Personal Choice and Challenge Questions: A Security and Usability Assessment

13 July 2009
Mike Just
University of Edinburgh
(joint work with David Aspinall)
What are Challenge Questions? (1 of 3)

- What are 'Challenge Questions?'
  - Type of 'authentication credential'
  - Users register Question & Answer
  - To authenticate later, user is posed Question and asked to provide Answer

Authentication Credentials

'Something You Have'
- Access card
- Smartcard
- Mobile

'Something You Are'
- Fingerprints
- Iris/retinal scan
- Facial scan

'Something You Know'
- Passwords
- PINs
- Images

'Something You Memorize'
- Challenge questions
- Images

'Something You Already Know'

What are Challenge Questions? (2 of 3)

- **Common Examples**
  - 'What is my Mother's Maiden Name?'
  - 'What was the name of my first pet?'
  - 'What was the name of my primary school?'

- **How do Challenge Questions support authentication?**
  - The answers to the questions should be known only to the users that registered the questions, similar to how passwords should be uniquely known.
How and why do we use Challenge Questions?

- Almost exclusively as secondary/fallback authentication in case of lost primary credential
- Sometimes used to complement primary credential
- Often driven by desire to avoid costly help-desk calls
- In some cases, 're-registration' is possible, but not always
  - Too expensive or takes too much time
  - Not all sites have a registration phase (that includes user identification with shared secrets)
- So, some form of secondary authentication is desireable
  - Challenge Questions are today's ubiquitous choice
What is studied w.r.t. Challenge Questions?

1. Security (Attacker's Point-of-View)
   - How difficult is it to determine the answers to the questions?
   - Demonstration of security often involves *quantitative analysis*

2. Usability (User's Point-of-View)
   - How easy is it to choose questions?
   - How easy is it to remember the answers?
   - Demonstration of usability often involves *qualitative research*
Related Work (1 of 5)

- **Applications**: Applications of challenge question authentication
- **Alternatives**: Alternatives to traditional question-answer model
- **Assessments**: Assessments of security and usability
Related Work (2 of 5)

- Introduced as means of authentication of client to server (i.e., password replacement)
- Challenge questions to protect secret keys
  - Secret sharing to tolerate forgetfulness
    - Ellison *et al.*, *JFGCS* 2000
    - Frykholm and Juels, *ACM CCS* 2001
- Group authentication
  - Shared knowledge between two or more users
    - Toomim *et al.*, *CHI* 2008
Related Work (3 of 5)

- User preferences
  - O'Gormann et al., *Financial Crypto*. 2004

- Browsing history
  - Asgharpour and Jakobsson, *IWSSI* 2007

- Digital objects as passwords
  - Mannan and van Oorschot, *HotSec* 2008

- First two: *Something you (already) know*
- Last two: *Something you have (access to)*
Usability

- Several studies of the applicability, memorability and repeatability of both system- and user-chosen questions
  - Pond et al., *Comp. & Sec.* 2000
  - Rabkin, SOUPS 2008 (Subjective assessment)
  - Just and Aspinall, *Trust* 2009
  - Schechter et al., *IEEE S&P* 2009
- Results indicate that users have difficulty remembering or repeating their answers
Related Work (5 of 5)

**Security**
- Assessment using 'live' attacks by friend & family, acquaintances and strangers
  - Pond *et al.*, *Comp. & Sec.* 2000
  - Toomim *et al.*, *CHI* 2008
  - Schechter *et al.*, *IEEE S&P* 2009
- Assessment using 'likelihood' measures
  - Griffith and Jakobsson, *ACNS* 2005
  - Rabkin, *SOUPS* 2008
  - Just and Aspinall, *Trust* 2009
- Results indicate that many questions are at risk
Recent research suggests significant problems with both the security and usability of challenge question authentication systems.

- How can we begin to improve?

A systematic and repeatable way to analyze the security and usability of challenge questions.

