Syllabus for CSCI 5352
Network Analysis and Modeling
Spring 2025

Lectures: Tuesday and Thursday from 11:30am – 12:45pm in JSCBB E125

Lecturer: Aaron Clauset

Email:	aaron.clauset
Zoom Office Hours:	Thursday 1:30–2:30pm or by appointment
Course Web Page:	Canvas
Q&A Forum	Piazza

Description: This graduate-level course will examine modern techniques for analyzing and modeling the structure and dynamics of complex networks. The focus will be on statistical algorithms and methods, and both lectures and assignments will emphasize model interpretability and understanding the processes that generate real data. Applications will be drawn from computational biology and computational social science. No biological or social science training is required. (Note: this is not a scientific computing course, but there will be plenty of computing for science.)

Prerequisites (recommended): CSCI 3104 (undergraduate algorithms) and APPM 3570 (applied probability), or equivalent preparation. An understanding of basic statistical concepts (distributions, summary statistics, Monte Carlo), scientific programming, and calculus (integration and differentiation).

Textbooks:
(Required) *Networks: An Introduction* by M.E.J. Newman (2nd edition)

Learning Objectives:
- develop network intuition, and understand how to reason about network phenomenona
- understand network representations and their implications for analysis and modeling
- learn principles and methods for describing and clustering network data
- learn to predict missing network information
- understand how to conduct and interpret numerical network experiments
- analyze and model real-world network data

Overview:
- lectures 2 times a week, some guest lectures and some class discussions
- problem sets (5 total) due every 2 weeks throughout the semester
- a class project, presentation, and final report
- this will be a challenging and fun course; plan accordingly

Schedule:

Week 1	Fundamentals of networks	—
Week 2	Ethics and networks	—
Week 3	Network representations and descriptive statistics	PS1 due
Week 4	Random graphs and network intuition	—
Week 5	Random graphs and null models	PS2 due
Week 6	Network prediction, node attributes	—
Week 7	Network prediction, missing links	PS3 due
Week 8	Community structure and mixing patterns	project proposal due
Week 9	Community structure models	—
Week 10	Spreading processes and cascades	PS4 due
Week 11	Spring break	—
Week 12	Spreading processes with structure	—
Week 13	Advanced link prediction	PS5 due
Week 14	Ranking in networks	—
Week 15	Project presentations	—
Week 16	Project presentations	—
Week 17	Finals (no class)	project writeup due

Problem Set Deadlines

There are 6 problem sets, due every 2 weeks
Problem sets are always due on Friday, by 11:55pm MT
Problem set files will be provided through the class Canvas page
All submitted work (problem sets and projects) is via Canvas

Class Project Deadlines

Project proposals are due Friday, March 7, by 11:55pm MT (Week 8)
Project presentations are in class, April 21 – May 1 (Weeks 15–16)
Final project report due Monday, May 5 by 2:00pm MT

Examinations

There are no exams this semester

Getting help

- **Attend the lectures and come to office hours.** Lectures are where all material will be introduced. Office hours are for any student wanting clarification, advice, support, feedback, or just a conversation. (If no one shows up, I do email or something for an hour, which is way more boring than talking to students about science and networks.)

- **Use the Canvas page**. All class materials, including lecture notes, supplementary videos, problem sets, and submissions will be done via the class Canvas. Grades will not be tracked there; you are responsible for tracking your own progress in the course.

- **Use the Piazza forum**. Asynchronous Q&A will be done on the class Piazza, which is

intended to help you get help from classmates and myself. Rather than emailing questions to me, please post your questions on our Piazza forum.

- **Email is for urgent matters**. Don't ask for help with specific parts of problem sets via email; use office hours or Piazza for those questions. Please reserve email for high-priority personal matters.

Grading

Grades will be assigned based on problem set scores (50%) and class project components (50%).

Letter grades will be calculated only after the project reports are scored. Prior to that, only numerical scores will be tracked. Don't ask me to estimate your final letter grade! There may be extra credit on the problem sets; doing the extra credit can only increase your final grade.

Class project

- In the class project, students will explore a class topic (of their choice) more deeply.

