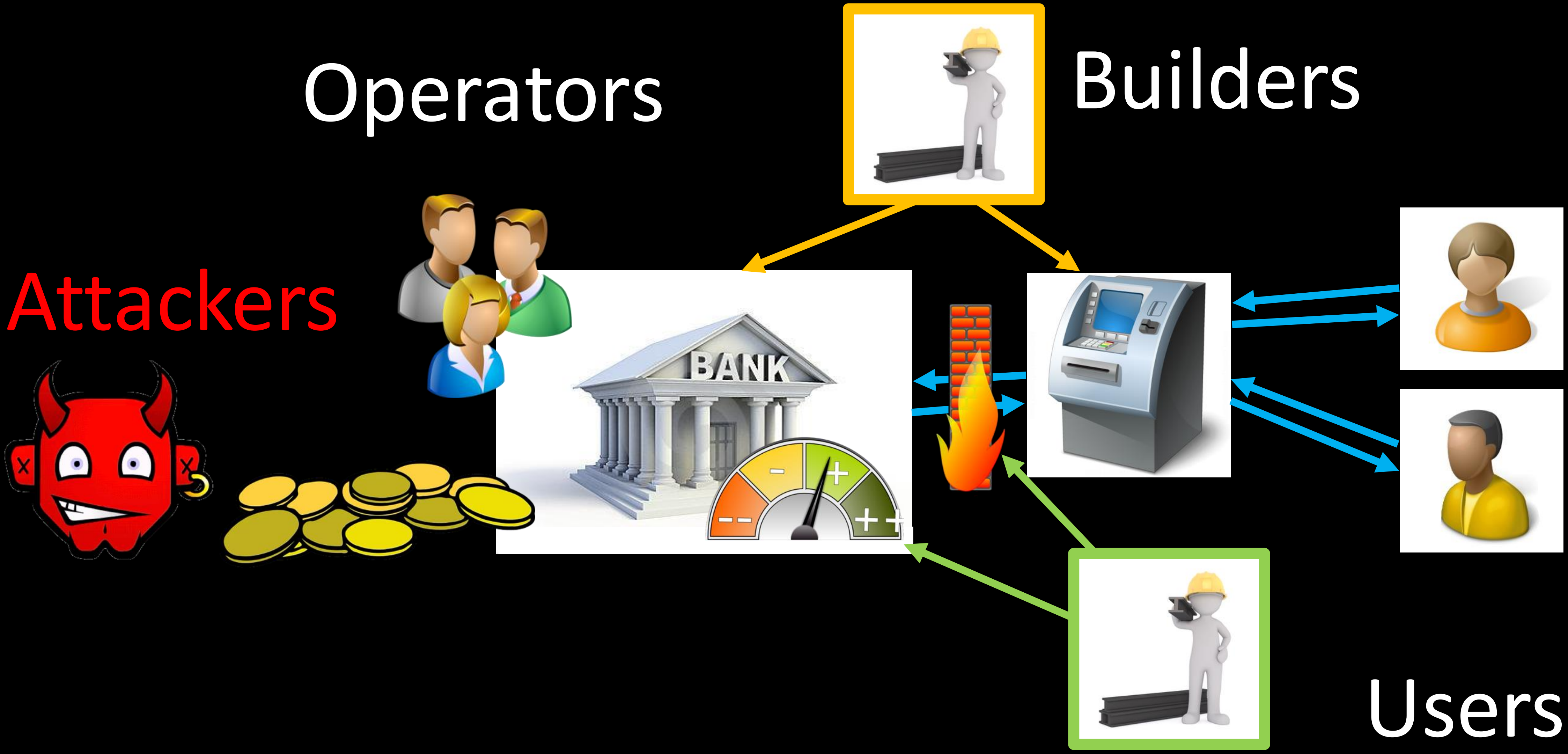
Secure Systems Engineering and Management