- To continue to assess current systems
- To allow assessment of future systems

Our focus was on user-chosen questions ('personal choice')

Along the way, we discovered an interesting experimental method.
1. Devised novel experiment for collecting authentication information
2. Created a security model for question assessment
3. Assessed the security and usability of 180 user-chosen challenge questions
   - Experiment with 60 first-year Biology students at the University of Edinburgh
Collecting Data (1 of 3)

- Ethically challenging, but users readily submit
- Issues regarding participant behaviour
  - Equate credentials with other private information?
  - Contribute *real* information?
  - Degree of freedom with user-chosen questions
- Opportunities for improved Collector behaviour
  - Challenge to ourselves: Don't collect!
  - Avoid having to maintain information
  - Consistent message: Keep credentials to yourself!
Collecting Data (2 of 3)

Stage 1
- Participant
- Questions
- Answers
- MATCH?
- Usability Analysis
- Security Analysis

Stage 2
- Experiment
- Questions
- Answers
- Answers

Questions
Answers
MATCH?
Collecting Data (3 of 3)

- Participants use of 'real' Questions and Answers
  - We asked if participants would use same Questions and Answers in real applications (e.g. Banking)
  - Of the respondents (94%) indicating that they would likely re-use their questions, 45% indicated some influence from not submitting their answers

- Participants and personal privacy
  - We asked participants if they would be concerned if their friends or family members knew their Questions and Answers
  - More than two-thirds of the questions raised 'no concern' at all for participants with < 10% meriting strong concern

- Results are similar to our earlier trials (Trust 2009)
Security Model (1 of 2)

- Existing security analysis of Challenge Questions is ad hoc
- There are no clear guidelines for choosing 'good' questions and answers
- We wanted a more systematic and repeatable approach that would
  - Provide some guidance for secure design
  - Allow continued assessment of new solutions
- We encourage further refinement of our model
- Assessment results depend upon context
Security Model (2 of 2)

Increasing Information for Attacker

Answer alphabet and distribution, common answer sets

Questions, distributions of likely answers

User account, published data, social networks, friends, family, ...

Attack Methods

Blind Guess

Focused Guess

Observation

Answer Guess

13 July 2009 Challenge Question Security and Usability
Security Analysis – Blind Guess (1 of 6)

- Brute force attack
- Security Levels based on equivalence to passwords
  - 6-char alphabetic password ($2^{34}$)
  - 8-char alphanumeric password ($2^{48}$)
- Answer entropy: 2.3 bits ($1^{st}$ 8 chars), then 1.5 bits
- Results (by question)
  - Average answer length: 7.5 characters
  - 174 Low, 4 Medium, 2 High
- Results (by user)
  - Q1 – 59 Low, 1 Medium, 0 High
  - Q1, Q2 – 38 Low, 13 Medium, 9 High
  - Q1, Q2, Q3 – 5 Low, 19 Medium, 36 High
Blind Guess (cont'd)

- Unlike passwords, the alphabet for answers is just 26 lowercase letters (plus 10 digits in some cases)
- Use of a single question seems to provide insufficient protection against the simplest attack
- But, multiple questions seem to help (only considering Blind Guess Attack)
- Offline attacks would require more security \(2^{80}\)
- Might consider VeryLow and VeryHigh categories as well
Attacker knows the Challenge Questions

Security Levels same as for Blind Guess

Answer types and space

Results (by question)
- 167 Low, 0 Medium, 13 High

Results (by user)
- Q1 – 58 Low, 0 Medium, 2 High
- Q1, Q2 – 46 Low, 11 Medium, 3 High
- Q1, Q2, Q3 – 5 Low, 28 Medium, 27 High

Much room for refinement of 'Space'

