Lecture 27: Wrap Up

Fall 2022 - MIT

6.1600

C-G, Kalai, Zeldovich

Wrap Up

- Case studies
  - * Authentication: OPM Hack
  - * Transport: Ps and Qs
  - * Platform: PS3
  - * Software: WannaCry
  - * Privacy: U.S. Census
- What's next?

Logistics

* Final exam: 9am - noon in 56-154
  - → Open laptop
  - → No network
* Course evaluations. PPP

Plan for this class

- Five case studies, one from each module of the course

- Goals
  1. To show you that you really have learned something this semester!
  2. To show how class topics intersect w/ real world.
  3. To entertain you. 

Let's start off with an authentication failure...

2014 - 2015

Office of Personnel Mgmt Hack: History

* To get sec clearance to see classified USG docs, fill out SF86
  - ↳ 136-page PDF!
  - Info on relationships, mental health, drug use, $, etc.
  - VERY invasive — one ostensible goal: understand blackmail risk
    - ↳ Does apparently happen!
* Roughly 2.8m have sec clearance, 1.6m conf/sec, 1.2m TS (CNN)
* Records stored in mainframe computer at OPM
  - "USG's HR dept"

June 2015: OPM announced that ≈20m background-check records breached

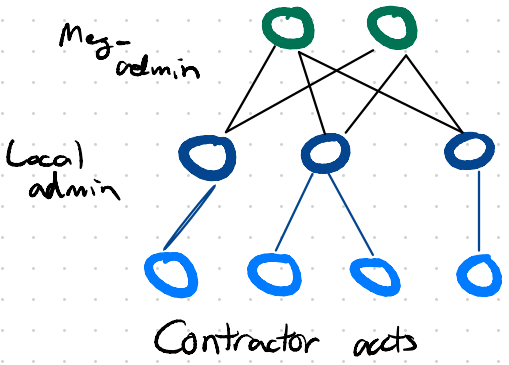
↳ NYT and others attribute to PRC govt

* Big problem for USG for two reasons:
  1. Blackmail info leaked
  2. CIA records NOT stored in OPM database.

OPM Hack: What Happened

[Various - Duo Blog, CS 558 write-up]

Actually two hacks — OPM discovered one, missed the second



1. Attacker got contractor's credentials.
   - ↳ Not clear how. Could be spear-phishing, etc.
   - ↳ No 2FA, smart card ~1% of OPM users
     - Why? Old systems. Expensive to upgrade

After getting credentials, attacker's goal is to compromise admin accts ("lateral movement")

Likely path...

2. Compromise root acct on local machine. ← R/W all files, etc.
   - ↳ Easy if old version of Windows

3. Work up to compromise of top admin credentials.
   - ↳ Exfiltrate data
   - "Pass the hash attack"

OPM Hack: Technical Details

STEP 2: Privilege escalation

- After compromise user acct, how to get root access?

```
at 16:05 /interactive "cmd.exe"
```

"at" is like a cronjob on Linux
↳ Run job at specified time. Runs as "system"!
(apparently only on very old versions of windows)

