Let’s Go in for a Closer Look: Observing Passwords in Their Natural Habitat

Sarah Pearman, Jeremy Thomas, Pardis Emami Naeini,
Hana Habib, Lujo Bauer, Nicolas Christin,
Lorrie Faith Cranor, Serge Egelman, Alain Forget

To be presented at ACM CCS 2017

Carnegie Mellon University
Societal Computing

CyLab
Security and Privacy Institute

CUPS
CyLab Usable Privacy & Security Laboratory

How do people manage all their passwords?

Risks of password reuse

CRACKED PASSWORDS

UserID	Password
jane	iloveyou89
john	godoggo!
jur	monkey1
kar	pa$$word
katie	princ3ss2

Risks of password reuse

CRACKED PASSWORDS

UserID	Password
jane	iloveyou89
john	godoggo!
jur	monkey1
kar	pa$$word
katie	princ3ss2

Online Store

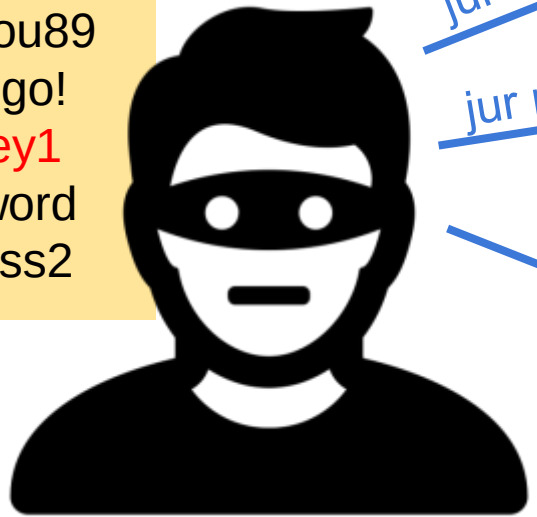
Bank

Employer

Risks of password reuse

CRACKED PASSWORDS

UserID	Password
jane	iloveyou89
john	godoggo!
jur	monkey1
kar	pa$$word
katie	princ3ss2



jur monkey1 → Online Store

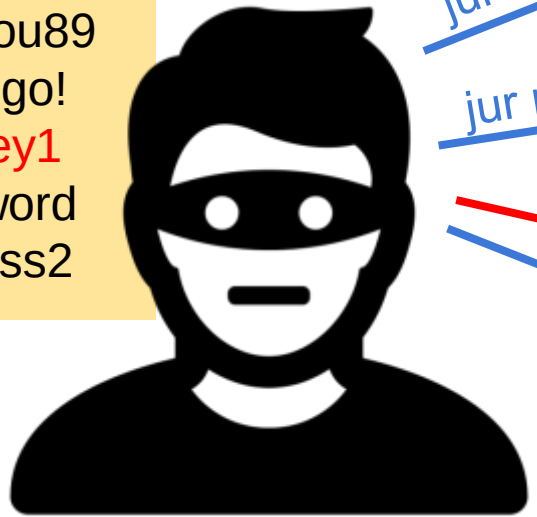
jur monkey1 → Bank

jur monkey1 → Employer

Risks of password reuse

CRACKED PASSWORDS

UserID	Password
jane	iloveyou89
john	godoggo!
jur	monkey1
kar	pa$$word
katie	princ3ss2



jur monkey1 → Online Store

jur monkey1 → Bank

jur monkey2 → Employer

jur monkey1 → Employer

Difficult to observe password use across all of a user’s accounts

Difficult to observe password use across all of a user's accounts

- Most past work relies on indirect or incomplete data sources
  - Interviews and surveys rely on self reports
  - Leaked password databases allow examination of one password per user
  - Passwords created specifically for studies don't provide insights into password portfolio

Difficult to observe password use across all of a user's accounts

- Most past work relies on indirect or incomplete data sources
  - Interviews and surveys rely on self reports
  - Leaked password databases allow examination of one password per user
  - Passwords created specifically for studies don't provide insights into password portfolio
- Wash et al 2016 analyzed password use over 6 weeks for 134 students

