patterns of information flow, cooperation and social dynamics among students. On the other hand, the co-curricular network can provide insights into academic interests, course popularity, and the interdisciplinary nature of learning.
The importance of analyzing these two networks together lies not only in the academic implications but also in their practical implications for professors, managers and policymakers. By understanding the interplay between communication networks and common courses, universities can adjust their approaches to foster collaboration, foster interdisciplinary communication, and optimize resource allocation.
At the heart of this discovery lies the recognition that students do not travel their academic paths in isolation. They are connected by overt and covert communication channels and shape their educational experiences in ways that are often overlooked. Using the power of network science, we aim to uncover the nodes and edges that define these communication networks and provide insights into the underlying structures that foster collaborative learning and knowledge diffusion.
At the same time, we turn our attention to the network of co-curricular activities and recognize that academic activities are not individual endeavours but joint ventures. Courses act as a bridge between students and facilitate the exchange of ideas, knowledge and skills. Through a close examination of the courses chosen by students in the current semester, we seek to discern patterns that emerge from students' collective academic pathways.

Objectives of the report

By calculating and analyzing these networks, we aim to discover their key features and important nodes. This could potentially provide strategies to increase student engagement, optimize course delivery, and create a conducive learning environment.
In addition, examining the relationship between these two networks can yield interesting insights. For example, how do students communicate with each other in an academic setting? What are the central nodes in the communication network between students? Do students who communicate a lot tend to take similar courses? Or vice versa, do joint courses lead to increased communication?

LITERATURE REVIEW OF THE REPORT

Network science in education

In the field of education, there is a rich body of literature that emphasizes the importance of student interactions, cooperative learning, and the role of courses in shaping educational experiences. The study of networks has been a topic of interest in various disciplines, from sociology and computer science to biology and physics. Network science has given rise to studies that apply network analysis to educational contexts. This includes research on social network analysis in classrooms, collaborative networks between researchers, and communication patterns in online learning platforms.

Previous studies in this field

Several studies have investigated the complex network of communication among students and shed light on the dynamics of communication, collaboration, and joint academic activities, particularly on the use of social networks and their impact on student's academic performance and social well-being. This section reviews key studies that have paved

the way for understanding student networks.

- "Analysis of Social Networks in Higher Education: a systematic literature review" written by Smith, A. and Jones, B.: It examines various topics including friendship networks and cooperative learning.
- "Communication Patterns in Academic Settings: A Social Network Perspective" Wang, C. and Johnson, R.: Using social network analysis, it examines the communication pattern among students in academic environments and identifies key nodes and clusters in the communication network.
- "Exploring academic collaboration: a network analysis approach" Garcia, M. and Patel, S.: By examining joint courses and joint projects, it identifies patterns that contribute to the formation of academic networks.
- "The role of social media in shaping student networks: a case study of using Facebook" by Chen, L. and Kim, J.: It examines how online interactions affect offline communication and academic collaboration.

Collectively, these previous studies contribute to the knowledge base surrounding student networks and pave the way for the current research effort.

There is a gap in studies

However, there is a gap in studies when it comes to simultaneously analyzing the communication network and the co-curricular network between students and understanding the interaction between these two networks. The purpose of this article is to fill this gap.
While studies have separately analyzed social networks and academic networks among students, few have examined the interaction between these two networks.
For example, how does the communication network between students affect their choice of major and vice versa? Or registration of joint courses increase communication? Additionally, most studies have relied on digital data from online platforms, which may not accurately represent all student interactions or shared academic interests. There is a need for more robust data collection methods that can provide a comprehensive view of academic and communication networks.
By addressing these gaps, we can gain a comprehensive understanding of the student experience that encompasses both their social interactions and academic activities.

