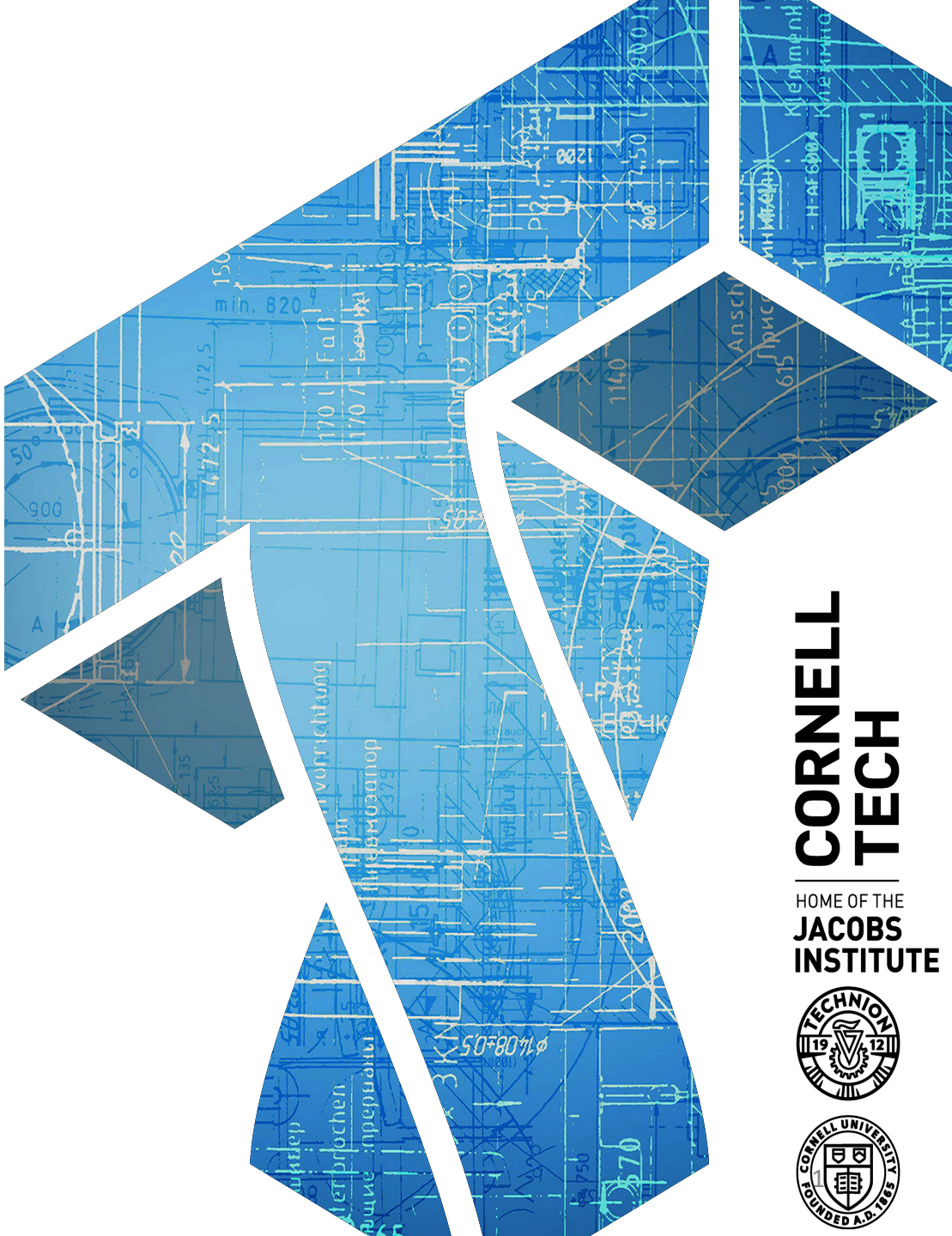
CS 5435: Computer fraud and abuse

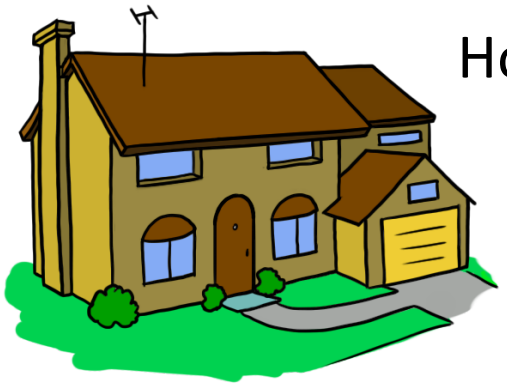
Instructor: Tom Ristenpart



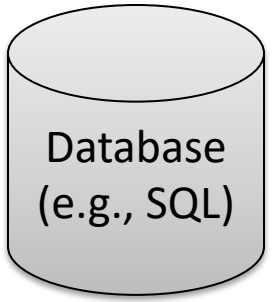
CORNELL TECH

HOME OF THE JACOBS INSTITUTE

So far: authentication



Hosts



Guests

Today: abusing services with (authenticated) access to them

A useful dichotomy

- Financially-motivated abuse

- Non-financially-motivated abuse

Early view on tech fraud: email spam and scams

- Spam
  - unsolicited bulk emails
  - Illegal in USA since CAN-SPAM act of 2003
- Scams
  - Nigerian emails (advanced fee fraud / confidence trick)
- Phishing
  - trick users into downloading malware, submitting password to attacker, CC info to attacker, etc.
  - Spear phishing: targeted on individuals (used in high-profile intrusions)

