CeasNotesJustin

Writing from CEAS with quick notes on each talk. Proceedings are at http://www.ceas.cc/papers-2004/papersbytopic.htm .

All ports except 80 and 443 are blocked! Very annoying

Chung-Kwei:

- Teiresias is IBM's pattern-discovery tool from bioinformatics
- looks *directly* transferrable to SpamAssassin's "regexp rules" approach
- probably heavily patented and hard to license though
- but a Google search for "pattern discovery algorithm" looks like a promising source

Social network talk:

- pretty useless spamfiltering-wise at least; not any spam orientation at all

Joshua Goodman Received talk:

- talking about parsing Received lines
- basically reimplementing spamcop algorithm
- looking for "last external IP address"
- thinks this will be useful for SenderID
- SenderID example uses HELO data, looks like, instead of PRA or SMTP MAIL FROM; due to multiple intervening hops
- try to use heuristics to find last external IP address:
  - using MX data fails due to load-balancing edge router
  - also the msn.com/hotmail.com problem
- proposed algo:
  - IP addr is 192.168
  - HELO matches user's domain and forward DNS lookup of HELO matches IP address
  - find an IP that matches MX record, next is external
- Bob Atk suggested putting external IP addrs in a DNS record?!
- interesting that they'd never checked SpamAssassin or Spamcop's algorithms, but that's MS for you

Brett Watson: beyond identity: problems even with sender id

- economics of whitelisting/blacklisting based on a reliable sender identification (ie. forging is no longer possible)
- mostly a philsophical talk

Multiple email addresses:

- about 50% of surveyed users had multiple email addresses
- "identities"; separation of work, personal, social groups; pseudo-anonymity; affiliation, status, prestige (alumni accts)
- mobility (available on the road)
- people now frequently have different "role" accounts
- typically once people go over 3 accts, they set them up to forward to a smaller number
- 20-30% of all email addrs change annually
- this talk is really oriented towards MUA UI developers
- another talk with not a whole lot of antispam relevance

Panel discussion of monetary spam filtering:

- Cynthia Dwork's talk:
  - 16 seconds per message computation time doubles spam cost
  - 56 seconds per message " means $36 per message for spammers
  - cycle theft arguments (zombies are illegal; spyware can be combatted with user @+ education) *already don't work* in the real world
- MailFrontier:
  - some kind of marketroid noise about how they're "third generation" because they have grey areas, or something; combination of multiple tests means "definitely spam, no false positives". riiight
  - "Reverse Turing Test": C-R as usual, with pictures of puppies
  - except the C-R page has some kind of plugin which will burn CPU cycles instead, woo
- The naysayer:
  - http://www.cl.cam.ac.uk/~rnc1/
  - going rate to solve puzzles is about $.11/hr in South India
  - Real Money systems: people will regulate it; EU Directive on E-Money (2000/46/EC)
  - people will walk away with 2.5% of it (cost of running + greed)
  - people will steal it (e.g. sysadmin skimming x% of incoming mails and stealing their tokens)
  - Payment systems: settlement: see taugh.com
  - also compares with the telco system (~1200mill ham mails/day, ~2000mill phone calls per day) – much fewer calls on telco system, most local, diff trust model
  - how much payment:
    - 30 responses per mill: .1c/mail mean $33 per sale to be viable
    - if .05c/mail, $16
    - at a 0.7% response rate, $33 profit means 23c/mail
- questions:
  - to Ironport: "why can't I nominate a charity?" to avoid interested parties
  - Dan Kohn to Ironport: "how much bonds debited?" not very much

- question from an Indian querier: "any documented cases of South Indian kids clicking on CAPTCHAs?" MailFrontier guy, naturally, says "nope". In reality, the answer is "yes", but that was in Thailand
- Yahoo! guy on CAPTCHAs: "seen everything: porn sites, people paid to type them; sites in Russia with full pages of CAPTCHAs, 10 hour turnaround after a new fix is deployed"
- Vanquish guy says they use CMU's CAPTCHA code
- question on CPU time stamp inflation: Cynthia Dwork says "memory cycles much more stable over time"
- Daniel: annoyed about senders having to "prove they are real" when they're doing the recipient a favour: MailFrontier guy: "we just want the problem to go away"
- Dave Crocker: "why didn't anyone on the panel take any notice of the naysayer's presentation and its points?"
- panel: "but we have only 5 minutes!"
- Vanquish guy: "he doesn't understand how PKI works" (!!!) then some advertising for Vanquish (again)
- Ironport: "Bonded Sender is working right now"
- MailFrontiers guy: "mostly agreed with his presentation, but we'll do whatever works (titters from audience); C-R is an atomic bomb against spam, but with some collateral damage against ham, but it can be turned off"
- naysayer on pay-to-send: "not only is my machine insecure, my email is insecure, but I don't want my *money* to be insecure" (applause)
- panel mod: there will be coevolution between attacker and defender, a lesson from the Cold War

