CSE 484 / CSE M 584: Computer Security and Privacy

Autumn 2019

Tadayoshi (Yoshi) Kohno
yoshi@cs.Washington.edu

Thanks to Dan Boneh, Dieter Gollmann, Dan Halperin, John Manferdelli, John Mitchell, Franzi Roesner, Vitaly Shmatikov, Bennet Yee, and many others for sample slides and materials ...

Announcements

- My office hours
  - 11/13 (Wed), 11:30am, CSE1 403
  - 11/20 (Wed), 2:30pm, CSE1 403
  - 11/27 (Wed), None
  - 12/4 (Wed), 12:30pm, CSE1 403
- HW 2 available (due 11/15)
- Final Project checkpoints looked great!

XML External Entities

XML External Entities

- Consider a web application that accepts XML input, parses it, and outputs the result (or includes untrusted input in XML documents)

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE foo [
  <!ELEMENT foo ANY>
  <!ENTITY bar "World">
]>
<foo>
  Hello &bar;
</foo>
```

- Parses as

  Hello World

But What About

- Consider an attacker uploading this XML document

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE foo [
<!ELEMENT foo ANY >
<!ENTITY xxe SYSTEM "file:///etc/passwd" >]>
<foo>&xxe;</foo>
```

- Attacker attempting to extract information from server

But What About

- Consider an attacker uploading this XML document

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE foo [
<!ELEMENT foo ANY >
<!ENTITY xxe SYSTEM "https://192.168.1.1/private" >]>
<foo>&xxe;</foo>
```

- Attacker attempting to probe a private network

But What About

- Consider an attacker uploading this XML document

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE foo [
<!ELEMENT foo ANY >
<!ENTITY xxe SYSTEM "file:///dev/random" >]>
<foo>&xxe;</foo>
```

- Attacker attempting a DoS by including a potentially never-ending file

Why Call “Server Side Request Forgery?”

What to Do?

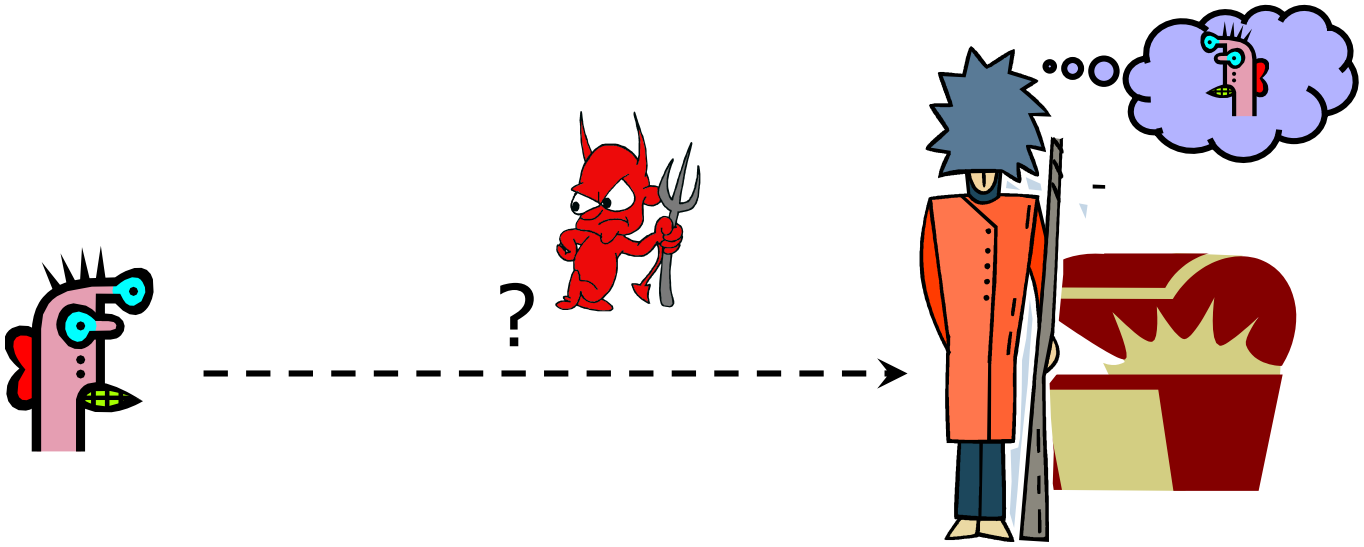
- Use less complex data formats, such as JSON
- Disable XML external entities and DTD processing in all XML parses
- Whitelist-based server-side input validation

