

**Syllabus for CSCI 5352**  
**Network Analysis and Modeling**  
**Fall 2013**

**Lectures:** Tuesdays and Thursdays from 12:30pm – 1:45pm in ECCS 1B12

**Lecturer:** Aaron Clauset

Office: ECOT 743  
Email: [aaron.clauset@colorado.edu](mailto:aaron.clauset@colorado.edu)  
Web Page: <http://tuvalu.santafe.edu/~aaronc/courses/5352/>  
Office Hours: Tuesday 2:00–3:30 or by appointment

**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 adequate mathematical and programming background is mandatory. The concepts and techniques covered in this course depend heavily on basic statistics (distributions, Monte Carlo techniques), scientific programming, and calculus (integration and differentiation). Students without sufficient preparation will struggle to keep up with the lectures and assignments. Students without proper preparation may audit the course.

**Required Texts:**

- (1) *Networks: An Introduction* by M.E.J. Newman
- (2) *Pattern Recognition and Machine Learning* by C.M. Bishop

**Overview:**

- Mostly lecture-style class, with some guest lectures and some class discussions.
- Problem sets (6 total, worth 50% of grade) due every 2 weeks throughout the semester.
- Independent project (worth 30% of grade) due at end of semester.
- No exams.
- Networks are cool.

**Tentative schedule:**

Week 1	Introduction and overview
Week 2	Network basics, and centrality
Week 3	Transitivity, assortativity, degrees
Week 4	Network distances
Week 5	Large-scale network structure
Week 6	Random graphs
Week 7	Configuration model
Week 8	Social and biological networks
Week 9	Preferential attachment
Week 10	Generative models and community structure
Week 11	More generative models
Week 12	Advanced generative models
Week 13	Dynamic networks
Week 14	Fall break
Weeks 15–16	Project presentations

**Deadlines:**

	assigned	due
PS 1	Aug. 27 (Tuesday)	Sept. 9 (Monday)
PS 2	Sept. 10 (Tuesday)	Sept. 23 (Monday)
PS 3	Sept. 24 (Tuesday)	Oct. 7 (Monday)
IP, topic proposal		Oct. 15 (Tuesday)
PS 4	Oct. 8 (Tuesday)	Oct. 21 (Monday)
PS 5	Oct. 22 (Tuesday)	Nov. 4 (Monday)
PS 6	Nov. 5 (Tuesday)	Nov. 18 (Monday)
IP, short presentation		Dec. 3, 5, 12
IP, final write up		Dec. 13 (Friday, 11:00am)

**Coursework and grading:** Attendance to the lectures is required.

Most of the class will be standard graduate-style lectures by me. These will be supplemented by guest lectures on special or advanced topics, and class discussions of selected papers drawn from the network science literature. Many lectures will have associated lecture notes, which I will post on the class website after class.

Grades will be assigned based on (i) performance on the problem sets, (ii) performance on the independent project, and (iii) class attendance and participation. Problem sets will develop and extend selected class topics and will introduce additional topics not covered in class. There are no written examinations in the course. Students are expected to spend serious quality time on the problem sets and project.

*Problem sets (PS):* The 6 problem sets, due every 2 weeks throughout the semester, will include both mathematical and programming problems.

Programming problems may be completed in a programming language of your choice. I recommend using something like `Matlab` or `Python`, which have data analysis and visualization capabilities built-in. Familiarity with `Mathematica` may be useful for some of the mathematical problems, and you are free to use `Mathematica` in any way to complete calculations.

- Problem sets will be due roughly two weeks after they are assigned (schedule given above).
- Solutions **must** be in PDF format (e.g., typeset using  $\text{\LaTeX}$ ), should include all necessary details for me to follow the logic. Non-PDF files will not be accepted.
- You **must** submit your solutions to me via email by 11:59pm the day they are due.
- No late assignments will be accepted, regardless of how small the lateness is.
- Collaboration is encouraged on the problem sets. However, you may not copy (in any way) from your collaborators and you must respect University academic policies at all times. To be clear: you may discuss the problems verbally, but you must write up your solutions separately. If you do discuss the problems with someone (and you are encouraged to!), you must then list and describe the extent of your collaboration in your solutions (a footnote is fine). Copying from any source, in any way, including the Web but especially from another student (past or present), is strictly forbidden.
- If you are unsure about whether something is permitted under these rules, ask me well before the deadline.
- You must submit your source code for your algorithms with your solutions. Source code should be attached to the end of the PDF.
- I do not expect you to code every algorithm from scratch in this class, although I do expect you to do a substantial amount of coding yourself. It is okay with me to use libraries, in conjunction with a regular programming language, that do standard numerical network calculations. If you use such libraries, *you must state in your solution to the corresponding problem* which libraries you used and what you used them for. You are still required to submit the rest of your code; do not submit code for the libraries.