A Data-driven Approach

Exploiting the Human

Michael Hicks

UPenn CIS 7000-003
Spring 2026



Operators

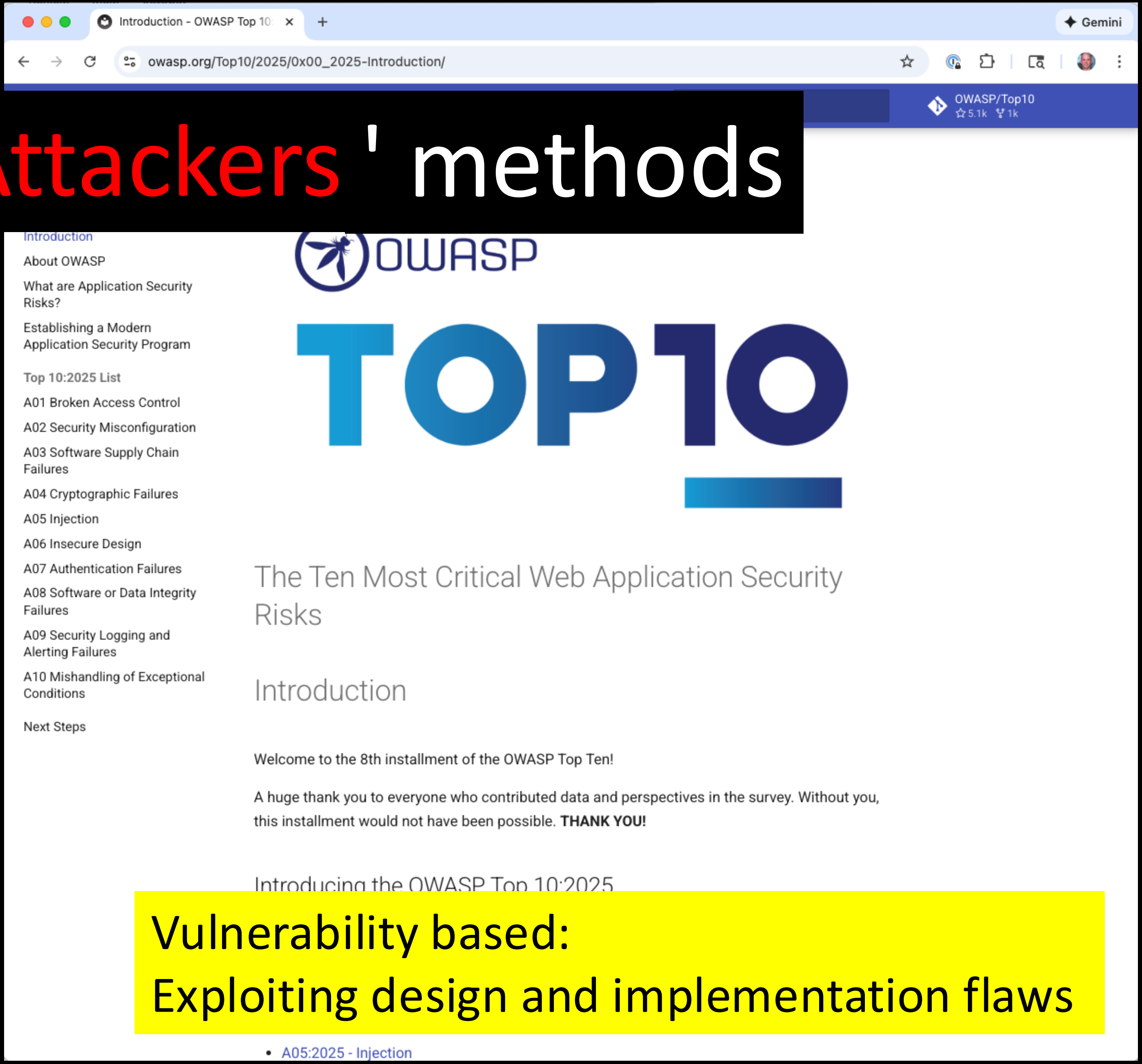
Builders

Attackers

Users

Cyber Builders

Attackers ' methods


Introduction - OWASP Top 10
owasp.org/Top10/2025/0x00_2025-Introduction/
Gemini
OWASP/Top10
☆5.1k 1k

Introduction

About OWASP

What are Application Security Risks?