Jon

Junk - Exchange August 24, 2019 at 3:16 PM

J

Official request

Reply-To: jocun01250@mail2banker.com

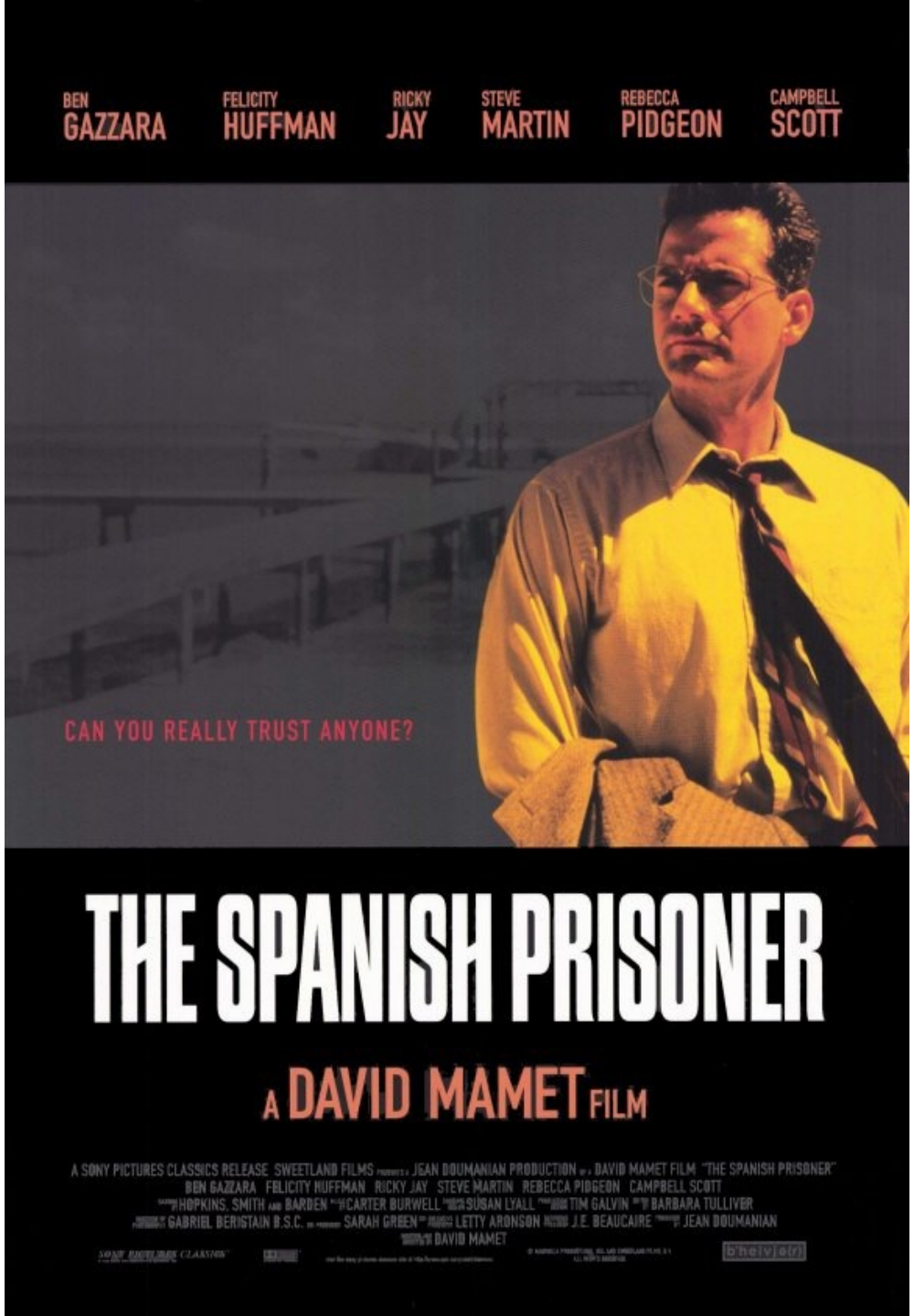
Hello,

This is an official request to you, we have a deceased client whose inheritance we wish to payout to avoid seizure. We are no longer paying for private searches to trace his relatives after today. We contacted you knowing you could be. Please treat urgent.

Regards,
Jon

Spanish Prisoner confidence trick

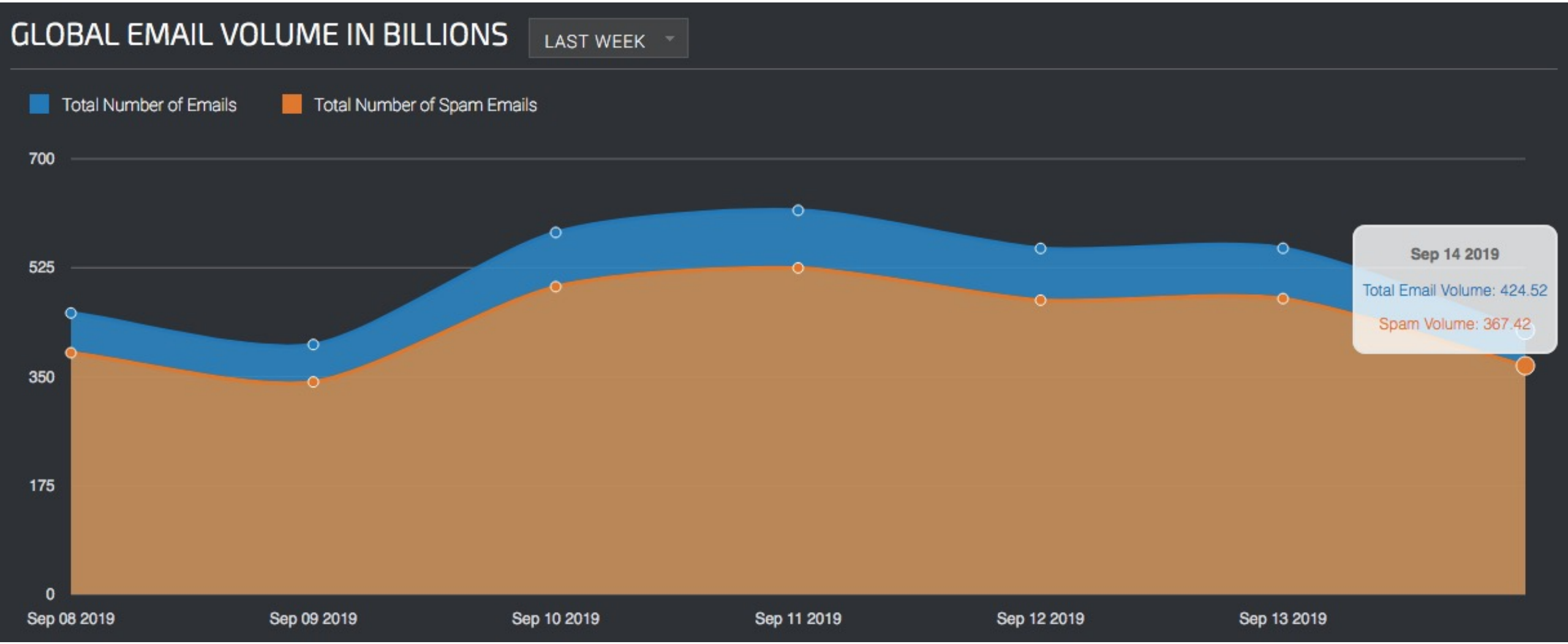
- 19th century
- In contact with rich guy in Spanish prison
- Just need a little money to bribe guards, he'll reward you greatly
- ***Advance-fee scam***