- OWASP very useful source here as well

Authentication

Another “Ten Most Critical Web Application Security Risks”

Basic Problem



How do you prove to someone that you are who you claim to be?

Any system with access control must solve this problem.

Many Ways to Prove Who You Are

- What you know
  - Passwords
  - Answers to questions that only you know
- Where you are
  - IP address, geolocation
- What you are
  - Biometrics
- What you have
  - Secure tokens, mobile devices

Passwords and Computer Security

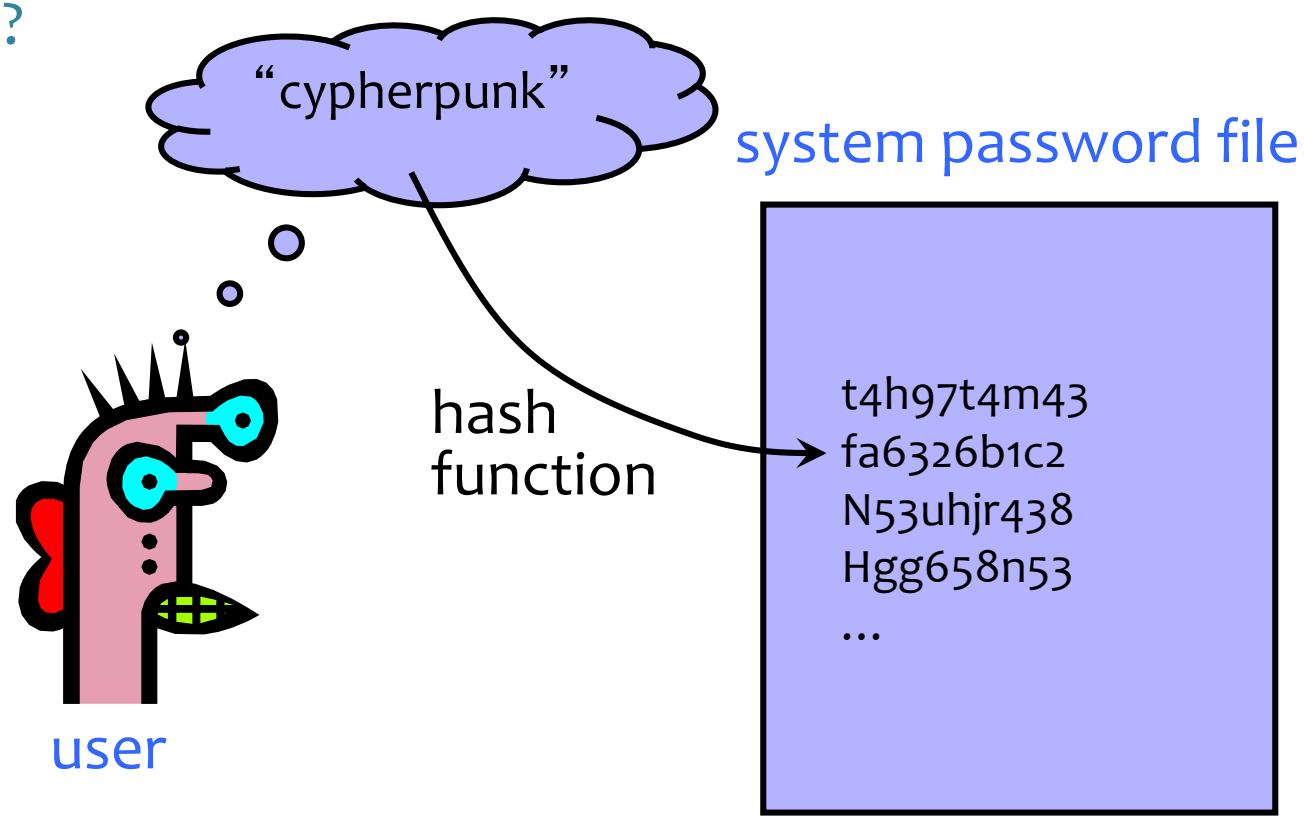
- In 2012, 76% of network intrusions exploited weak or stolen credentials (username/password)
  - Source: Verizon Data Breach Investigations Report
- A first step after any successful intrusion: install sniffer or keylogger to steal more passwords
- A second step: run cracking tools on password files
  - Cracking needed because modern systems usually do not store passwords in the clear (how are they stored?)
- In Mitnick's "Art of Intrusion" 8 out of 9 exploits involve password stealing and/or cracking

