

Nicolas Christin
Carnegie Mellon University
nicolasc@cmu.edu

Peer-to-Peer Networks: Interdisciplinary Challenges for Interconnected Systems

Agenda

- Outline
 - Introduction/Motivation
 - Conflicting incentives
 - Copyright law primer
 - Legal strategies
 - Technological strategies: P2P Network poisoning
 - Discussion lead-in: Social norms and the need for change?
- Objective
 - Provide a basic level of technological and legal literacy to inform ethics discussion
 - **Disclaimer: I am a technologist, not a lawyer**
 - Although I try to keep informed of the body of law relevant to my field of research

Motivation

- Peer-to-peer traffic represents 30% to 70% of Internet traffic (Madhukar, 2005)
- “Killer app” of the early 21st century
 - Arguably fostered rise of broadband networks
 - Many beneficial applications
 - Software distribution (Free – e.g., Linux distributions, or proprietary – World of Warcraft patches)
 - Censorship resilience (e.g., Freenet, Tor)
 - ... and also, dissemination of copyrighted materials
 - Songs, Movies, Software...

The roots of the P2P phenomenon: More and more information...

Source: <http://www.sims.berkeley.edu/research/projects/how-much-info-2003/>

- 5 exabytes of new information in 2002 with over 90% stored on magnetic media, mostly in hard disks. (1 exabyte = 2^{60} B or roughly 10^{18} B)
- Amount of information stored on paper, film, magnetic and optical media has doubled over 2000-2002

The roots of the P2P phenomenon: ... that is easy to replicate...

- With modern hardware and free software, it takes about:^{*}
 - 5-10 minutes to rip a music CD and encode all tracks in MP3 format
 - 10-25 minutes to decrypt and copy a 9 GB DVD onto a hard drive
 - 4-8 hours to compress a 90-minute movie into a format suitable (e.g., DivX) for storage on one CD

* Numbers are c. 2004, but have not changed significantly since

The roots of the P2P phenomenon: ... and even easier to distribute!

- Google indexed 8,168,684,336 web pages as of September 2005
 - (they do not publish that number anymore)
- eDonkey2000
 - 3,000,000+ users
- KaZaA
 - 2,000,000+ users
 - 4+ petabytes of data (1 PB = 2^{50} B or roughly 10^{15} B)
- BitTorrent
 - Pirate Bay trackers: 12,000,000 users in 2008 (total higher)
- Many others
 - (SoulSeek, DirectConnect, Freenet, ...)

Commonly held fallacy

- “The P2P dilemma pits end users vs. Copyright holders”
- End users
 - want free contents
 - are a bunch of uneducated savages happy to steal stuff
 - (depending on whom you ask)
- Copyright holders
 - Are almost broke, have a hard time making ends meet when “competing with free”
 - Are reckless corporate predators without an ounce of concern for the end users well-being
 - (again depending on whom you ask)



One wish it would be that simple...

A five-way tussle

- **End users:** want to obtain content at the smallest cost possible
- **Content providers:** live from selling content, arguably suffer from unauthorized replication
- **Electronics manufacturers:** often benefit from digital media portability; often have relationships with content providers
- **Software developers (and P2P service indexers):** benefit, potentially profit from P2P rise
- **Internet Service Providers:** benefit from P2P clientele, but suffer from traffic overload; sometimes sister companies of content providers!

Unclear set of incentives – ripe for negative externalities and incentives misalignment

Copyright policy

- Most copyright laws predate availability of digital content
- How does “digital world” affect copyright policy?
- History of evolution of copyright laws
 - Piano rolls
 - Sony Betamax
 - MGM vs. Grokster

Copyright elements

Source: <http://www.sims.berkeley.edu/~pam/coptutor/>

- Subject matter: works of authorship
 - E.g., literary works, musical works, pictorial works.
 - NB: software is a “literary work”
- Qualifications:
 - Who: the author (but in US, work for hire rule)
 - Procedure: rights attach automatically (but US authors must register to sue; remedies depend on regis.)
 - Criteria: “**originality**” (some creativity); [in US] works must also be “**fixed**” in some tangible medium