Security Behavior Observatory (SBO)

- Home Windows computers instrumented with SBO client and browser extensions
- Data from >500 participants since 2014
- Current active participants: ~200
- Allows for empirical observation, scientific analysis of security and privacy behavior

Password data collection

Deployed new SBO browser extensions in January 2017 to collect

- Hashes of passwords and 4+ character substrings
- Length, strength, characters in each class (uppercase, lowercase, digits, special characters)
- Browsing metadata (including URL)

Accurate strength measurement

- Neural network guesser (Melicher et al. 2016)
- Accurate measurement of guessability by sophisticated attacker
- Runs in browser plugin (no need to transmit password)
- No noticeable delay to user

Domain	Password	Reuse Type
nytimes.com	**Usabl3!!**	
yahoo.com	**s3curity**	
google.com	**security123**	
cmu.edu	**p4$$w0rd**	
facebook.com	**p4$$w0rd**	
amazon.com	**security?**	
twitter.com	**security?**	

*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com	**Usabl3!!**	Not reused
yahoo.com	**s3curity**	
google.com	**security123**	
cmu.edu	**p4$$w0rd**	
facebook.com	**p4$$w0rd**	
amazon.com	**security?**	
twitter.com	**security?**	

*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com	**Usabl3!!**	Not reused
yahoo.com	**s3curity**	
google.com	**security123**	
cmu.edu	**p4$$w0rd**	
facebook.com	**p4$$w0rd**	
amazon.com	**security?**	
twitter.com	**security?**	

*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com	Usabl3!!	Not reused
yahoo.com	s3curity	Partially reused
google.com	security123	
cmu.edu	p4$$w0rd	
facebook.com	p4$$w0rd	
amazon.com	security?	
twitter.com	security?	

*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com	Usabl3!!	Not reused
yahoo.com	s3curity	Partially reused
google.com	security123	
cmu.edu	p4$$w0rd ←	
facebook.com	p4$$w0rd ←	
amazon.com	security?	
twitter.com	security?	

*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com	Usabl3!!	Not reused
yahoo.com	s3curity	Partially reused
google.com	security123	
cmu.edu	p4$$w0rd	Exactly reused
facebook.com	p4$$w0rd	
amazon.com	security?	
twitter.com	security?	

*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com	`Usabl3!!`	Not reused
yahoo.com	`s3curity`	Partially reused
google.com	`security123`	
cmu.edu	`p4$$w0rd`	Exactly reused
facebook.com	`p4$$w0rd`	
amazon.com	`security?` ←	
twitter.com	`security?` ←	

*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com	Usabl3!!	Not reused
yahoo.com	s3curity	Partially reused
google.com	security123	
cmu.edu	p4$$w0rd	Exactly reused
facebook.com	p4$$w0rd	
amazon.com	security?	
twitter.com	security?	

*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com	Usabl3!!	Not reused
yahoo.com	s3curity	Partially reused
google.com	security123	
cmu.edu	p4$$w0rd	Exactly reused
facebook.com	p4$$w0rd	
amazon.com	security?	Partially AND exactly reused
twitter.com	security?	

*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com	`Usabl3!!`	Not reused
yahoo.com	`s3curity`	Partially reused
google.com	`security123`	
cmu.edu	`p4$$w0rd`	Exactly reused
facebook.com	`p4$$w0rd`	
amazon.com	`security?`	Partially AND exactly reused
twitter.com	`security?`	

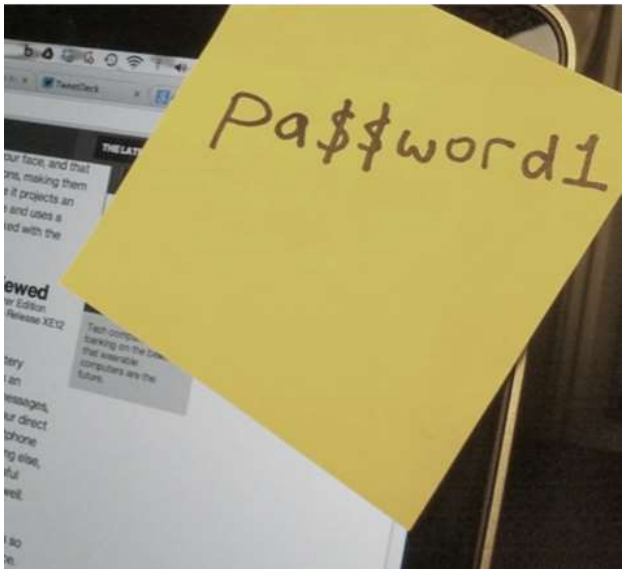
*These are fictitious passwords; we do not record actual plaintext passwords

