Do Users' Perceptions of Password Security Match Reality?

Carnegie Mellon University

Blase Ur, Jonathan Bees, Sean M. Segreti, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor

Perception vs. Reality

123456789
12345
password
iloveyou
princess
rockyou
1234567
12345678
abc123
nicole
daniel
babygirl
monkey
lovely
jessica
654321
michael
ashley
qwerty
111111
iloveu
000000
michelle
tigger
sunshine
chocolate
password1
soccer
anthony
friends
butterfly
purple
angel
jordan
liverpool
justin
loveme
football
basketball
superman
tinkerbell
pretty
secret
hottie
bubbles
jennifer
lovers
spongebob
angels
andrew
playboy
666666
jasmine
andrea
carlos
eminem
flower
shadow
melissa
summer
chelsea
barbie
elizabeth
brandon
robert
dragon
cookie
matthew
softball
whatever
computer
hannah
samantha
william
sakura
christian
thomas
steven
america

Compare actual strength of passwords to users' perceptions

How strong is a particular password actually?

How strong is a particular password actually?

Modeling Guessing Attacks

Modeling Guessing Attacks

- Data-driven password-guessing attacks

Modeling Guessing Attacks

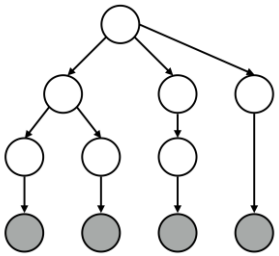
- Data-driven password-guessing attacks
  - Previously stolen passwords

Modeling Guessing Attacks

- Data-driven password-guessing attacks
  - Previously stolen passwords
  - Natural-language corpora

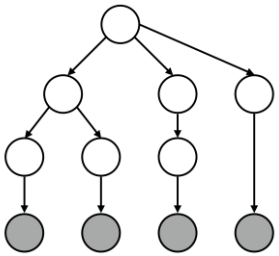
Modeling Guessing Attacks

- Data-driven password-guessing attacks
  - Previously stolen passwords
  - Natural-language corpora
- Simulated cracking software & algorithms



Modeling Guessing Attacks

- Data-driven password-guessing attacks
  - Previously stolen passwords
  - Natural-language corpora
- Simulated cracking software & algorithms
  - CMU Password Guessability Service



How strong do people think a password is?

Measuring Perceptions

- Online study
  - Compensated $5 for ~30 minutes

Measuring Perceptions

- Online study
  - Compensated $5 for ~30 minutes
- 165 participants from Mechanical Turk

Measuring Perceptions

- Online study
  - Compensated $5 for ~30 minutes
- 165 participants from Mechanical Turk
  - Age 18+, live in United States

Measuring Perceptions

- Online study
  - Compensated $5 for ~30 minutes
- 165 participants from Mechanical Turk
  - Age 18+, live in United States
  - Median age 33
  - 49% female, 51% male

Measuring Perceptions

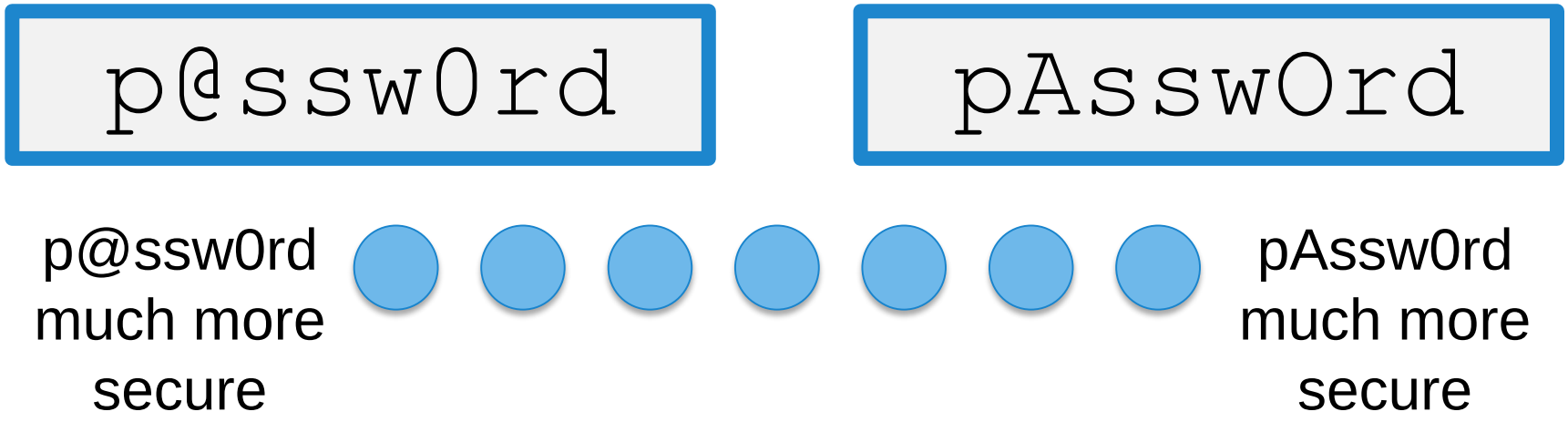
- Online study
  - Compensated $5 for ~30 minutes
- 165 participants from Mechanical Turk
  - Age 18+, live in United States
  - Median age 33
  - 49% female, 51% male
  - 16% CS or related degree or job
  - 4% student/professional in computer security