By Victim Loss

Crime Type	Loss	Crime Type	Loss
BEC/EAC	$1,297,803,489	Tech Support	$38,697,026
Confidence Fraud/Romance	$362,500,761	Harassment/Threats of Violence	$21,903,829
Investment	$252,955,320	Misrepresentation	$20,000,713
Non-Payment/Non-Delivery	$183,826,809	IPR/Copyright and Counterfeit	$15,802,011
Real Estate/Rental	$149,458,114	Civil Matter	$15,172,692
Personal Data Breach	$148,892,403	Malware/Scareware/ Virus	$7,411,651
Corporate Data Breach	$117,711,989	Health Care Related	$4,474,792
Identity Theft	$100,429,691	Ransomware	*$3,621,857
Advanced Fee	$92,271,682	Denial of Service/TDos	$2,052,340
Credit Card Fraud	$88,991,436	Re-Shipping	$1,684,179
Extortion	$83,357,901	Charity	$1,006,379
Spoofing	$70,000,248	Gambling	$926,953
Government Impersonation	$64,211,765	Crimes Against Children	$265,996
Other	$63,126,929	Hacktivist	$77,612
Lottery/Sweepstakes	$60,214,814	Terrorism	$10,193
Overpayment	$53,225,507	No Lead Value	$0.00
Phishing/Vishing/Smishing/Pharming	$48,241,748		
Employment	$45,487,120		

https://pdf.ic3.gov/2018_IC3Report.pdf



https://www.talosintelligence.com/reputation_center/email_rep#tab=1

Storm botnet (2007-08)

- September 2007
  - Media: 1 – 50 million bots
  - More likely: 10,000s to 100,000s
- Early spam campaigns used titles such as "230 dead as storm batters Europe."
- Propagated via spam linking to malware
- Thought to be controlled by Russian Business Network

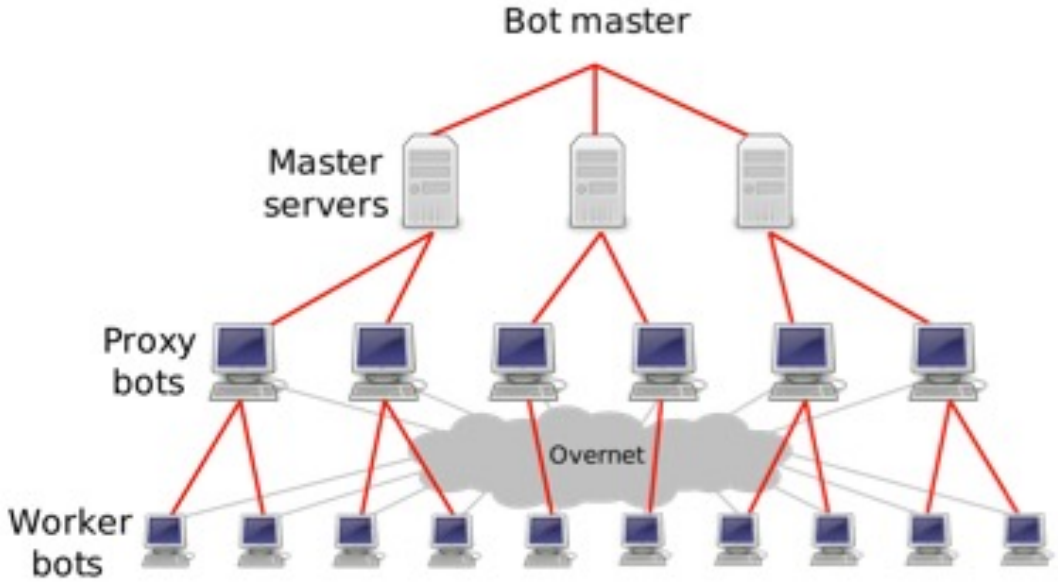


Figure 1: The Storm botnet hierarchy.

Features:

- Uses P2P (Overnet/Kademlia)
- Uses fast-flux DNS for hosting on named sites
- Binary has gone through many revisions
- Features of P2P network have evolved with time
- Hides on machine with rootkit technology

[Enright 2007]

See also: http://cseweb.ucsd.edu/~savage/papers/login08.pdf

How to make money off a botnet?

- Rental
  - "Pay me money, and I'll let you use my botnet… no questions asked"
- DDoS extortion / DDoS for hire
  - "Pay me or I take your legitimate business off web"
  - "Pay me and I'll take someone else off Internet"
- Bulk traffic selling
  - "Pay me to direct bots to websites to boost visit counts"
- Click fraud, SEO
  - "Simulate clicks on advertised links to generate revenue"
  - Cloaking, link farms, etc.
- Theft of monetizable data (eg., financial accounts)
- Data ransom (now called ransomware)
  - "I've encrypted your harddrive, now pay me money to unencrypt it"
- Spam to advertise products

Underground forums

Category	Threads B	Threads S	Users B	Users S	Top Subcategory
payments	8,507	8,092	1,539	1,409	paysafecard
game-related	2,379	2,584	924	987	steam
accounts	2,119	2,067	850	974	rapidshare
credit cards	996	1160	467	566	unspecified cc
software/keys	729	1410	422	740	key/serial
fraud tools	652	1155	363	601	socks
tutorials/guides	950	537	562	393	tutorials
mail/drop srvs	751	681	407	364	packstation
merchandise	493	721	264	404	ipod
services	266	916	176	555	carder