Password Storage

- Recall discussions from crypto section
  - Don't store plaintext passwords
  - Don't use encrypted passwords
  - Use hashed passwords
  - Hash a salt along with the password, and store the salt and the hashed salt+password on the server

UNIX-Style Passwords

- How should we store passwords on a server?
  - In cleartext?
  - Encrypted?
  - Hashed?



Password Hashing

- Instead of user password, store H(password)
- When user enters password, compute its hash and compare with entry in password file
  - System does not store actual passwords!
  - System itself can't easily go from hash to password
    - Which would be possible if the passwords were encrypted
- Hash function H must have some properties
  - One-way: given H(password), hard to find password
    - No known algorithm better than trial and error
  - "Slow" to compute

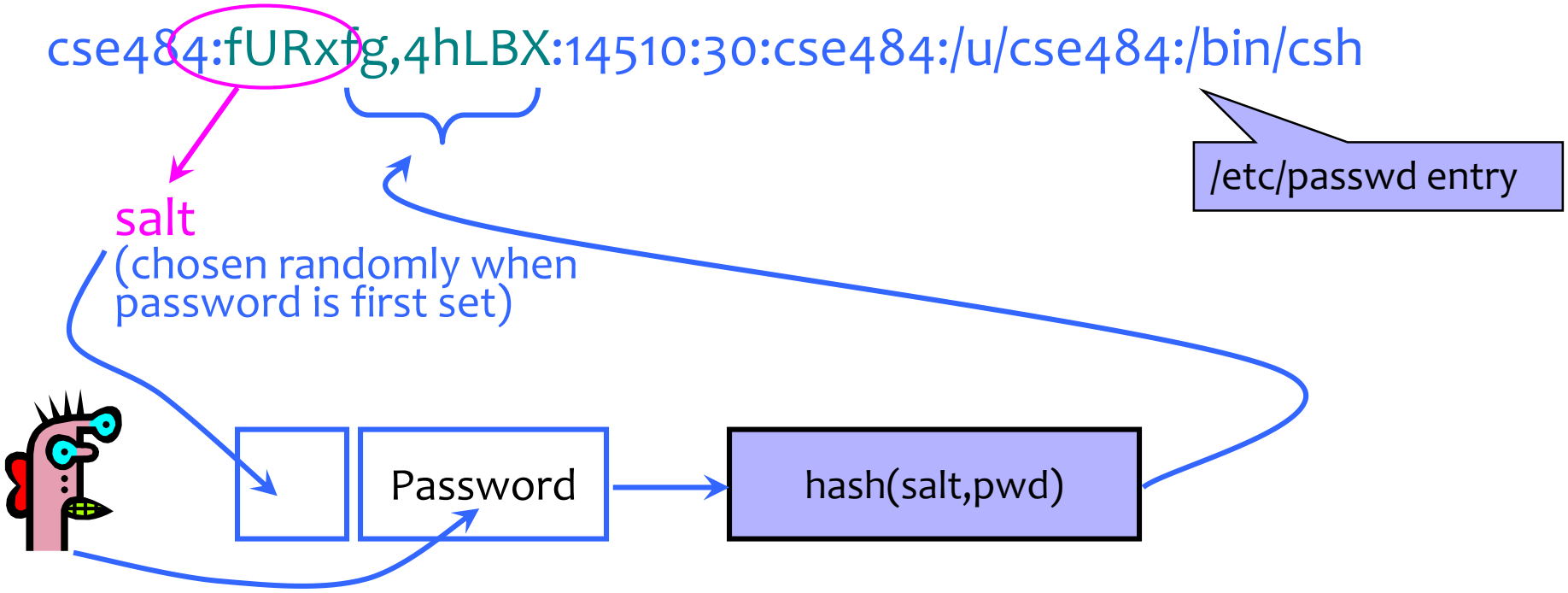
UNIX Password System

- Approach: Hash passwords
- Problem: passwords are not truly random
  - With 52 upper- and lower-case letters, 10 digits and 32 punctuation symbols, there are 94^8 --- 6 quadrillion possible 8-character passwords (~2^{52})
  - **BUT:** Humans like to use dictionary words, human and pet names --- 1 million common passwords

Dictionary Attack

- Dictionary attack is possible because many passwords come from a small dictionary
  - Attacker can pre-compute H(word) for every word in the dictionary – this only needs to be done once!
    - This is an offline attack
    - Once password file is obtained, cracking is instantaneous
  - Sophisticated password guessing tools are available
    - Take into account freq. of letters, password patterns, etc.

Salt



- Users with the same password have <u>different</u> entries in the password file
- Offline dictionary attack becomes much harder

Advantages of Salting

- Without salt, attacker can pre-compute hashes of all dictionary words once for all password entries
  - Same hash function on all UNIX machines
  - Identical passwords hash to identical values; one table of hash values can be used for all password files
- With salt, attacker must compute hashes of all dictionary words once for each password entry
  - With N-bit random salt, same password can hash to 2^N different hash values
  - Attacker must try all dictionary words **for each salt value** in the password file
- Pepper: Secret salt (not stored in password file)

Other Password Security Issues

- Keystroke loggers
  - Hardware
  - Software (spyware)
- Shoulder surfing
- Same password at multiple sites
- Broken implementations
  - TENEX timing attack

Examples from One Company

KeyGrabber Forensic Keylogger Cable & Module

Hardware keylogging for pros

The **KeyGrabber Forensic Keylogger** is a series of specialized hardware keyloggers with flash drive access, aiming at minimizing the risk of exposure. They diverge from the classic USB adapter shape, making them nearly impossible to locate. Feature-wise they inherit the entire functionality of the KeyGrabber Keylogger.

The **KeyGrabber Forensic Keylogger Cable** is an ultra-compact USB keylogger hidden inside a USB extension cable. From the outside, the USB cable doesn't differ from any ordinary cable in common use, and draws no attention. Available as the standard and Pro version.

Even More Issues

- Usability
  - Hard-to-remember passwords?
  - Carry a physical object all the time?
- Denial of service
  - Attacker tries to authenticate as you, account locked after three failures
- Social engineering

Default Passwords

- Examples from Mitnick's "Art of Intrusion"
  - U.S. District Courthouse server: "public" / "public"
  - NY Times employee database: pwd = last 4 SSN digits
- Mirai IoT botnet
  - Weak and default passwords on routers and other devices

Weak Passwords

- RockYou hack
  - "Social gaming" company
  - Database with 32 million user passwords from partner social networks
  - Passwords stored in the clear
  - December 2009: entire database hacked using an SQL injection attack and posted on the Internet
  - One of many such examples!