Study Tasks

1. Evaluating password pairs

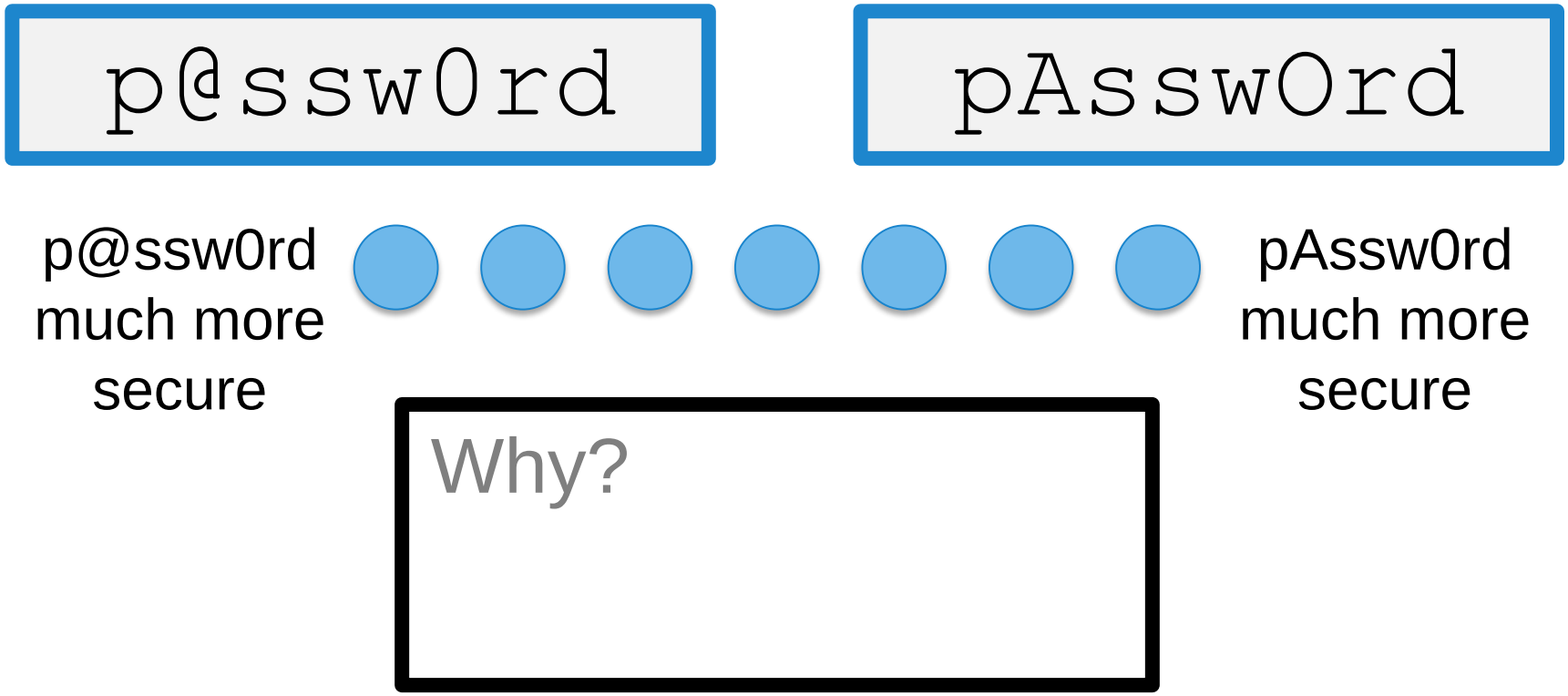
Study Tasks

1. Evaluating password pairs



Study Tasks

1. Evaluating password pairs



Task 1 Hypotheses

- 25 common characteristics, e.g.,
  - Capitalization
  - Letters vs. digits vs. symbols
  - Choice of words and phrases

