Free Speech on the Internet

CS @CU

Free Speech Online

- Does the First Amendment apply to online speech? Of course.
  - Is a blogger “the press”? Currently irrelevant under the First Amendment; might matter for some states’ “shield” laws.
- What about unprotected speech? (Obscenity, true threats, libel, etc.)
  - The desire to deal with online obscenity motivated the Communications Decency Act of 1996 (CDA)
- Who enforces any such limits?
- How do they enforce it?

Basics

- Technical environment in 1996
- Corporate influence
- Filtering technologies

Technical Environment, 1996

Architectural Issues

- As noted, architecturally all Internet nodes are equal
- In theory, I can run a web server and mail server in my apartment, just like Google's
- "It is no exaggeration to conclude that the Internet has achieved, and continues to achieve, the most participatory marketplace of mass speech that this country—and indeed the world—has yet seen." (ACLU v. Reno, 929 F. Supp 824, E.D. Penn., 1996)
- That's theory…

The Internet in 1996

- The (popular) Worldwide Web had only been around for ~2 years
  - Windows 3.1, which did not include standard TCP/IP, was still pretty common
  - Many people used AOL's proprietary service, which had a gateway to the Web
- Most people used dial-up modems, generally at 33.6K bps maximum
  - Cable modems and DSL were just starting to be deployed
- There were (effectively) no consumer digital cameras, and hence no consumer JPGs
- Software MP3 players were first available in 1995, i.e., there was little (compressed) digital music

Search Engines, 1996

- AltaVista, 1995
  - The first true search engine
  - Wildly popular for a while
  - Eventually displaced by Google, which gave more useful answers
- Yahoo!
  - Manual classification of web sites
  - More like a library card catalog
  - Couldn’t cope with the growth of the web

Mike Less: “The last battle between Melvil Dewey and Vannevar Bush”

A “Card Catalog”?

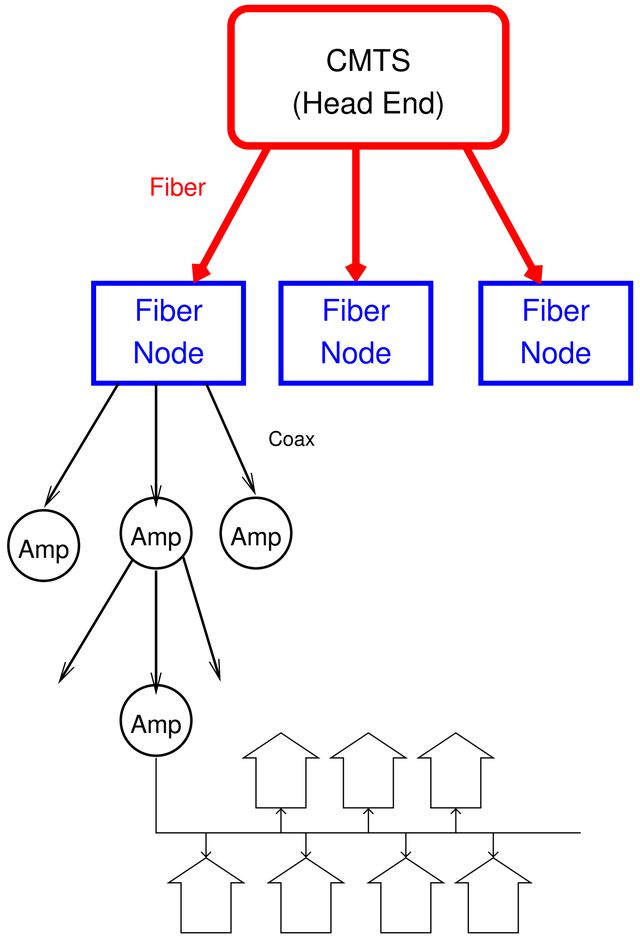
Butler Library, February 2018

You Can’t Run a Server via Dial-up

- You don’t have a stable IP address
- You’re not always online
- You don’t have enough bandwidth

