Proposal: Information Flow on the Internet

Jeremy Foote
jdfoote@u.northwestern.edu

Overall Goal

The goal of this session will be for students to gain an intuitive understanding of some of the technologies that underlie the Internet, such as packets, packet routing, encryption, VPN, Tor, man-in-the-middle attacks, and government surveillance. Each of these technologies will be explained, and then enacted through an in-class packet routing exercise.

This proposal is based on feedback from the previous class that I taught. I have tried to make the discussion much more about the technical aspects of the Internet rather than the social implications.

IP Addresses and Packet Routing - 30 minutes

We will talk a little bit about the history of the Internet - how and why computer internetworking started.

You are the Internet Activity

At this point, I will introduce an activity adapted from Aaron Shaw's class. I will have a few giant nametags to hang around people's necks: "Instagram", "Facebook", "User A", "User B", and "User C". Each of the "users" will be given a picture, a pair of scissors, a pen, and a stack of envelopes, and told that their goal is to send a picture to "Instagram" or "Facebook", with some constraints:

1. No talking
2. Only a 1 inch x 1 inch square can fit in the envelope (no bending)

3. They have to pass the envelopes to a neighbor

And a few rules for the "routers" (everyone else):

1. Can only have 1 envelope at a time
2. Must pass the envelope to a neighbor

After one round of this, we'll go in the reverse direction. I'll ask, "How would User B see User A's photos?" We'll talk about sending requests, and the need to make a copy of the photo in order to send it, and Instagram will pass the photo pieces back to the users. For the next round, we'll add a few complications. I will pass a note with instructions to a few nodes: one to pass envelopes the wrong direction, and one to keep envelopes and not pass them on.

Discussion Questions

1. What information did routers need? (Address, best next hop)
2. How did you deal with congestion? (Routing around slow areas)
3. How did you deal with broken or malicious nodes? (Routing around them)
4. Who was the leader in charge of making changes to the path? (No one - just local decisions)
5. What are the problems with this approach? (packets can be lost)
6. How could we solve it so that packets aren't lost? (TCP)

During the discussion, I will present a few slides about how they were passing around "packets", which had address "headers". We will discuss how IP addresses are assigned and a very brief discussion about how routers make forwarding decisions. We will also talk a little bit about the power of a decentralized, end-to-end, "dumb" network.

TCP - 20 minutes

Hopefully, the students can help me to design TCP at this point. :)

I will break the class into small groups to discuss how to make sure that packets aren't lost along the way. They can present their solutions, and we'll talk about how

TCP solves this problem: namely, by numbering each packet, and having the server respond when packets are received.

Server-side Security - 20 mins

What if User A and User B aren't friends on Instagram, but User A sends a request to see User B's photos, how does Instagram know whether or not to show them?

How could we use this system for User A to prove they are who they say they are?

Slides

Show slides explaining authentication, authorization, and cookies. These all happen on the server - the user sends a password, which proves they are who they say they are. The server passes back a session ID cookie, which is then sent in all future requests, proving that they came from the user.

Discussion

There are still some vulnerabilities in our system. What are they? (Intermediate nodes could intercept images or session cookies, could impersonate the user.)

Network Security - 30 mins

The big danger of the Internet is that all of the data is passed through intermediaries who may or may not be trusted. Let's run the simulation again, with a few new nodes.

Now, I will give one student a piece of paper with the assignment to be an "imposter". When image packets come through, they will substitute them with a different, embarrassing photo. Another student will be "The NSA". Their assignment will be to make a note about everything that comes through (Ideally, they'd make a "copy", but I can't figure out how to do this)

Discussion

1. How could you stop these sorts of attacks in our system? Is it possible?
2. I will bring in a few small boxes with locks and ask: How could locks and boxes be used to send messages privately? Is there a way to know who sent each message, and that no one saw it along the way?

Slides

Public Key Encryption and HTTPS - allows for encryption of data - including the session cookie, so that User A can prove that they are User A, and so that no middlemen can intercept or copy the data that is being sent.

There are a few good videos about how this works on YouTube, that I could show.

I'll describe how the basic idea is that Facebook has an empty box with an open lock on it. User A can put a key and lock in that box and lock it, knowing that only Facebook can open it. Then, Facebook sends the box back, this time locked with User A's lock. Now, Facebook and User A each have a key to that lock, and they can send messages back and forth.

So, now User A knows that Facebook is Facebook, and no one is looking at her messages, but how can Facebook know that User A is really User A? Passwords - User A sends their password, locked in the box.

Proxies, VPN, and Tor - 20 mins

We will talk about another limitation of TCP/IP. The NSA student will now be a despotic government, who wants to know everyone who visits an anti-government website. We will talk about how even encrypted information still contains an address, which lets men-in-the-middle to know who has been talking to whom.

Discussion

1. How is that a problem?
2. What are some potential solutions?

Slides

- Proxies - Deal between you and a node on the network to change the address on the envelope to their address, and to forward everything to you.
- VPN - Similar, but traffic gets encrypted before sent to the proxy
- Tor - A bunch of proxies in the middle, so that nodes don't know where message originated. Locked boxes in locked boxes.

We will then run things one more time - and show how an image is sent using SSL and Tor.

Discussion

What are the implications of having an anonymous internet? Where is there still a weakness? (FB or Instagram has the unencrypted data)

Philosophical Musings - If there's time (there won't be)

I'd love to talk about how the design decisions of the Internet were made to solve the small, initial problems of the network, but they have had incredible ramifications. Perhaps we could end with discussion about what sorts of technological decisions being made now might have similar long-term consequences (e.g., self-driving car regulations, gene manipulation, etc.)