Task 1 Hypotheses

- 25 common characteristics, e.g.,
  - Capitalization
  - Letters vs. digits vs. symbols
  - Choice of words and phrases
- Created 3 pairs per hypothesis
  - Randomly chose 1 pair per participant

Task 1 Hypotheses

- 25 common characteristics, e.g.,
  - Capitalization
  - Letters vs. digits vs. symbols
  - Choice of words and phrases
- Created 3 pairs per hypothesis
  - Randomly chose 1 pair per participant
  - At least one password per pair from rockyou

Study Tasks

1. Evaluating password pairs
2. Rating selected passwords

Study Tasks

1. Evaluating password pairs
2. Rating selected passwords

Please rate the **security** of the following password: `rolltide`

Please rate the **memorability** of the following password: `rolltide`

Study Tasks

1. Evaluating password pairs
2. Rating selected passwords
3. Rating creation strategies

Study Tasks

1. Evaluating password pairs
2. Rating selected passwords
3. Rating creation strategies
4. Describing attackers
   - Who, why, how

Results

1. Evaluating password pairs
2. Rating selected passwords
3. Rating creation strategies
4. Describing attackers

Evaluating Password Pairs

`iloveyou88`

`ieatkale88`

Evaluating Password Pairs

`iloveyou88`

`ieatkale88`

Evaluating Password Pairs

`iloveyou88` `ieatkale88`

Evaluating Password Pairs

`iloveyou88`

`ieatkale88`

4,000,000,000 × more secure!

Evaluating Password Pairs

`brooklyn16`

`brooklynqy`

Evaluating Password Pairs

`brooklyn16`

`brooklynqy`

Evaluating Password Pairs

`brooklyn16`

`brooklynqy`

Evaluating Password Pairs

`brooklyn16`

`brooklynqy`

300,000 ×
more secure!

Ways People Were Wrong

- Overstated security benefits of:
  - Digits
  - Character substitutions (e.g., a→@)
  - Keyboard patterns (e.g., 1qaz2wsx3edc)
- Did not recognize common words/phrases

Many Ways People Were Right

- Capitalize letters other than the first
- Put digits and symbols in middle, not end
- Use symbols rather than digits
- Avoid:
  - Common first names
  - Words related to account
  - Years and sequences

If perceptions of many individual characteristics are correct, then why do people make bad passwords?

Perceptions of Attackers

Perception: How Many Guesses?

Perception: How Many Guesses?

- 2 guesses (Min)

Perception: How Many Guesses?

- 2 guesses (Min)
- 100,000,000,000,000,000,000,000,000, 000,000,000,000,000,000,000,000,000, 000,000 guesses (Max)

Perception: How Many Guesses?

- 2 guesses (Min)
- 100,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000 guesses (Max)
- 34% ≤ 50 guesses (manual attack)

Perception: How Many Guesses?

- 2 guesses (Min)
- 100,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000 guesses (Max)
- 34% ≤ 50 guesses (manual attack)
- 67% ≤ 50,000 guesses (small-scale)

Perception: How Many Guesses?

- 2 guesses (Min)
- 100,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000 guesses (Max)
- 34% ≤ 50 guesses (manual attack)
- 67% ≤ 50,000 guesses (small-scale)
- 7% ≥ 10^{14} guesses (large-scale)

Reality: How Many Guesses?

