THE OHIO STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

SYLLABUS: STAT 4690
STATISTICAL ANALYSIS OF NETWORKS
SPRING 2024

Course overview

Instructor

Instructor: Subhadeep Paul
Email address: paul.963@osu.edu
Lectures: Tuesdays and Thursdays, 12:45pm–2:05pm in Mendenhall Lab 173
Office hours: Cockins Hall 231, Tuesdays 4:30 – 5:30 PM.

Teaching Assistant

Name: Fangyi Wang

Email address: wang.15022

Office hours: Cockins Hall 332, Wednesdays 1-2 PM.

Prerequisite:

Prerequisites: Stat 3201, 3202 and Stat 3301 or equivalent.

Course description

The course is intended to introduce the field of statistical analysis of network data at the undergraduate level. The course will have a good mix of statistical methods and applications. We will discuss a variety of methods for analyzing network data and learn to apply those methods to datasets using R packages. Applications of these methods to network data from social sciences, economics, psychology, neuroscience, and biology will be discussed.

Module 1: Introduction

- Basic concepts and properties of graphs, motivation for network analysis,
- Historical context of network analysis (in statistics, sociology, math, physics, computer science etc.)
- Application areas of network data
- Community Detection
- Small-world property and heterogeneous degree distribution
- Data structures - single network, multi-layer network, time varying network, multimodal network

Module 2: Stochastic Block Models

- Erdos-Renyi random graphs and their properties.
- Stochastic Block Models – definition and properties, estimation methods - maximum likelihood, modularity optimization, variational methods, spectral methods
- Characterizing the error of community detection and evaluation of the methods
- Extensions - degree corrected, mixed membership, superimposed models, and graphons.

Module 3: Latent Space Models

- Modeling choices – social relations model, additive and multiplicative effects model
- Estimation and Inference
- Extensions - hyperbolic space, more complex data structures

Module 4: Dynamic and Multilayer networks

- Multilayer networks: models and methods
- Discrete time dynamic networks
- Continuous time interaction networks

Course learning outcomes

Upon successful completion of the course, students will be able to

1. Compute various summary statistics and properties of network data
2. Fit statistical models to network data and assess model fit
3. Quantify accuracy of various methods of network analysis
4. Perform statistical learning tasks on network data including community detection, link prediction, and classification
5. Identify appropriate analysis strategies for more complex networked data structures including dynamic and multilayer networks
6. Apply the network analysis methods to network data obtained from a variety of disciplines.

Course materials and technologies

Textbooks

Required

Statistical Analysis of Network Data with R (2020) by Eric D. Kolaczyk and Gábor Csárdi, Springer New York, NY

The electronic version of this book is freely available to OSU students from the OSU library at this link (please log in with your OSU credentials)

https://library.ohio-state.edu/record=b9007501

(click on “connect to resource SpringerLink” to download a PDF copy).

Course material will be supplemented with typed lecture notes that will be provided regularly.

Necessary software

- This class requires you to use the statistical software package called R (The R Project for Statistical Computing; http://www.r-project.org/). This software package is available as Free Software.

    - You can download R for Windows, Mac, and Linux, from the CRAN archive at https://cran.r-project.org.
    - An in-depth introduction to R is available at http://cran.r-project.org/doc/manuals/R-intro.pdf
- An easier to use interface to R is available in the software package RStudio (now posit). This package is available for Windows, Mac, and Linux and can be downloaded for free from https://posit.co/. **Note that RStudio requires R to be installed.**

Grading and faculty response

Grades

Assignment or category	Percentage
Homework Assignments	**50**
Mid-term in-class Exam	**20**
Final Project	**30**
Total	**100**

Grading scale

93–100: A
90–92.9: A-
87–89.9: B+
83–86.9: B
80–82.9: B-
77–79.9: C+
73–76.9: C
70 –72.9: C-
67 –69.9: D+
60 –66.9: D
Below 60: E

