Teaching statement

Andreas Haeberlen

http://www.cis.upenn.edu/~ahae/

Scientific inquiry has two important aspects – discovering new knowledge, and helping to spread it – and the opportunity to pursue both is one of the reasons I was drawn to a career in academia.

I am fortunate to be affiliated with Penn’s new undergraduate program in Networked and Social Systems Engineering (NETS). Being a part of NETS has given me the (rare) opportunity to help shape an entirely new curriculum, and to develop courses on emerging topics for which the textbook had yet to be written. I am delighted that Penn was one of the first universities to offer a course in cloud computing (NETS 212), and it was wonderful to see the course double in size each year during its first two years.

My courses have a substantial practical component. I use the lectures to teach the concepts, and I rely on a series of projects to give the students an opportunity to gain practical experience with real systems, such as Amazon’s cloud. The assignments build on each other and lead towards a larger final project, which is completed in small teams.

At Penn, I have had the opportunity to teach four courses: two seminar courses about current research in distributed systems, networking, and security, as well as two regular courses, NETS 212 and CIS 455/555, which I will describe in more detail below. Enrollment in these courses has been growing quickly; both are now over capacity and have long waiting lists.

1 Scalable and Cloud Computing (NETS 212)

NETS 212 is one of the flagship courses of the new NETS curriculum, and Penn was one of the first universities to offer a course of this type. The NETS program focuses on networks and networked interactions, including the interplay of technology, algorithms, economics, and sociology. NETS 212 is primarily about the technical aspects of these networks – the large-scale software systems that are enabling services like Google, Facebook, Amazon, or eBay. However, building a system of this type presents not only technical challenges – such as scaling to millions of users and handling petabytes of data – but also many nontechnical challenges, such as creating incentives for users, preventing manipulation of data, or protecting the privacy of users. NETS 212 is meant to offer a broad view that integrates all of the above challenges; however, when Zachary Ives and I co-developed the course in 2010, there were no textbooks or other courses that offered this perspective, so we had to essentially start from scratch.

The technical topics on the syllabus include the basics of programming at scale, dealing with faults, relaxed consistency models, etc., as well as the MapReduce programming model. We look at several “big data” algorithms, such as PageRank and adsorption, and I present case studies of commercial cloud services. Also, I try to have at least one guest lecturer each semester; for instance, JJ Geewax from Google (a Penn alum) gave a lecture about ad exchanges. The semester ends with a discussion of some advanced topics, such as differential privacy or security in decentralized systems.

To give the students some “hands-on” experience with real cloud services, I assign four projects over the course of the semester that each expose the students to a new cloud-related technology, such as key-value stores, MapReduce, Spark, or Amazon EC2. The students work with Amazon’s live cloud infrastructure; this is made possible by Amazon’s AWS Educate program, which provides each student with some free credit that they can spend on Amazon’s cloud. Other parts of the projects are completed using a virtual machine image I distribute at the beginning of the course. This not only provides the student with the necessary software, such as a Hadoop installation, it also exposes them to virtualization, which is a key enabling technology for the cloud. To maintain student motivation, each assignment results in a simple but usable service (cloud-based image search, geocoding application, SocialRank, and a social network viewer). At the beginning, I provide a lot of support code, but I reduce it step by step as the semester progresses.

1
The semester ends with a final project, which is to build a small Facebook clone in teams of two. The required features include a profile page, a “wall”, and friend recommendations, but I also give extra credit for creativity and special features, and each year there are plenty of opportunities to award it. For instance, I have seen teams implement chat rooms, interactive visualizers that show the user’s social graph, and a Facebook-style timeline; one solution even acquired a small user base!

NETS 212 has proven to be very popular with the students: enrollment went from 11 students in the first year to 31 the year after that, and has grown steadily since then; the enrollment for Fall 2015 is 73, and we even had a waiting list. NETS 212 is also an unusually diverse course; over the years, I have counted students from no less than 20 different majors. The feedback from the course evaluations has been positive; students thought that the course was “very interesting”, or even “extremely fascinating”.

2 Internet and Web Systems (CIS 455/555)

CIS 455/555 was originally developed by Zachary Ives before I took it over in the spring of 2011. The course is open both to undergraduate students (seniors) and to graduate students, and it is now one of the seven core courses in our MSE curriculum.

CIS 455/555 focuses on the issues encountered in building web systems: scalability, interoperability, consistency models, replication, naming and resource location, etc. Although the course is about the web and web services in general, it uses web search as a case study; hence, the students also learn the basics of web search, crawling, distributed storage, fault tolerance, indexing, and ranking. Thus, at the end of the course, the students should be able to understand all the major elements of a search engine like Google – starting with the protocols used to download the search page and to submit requests, and ending with the details of PageRank and defenses against manipulation of rankings. The course ends with some advanced topics, such as accountability or Google’s incremental index updates.

As with NETS 212, the course includes a series of programming assignments that expose the students to various aspects of a web system and that lead towards a large final project at the end. The first assignment is to write a multithreaded web server and servlet container essentially from scratch; this is an opportunity to learn the basics of concurrent programming and client/server systems. The second assignment is to build a web crawler and a pub/sub system, and the third assignment is to build a decentralized web cache based on a distributed hashtable. The final project then is to integrate the solutions from the first three assignments and to build a small web search engine – not unlike Google in its early days – in teams of four. For several years now, Google has sponsored a special award (cell phones or tablets) for the best project. Choosing a team for the award has been difficult because, in every year, there have been several teams that went far beyond the project requirements. The additional features I have seen include, for instance, page previews, a “safe search” feature, and a special mode for the visually challenged that allows the user to control the search entirely with spoken commands.

In the four years I have taught the course, enrollment has more than doubled, from 39 students in 2010 to 100 students in 2016 (with some additional students on the waiting list). I have worked hard to improve the course, and I was delighted to see that the ratings have climbed steadily every year, from a 3.17 instructor rating in the first year to a 3.61 rating in 2014; course quality has gone up similarly, from 3.14 to 3.57. (Both are on a scale from 0 to 4.) The feedback that the students provided on their evaluations has been overwhelmingly positive and has included comments such as “This course was awesome!” or “One of the best courses I’ve ever taken".