Reality: Small-Scale Guessing

Reality: Small-Scale Guessing

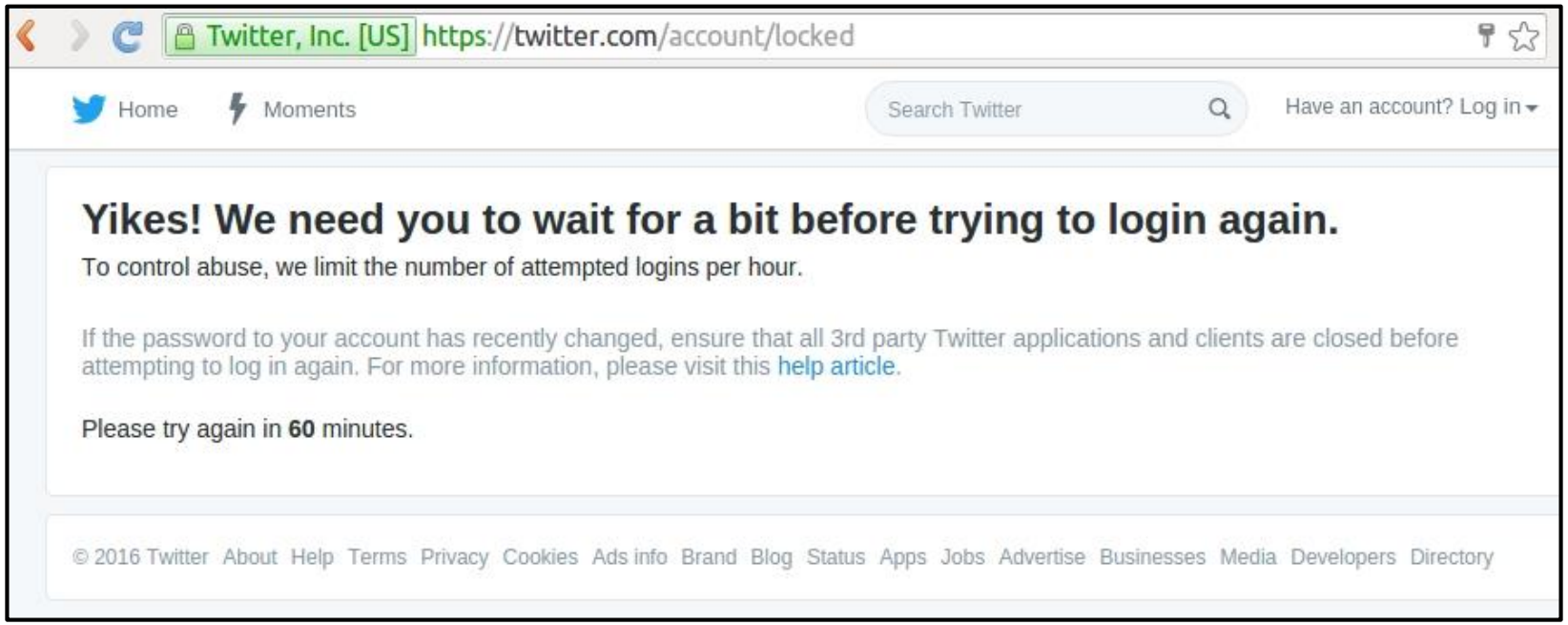
- Targeted guessing by someone you know

Reality: Small-Scale Guessing

- Targeted guessing by someone you know
- Automated attack by a stranger

Reality: Small-Scale Guessing

- Targeted guessing by someone you know
- Automated attack by a stranger



Reality: Small-Scale Guessing

- Targeted guessing by someone you know
- Automated attack by a stranger
- 1 – 1,000,000 guesses

Reality: Large-Scale Guessing

Reality: Large-Scale Guessing

- Against stolen database of passwords

Reality: Large-Scale Guessing

- Against stolen database of passwords

Domino's Pizza

LinkedIn

YAHOO!

Adobe

GAWKER

000webhost.com
better than paid hosting

SONY

STRATFOR
GLOBAL INTELLIGENCE

Reality: Large-Scale Guessing

- Against stolen database of passwords
- Against password-protected file

Reality: Large-Scale Guessing

- Against stolen database of passwords
- Against password-protected file
- 1,000,000 guesses (best practices)

Reality: Large-Scale Guessing

- Against stolen database of passwords
- Against password-protected file
- 1,000,000 guesses (best practices)
- 10^{14} or more (common reality)

Perception

Small-scale

67% ≤ 50,000

Reality

Small-scale…

…and large-scale

≥ 10^{14} guesses

Limitations

- MTurk sample not generalizable
  - Younger, more technical

Limitations

- MTurk sample not generalizable
  - Younger, more technical
- Password security context-dependent
  - Account value
  - Expectations of attack