- It may be done individually or in a team of two (2). If you choose to work with another student, you are responsible for finding that person and for evenly dividing the work.

- There are three (3) deliverables associated with the class project:

  1. *Class project proposal* (due Week 8, via Canvas, PDF only)
     Must include: (1) the names of the individual(s) on the team, (2) no more than 800 words describing the (i) background material, (ii) research question, and (iii) anticipated findings, and (3) a brief description of any data and algorithms you plan to use.

  2. *Project presentation* (due Weeks 15–16, in class)
     A roughly 10 minute-ish presentation of the project idea and any results found.

  3. *Project paper* (due Week 16, via Canvas, PDF only)
     A 10-page scientific report on the project idea, results and what you learned. The report should be formatted like a scientific paper (11pt font, 2-column, single spaced, 1-inch margins), with Introduction, Methods, Results, Discussion, and Bibliography sections, with an optional Appendix section for code.[1] The Bibliography should include appropriate references for your work (primary literature, course materials, websites, source data, methods, etc.). If you would like advice about how to write your report, I'm happy to provide it.
     Warning: Do not plagiarize any part of the report. Suspected plagiarism will receive a 0.

Advice: Try to choose a project that you can complete in 7 weeks, and which takes about $3\times$ the time you spend on one problem set in the course. Here are four examples of projects:

1. Reproduce the results of some paper on networks in the scientific literature.

2. Choose one or more empirical networks and then interpret the outcome of applying a variety of methods from the class.

---

[1]Bibliography and Appendix sections do not count against your page limit. A link to a `github` repo for code is okay, but you must also include your code in the PDF, unless you request and obtain from me permission not to.

3. Contribute a novel implementation of methods covered in the class to `networkx`.

4. Make up your own idea!

The one key requirement of the project is that it involves networks. I'm happy to provide feedback on your ideas *before the project proposal deadline*—my office hours are an excellent time to chat about it. Note: if your proposal is not well scoped, I may ask you to revise it.
Warning: you are responsible for pacing yourself on your project; the class has no "progress report" deadlines midway through the project; plan accordingly.

Problem sets

- Biweekly deadline : Friday, by 11:55pm MT, via the class Canvas page
    - *No credit for solutions submitted any other way*
    - *Late submissions will not be accepted*
    - *Solutions must be submitted as a single PDF file*
    - *Solutions must include your source code, appended to the end of your PDF*
    - *The lowest 1 problem set grade will be dropped at the end of the semester*
    - *Figures: No credit for any figure with unlabeled axes or data series.*

- Programming and data analysis problems may be completed in a programming language of your choice. Many students find great success with Jupyter notebooks, which have good support for data analysis and visualization. You may use standard libraries, e.g., in Python: `networkx`, `numpy`, `scipy`, `webweb`, as needed. If you're not sure about a library, ask.

- Partial credit will be awarded; to maximize your chances of receiving some partial credit, show your work and explain your thinking.

Working in teams

- Problem sets are to be completed individually.

- The class project may be completed in a team of 1 or 2.

- Students are *encouraged* to form study groups to discuss the problems with each other.

- **How to avoid cheating?** Simple: each student must *independently* write up their *own* solution—the one they submit for credit for themselves. Copying any part (words or code) is strictly forbidden. Submitting lightly edited work by someone else, e.g., changing a few words or variable names, is cheating. Instead, work together at a whiteboard or on paper, but then, even if you solved the problem together, you must write up and submit your own version (your own words!) of it.

Intellectual Honesty and Plagiarism

- Intellectual dishonesty or plagiarism of any form, at any level, *will not be tolerated.*

- **Discussing questions with other students is encouraged, but** you must list your collaboration in the solutions you submit. If you discussed it with 10 other people, then all 10 names should appear in your solution. If someone was particularly helpful, say so. Be generous! If you're not sure whether someone should be included in your list of collaborators, include them. For discussions in class, in section, or in office hours, where collecting names is impractical, it's okay to write something like "discussions in class." There is never a penalty for discussing problems with other students.

- **Copying from any source, in any way, is strictly forbidden**. This includes the Web, chatGPT or other generative AI systems, and other students (past or present). Asking or paying someone on the internet for answers is dishonest. If you are unsure about whether something is permitted, I'm happy to talk over what you're thinking about.

