CSci 599: Internet Measurement
Spring 2012

Basic Information

Place and time: TBA, Wed 3:30pm-6:20pm

Instructor: Ethan Katz-Bassett
email: ethan.kb@usc.edu
phone: --
office: SAL 236
office hours: Wed 5:00pm-6:00pm

Course homepage: TBA

Prerequisites: CS551 of permission of the instructor. This class is appropriate for graduate students or advanced undergraduates with previous classwork in networking. Students from non-systems/networking areas are welcome.

Course Description

The Internet now plays a central role in many aspects of our lives. Despite the myriad ways we have come to depend on it, many aspects of it can be opaque even to network operators. Internet measurement as a field seeks to understand the Internet by assessing its operation. The field is interesting for a range of reasons:

- We depend on the Internet, so we need to understand its operation, and we need to discern its problems in order to improve it. Engineers at Google and Facebook cite the difficulty of Internet measurement as one of the impediments they face.
- The Internet is a loose federation of networks that must cooperate to provide global connectivity, even as they compete for business. Some of the opacity arises from this tension, as networks may lack incentive to expose their inner workings and may lack the visibility necessary to optimize their performance.
- The Internet is one of the largest systems humankind has ever engineered, and it has emergent properties. Measurement provides a basis for determining these properties.
- The Internet protocols were designed in an era when networks looked very different than they do today. They do not naturally expose some of the information that we want to learn. Some of the fun of Internet measurement comes from finding novel ways to manipulate the protocols to reveal information.

In this course, we will investigate important problems, techniques, results, and challenges from the field. We will explore both what measurements tell us about the Internet and how we can leverage what they tell us to improve systems, including peer-to-peer file sharing and Google. We will focus on why certain questions are hard to answer, how we might start to answer them, and why different measurements might reveal what seem to be conflicting answers. We will learn to measure various aspects of the Internet, including topology, routing and routing policies, performance, failures, traffic, and applications. Researchers often talk about Internet

measurement as being analogous to astronomy, in that we take observations from afar in order to understand how a system works. We will learn to leverage and integrate the various sources of information that leak out from services about their internal operations.

Textbooks and Course Materials

We will read 2-4 research papers a week.
Recommended supplemental textbook: Crovella and Krishnamurthy's textbook Internet Measurement: Infrastructure, Traffic & Applications.

Grading

There are no exams in this class.

The course grade will be determined based on:

- Written paper responses and class presentations/discussion (50%): Students are expected to write responses to 2-4 papers a week. Each week, one or two students will give a presentation on the papers and lead the discussions in each class. Other students are expected to participate in the discussion.
- A research project, including ~6 pg writeup and ~20 minute presentation (50%): The semester-long project is an open-ended Internet measurement project. The instructor will provide some possible project topics, or (with instructor approval) you can work on a project of your own devising. Most projects will likely involve either making new measurements or performing new analysis of existing data. Projects should be done in groups of two.

Schedule by Week

- Week 1: Course overview. Why measure the Internet? How to conduct measurements?
    - Strategies for Sound Internet Measurement. Vern Paxson, IMC 2004.
    - Using PlanetLab for Network Research: Myths, Realities, and Best Practices. Neil Spring, Larry Peterson, Andy Bavier, and Vivek Pai. ACM SIGOPS Review, January 2006.
    - Reverse Engineering the Internet. Neil Spring, David Wetherall, and Thomas Anderson. HotNets 2003.
    - Tussle in Cyberspace: Defining Tomorrow's Internet. David D. Clark, John Wroclawski, Karen R. Sollins, Robert Braden. SIGCOMM 2002.
- Weeks 2-3: Internet topology
    - What is the Internet's topology?
        - Heuristics for Internet map discovery. R. Govindan and H. Tangmunarunkit. IEEE INFOCOM 2000.
        - On power-law relationships of the Internet topology. Faloutsos, Faloutsos, and Faloutsos. SIGCOMM 1999.
        - Sampling Biases in IP Topology Measurements. Anukool Lakhina, John Byers, Mark Crovella, and Peng Xie. INFOCOM 2003.