Limitations

- MTurk sample not generalizable
  - Younger, more technical
- Password security context-dependent
  - Account value
  - Expectations of attack
- No model is perfect

Conclusions

Conclusions

- Perceptions of individual characteristics
  - Often consistent with current attacks
  - Some crucial differences

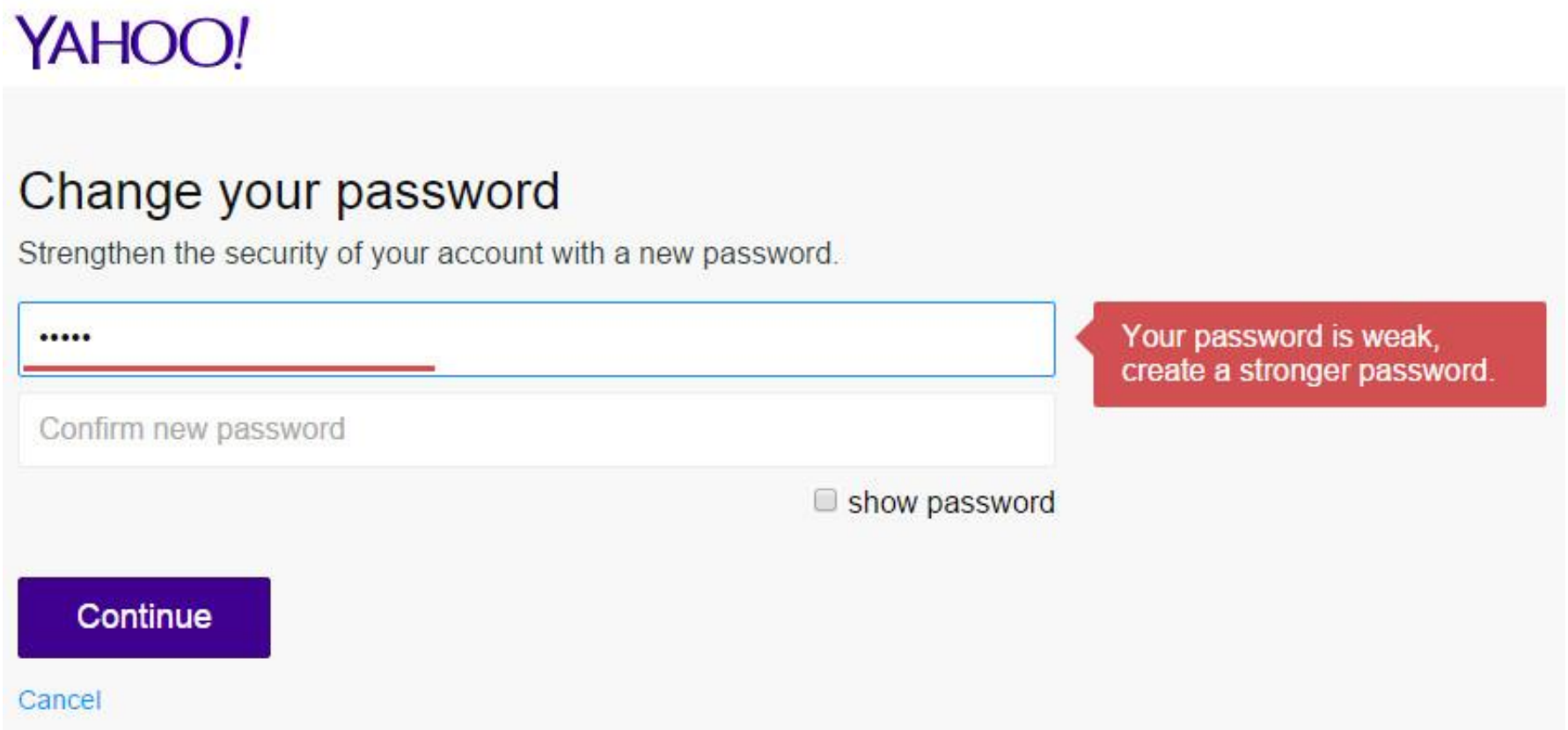
Conclusions

- Perceptions of individual characteristics
  - Often consistent with current attacks
  - Some crucial differences
- Huge variance in perceptions of attackers