Sharing communication and academic networks for future studies

The intersection of the communication network and the common course network provides a unique perspective on the interrelationship of social and academic life among students. This intersection can be analyzed in several ways:

- Correlation analysis: We can check the correlation between the student's position in the communication network and their position in the common course network. For example, do students who are central in the communication network tend to be central in the co-curricular network?
- Community detection: We can identify communities that are common to both networks. These can represent a group of students who not only frequently interact with each other but also share similar academic interests.
- The effect of common subjects on communication: We can investigate whether common subjects affect communication patterns or not. For example, are students enrolled in similar courses more likely to communicate?
- The influence of communication on course choice: Conversely, we can examine whether communication patterns influence course choice or not. For example, do students who are friends tend to enrol in joint courses?

By analyzing the commonalities of these two networks, we can gain a comprehensive understanding of the student experience that includes both their social interactions and academic activities. It can provide strategies for fostering a vibrant and connected learning community. However, it is important to consider the dynamic nature of these networks as they evolve as student interactions and course enrollments change. Therefore, findings should be interpreted with this dynamic context in mind.

RESEARCH METHOD

data collection

The data of this research was collected using a questionnaire that was distributed among students. This questionnaire is designed to collect information on two key aspects:

- Communication Patterns: Students were asked to write down how many students they have made friends with.
- Course registration: Using the data in the Sharif courseware, we obtained the registration of the students for the courses of this semester.

Advantages of this method:

- Direct information: This method provides direct information from students, which can be more accurate and relevant than indirect measures.

- Flexibility: questionnaires can be easily adapted to collect a wide range of information.
- Scalability: This method is scalable and can be used to collect data from a large number of students.
- Structuredness: Questionnaires provide structured and standardized data, making quantification and analysis easier. This facilitates the application of quantitative methods, including network analysis, ensuring consistency in data interpretation.
- Effective for cross-sectional studies: Questionnaires are suitable for cross-sectional studies, such as examining communication patterns and common subjects in a particular semester.
- Cost-effectiveness: Compared to other methods, such as interviews or observations, questionnaires are generally more cost-effective. They require less resources to manage and collect data.

Disadvantages of this method:

- Reliability: The reliability of the data depends on the honesty and correctness of the student's answers.
- Response bias: Students may not remember all of their interactions or may choose not to disclose certain information. There is a risk of response bias. This can affect the validity of the results, especially if certain groups of students responded too little or too much.
- Low response rate: There may be a low response rate which can lead to incomplete data.
- Limited depth of information: Questionnaires may provide quantitative data, but may lack the depth that qualitative methods provide.
- Inability to move in time: Questionnaires provide a picture of students' behaviour at a specific point in time. They may not understand the dynamics of time, such as fluctuations in communication patterns or changes in course enrollment during the semester.

roving the accuracy of data collected for network analysis is critical to validating findings. Here are some of the strategies we used for this improvement:

- Clear questions: We tried to express the questions and their meaning as clearly as possible, and in some questions, we provided more explanations for better understanding. This reduces the possibility of misinterpretation and wrong answers.
- Anonymity and Confidentiality: We assured the participants that their responses would be anonymous and confidential. This can encourage them to provide honest and accurate information.
- Multiple data points: instead of relying only on the questionnaire, we tried to obtain the course registration data from the university records in the university courseware.
- Pilot questionnaire: We conducted a pilot questionnaire to identify and correct possible issues that could affect the accuracy of the data.
- Incentives for participation: We tried to increase the response rate by trying and create friendly motivation among students.

questionnaire design

The questionnaire should be carefully designed to ensure clarity and ease of understanding. It is also very important to ensure the privacy and confidentiality of students' data. We also tried to ask questions that advance the research topic and cause a better analysis. To avoid abnormal results, we tried to ensure that the questionnaire reached a diverse and representative sample of the student community and not only in the hands of a specific group. , did not take place completely).
At the beginning of the questionnaire, we mentioned some points so that the students were aware of the purpose of the study, how their data would be used, and the necessary measures to protect their confidentiality.

participant participation

As we stated in the previous section, the effort was to increase the number of participants by providing more information and getting more diverse communications in addition to diversifying the participating groups. Of course, this could not be done very well due to reasons such as time and access limitations. But very good data was collected from different input groups as many as 51 input data.