Establishing a Modern Application Security Program

Top 10:2025 List

A01 Broken Access Control

A02 Security Misconfiguration

A03 Software Supply Chain Failures

A04 Cryptographic Failures

A05 Injection

A06 Insecure Design

A07 Authentication Failures

A08 Software or Data Integrity Failures

A09 Security Logging and Alerting Failures

A10 Mishandling of Exceptional Conditions

Next Steps

OWASP TOP 10

The Ten Most Critical Web Application Security Risks

Introduction

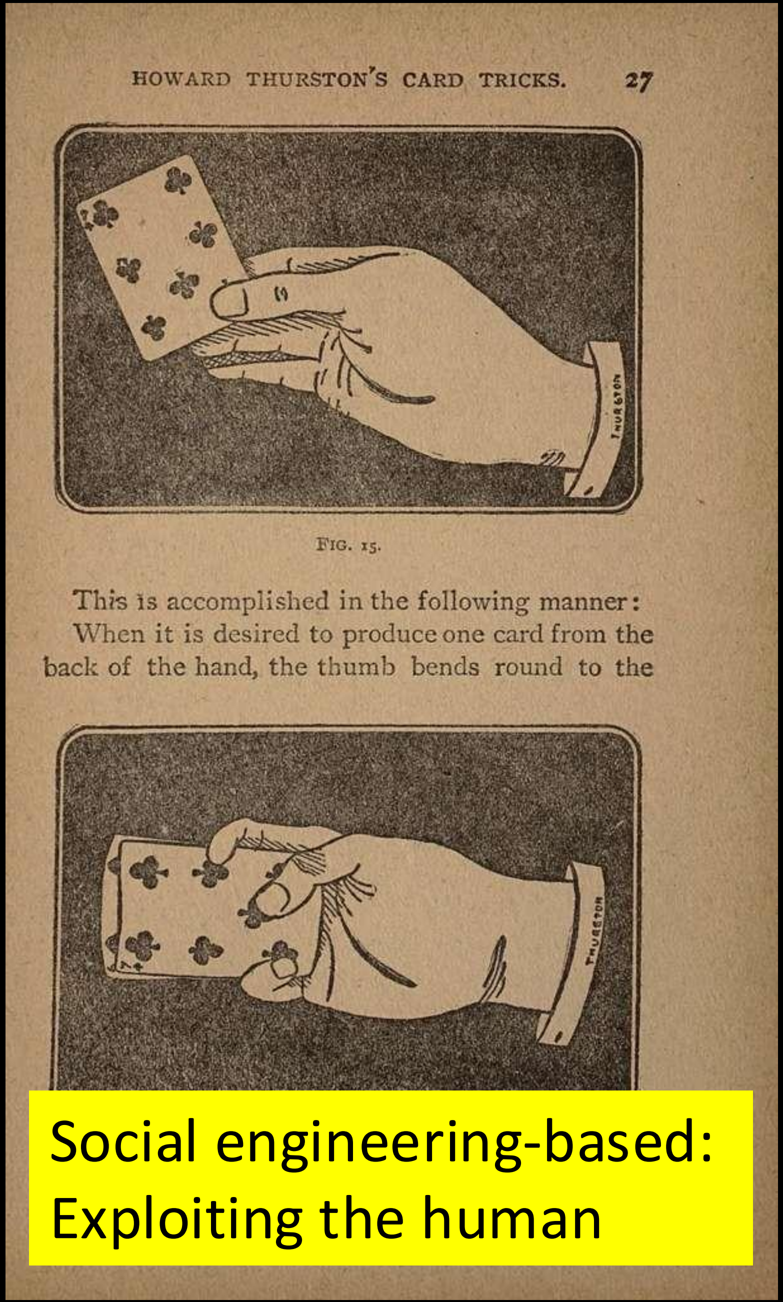
Welcome to the 8th installment of the OWASP Top Ten!

A huge thank you to everyone who contributed data and perspectives in the survey. Without you, this installment would not have been possible. **THANK YOU!**

Introducing the OWASP Top 10:2025

- A05:2025 - Injection

Vulnerability based:
Exploiting design and implementation flaws



HOWARD THURSTON'S CARD TRICKS. 27

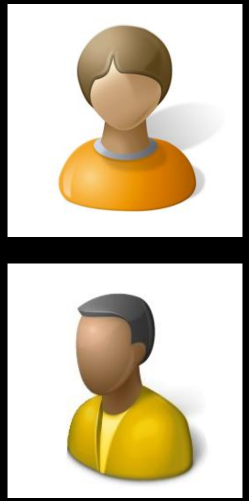
FIG. 15.

