Telepathwords: preventing weak passwords by reading users’ minds

Saranga Komanduri, Richard Shay, Lorrie Faith Cranor, Cormac Herley, Stuart Schechter
Authentication ecosystem
Authentication ecosystem

System administrators

Users

Adversaries
Defending users

• System administrators set *password policies*
  – Constraints on passwords
Character requirements

• Common component of policies
  – Length
  – Uppercase, digit, symbol
Character requirements

• Common component of policies
  – Length
  – Uppercase, digit, symbol

• Can’t prevent weak passwords:
Character requirements

• Default policies often use only character requirements

• In Microsoft Active Directory (3class8)
  – 3 of the following: uppercase, lowercase, digit, symbol
  – 8 character minimum

These requirements don’t improve security, and they make passwords harder to type
Goal

• Focus on weakest passwords
  – Threat model: online attack of an organization
  – Policies should make the weakest passwords harder to guess
Contributions

• Show that character requirements don’t prevent weak passwords

• Introduce Telepathwords
  – Detects weak passwords while providing real-time feedback

• Show that real-time feedback coupled with prevention of common patterns works well
  – Significantly better security than character requirements
  – Not significantly different in usability (after creation)
Telepathwords

• Real-time predictions

To help you avoid a weak password, we will try to guess each character of your password before you type it. 
\(\times\) indicates that the character below it was one we guessed, and does little to improve your password. 
\(\checkmark\) indicates that the character below it was hard to guess, and is more likely to improve your password.

Choose a password: 

Re-enter your password:

Continue
Telepathwords

- Explanations show why password is guessable

Best guesses for the next key you'll type:
- **M** as in *showmethemoney*
- **C** as in *showmethecurry*
- **W** as in *showmetheway*
Telepathwords

• Prediction display can be turned off
Telepathwords

- Feedback bar scores password by predictions

<table>
<thead>
<tr>
<th>M</th>
<th>as in showmethemoney</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>as in showmethecurry</td>
</tr>
<tr>
<td>W</td>
<td>as in showmetheway</td>
</tr>
</tbody>
</table>

= predicted

= unpredicted

Best guesses for the next key you’ll type

- Can enforce a policy by requiring a minimum number of unpredicted characters
Related work

• Character requirements haven’t changed much in 35 years since Morris and Thompson
  – Proposed 6-character minimum [1979]
  – Very little security improvement [Bonneau 2012]
Related work

• Password meters
  – Meters are typically based on character requirements [Ur et al. 2012]
  – No consistency across meters [de Carnevalet and Mannan 2013]
Related work

• Password meters
  – Meters are typically based on character requirements [Ur et al. 2012]
  – No consistency across meters [de Carnevalet and Mannan 2013]

• *zxcvbn* entropy estimator [Wheeler 2012]

Password meters don’t explain their scores
Generating predictions

• Multiple, weighted “predictors” produce next-character guesses and likelihood scores
Generating predictions

- Search query n-grams

Choose a password: showmethe\text{\underline{\textcolor{red}{M}}as in showmethemoney}
\text{\underline{\textcolor{red}{C}}as in showmethecurry}
\text{\underline{\textcolor{red}{W}}as in showmetheway}
Generating predictions

• Password sets

Choose a password:

Best guesses for the next key you'll type:
Generating predictions

- Common substitutions (s -> $, a -> @, etc.)

Choose a password:

P@$$

Was in password

W as in password

I as in passion

P as in passport
Generating predictions

• Keyboard patterns

Choose a password:

zaqwsxcde

Best guesses for the next key you'll type:

R
RFVBGT
3-v

(4 more ✔ marks required)
Generating predictions

• Keyboard patterns

Choose a password:

zaqw

Best guesses for the next key you'll type:

S as in zaqwsx
E as in zaqwer
A as in zaqwan

(4 more ✔ marks required)
Generating predictions

• Keyboard patterns

Choose a password:

zaqwsxcde

Best guesses for the next key you'll type:

4 more ✓ marks required

Rrfvbgt

RFVBGT

3-v
Generating predictions

• Repeating patterns

Choose a password:

\[ \text{abababab} \]

(4 more ✓ marks required)

Best guesses for the next key you'll type:

- Repeating ab (ab)
- AAB
- Y as in babababy
Generating predictions

• Interleaving strings

Choose a password:

(3 more ✓ marks required)

Best guesses for the next key you’ll type:

- O o-r-d
- O o-r-d
- E e-r-d
Generating predictions

• Can cover many behaviors and easily add

• Many possible ways to implement, ours is just one example
Evaluation

• December 2013 - deployed as a public website
  https://telepathwords.research.microsoft.com

• February 2014 - Mechanical Turk study
  – CMU branded, using Javascript API
Policies

- 6 policy conditions (2 Telepathwords)

- All conditions included some visual feedback
Condition: basic8

Requirements
– Minimum 8 characters
Condition: 3class8

- Minimum 8 characters
- Must contain at least 3 of the following: lowercase, uppercase, digit, symbol
Condition: 3class8-d

- Minimum 8 characters
- Must contain at least 3 of the following: lowercase, uppercase, digit, symbol
- Letters in password must not be in a dictionary

word for your email account.