  A note about generative AI: OpenAI's `chatGPT` and other LLMs are "premium bullshit" generators that don't actually know anything and cannot reason like you can. But they are very good at stringing together and reshaping bits of text or code from their training data, into verbally pleasing or syntactically correct output. If interacting with one helps you learn the course material or figure out how to do the work correctly, that's great and permissible! But, under no circumstances should you submit LLM answers (text or code), in whole or in part, as something you yourself wrote. Because: they make mistakes, over-confidently, and it's your responsibility as a user to know the difference between bullshit and a correct answer. Moreover, for scientific topics like biological networks, its training data doesn't include examples of the kind of programming tasks or scientific interpretation I'll ask you to do in this class.

  There will be a zero-tolerance policy to violations of this policy. Violators will be removed from the class, given a grade of F, and reported to the CU Honor Council.

- **Write everything in your own words and cite all outside resources**. You are strongly encouraged to use outside resources, e.g., the Internet, other textbooks, the primary literature, etc.[2] to teach yourself how to solve any problem. But, you must write your solutions *in your own words*. I'm not interested in seeing Wikipedia's answer, or chatGPT's answer, or anyone else's, in whole or in part. The only sources you are not required to cite are the course materials. As a small bonus for having read the syllabus carefully, I will award some extra credit points if, within the first two weeks of class, you send me an email containing a picture of your favorite scientist (living or dead) and a blurb about why you admire them.

Advice for writing up your solutions:
Your solutions for the problem sets should have the following properties. I will be looking for these when I grade them:

---

[2]This is the 21st century after all. It would be madness not to encourage you to use the greatest tool ever created for scholarship. But, with great power comes great responsibility. Use it wisely, you must.

1. **Clarity**: Your solutions should be both clear and concise. The longer it takes me to figure out what you're trying to say, the less likely you are to receive full credit. The more clear you make your thought process, the more likely you are to get full credit.

2. **Completeness**: Full credit is based on (i) sufficient intermediate work and (ii) the final answer. For many problems, there are multiple paths to the correct solution, and I need to understand exactly how you arrived at the solution. A heuristic for deciding how much detail is sufficient: if you were to present your solution to the class and everyone understood the steps and could repeat them themselves, then you can assume it is sufficient.

3. **Succinctness**: Solutions should be long enough to convince me that your answer is correct, but no longer. More than half a page of dense algebra, more than a few figures or more than a page or two per problem is probably not succinct. Clearly indicate your final answer (circle, box, underline, whatever). Rewriting your solutions, with an eye toward succinctness, before submitting will help. Strive for maximum understanding in minimum space.

4. **Numerical experiments**: Some programming problems will require you to conduct numerical experiments using random number generators. One run is not a result. Your goal is to produce beautifully smooth central tendencies and you should average your measured quantity over as many independent trials as is necessary to get something smooth. Further, your results should span several orders of magnitude. I recommend a dozen or so measurement values across the x-axis, distributed logarithmically, e.g., $n = \{2^4, 2^5, 2^6, \ldots\}$.

    Solutions that use a numerical experiment but fail to adequately explain the experimental design will receive an automatic 0.

5. **Source code**: Your source code for all programming problems must be included at the *end* of your solutions. Code must include copious comments explaining the sub-algorithms and must be run-able; that is, if I try to compile and run it, it should work as advertised.

6. **Data analysis**: In presenting results from analyzing real data, you should always briefly describe the data to the reader. Explain what the network is (what is a vertex and when are two vertices connected) and what any network meta-data (vertex attributes, edge weights, etc.) means. Try also to explain what questions you are investigating, and how your results address those questions.

7. **Figures**: Always label your axes and always label your data series. Avoid having a lot of wasted whitespace in your figures (choose appropriate x- and y-ranges). Know what message you want the reader to take away from your figure, and be sure your figure accomplishes it clearly.

    Figures that lack axis or data series labels will receive an automatic 0.

8. **Solutions should be detailed and clear.** Each of your answers should include a clear, written explanation of how it answers the question. Code only, without a written explanation of what it is supposed to do and how, is almost never sufficient for full credit. The more clear the explanation, the more likely you are to receive full credit.