This is accomplished in the following manner: When it is desired to produce one card from the back of the hand, the thumb bends round to the

Social engineering-based:
Exploiting the human

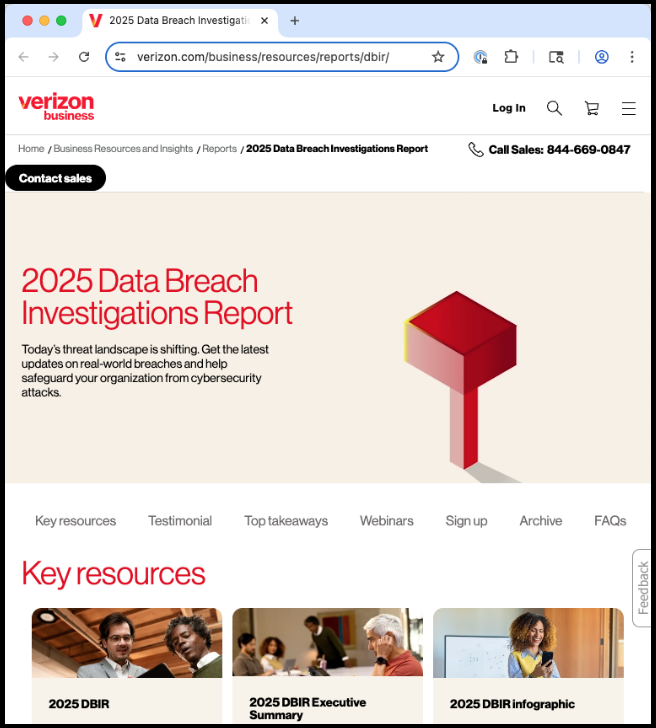
Attackers ' methods

Social engineering-based:
Exploiting the human



Users

Top attack vectors



2025 Data Breach Investigati
verizon.com/business/resources/reports/dbir/
verizon business
Log In

Home / Business Resources and Insights / Reports / 2025 Data Breach Investigations Report

Call Sales: 844-669-0847

Contact sales

2025 Data Breach Investigations Report

Today's threat landscape is shifting. Get the latest updates on real-world breaches and help safeguard your organization from cybersecurity attacks.

