Teaching Statement

Bruno Ribeiro
ribeiro@cs.cmu.edu
http://www.cs.cmu.edu/~ribeiro/
Phones: 617-945-6927 (cell)
412-256-4935 (office)

Networks are at the core of many of today's highest impact computing applications, with the mining and modeling of this networked data increasingly enabling new disruptive technologies and driving scientifically minded approaches to business decisions across all industries. Graduate and undergraduate students, as future industry and academic leaders, need to understand how this data is collected, how large datasets can be efficiently processed, and what can be reliably ascertained from the output. As an educator, I aim to:

- Develop courses centered on data mining, network science, data science, and advanced computer networks, emphasizing collective networked behavior;
- Mentor graduate students towards designing accurate models and principled statistical and data mining tools that will support new applications and better decision making;
- Teach traditional courses in data mining and computer networking with a focus on fundamental principles;
- Motivate undergraduate students in courses such as data structures and programming with state-of-the-art social network app development examples. Such examples are not only fun for undergrads but also teach fundamental principles of statistics of graphs and social network research.

TEACHING PHILOSOPHY

Richard Feynman is one of my sources of inspiration for teaching. Feynman revolutionized the way science was taught in Brazil in the 60s. Feynman was a gifted scientist and an inspiring instructor who was able to make his students understand complex ideas through basic principles and familiar analogies that had a life-long impact on them. I aspire to do the same for my students.

In CS, statistics, and applied probability, the majority of the techniques and algorithms (even the most advanced ones) are based on simple concepts. My pedagogic approach is to present these basic concepts upfront (which have a long half-life) and then introduce the techniques as examples of that concept. I find teaching at all levels enjoyable and I also believe in learning through hands-on experimentation with real datasets.

TEACHING EXPERIENCE

I have taught and developed the following courses. Full syllabi can be provided upon request. The first two courses were *first-time* graduate offerings at UMass Amherst.

CS691NW Network Science (Fall 2011) (UMass Amherst, Graduate Course)

Instructors: Bruno Ribeiro & Don Towsley

This course, introduced for the first time at UMass, explored the mathematical theory of networks, particularly social and technological networks. The course showed applications to search networks, identify influential nodes, and examine network resilience, routing, epidemiology, and social influence. Topics included: experimental studies of social networks, the World Wide Web and information and biological networks; methodologies for the analysis of network data; graph theory; probabilistic models of networks, preferential attachment models and the small-world model; and network dynamics. Homework assignments were designed to further instill the material through theoretical exercises and hands-on experiments with real datasets.

Responsibilities: I took the lead in developing the course material, class schedule, syllabus, exams, and homework assignments. I was also responsible for grading the exams and supervising the teaching assistant (TA) assigned to the course. Lectures and office hours were equally shared with my co-instructor.

CS691SG (STAT691SG) Statistics of Graphs and Networks (Fall 2012)
(UMass Amherst, Graduate Course)
Instructors: Bruno Ribeiro, Krista Gile, Don Towsley

Mining and modeling complex systems using large amounts of networked data brings new possibilities but also new challenges. Statistics of Graphs and Networks was a research seminar designed to look into the problem of inferring network relationships and user characteristics from incomplete data. The topics were wide ranging, including models for cross-sectional and dynamic networks, partially-observed networks, and networks with latent structures. Also included were network sampling, power laws, non-parametric statistical methods for networks, and applications to epidemiology, phylogenetics, and network tomography. The course discussion highlighted statisticians' and computer scientists' methods for the statistical analysis of networks. Students were given required reading lists of research articles and were required to present them, as well as listen to their peers' presentations and discuss them, providing critical feedback. Students taking three credits presented a research project at the end of the semester.

Responsibilities: All three instructors had an equal share of responsibility for the course; however, my differentiated role was helping students define project topics and presenting alternative points of view to challenge students during classroom discussions.