Domain	Password	Reuse Type
nytimes.com 1	Usabl3!!	Not reused
yahoo.com 2	s3curity	Partially reused
google.com 3	security123	
cmu.edu 4	p4$$w0rd	Exactly reused
facebook.com	p4$$w0rd	
amazon.com 5	security?	Partially AND exactly reused
twitter.com	security?	

*These are fictitious passwords; we do not record actual plaintext passwords

Overview of data collected

- 154 participants
- 31-217 days in study (Mean 147 days)
- 4,057 passwords
- Mean 26.3 passwords/participant
- 2,077 different web domains



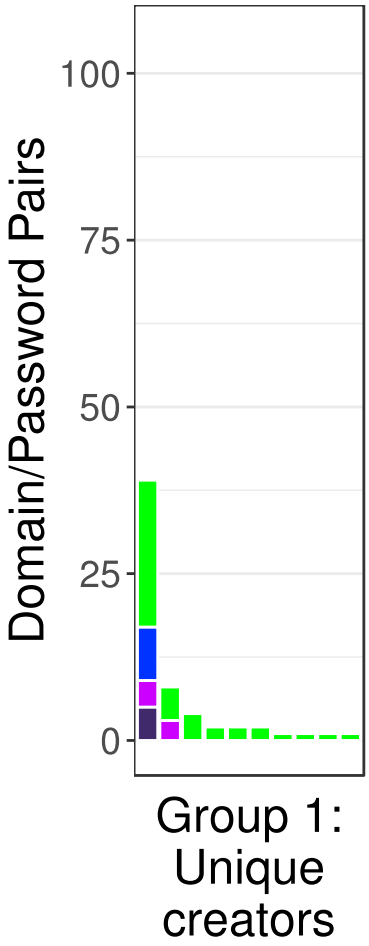
Data continues to be collected

Password reuse per participant

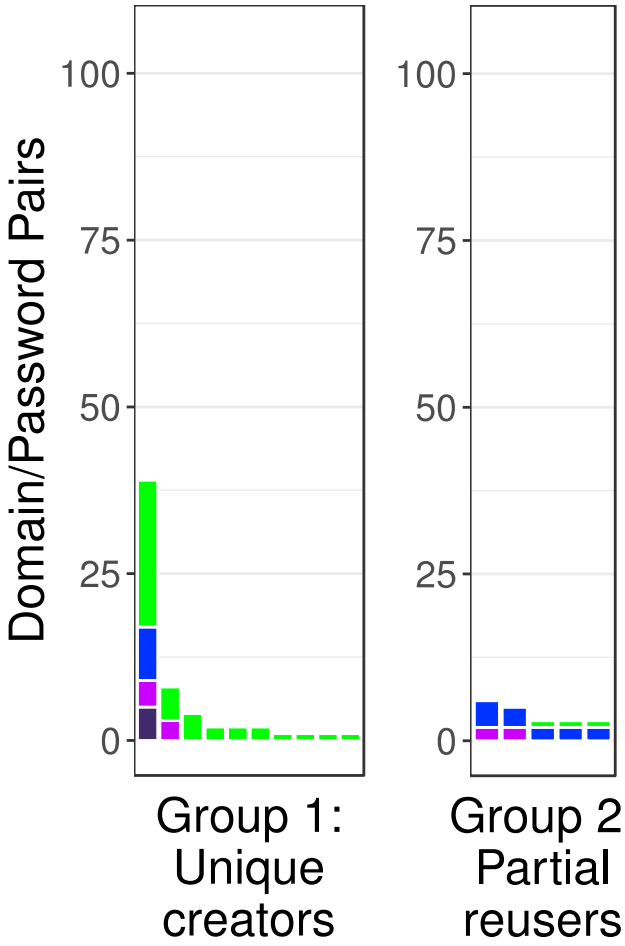
- 26.30 different accounts (domain/password pairs)
- 9.88 distinct passwords

