Some Security and Privacy issues in the 21st Century Internet

stephen.farrell@cs.tcd.ie

2022-01-12

(based on HEANET Conference talk from 2016)

TL;DR

The talk will describe some of the ways in which Internet security and privacy have evolved over the past couple of decades. In brief, we have seen improved deployment of security technologies in an increasingly hostile environment, yet we also continue to see the same mistakes being made e.g. the absence of small-device software update. The conclusion can be optimistic or pessimistic, depending on one's point of view. However, it is clear that Internet security and privacy issues will continue to create employment opportunities for defenders and attackers.

Summary

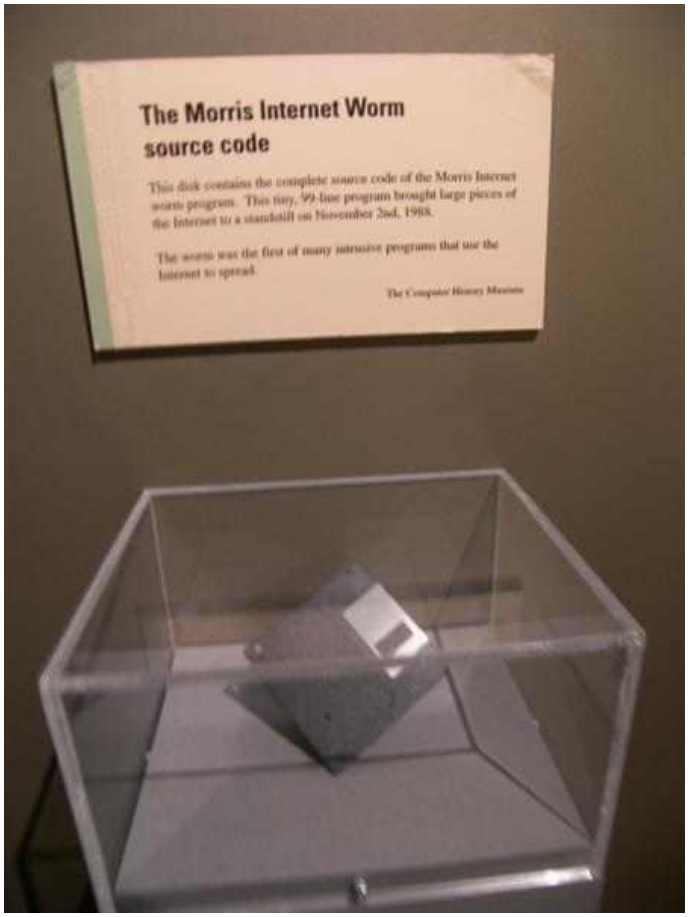
- Same mistakes get made over and over
- But we are finally starting to gain real experience in deploying deployable security technologies
- Same mistakes get made over and over
- Privacy is the real next challenge
- Same mistakes get made over and over
- But there are things you can do to help...

Let's start in the 19th Century

- A little before the Internet but…wires were tapped
  - http://bugsweeps.com/info/wiretap_short_history.html
  - https://www.counterpunch.org/2013/08/09/a-social-history-of-wiretaps-2/
- Basic law enforcement requirement:
  - Everything needs to be tappable
- Same as current lawful intercept
  - Not clearly a great plan

1988 – Morris worm

- First widespread worm in the wild
- Partitioned the Internet for days
- Sendmail debug mode, fingerd buffer overrun, password guessing
  - Buggy password guessing CPU consumption caused the DoS
- Worth reading the initial report as it describes things from first principles
  - https://spaf.cerias.purdue.edu/tech-reps/823.pdf



https://en.wikipedia.org/wiki/File:Morris_Worm.jpg

1989-date - Ransomware

- AIDS trojan
  - distributed via floppy disk to 20,000 attendees at WHO AIDS conference
  - @90th reboot, encrypted file names and demanded US$189 sent to Panama post office box

