

# Advanced Topics in Computer Systems: Internet Reverse Engineering

Neil Spring

Fall 2005

**Instructor** Neil Spring

**E-mail** nspring@cs (include "818" in your subject line)

**Phone** x52909

**Class** TuTh 3:30-4:45pm

**Office hours** TuTh 2:30-3:30, and by appointment AVW 4133

**Web** <http://www.cs.umd.edu/class/fall2005/cmsc818s>

NOTE: Class will not be held on 9/27, and 11/15, and may not be held 9/29, 10/25, 10/27. I have yet to learn to say "no" when invited to give my opinion at workshops and on panels at conferences in France. I will have assignments for you to complete with your project group on days when class does not meet. Subscribe to the ical [webcal://ical.mac.com/neil.spring/CMSC%20818S.ics](http://ical.mac.com/neil.spring/CMSC%20818S.ics) or <http://ical.mac.com/neil.spring/CMSC%20818S>.

## 1 Goals of this course

The primary goal of this course is to explore emerging research areas in networking, specifically in network measurement, and currently hot topics linking networking with artificial intelligence, databases, programming languages, etc. The secondary goal is to learn about the true (non-textbook) operation of the Internet and to learn how to use tools to discover more on your own.

## 2 Summary

The Internet is a complex and unintelligible artifact, outgrowing our attempts to understand and manage it. Government would like assurances that the network is not vulnerable; enterprises and users would like assurances that the service they pay for is what they are delivered; service providers would like assurances that their neighbors behave correctly.

In this course, we will examine the breadth of research in network measurement and in collaborative data storage. The course will focus on four parts:

1. New tools for studying the network from its edges: RPT, Spruce, pathchar; and challenges for the future. These tools are often clever at the core but difficult to get right.
2. Platforms for studying the network: "telescopes", surveyor, nimi, Scriptroute, PlanetLab; and the special problems of hosting "measurement" as a service. Measurement platforms are challenging systems involving mobile code and wide-area deployment.

3. Analyses of measurement data: Rocketfuel, policy inference, failure detection, overlay scheduling; and how to make the best guesses about features we cannot measure. Internet measurements provide many large graphs that lack structure.
4. Storage, coordination, and scheduling of collaborative data collection. There's more data than we know how to manage. And sometimes people think it's private.

### 3 Prerequisites

You should understand basic networking concepts of packet switching, addressing, retransmission, routing, etc. Grab an undergraduate text and read it over the weekend if you aren't yet familiar.

### 4 Grading

#### 4.1 Participation: 20%

Speak in class. Raise questions from your reading of papers, especially thought-provoking "would X be a good idea?" or "has Y been done before?" style questions. Be critical. Tell me I'm wrong. Shape the course. If you share your experience and outlook with the class enough that I can describe who you are at the end, you'll be fine.

#### 4.2 Presentation: 20%

Your presentation of material (one or two lectures) will also factor. I expect a discussion-dominated presentation: everyone should read the main paper being presented (but be perhaps unfamiliar with the background). Your task will be to show hard-to-understand graphs and results, explain methodology, put the work in context, raise criticisms, and summarize lessons, *not* to defend the work or present it as if at a conference.

#### 4.3 Exam: 20%

Two in-class exams will test your understanding of course material. The first will (likely) cover the vocabulary of background material covered in the first few weeks. The second (final) will consist of an open-ended two page written evaluation of proposed research.

#### 4.4 Project: 40%

Individual or pair projects in network measurement as components of reverse engineering. Topics that might get you started include:

1. Using measurements to prove the absence of connectivity. Determine which links don't exist, and what coverage is needed to "prove" they aren't present (or are unused backups).
2. Internet mapping with DNS: reverse-lookup all IP addresses, then infer connectivity between routers using address allocation policies and host names, verify (spot-check).
3. Exploit NTP (network time protocol) for one-way measurement. Contact pool.ntp.org servers to collect precise round-trip-time measurements, measuring jitter, delay, etc.
4. Mobile Cyclone objects for measurement execution (replace Ruby in Scriptroute to more easily port existing C-based tools).
5. Abstract measurement specification (measure these things about these links). Only attempt if you have insight.

6. PlanetLab mapping with Doubletree and incremental update. Generate a service to replace Stribling's all-pairs ping.
7. Uncertainty of link-measurements (e.g., clink) as a function of distance, latency, and other properties. How predictable is the "badness" of a performance (bandwidth) measurement from more easily measured properties?
8. Generic OWAMP receiver interface for Scriptroute. What are the right high-level primitives for accessing this facility?
9. In-kernel packet train scheduling (extended PeriScope). What are the limits from user space? and what are the limits if packet probe scheduling is managed by the kernel?
10. A project in education, for example, to help students understand BGP, TCP, OSPF, or BitTorrent through measurement.
11. A project of your design – bridging your research interests with network measurement: conducting a measurement to prepare for a new protocol you're interested in, for example.
12. A project inspired by a reading in the class.

The project will be evaluated based on a presentation and paper at the end, but also on intermediate progress points, which will likely consist of choosing and understanding a topic by the third week, initial results or insight by the fifth, and a draft report by the tenth.

Written material at the graduate level should never include 12 point (11 point or smaller is okay), helvetica (times or similar is okay), or easily detected spelling errors. A warning: I have little tolerance for unsupported assertions (e.g., "The widget may be useful for many applications" without defining the applications or citing another paper) or verbose phrasing ("Note that there is the potential that a large number of applications will be able to leverage the widget," which could easily be rephrased to the last example). Find a copy of Strunk and White or read it on-line.

## 5 Lateness

All assignments can be turned in electronically. I will permit one project milestone to be turned in after the weekend (when due Friday, it can be turned in on Monday).

There are exceptions for excused absences defined by the university. If you're sick, get a note. If you need accommodation for disability, illness, family emergency, etc., ask me as early as you can. Finally, remind me of religious observances if they interfere with the class.