- Lots of other equally simple tricks. Some not as simple

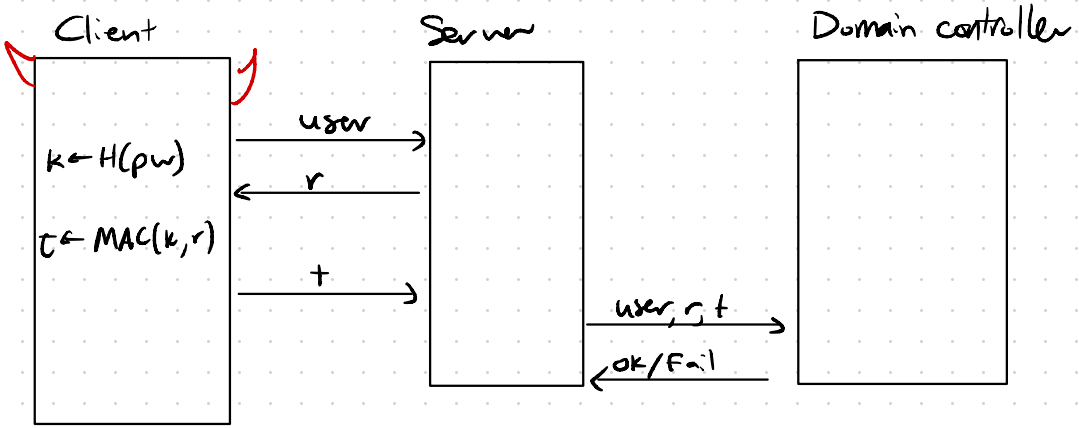
STEP 3a: Getting OPM admin password

- Once you have root (SYSTEM) privileges on local machine, need to get admin access on another
- Windows stores (user, Hash(password)) pairs of all logged-in users... like Kerberos.
  ↳ If admin is logged in, you can get their hashed password!!

STEP 3b: "Pass the hash"

(See CSS58 blog post, many other NTLM)

NTLM Authentication protocol...



Problem: Client doesn't need cleartext pw to log in!
↳ Compromising local machine allows lateral movement

→ Not super-clear how to defend against these while keeping backwards compatibility...

Lessons?

- Passwords are a terrible form of authentication
  - ↳ Ideally, ONLY used to auth user to their phone/laptop
- Always use 2FA
- Use signatures whenever possible

Also...

- Backwards compatibility is enemy of security
  - ↳ Isolation. (e.g. CT scanners running XP)
- Audit logging could have made cleanup easier

Transport: Mining your Ps & Qs

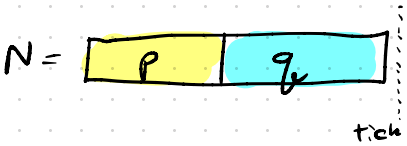
[Heninger, Durumeric, Wustrow, Halderman 2012]

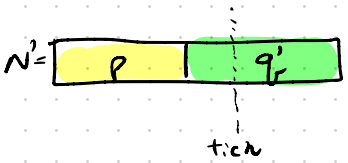
- Many hardware devices run SSH/TLS servers: routers, mgmt interfaces, doorbells, ...
- To run SSH/TLS, these devices need public keys.
- Most popular sig alg used to be RSA

Pub key: $N = p \cdot q$ ← Big primes

* Doesn't matter how RSA works — key idea (ha!) is that no one but signer knows factorization of N.

- When device boots for first time, gens RSA key
  - ↳ Where is randomness from?
    - ↳ Keyboard, Hard disk timings, clock, entropy saved from last boot
    - ↳ Embedded device may have few/none of these!
- Researchers scanned web, found many duplicate keys
  - ↳ Two devices start in same state → gen same key
- Also, some keys N and N' that share exactly one prime factor.





Transport: Ps and Qs

Given $N = p \cdot q$, $N' = p \cdot q'$ can factor both as

$$p = \gcd(N, N')$$

efficient via Euclid's algorithm (300 BC)

Lessons?

- Transport sec. is one of the triumphs of crypto
- When attacker gets data on the wire, often they get it b/c
  - * Implementation bug
  - * Non-use of encryption
  - * Compromise of endpoint

Remember: Encryption still leaks who you're talking to, when you're talking, what you're saying

Platform: Sony PS3 Hack [fail0verflow 27C3 talk]

Software bug / crypto implementation failure → Platform sec failure

- Sony PS3 originally could boot Linux/windows
  - ↳ Theory: avoid tariffs
  - * Popular in pre-GPU era for cheap HPC (dlog)
- Later on, Sony shipped update that disabled ability to run custom OS
  - * Used secureboot, much like iPhone
  - * Only boots Sony-signed OS

Sony used EC-DSA... morally equivalent to:

Should be a random long #

$$\sigma = (g^r,\ r + \text{Hash}(pk \| g^r \| m) \cdot sk) \pmod q$$

$$\sigma' = (g^r,\ r + \text{Hash}(pk \| g^r \| m') \cdot sk)$$

⇒ Sony's signatures leak their secret key.