Copyright elements

Source: <http://www.sims.berkeley.edu/~pam/coptutor/>

- Set of exclusive rights (right to exclude others):
 - to reproduce work in copies,
 - to prepare derivative works, including translations
 - **to distribute copies to the public,**
 - to publicly perform or display the work, or communicate it to the public (broadcast)
 - “moral rights” of integrity & attribution
 - some rights to control acts of those who facilitate or contribute to others’ infringement (e.g., ISPs)
- Limitations on exclusive rights include
 - **Fair use** (e.g., *Sony Betamax*, *Acuff-Rose*) in US
 - Personal backup copies (e.g., software backups in EU)
 - Library-archival copying (e.g., ILL, course reserves)
 - Other (e.g., playing radio in fast food joint)

Digital Millennium Copyright Act

- Controversial US law passed in 1998 to address shift in technological means
- Not merely infringement of copyright itself!
- **Criminalizes production and dissemination of technology that can circumvent measures taken to protect copyright**
 - Should outlaw reverse engineering of DVD encryption for instance
 - Conflicts with free speech provisions (First Amendment)!
- Heightens the penalties for copyright infringement on the Internet
 - While at the same time limiting penalties against ISPs
- European Union passed EUCD in 2004
 - Cousin of DMCA, many similarities

MGM vs. Grokster

- Can developers of software that facilitates infringement be held liable for infringement?
- Debate about fair use and non-infringing uses
 - Sony Betamax case
- US Supreme Court opinion (unanimous)
 - "We hold that one who distributes a device with the object of promoting its use to infringe copyright, as shown by the clear expression or other affirmative steps taken to foster infringement, is liable for the resulting acts of infringement by third parties."
- Relatively loosely defined set of rules
 - How do you define "clear expression"?
 - Private emails within the company?
 - Aggressive marketing campaign?
 - Likely to result into more litigation

Alternatives to copyright

- DMCA (1998) was conceived before Napster (1999)
- Some argue that copyright itself became inappropriate given the scale of the infringement
 - “Everybody breaks the law, so we should change the law”
- Compulsory licensing
 - Pay a fee for use and (possibly unlimited) replication
 - Facilitates enforcement and regulation
 - Apple’s iTunes is a form of compulsory licensing

Legal strategies for © holders

- Suing the ISP
 - Secondary liability (facilitator)
 - Not very successful – ISPs are not even supposed to know what goes over their network (Federal Wiretapping Act, 18 USC §119)
 - Provisions for “maintenance”
 - Take down notice
- Suing the hardware manufacturer
 - Hard case to make – safe harbors from Sony Betamax (non infringing uses)
- Suing the software developers and/or indexing sites
 - See MGM vs. Grokster
 - Pirate Bay take-down
- Suing individual infringers
 - Primary infringement
 - Also quite striking with the public
 - Pop quiz: How many cases successfully prosecuted? How many cases settled?

Technological strategies for © holders: Network DRM

- Most revenue losses seem to come from P2P diffusion
- Crippling physical CD is dangerous, and doesn't really work
 - Different audio standard?
 - See DVDs...
 - TCG?
 - **Apparently mostly abandoned since 2005**
- Wouldn't it be more sensible to try and limit P2P diffusion?
 - Faces the same problems:
 - Dangerous from an ethical standpoint
 - Relatively easy to thwart

Perceived content replication

The screenshot shows the eMule v0.45b interface. The search results table is as follows:

