I Know What You Did Last Login: Inconsistent Messages Tell Existence of a Target’s Account to Insiders

Ayako Hasegawa, Takuya Watanabe, Eitaro Shioji, Mitsuaki Akiyama
NTT Secure Platform Laboratories

December 13, 2019
The Annual Computer Security Applications Conference (ACSAC)
Background: Account-Existence Attack

Attacker

Online Service

Login

User ID

Password

Service Response

Login

That user ID doesn’t exist!

The target does not have an account

Login

Incorrect password!

The target has an account

target.user@example.com

Threat Model: Mass account enumeration by strangers
Account-Existence Privacy Attack by Insiders

- Threat model: **Targeted privacy attack by insiders** (e.g., partners, family, friends, co-workers)

- Account existence on certain online services may imply one’s personal preferences or situations
  - Career change services → Looking for a new job
  - Payday loan services → Falling into financial troubles

- Practical Implications
  - Attackers know a target’s user ID (email address or phone number)
  - The attack flow does not need highly-technical skills (manual attack)
  - Rate-limiting defenses (e.g., CAPTCHA) are not effective
Key Contributions

• Proposal of a new privacy threat, “Account-Existence Privacy Attack by Insiders”, confirmed through user studies

• Comprehensive measurement of the threat on actual online services with login-related messages
Research Questions

- RQ1: What services do users consider sensitive?
  ⇒ Exploratory user study

- RQ2: Are such sensitive services secure against our attack to identify the existence of a target’s account?
  ⇒ Measurement study

- RQ3: How much does our account-existence attack actually impact user privacy?
  ⇒ Main user study
Exploratory User Study (Methodology)

RQ1: What services do users consider sensitive?

• Online survey
  – Amazon Mechanical Turk
  – N=614 (U.S. residents)
  – Reward: $1

• Questionnaire
  – Sensitive service categories (multiple choices allowed)
  – Sensitive service names (optional)

Question:
Among online services, are there any that you would feel uncomfortable if other people find out that you have an account on?
**Exploratory User Study (Results)**

### Sensitive service categories (multiple choices allowed)

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dating</td>
<td>54.4%</td>
</tr>
<tr>
<td>Porn</td>
<td>50.5%</td>
</tr>
<tr>
<td>Social Networking</td>
<td>19.9%</td>
</tr>
<tr>
<td>Career change</td>
<td>17.6%</td>
</tr>
<tr>
<td>Forum</td>
<td>14.3%</td>
</tr>
<tr>
<td>Financial</td>
<td>12.5%</td>
</tr>
<tr>
<td>Shopping</td>
<td>8.8%</td>
</tr>
<tr>
<td>Healthcare</td>
<td>4.6%</td>
</tr>
<tr>
<td>Cloud storage</td>
<td>4.4%</td>
</tr>
<tr>
<td>Other</td>
<td>3.9%</td>
</tr>
<tr>
<td>Never</td>
<td>18.4%</td>
</tr>
</tbody>
</table>

81.6% of participants selected one or more sensitive service categories
Exploratory User Study (Results)

- Highly-sensitive services given by participants
  - Dating services provided for sexual minorities
  - Porn services provided for particular sexual propensities
  - Forum services regarding sex life, sperm banks, hate speech
  - Payday loan services
  - Sexually transmitted diseases (STDs) testing services
  - Supplemental nutrition assistance program (SNAP) benefits services
Measurement Study (Methodology)

RQ2: Are such sensitive services secure against our attack to identify the existence of a target’s account?

- Evaluated login-related messages on actual services
  - **69 sensitive services** provided in our exploratory user study
  - **35 popular services** from the top of Alexa Top Global Sites
  - Excluded services
    - Services in non-English languages
    - Services sharing the same auth platform (1 chosen)
    - Services not using email addresses as user IDs (i.e., usernames, phone numbers)
Attack Procedure Overview (e.g., in Login Function)

- **Attacker’s email address** (unregistered)
  - Step 1: Service Vulnerability Test (Vulnerable if inconsistent)
  - Login
    - That user ID doesn’t exist!

- **Attacker’s email address** (registered)
  - Step 2: Account Existence Test (Account exists if matches)
  - Login
    - Incorrect password!

- **Target’s email address** (unknown)
  - Login
    - Incorrect password!