- Ransomware now a much bigger problem

https://www.gov.ie/en/news/ebbb8-cyber-attack-on-hse-systems/

https://ransomware.org/what-is-ransomware/the-history-of-ransomware/#evolution-of-ransomware

1993-today Name fun: digital.com

- Owned by: DEC, Compaq, HP, and now “Quality Nonsense Ltd.” of London, UK
  - https://digital.com is now some kind of review site
    - Even has a bit of history of the name https://digital.com/about/#section-7
  - https://betanews.com/2015/03/17/30-years-of-dotcom-what-became-of-the-first-100-domains/
- In 2015, of 1st 100 .com domains:
  - 52% still same
  - 24% dead
  - 18% redirect new company
  - 6% redirect same company

https://en.wikipedia.org/wiki/File:Digital_556-flattened4.svg

- One might speculate that in the long term less than 50% of DNS names will remain “good”
  - but almost all Internet security mechanisms depend on DNS names!

1999, 2016 – Crypto product survey

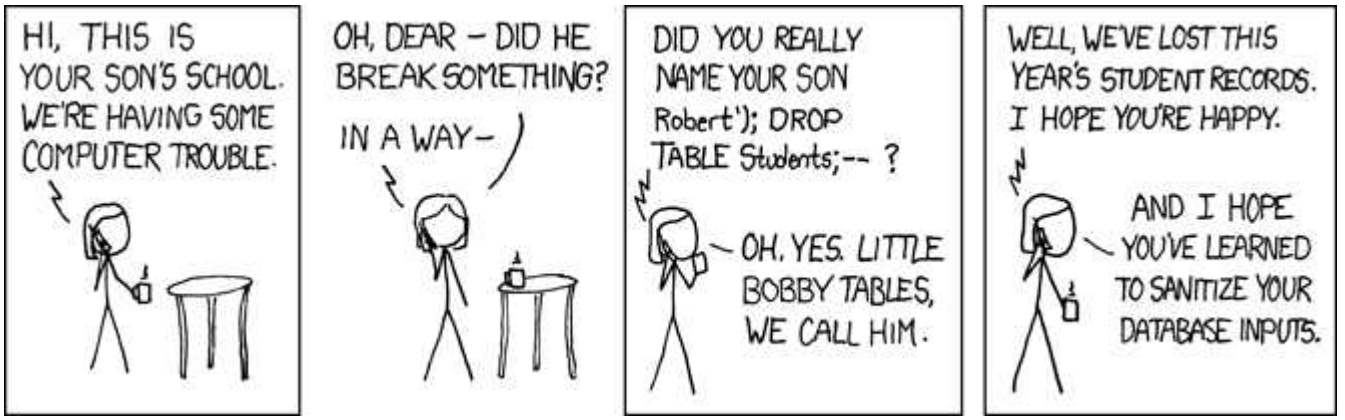
- Surveys done in 1999 and 2016 identifying cryptographic products (incl. OSS) available worldwide
  - Fewer in 2016, 546 vs. 805 “foreign,” but crypto is now a mainstream feature more than a product category
- Not clear surveys are commensurate, except for the intended affect on US policy related to cryptography
  - Any such laws are ultimately not a problem as mathematics is not nationalist!
  - They can be a PITA though
- https://cryptome.org/cpi-survey.htm
- https://www.schneier.com/academic/paperfiles/worldwide-survey-of-encryption-products.pdf

2003-date – DDoS Galore

- 2003: SCO case...
  - https://news.netcraft.com/archives/2003/12/10/ddos_takes_sco_site_down.html
  - Web site “down for 3 days” after ~64Mb/s (50,000 packets/s) SYN flooding
  - Our sympathies are where?
- 2016: Brian Krebs case…
  - https://krebsonsecurity.com/2016/09/krebsonsecurity-hit-with-record-ddos/
  - Journalist attacked by subject of article (who is in the business of DDoSing folks:-)
  - 620Gb/s attack, reportedly later 1Tb/s vs. Dyn
  - Botnet of crap devices, not a reflection attack
- DDoS is a cause of Internet centralisation forcing sites to use CDNs and similar, not really such a desirable thing