MailFrontier presentation: anatomy of a phishing email

- Bank of America sends email from bankofamerica1.com, Sony from sonystar.com; this screws up the notion of a trusted domain name
- the MSIE %00 vulnerability
- high-numbered ports mean that websites can be run unnoticed, even if a HTTP server is already running
- the fake address-bar window trick
- fraudulent pop-ups over real site: goes to fraud site, create popup, go to fraud site: pop-ups are a phishing risk (yay!)
- "your submitted information will be verified by eBay staff within 24 hours"; buys more time
- A survey, based on results from over 83,450 respondents (subset of total responses), in diagnosing which sites were frauds and which were real:
  - 26.7% got everything wrong
  - only 13.8% of respondents got all correct
- da.ru is a frequent hosting site for phishing scams
- hasn't looked at the Active/X malware on the phishing sites, for some reason!
- Consumer Reports sends from some domain called "d1sub.com"; Fortune 500's should really improve their practices
- q: "are we getting to a stage where we won't be able to tell phish from ham?" a from audience: "use pine"
- q: "why haven't the arrests of phishers been publicised better?" suggests including some support in web browser for a "trusted logos" area on-screen, for certifications
- Dave Crocker: don't map to domain names, "domain names are not good enough, they do *not* map to trademarks".

Geoff Hulten, MS: Trends in Spam Products and Exploits

- corpus analysis, from Hotmail's feedback loop
  - volunteers classify random samples of their mail as spam or good; tens of thousands of hand-classified messages per day; large "unbiased" (???) sample of spam
- additional analysis on two sets of spam:
  - about a year between the two
  - products sold, exploits used, trends
- viagra types: 17% 2003, to 34% 2004
- graphic porn down: 13% to 7%
- exploits: increasing rapidly, 1.33 exploits 2003 to 1.73 in 2004
- word obscuring: up to 20% in 2004
- URL chaffing, adding good URLs to spam: not there in 2003, 10% in 2004 – anti-SURBL attack 😷
- Spammers are putting more work into each spam

Introducing the Enron Corpus:

- 1.3million messages originally; removed msgs with "integrity problems", replaced usernames etc
- http://www-2.cs.cmu.edu/~enron
- 200,399 useful, non-dupe messages
- 158 messages, 1,268 msgs/user
- missing message headers, so not much use for spam filtering; Exchange-mangled; no HTML. still, maybe good for "body" rules and FP avoidance
- no mention how much of the corpus was spam 😷

Larry Lessig:

- extraordinary amount going to tech fixes; very little going to how the law could address it
- compares govt attention to "pirate radio" creating static for large commercial stations, vs the spam problem
- multiple types of regulators: the law, social norms, the market, and architecture (example: windows in lecture theatre are closed to enforce paying attention to speakers)
- the law also regulates the other three
- (that was the wrong talk! starts again!)
- 1. "regulation is always multiple modalities"
- 2. "interests will react"
- 3. "special interests defeat general interests"
- in the old days, we had norms to defeat spam; that failed
- using code to fix; so far that's failed
- "the market will fix the problem"; ISPs trying to be the spam-free email provider; that's also failed
- CAN-SPAM: totally failed – even displaced effective state legislation
- not any single modality alone can fix it
- regulation is a restriction, plus somebody to enforce it
- CAN-SPAM: wanted truthful headers
- opt-out doesn't provide any way for you to know if you've really been opted-out
- enforcement: state AGs, ISPs, federal - centralised; too big though. they have better things to do with their time than bust spammers