        - What are our standards for validation of measurement-based networking research? Balachander Krishnamurthy and Walter Willinger. HotMetrics 2008.
        - On inferring autonomous system relationships in the Internet. L. Gao. In IEEE Global Internet Symposium, 2000.
    - Missing links
        - The (in)completeness of the observed Internet AS-level structure. Ricardo Oliveira, Dan Pei, Walter Willinger, Beichuan Zhang, Lixia Zhan. IEEE/ ACM ToN 2010.
        - Bigfoot, Sasquatch, the Yeti and Other Missing Links: What We Don't Know About the AS Graph. Matthew Roughan, Jonathan Tuke, and Olaf Maennel. IMC 2008.
        - Where the Sidewalk Ends: Extending the Internet AS Graph Using Traceroutes From P2P Users. Kai Chen, David R. Choffnes, Rahul Potharaju, Yan Chen, Fabián E. Bustamante, Dan Pei, Yao Zhao. CoNEXT 2009.
        - IXPs: Mapped? Brice Augustin, Balachander Krishnamurthy, Walter Willinger. IMC 2009.
- Week 4-8: Internet routing and performance
    - Traceroute and Its Limitations
        - How to Accurately Interpret Traceroute Results. Richard Steenbergen. NANOG 47, 2009.
        - Measuring multipath routing in the Internet. Brice Augustin, Timur Friedman, and Renata Teixeira. IEEE/ACM ToN, June 2011.
        - Reverse traceroute. Ethan Katz-Bassett, Harsha V. Madhyastha, Vijay K. Adhikari, Colin Scott, Justine Sherry, Peter van Wesep, Tom Anderson, Arvind Krishamurthy. NSDI 2010.
        - Internet-Scale IPv4 Alias Resolution with MIDAR. K. Keys, Y. Hyun, M. Luckie, and k. claffy, IEEE/ACM ToN, 2012.
    - Understanding Routes
        - Quantifying the Causes of Path Inflation. Neil Spring, Ratul Mahajan, Thomas Anderson. SIGCOMM 2003.
        - Moving Beyond End-to-End Path Information to Optimize CDN Performance. Rupa Krishnan, Harsha V. Madhyastha, Sridhar Srinivasan, Sushant Jain, Arvind Krishnamurthy, Thomas Anderson, and Jie Gao. IMC 2009.
    - Path changes
        - Predicting and Tracking Internet Path Changes. Í. Cunha, R. Teixeira, D. Veitch, and C. Diot. SIGCOMM 2011.
        - Delayed Internet Routing Convergence. Craig Labovitz, Abha Ahuja, Abhijit Bose, and Farnam Jahanian. SIGCOMM 2000.
        - A Measurement Study on the Impact of Routing Events on End-to-End Internet Path Performance. Feng Wang, Zhuoqing Morley Mao, Jia Wang, Lixin Gao, and Randy Bush. SIGCOMM 2006.