9. **Exceptional circumstances**: Only in exceptional circumstances, e.g., incapacitation due to illness or injury, may assignments be forgiven. *Unexceptional* circumstances include registering late, travel for job interviews or conferences or fun, forgetting the homework deadline, or

simply not finishing on time. Final grades will be computed as if forgiven assignments did not exist; this increases other assignments' weight.

Honor Code: As members of the CU academic community, we are all bound by the CU Honor Code. I take the Honor Code very seriously, and I expect that you will, too. Any significant violation will result in a failing grade for the course and will be reported. Here is the University's statement about the matter:

All students enrolled in a University of Colorado Boulder course are responsible for knowing and adhering to the Honor Code. Violations of the Honor Code may include but are not limited to: plagiarism (including use of paper writing services or technology [such as essay bots]), cheating, fabrication, lying, bribery, threat, unauthorized access to academic materials, clicker fraud, submitting the same or similar work in more than one course without permission from all course instructors involved, and aiding academic dishonesty. Understanding the course's syllabus is a vital part in adhering to the Honor Code.

All incidents of academic misconduct will be reported to Student Conduct & Conflict Resolution: StudentConduct@colorado.edu. Students found responsible for violating the Honor Code will be assigned resolution outcomes from the Student Conduct & Conflict Resolution as well as be subject to academic sanctions from the faculty member. Visit the Honor Code website for more information on the academic integrity policy.

Accommodation for Disabilities, Temporary Medical Conditions, and Medical Isolation: If you qualify for accommodations because of a disability, please submit your accommodation letter from Disability Services to your faculty member in a timely manner so that your needs can be addressed. Disability Services determines accommodations based on documented disabilities in the academic environment. Information on requesting accommodations is located on the Disability Services website. Contact Disability Services at 303-492-8671 or DSinfo@colorado.edu for further assistance. If you have a temporary medical condition, see Temporary Medical Conditions on the Disability Services website.

If you have a temporary illness, injury or required medical isolation for which you require adjustment, please notify me via a direct email so that I can work with you to find a suitable adjustment.

Accommodation for Religious Obligations: Campus policy requires faculty to provide reasonable accommodations for students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. Please communicate the need for a religious accommodation in a timely manner. In this class, I will make reasonable efforts to accommodate such needs if you notify me of their specific nature by the end of the 3rd week of class. See the campus policy regarding religious observances for full details.

Preferred Student Names and Pronouns: CU Boulder recognizes that students' legal information doesn't always align with how they identify. Students may update their preferred names and pronouns via the student portal; those preferred names and pronouns are listed on instructors' class rosters. In the absence of such updates, the name that appears on the class roster is the student's legal name.

Classroom Behavior: Students and faculty are responsible for maintaining an appropriate learning environment in all instructional settings, whether in person, remote, or online. Failure to adhere to such behavioral standards may be subject to discipline. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with race, color, national origin, sex, pregnancy,

age, disability, creed, religion, sexual orientation, gender identity, gender expression, veteran status, marital status, political affiliation, or political philosophy.

For more information, see the policies on classroom behavior and the Student Conduct & Conflict Resolution policies.

Sexual Misconduct, Discrimination, Harassment, and/or Related Retaliation: CU Boulder is committed to fostering an inclusive and welcoming learning, working, and living environment. University policy prohibits protected-class discrimination and harassment, sexual misconduct (harassment, exploitation, and assault), intimate partner abuse (dating or domestic violence), stalking, and related retaliation by or against members of our community on- and off-campus. The Office of Institutional Equity and Compliance (OIEC) addresses these concerns, and individuals who have been subjected to misconduct can contact OIEC at 303-492-2127 or email CUreport@colorado.edu. Information about university policies, reporting options, and OIEC support resources including confidential services can be found on the OIEC website.

Please know that faculty and graduate instructors are required to inform OIEC when they are made aware of incidents related to these concerns regardless of when or where something occurred. This is to ensure that individuals impacted receive outreach from OIEC about their options and support resources. To learn more about reporting and support for a variety of concerns, visit the Don't Ignore It page.