File Name	Size	Availability	Complete Sources	Type
U2 - All Because Of You.mp3	6.31 MB	32	93%	Audio
06.- U2 - all because of you - [EMG].mp3	6.31 MB	20	90%	Audio
U2 - How To Dismantle An Atomic Bomb - 06 -All Because Of You.mp3	6.31 MB	11	54%	Audio
U2 - All because of you.mp3	6.36 MB	7	85%	Audio
U2 - All Because of You.mp3	6.31 MB	5	80%	Audio
U2 - All Because of you.mp3	B	2	100%	Audio
U2 - All Because of you.mp3	B	2	100%	Audio
U2 - All Because of you.mp3	B	2	100%	Audio
U2 - All Because of you.mp3	B	1	100%	Audio
U2 - All because of you.mp3	5.01 MB	1	100%	Audio
U2_-_ [how_to_dismantle_an_atomic_bomb]_-_06_-_all_because_of_you...	2.51 MB	1	100%	Audio
06.- U2 - all because of you - [EMG].mp3	6.32 MB	1	100%	Audio
U2 - All Because of You.mp3	3.35 MB	1	0%	Audio
U2 - How To Dismantle An Atomic Bomb - 06 - All Because of You.mp3	6.31 MB	1	100%	Audio
U2 - How To Dismantle An Atomic Bomb - 06 - All Because of You.mp3	6.31 MB	1	100%	Audio

A red box highlights the 'Complete Sources' column, and a red oval highlights the 'Availability' column. A red callout box points to the 'Availability' column with the text: "Content replication = Number of peers that share a given file on the network".

Perceived content replication

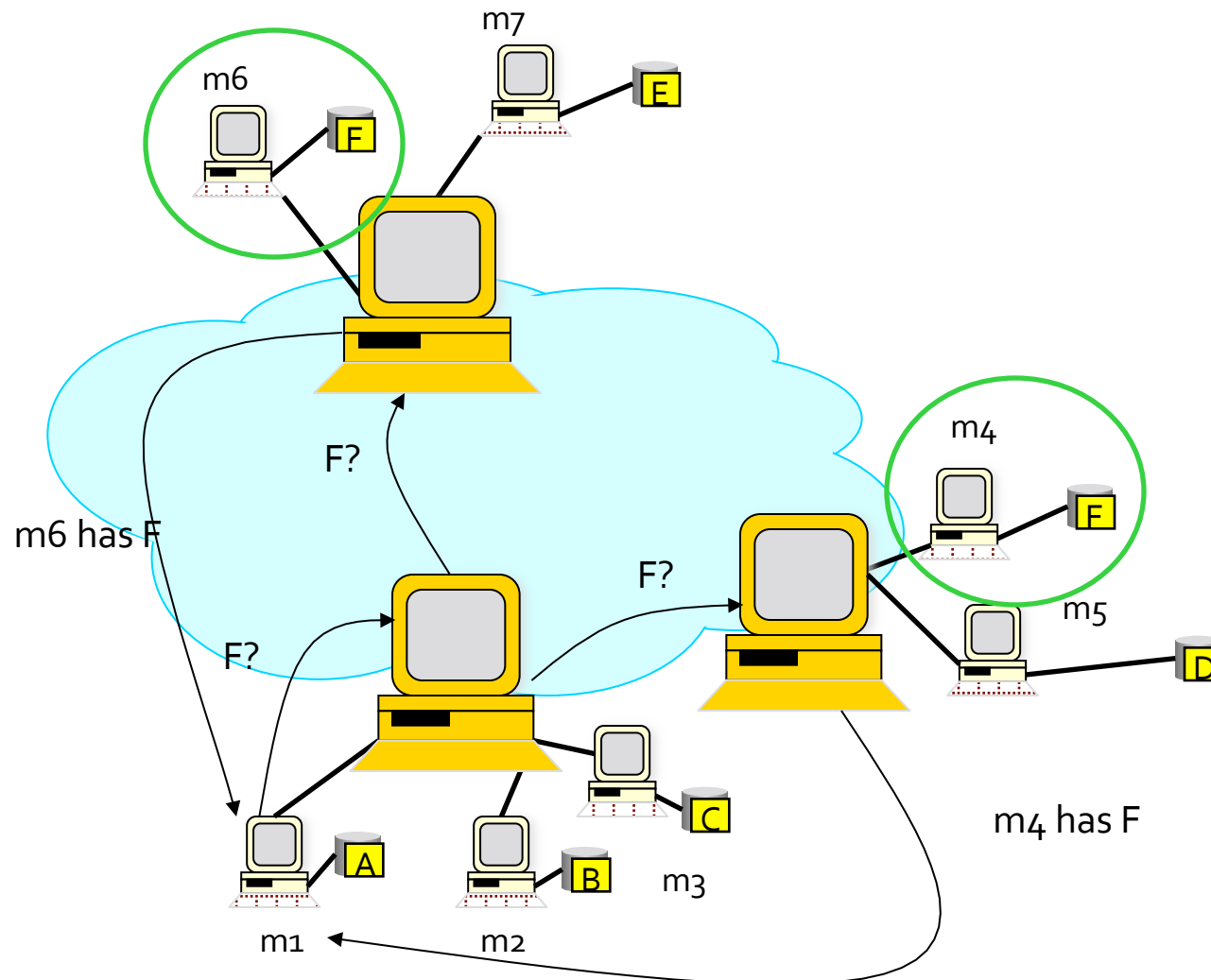
The screenshot shows the eMule v0.45b interface. The search results table is as follows:

