Weaving Empirical Research Into Computer Science Education

Teaching Statement

Jasleen Kaur
Department of Computer Science
University of North Carolina at Chapel Hill

1 Guiding Philosophy

Systems researchers face a challenging responsibility created by the tremendous growth and commercialization of the Internet. On one hand, to keep up with the Internet's rapid growth rate, commercial interests have relied on "quick fixes" to existing networking technologies. On the other hand, the success of the Internet has greatly raised the expectations of users, which can not be met by relying simply on minor tweaks to legacy network designs.[1] Researchers need to continually question whether the assumptions and goals used for designing networking solutions in the past are valid any more. Furthermore, systems professionals must find the courage to stop treating large distributed systems such as the Internet as black-boxes and use measurement and analysis techniques to look inside these. The *guiding principle* of my faculty career is to adopt this attitude in my research and inculcate it in the students I interact with.

Unfortunately, a majority of current *systems* courses deal mostly with the design and engineering aspects of systems technology, but rarely go into depths of rigorous methodologies for measurement, controlled experimentation, or scientific data analysis. Furthermore, systems education, like most Computer Science (CS) education, is based mostly on classroom teaching and simplified course projects. These pedagogical techniques are not adequate to understand and appreciate the intricacies of large-scale systems such as the Internet. I have used my research interests to advance the current state of CS education in three main ways: (i) by introducing the treatment of scientific experimentation, measurements and analysis into CS curricula, (ii) by using the experimental tools developed in my research and elsewhere to improve the efficacy of systems education, and (iii) by encouraging participation of undergraduates in systems research projects. I elaborate on my contributions below.

2 Curriculum Development

My curriculum development efforts have resulted in the introduction of three new courses in the CS department.

- A new undergraduate course titled *Systems Performance Analysis* first offered in Spring 2005.

  The prime goal of this course is to fill a significant need for familiarizing CS students with conducting rigorous performance analysis of computer systems. The course design serves two important purposes. First, it introduces concepts in systems performance analysis techniques, including modeling, simulations, measurements, and tracing. It integrates and uses case studies of hardware and software systems—such as CPUs, disks, and a web server—to illustrate most concepts. Second, through projects and assignments, it familiarizes students with several tools that can be used to study stand-alone as well as distributed applications/systems. Course projects pursued in this course include the analysis of a web server, a database server, a gaming server, a media-streaming server, and a kernel process-creation API.

  More details are available at: *http://www.cs.unc.edu/∼jasleen/Courses/Spring05/.*

[1] Some well-argued examples can be found in [3, 6].

- A new introductory graduate course in networking titled *Computer Networks* first offered in Fall 2003.

  This course has been approved to satisfy the *systems distribution requirement* of the Computer Science graduate program at UNC. The course is designed to serve two purposes. First, it reviews traditional topics and introduces students to current research areas in the field of networking. Second, through projects and assignments, it builds in students several experimental and analysis skills required for the empirically-oriented networking and systems research pursued in the UNC CS department.

  More details about the latest offering are available at: *http://www.cs.unc.edu/∼jasleen/Courses/Fall08/*.

- A graduate seminar course titled *Research Topics in Networking* first offered in Spring 2003.

  In this course, students work on semester-long research projects and are involved in the critical review and discussion of several recent publications in special topics from the field of networking. Past offerings have considered the areas of network measurements, and the design and analysis of transport protocols. Students are required to work on open research problems and present a workshop-quality research paper at the end of the semester.

  All previous offerings of this course have been well received and have been fairly successful in giving students a rich hands-on research experience. Several research projects pursued in this course have resulted in refereed publications [2, 4, 5, 7, 8].

In addition to designing and teaching the above courses, I have also assumed the prime responsibility for teaching the undergraduate course, *Internet Services and Protocols* (COMP 431). This course fulfills the systems distribution requirement of the Computer Science undergraduate program at UNC. The course is designed to serve two main purposes. First, it adopts a top-down approach to understanding Internet protocols and services, focusing heavily on application-layer and transport-layer protocols. Second, through several hands-on assignments, it exposes students to issues that affect the performance of Internet applications and protocols.

More details about the latest offering are available at: *http://www.cs.unc.edu/∼jasleen/Courses/Spring08/*.

3 Pedagogical Approach

Deductive-style Instruction Interaction in my classrooms is explicitly designed to involve students in critically thinking and deriving sound principles on which to base system designs and analysis. This implies while preparing for each topic to be discussed in class, I prepare a *deductive* style for discussing it—for every new idea discussed, this involves asking pre-prepared questions to students to encourage them to think critically and "derive" the idea, rather than simply understand it when I present it myself. I have found this style to be fairly effective in arousing curiosity in students, helping them build a sound understanding of fundamentals, and making the classes enjoyable. I have often found that with this style students think much beyond the classroom and come back with questions and suggestions on how to improve upon an idea derived in the class.