A Four-Semester Series of Graduate Student Seminars
As a graduate student at UMass Amherst, I developed and ran a series of student seminars tackling themes as diverse as data mining, systems biology, network science, sensor management, and wireless networks during a total of four semesters. From this I learned that the key to a lively seminar lies in the first day, when papers are assigned. I always proposed a long list of interesting papers (asking for input from the faculty, senior graduate students, and postdocs), read all papers, and quickly described what was interesting in the paper in the first lecture. Then, I would ask for volunteers, mildly applying some arm-twisting and peer pressure to keep the schedule full and in check. When the presenter worked in the same area of knowledge as the paper he or she presented, I would ask the presenter to propose two possible extensions to the work.

COURSE PREFERENCES

My background and interests make me well suited to teach a variety of graduate and undergraduate courses including data mining, computer networking, statistics of networks, probability theory, and network science.

In addition to these disciplines, I enjoy teaching introductory CS courses, especially introductory data structures and advanced programming courses, where I can challenge undecided students to build fun systems that embody some of the principles of my data mining and social networking research. This approach should attract undecided students with a special affinity for CS and a desire to build real systems (e.g., Facebook apps) that can have a real impact on society. Through this approach, I plan to mentor undergraduate students towards pursuing further CS coursework.

I am also interested in teaching seminar courses. I find them to be ideal conduits to exploratory research of new and exciting research topics.

STUDENT MENTORING

My research is centered on deeply understanding statistical problems and probability models through real-world problem solving. I find working together with students on compelling social problems an effective way of teaching them how to approach problems in a principled way.

- **Fabricio Murai** (3rd year Ph.D. Candidate, UMass Computer Science). Since his arrival at UMass in 2011, Fabricio has received continuous support from one of the grants that I co-lead. Since then, Fabricio and I have written four papers together: one journal paper (IEEE JSAC), one conference paper (IEEE INFOCOM), a workshop paper (SIMPLEX @ WWW'13), and a new paper under submission. I proposed the problem and

mentored Fabricio on his outstanding Ph.D. candidacy project (UMass CS's qualifying exam is a project) that won him the *2013 UMass CS Outstanding Synthesis Award*. I have also mentored Fabricio in selecting the right courses that match his research interests and career aspirations. Fabricio passed his Ph.D. Candidacy Exam with distinction in 2013.

- **Kun Tu** (5th year Ph.D. Candidate, UMass Computer Science). Since Fall 2013, I have had the pleasure of mentoring Kun in his work on the analysis and design of recommendation systems for online dating websites. I proposed the topic and mentored Kun through his Ph.D. candidacy project that cleared the way for him to proceed as a Ph.D. candidate. The two resulting works of our efforts were covered in the MIT Technology Review (Emerging Technology From the arXiv) [*under review*] and published at *5th International Workshop on Social Recommender Systems* in 2014.
- **James Atwood** (3rd year Ph.D. Candidate, UMass Computer Science). Since Fall 2013, I have worked closely with James. Our joint work on efficient data structures for random graph generation appeared at the *SIMPLEX Workshop* in 2014.
- **Yeon-sup Lim** (5th year Ph.D. Candidate, UMass Computer Science). I mentored Yeon-sup during his Ph.D. Candidacy project. Yeon-sup was able to fulfill his requirements ahead of the deadline and the resulting work appeared at the *SIMPLEX Workshop* in 2014.
- **Ting-Kai Huang** (UC Riverside Computer Science Ph.D., now at Google). I proposed the problem and mentored Ting-Kai through the work that forms part of his Ph.D. thesis. Ting-Kai successfully defended his thesis in the fall of 2013 and is now at Google. The resulting work was published at the *ACM Conference On Online Social Networks* (*COSN)* 2014 (a relatively new conference with a competitive acceptance rate of 15.9%).
- **Minh Hoang** (4th year Ph.D. Candidate, UCSB Computer Science). I have worked closely with Minh on the challenging project of designing an early warning system that can predict large information cascades over online social networks. Minh is partially funded by a grant that I Co-PI. Our work develops CAMPS, the first principled data-driven approach to predict properties of complex contagion processes (e.g. the size of the infected population after four weeks).