Key resources

2025 DBIR

2025 DBIR Executive Summary

2025 DBIR infographic

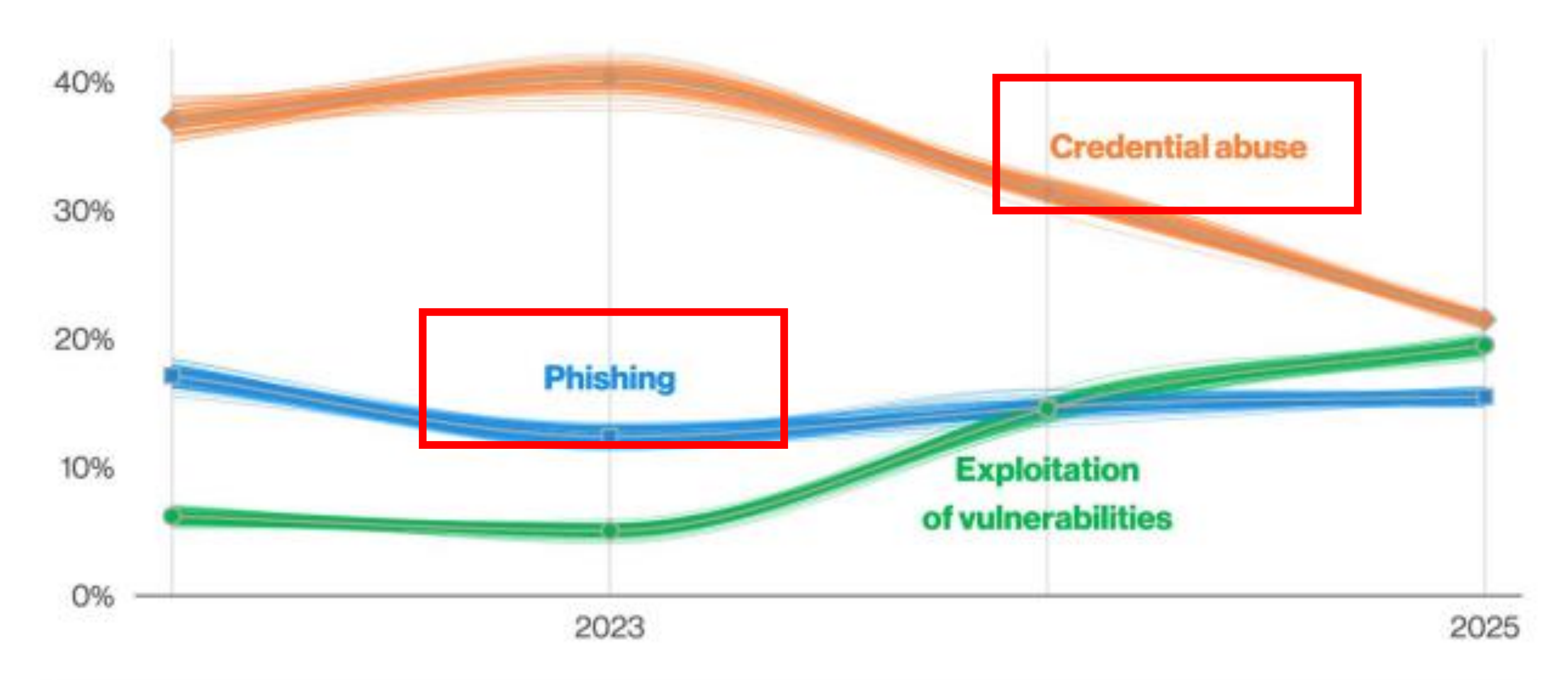


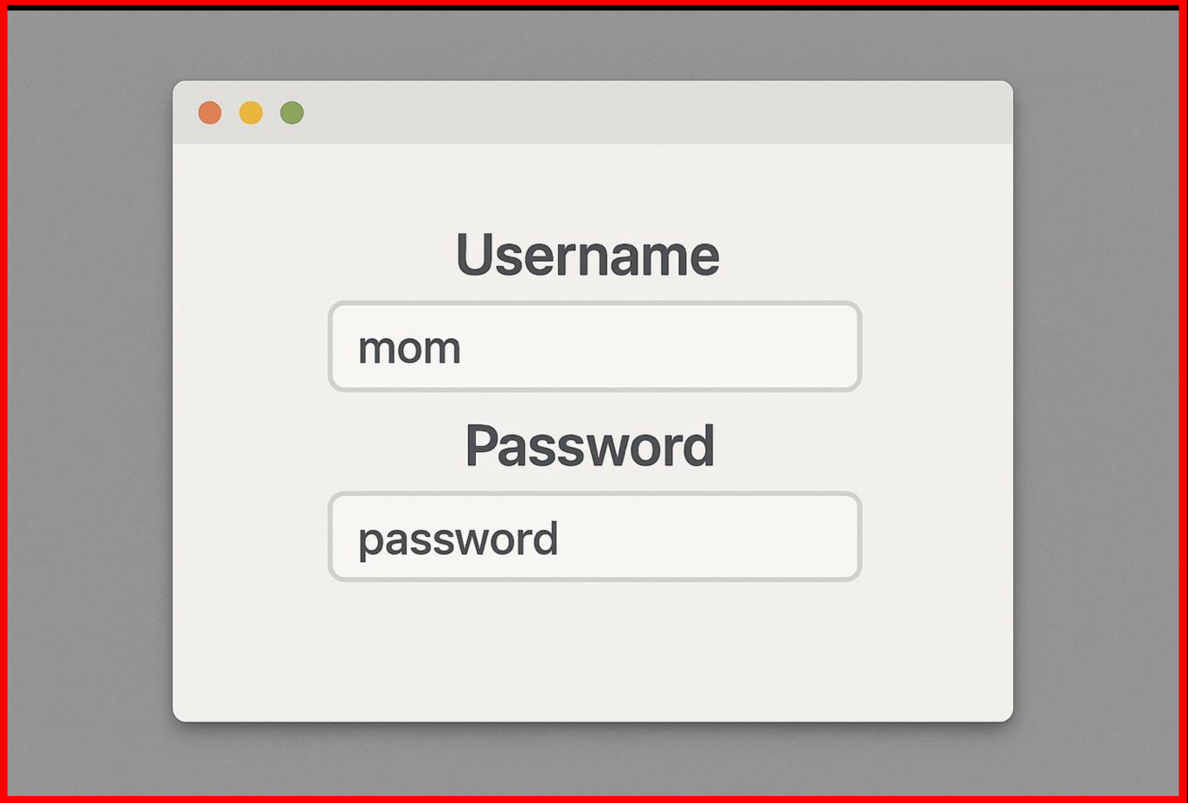
Figure 16. Known inital access vectors over time in non-Error, non-Misuse breaches (n in 2025 dataset=9,891)