Hands-on Experimentation Classroom instruction in most systems courses today is complemented with only simplified course projects. I believe that the intricacies of large-scale systems can be appreciated only by hands-on measurement and analysis. Within the context of networking, for instance, in order to explain to students the properties of Autonomous Systems and network topologies, an extremely effective technique is to show and study the output of *traceroute*-based measurement and analysis tools. Such pedagogical techniques may have been inaccessible for use in the past, but with the advent of wide-area testbeds like PlanetLab and EmuLab, it is becoming increasingly possible to use the Internet as the world's largest and most realistic networking laboratory. I have used several publicly-available measurement and analysis tools—including some that have been developed by my research group—in conjunction with these testbeds, to enhance the effectiveness of systems instruction. I have adopted this approach in each of the courses offered till date.

In addition to the above 'in the wild" measurement and analysis experience, I design assignments and projects that focus on the science of rigorous and controlled experimentation. For instance, the graduate net-

working course that I offer (COMP 631) include a significant experimental component using the publicly-available network simulator, NS-2 [1]. Students are required to design and conduct experiments, and analyze and present observations that help isolate and explain the performance of targeted protocol mechanisms.

I am also collaborating with researchers at the Renaissance Computing Institute (RENCI) at UNC, for using the recently-developed netFPGA platform from Stanford University [9] as the basis for designing course projects for future offerings of my graduate courses. The goal is to have students use these platforms to rapidly develop prototypes of protocols and applications and evaluate them in a laboratory setting. The projects will be scoped to especially encourage cross-group interaction and communication skills by requiring different prototypes to be inter-operable.

4 Student Advising

Graduate Student Advising To date, I have graduated two PhD students—a third student is scheduled to defend his dissertation in January of 2009. I am also currently advising two students who are in their first- and second-years of the graduate program.

Perhaps of all my responsibilities as a faculty member, I take one-on-one advising of students most seriously. To me, this is the only way of repaying the debt of excellent and valuable advising that I myself have received from my former advisers. My aim is to help develop these students into well-rounded researchers—who have the ability to initiate and conduct top-rated research, effectively communicate their ideas, as well as stick to a high level of professional ethics. I work closely with students in their first two years to provide them with a fruitful research experience—this often includes publishing research results at peer-reviewed venues, which goes a long way in encouraging students to pursue a career in research. In subsequent years, I work with my students in developing the critical ability of clearly articulating and presenting their research ideas, both in terms of technical writing as well as oral presentation. I also encourage and support my students in presenting results at professional meetings.

I have also had the good fortune of interacting with several other motivated graduate students by virtue of my offerings of graduate courses (and pursuing semester-long research projects with them), by serving on their dissertation committees, as well as by actively participating in the NetLunch reading group (which typically meets once every two weeks).

Research Experience for Undergraduates I believe that by involving undergraduates in independent research projects at the senior undergraduate level, we can go a long way in increasing their confidence and interest in pursuing graduate studies and CS research. Towards this end, I attempt to engage students in my undergraduate courses in "investigative" measurement and analysis projects by using exciting applications such as web servers, email servers, gaming servers, and streaming media servers. I am continuing to adapt my offering of COMP 431 based on the responses received from previous course offerings.

I also attempt to identify talented students in my courses and encourage them to pursue research projects beyond the course boundaries. I have worked with one such student in the summer of 2008 (supported by REU funds from NSF), and am continuing working with him in developing his research into an honors project. This has been a rewarding research experience and has been instrumental in encouraging this student to apply to grad school.

5 Dissemination

My research group has made the products of our research publicly available to students as well as other researchers, both within as well as outside UNC. Most prominently, this includes TCPdebug, which is a set of OS-sensitive passive network monitoring tools for analyzing real-world TCP traces. These tools have been developed as part of our research on the performance of Internet TCP transfers, and significantly advance the state-of-the-art in passive TCP analysis tools.

The tool-set is available for download at: *http://www.cs.unc.edu/∼jasleen/research/tcp-analysis/*.

We have also freely distributed the protocol implementations, simulation scripts, as well as data-sets developed or used in our evaluation studies.

References

[1] NS (Network Simulator). 1999. http://www-mash.cs.berkeley.edu/ns.

[2] J. Aikat, J. Kaur, D. Smith, and K. Jeffay. Variability in TCP Round-trip Times. In *Proceedings of the ACM SIGCOMM Internet Measurement Conference*, October 2003.

[3] D. Clark, C. Partridge, J.C. Ramming, and J. Wroclawski. A Knowledge Plane for the Internet. In *Proceedings of ACM SIGCOMM*, August 2003.

[4] R. Kumar and J. Kaur. Efficient Beacon Placement for Network Tomography. In *Proceedings of the ACM SIGCOMM Internet Measurement Conference*, October 2004.

[5] R. Kumar and J. Kaur. Practical Beacon Placement for Link Monitoring Using Network Tomography. *IEEE Journal on Selected Areas in Communication*, 2006.

[6] P. Molinero-Fernandez, N. McKeown, and H. Zhang. Is IP Going to Take Over the World (of Communications)? In *Proceedings of the 1st HotNets Workshop, Princeton, NJ*, October 2002.

[7] S. Rewaskar and J. Kaur. Testing the Scalability Limits of Overlay Routing Infrastructures. In *Proceedings of the Fifth Passive and Active Measurement Workshop*, April 2004.

[8] A. Shriram and J. Kaur. Identifying Bottleneck Links Using Distributed End-to-end Available Bandwidth Measurements. *First ISMA Bandwidth Estimation Workshop*, December 2003.

[9] Standford University. http://www.netfpga.org.