    - Outages and anomalies
        - End-to-end routing behavior on the Internet. Vern Paxson. IEEE/ACM ToN, October 1997.
        - End-to-End Internet Packet Dynamics. Vern Paxson. IEEE/ACM ToN, June 1999.
        - User-level Internet Path Diagnosis. Ratul Mahajan, Neil Spring, David Wetherall, and Thomas Anderson. SOSP 2003.
        - Internet Optometry: Assessing the Broken Glasses in Internet Reachability. Randy Bush, Olaf Maennel, Matthew Roughan, and Steve Uhlig. IMC 2009.
    - Tomography
        - Server-based Inference of Internet Link Lossiness. Venkata N. Padmanabhan, Lili Qiu, and Helen J. Wang. INFOCOM 2003.
        - Active Measurement for Multiple Link Failures Diagnosis in IP Networks. Hung X. Nguyen and Patrick Thiran. PAM 2004
        - Improving Accuracy in End-to-End Packet Loss Measurement. J. Sommers, P. Barford, N. Duffield and A. Ron. SIGCOMM 2005.
        - Shifting Network Tomography Toward a Practical Goal. Denisa Ghita, Can Caracus, Katerina Argyraki, and Patrick Thiran. CoNEXT 2011.
- Week 9-11: Internet traffic
    - Backbone traffic
        - Internet inter-domain traffic. Craig Labovitz (Arbor Networks), Scott Iekel-Johnson (Arbor Networks), Danny McPherson (Arbor Networks), Jon Oberheide (University of Michigan), Farnam Jahanian (University of Michigan). SIGCOMM 2010.
        - Anatomy of a Large European IXP. SIGCOMM 2012.
        - Packet-Level Traffic Measurements from the Sprint IP Backbone. Chuck Fraleigh, Sue Moon, Bryan Lyles, Chase Cotton, Mujahid Khan, Deb Moll, Rob Rockell, and Ted Seely, Christophe Diot. IEEE Network November/ December 2003.
    - Edge traffic
        - 2 of 3 papers on measuring edge networks: Broadband Internet Performance: A View From the Gateway, Characterizing Residential Broadband Networks, or Netalyzer
        - 2 of 5 papers on traffic differentiation: NetPolice, Glasnost, DiffProbe, ShaperProbe, or Nano (Detecting Network Neutrality Violations with Causal Inference)
    - Anomaly detection
        - A Framework for Classifying Denial of Service Attacks. Alefiya Hussain, John Heidemann, Christos Papadopoulos. SIGCOMM 2003
        - Diagnosing Network-wide Traffic Anomalies. Anukool Lakhina, Mark Crovella and Christophe Diot. SIGCOMM 2004.
        - ASTUTE: Detecting a Different Class of Traffic Anomalies. Fernando Silveira, Christophe Diot, Nina Taft, Ramesh Govindan. SIGCOMM 2010.

- Week 12-14: Internet applications
    - P2P file sharing
        - Taming the Torrent: A practical approach to reducing cross-ISP traffic in P2P systems. David R. Choffnes and Fabián E. Bustamante. SIGCOMM 2008.
        - Pitfalls for ISP-friendly P2P design. Michael Piatek, Harsha V. Madhyastha, John P. John, Arvind Krishnamurthy, Thomas Anderson. HotNets 2009.
        - Pitfalls for Testbed Evaluations of Internet Systems. David R. Choffnes and Fabián E. Bustamante. SIGCOMM CCR, April, 2010.
    - Spam, botnets, and the underground economy
        - Click Trajectories: End-to-End Analysis of the Spam Value Chain. Kirill Levchenko, Andreas Pitsillidis, Neha Chachra, Brandon Enright, Márk Félegyházi, Chris Grier, Tristan Halvorson, Chris Kanich, Christian Kreibich, He Liu, Damon McCoy, Nicholas Weaver, Vern Paxson, Geoffrey M. Voelker, and Stefan Savage, IEEE Symposium and Security and Privacy Oakland 2011.
        - Understanding the Network-Level Behavior of Spammers. A. Ramachandran and N. Feamster. SIGCOMM 2006.
        - Studying Spamming Botnets Using Botlab. John P. John, Alexander Moshchuk, Steven D. Gribble, and Arvind Krishnamurthy. NSDI 2009.
    - Topic TBD: Online social networks, mobile, or DNS.

- Week 15: Project presentations

Statement for Students with Disabilities

Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776.

Statement on Academic Integrity

USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one's own academic work from misuse by others as well as to avoid using another's work as one's own. All students are expected to understand and abide by these principles. Scampus, the Student Guidebook, contains the Student Conduct Code in Section 11.00, while the recommended sanctions are located in Appendix A: http://www.usc.edu/dept/publications/SCAMPUS/gov/

Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty. The Review process can be found at:
http://www.usc.edu/student-affairs/SJACS/

Emergency Preparedness/Course Continuity in a Crisis

In case of a declared emergency if travel to campus is not feasible, USC executive leadership will announce an electronic way for instructors to teach students in their residence halls or homes using a combination of Blackboard, teleconferencing, and other technologies.