MODEL

final data of friendship

After the deadline for filling out the questionnaire, we entered the number of 51 input data into Excel and completed the Excel in the format of Figure 1. Due to the problems in some data, such as spelling mistakes, naming only the first name or only the last name, moving the first and last name, etc., we cleaned the data manually and reduced such errors to zero.

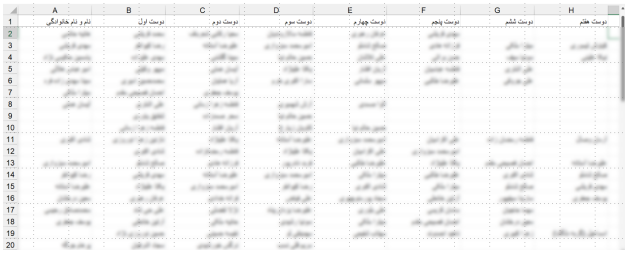


FIG. 1. Output Excel format for friendship data (information has been blurred due to confidentiality)

FIG. 2. Schematic of the friendship network for the first model (only considering common friendships)

FIG. 3. Schematic of the friendship network for the second model (considering friendships as directional relationships)

Building a friendship network

The first step is to build networks based on the data collected from the questionnaire. Each student named in Excel is represented as a node in the friendship network. Nodes are connected by edges, which reflect friendship relationships between students.
In this section, we have used two undirected and directed network models, and we have calculated all the characteristics of the friendship network for the two models.

- The first model (non-directional): In this model, the edge was placed between two students if both students jointly introduced each other as friends.
- The second model (directed): In this model, Yali moves from the first student who introduced the second student as his friend to the second student. That is, the first student may consider the second student as his friend, but the second student does not consider the first student as his friend. As a result, there are input and output edges in this model.

Analysis and celebration of the friendship network

Several network metrics can be calculated to identify key characteristics in a communication network:
1. Network drawing: Here, we first draw the network for both models to have an overview of the network schematic.
As we can see, in this network model, there are 5 correlation components. There are many leaves in this network, which means that many students have few friends in this small community. There are also a few triangles in this network.
In the second model, the correlation components are less, but they exist, because of the network drawing model, they are not well seen. In this network, some heads who have filled the form themselves and other students who have mentioned their names as friends, have input and output edges, but those who have not filled the form themselves and have only been mentioned by other people. They have entrance ridges.
2. The number of vertices and the number of edges: In this network, we have 175 vertices, meaning 175 student names, and the number of connections, or edges, is 239.
3. Degree distribution:
Degree distribution is a concept that is usually used in the study of network theory. In the field of networks, the degree of a node is the number of connections it has with other nodes, and the degree distribution is a statistical description of these node degrees in the entire network.
The degree distribution is a key measure that helps describe the overall structure of a network. It provides insights into how the vertices are connected and whether the network exhibits certain patterns such as random, regular, or scale-free.
Also, the degree distribution provides important information about the structure, strength and possible vulnerabilities of networks. 4. Distance or diameter:
The diameter of a graph is the length of the shortest path between nodes that are the most distant. In this network, the diameter of the graph is infinite and cannot be calculated numerically.
5. Degree centrality:
It measures the number of direct connections of a node. In the context of a friendship network, it can show how many peers a student communicates with. Students with a high degree of centrality are often well-connected and influential within the network. The degree centrality in this network for the first model is 0.01569786535303774 on aver-