Measured in the experiment
## Login-related Functions / (In)consistent Messages

<table>
<thead>
<tr>
<th>Function</th>
<th>Input</th>
<th>Inconsistent (Insecure) Output</th>
<th>Consistent (Secure) Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Login</strong></td>
<td>Registered user ID with incorrect password</td>
<td>“Incorrect password”</td>
<td>“Incorrect user ID or password”</td>
</tr>
<tr>
<td></td>
<td>Unregistered user ID with arbitrary password</td>
<td>“That user ID doesn’t exist”</td>
<td></td>
</tr>
<tr>
<td><strong>Password</strong></td>
<td>Registered email address</td>
<td>“We just sent you a password-reset link”</td>
<td>“If that email address is in our database, we’ll send you an email to reset your password”</td>
</tr>
<tr>
<td>recovery</td>
<td>Unregistered email address</td>
<td>“This email address doesn’t exist in our database”</td>
<td></td>
</tr>
<tr>
<td><strong>Account</strong></td>
<td>Registered user ID</td>
<td>“This user ID is already in use”</td>
<td>“A link to activate your account has been emailed to ⟨input email address⟩”</td>
</tr>
<tr>
<td>creation</td>
<td>Unregistered user ID</td>
<td>“Welcome! You have signed up successfully”</td>
<td></td>
</tr>
</tbody>
</table>
Measurement Study (Methodology)

- Collected messages on each service
  - 4 stages of account life cycle
    - Before registration
    - After registration
    - After email-address update
    - After account closure
  - 3 login-related functions
    - Login
    - Password recovery
    - Account creation
  - 2 platforms
    - Websites
    - Mobile apps

4*3*2 = 24 messages/service (maximum)

Collected over 1.1k messages in total
## Measurement Study (Methodology)

- Conducted 4 types of analysis

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Purpose</th>
<th>Target</th>
<th>Compared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis-1</td>
<td>Basic security check</td>
<td>Websites</td>
<td>Before registration</td>
</tr>
<tr>
<td>Analysis-2</td>
<td>Difference of security level on websites and apps</td>
<td>Websites and apps</td>
<td>Before registration</td>
</tr>
<tr>
<td>Analysis-3</td>
<td>Effectiveness of email-address update</td>
<td>Websites</td>
<td>Before registration</td>
</tr>
<tr>
<td>Analysis-4</td>
<td>Effectiveness of account closure</td>
<td>Websites</td>
<td>Before registration</td>
</tr>
</tbody>
</table>
Measurement Study (Methodology)

• Evaluation of each function

Before registration

Login
That user ID doesn’t exist!

Login
Incorrect password!

After registration
After email-address update
After account closure

Messages are inconsistent ⇒ the function is insecure

• Evaluation of each service

<table>
<thead>
<tr>
<th>Service</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Login</td>
<td>Secure</td>
</tr>
<tr>
<td>Password recovery</td>
<td>Secure</td>
</tr>
<tr>
<td>Account creation</td>
<td>Insecure</td>
</tr>
</tbody>
</table>

Any one of three functions is insecure ⇒ the service is vulnerable
# Measurement Study (Results of Analysis-1)

## Basic security check

<table>
<thead>
<tr>
<th></th>
<th># total</th>
<th>% secure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Login</td>
<td>Password Recovery</td>
</tr>
<tr>
<td>Sensitive</td>
<td>69</td>
<td>72.5%</td>
</tr>
<tr>
<td>Popular</td>
<td>35</td>
<td>54.3%</td>
</tr>
<tr>
<td>Total</td>
<td>87</td>
<td>71.3%</td>
</tr>
</tbody>
</table>

- 98.9% of services we investigated were vulnerable
- Almost all services had an insecure account-creation function
  - Our attack works most efficiently in account-creation function
Measurement Study (Results of Analysis-2)

- Almost all services were vulnerable regardless of platforms (web or app)
- \(...\) are there differences at function-level granularity?
  - compared in terms of the number of secure functions on each platform

<table>
<thead>
<tr>
<th></th>
<th>#</th>
<th>Web &gt; App</th>
<th>Web == App</th>
<th>Web &lt; App</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitive</td>
<td>46</td>
<td>19.6%</td>
<td>71.7%</td>
<td>8.7%</td>
</tr>
<tr>
<td>Popular</td>
<td>31</td>
<td>22.6%</td>
<td>71.0%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Total</td>
<td>63</td>
<td>19.0%</td>
<td>71.4%</td>
<td>9.5%</td>
</tr>
</tbody>
</table>

- Most services had the same security level on their web and app platforms
- Apps tended to be less secure for services with a difference
Effectiveness of email-address update and account closure

<table>
<thead>
<tr>
<th></th>
<th>Email-address update</th>
<th>Account closure</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Secure services</td>
<td>92.5%</td>
<td>47.6%</td>
</tr>
<tr>
<td>(sensitive and popular)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Email-address update was effective against our attack on most services
  - Services behave as if the previous email address is unregistered

- Account closure was much less effective than email-address update
  - Account closure was not properly reflected to the messages
Main User Study (Methodology)

RQ3: How much does our account-existence attack actually impact user privacy?