Weak Passwords

- RockYou hack

Password Popularity – Top 20

Rank	Password	Number of Users with Password (absolute)
1	123456	290731
2	12345	79078
3	123456789	76790
4	Password	61958
5	iloveyou	51622
6	princess	35231
7	rockyou	22588
8	1234567	21726
9	12345678	20553
10	abc123	17542

Rank	Password	Number of Users with Password (absolute)
11	Nicole	17168
12	Daniel	16409
13	babygirl	16094
14	monkey	15294
15	Jessica	15162
16	Lovely	14950
17	michael	14898
18	Ashley	14329
19	654321	13984
20	Qwerty	13856

Password Usability

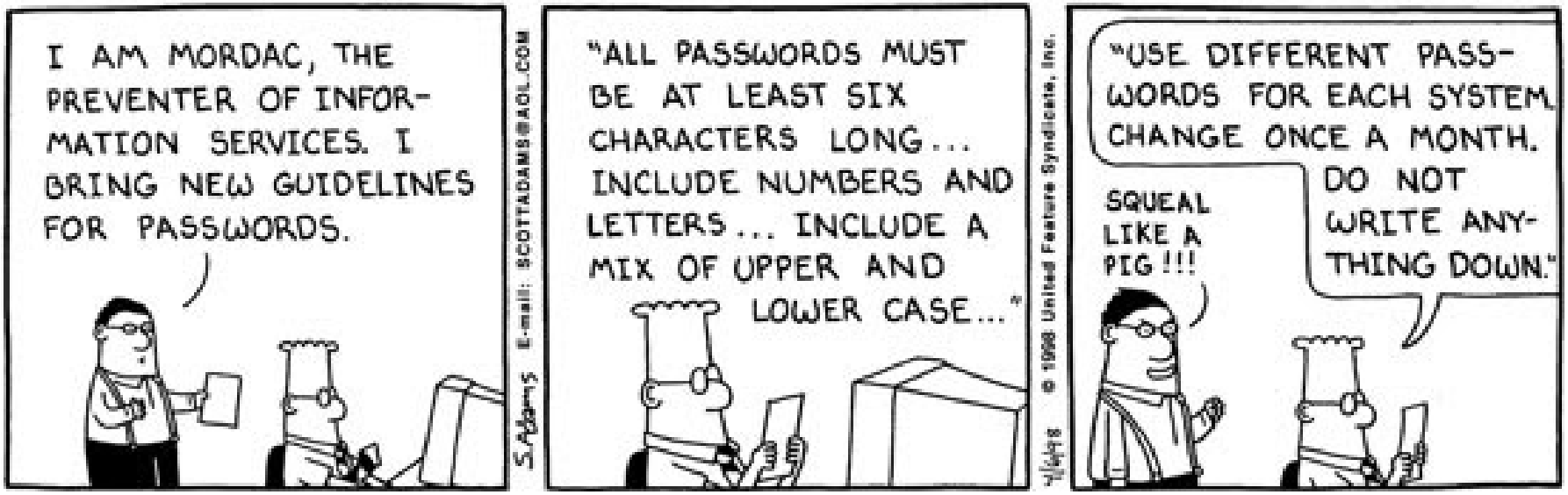


Image from http://www.interactivetools.com/staff/dave/damons_office/

More Password / Authentication Issues

- Credential Stuffing (using stolen credentials on other sites)
- Website permits brute force / automated guesses
- Not supporting multi-factor authentication (upcoming slides)
- Weak password recovery mechanisms (next slides)
- Application timeouts too long

Recovering Passwords

Palin E-Mail Hacker Says It Was Easy

By Kim Zetter ✉ September 18, 2008 | 10:05 am | Categories: Elections, Hacks and Cracks