File Name	Size	Availability	Complete Sources	Type
U2 - All Because Of You.mp3	6.31 MB	32	93%	Audio
06.- U2 - all because of you - [EMG].mp3	6.31 MB	20	90%	Audio
U2 - How To Dismantle An Atomic Bomb - 06 -All Because Of You.mp3	6.31 MB	11	54%	Audio
U2 - All because of you.mp3	6.36 MB	7	85%	Audio
U2 - All Because of You.mp3	6.31 MB	5	80%	Audio
U2 - All because of you.mp3	5.01 MB	2	100%	Audio
U2 - All because of you.mp3	2.51 MB	2	100%	Audio
U2 - All because of you.mp3	6.32 MB	2	100%	Audio
U2 - All Because of You.mp3	3.35 MB	1	0%	Audio
U2 - How To Dismantle An Atomic Bomb - 06 - All Because of You.mp3	6.31 MB	1	100%	Audio
U2 - How To Dismantle An Atomic Bomb - 06 - All Because of You.mp3	6.31 MB	1	100%	Audio

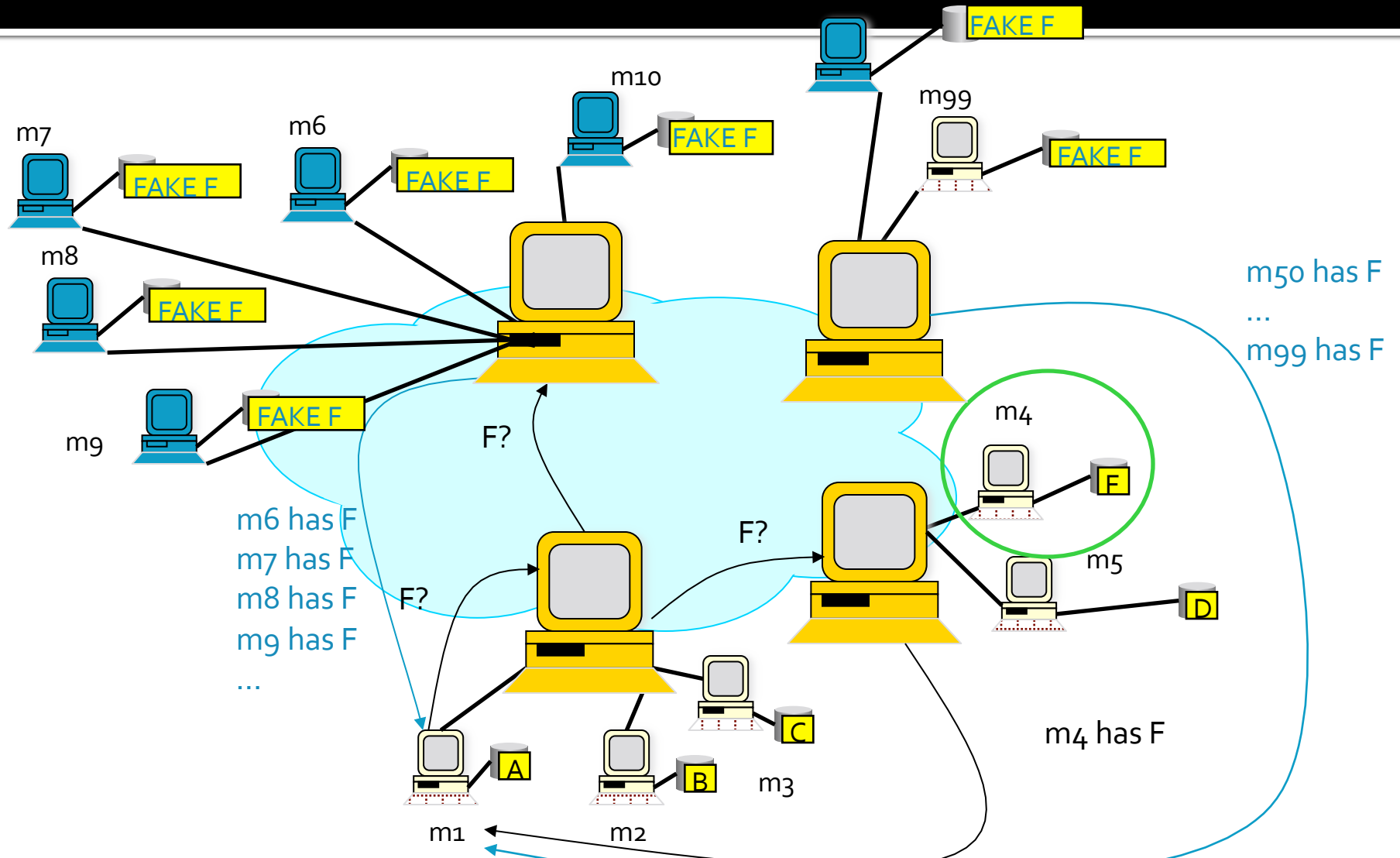
A red box highlights the 'Availability' column, and a red oval highlights the values in this column. A red callout box points to the value '32' in the first row.

Perceived content replication = Number of peers that I see sharing a given file on the network

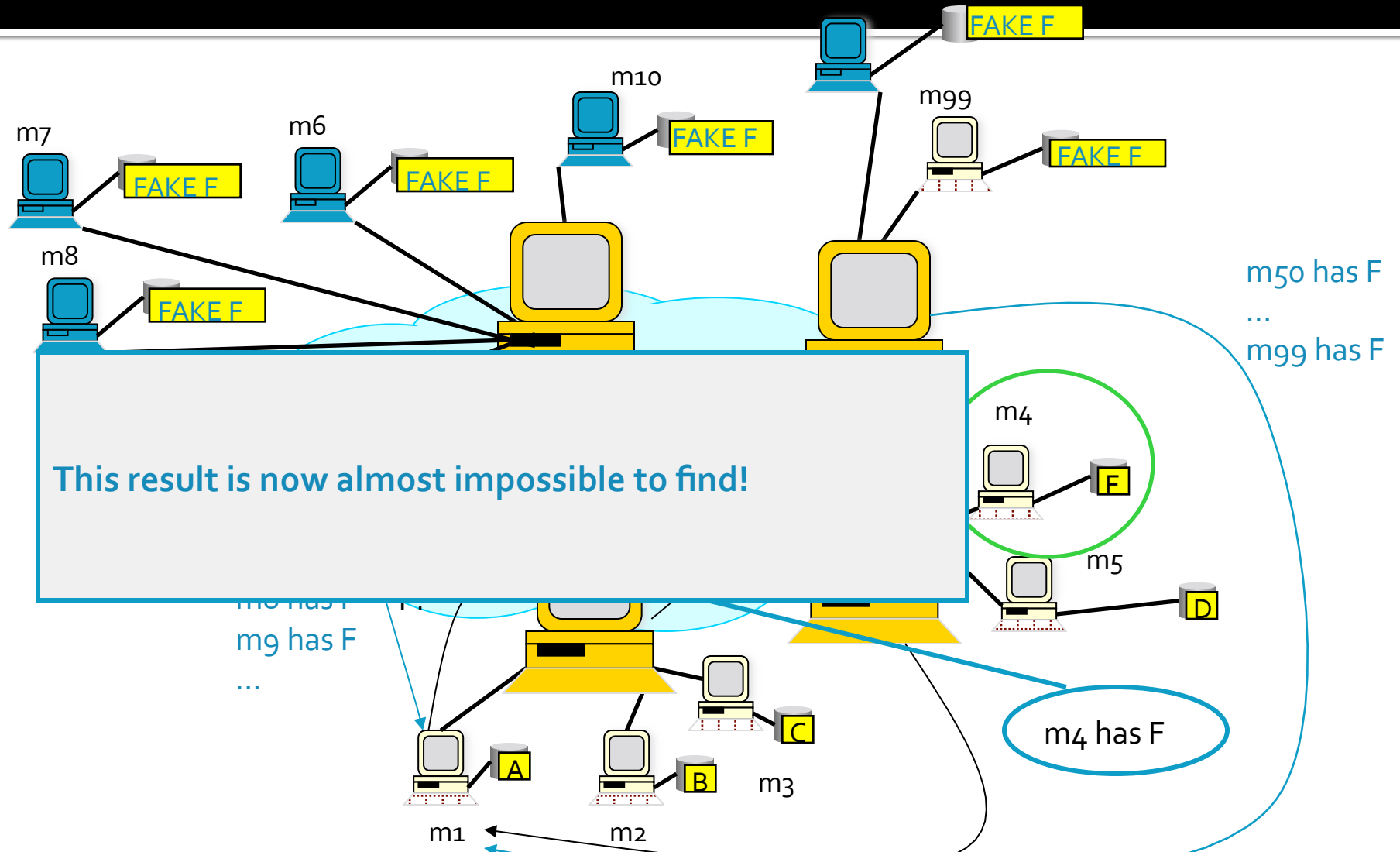
P2P poisoning



P2P poisoning



P2P poisoning



Poisoning techniques

	Item poisoning	Index poisoning
Strategy	Inject unusable files	Advertise nonexistent files
Pros	<ul style="list-style-type: none">• Higher level of user frustration• Self-sustaining<ul style="list-style-type: none">– Second order propagation	<ul style="list-style-type: none">• Low resource usage• Harmless to network
Cons	<ul style="list-style-type: none">• High resource usage<ul style="list-style-type: none">- Practical only for small files	<ul style="list-style-type: none">• Mostly ineffective against greedy downloads

Other censorship techniques

- DDoS content provider (m6 and m4)
- DDoS indexing nodes (yellow supernodes)
- Network partitioning
 - DDoS neighbors of the node that is serving content

All very questionable from a legal standpoint!

How the law may protect infringers

- CFAA 18 USC §1030
 - Makes it an offense to “*intentionally access a protected computer without authorization, and, as a result of such conduct, cause damage*”
 - Pretty much all machines fall under that category

How technology evolves

- BitTorrent
 - Decentralized network
 - Essentially “one network per file shared”
 - No global indexing service
 - (until recently, BitTorrent trackerless systems)
 - Are copyright holders going to target Google???
 - No but they did target The Pirate Bay
- Relatively easy to go after “trackers”
 - And yet, this is usually bungled (Piatek et al. 2008)
 - Printers receiving take-down notices
 - Only guilty of requesting names and addresses of peers!

Preliminary discussion points

- Evolving social norms: if everybody breaks the law, then the law must be broken
 - Pirate Party of Sweden
 - Alternatives?
- Ethics

Players Ethics

	Utilitarian	Deontological	Virtue
End users			
Content providers			
Hardware manufacturers			
ISP			
Software manufacturers/ indexing services			