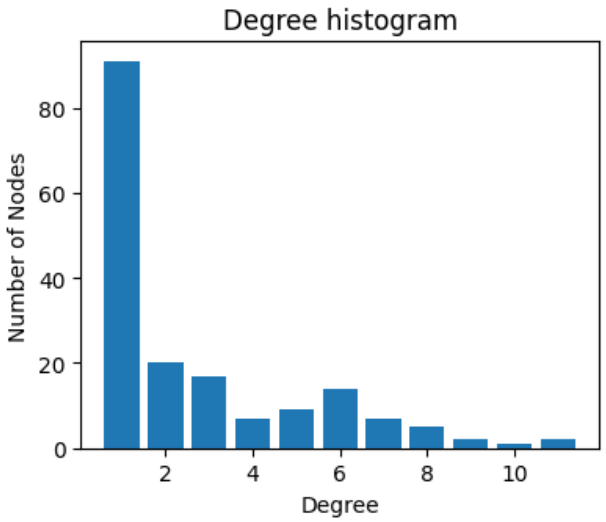


FIG. 4. The histogram of the degree distribution of the first model

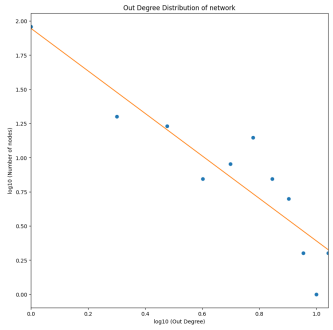


FIG. 5. Distribution of the degree of the first model in log-log form

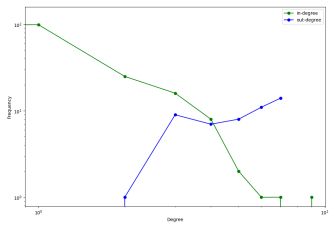


FIG. 6. Second model degree distribution as log-log for input and output edges

age, which is a very small number and shows that there are not very good connections. Also for the second model for equal inputs 0.008571428571428598 and equal to outputs 0.00857142857142857 which, as we can see, is much less than the first model.

6. Centrality between nodes(betweenness centrality):

It measures the placement of a node in the paths between other nodes. Students with high between-node centrality can connect different student groups as important bridges in the network. This value for this network is 0.020248108810994994 on average.

7. Seeking or avoiding co-identification:

Peer-to-peer is the preference for nodes in a network to connect to other nodes that are similar in some way. Although the specific measure of similarity may vary, network theorists often consider ranking based on the degree of a node. This value in this network for the first model equals 0.17029271244080021 For the second model, it is equal to 0.2188344043206471, which means that this network is convergent. This happened because not all people filled the form and most people were friends of others.

8. Modularity:

Bricking is a measure of network structure that measures the strength of dividing a network into modules (also called groups, clusters, or communities). High modularity networks have dense connections between nodes within modules but sparse connections between nodes in different modules. In this network, this value is equal to 0.5475481871815969, which indicates that there is almost a good association in this network.

9. Cluster coefficient:

In graph theory, the clustering coefficient is a measure of the degree to which nodes in a graph tend to cluster together. Evidence shows that in most real-world networks, and especially social networks, nodes tend to form tight clusters characterized by a relatively high density of ties. For this network, the average clustering coefficient is equal to 0.126.

10. Average neighbor degree:

This value represents the average degree of neighbours of the vertices, which is equal to 0.48918367346938785 for this network because many vertices who were friends of people did not fill out the questionnaire.

Other things like node connectivity, k components, max clique, maximum independent set, k core, density, adjacency matrix and normalized laplacian matrix

It has been calculated for this network, which is not included in the description of this article, and you can see them in the code.[1]

final course data

Using the university's courseware, where the registration data for the current semester's courses is known by doing WebScraping. We started to extract the names of the students who enrolled in the courses and gave the output of the code to Excel, which was finally shown below.

Creating a learning network

Every student whose name appears in one of the courses is one of the vertices of the network, and there is a link between any two students who are jointly enrolled in a course.