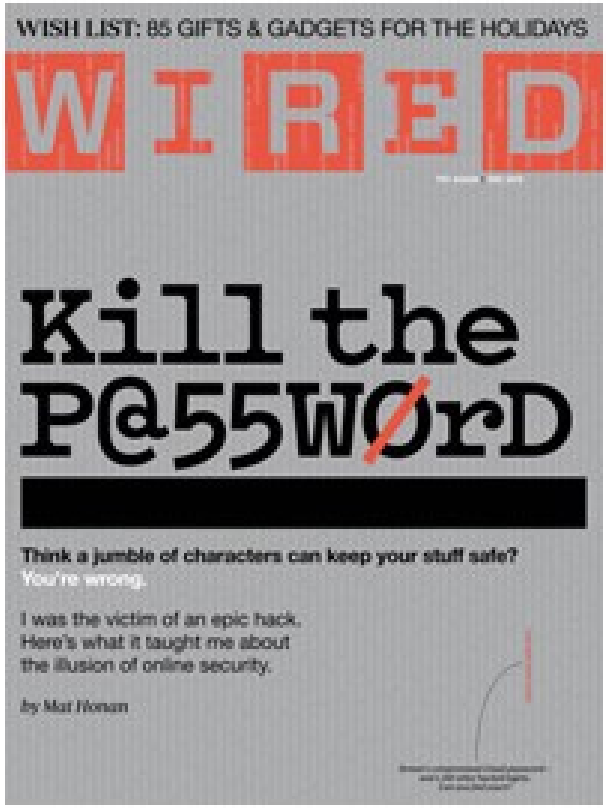
A p
obt
priv
sup
rev
too
Re

after the password recovery was reenabled, it took seriously 45 mins on wikipedia and google to find the info, Birthday? 15 seconds on wikipedia, zip code? well she had always been from wasilla, and it only has 2 zip codes (thanks online postal service!)

the second was somewhat harder, the question was "where did you meet your spouse?" did some research, and apparently she had eloped with mister palin after college, if youll look on some of the screenshits that I took and other fellow anon have so graciously put on photobucket you will see the google search for "palin eloped" or some such in one of the tabs.

I found out later though more research that they met at high school, so I did variations of that, high, high school, eventually hit on "Wasilla high" I promptly changed the password to popcorn and took a cold shower…

Wired Cover Story (Dec 2012)