2003, 2013 – SQL Injection

- US FTC threatens petco.com with sanctions for leaking 500,000 user credit card details
  - https://news.netcraft.com/archives/2003/12/10/us_regulators_probe_security_lapses_at_retailers.html
  - In 2016, https://www.petco.com re-directed to http://www.petco.com even though they had TLS in place since at least 2013 (https://crt.sh/?q=petco.com) - They've since fixed at least that
- 2021: OWASP top-10: Still there @ #3!
  - https://owasp.org/Top10/



https://www.xkcd.com/327/

2003,2016 – openssl vulns/updates

- 2003: Various CVEs (bugs) in openssl reported, openssl updated, 50k web sites still using vulnerable old versions ~1 year after CVE
  - https://news.netcraft.com/archives/2003/11/03/vulnerable_versions_of_openssl_apparently_still_widely_deployed_on_commerce_sites.html
- Same old, same old in 2016
  - https://news.softpedia.com/news/big-surprise-companies-are-slow-to-patch-latest-openssl-flaw-504579.shtml
- OpenSSL team now much better supported
  - And… I might make that worse:-)
    - https://github.com/sftcd/openssl
- Not just OpenSSL though:
  - This one is from Jan 29 2021:
  - https://www.theregister.com/2021/01/29/severe_libgcrypt_bug/

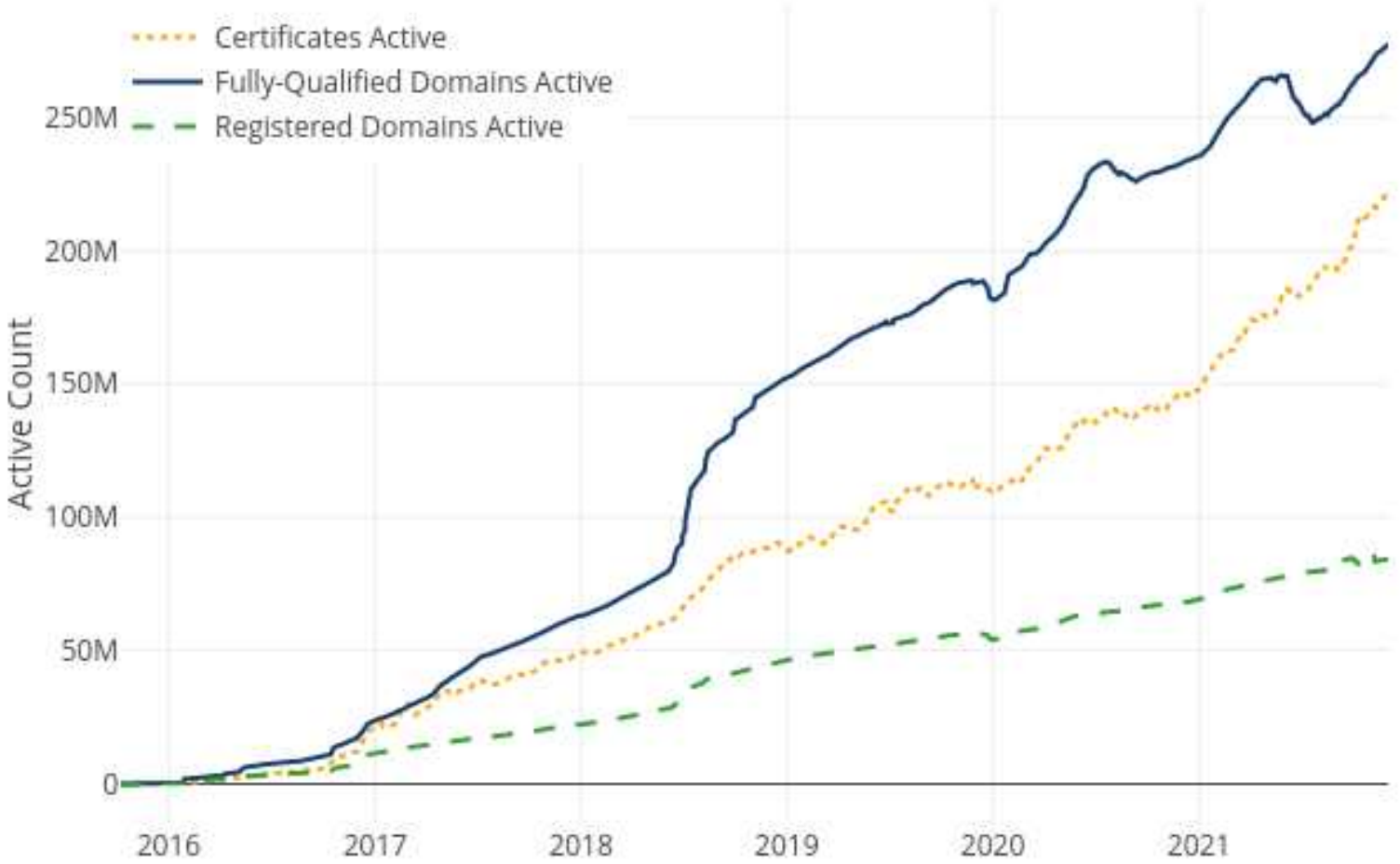
https://threatpost.com/openssl-fixes-critical-bug-introduced-by-latest-update/120851/