Percentages of accounts:

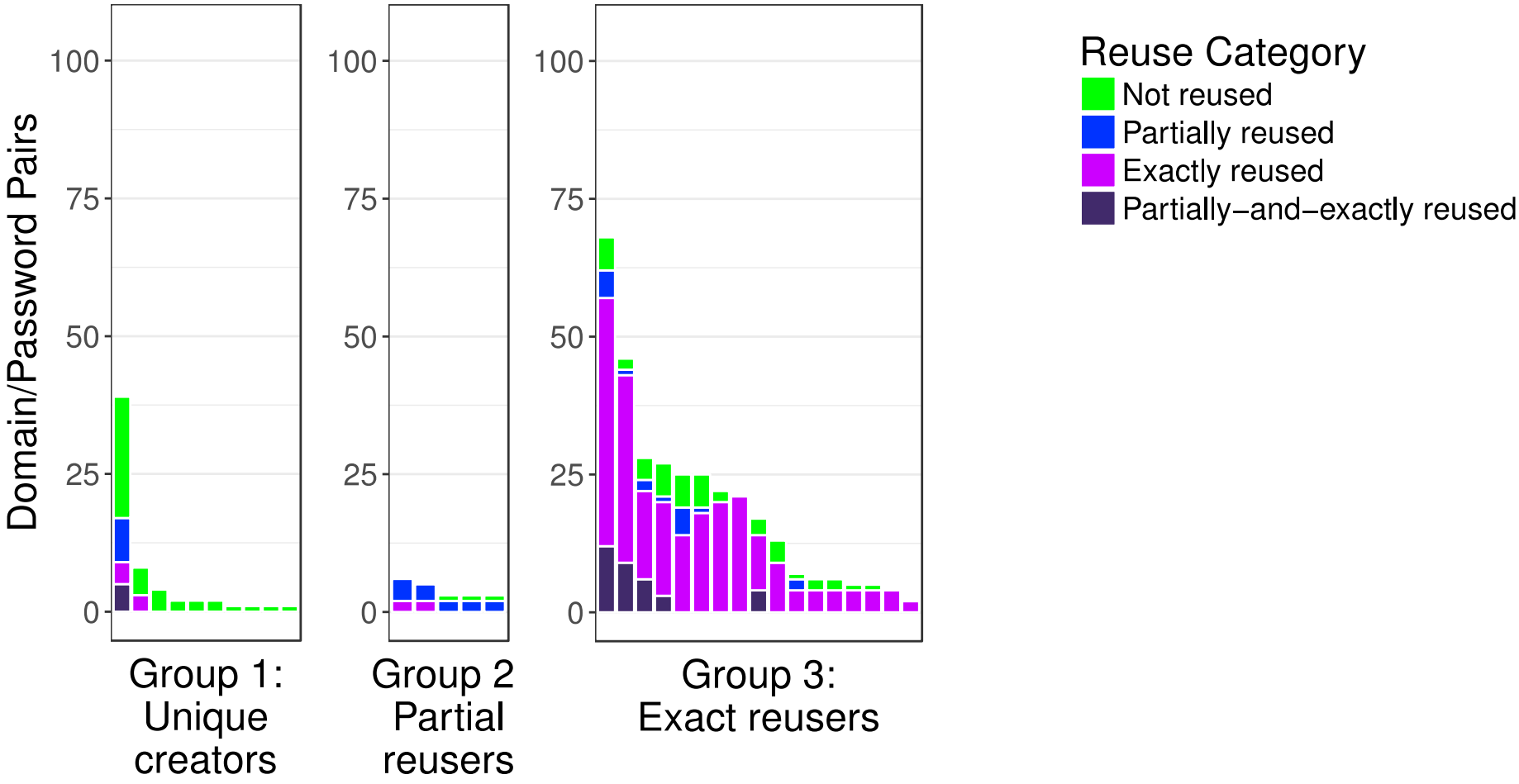
- With non-reused passwords: 21%
- With only-partially-reused passwords: 12%
- With only-exactly-reused passwords: 16%
- With exactly-and-partially-reused passwords: 51%