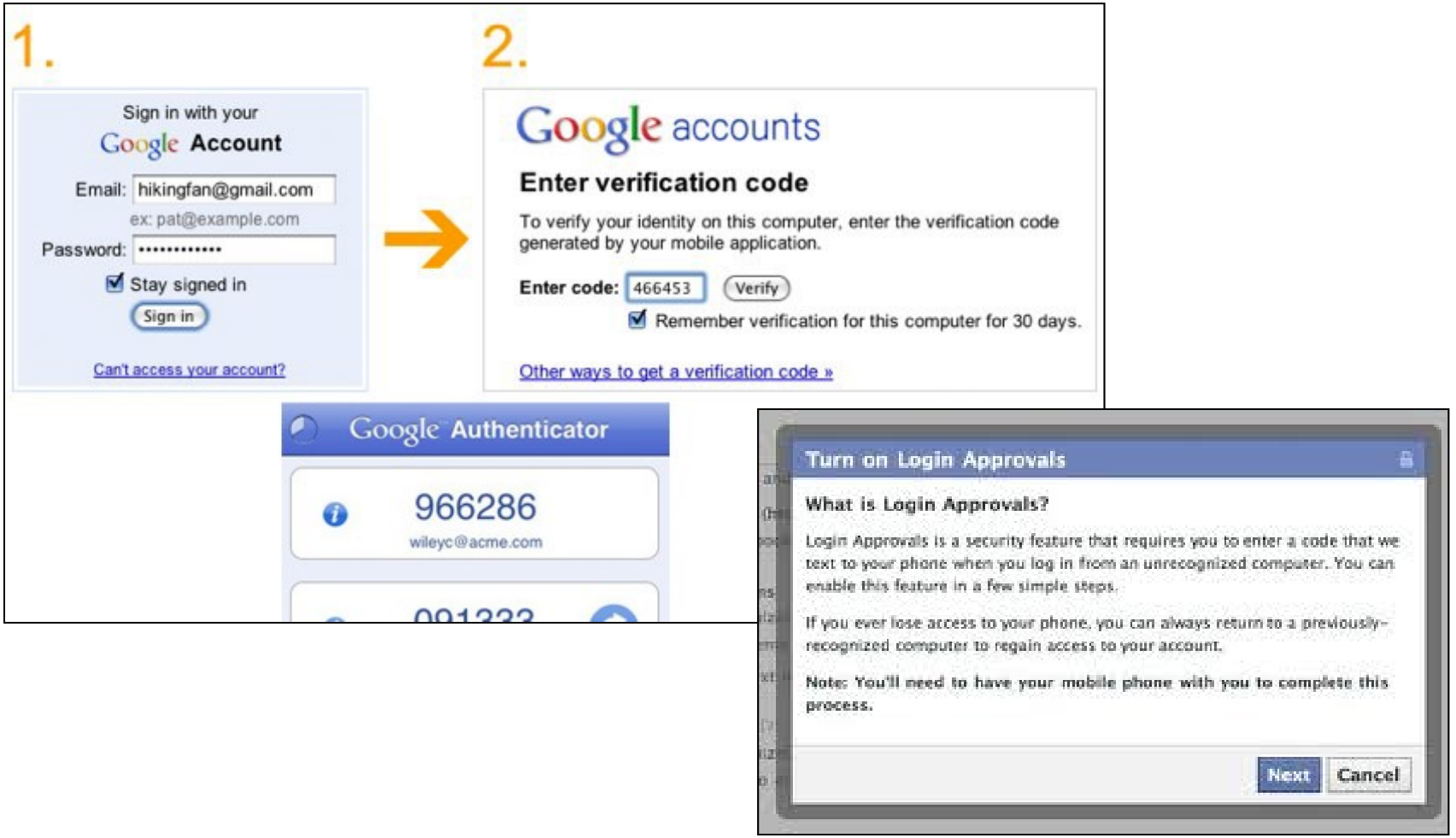


"This summer, hackers destroyed my entire digital life in the span of an hour. My Apple, Twitter, and Gmail passwords were all robust—seven, 10, and 19 characters, respectively, all alphanumeric, some with symbols thrown in as well—but the three accounts were linked, so once the hackers had conned their way into one, they had them all. They really just wanted my Twitter handle: @mat."

Improving(?) Passwords

- Add biometrics
  - For example, keystroke dynamics or voiceprint
- Graphical passwords
  - Goal: easier to remember? no need to write down?
- Password managers
  - Examples: LastPass, built into browsers
