

Web Security and Special Topics

Question 1 *Boogle*

Boogle is a social networking website that's looking into expanding into other domains. Namely, they recently started a map service to try their hand at fusing that with social media. The URL for the main website is <https://www.boogle.com>, and they want to host the map service at <https://maps.boogle.com>.

- (a) Describe how to make a cookie that will be sent to only Boogle's map website and its subdomains.

- (b) How can Boogle ensure that cookies are only transmitted encrypted so eavesdroppers on the network can't trivially learn the contents of the cookies?

- (c) Boogle adds the ability for users to check in to locations on maps.boogle.com, but discovers an XSS vulnerability that slipped through QA. Name a hotfix they can do to prevent scripts from stealing cookies using XSS.

- (d) Some of the XSS attacks are scraping sensitive information from the map site, like user emails. The security team wants to know the scope of the vulnerability. Can attackers use XSS to also scrape sensitive information from the main site, <https://www.boogle.com>? Explain why or why not.

- (e) Boogle wants to be able to host websites for users on their servers. They decide to host each user's website at [https://\[username\].boogle.com](https://[username].boogle.com). Why might this not be a good idea?

- (f) Propose an alternate scheme so that Boogle can still host other users websites with less risk, and explain why this scheme is better.

Note: It is okay if the user sites interfere with each other, as long as they cannot affect official Boogle websites.

Question 2 *Tracking*

Let's say the web-page at `http://cute-puppies.com` looks like the following:

```
<html>
  <body>
    <p>Here is a GIF of puppies</p>
    
    <script type="text/javascript"
      src="http://yahoo.com/analytics.js"></script>
    <script type="text/javascript"
      src="https://google.com/analytics.js"></script>
  </body>
</html>
```

Note that `google.com` is loaded over HTTPS, whereas `yahoo.com` is loaded over HTTP.

Alice uses Mozilla Firefox on her laptop running Microsoft Windows. In her first browser tab, she has `https://berkeley.edu` open. In a second tab, she opens `http://cute-puppies.com`. In a third tab, she opens `http://cute-puppies.com` once again.

Assume that no two entities share information out of band and `cute-puppies.com` and `image-host.com` keep cookies for users. Each of the parts below are independent.

- (a) Assuming Alice does not use any tracking protection, which entities know that the same person visited `cute-puppies.com` twice?
- | | |
|--|--------------------------------------|
| <input type="checkbox"/> <code>cute-puppies.com</code> operators | <input type="checkbox"/> Microsoft |
| <input type="checkbox"/> <code>yahoo.com</code> operators | <input type="checkbox"/> Mozilla |
| <input type="checkbox"/> <code>google.com</code> operators | <input type="checkbox"/> Alice's ISP |
| <input type="checkbox"/> <code>image-host.com</code> operators | <input type="checkbox"/> UC Berkeley |
- (b) Assume Alice opted in for a privacy service run by her ISP. This privacy service blocks analytics scripts based on a URL-based blacklist (not host-based). Which entities know that the same person visited `cute-puppies.com` twice?
- | | |
|--|--------------------------------------|
| <input type="checkbox"/> <code>cute-puppies.com</code> operators | <input type="checkbox"/> Microsoft |
| <input type="checkbox"/> <code>yahoo.com</code> operators | <input type="checkbox"/> Mozilla |
| <input type="checkbox"/> <code>google.com</code> operators | <input type="checkbox"/> Alice's ISP |
| <input type="checkbox"/> <code>image-host.com</code> operators | <input type="checkbox"/> UC Berkeley |
- (c) Assume Alice uses a browser plugin. The browser-plugin blocks the analytics scripts based on a URL-based blacklist (not host-based). Which entities know that the same person visited `cute-puppies.com` twice?

- | | |
|--|--------------------------------------|
| <input type="checkbox"/> <code>cute-puppies.com</code> operators | <input type="checkbox"/> Microsoft |
| <input type="checkbox"/> <code>yahoo.com</code> operators | <input type="checkbox"/> Mozilla |
| <input type="checkbox"/> <code>google.com</code> operators | <input type="checkbox"/> Alice's ISP |
| <input type="checkbox"/> <code>image-host.com</code> operators | <input type="checkbox"/> UC Berkeley |

(d) Assume Alice uses a VPN run by UC Berkeley. Which entities know that the same person visited `cute-puppies.com` twice?

- | | |
|--|--------------------------------------|
| <input type="checkbox"/> <code>cute-puppies.com</code> operators | <input type="checkbox"/> Microsoft |
| <input type="checkbox"/> <code>yahoo.com</code> operators | <input type="checkbox"/> Mozilla |
| <input type="checkbox"/> <code>google.com</code> operators | <input type="checkbox"/> Alice's ISP |
| <input type="checkbox"/> <code>image-host.com</code> operators | <input type="checkbox"/> UC Berkeley |

Question 3 *A Tour of Tor*

As a reminder, when connecting to a normal website through Tor, your computer first queries the Tor “consensus” to get a list of all Tor nodes, and using this information it connects to the first Tor node and, from there, creates a circuit through the Tor network, eventually ending at an exit node.

- (a) (4 min) Consider the scenario where you are in a censored country and the censor chooses not to block Tor, the censor is the adversary, and no Tor relays exist within this country. How many Tor relays must your traffic pass through, including the exit node, to prevent the censor from blocking your traffic.

- One Four
 Two Tor doesn't stop this adversary
 Three

- (b) (4 min) Consider the scenario where you are the only user of Tor on a network that keeps detailed logs of all IPs contacted. You use Tor to email a threat. The network operator is made aware of this threat and that it was sent through Tor and probably originated on the operator's network. How many Tor relays must your traffic pass through, including the exit node, to guarantee the network operator can't identify you as the one who sent the threat?

- One Four
 Two Tor doesn't stop this adversary
 Three

- (c) (4 min) Consider the scenario where there is a single hostile Tor node but you don't know that node's identity, and that node can be an exit node. You want to keep confidential from this node what HTTP sites you are visiting through Tor. How many Tor relays must your traffic pass through, including the exit node, to guarantee this adversary can't know what sites you visit?

- One Four
 Two Tor doesn't stop this adversary
 Three

- (d) (4 min) Consider the scenario where there are multiple independent hostile Tor nodes but you don't know their identities, and these nodes can be exit nodes. You want to keep confidential from all these nodes what HTTP sites you are visiting through Tor. How many Tor relays must your traffic pass through, including the

exit node, to guarantee that every independent hostile node can't know what sites you visit?

- One
- Two
- Three
- Four
- Tor doesn't stop this adversary

(e) (4 min) Consider the scenario where there are multiple colluding hostile Tor nodes but you don't know those nodes identities, and these nodes can be exit nodes. You want to keep confidential from all these nodes what HTTP sites you are visiting through Tor. How many Tor relays must your traffic pass through, including the exit node, to guarantee that the colluding system of hostile nodes can't know what sites you visit?

- One
- Two
- Three
- Four
- Tor doesn't stop this adversary

(f) (4 min) Consider the scenario where there is a single hostile Tor node but you don't know that node's identity, and that node can be an exit node. You want to have data integrity for the HTTP sites you are visiting through Tor. How many Tor relays must your traffic pass through, including the exit node, to guarantee this adversary can't manipulate the data you receive from the sites you visit?

- One
- Two
- Three
- Four
- Tor doesn't stop this adversary