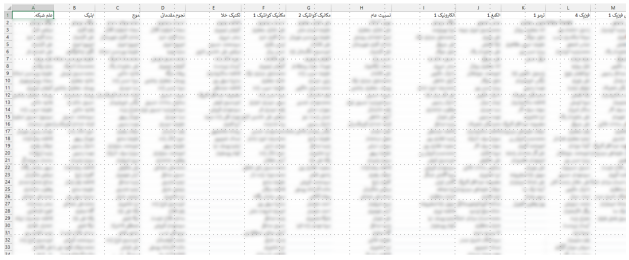


FIG. 7. Output Excel format for lesson data (information has been blurred due to confidentiality)

Analysis and evaluation of curriculum network

Several network metrics can be calculated to identify key features in the learning network:
1. Network drawing:
Here, we first draw the network to have an overview of the network schematic.

As we can see, in this network model, there is a clear

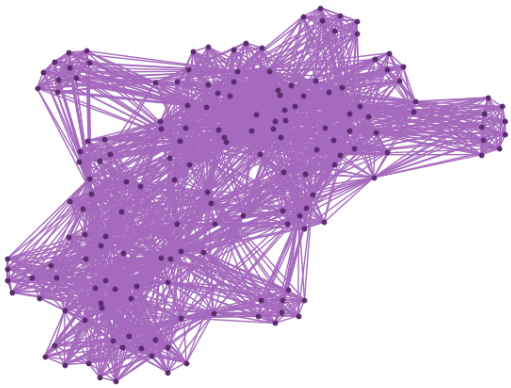


FIG. 8. Curriculum schematic

category that shows common subjects.
2. The number of vertices and the number of edges: In this network, we have 168 vertices, meaning 168 student names, and also the number of common courses, or edges, is 2247.
3. Degree distribution:
4. Distance or diameter:
The diameter of a graph is the length of the shortest path between nodes that are the most distant. In this network, the diameter is 3. It means that the maximum distance of Afra according to common courses is 3 ridges.

5. Degree centrality:
The degree centrality in this network is equal to 0.1601796407185628 on average.

6. Centrality between nodes(betweenness centrality):
This value for this network is 0.006426046358230176 on average.

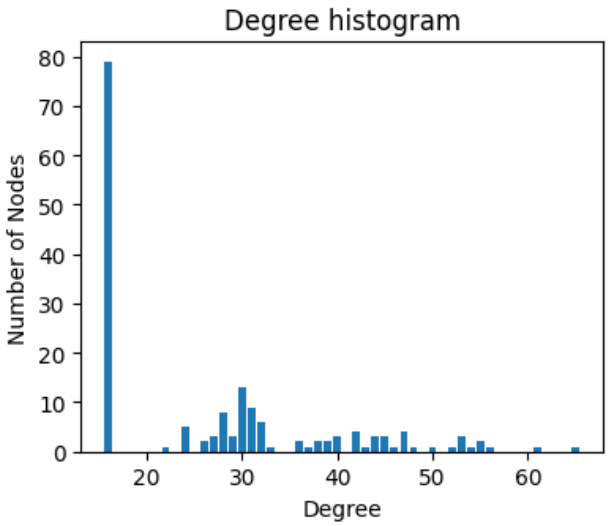


FIG. 9. The histogram of the degree distribution of the curriculum

7. Seeking or avoiding co-identification:
This value in this network is equal to -0.045519523410095096, which means that this network is asynchronous.

8. Modularity:
In this network, this value is equal to 0.3681180802014811, which indicates that there is not such good association in this network.

9. Cluster coefficient:
For this network, the average clustering coefficient is equal to 0.779, which is a very good coefficient and shows that relatively complete clustering has occurred in the network. This could be due to common subjects and common inputs.
Other things like node connectivity, k components, max clique, maximum independent set, k core, density, adjacency matrix and normalized laplacian matrix; have been calculated for this network, which is not included in the description of this article, and you can see them in the code.[2]

RESULTS AND FINDINGS

Summary of key findings

This research shows us about the friendship network that students are not isolated beings, but exist in a rich set of social connections. Through the lens of network science, we identified central nodes, influential communicators, and close-knit communities that contribute to the vibrant social fabric of the academic community.
Understanding the importance of these connections opens up ways to foster a collaborative learning environment, where students can benefit from the exchange of ideas and support.