  - Can have security vulnerabilities…
    - https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/silver
    - https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/li_zhiwei
- Two-factor authentication
  - Leverage phone (or other device) for authentication

Multi-Factor Authentication



FIDO + Hardware Two Factors

Google Advanced Protection

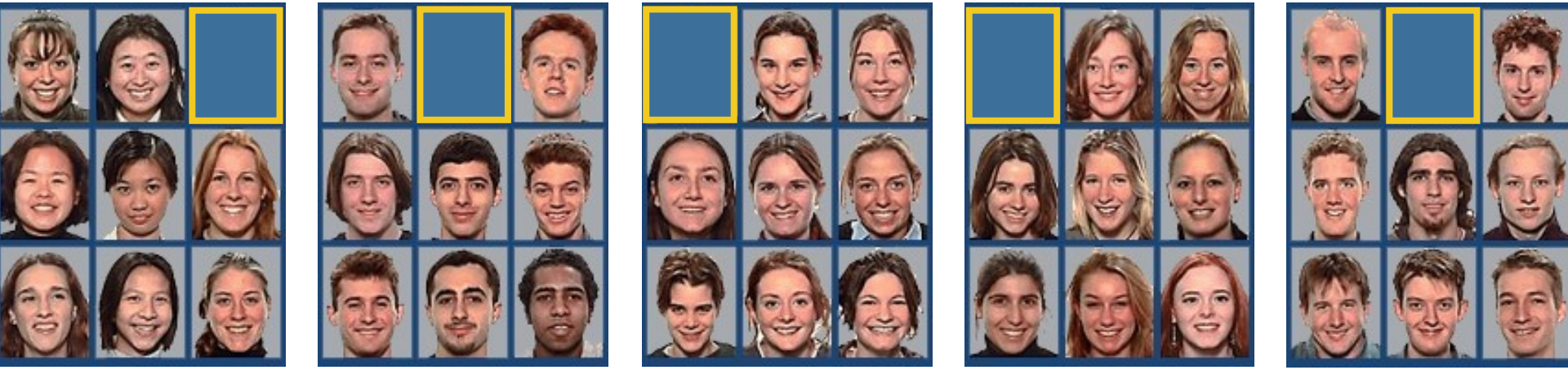
Google's strongest security for those who need it most

The Advanced Protection Program safeguards the personal Google Accounts of anyone at risk of targeted attacks – like journalists, activists, business leaders, and political campaign teams.