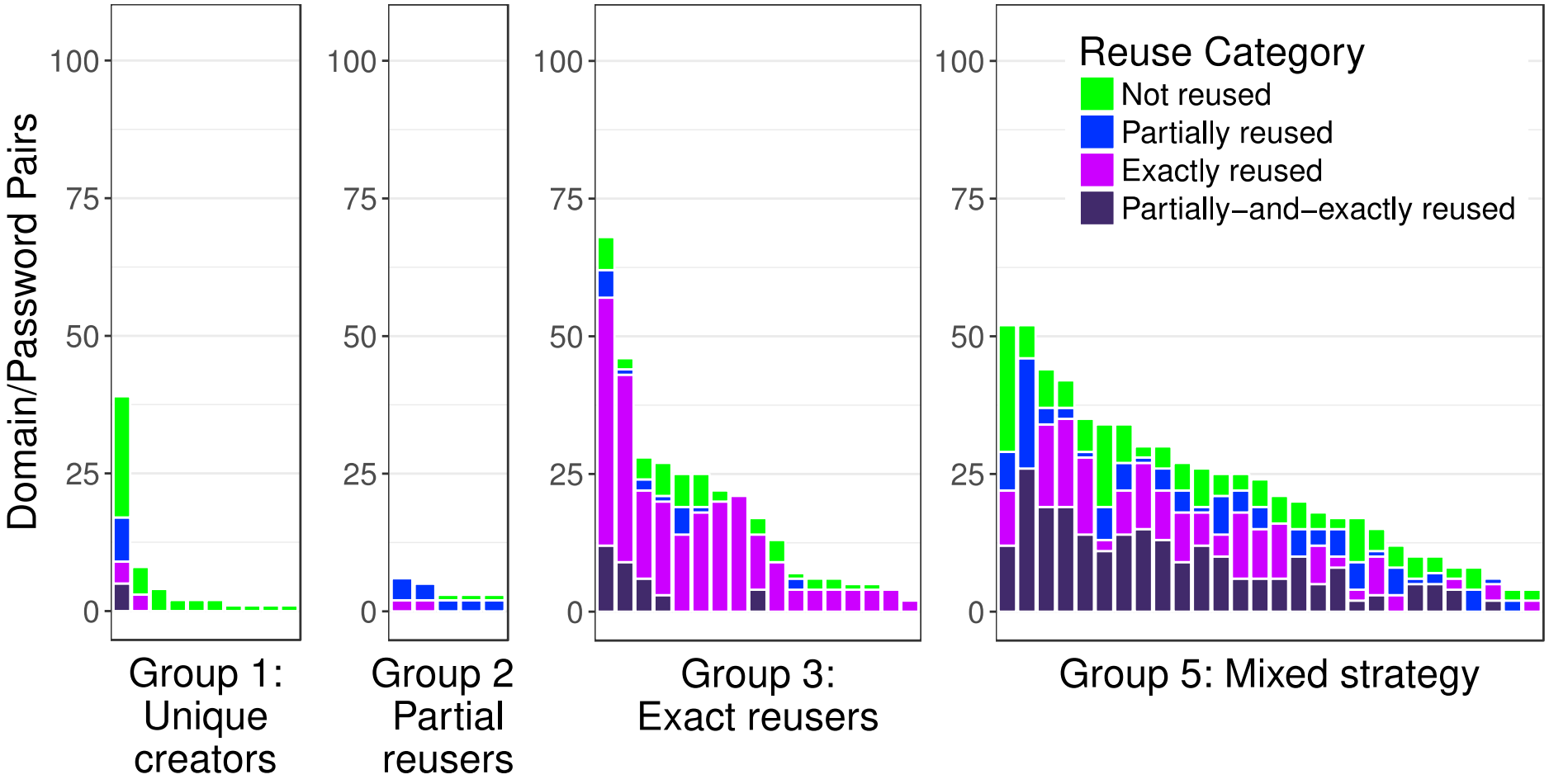


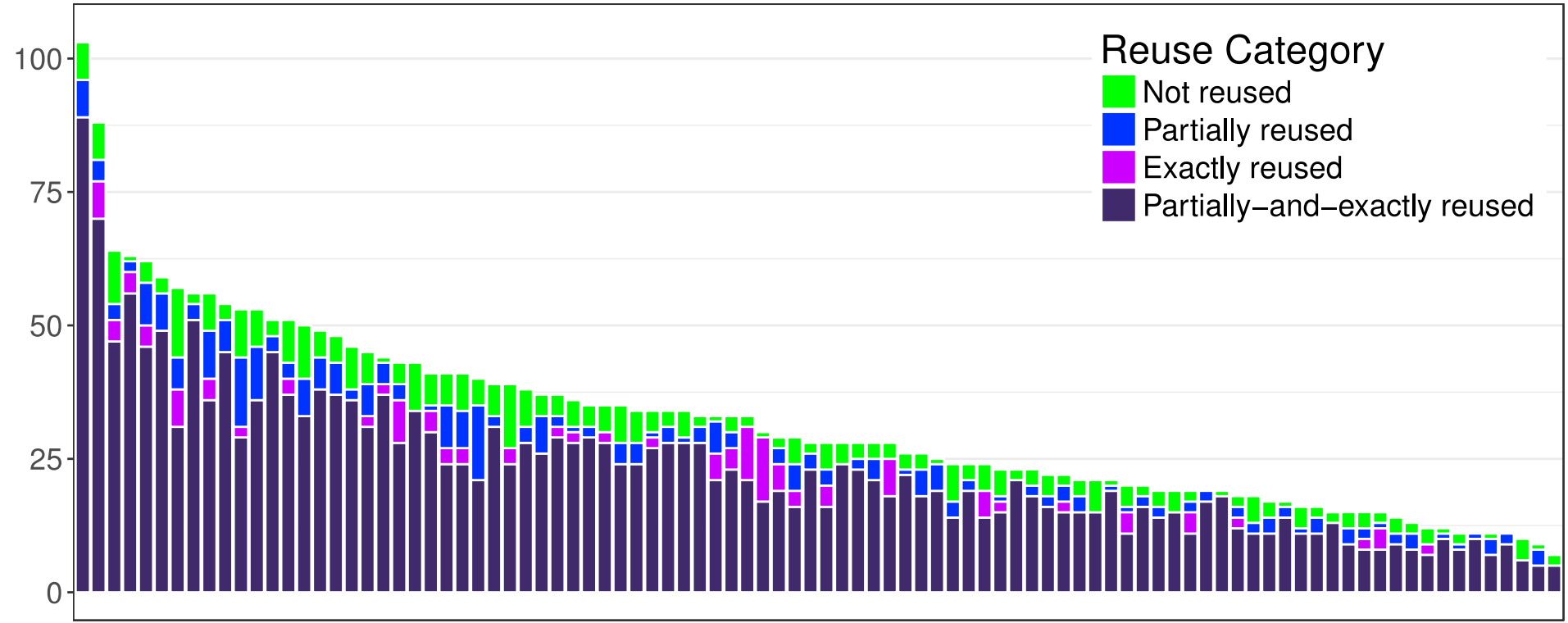
Reuse Category
- Not reused
- Partially reused
- Exactly reused
- Partially–and–exactly reused



Reuse Category
- Not reused
- Partially reused
- Exactly reused
- Partially–and–exactly reused