In parallel, our analysis of the co-curricular network sheds light on the academic trajectories woven by students. Courses act as bridges, connecting people with diverse interests and creating clusters of shared academic activities. Our findings highlight key courses that play pivotal roles in fostering interdisciplinary collaborations, guiding curriculum development, and influencing students' academic journeys.
Our data was too little to establish interaction, but this little data showed us how our two networks are structured, what their degree distribution is, and the network type with its parameters.

Using the results in the real world

Our findings reveal complex patterns of student interactions and shared academic interests, and provide valuable insights that can inform strategies to increase student engagement, optimize course offerings, and foster a vibrant learning community. Furthermore, exploring the relationship between communication networks and co-curricular activities opens up new ways of understanding the complex dynamics of academic life.
However, it is important to note that implementing these insights involves challenges related to data privacy, accuracy, and the dynamic nature of networks. Therefore, while network-based strategies have significant potential, they should be used with careful consideration and refinement.
Considering the limitations of our study, such as the reliance on small data and the static nature of one semester, there is an urgent need for further research to collect more data over a longer time. Future studies could examine longitudinal trends, examine the influence of external factors on network dynamics, and examine the role of emerging technologies in shaping student interactions.

MORE TIPS

Future Research

Based on the findings of this paper, there are several potential directions for future research:
1. Time analysis: This study has presented a picture of the networks at a specific point in time. Future studies could conduct a temporal analysis to understand how these networks evolve. For example, how does the friendship network change as students progress through their studies?

2. Comparative analysis among institutions: Extending the study to different educational institutions to compare student networks. Future studies could examine how network dynamics differ in different settings, including differences in organizational size, structure, and culture.

3. Analysis of professors and students' network: Examining the networks formed between faculty members and students. How do these networks contribute to academic support, mentoring and collaborative research opportunities? **4. Integration with other data:** Networks can be enriched by integrating additional data, such as academic performance, extracurricular activities, or demographic information. This can provide a more comprehensive understanding of the student experience.
5. Forecast: Network metrics can be used to predict various outcomes of interest. For example, can a student's academic performance be predicted based on his/her position in the friendship network or shared courses?

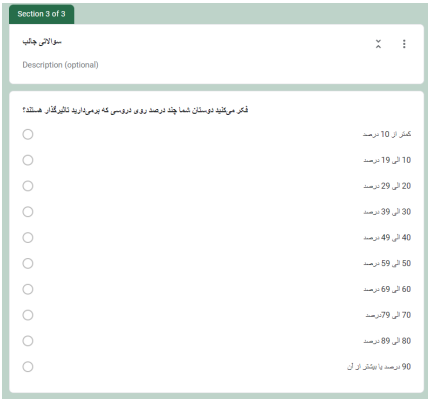
appendices

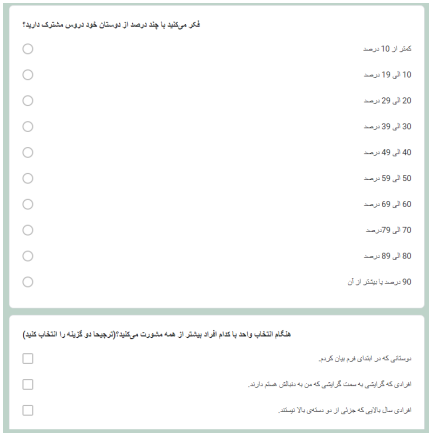
network analysis code

I uploaded the code in my GITHUB with ID: haniehtm.

Questionnaire used in the study

[3]Here we have summarized the image of the questionnaire that we sent to people:





[1] The code is placed in Appendix A-1.
[2] The code is placed in Appendix A-1.
[3] Using the final questions, we wanted to reach a series of shared analyzes at the end of the article, but unfortunately we didn't have the time.