Source: 2025 Verizon Data Breach Investigations Report

Credential abuse

Unauthorized use of **stolen**, **leaked**, or **otherwise compromised login credentials** (usernames and passwords). Acquired by:

- **Phishing** – tricking users into entering credentials on fake pages
- **Data breaches** – credentials stolen from one service and sold or leaked
- **Credential stuffing** – using breached username/password pairs from one site to try logging into other sites, exploiting password reuse
- **Brute forcing** – systematically guessing weak or common passwords
- **Infostealers** – malware that harvests saved passwords from browsers or password managers
- **Purchasing on dark web markets** – credentials are actively traded commodities



Username

mom

Password

password

Are you capable of remembering a unique strong password for every account you have?

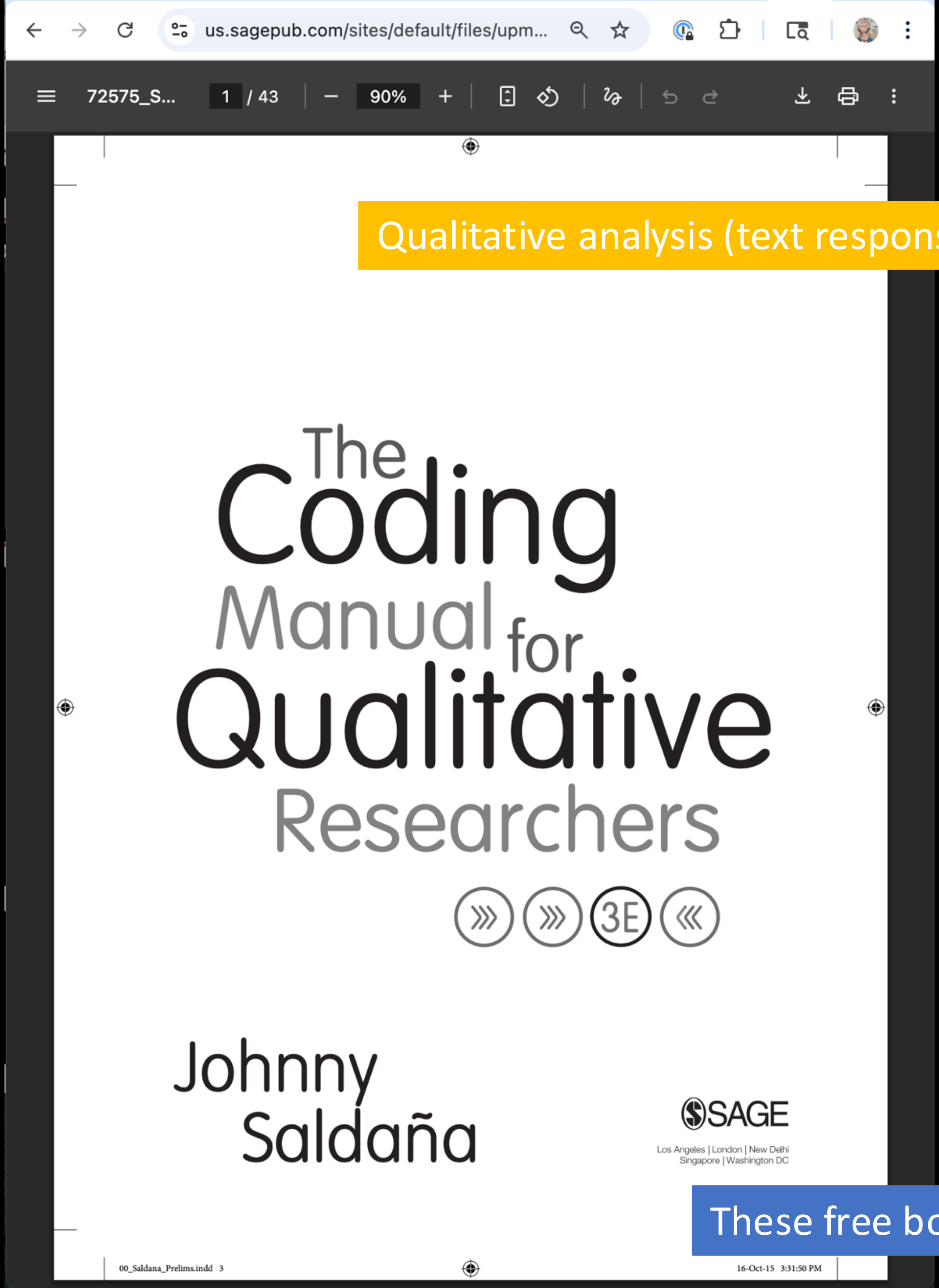
Passwords

- Goal**: easy to remember** but **hard to guess**
  - Turns out to be **wrong** in many cases!
    - Hard to guess = Hard to remember!
  - **Compounding problem**: repeated password use
- **Password cracking tools** train on released data to quickly guess common passwords
  - John the Ripper, http://www.openwall.com/john/
  - Project Rainbow, http://project-rainbowcrack.com/
  - many more …
- Top 10 worst passwords of 2023: 123456789 Qwerty Password 12345 Qwerty123 1q2w3e 12345678 111111 1234567890

https://wesecureapp.com/blog/worlds-worst-passwords-is-it-time-to-change-yours/

Hypothesis: Password reuse is common, improves chances of breach

- How would we know if this is true?
- Do an experimental study!
  - Find a good source of data
  - Consider what you'd find in that data if the hypothesis were true
    - Ideally: Very unlikely to be another explanation
  - Analyze the data, report the results
  - Consider consequences of the outcome; do follow-on work



us.sagepub.com/sites/default/files/upm...

