Web Authentication using Third-parties in Untrusted Environments

Anna Vapen
PhD Thesis Presentation 2016-09-30
Supervisors: Nahid Shahmehri, Niklas Carlsson
Agenda

1. Background
2. Research problems
3. Analysis
   - Web authentication and untrusted computers
   - The third-party authentication landscape
   - Third-parties and privacy risks
4. Contributions
Background
Web Authentication

- Method to prove that you are a specific person
- Personal web experience
  - User accounts require authentication

Example: Signing in to Google with username and password
Password Challenges

Most common web authentication method
Simple setup
Reused on several sites
Written down

Replay attacks
Forgotten by the user

Alternative methods
Time consuming
Additional equipment
Mobile Users and Untrusted Environments

- Mobile users
  - Different devices
  - Different places

- Untrusted environments
  - Infected computer
  - Untrusted WiFi network
Third-party Web Authentication

- Use an **IDP** (identity provider) account to access many **RPs** (relying parties)
- Fewer logins – simplify authentication
- Information sharing between websites
  - Privacy leaks!
Third-party Authentication Scenario

Relying party (RP)

Identity provider (IDP)

Redirect
Logged in

Background
Research Problems
Research Problems

1. Web authentication
   – For mobile users in untrusted environments?

2. Third-party authentication
   – Usage over time?
   – How to measure?

3. Privacy risks
   – Information flows between parties?
Web Authentication and Untrusted Computers
Mobile Phones as Authentication Devices

- Strong authentication
- Carried by the user
- Security problems
- Comparing solutions?
Design and Evaluation Method

• Design
  – Select requirements
  – Get design suggestions

• Evaluation
  – Start with an existing design
  – Get a security rating of the design

Web authentication

Requirements
• Security
• Availability
• ...

Login

Security rating

PrimeLife’11
Web authentication

Optical Authentication Proof-of-Concept

(1) Challenge barcode shown on screen
(2) Take a picture of the challenge
(3) Response generated
(4) Show response to webcam

Logged in!

IJMCMC’11
The Third-party Authentication Landscape
Data Collection

• Popularity-based logarithmic sampling
  – 80,000 points uniformly on a logarithmic range
  – Pareto-like distribution
  – Capturing data from different popularity segments
Large-scale Crawling

- Selenium-based crawling and relationship identification
- Able to process Web 2.0 sites with interactive elements
- Low number of false positives
- Validation with semi-manual classification and text-matching
Collected Data

1.6 terabyte analyzed data

25 million analyzed links

3,329 unique relationships

50 IDPs and 1,865 RPs

WHOIS, server location, and audience location

Total site size and number of links and objects
IDPs vs Content Sharing Services

Content sharing: Importing images, scripts etc. from other sites (third-party content providers)

IDPs are selected locally, in contrast to content services.
Service-based Analysis of RPs

Early adopters, using several IDPs

Likely to be IDPs

Social/portal

Video

Tech

File sharing

Info

News

Using social/portal IDPs

Ad services, CDNs

Manual analysis: Top 200 websites in April 2012

3rd-party authentication
Third-parties and Privacy Risks
Privacy risks

App Rights and Information Flows

App rights example

Actions:
- Read
- Write
- Update/remove

SEC’15, UEOP’16
Our Studies on Privacy Risks

- Categorization app-rights data
  - Manual study on the top 200 most popular websites
  - Longitudinal approach: three years
- Targeted login tests
- Privacy risk categorization
  - Data types in app rights
  - Combinations of types
Protocol Selection

- **OpenID**
  - Authentication protocol
  - Decreasing in popularity

- **OAuth**
  - RP may use actions on IDP
  - Rich user data is shared
  - Increasingly popular

Privacy risks

SEC’15, UEOP’16
IDP Selection

- Top 200 April 2012: 69 RPs and 180 relationships
- Same sites, April 2015: +15 RPs and +33 relationships
- 75% of these RPs are selecting all their IDPs from the top 5 most popular IDPs

Top IDPs: Facebook, Twitter, Sina, Yahoo, Google

Privacy risks

SEC’15, UEOP’16
Risk Types

Facebook, Twitter and Google:

• Only a few relationships in the most privacy preserving category

• 2+ IDPs: More than half are using actions
  – Dangerous when having several IDPs
  – Potential multi-hop leakage
Multi-account Information Risks

- Cross account leakage
- Unwanted combinations of conflicting information
- RPs handle multi-IDP usage badly

Connecting several IDPs to an RP
Structures in the RP-IDP Landscape

High-degree IDP case
- IDP having many RPs
- Top IDPs

High-degree RP case
- RP having many IDPs
- Specialized IDPs

Hybrid case
- Hybrids are both RP and IDP
RP-to-RP Leakage Example

- Potential RP-to-RP leaks
  - Data posted to IDP from RP1
  - Data read from IDP to RP2

<table>
<thead>
<tr>
<th>IDP</th>
<th>February 2014</th>
<th>April 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Severe</td>
</tr>
<tr>
<td>Facebook</td>
<td>645</td>
<td>150</td>
</tr>
<tr>
<td>Twitter</td>
<td>110</td>
<td>110</td>
</tr>
<tr>
<td>Google</td>
<td>91</td>
<td>0</td>
</tr>
</tbody>
</table>

Dataset with 44 RPs using Facebook, 14 using Twitter and 12 using Google
Contributions
Contributions

- Design and evaluation method
- Large-scale RP-IDP measurements
  - Novel measurement method
  - Categorization of RP-IDP relationships
- Privacy risks and information sharing
  - Protocol analysis
  - Structural properties
Web Authentication using Third-parties in Untrusted Environments

Anna Vapen

Papers included in this thesis:

- Security Levels for Web Authentication using Mobile Phones, *PrimeLife*’11
- 2-clickAuth - Optical Challenge-Response Authentication using Mobile Handsets, *IJMCMC*’11
- Third-party Identity Management Usage on the Web, *PAM*’14
- A Look at the Third-Party Identity Management Landscape, *IC*’16
- Information Sharing and User Privacy in the Third-party Identity Management Landscape, *SEC*’15
- Longitudinal Analysis of the Third-party Authentication Landscape, *UEOP*’16