Assignment information

Homework: Homework will be assigned regularly (about 5-6 assignments over the semester). It will consist of both written mathematical and numerical problems as well as data analysis problems that require you to write R codes. You are encouraged to work together on the problems, but each student must hand in their own work, written in their own words. Do not copy any part of another student's homework including computer output. Use of homework solutions distributed in previous offerings of the course or available on the web constitutes academic misconduct and will be handled according to university rules. All homework must be submitted online as a PDF file through the class website (carmen). For written problems you may take a picture of or scan your written solutions and upload them as a PDF file. For R based problems please submit typed solutions. Please be sure that the questions are clearly labeled, all supporting work (including computer code) can be easily identified, and that all figures/tables are referenced and interpreted in the text.

Please note late submission of assignments will be penalized at a rate of 10% of points received in the assignment per day they are late. The late penalty may be waived if prior exception has been sought**.** if you are unable to complete an assignment on time and need extra time, please get in touch with me as soon as possible so we can discuss your situation.

In-class midterm Exam: There will be one in-class midterm exam. This course has no final exam. Coverage of the midterm exam includes lecture material, assigned reading, and homework. The exam will be in-person at the usual lecture room during regular class hours. This will be a closed book/closed notes exam. Further details will be given in advance of the exam. Statistical tables will be provided as needed. Calculators may be used, but no communication devices are allowed (e.g. mobile phones). You may use one 8.5x11 inch handwritten sheet of paper (both sides) with formulas. Makeup exam requires a valid excuse and official proof if I am notified in advance or as soon as possible.

Final Project: There is one group project in this course. Students will be expected to work in groups of 3-4 people to complete the project. Each group will analyze a network dataset of their choice and present the findings during the final project presentation on the last day of classes.

Tentative Dates for exams and presentations:
In-class midterm exam: Tuesday, March 05, 2024, during class time.
Final project presentation: Thursday, April 18, 2024, during class time.

Faculty feedback and response time

I am providing the following list to give you an idea of my intended availability throughout the course.

Grading and feedback

For bi-weekly assignments, you can generally expect graded homeworks back within **14 days**.

E-mail

I will reply to e-mails related to course logistics and technical questions that can be answered over email within **24 hours on school days**. However, if you have questions on understanding the course materials or regarding the homework assignments, please consider visiting my office hours or the TA's office hours.

Other course policies

Academic integrity policy

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct **http://studentlife.osu.edu/csc/**.

If I suspect that a student has committed academic misconduct in this course, I am obligated by university rules to report my suspicions to the Committee on Academic Misconduct. If COAM determines that you have violated the university's *Code of Student Conduct* (i.e., committed academic misconduct), the sanctions for the misconduct could include a failing grade in this course and suspension or dismissal from the university.
If you have any questions about the above policy or what constitutes academic misconduct in this course, please contact me.
Other sources of information on academic misconduct (integrity) to which you can refer include:

- Committee on Academic Misconduct web page (go.osu.edu/coam)
- *Ten Suggestions for Preserving Academic Integrity (*go.osu.edu/ten-suggestions*)*

Copyright for instructional materials

The materials used in connection with this course may be subject to copyright protection and are only for the use of students officially enrolled in the course for the educational purposes associated with the course. Copyright law must be considered before copying, retaining, or disseminating materials outside of the course.

Statement on Title IX

Title IX makes it clear that violence and harassment based on sex and gender are Civil Rights offenses subject to the same kinds of accountability and the same kinds of support applied to offenses against other protected categories (e.g., race). If you or someone you know has been sexually harassed or assaulted, you may find the appropriate resources at http://titleix.osu.edu or by contacting the Ohio State Title IX Coordinator at titleix@osu.edu

Commitment to a diverse and inclusive learning environment

The Ohio State University affirms the importance and value of diversity in the student body. Our programs and curricula reflect our multicultural society and global economy and seek to provide opportunities for students to learn more about persons who are different from them. We are committed to maintaining a community that recognizes and values the inherent worth and dignity of every person; fosters sensitivity, understanding, and mutual respect among each member of our community; and encourages each individual to strive to reach his or her own potential. Discrimination against any individual based upon protected status, which is defined as age, color, disability, gender identity or expression, national origin, race, religion, sex, sexual orientation, or veteran status, is prohibited.