- More than just 2-factor; additional protection
- Does need two hardware tokens

Graphical Passwords

- Many variants… one (very early!) example: Passfaces
  - Assumption: easy to recall faces



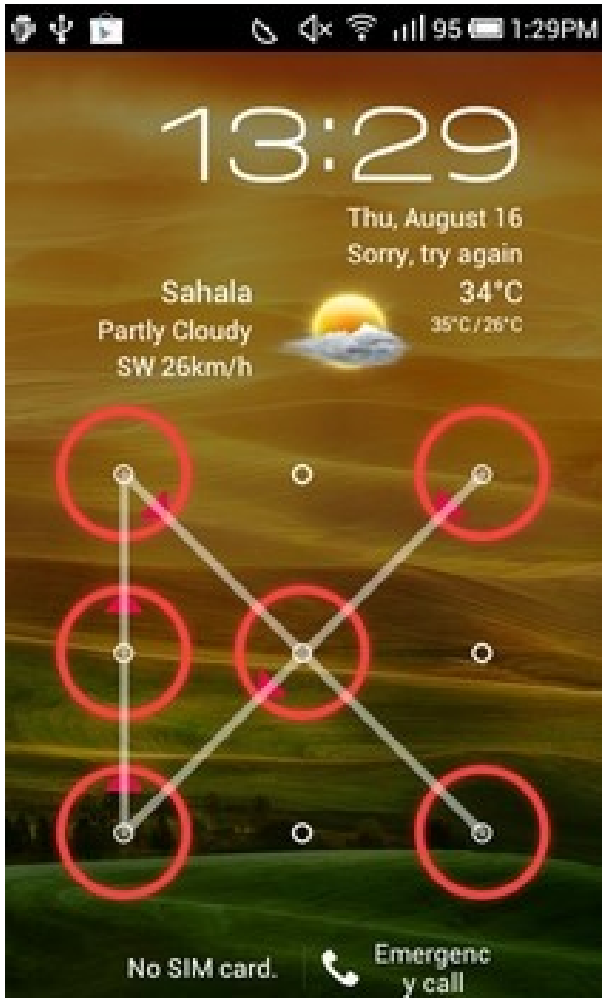
  - Problem: to make passwords easy to remember, users choose predictable faces

Graphical Passwords

- Another variant: draw on the image (Windows 8)

- Problem: users choose predictable points/lines

Unlock Patterns



- Problems:
  - Predictable patterns (sound familiar by now??)
  - Smear patterns
  - Side channels: apps can use accelerometer and gyroscope to extract pattern!

What About Biometrics?

- Authentication: **What you are**
- Unique identifying characteristics to authenticate user or create credentials
  - Biological and physiological: Fingerprints, iris scan
  - Behaviors characteristics - how perform actions: Handwriting, typing, gait
- Advantages:
  - Nothing to remember
  - Passive
  - Can't share (generally)
  - With perfect accuracy, could be fairly unique

Issues with Biometrics

- Private, but not secret
  - Maybe encoded on the back of an ID card?
  - Maybe encoded on your glass, door handle, …
  - Sharing between multiple systems?
- Revocation is difficult (impossible?)
  - Sorry, your iris has been compromised, please create a new one…
- Physically identifying
  - Soda machine to cross-reference fingerprint with DMV?
- Birthday paradox
  - With false accept rate of 1 in a million, probability of false match is above 50% with only 1609 samples

Stepping Back

- What is the threat model?
  - Someone with access to your physical possessions (e.g., key logger, steal written password book)
  - Someone across the Internet (e.g., who compromises one or multiple sites)
- What "costs" are one willing to expend?
  - Usability
  - Legal protection (e.g., passwords vs biometrics and the law)
- Keep in mind password recovery mechanisms