- **Online survey**
  - Amazon Mechanical Turk
  - N=447 (U.S. residents)
  - Reward: $3

- **Questionnaire**
  - Demographics
  - Reasons for not wanting others to know their use of sensitive services (optional)
  - Motivation for violating intimates’ or acquaintances’ privacy (optional)
  - Characteristics of email addresses registered on sensitive services
### Main User Study (Results)

#### Reasons for not wanting others to know their use of sensitive services

(Results of thematic analysis by two coders)

<table>
<thead>
<tr>
<th>Theme</th>
<th>%</th>
<th>Examples of participants’ comments</th>
</tr>
</thead>
</table>
| Embarrassment          | 85.4% | • “Dating sites mean that I can’t find a date normally […]”  
|                         |     | • “Because porn is still taboo and especially because I’m a woman”  
|                         |     | • “They might laugh at my interests” |
| Effect on work         | 11.6% | • “It might damage my reputation and how they view me professionally, weakening my potential promotions and contacts in the future” |
| Immoral behavior       | 3.0%  | • “I’m married so I shouldn’t be on dating sites […]” |
Main User Study (Results)

- As potential perpetrators, 25% of participants represented such a desire

Motivation for violating intimates’ or acquaintances’ privacy
(Results of thematic analysis by two coders)

<table>
<thead>
<tr>
<th>Theme</th>
<th>%</th>
<th>Examples of participants’ comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jealousy</td>
<td>63.2%</td>
<td>• “[...] I wanted to know if a current boyfriend had a profile on an online dating site”</td>
</tr>
<tr>
<td>Curiosity</td>
<td>34.2%</td>
<td>• “I would just want to see what friends I have use what sites so I can talk to them more about it”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “I have wanted to know whether coworkers were considering career changes/looking at job-search sites”</td>
</tr>
<tr>
<td>Worry</td>
<td>2.6%</td>
<td>• “I have wanted to know my son’s information [...] as to make sure he isn’t into things that might cause him harm.”</td>
</tr>
</tbody>
</table>
Main User Study (Results)

Characteristics of email addresses participants register on sensitive services

<table>
<thead>
<tr>
<th>Email address is..</th>
<th>%</th>
<th>Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>- the same as the one used for a non-sensitive service, and</td>
<td>30.5%</td>
<td>Potential victims (45.2%)</td>
</tr>
<tr>
<td>- anyone who knows me may know it</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- different from any used for non-sensitive services, and</td>
<td>14.7%</td>
<td>Appropriate behavior (54.5%)</td>
</tr>
<tr>
<td>- anyone who knows me may know it</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- the same as the one used for a non-sensitive service, and</td>
<td>15.0%</td>
<td></td>
</tr>
<tr>
<td>- no one who knows me knows it</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- different from any used for non-sensitive services, and</td>
<td>39.5%</td>
<td></td>
</tr>
<tr>
<td>- no one who knows me knows it</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Privacy and Usability Tradeoff

• Adopting consistent (secure) messages may be less user-friendly
  – e.g., “Incorrect email address or password” interferes with the user’s understanding of the reason for login failure

• At least the services considered sensitive by users should adopt consistent messages

• Open question: how to design a system with consistent messages while not sacrificing usability?
Responsible Disclosure

• These defects are not due to software or service specific vulnerabilities but rather a design issue across existing services

• We reported this defect to all sensitive services which we found vulnerable through our measurement study

• We improved OWASP’s Authentication Cheat Sheet and Application Security Verification Standard (ASVS)
Summary

• The first comprehensive study of a privacy attack by insiders that identifies the target’s account existence by abusing inconsistent login-related messages

• RQ1: What services do users consider sensitive?
  ⇒ Dating, porn, social networking, career change, forum, financial ...
  82% feel account existence disclosure is a privacy issue in certain services

• RQ2: Are such services secure against our attack to identify the existence of a target’s account?
  ⇒ Almost all services we investigated were vulnerable
    Possibly because they assume account enumeration (rate-limiting defenses)?

• RQ3: How much does our account-existence attack actually impact user privacy?
  ⇒ Potential perpetrators (25%) and potential victims (45%)