Towards Reliable Storage of 56-bit Secrets in Human Memory

Joseph Bonneau
Princeton University

Stuart Schechter
Microsoft Research

Presented By:
William Hollingsworth
Problem Definition

• Humans do not pick good passwords

“Humans are incapable of securely storing high-quality cryptographic keys, and they have unacceptable speed and accuracy when performing cryptographic operations. They are also large, expensive to maintain, difficulty to manage, and they pollute the environment. It is astonishing that these devices continue to be manufactured and deployed. But they are sufficiently pervasive that we must design our protocols around their limitations.”

--Kaufman, Perlman, and Speciner (2002)
History of the Password

- User-chosen passwords relatively new
  - Spoken passwords
- A way to authenticate ourselves to a computer
- Weaknesses quickly discovered
History of the Password

• Multics
  – qua-vu ri-ja-cas te-nort oi-boay fleck-y

• Air Force Data Services
  – Gave user option of random or user-chosen
  – 50/50 split
History of the Password

- Multics
  - qua-vu ri-ja-cas te-nort oi-boay fleck-y
- Air Force Data Services
  - Gave user option of random or user-chosen
  - 50/50 split
- More likely to write down random passwords
  - Less secure
An Optimistic View

- **Observation**: Humans tend to *forget* random information they only see once
- What if we try to teach the user their secret over time?
An Optimistic View

• **Observation**: Humans tend to *forget* random information they only see once

• What if we try to teach the user their secret over time?

  *Spaced Repetition*

• **Examples**: Vocabulary, Anatomy
Outline

- Background
- Proposed Design
- Methodology
- Results
- Conclusion
Proposed Design

- Choose password on account creation
- Teach random password at login

- Two groups
  - 12 Letters: $\log_2 26^{12} \approx 56.4$ bits
  - 6 Words: $\log_2 676^6 = \log_2 26^{12} \approx 56.4$ bits
Proposed Design

- At first login, display the first chunk above text field
- Wait a little longer each time to display the “hint”
- After three consecutive logins without the “hint”, give a new chunk
Proposed Design

- At first login, display the first chunk above text field
- Wait a little longer each time to display the “hint”
- After three consecutive logins without the “hint”, give a new chunk
- Repeat
Outline

- Background
- Proposed Design
- Methodology
- Results
- Conclusion
But we just said humans tend to write such passwords on paper.
Methodology

• But we just said humans tend to write such passwords on paper

• Withhold the purpose of the study from test participants
  - 223 participants recruited through the Amazon Mechanical Turk
Methodology: Distractor Task

- The Stroop Effect
Methodology: Distractor Task

Instructions
Watch for a word to appear in one of the two boxes below.
- If the word "left" appears in either box, type 'f'.
- If the word "right" appears in either box, type 'j'.
Lower scores are better. Keep your score low by responding as quickly and as accurately as possible.

Figure 3: The Attention Game, our distractor task
Methodology: Limitations

- Possible to receive hints after learning code
- Only English words
  - Under-studied aspect of passwords
- Validity concerns
  - 70 “suspected” true purpose
  - 28 “certain”
Methodology: Treatments

- Three groups
  - *Words* (~40%)
  - *Letters* (~40%)
  - *Control* (~20%)

- Do the random passwords cause participants to drop out?
Methodology: Sessions

- 90 Sessions
  - ~5 minutes ea.
  - >= 30 minutes between sessions

Figure 4: Participants were asked to fill out this two-question survey before every attention game.
Methodology: Completion

- Participants debriefed during 90th session
- Not informed of follow-up
- Those who finished early did not appear to “spoil” this for others
Methodology: Completion

- Participants debriefed during 90th session
- Not informed of follow-up

- Most participants truly learned the password
  - 135 *words* and *letters* participants completed
  - 24 admitted to writing down their password
  - Asked just *before* debriefing
Methodology: Completion

- Participants debriefed during 90th session
- Not informed of follow-up
- Most participants truly learned the password
  - “the words are branded into my brain”
Methodology: Completion

- Participants debriefed during 90th session
- Not informed of follow-up

- Asked authentication preferences for various scenarios
  - Most said learned random password
  - After debriefing (possible they lied)
Outline

- Background
- Proposed Design
- Methodology
- Results
- Conclusion
### Results: Summary

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Letters</th>
<th>Words</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signed up for the ‘attention’ study</td>
<td>41</td>
<td>92</td>
<td>90</td>
<td>223</td>
</tr>
<tr>
<td>Quit after 2 or 3 games</td>
<td>0/41 0%</td>
<td>9/92 10%</td>
<td>12/90 13%</td>
<td>21/223 9%</td>
</tr>
<tr>
<td>Otherwise failed to finish</td>
<td>6/41 15%</td>
<td>14/92 15%</td>
<td>12/90 13%</td>
<td>32/223 14%</td>
</tr>
<tr>
<td>Completed the ‘attention’ study</td>
<td>35/41 85%</td>
<td>69/92 75%</td>
<td>66/90 73%</td>
<td>170/223 76%</td>
</tr>
<tr>
<td>Received full security code</td>
<td>—</td>
<td>63/68 93%</td>
<td>64/65 98%</td>
<td>127/133 95%</td>
</tr>
<tr>
<td>Typed entire code from memory</td>
<td>—</td>
<td>62/63 99%</td>
<td>64/64 100%</td>
<td>126/127 99%</td>
</tr>
<tr>
<td>Participated in first follow-up</td>
<td>—</td>
<td>56/63 89%</td>
<td>56/64 88%</td>
<td>112/127 88%</td>
</tr>
<tr>
<td>Recalled code correctly</td>
<td>—</td>
<td>46/56 82%</td>
<td>52/56 93%</td>
<td>98/112 88%</td>
</tr>
<tr>
<td>Participated in second follow-up</td>
<td>—</td>
<td>52/56 93%</td>
<td>52/56 93%</td>
<td>104/112 93%</td>
</tr>
<tr>
<td>Recalled code correctly</td>
<td>—</td>
<td>29/52 56%</td>
<td>32/52 62%</td>
<td>61/104 59%</td>
</tr>
</tbody>
</table>