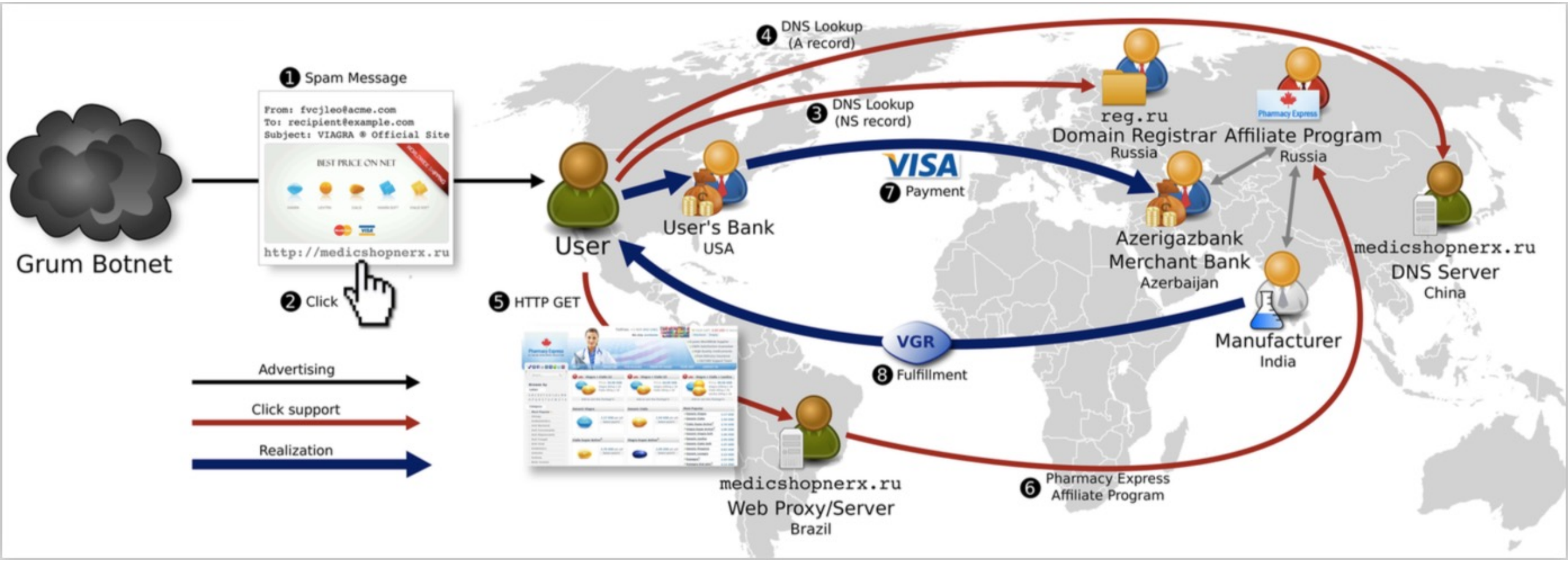
Table 6: Top 10 most commonly traded merchandise categories on LC.

[Motoyama et al., An Analysis of Underground Forums, 2011]

How to make money off financial credentials?

- Money mules
  - Deposits into mules' account from the victim's
  - Mule purchases items using stolen CCN, sells them online
  - Mule withdraws cash from ATMs using victim credentials
- Wires money to (for example) former Soviet Union

Example spam-advertised goods backend

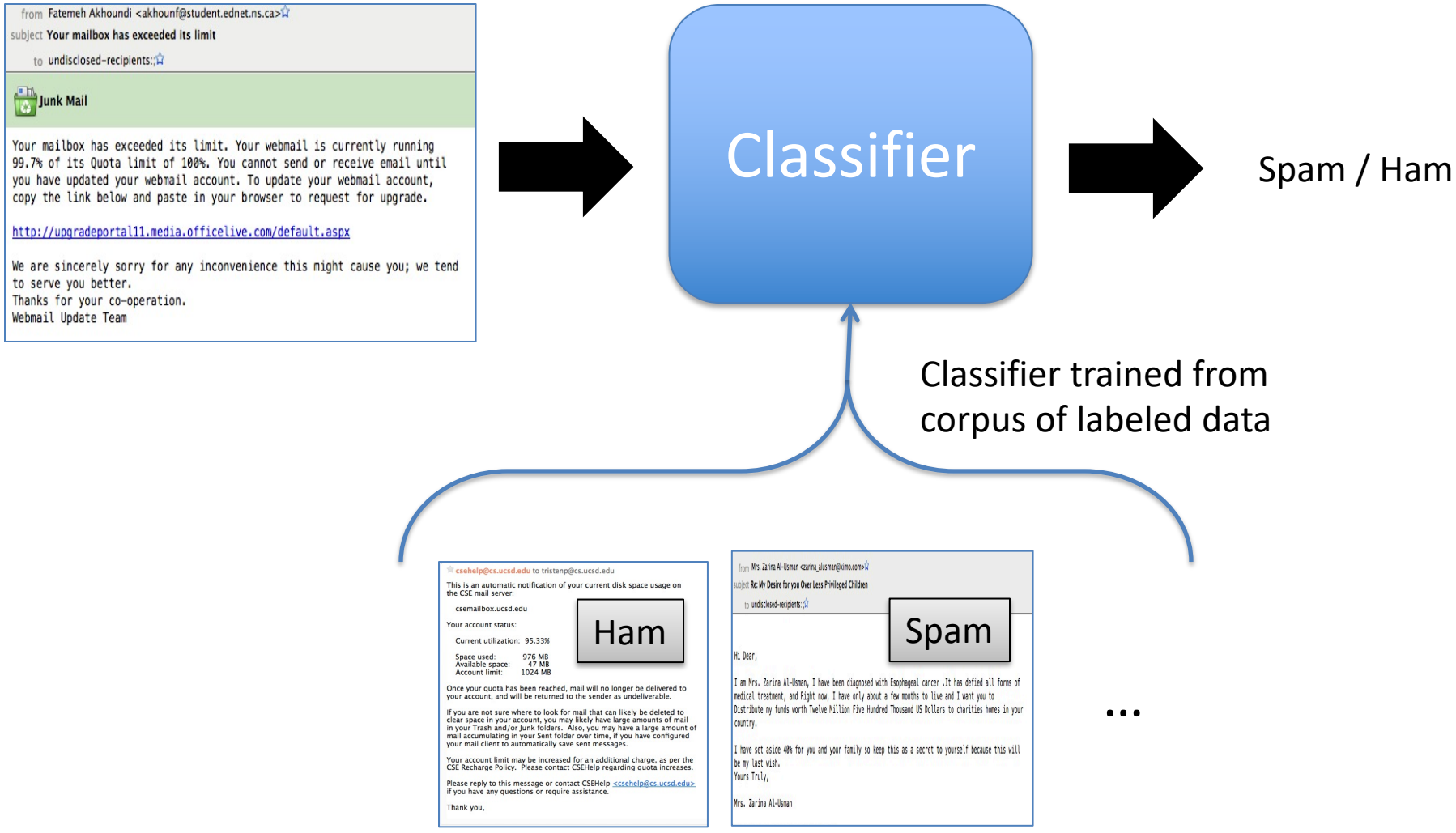


From Levchenko et al., “Click Trajectories: End-to-End Analysis of the Spam Value Chain”, IEEE Symposium on Security and Privacy, 2011

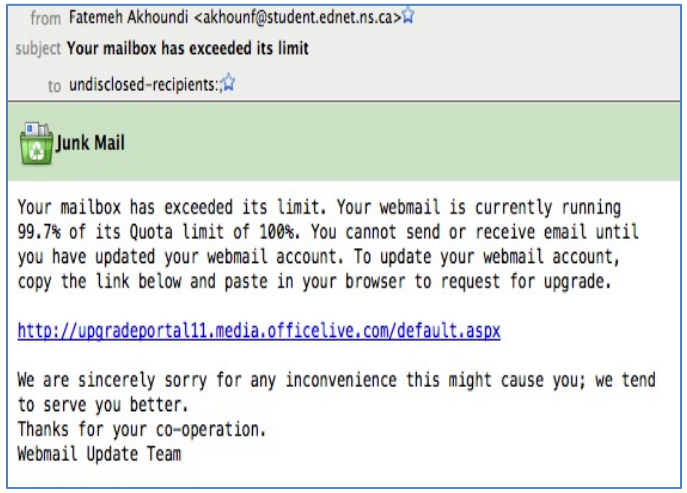
How to prevent fraudulent actions?

- **Content analysis**
  - Spam filters
- **Identify bad accounts**
  - Correlate bad actions with accounts, try to detect bots
  - Sock-puppet accounts (many controlled by one person)
- **Identify bad devices**
  - Device cookies to correlate different account accesses
  - IP blacklists, ISP blacklists
- **Target fraud-support infrastructure**
  - Botnet takedowns, bank accounts, jail sentences

Spam Classifiers



Naïve Bayes Classifier



Represent email as "bag of words"

quota	1	x_1
webmail	4	x_2
cornell	0	x_3
fee	1	x_4
⋮	⋮	

Intuition: spam and ham have different distribution of keywords

$$\Pr[\text{spam} \mid x_1, x_2, \ldots, x_n] = \frac{\Pr[x_1, x_2, \ldots, x_n \mid \text{spam}]\ \Pr[\text{spam}]}{\Pr[x_1, x_2, \ldots, x_n]}$$

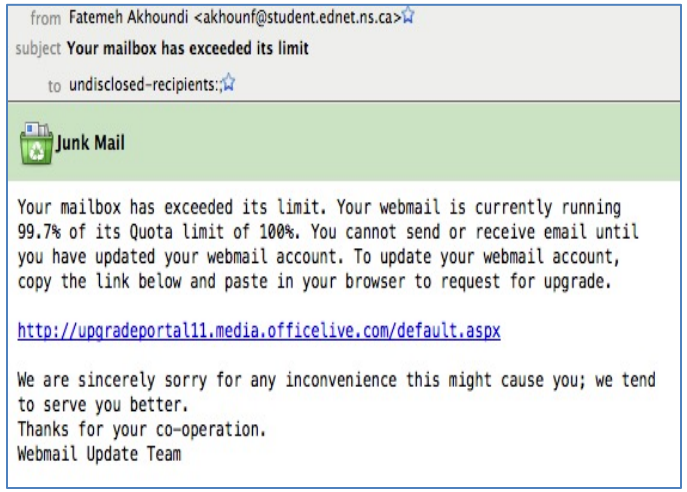
Bayes' theorem

$$= \frac{\Pr[\text{spam}] \prod \Pr[x_i \mid \text{spam}]}{\Pr[x_1, x_2, \ldots, x_n]}$$

"Naïve": assume words independent

$$\Pr[\text{ham} \mid x_1, x_2, \ldots, x_n] = \frac{\Pr[\text{ham}] \prod \Pr[x_i \mid \text{ham}]}{\Pr[x_1, x_2, \ldots, x_n]}$$

Naïve Bayes Classifier



Represent email as "bag of words"

quota	1	x_1
webmail	4	x_2
cornell	0	x_3
fee	1	x_4
⋮	⋮	

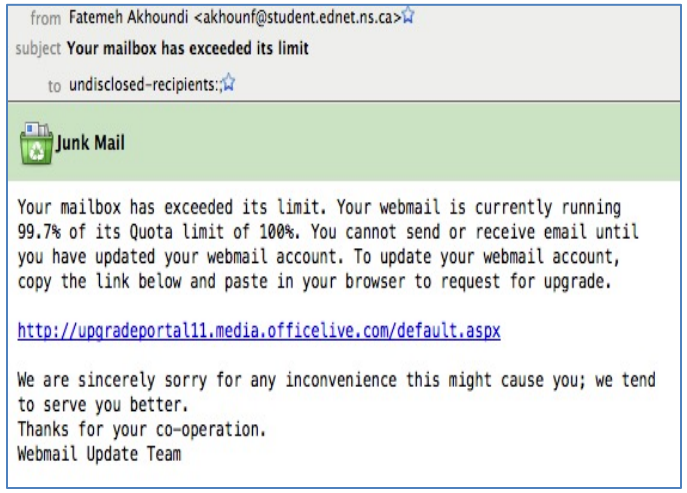
Intuition: spam and ham have different distribution of keywords

$$Pr[\text{ spam } | \ x_1 , x_2, \dots , x_n] = \frac{Pr[\text{spam}] \prod Pr[x_i | \text{ spam }]}{Pr[x_1 , x_2, \dots , x_n]}$$

$$Pr[\text{ ham } | \ x_1 , x_2, \dots , x_n] = \frac{Pr[\text{ham}] \prod Pr[x_i | \text{ ham}]}{Pr[x_1 , x_2, \dots , x_n]}$$

Classify as spam if: $Pr[\text{ spam } | \ x_1 , x_2, \dots , x_n] > Pr[\text{ ham } | \ x_1 , x_2, \dots , x_n]$

Naïve Bayes Classifier


from Fatemeh Akhoundi <akhounf@student.ednet.ns.ca>
subject **Your mailbox has exceeded its limit**
to undisclosed-recipients:;

Junk Mail

Your mailbox has exceeded its limit. Your webmail is currently running 99.7% of its Quota limit of 100%. You cannot send or receive email until you have updated your webmail account. To update your webmail account, copy the link below and paste in your browser to request for upgrade.

http://upgradeportal11.media.officelive.com/default.aspx

We are sincerely sorry for any inconvenience this might cause you; we tend to serve you better.
Thanks for your co-operation.
Webmail Update Team

Represent email as "bag of words"

quota	1	x_1
webmail	4	x_2
cornell	0	x_3
fee	1	x_4
⋮	⋮	

Intuition: spam and ham have different distribution of keywords

Estimate these from labeled training data

$$\Pr[\text{spam} \mid x_1, x_2, \ldots, x_n] = \frac{\Pr[\text{spam}] \prod \Pr[x_i \mid \text{spam}]}{\Pr[x_1, x_2, \ldots, x_n]}$$

$$\Pr[\text{ham} \mid x_1, x_2, \ldots, x_n] = \frac{\Pr[\text{ham}] \prod \Pr[x_i \mid \text{ham}]}{\Pr[x_1, x_2, \ldots, x_n]}$$

Classify as spam if: $\Pr[\text{spam}] \prod \Pr[x_i \mid \text{spam}] > \Pr[\text{ham}] \prod \Pr[x_i \mid \text{ham}]$

Spam classifiers

- Nowadays classifiers more complex than this
  - Other features: Who is sender? How many links embedded?
  - Can update in real-time given labeling by user(s)
  - For larger orgs, can leverage wide view across many email recipients (Gmail)
- Nowadays some companies do pretty good job of making sure spam doesn't hit your inbox
  - 95% of email gets filtered as spam (2009, ENISA Spam Survey)

Botnet countermeasures?

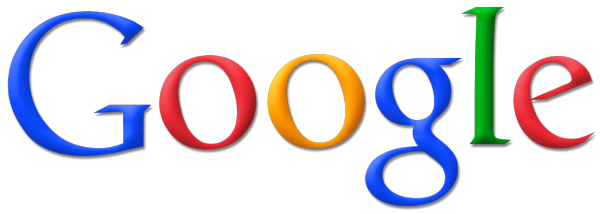
- Infection prevention
- Infection detection
- C&C take-down
- Undermine the economics
  - Banking take-down. See

Microsoft Seizes ZeuS Servers in Anti-Botnet Rampage

BY KIM ZETTER 03.26.12 2:45 PM

http://www.wired.com/threatlevel/2012/03/microsoft-botnet-takedown/

Web services in-class exercise



In-class 5-minute exercise:

Discuss your favorite web service:

1. What commercially-motivated abuse might it have to contend with?
2. How might you prevent it?

Removing bad actors from services is *hard*

- **Facebook**: spammer accounts, fake fraud accounts, …
- **Twitter**: illicit promotional accounts, …
- **Yelp:** fake reviews, fake restaurants, …
- **AirBNB**: scam rental or experience ads
- **Lyft**: colluding drivers + riders
- **Google**: spammers using Gmail, advertisers violating terms of service, SEO, Play store bad apps, …

Examples of other forms of abuse

- Online harassment and bullying
  - Coordinated campaigns
  - Doxxing, raiding (other websites), …
- Misinformation campaigns
- Targeted attacks and RATs (remote access trojans)
- Tech abuse in intimate partner violence

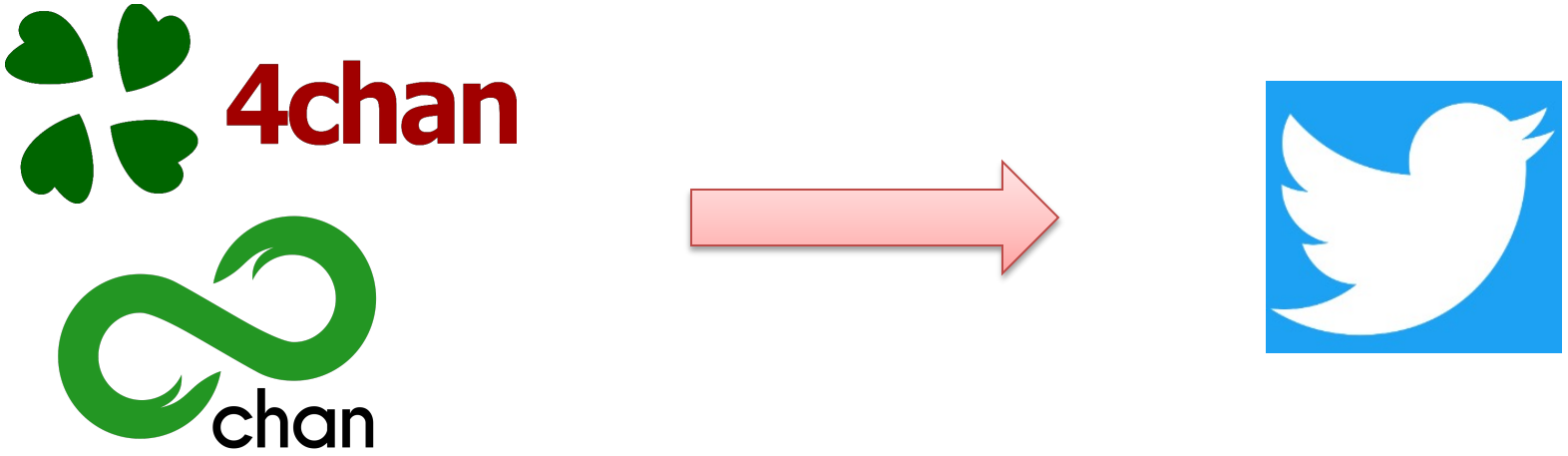
Rest of lecture will discuss sexual violence, physical violence, and other forms of abuse

Coordinated harassment campaigns

- Anonymous bulletin board
- Generated lots of memes, known to host some child sexual abuse media (CSAM) (though technically against site policy)
- Associated with Anonymous hacker group
- Used to coordinate harassment, bullying, hacking
- Associated with alt-right groups
- Involved in Gamergate, banned Gamergate discussion

- Similar to 4chan (different site operator)
- No longer available due to CloudFlare and other providers dropping it

Coordinated harassment campaigns: Gamergate



- Sustained harassment campaign against female gaming developers and others
  - Rape threats, murder threats, doxing (home address, other personal information disclosed)
  - Astroturfing, sock-puppet campaigns on Twitter
- Organized on 4chan & (after banning on 4chan) 8chan, most harassment played out on twitter (#gamergate)
- Why? Ex-boyfriend of initial victim posted false claims about victim
  - "Right-wing backlash against progressivism" (in gaming)

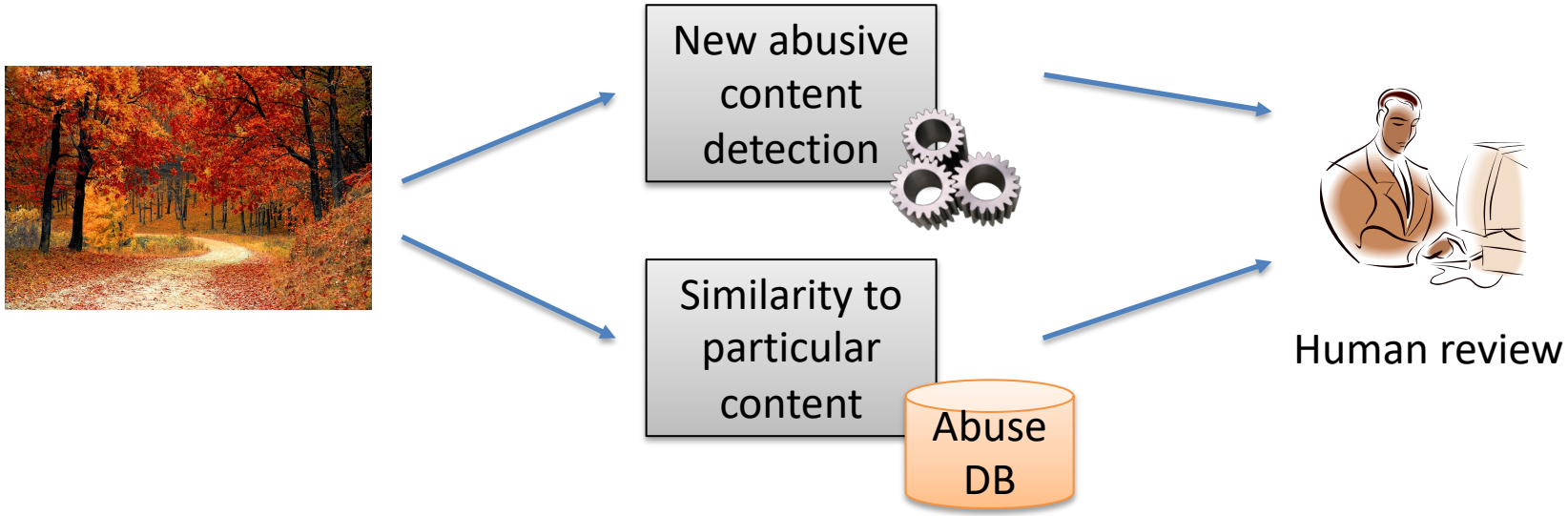
Mininformation campaigns

- Pizzagate conspiracy
- QAnon conspiracy
- 2016 election interference
  - Cambridge Analytica
  - Russian influence operations
- Research studies indicate falsehoods spreads faster than truth on social media

"there is a worldwide cabal of Satan-worshiping pedophiles who rule the world, essentially, and they control everything. They control politicians, and they control the media. They control Hollywood, and they cover up their existence, essentially. And they would have continued ruling the world, were it not for the election of President Donald Trump,"

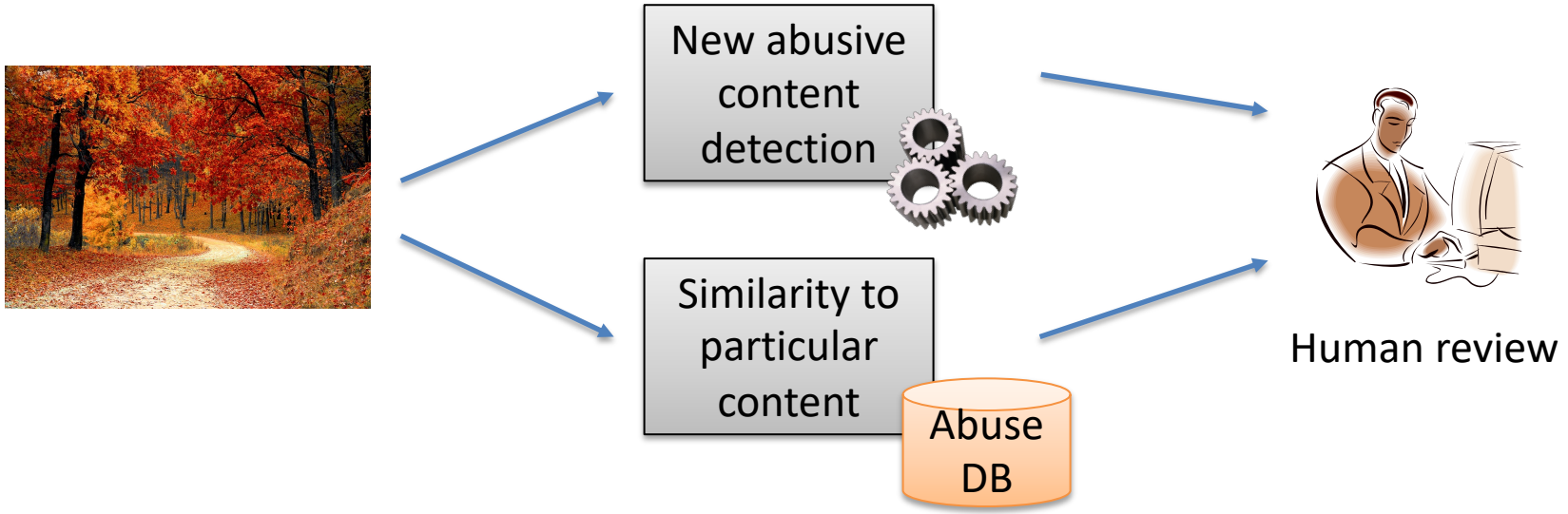
-Travis View

Identifying abuse content



- Use locality-sensitive hashing to check if similar to known bad content
  - H such that |H(image) – H(image')| < threshold if image, image' very similar
- Use machine learning to try to identify patterns indicative of abusive content
  - Is image a picture of a naked child?
- Refer out to human for review (Facebook has 10,000s of moderators)
- Content flagged by users also sent through reviewing pipelines
  - Flagging mechanisms also subject to abuse (illicit takedowns)

Identifying abuse content



- PhotoDNA is system for similarity matching of child sexual abuse media (CSAM) run by Microsoft
  - Used widely in industry
- Legal requirements around CSAM strict: must report detected content to National Center for Missing and Exploited Children (NCMEC)

Targeted attacks

- Dissidents, journalists, activists targeted by nation-states
  - Phishing attacks, botnet-style C&C servers to collect data
  - Remote Access Trojans (RATs)
- Small industry of companies providing “lawful access” tools


From: Melissa Chan <melissa.aljazeera@gmail.com>
To:
Sent: Tuesday, 8 May 2012, 8:52
Subject: Torture reports on Nabeel Rajab

Acting president Zainab Al Khawaja for Human Rights Bahrain reports of torture on Mr. Nabeel Rajab after his recent arrest.

Please check the attached detailed report along with torture images.

1 attachment: Rajab.rar 1.4 MB Save

Figure 1: E-mail containing FinSpy.

https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-blond.pdf
https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-marczak.pdf
https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-hardy.pdf

Targeted personal attacks

25% of women
11% of men

suffered rape, physical violence, and/or stalking **by an intimate partner**

[National Intimate Partner and Sexual Violence Survey 2010-2012]

Intimate partner violence (IPV) abusers exploit technology:

- Harassing texts/messages
- GPS devices & spyware apps
- Victim accounts being “hacked”
- Physical device access
- …

Four categories of common IPV tech attacks

Ownership-based

- Abuser owns device/account
- Shared account/device
- Buying children device

- Prevent use / destroy device
- Digitally control access
- Track location, monitor usage

Account/device compromise

- Physical access to unlocked device
- Force password / pin revelation
- Remotely "hack" via security questions / passwords

- Install spyware / "dual-use" app
- Track location, monitor victim
- Steal or delete info
- Lock victim out of account
- Impersonate victim

Harmful messages or posts

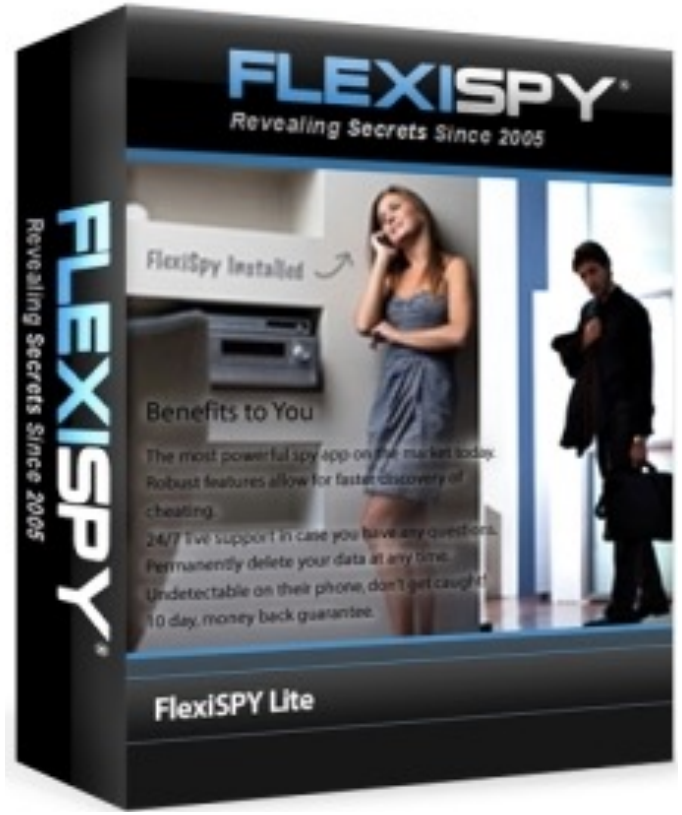
- Call/text/message victim (from spoofed account)
- Post harmful content (e.g., threaten violence)
- Harass victim's friends/family
- Proxy harassment

Exposure of private information

- Blackmail by threat of exposure
- "Doxxing" victim
- Non-consensual intimate images
- Fake profiles/advertisements of sexual services

The spyware (aka stalkerware) problem

"An abusive partner kicked in our front door and wound up in the lobby of our building by tracking her phone . . . ***it was some secondary application that the abuser had put on it and knew exactly where she was****. He literally kicked our front door open. We called the police … it was scary."* – Case manager



https://www.spymasterpro.com/

Clinic to End Tech Abuse

- Research on tech abuse in IPV led to new clinical model for helping survivors
  - Direct support of survivors on tech abuse by volunteer tech consultants
  - Partnership with New York City Mayor's Office to End Domestic and Gender-Based Violence (ENDGBV)
- Work on improving technology

- https://www.ceta.tech.cornell.edu/