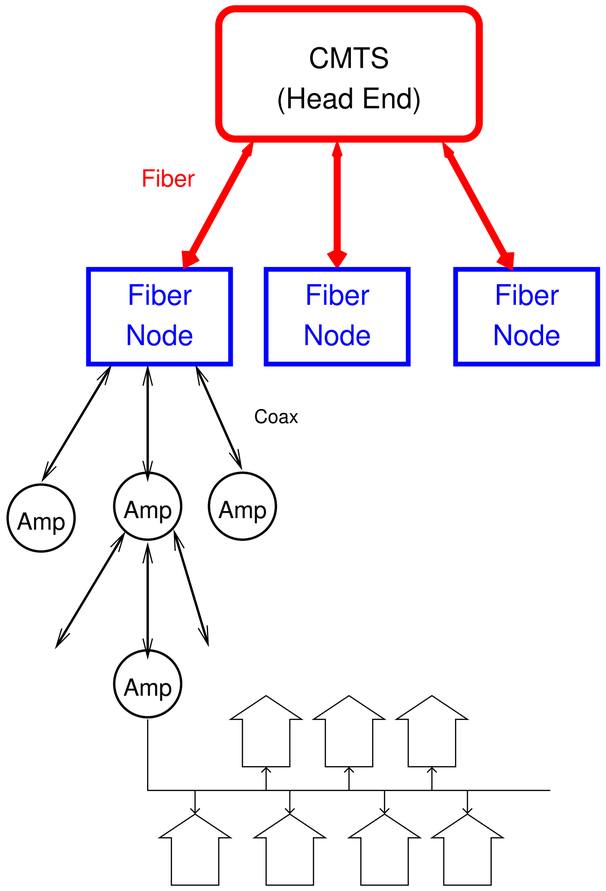
The Cable TV Plant

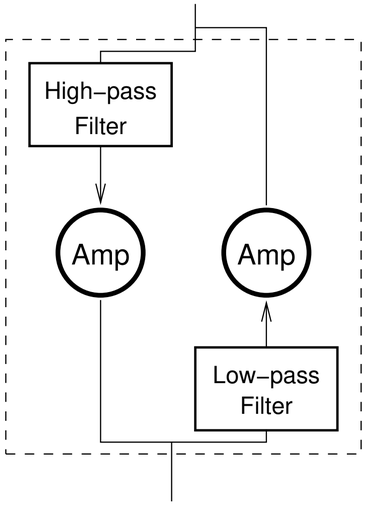
The cable TV infrastructure was designed for one-way broadcasts



Cable ISPs

- Every amplifier has a “high-pass filter”, to send TV channels downstream, and a “low-pass filter”, to pass frequencies *below channel 1* upstream
- You can use many TV channels to send Internet text downstream
- You have only one channel going upstream
- *The design inherently has asymmetric bandwidth*
- The only way to increase upstream bandwidth is to replace amps with (expensive) fiber nodes





Consumer Computers aren't Equal

- They don't have the bandwidth to run servers
- They don't have stable IP addresses
  - Some cable/DSL/fiber ISPs *deliberately* change IP addresses periodically, to make it hard technically
- Some ISPs even block standard server ports, to conserve scarce upstream bandwidth
- (Direct outbound email from residential computers is generally blocked, too, because of spam)
- Conclusion: consumers must rely on corporate nodes to speak

Corporate Influence

Corporate Players

- ISPs—primarily cable TV providers or phone companies
- The tech giants: Google, Facebook, etc.
- Content distribution networks
- Cloud providers
- Hosting services

ISPs

- Local Internet service is (at least close to) a natural monopoly
  - Running cable or fiber is *very* expensive; profitability depends on the subscriber rate
- Very few consumers have a choice of providers
  - (Regulations implementing competition-providing provisions of the Telecommunications Act of 1996 were struck down, 290 F.3d 415, 2002)
- There has been consolidation in the telecom/ISP industry: fewer, larger providers

The Tech Giants

- Search engines are vital for finding information
  - Without them, you'd have to rely on influencer sites and the like
- Facebook, Twitter, etc: the "network effect"
- Metcalfe's Law: "The value of a network is proportional to the square of the number of nodes"
- Translation: because Facebook has grown, it will grow even larger

Content Distribution Networks

- There is too much load for any single web site to serve large content to the world
- Solution: *content distribution networks* (CDNs), which replicate content on modes closer to end-users
- Unless you're the size of Google, Apple, Facebook, etc., you can't be a busy site without CDNs

Cloud and Hosting Services

- Bandwidth costs money, especially if you want links to multiple providers
- Big data centers are expensive, and need power, cooling, security, etc.
- Solution: cloud and hosting providers, e.g., Amazon Web Services
- Plus: these companies offer extra services, worry about network connectivity, and more

The Corporate Gatekeepers

If these large providers want you offline, you're functionally offline

- You need a CDN for load
- You need a data center or a cloud provider
- If you're not indexed by search engines, few people will find your content
- If you're not on a popular social network, few will hear your voice

Whither the First Amendment?

The Facebook Oversight Board

- To deal with complaints about some of its decisions, Facebook created an external oversight board
- In theory, it can even overrule Zuckerberg
- But…
  - Are they looking at the right questions?
  - To whom are they accountable?
- Is this the right model?

Filtering Technologies

Where to Filter?

- Assume that corporate gatekeepers wish to filter content. How?
- It’s hard at CDNs—they don’t have full context
  - It’s even worse for ISPs
- For cloud and hosting services, it’s all-or-nothing
- Can Big Tech filter?

Implementing Filtering

- Manual
  - Not feasible—too much content
  - In Stratton Oakmont v. Prodigy (23 Media L Rep 1794), 1995, Prodigy said that it was receiving 60,000 new posts per day
  - Imagine today's load…
- Reactive, i.e., after abuse reports
  - Subject to abuse by partisan groups
- Can automatic filtering work?

Automatic Filtering

The only vaguely feasible approach uses machine learning—can it work?

- Can machine learning understand language well enough?
  - What about sarcasm?
  - Broader context?
  - Slang, changing language, online conventions?
- What about video, e.g., YouTube?
- What about voice, e.g., Clubhouse?

Standards

What standards should they filter to?

- Illegal content? That's often subtle or controversial
  - It's also locale-dependent, e.g., United States v. Thomas (74 F.3d 701), 1994 (obscenity), or Nazi-related material in many countries
  - Anti-government material in some countries?
- Copyright infringement? What about fair use, license, jurisdiction, etc.?
- Internal standards

Standards Change

- Twitter, 2012: “The free speech wing of the free speech party” (but also see this)
- Twitter, 2018: “We can only serve the public conversation, we can only stand for freedom of expression if people feel safe to express themselves in the first place”
  - “Our enforcement system mainly relies on reports”
  - “We will be implementing more AI”
  - “Can these algorithms… scale?”
  - “Can they actually explain how they make decisions?”

Daily Bird

Red-bellied woodpecker, Central Park, December 25, 2019

CS@CU