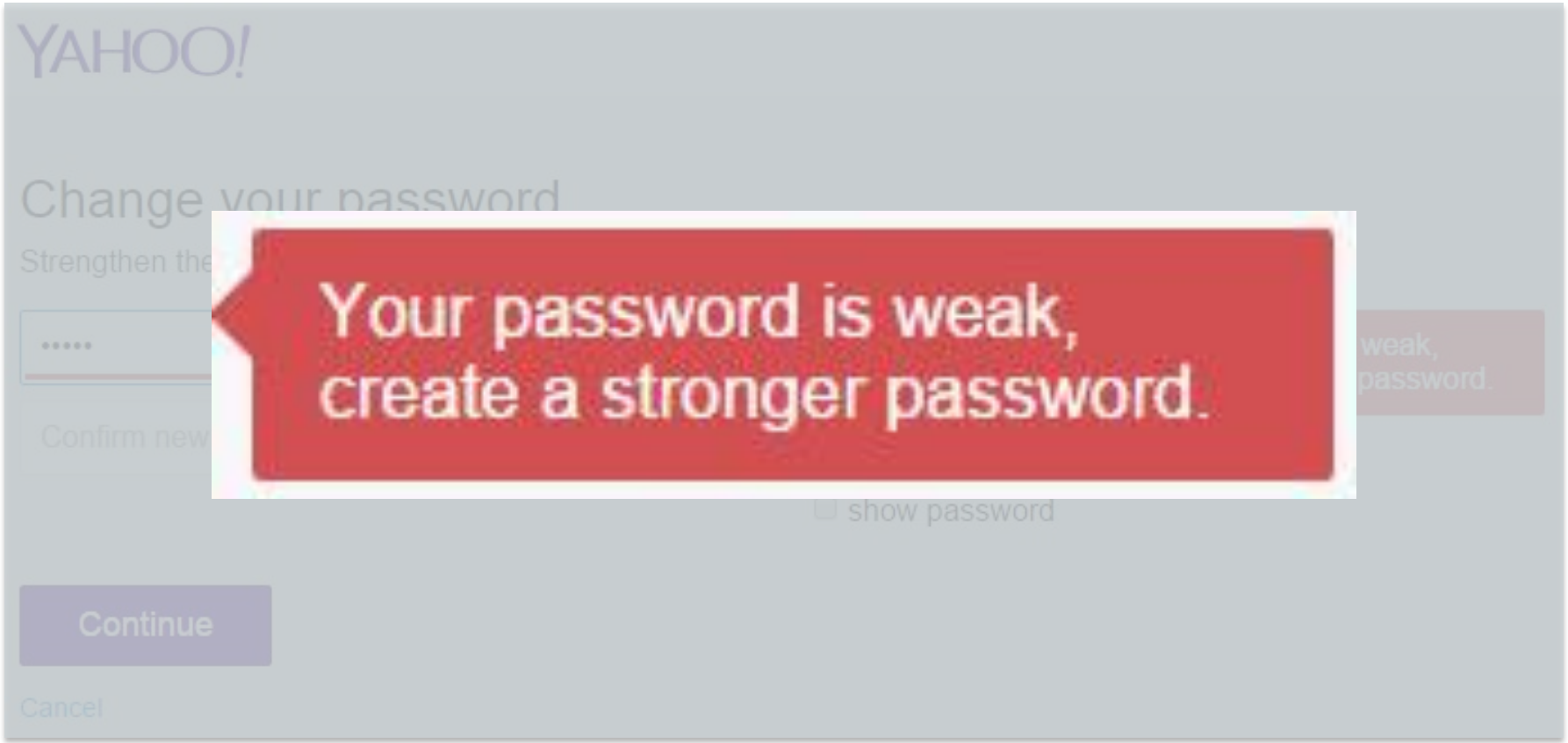
Conclusions

- Perceptions of individual characteristics
  - Often consistent with current attacks
  - Some crucial differences
- Huge variance in perceptions of attackers
- Current user feedback is insufficient

Current User Feedback Insufficient



Current User Feedback Insufficient



Conclusions

- Perceptions of individual characteristics
  - Often consistent with current attacks
  - Some crucial differences
- Huge variance in perceptions of attackers
- Current user feedback is insufficient

Do Users' Perceptions of Password Security Match Reality?

Carnegie Mellon University

Blase Ur, Jonathan Bees, Sean M. Segreti, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor