CS 485/ECE 440/CS 585 Lab 1

Due by 11:59pm on Thursday, 30 September, as an e-mail to the instructor (jedcrandall@gmail.com). Please send only PDF files.

The purpose of lab 1 is to make you familiar with the basic workings of ARP, IP addressing, routing, NAT, TCP, *etc.*, and to consider some of the issues that arise when tying addresses to hosts on the Internet. The basic setup is that I've allegedly committed some crime on your second virtual machine by downloading a copyrighted image, and then allegedly deleted the main piece of evidence: the file itself. Your first virtual machine, which provided NAT and routing for the second, has a tcpdump file in the root directory. You'll also be given a more complete tcpdump for the server the file was downloaded from.

They say that a man who represents himself in court has a fool for a client and a fool for a lawyer, so I'm not only representing myself in my defense but also hiring a fool (myself) as a technical consultant. Your group has been hired as technical consultants for the prosecution. Your job is to make a detailed case that shows exactly what I downloaded, beyond a shadow of a doubt. You should explain the technical details thoroughly, since the judge (me) will allow the defense to hire an independent networking expert (also me) to tell the jury (twelve of the voices in my head) whether your case that the evidence proves I downloaded the particular file has any merit. If you leave any loose ends, I'll use them to get my client off. Your job is to make sure your case is so solid that my client does hard time.

I expect to see lots of annotated tcpdumps, technical details from each of the different layers (ARP, IP, TCP, HTTP), and a packet-by-packet reconstruction of everything that happened. You'll need to do a lot of work before you can even start the writing, so get started early.

The defense has seen the evidence that you have and has filed an appeal to not let either tcpdump be allowed in court. As technical consultants for the prosecution you'll need to both rebut the appeal in a point-by-point fashion and also give a detailed reconstruction of the download of the file. Since I allegedly deleted the file I allegedly downloaded, you'll need to reconstruct the file somehow from the tcpdumps, if possible. If that's not possible, you'll need to explain why not. You may use existing tools or write your own code to analyze the tcpdump files and reconstruct the flow, but every member of your group will need to understand how the file was reconstructed when the time comes for cross examination.

The same is true of all technical details, each member of the group needs to be able explain any of the technical details of your case when I do my oral cross examination.

You'll write your case up as a writeup, with lots of figures (annotated tcpdumps, graphs, explanatory figures, and such). There is no minimum or maximum length, but, as a rule of thumb, to make a strong case you'll need at least five pages of text (without any figures) and about twice as much figures as text,

so you should shoot for a minimum of 15 pages total.

The rubric you'll be graded on, out of 100 points, is as follows:

Undergrads (CS 485 or ECE 440):

- 15 points presentation (English grammar, presentation quality, easy to understand, etc.)
- 40 points technical strength of your case
- 20 points for the completeness of your implementation of tools to analyze the evidence
- 25 points cross examination

Grads (CS 585)

- 35 points presentation (English grammar, presentation quality, easy to understand, etc.)
- 40 points technical strength of your case
- 25 points cross examination

The presentation will be given one grade that will be scaled accordingly for the different group members. Writing is the responsibility of the grad student, but the undergrad slaves are expected to help with the content and proofreading, and help with pieces of the writing as needed. Points for technical strength of your case will be based on how well you addressed each of the points in my appeal not to allow the evidence, how much detail and rigor you put into a packet-by-packet reconstruction of exactly what happened, and how much knowledge of the various layers (physical, link, routing, transport, and application) you incorporate into your argument.

Completeness of the implementation doesn't preclude you from using existing tools, but if there's some way that your case could have been made stronger that requires implementation (such as reconstructing the file) and you didn't do it, I'll dock the undergrads points in this part of the rubric.

Cross examination is an individual score, *i.e.*, it will not be the same score for the whole group but will be a score I assign to each individual student after I cross-examine them. So make sure that you understand everything that your group puts in the report. I'll be asking you detailed questions, like, “How did your group match the packets between the two TCP dumps?” or “What did you find out about how port numbers are rewritten in NAT?”

You may use Google as much as you want and you may discuss the lab at a high level with other groups. You can even share code with other class members. The work you present in your report must be your own work, however.

Appeal to not allow tcpdump evidence

The defense requests that the tcpdump evidence not be allowed for the following reasons:

1. The IP address of the system that the defendant allegedly downloaded the file onto does not appear in the server's tcpdump. If address translation had occurred then the source port numbers would not match, so it's highly suspicious that they do match.
2. None of the timestamps match exactly. Furthermore, there is evidence of packet delay and TCP retransmissions. Since shasta (that hosts the virtual machine the defendant allegedly used) and the alleged server the file was downloaded from are on the same subnet there should have been no packet delay or TCP retransmission.
3. Even if a file can be reconstructed, if the tcpdump closest to where the defendant was allegedly downloading the file does not have both IP and TCP checksums that match those of the tcpdump the reconstruction was performed on, then the reconstruction of the file is meaningless.
4. There is nothing in the tcpdump files that links the machine the download allegedly occurred on with the suspected IP address on the local network. So, even if that IP address can be linked to the crime, the machine itself cannot be linked to that IP address since anybody on that network could have used the IP address during that time.
5. Similarly, since the timestamps are suspect, there's no evidence linking the name of the web server to the IP address that the server was allegedly using when the download allegedly occurred.
6. None of the ARP traffic from either tcpdump file appears in the other, which is highly suspect since shasta and the alleged server are on the same virtual LAN.
7. That a server's name appears in the tcpdump is meaningless, since tcpdump resolves names at the time you run the tool to look at the dump, not at the time the tcpdump was created. I contend that if the prosecution's case uses any tcpdump analyses without the “-nnn” option then the case should be thrown out.
8. The router (the first virtual machine) does not have open the source port that was allegedly used, so could not have received the return packets from the server on that port.
9. Even if a TCP/IP connection between two machines can be implicated in a crime, it still can't be tied to a particular socket or a particular user, not even if that user was the only one logged into the machine at the time.

These should just get you started, you'll need to go well beyond just addressing the above points. There is a lot you can do (*e.g.*, run your own experiments for connections to the server in question and make a histogram of packet delays with an explanation of where the delays come from and how they compare to your evidence, *etc.*). For every hour and 15 minutes in class you should be spending another 2 or 3 hours outside of class working on this, if you're not sure what to do next email me. You'll be expected to work hard on this lab from the day it is assigned until the day it is due, if you feel like you're done or that you understand the assignment fully and think it should be easy enough to get done, then you're probably missing something important. If you need to prove that a certain thing works a certain way, the best way to do it is to set up an experiment. I allegedly deleted the images for each group from the

server, but there's an image called image.jpg there that you can request for experimental purposes. I have also set it up so that you can run tcpdump on shasta's interface to the vlan it shares with the alleged server. You'll be using sudo to do so and your tcpdump process will execute as root, any attempt to abuse this privilege will be a violation of university policy and will be dealt with accordingly so only use it to record tcpdumps and not to try to hack shasta or sniff any traffic unrelated to your experiment.