Table 1: Results summary: participants who signed up for the attention study, the fraction of those participants who completed the study, the fraction of the remaining participants who entered the first two chunks of their security code reliably enough to be shown the full security code (all three chunks), the fraction of those remaining who participated in the follow-up studies (after 3 and 17 days, respectively), and the fraction of those who recalled their security code correctly. The control group did not receive security codes and hence are excluded from the latter rows of the table.

$H_0$: The treatment had no effect on the drop out rate

$H_A$: The treatment does have an effect

$p = 0.2166$
## Results: Summary

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Letters</th>
<th>Words</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signed up for the ‘attention’ study</td>
<td>41</td>
<td>92</td>
<td>90</td>
<td>223</td>
</tr>
<tr>
<td>Quit after 2 or 3 games</td>
<td>0/41 0%</td>
<td>9/92 10%</td>
<td>12/90 13%</td>
<td>21/223 9%</td>
</tr>
<tr>
<td>Otherwise failed to finish</td>
<td>6/41 15%</td>
<td>14/92 15%</td>
<td>12/90 13%</td>
<td>32/223 14%</td>
</tr>
<tr>
<td>Completed the ‘attention’ study</td>
<td>35/41 85%</td>
<td>69/92 73%</td>
<td>66/90 73%</td>
<td>170/223 76%</td>
</tr>
<tr>
<td>Received full security code</td>
<td>—</td>
<td>63/68 93%</td>
<td>64/65 98%</td>
<td>127/133 95%</td>
</tr>
<tr>
<td>Typed entire code from memory</td>
<td>—</td>
<td>62/63 99%</td>
<td>64/64 100%</td>
<td>126/127 99%</td>
</tr>
<tr>
<td>Participated in first follow-up</td>
<td>—</td>
<td>56/63 89%</td>
<td>56/64 88%</td>
<td>112/127 88%</td>
</tr>
<tr>
<td>Recalled code correctly</td>
<td>—</td>
<td>46/56 82%</td>
<td>52/56 93%</td>
<td>98/112 88%</td>
</tr>
<tr>
<td>Participated in second follow-up</td>
<td>—</td>
<td>52/56 93%</td>
<td>52/56 93%</td>
<td>104/112 93%</td>
</tr>
<tr>
<td>Recalled code correctly</td>
<td>—</td>
<td>29/52 56%</td>
<td>32/52 62%</td>
<td>61/104 59%</td>
</tr>
</tbody>
</table>

Table 1: Results summary: participants who signed up for the attention study, the fraction of those participants who completed the study, the fraction of the remaining participants who entered the first two chunks of their security code reliably enough to be shown the full security code (all three chunks), the fraction of those remaining who participated in the follow-up studies (after 3 and 17 days, respectively), and the fraction of those who recalled their security code correctly. The control group did not receive security codes and hence are excluded from the latter rows of the table.

Most participants appeared to learn the full security code. Many were able to recall it during the follow up.
Results: Learning Rates

- Effect of treatment?
Results: Learning Rates

- Effect of treatment?

$H_0$: Encoding had no effect

$H_A$: Encoding had an effect

$p = 0.07$
Results: Learning Rates

- Effect of chunk no.?
Results: Learning Rates

- Effect of chunk no.?
  - b/w chunks 2, 3

H₀: Chunk # had no effect
Hₐ: Chunk # had an effect

p = 0.39
**Results: Login Speed**

- **Entry time**
  - *Words*: median 7.7s
  - *Letters*: median 6.0s

H₀: Treatment had no effect
Hₐ: Letters group faster
p < 0.01
Results: Errors

- Number of errors
  - *Words*: median 5
  - *Letters*: median 7

$H_0$: Treatment has no effect
$H_A$: More errors in letters
$p = 0.08$
Results: Follow-up Recall

- Follow-up #1:
  - Available 72 hours after
  - Median completion 90 hours
  - 88% completion

- Follow-up #2:
  - Available 14 days after
  - Median completion 16 days
  - 59% completion
Results: Follow-up Recall

- Did treatment group have an effect?

<table>
<thead>
<tr>
<th></th>
<th>Words</th>
<th>Letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Follow-up #1</td>
<td>82%</td>
<td>93%</td>
</tr>
<tr>
<td>Follow-up #2</td>
<td>56%</td>
<td>62%</td>
</tr>
</tbody>
</table>

$H_0$: Treatment had no effect

$H_A$: Worse recall in letters group

$p = 0.45$
Results: Follow-up Recall

- Many able to enter at least one chunk
- 10% of errors from reordering chunks
  - Could be accounted for
    - $\log_2(3!) \approx 2.6$ bits
Conclusion

- It is possible to teach users random passwords over time
- Authors developed a prototype system to do this
Thanks!

Questions?