<table>
<thead>
<tr>
<th>Q Type</th>
<th>%</th>
<th>$\log_{10}$ Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proper Name</td>
<td>50%</td>
<td>4 – 5</td>
</tr>
<tr>
<td>Place</td>
<td>20%</td>
<td>2 – 5</td>
</tr>
<tr>
<td>Name</td>
<td>18%</td>
<td>3 – 7</td>
</tr>
<tr>
<td>Number</td>
<td>3%</td>
<td>1 – 4</td>
</tr>
<tr>
<td>Time/Date</td>
<td>3%</td>
<td>2 – 5</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>6%</td>
<td>8 – 15</td>
</tr>
</tbody>
</table>
Attacker tries to obtain or observe the answer

Security Levels defined qualitatively
- Low – Answer publicly available
- Medium – Answer not public, but known to F&F
- High – Neither

Levels assigned to questions by
- Subjective analysis, and
- Participant input (provided upper bound only)

Results (by question)
- 124 Low, 54 Medium, 2 High

Results (by user)
- 24 Low, 34 Medium, 2 High
- Did not "sum" levels (used max)

Much room for refinement of levels and analysis
Security Analysis – Overall (5 of 6)

- Overall rating is a 3-tuple (Blind, Focused, Observation)
- Results
  - All Low – 1 participant
  - All High – 0 participants
  - No Lows – 31 participants (50%)
  - (H,M,M) or (M,H,M) – 15 participants (25%)
  - (H,H,M) – 11 participants (20%)
- Dependencies not (yet) considered
- Ability to perform observation attacks in parallel, and offline, is a significant advantage for attackers
Security Analysis – Overall (6 of 6)

- Perceived effort of Stranger to Discover Answers
  - Very difficult (47%)
  - Somewhat difficult (42%)
  - Not difficult at all (11%)
  - Users overestimate the difficulty of attack

- Perceived effort of Friend/Family to Discover Answers
  - Very difficult (11%)
  - Somewhat difficult (36%)
  - Not difficult at all (53%)
  - Users surprisingly aware of this risk
Usability Analysis (1 of 3)

- Usability often refers to 'usable interface design'
- For usable authentication, similar principles apply
  - The user should be able to understand and execute their task
  - We're dealing specifically with information
  - We're more concerned with mental capabilities, e.g., processing, memory
Usability Analysis (2 of 3)

- Applicability
  - Users have sufficient information to provide an answer to a question
  - E.g., 'What was my first pet's name?'
  - Relevant to administratively-chosen questions (not user-chosen)

- Memorability
  - Users can consistently recall the original answer to a question over time
  - Precise recall, 'blank'

- Repeatability
  - Users can consistently and accurately repeat the original answer to a question over time
  - E.g., 'Favourites' change over time, 'Street' versus 'Avenue'
Usability Analysis (3 of 3)

- Answer recall (180 questions)
  - 15 errors (8%)
  - Reduces to 7 errors (4%) if we exclude 'capitalization' errors
- Answer recall (60 users)
  - 11 users (18%) made at least one error
  - Reduces to 7 users (12%) if we exclude 'capitalization' errors
- Comments suggest that 'complicated answers' and allowance of free-form answers may be culprit
- Florêncio & Herley (2007) found that 4.28% of Yahoo! users forget their passwords
- Our results were after 23 days, with young students
What Does it All Mean? (1 of 2)

- Our results corroborate recent results regarding the security and usability of challenge questions.
- But, before we write-off challenge questions ...
  - Multiple questions seem to help (security at least)
  - Current implementations are terribly boring
    - Little research of challenge question auth
    - Most has been to assess security and usability
    - Less research into new designs
Potential paths forward

- Dynamic assessments of security and usability
- New types of information for authentication (new questions, 5 W's)
- Options of other methods: who you know, what you have access to, …
- Users are different – customize to meet their strengths (no 'one-size-fits-all')

But, how to improve usability …

- Fixed-form answers
- Tolerance for < 100% accuracy
Further Information

- Project web site
- Email
  - mike.just@ed.ac.uk