Your password needs 8 more characters and must include lowercase letters, uppercase letters, digits, symbols.
Condition: 3class12

- Minimum 12 characters
- Must contain at least 3 of the following: lowercase, uppercase, digit, symbol
Conditions: telepath, telepath-v

- Minimum 6 unpredicted characters
- “Show Password” checked by default in telepath-v

Please create a new password for your email account.

To help you avoid a weak password, we will try to guess each character of your password before you type it.

Indicates that the character below it was one we guessed, and does little to improve your password.

Indicates that the character below it was hard to guess, and is more likely to improve your password.

Choose a password: 

Re-enter your password: 

Continue
Conditions

Character requirements

- basic8
- 3class8
- 3class12
- 3class8-d
- telepath
- telepath-v

“Dictionary” policies
Evaluation

• N = 2,844 (started) / 2,560 (finished)
• Median age = 27 (limited to 18+)
• 60% male, 44% with Bachelor's or above
• Required 95% acceptance rate and U.S. location
Study design

• Hypothetical email scenario for password creation

Steps:

1. Create a password under a randomly assigned condition
2. Take a survey
3. Recall password
4. Return in two days
Policy metrics

Security

• Weir+ guessability
  – Minimum number of guesses needed for single password, 2.5%, 5%, and 10%

• zxcvbn entropy estimate
  – Min-entropy; 2.5th, 5th, 10th percentiles

• Probability metrics were not viable
Policy metrics

Usability

• Creation difficulty
• Did participants find it insightful?
• Recall difficulty
Security results

Percent covered vs. zxcvbn entropy scores

- Line color: basic8
Security results

Percent covered

0.1%  1.0%  10.0%

zxcvbn entropy scores

basic8
Security results

- Percent covered
- Basic8

Graph showing security results with entropy scores on the x-axis and percent covered on the y-axis.
Security results

Percent covered vs. zxcvbn entropy scores
Security results

Percent covered

zxcvbn entropy scores

Stronger policies

basic8
3class8
3class12
3class8–d
telepath
telepath–v
Security results

Percent covered

zxcvbn entropy scores

- green: basic8
- orange: 3class8
- purple: 3class12
- pink: 3class8–d
- green: telepath
- yellow: telepath–v
Security results

Percent cracked

Weir+ guess numbers

basic8
3class8
3class12
3class8–d
telepath
telepath–v
Security results

password

Password1

Asdfghjkl123

Percent cracked

0.1%

1.0%

10.0%

Weir+ guess numbers

10^0

10^3

10^6

10^9

basic8

3class8

3class12

3class8–d

telepath

telepath–v
Security results

![Graph showing security results with lines for different passwords: Pokemon91, iamabeliever, 1024scott. The x-axis represents Weir+ guess numbers on a logarithmic scale from $10^0$ to $10^9$, and the y-axis represents the percent cracked from 0.1% to 10.0%. The lines indicate the percentage of passwords cracked for different password types: basic8, 3class8, 3class12, 3class8–d, telepath, telepath–v.]

- Pokemon91
- iamabeliever
- 1024scott
Security results

Percent cracked vs. Weir+ guess numbers for different classes:
- basic8
- 3class8
- 3class12
- 3class8-d
- telepath
- telepath-v
Usability: Recall after 2—5 days

Recalled in < 5 tries

- basic8
- 3class8
- 3class12
- 3class8-d
- telepath
- telepath-v
Usability: Creation time

- basic8
- 3class8
- 3class12
- 3class8-d
- telepath
- telepath-v

Median Time to Create Password (s)
Usability: Creation sentiment

- **Creation Difficult**
- **Creation Annoying**

<table>
<thead>
<tr>
<th></th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>basic8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3class8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3class12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3class8-d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>telepath</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>telepath-v</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Usability

Feedback Gave Insight

- basic8
- 3class8
- 3class12
- 3class8-d
- telepath
- telepath-v

0%  Strongly agree  Agree  Neutral  Disagree  100%  Strongly disagree
Conclusions

• “Dictionary” policies with real-time feedback can help users avoid weak passwords
  – Usability cost on creation
  – Telepathwords’ feedback gave insight into password strength
Conclusions

• “Dictionary” policies with real-time feedback can help users make stronger passwords
  – Usability cost on creation
  – Telepathwords’ feedback gave insight into password strength

• Character-class requirements had little to no effect on security using our metrics
Returned after 2-5 days

Percent Returned

- basic8
- 3class8
- 3class12
- 3class8–d
- telepath
- telepath–v
Usability: Stored password

The diagram shows the percentage of stored passwords for different conditions. The conditions are labeled as basic8, 3class8, 3class12, 3class8-d, telepath, and telepath-v. The y-axis represents the percentage of stored passwords, ranging from 0% to 100%. The x-axis lists the conditions that were tested.
Usability: Recall after 2—5 days

Did not store password

Stored password
Usability: Toggled Show Password

Toggled "Show Password" Box

- basic8
- 3class8
- 3class12
- 3class8-d
- telepath
- telepath-v