- * Has happened to cryptocurrency wallets
- * Also embedded devices
- * Can also be if $r \leftarrow \text{Hash}(\text{time}())$

Sony PS3 Disk Encryption

* Store data on disk encrypted
  ↳ No integrity protection
* Sector data d at sector i written* as [AES-XTS]

* morally

$$Enc((k_1, k_2), i, d) := \begin{cases} r \leftarrow F(k_1, i) \\ ct \leftarrow r \oplus AES(k_2, r \oplus d) \end{cases}$$

→ No authentication!
→ Attacker gets "decryption oracle"

learn location of data on HD

  * Copy known plaintext (e.g. movie to drive)
  * Copy target ct on to these sectors
  * Read movie back

Why like this?

→ Full-disk enc is really for "stolen laptop" attack
→ Don't have extra space for auth info
  (also crash guarantees)

Lessons: - Have an update plan ... they did!
- Don't rely on secureboot for $
  ↳ Very hard to secure a device in "attacker's" hands

Software: WannaCry Ransomware

[Many articles in popular press]

Platform & Software sec failures

- Affected 100ks of computers
  - * Hospitals,
  - * Manufacturing (TSMC, Nissan, etc)
  - * Universities
  - * Telcos
  - ⋮
- Encrypts all juicy-looking files on all HDS (docx, pptx, etc.)
- Shows box demanding ransom payment in Bitcoin with "countdown" timer

↳ Caused LOTS of damage ($4bn?)
  ↳ Didn't raise much $$$ (maybe $300k) ← Possible to know b/c Bitcoin
    - * Hit mostly big enterprises
    - * Shoddy payment infrastructure (1 static Bitcoin addrs)

WannaCry: History

* Much speculation / unattributed sources

[See Checkpoint]

- Starts with NSA TAO (?)
  - * Developed an exploit "EternalBlue" in MSFT SMB server (used for file sharing)
    - * Combination of three bugs not reported to MSFT
      1) Invalid cast of struct
      2) Parser bug
      3) Allocation bug

      (see Checkpoint blog post ... Not at all trivial)
  - * Attacker can – over net – get RCE on Windows machines
  - * Possibly used for years (five years – report)
  - * Key component used to spread WannaCry

  ⇒ Software bug

- How did EternalBlue get out?

  [WSJ articles]

  - * NSA contractor (Harold Martin) took TBs of NSA data home with him ... motivation

    ⇒ Platform problem – least privilege?
  - * WSJ reports: Martin ran Kaspersky AV on computer with NSA data
  - * AV ships 'suspicious' files home for analysis
  - * WSJ reports: likely way exploit leaked

WannaCry: History (cont'd)

- After theft, Microsoft (March 14, 2017) issued patches for supported windows versions
  - ↳ Older windows unpatched for 2 months

Platform security - updates

- Shadow Brokers dump many exploits (including 0days) online — github repo
  - *Dump April 14, 2017
- Shows up in WannaCry May 12, 2017
  - *Suspected to be N. Korea. (why?)

WannaCry: Mechanics

(Logrhythm post)

1. Connects to website at random-looking addr
   - ↳ Exits if succeeds ("kill switch")
     - * Potentially used to check whether running in VM
     - * Used to help mitigate

2. Installs Tor, uses to connect to C2 infrastructure.
   - * C2 hosted at onion address

3. Encrypts all files that have fixed set of extensions.
   - ↳ Uses RSA + AES
     - Often source of bugs: * Power Worm - didn't save key
       - * Key reuse across users
     - ↳ Malware also updates itself!

4. Demands ransom be paid to one of four static Bitcoin addrs
   - * $300 then $600
   - * Problem: No automated way to match payment to machine/payer → No scale
   - * Problem: Spreading via SMB meant that it hit mostly enterprises — w/ better backups

5. Spreads itself
   - * Tries to connect to port 445 (SMB) on all IPs in local net (/24)
     - ↳ Random IPs on internet

What can we learn?