2003 – Weird CA Business Model

- A hoster offering “cheap” US$25/yr certs as a way to attract the kind of web site that uses SSL
  - https://news.netcraft.com/archives/2003/09/09/do_ssl_certificate_authorities_still_have_a_margin_generating_business_model.html
- The CA business model was always weird and still is
- Things like LetsEncrypt (2015) and acme have improved the web PKI a lot

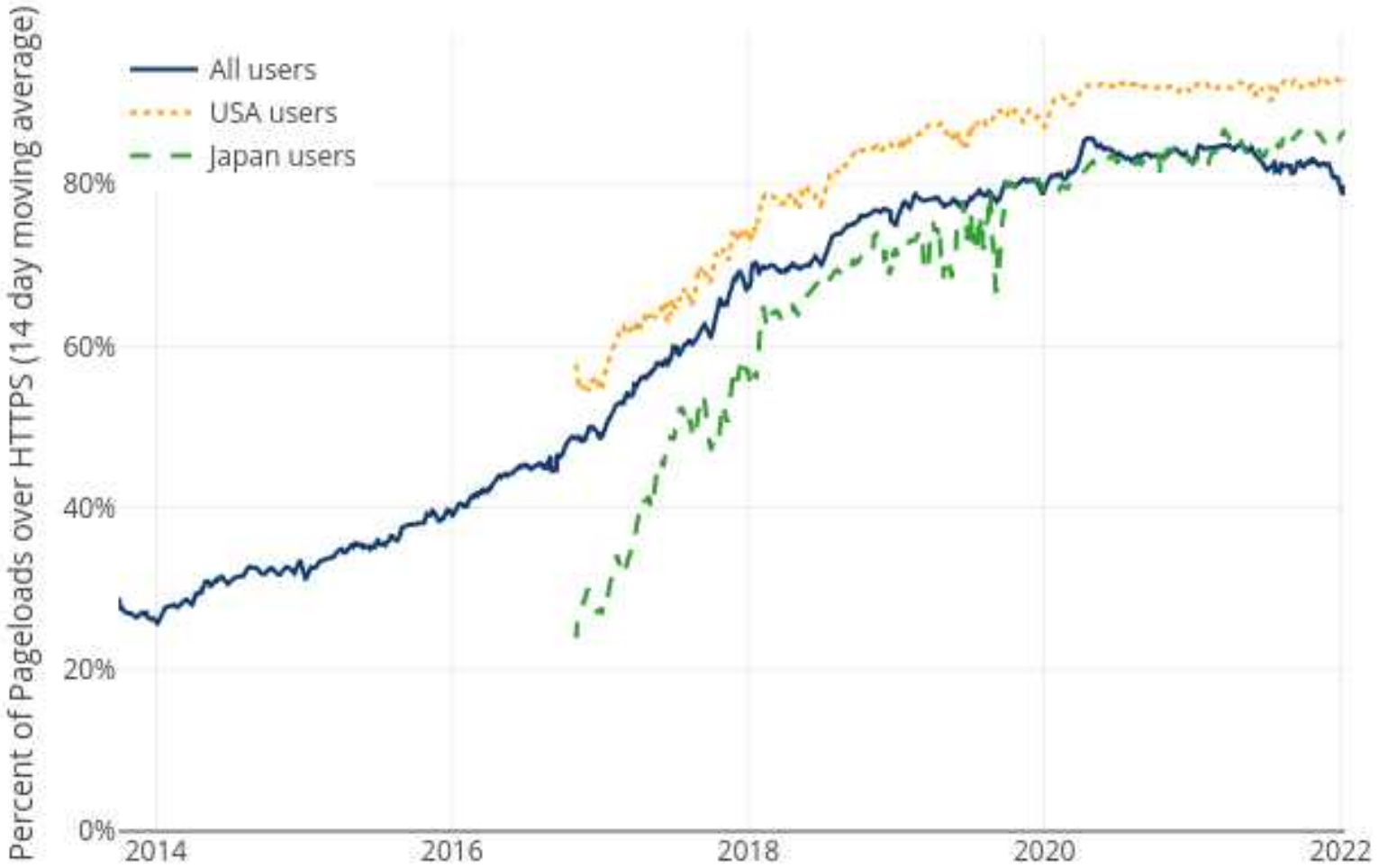
https://letsencrypt.org/

Letsencrypt Growth



https://letsencrypt.org/stats/

HTTPS Growth



https://letsencrypt.org/stats/ based on FF telemetry https://docs.telemetry.mozilla.org/datasets/other/ssl/reference.html

2003 – SSO around the corner (still)

- Large consortium of vendors establish a single-sign-on system and start to deploy that
  - https://news.netcraft.com/archives/2003/01/22/liberty_alliance_identity_server_launched.html
- This still happens, (Fido, WebAuthn) there'll always be another fashionable "federated" thing, maybe someday one will work out as planned
  - https://en.wikipedia.org/wiki/WebAuthn is I think the most recent
- To be fair, some stuff works: Eduroam, or login based on ${megacompany} credentials (FB, Google etc)

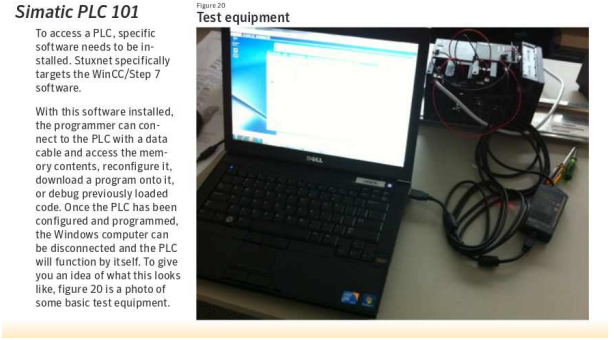
https://web.archive.org/web/19980529171444/http://www.sse.ie/prodserv.html

2007 - Netflix Competition

- Anonymised data sets published to allow researchers to improve delivery algorithms
- Correlation of review times with IMDB allow identification (with some embarrassment)
  - https://en.wikipedia.org/wiki/Netflix_prize#Privacy_concerns
  - https://arxiv.org/abs/cs/0610105
- Fine example of unexpected nature of some privacy issues. Many privacy issues however are utterly predictable if one spends a very short amount of time thinking about the topic.

2010-ish - Stuxnet

- Targeted controllers for (off-line) Iranian centrifuges involved in Uranium purification, so had to span air-gap
  - https://en.wikipedia.org/wiki/Stuxnet
  - https://www.wired.com/images_blogs/threatlevel/2010/11/w32_stuxnet_dossier.pdf
- Interesting part of Stuxnet is that the US essentially admitted it! That's pretty stupid really but also set an awful precedent for which we'll continue to pay for some time to come
- Malware attribution is almost never believable, even genetic linkage could/should be faked by those skilled in the art



Simatic PLC 101

To access a PLC, specific software needs to be installed. Stuxnet specifically targets the WinCC/Step 7 software.

With this software installed, the programmer can connect to the PLC with a data cable and access the memory contents, reconfigure it, download a program onto it, or debug previously loaded code. Once the PLC has been configured and programmed, the Windows computer can be disconnected and the PLC will function by itself. To give you an idea of what this looks like, figure 20 is a photo of some basic test equipment.

Figure 20
Test equipment

2013 - Snowdonia

- Partial timelines:
  - https://en.wikipedia.org/wiki/Global_surveillance_disclosure
  - https://www.theguardian.com/us-news/nsa
- My favourite:
  - https://www.theguardian.com/world/2014/feb/27/gchq-nsa-webcam-images-internet-yahoo
- My most interesting (politically):
  - https://theintercept.com/2018/02/17/gchq-belgacom-investigation-europe-hack/
  - Some political events between the UK and EU happened since;-)
- My most interesting (technically):
  - The short-range radar thing
  - https://en.wikipedia.org/wiki/NSA_ANT_catalog

https://en.wikipedia.org/wiki/Edward_Snowden

2016 – DNC spearphish

- Report on spearphising attempts aimed at US Democratic political party's 2016 presidential election campaign
  - https://www.secureworks.com/research/threat-group-4127-targets-hillary-clinton-presidential-campaign
- 10% of links clicked according to report
- Again, attribution is not convincing
- Role of DKIM signatures is interesting and not considered during development of DKIM

2018 – DNA Databases

- Law enforcement using stored DNA data to identify suspects' relatives
  - https://yro.slashdot.org/story/18/08/02/0253250/top-genetic-testing-firms-promise-not-to-share-data-without-consent?utm_source=rss1.0mainlinkanon&utm_medium=feed
- Genomic data (and similar) is sensitive now and may still be sensitive in >100 years
  - Very hard challenge for security mechanisms!

2018 - “Smart” speaker snooping

- Alexa “bug” causes bad stuff:
  - https://yro.slashdot.org/story/18/05/24/1633218/woman-says-alexa-device-recorded-her-private-conversation-and-sent-it-to-random-contact-amazon-confirms-the-incident?utm_source=rss1.0mainlinkanon&utm_medium=feed
- Religious practice via Alexa
  - https://www.theregister.co.uk/2018/05/24/pray_for_me_alexa/
  - What could possibly go wrong there? ;-)

2018 – Service provider “sharing”

- Cambridge Analytica
  - https://en.wikipedia.org/wiki/Cambridge_Analytica
- FB and tech “partners”
  - https://www.nytimes.com/2018/12/18/technology/facebook-privacy.html
- Android apps
  - https://privacyinternational.org/report/2647/how-apps-android-share-data-facebook-report
- US Telcos selling location data
  - https://arstechnica.com/?p=1438339

2018 leaks galore

- Marriott hotel group, losing up to 500M customer records, incl. passport data
  - https://www.theregister.co.uk/2018/11/30/marriott_starwood_hotels_500m_customer_records_hacked/
- Under armour: 150M records
  - https://www.reuters.com/article/us-under-armour-databreach/under-armour-says-150-million-myfitnesspal-accounts-breached-idUSKBN1H532W
- Too many to list really

2018 mega-scalers maybe not as secure as thought

- One natural reaction to security problems for smaller sites is to offload the problem to mega-scalers like Google or FB using their “login via” schemes
- Maybe they’re not so much better than others as we thought:
  - Google+ 500k user data leak
    - https://arstechnica.com/tech-policy/2018/10/google-exposed-non-public-data-for-500k-users-then-kept-it-quiet/
  - FB/Cambridge analytica again:
    - https://www.thesslstore.com/blog/facebook-data-leak-cambridge-analytica/

2020/2021 supply-chain again

- Solarwinds attack
  - https://www.csoonline.com/article/3601508/solarwinds-supply-chain-attack-explained-why-organizations-were-not-prepared.html
- Again – don't get hung up on attribution
- IMO the main fail was putting such a privileged device onto many networks
  - Did it need all those privs? Least-privilege is a very old security concept
- The build issue is secondary, but important

Moving on to the Future...

- Some people talk about the Internet of things but the Internet has always been made up of things
- Though maybe we can re-use the IoT marketing term…

Seen in a TCD cubicle...

The Other IoT

The Internet-of-Toilets will use the 5G network. In this IoT, each time a toilet is used, chemical (and perhaps DNA) analysis of the flushed content is done by the device and packets are sent to the network containing the results. IoT devices may be in the home, in businesses or provided by municipalities.

The data may be used for personalised healthcare services, for public health or, of course, advertising (imagine a pop up add over a pub urinal for just that medical condition you have;-). Insurance companies and lots of other businesses would likely be interested in the data. Service-selection and long term storage of the data present challenges.

These IoT devices are multi-user with no sophisticated user interfaces (except in Japan:-) and issues of identity, privacy, confidentiality and consent abound. Lawful intercept considerations would also arise - while societies may consider it ok for law enforcement to be able to listen to audio calls, it is not clear that the same is true for the packets emitted here, yet those are all bytes for the network."

Text is from about 2014 or so.

Some Security and Privacy Issues...

- Who controls the data generated?
- User interfaces and the lack thereof...
- Who is authorised to update devices, and how?
- Random numbers and crypto processing

Data Transport

- Devices generate data & send (secured via TLS?) to some host
- Today, there's no great way to get a (D)TLS server cert to use for that host unless the host has a DNS name
  - Leads to device → cloudy-server lock-in
- Challenge: find ways to authenticate and securely exchange keys between a small device and a host that the device-owner chooses
- Challenge: sometimes emitting a packet (encrypted or not) leaks privacy sensitive information
  - E.g. query sent to NTP server => person arrived home and stuff woke from suspend

Pervasive Monitoring

From RFC7258/BCP188: “Pervasive Monitoring is an Attack”

> Pervasive Monitoring (PM) is widespread (and often covert) surveillance through intrusive gathering of protocol artefacts, including application content, or protocol meta-data such as headers. Active or passive wiretaps and traffic analysis, (e.g., correlation, timing or measuring packet sizes), or subverting the cryptographic keys used to secure protocols can also be used as part of pervasive monitoring. PM is distinguished by being indiscriminate and very large-scale, rather than by introducing new types of technical compromise.

PM is not everything

- PM is far from the only security or privacy issue on which we need to work
  - Spam, malware, DDoS, …
  - But mitigations for PM can also help a lot with other problems
- Hypothesis: If we work to address PM, and prioritise services and mechanisms that mitigate PM and that are also effective against other attacks then we will be doing the “right thing”
- And the “we” there means us all, not just IETF

Data Storage

- So my devices produce a trickle/ton of data every day, what happens to that?
  - Helps some vendors monetize me?
  - Leaks to some bad actors eventually?
  - Gets deleted when vendors end-of-life service?
- Challenges:
  - Minimise the data that is ever sent/stored
  - Scrub stored data regularly (with what guarantee?)
  - Data portability?

(Lack of) User Interfaces

- Device discovery and provisioning are just hard unless you also have the device call-home to a vendor-selected site
  - Device offering web server? See “data transport”
- Challenge: we need ways to introduce devices to our networks that are acceptable to the owners/operators of those networks
  - While we can all develop some of these, it’s not clear what’s really going to work well enough at big enough scale for the range of devices that will be developed

Opportunistic Security RFC 7435

- Security mafia modus operandi has (in practice) been to define and implement security that works for higher security environments

  => often hard/expensive to deploy => often not used => cleartext often sent even when better options exist

- Opportunistic Security (OS) aims to evaluate these trade-offs on a connection-by-connection basis, explicitly allowing for e.g. unauthenticated endpoints for confidentiality (open-channel key exchange) as an option that is better than cleartext
- I (personally) hope that this concept is followed very often and is fleshed out to the point where we end up with a new security development approach that is based around OS
  - Not there yet: TLS deprecation of RC4 was interesting because of differing perspectives from web and mail folks about what conclusion to draw when following the OS approach

OS example: Deprecating RC4

- RC4 past sell-by date: agreed by all
- For the web ~15% of https sites were using TLS/RC4 (FF 2014 measurement)
  - When RC4 zapped 99% of those just picked a better option (AES, 3DES)
- SMTP+STARTTLS between MTAs
  - There is a widely deployed MTA that only does RC4, 3DES is buggy and won't work (so I'm told)
  - Zapping RC4 means emails will be sent in clear between MTAs when one is the buggy one
- So – which is better: deprecate RC4 entirely or add this and possibly other caveats?
  - IETF rough consensus was to deprecate entirely, but some mail folks were in the rough
- Interesting example implying conclusion from following OS protocol design pattern will depend on scope
  - OS requires us each to figure out some kind of utility or objective function and where those differ enough, different well meaning folks will reach different conclusions
- It is OK that it is harder to figure out what to do when following the OS approach

Updates...

- Non-updatable devices are a recipe for disaster
  - cf. Mirai – plenty of badly engineered devices will continue to be added to (and found on) the public Internet every day
- Challenge(s): Many issues with s/w update in this context, we hosted a workshop on this in TCD in June 2016, see RFC 8240 for details:
  - https://tools.ietf.org/html/rfc8240
  - Some progress since then – vendors now mostly seem to accept updates are needed at least
- Interesting “conclusion”: device update seems to call for all devices to support a way to “root” the device – both dangerous and seemingly necessary!

Crypto for small devices

- Some very small devices can’t play (D)TLS with the Internet, what then? Roll-your-own crypto? Urgh
- Issue with assurance that the crypto is “good”
  - Dual-ec fiasco http://dualec.org/
    - Bit sad that https doesn’t work for that!
- Challenge: initiatives like https://cryptech.is/ needed for smaller platforms
  - More generally: We may need larger players to help fund OSS that makes life easier for small developers of small devices
  - Key part of this: good (P/T)RNGs

What to do? (1)

- Consider privacy issues in your systems and networks and the data you make available
  - Avoid logging potentially sensitive data if you can
  - Find and delete old crap you no longer need
  - That means more work! But you should do it
- Encourage target diversity - Don't all use the same services all the time
  - Even if you're not a huge population, you can start trends

What to do? (2)

- Turn on crypto – ciphertext should be base assumption for new things
    - Consider the OS approach to make that easier
- Don't use new stuff without considering privacy implications
    - Data minimisation will save you some later leaks
- Help with better implementations

What to do? (3)

- Don't demand the impossible (and do nothing in the meantime)!
  - Encourage clean-slate work, but don't imagine it can all be deployed now – and only deployed things help
- Agitate (if that's your kind of thing:-)
- Consider privacy trade-offs when deploying e.g. IDS, anti-spam or malware detection technologies
- Be responsible and take the broader implications of your work into account before, while and after doing it
  - That includes research!

Summary

- There’s a bit of history, mostly bad:-)
- But there has been progress too, and we are getting better in some important respects
- IETF and others have consensus PM is an attack (RFC7258) and are working that problem, as a way to help get more and better deployment of security and privacy technologies
- When/if societies do decide that PM is as bad as it is, then the technical community should have in place the tools to effect that decision – you can help put those in place!

Thanks

Offline questions welcome too
stephen.farrell@cs.tcd.ie

2016 version:
https://down.dsg.cs.tcd.ie/heanet/
2022 version:
https://down.dsg.cs.tcd.ie/cs7053/lectures/heanet-farrell.pdf