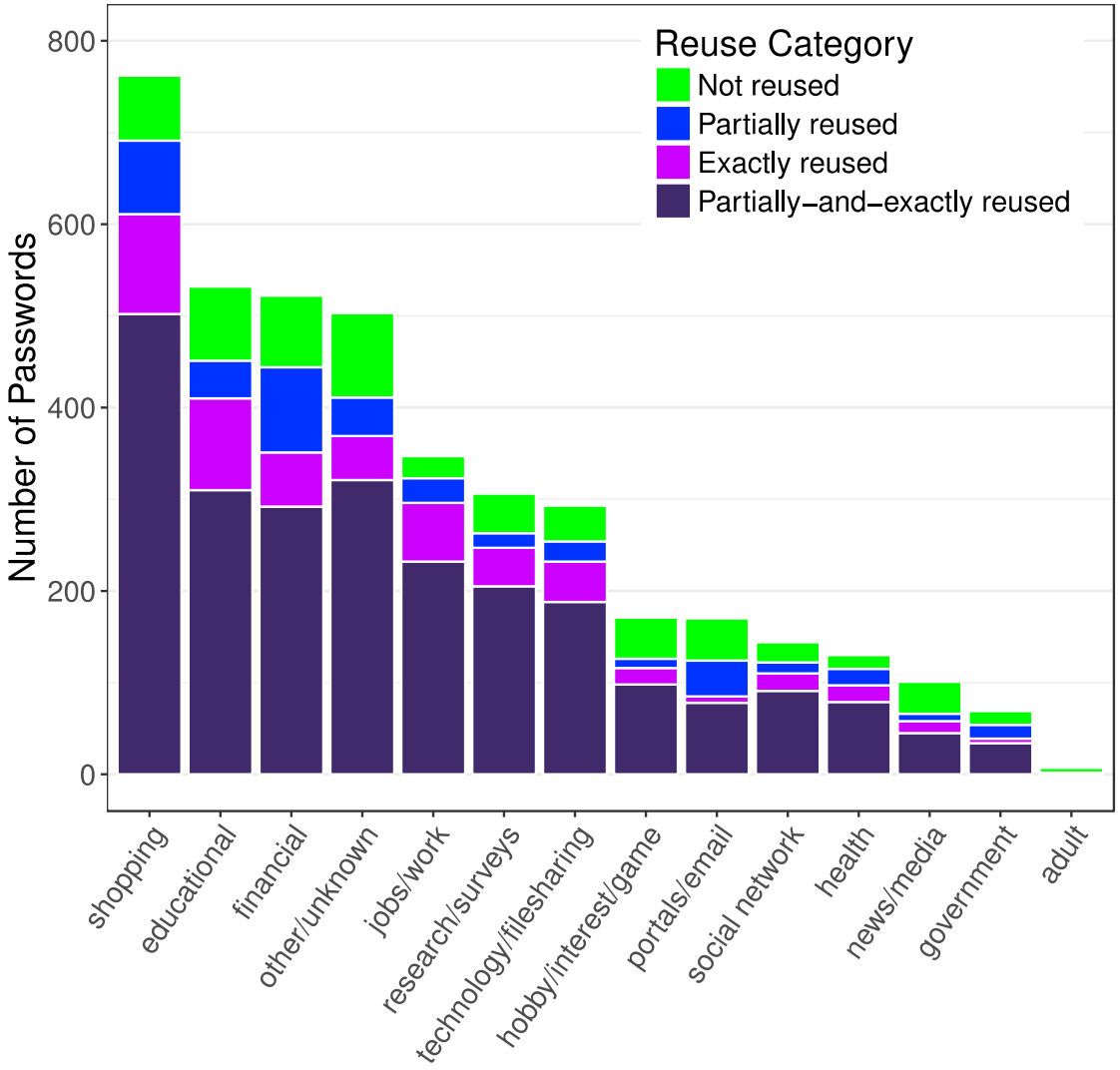

Domain/Password Pairs
100
75
50
25
0
Group 1:
Unique
creators
Group 2
Partial
reusers
Group 3:
Exact reusers
Reuse Category
Not reused
Partially reused
Exactly reused
Partially–and–exactly reused


Domain/Password Pairs
100
75
50
25
0
Group 1: Unique creators
Group 2 Partial reusers
Group 3: Exact reusers
Group 5: Mixed strategy
Reuse Category
Not reused
Partially reused
Exactly reused
Partially–and–exactly reused