- solution: marries legal/architectural/market
- legal: has two parts: (1) labels ("ADV" in the subject line)
- (2) a bounty
- (q: SEXUALLY-EXPLICIT tag is a label, already massively flouted by spammers. other labels would be flouted just as much.)
- architecture: filter code then blocks mails with "ADV"
- market: spammers would then have to incentivise people to receive their mail by sending offers they want (yeah right 😷
- enforcement: spam will only be sent if you can be paid, so "follow the money" – part of CAN-SPAM states "the business that benefits is responsible"
- market in enforcement: bounty hunters who identify label-less spam (ah). amateurs, not law enforcement, large population
- during CAN-SPAM development: labels were undesirable. Reason: "labels are too effective", because e.g. Amazon would have to have labelled their ads (because there was no distinction between opt-in and opt-out) and would be filtered
- fundamental problem: corruption due to vested interests lobbying (cf CAN-SPAM)
- sees difficulties in differentiating
- q: tracing spam to the business that benefits often involves getting forwarding addresses from e.g. a CGI script running on a server in the Ukraine. *needs* law-enforcement power to get that IMO. a: "yes, and law-enforcement power is available, and jurisdiction problems are easy" (not sure about that! at least for the non-LE bounty-hunter case)
- q: opt-in would have fixed it, like it has in Australia; but DMA keeps emasculating the laws into YOU-CAN-SPAM. a: agrees that there are multiple answers, but prefers not requiring opt-in across the board and uses the UCE definition as it allows political speech without adding to their costs. (I disagree, personally; the "UBE" definition works for me --jm)
- Jon Praed: enforcement requires tremendous resources, and in some cases you've got to get to that IP address within 7 days to get those logs, with LE power. This is not easy. Notes that spammer margins are incredibly low, and those bounties as a result would be small and/or hard to get.
- JP again: also suggests labels to label "good" commercial mail, personal mail, and then leave over "unknown" mail – which is then suspect. also suggests that the *headers* are the labelling, in reality.
- q: "special interests always seem to wipe out general interest on this issue in laws. what can we do, law-wise?" "my brand is pessimism", "there was this moment, when they passed CAN-SPAM, when legislators were keen to fix it – then the special interests came in".
- observation from audience: spots the parallel between UK and Pirate radio in the late 60's, which also passed a McCain anti-advertiser provision to deal with it.
- Dave Crocker: believes that the suggestion would result in little real effect on spammers, and quite a heavy hit on legit businesses

Hal Varian:

- "who is annoyed by spam?" one-fifth of US residents acknowledge buying products from spam; 77% considered spam an annoyance. That means 23-30% don't find it annoying! who are they?
- before federal DNC: multiple lists, state, DMA, company-specific
- federal DNC: lots more teeth; major fines, enforcement
- mapped DNC lists (with last 4 digits redacted, obtained via FOIA) against census data
- very popular in predom-asian areas
- income under $10k/yr: very low prob of signup
- income over $100k/yr: very high
- lots more interesting correlations, too many to write up
- income, education, number of kids are main significant determinators
- almost no corr between having internet and signing up for the DNC!
- est. signup rate for do-not-email: 31.5% (iirc)
- summary: telemktg, spam annoy same people
- 70% of variation in signups can be expl'd with only 4 vars: median income, presence of teens, education, presence of state list that was merged into fed DNC
- not many ppl used state lists, even though some were effective, and cheap. seemingly small costs can be big barriers
- DNC: effective because it had teeth ($11k fine), lots of publicity, and nationwide.
- q: "how do income and education correlate?" a: regression determines independent effects, accounts for correlation
- q: "do tmers target upper-middle-class?" a: yes, very much! q: "different to spam then, since spam is a lot more scattershot" a: agreed. also note that targeting sometimes aimed at *less* middle-class consumers, on the basis that they want to sucker people sometimes
- q on statistical reliability, answered, seems satisified
- q: about getting the NPA data via FOIA: apparently Telcordia asks for a fortune for that! also wondered if he'd checked in Europe. a: not yet, interested in the idea
- q: "what about using same techniques to find who benefits from spamming?" a: lots of interesting questions about looking at the spamming industry from an economic POV