- That you should fill out your course evals?
- Less software => fewer bugs
  - ↳ Do you really need an SMB server?
- Any bug is a security bug!
  - ↳ Not obvious that a parser bug could cause such chaos
- Design for fast updates
  - ↳ Most machines affected were old (XP)
  - ↳ Many didn't get patch in time
  - ↳ Most secure s/w (e.g. Chrome) has a very aggressive update plan — not an accident
- Having a recovery plan (backup) is as important as trying to prevent attack.

Privacy: U.S. Census

- Performed every 10 years
  - ↳ Data used to allocated House seats
  - ↳ Used for redistricting

  13 USC §9: Census Bureau may not "make any publication... whereby the data furnished by ... any individual under this title may be identified."

- In 2020 census, bureau used D.P. to protect released data from de-identification
  - ↳ Used $\varepsilon = 19.61$.

  ⇒ If a bad event B was going to happen to you absent data release w.p. p, it will now happen w.p. $\leq e^{19.61} \cdot p$

$$\Pr[\text{Bad in world with data included}] \leq e^{19.61} \Pr[\text{Bad in world w/o data included}]$$

$e^{19.61} \approx 1{,}000{,}000$

$\Pr[\text{Bad in world w/o data included}] \leq \Pr[\text{struck by lightning next yr}] \approx 2^{-19}$

This ε ignores some non-private data releases (e.g. state pop)

- Still, amazing to see sophisticated privacy tech used in practice
- Alabama sued in March 2021 over use of D.P.
  - ↳ Tossed out... still may come up again

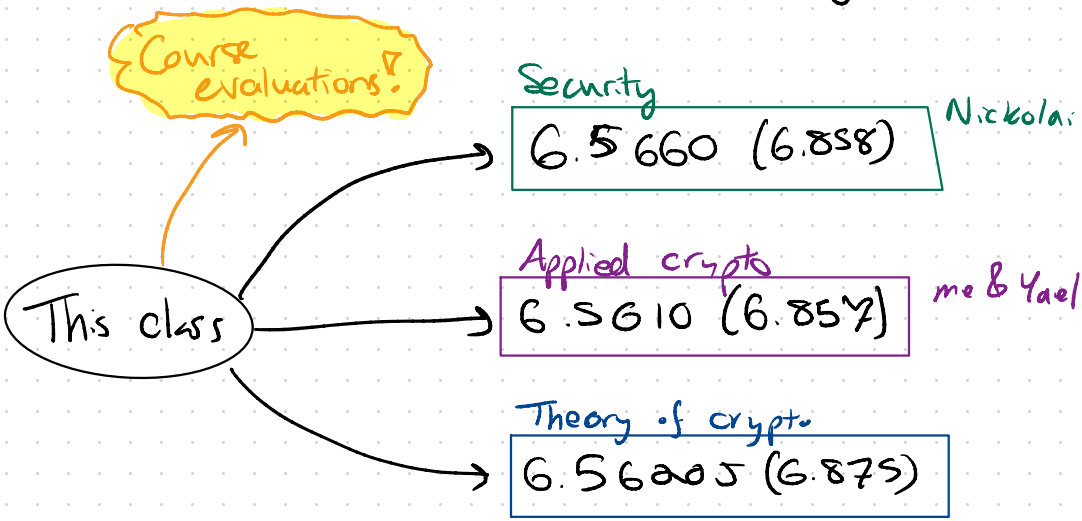
Lessons?

* Cryptography can help in many places
* Publishing data sets might not be one of them
  ↳ Tough trade-offs

* Though, I worry less about intentionally published datasets and more about unintentionally published ones (data breaches, etc.)
  ↳ There, secure systems-building tools + cryptography can help!

What's next for you?

If you're interested in learning more...



+ Lots of offerings at Harvard on privacy & security policy.
+ OS, randomized algs, ...

- Charles River Crypto Day
- CIS seminar (F 10:30am)
- Security seminar (Th 4pm)

↳ Research! Feel free to ping us