Land Acknowledgement

We would like to acknowledge the land that The Ohio State University occupies is the ancestral and contemporary territory of the Shawnee, Potawatomi, Delaware, Miami, Peoria, Seneca, Wyandotte, Ojibwe and Cherokee peoples. Specifically, the university resides on land ceded in the 1795 Treaty of Greeneville and the forced removal of tribes through the Indian Removal Act of 1830. I/We want to honor the resiliency of these tribal nations and recognize the historical contexts that has and continues to affect the Indigenous peoples of this land.

More information on OSU's land acknowledgement can be found at https://mcc.osu.edu/about-us/land-acknowledgement

Your mental health

As a student you may experience a range of issues that can cause barriers to learn, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance or reduce a student's ability to participate in daily activities. The Ohio State University offers services to assist you with addressing these and other concerns you may be experiencing. If you or someone you know are suffering from any of the aforementioned conditions, you can learn more about the broad range of confidential mental health services available on campus via the Office of Student Life's Counseling and Consultation Service (CCS) by visiting ccs.osu.edu or calling 614-292-5766. CCS is located on the 4th Floor of the Younkin Success Center and 10th Floor of Lincoln Tower. You can reach an on call counselor when CCS is closed at 614-292-5766 and 24 hour emergency help is also available 24/7 by dialing 988 to reach the Suicide and Crisis Lifeline.

Accessibility accommodations for students with disabilities

Requesting accommodations

The university strives to maintain a healthy and accessible environment to support student learning in and out of the classroom. If you anticipate or experience academic barriers based on your disability (including mental health, chronic, or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion.

If you are isolating while waiting for a COVID-19 test result, please let me know immediately. Those testing positive for COVID-19 should refer to the Safe and Healthy Buckeyes site for resources. Beyond five days of the required COVID-19 isolation period, I may rely on Student Life Disability Services to establish further reasonable accommodations. You can connect with them at slds@osu.edu; 614-292-3307; or slds.osu.edu.

Accessibility of course technology

This course requires use of CarmenCanvas (Ohio State's learning management system) and other communication and multimedia tools. If you need additional services to use these technologies, please request accommodations with your instructor.

- Canvas accessibility (go.osu.edu/canvas-accessibility)
- Streaming audio and video
- CarmenZoom accessibility (go.osu.edu/zoom-accessibility)
- Collaborative course tools

Course schedule (tentative)

Refer to the Carmen course for up-to-date assignment due dates.

Week	Dates	Topics
1	01/09, 01/11	Descriptive summaries and properties of graphs, visualization of network data
2	01/16, 01/18	Centrality measures, clustering coefficient, small-world property
3	01/23, 01/25	Community structure and modularity
4	01/30, 02/01	Properties of Erdos-Renyi random graphs and simulation
5	02/06, 02/08	The stochastic block model (SBM), maximum likelihood estimation
6	02/13, 02/15	SBM variational EM method, fitting SBM to networks
7	02/20, 02/22	Spectral methods for community detection and SBM
8	02/27, 02/29	Accessing accuracy of community detection – overview of theory and simulation-based evidence
9	03/05, 03/07	In-class midterm exam, Extensions of SBM- degree corrected and mixed membership
10	03/12, 03/14	Spring break – no classes
11	03/19, 03/21	Latent space models (LSM) – motivation and desctiption
12	03/26, 03/28	Estimation and inference in LSMs, fitting LSMs to network data
13	04/02, 04/04	LSM – additive and multiplicative effects, assessing model fit
14	04/09, 04/11	Multilayer networks – examples, descriptive properties, models and methods
15	04/16, 04/18	dynamic network models and methods, Project presentation