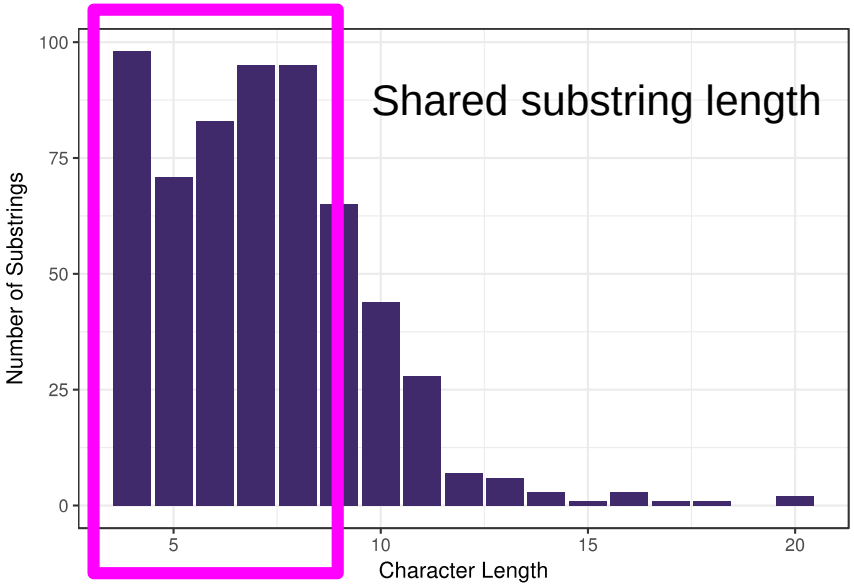


Group 4: Exact–and–partial reusers

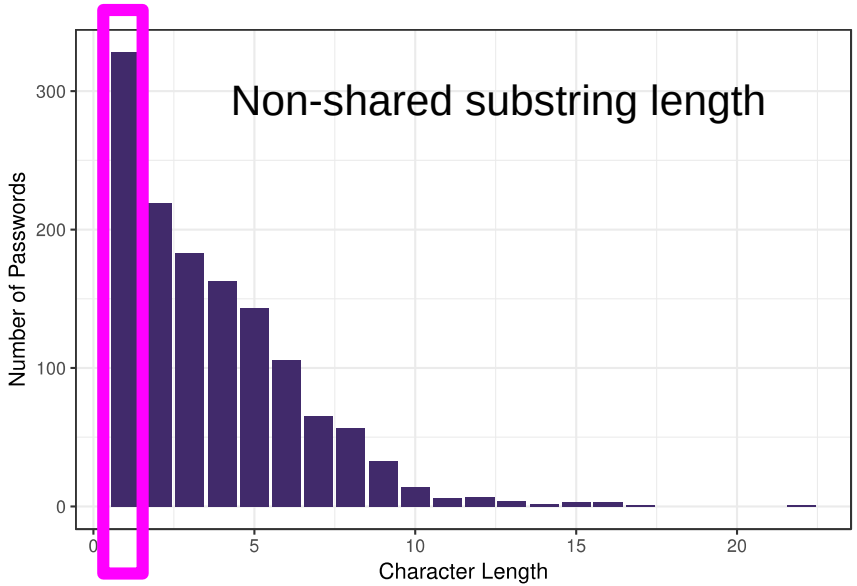
Lots of reuse across almost all categories of websites



Typical partial reuse: 4-8 characters



4-8 characters shared



1character not shared

Password managers not helping

Password managers not helping

- Password managers and use of autofill had no significant effect on
  - frequency of password reuse
  - password strength

Password managers not helping

- Password managers and use of autofill had no significant effect on
  - frequency of password reuse
  - password strength
- Users may be using password managers to store and autofill passwords, but not to generate random passwords

Reuse is rampant

- SBO provides unique opportunity to study how users manage large numbers of passwords
- Users cope with password demands through a mix of reuse strategies
- Password managers may not be helping very much

SBO project: **http://cups.cs.cmu.edu/sbo.html**