Nicola Lugaresi: EU vs Spam - a legal response

- why legislate? social norms, market, self-regulation, self-help, have all failed; "code has failed"; law wants it chance to fail
- law probs: lack of jurisdiction and/or intl cooperation, lack of enforcement, lack of coord with other tools, bad laws
- EU law: 3 main goals: practical: fight spam; ethical: protect privacy, and state its relevance; political: don't just trail the US, lead sometimes!
- approach: optin and "soft opt-in" (transactional – in ctx of sale, similar products, same company; opp to object when email collected); no disguising identity; valid address to cease further comms
- other tools: labels, registries, spam boxes, codes of conduct – not convinced by any of these
- compared 2 antispam cases, one in Paris, a property case between spammer and AOL,MS; one in Napoli, a privacy case for an individual – E1000 damages and E750 costs paid to the individual!
- spam definition in EU: not bulk, just direct mktg not just "commercial". too narrow in my opinion
- not great against hard spam (proxy-abusers etc.): needs other approaches
- q: worried about definition of spam narrowed to direct mktg, and not bulk: a: agreed
- q: if I recall correctly not protecting corporate accounts, just individual "natural person" accounts at an ISP? a: not in Italian transposition at least, may be just in the Irish version (or my misreading 😷
- q: role accounts considered? a: his opinion, yes, in Italy, that's the case. good news

No-Email-Collection flag: Matthew Prince, Unspam LLC

- Lessig-influenced presentation style – "law professors get paid based on the number of slides produced"
- CAN-SPAM removes individual rights to sue spammers, one thing that's really been effective so far. not good news
- proposes a new meta tag: no-email-collection

- cute cat photo!
- missing a huge chunk of the spam pipeline; we're focussing on the proxy-to-recipient part of the chain. focus on the address scraping!
- definition of spam: hard. but harvesting email addresses: everyone agrees that's a no-no

- <meta name="no-email-collection" terms="[url of terms page] />

- http://www.unspam.com/noemailcollection
- project honeypot: like my script – email cookie addresses for scrapers
- create subdomains for honeypots and point them at Unspam's server! they'll collect the data
- generating a public corpus!
- code licensed under GPL
- http://www.projecthoneypot.org/ , http://www.unspam.com/
- q: "scraping through zombies?" a: yes, but that'll increase potential costs for spammers (hmm)
- q: "two classes of spammer lists: resold email addrs as well as scraped"? yes, but getting one class is useful
- q: "can a meta tag be enforceable? is a clickwrap license legally viable when it occurs between two computers?" a: if it's a community norm, that can improve legal viability; also CAN-SPAM specifically forbids scraping; also it may cause the spammer to think twice about this

Paula Bruening, CDT: Tech Responses to the Problem of Spam: Preserving Free Speech and Open Internet Values

- CAN-SPAM not entirely working
- worried about antispam tech hurting speech capabilities on the net
- concerned that "only popular speech will be delivered"
- what's key: tech must not be only part; devs must think of access issues; "let a thousand flowers bloom"
- q: "spam filters won't block websites" a: yes, but urgent updates do require email
- q: "is the CDT advocating political UBE?" a: good question, no answer

Barry Leiba, IBM: a multifaceted approach to spam filtering

- cowritten with Nathaniel Borenstein (woo!)
- "The Anti-Spam Gauntlet": describes exactly the SpamAssassin philosophy
- cooperation required with others
- open standards are required and are key to implementing anti-spam measures
- (I suggested open source as well as open stds
- Daniel: patents kill open standards; a from NSB: IBM are committed to "respecting IP rights" but not letting these stop open standards use by open source and other parties

SpamGuru:

- "enterprise-class anti-spam filter", but aren't we all
- centralized filter with personalized performance
- includes a "Bulk Mail Manager" for outbound *bulk* mail, interesting
- uses a "DNS analysis" step which sounds like it performs SPF checks
- DNS and domain analysis: check open relays, reverse DNS lookups and static IP tables; mail from dyn IPs; recency of dom registration; probabilistic analysis of Received trail
- bayes learning also feeds blacklist/whitelist; AWL is actually probabilistic
- "plagiarism detection": signature based really: "fast analysis of common k-grams"; learns from few examples; almost guaranteed not to be a FP; high FN rate though
- text classifier: Linear Discriminant: regularized linear classifier; approximates SVM
- Chung-Kwei (which rocks): really really effective: 86% with < 0.01% FPs on their test corpus
- test: corpus: 173k msgs, 130k spam, 42k good
- spam defn: UCE (not UBE). cleaned repeatedly
- combining algorithms: right with SpamAssassin dogma
- nice graph of aggregated performance; 96% with < 0.01% FPs
- SpamAssassin TODO: we need to add short-circuiting again!
- http://www.research.ibm.com/spam
- q: "what period, who were the 100 users?" a: users at IBM Watson
- q: how do you get your "recency of domain registration" data? a: straight from WHOIS

Richard Clayton: Stopping Spam by Extrusion Detection

- from demon.co.uk
- ISPs can spot smarthost load going up, and suspect that there's a spammer active
- insecure customers main problem for UK ISPs
- ISP's real problem: blacklisting of IP ranges and smarthosts; rapid action is req'd
- hard problem to solve: expensive to examine outgoing content; legal issues with blocking, and FP may cost you customers; volume is not good indicator!; "incorrect" sender domain doesn't indicate spam
- solution: spot delivery failure errors (due to user unknown, remote blocks) in smarthost logs
- heuristics: "too many" delivery failures (40/day sufficient); ignore "bounces" – have null <> return-path; ignore "mailing lists" (most dests work, few fail)
- when first turned on, was finding 40 infected customers *per day*!
- http://www.cl.cam.ac.uk/~rnc1/
- q: "direct-to-MX spam? trapping port 25?" a: no we don't do that and don't mind about that, as much as spammers using our smarthost and getting that blocklisted
- q: "sending outbound (or parts thereof) through SpamAssassin?" a: SpamAssassin is too expensive (in terms of load)
- q: "hair-trigger nature of listing?" a: it's not automatic. there's always a manual verification, and it's usually very obvious at that step

Resisting Spam Delivery through TCP damping:

- by default, TCP allows sender to control rate of flow; sender can achieve highest speed permitted by network
- TCP damping tries to reduce net efficiency at the receiver side; more time, more bandwidth, more CPU cycles

- low pain for recipients, high aggregated pain to spammers
- need to do this at TCP layer; higher and lower aren't useful
- even with tarproxy or similar, a smart spammer can blast the entire message to your TCP layer in one blat, even if you're tarpitting at the application layer
- damping: increase sending time (delaying TCP packets); consume network bandwidth (request more packets)
- increase delay: set adv_win = 0; fake congestion; delay outgoing ACKs (TCP conn terminates after 14 retries). cost at receiver: long idle TCP conn
- increase bandwidth costs: request more retrans.; request more ACKs – reuse sequence numbers, use seqs that won't be used in this conn; send packets in reverse order. cost: about 1:1 ratio
- used SpamAssassin at delivery time to estimate spamminess! mostly headers during early SMTP conversation, but you can use body rules before "250 Message Accepted for Delivery"
- q: economics. "increases senders costs, but not a transfer to the recipient." a: there are no existing techniques to do this, and TCP damping must work in existing system.
- q: if I was a spammer, and I figured out you were TCP damping, I'd ignore your advertised windows and blat entire message, hurting the network overall. a: sure, but hurting the spammer's bandwidth like this is worth it
- q: but this encourages broken TCP implementations. a: but a broken TCP stack still won't get their spam delivered
- q from John Levine: TurnTide does exactly this technique by narrowing the TCP window on the spammer's connections.
- q: why not just use delayed ACKs? a: because it's not entirely as effective as the other techniques

AOL hashing:

- I-Match: large corpus; lexicon generation
- intersection of document and lexicon gives signature
- trad I-Match lexicon generation: reject v frequent and hapaxes
- use "Mutual Information" as a measurement of fitness to avoid overlapping rules
- use multiple lexicons to avoid randomization from having an effect
- generate multiple lexicons, by removing random entries from an original lexicon
- also: distributional word clustering (Information Bottleneck) for lexicon selection (Terms with similar class distribution of P(spam|term))
- q: "'cluster' selection" – is that reports from live users? yep
- q: "FP rate?" a: very very low

Distributed, collaborative spam filtering:

- TCD, yay
- definition: "spam is email that the recipient is interested in receiving". we disagree, of course 😊
- P2P approach

Reputation network analysis for mail filtering:

- 75% of semweb data is FOAF files
- using web of trust
- a bit like http://web-o-trust.org/ , but not yet workable with email addrs since there's no spoofing protection

On attacking statistical spam filters:

- spammers wanted to evade bayes
- tokenization/obfuscation: turn out to be good spamsigns
- should not have used SpamArchive spam, due to its lack of headers, in my opinion; headers improve spam recognition greatly
- (correction: SpamArchive spam now does include headers, I missed that change – so that's not a big deal. Also, from talking to one author post-talk, he noted that they omitted the hdrs since the spam and ham each came from a different corpus, therefore a different set of hosts. If not ignored, those tokens would have been very obvious clues for the classifier.)
- pretty similar to http://www.cs.dal.ca/research/techreports/2004/CS-2004-06.pdf 😊