### **Cookies That Give You Away** Evaluating the surveillance implications of web tracking

### Steven Englehardt

Dillon Reisman, Christian Eubank, Peter Zimmerman, Arvind Narayanan, and Ed Felten



Jonathan Mayer



# **Motivation and Attack**

### **Measurement Infrastructure**

## **Results and Lessons**

# Motivation





QUANTUMCOOKIEXKEYSCORE

- MUTANT BROTH
- QUANTUMINSERT
- BADASS

#### THE//INTERCEPT GREENWALD FEATURES BLOG DOCUMENTS STAFF CONTACT // @

### **OPERATION SOCIALIST**

#### THE INSIDE STORY OF HOW BRITISH SPIES HACKED BELGIUM'S LARGEST TELCO

BY RYAN GALLAGHER Serigallagher

12/13/2014

When the incoming emails stopped arriving, it seemed innocuous at first. But it would eventually become clear that this was no routine technical problem. Inside a row of gray office buildings in Brussels, a major hacking

POPULAR







#### Adversary

#### Page vists:

- nytimes.com

#### **Cookies:**

- e3d0836d-4a3...
- 6076ff86-f7b...
- e7eff9af-8e4...

Source: Mayer & Mitchell; Third-Party Web Tracking: Policy and Technology



#### Adversary

#### Page vists:

- nytimes.com

#### **Cookies:**

- e3d0836d-4a3...
- 6076ff86-f7b...
- e7eff9af-8e4...



#### Adversary

#### Page vists:

- nytimes.com
- imdb.com

#### **Cookies:**

- e3d0836d-4a3...
- 6076ff86-f7b...
- e7eff9af-8e4...

#### Identity:

- Enguerrand Quarton



#### Adversary

#### Page vists:

- nytimes.com
- imdb.com

#### **Cookies:**

- e3d0836d-4a3...
- 6076ff86-f7b...
- e7eff9af-8e4...

#### Identity:

- Enguerrand Quarton



#### Adversary

#### Page vists:

- nytimes.com
- imdb.com
- arstechnica.com

#### **Cookies:**

- e3d0836d-4a3...
- 6076ff86-f7b...
- e7eff9af-8e4...

#### Identity:

Enguerrand Quarton



#### Adversary

#### Page vists:

- nytimes.com
- imdb.com
- arstechnica.com

#### **Cookies:**

- e3d0836d-4a3...
- 6076ff86-f7b...
- e7eff9af-8e4...
- d5j20xnd-23d...

#### Identity:

Enguerrand Quarton

# Assumptions about the user

 A user's identifiers stay constant within a page visit

2. A user's cookies stay constant **across page visits** 

# **Adversary's Abilities**

1. Can disambiguate traffic between multiple users on the same IP within a page visit

2. Can utilize cookies to link requests across page visits

3. Adversary is passive

# **Motivation and Attack**

### **Measurement Infrastructure**

# **Results and Lessons**



# **OpenWPM**

- Web measurement framework
- Real browser automation
- Fully instrumented
- Support for user profiles





# We use two models of browsing

### • AOL Query Log

- $\circ~$  25 AOL users with 50 100 queries
- Issue queries on Google
- Visit top 5 results

### • Alexa Top 500

- Random sample of 200 urls for each user
- Localized to user's browsing location

# An adversary's restrictions

- Policy restrictions
  - NSA: One-End-Foreign
- Geographic restrictions
  - Collection outside US at undersea cables





# **Summary of Measurements**

- 1. Simulate users in 3 locations with 2 browsing models
  - a. United States
  - b. Ireland
  - c. Japan
- 2. Simulate a restricted adversary
- 3. Test effectiveness of blocking
- 4. Average the size of the largest connected component for 25 OpenWPM measurement instances in each case

### **Motivation and Attack**

### **Measurement Infrastructure**

# **Results and Lessons**

### Average percentage of first-party Average number of sites linked identity leakers



# **US Linking under "One-End-Foreign"**



# What can users do?

- Enable DNT
- Block 3rd party cookies for non-1st party sites
- Block all 3rd party cookies
- Ghostery
- HTTPS Everywhere

# Average of 25 simulated users from a US location browsing 200 sites under AOL model



# Takeaways

- 1. Cookies should be treated like identities
- 2. HTTPS by default
- 3. Policy debates happen in a vacuum, tools like ours can help fix that

Infrastructure: <u>https://github.com/citp/OpenWPM</u>

Code and Data: <a href="https://github.com/englehardt/cookies-that-give-you-away">https://github.com/englehardt/cookies-that-give-you-away</a>

Icon Credits:

User by Jose Campos, database by Stefan Parnarov, Hard Drive by Edward Boatman, programmer by Hadi Davodpour from the Noun Project