*Independent project (IP):* The purpose of your independent project is to explore a research question of your own devising related to network analysis and modeling. You must work independently, but I encourage you to solicit feedback from other students on your ideas and progress. The deliverables here are a short (10 minute) in-class presentation of your project results and a 10-page writeup due via email to me by 11:00am on Friday, December 13.

A short project proposal is due via email on Tuesday, October 15. (This gives you about 8 weeks to complete the project.) Your proposal should be two paragraphs, describing (i) any background material, including necessary references to the scientific literature, and (ii) what you are specifically going to do. The best project topics are those you can make good progress on in 6 weeks. I am happy to provide feedback on your project ideas, and if your proposal is inappropriately scoped, I may ask you to revise it.

*Grading:* Grades will be assigned based on (5352-002) a 20% attendance, 50% problem sets, and 30% project division, or (5352-740) a 63% problem sets, and 37% project division.

---

### **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:

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.,  $n = \{2^4, 2^5, 2^6, \dots\}$ .

No credit will be given if you fail to label your axes and data series, or if you fail to explain your experimental design.

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.
- 

**Suggestions:** Suggestions for improvement are welcome at any time. Any concern about the course should be brought first to my attention. Further recourse is available through the office of the Department Chair or the Graduate Program Advisor, both accessible on the 7th floor of the Engineering Center Office Tower.

**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 of the University of Colorado at Boulder are responsible for knowing and adhering to the academic integrity policy of this institution. Violations of this policy may include: cheating, plagiarism, aid of academic dishonesty, fabrication, lying, bribery, and threatening behavior. All incidents of academic misconduct shall be reported to the Honor Code Council ([honor@colorado.edu](mailto:honor@colorado.edu); 303-735-2273). Students who are found to be in violation of the academic integrity policy will be subject to both academic sanctions from the faculty member and non-academic sanctions (including but not limited to university probation, suspension, or expulsion). Other information on the Honor Code can be found at <http://www.colorado.edu/policies/honor.html> and at <http://www.colorado.edu/academics/honorcode/>

**Special Accommodations:** If you qualify for accommodations because of a disability, please submit to your professor a letter from Disability Services in a timely manner (for exam accommodations provide your letter at least one week prior to the exam) so that your needs can be addressed. Disability Services determines accommodations based on documented disabilities. Contact Disability Services at 303-492-8671 or by e-mail at [dsinfo@colorado.edu](mailto:dsinfo@colorado.edu).

If you have a temporary medical condition or injury, see Temporary Injuries under Quick Links at Disability Services website and discuss your needs with your professor.

Campus policy regarding religious observances requires that faculty make every effort to deal reasonably and fairly with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. 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 full details at [http://www.colorado.edu/policies/fac\\_relig.html](http://www.colorado.edu/policies/fac_relig.html)

**Classroom Behavior:** Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail 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 differences of race, color, culture, religion, creed, politics, veterans status, sexual orientation, gender, gender identity and gender expression, age, disability, and nationalities. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records. See policies at <http://www.colorado.edu/policies/classbehavior.html>  
[http://www.colorado.edu/studentaffairs/judicialaffairs/code.html#student\\_code](http://www.colorado.edu/studentaffairs/judicialaffairs/code.html#student_code)

**Discrimination and Harrassment:** The University of Colorado at Boulder Discrimination and Harassment Policy and Procedures, the University of Colorado Sexual Harassment Policy and Procedures, and the University of Colorado Conflict of Interest in Cases of Amorous Relationships policy apply to all students, staff, and faculty. Any student, staff, or faculty member who believes s/he has been the subject of sexual harassment or discrimination or harassment based upon race, color, national origin, sex, age, disability, creed, religion, sexual orientation, or veteran status should contact the Office of Discrimination and Harassment (ODH) at 303-492-2127 or the Office of Student Conduct (OSC) at 303-492-5550. Information about the ODH, the above referenced policies, and the campus resources available to assist individuals regarding discrimination or harassment can be obtained at <http://www.colorado.edu/odh>