72575_S... 1 / 43 90%

The Coding Manual for Qualitative Researchers

3E

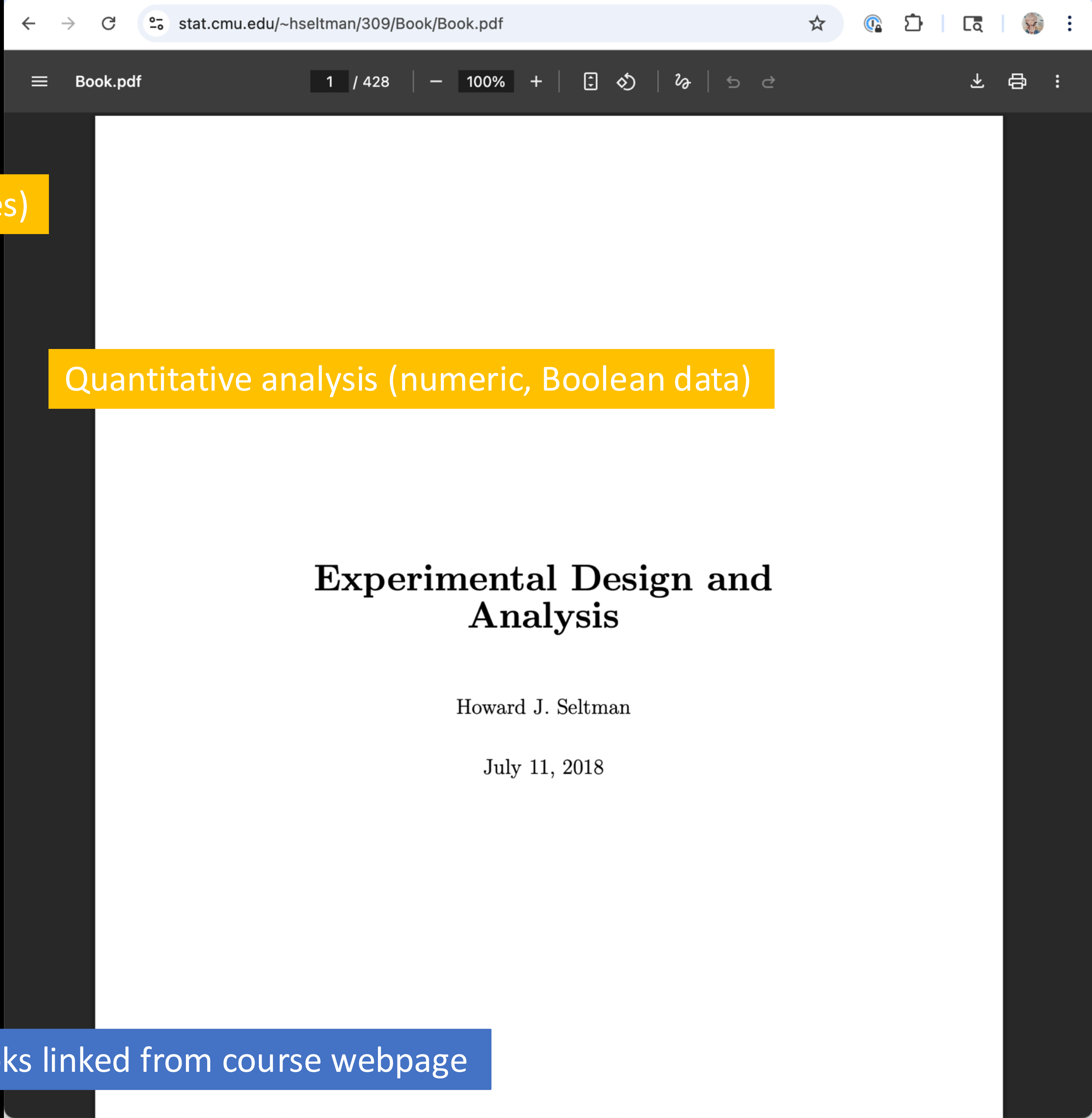
Johnny Saldaña

SAGE
Los Angeles | London | New Delhi
Singapore | Washington DC

00_Saldana_Prelims.indd 3

16-Oct-15 3:31:50 PM

Qualitative analysis (text responses)



stat.cmu.edu/~hseltman/309/Book/Book.pdf

Book.pdf 1 / 428 100%

Experimental Design and Analysis

Howard J. Seltman

July 11, 2018

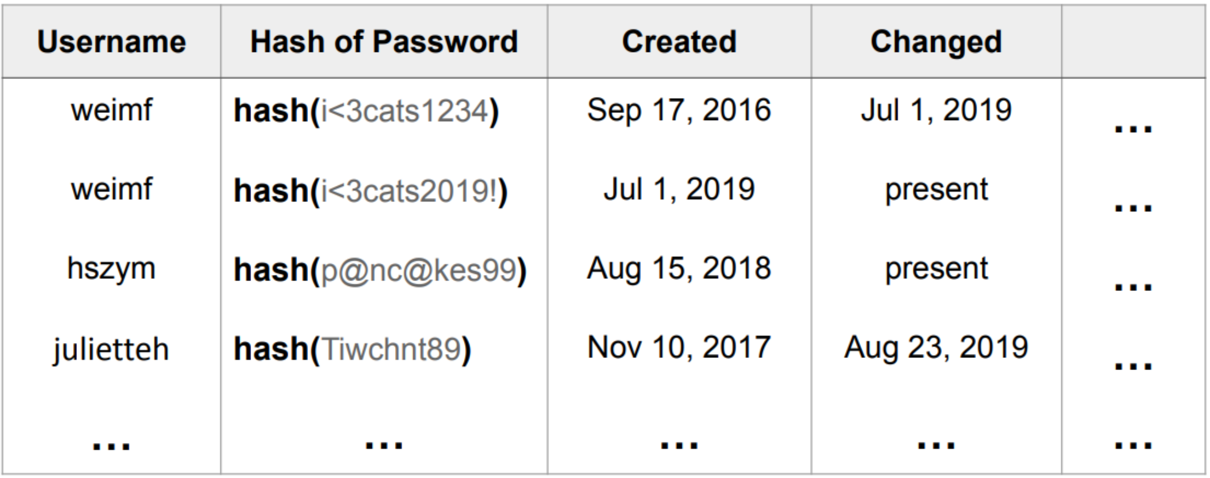
Quantitative analysis (numeric, Boolean data)

These free books linked from course webpage

Research study on password reuse

Approach:

1. Find credentials in leaked data
2. Generate guesses against university accounts
   - Using normal cracking, but also variations of leaked passwords for similar accounts
   - Check against historical password database:



Username	Hash of Password	Created	Changed	
weimf	hash(i<3cats1234)	Sep 17, 2016	Jul 1, 2019	...
weimf	hash(i<3cats2019!)	Jul 1, 2019	present	...
hszym	hash(p@nc@kes99)	Aug 15, 2018	present	...
julietteh	hash(Tiwchnt89)	Nov 10, 2017	Aug 23, 2019	...
...	...	...	...	...

3. Ask: Cracking easier with cross-account data?

https://www.usenix.org/conference/usenixsecurity23/presentation/nisenoff-retrospective

A Two-Decade Retrospective Analysis of a University's Vulnerability to Attacks Exploiting Reused Passwords

Alexandra Nisenoff[†*], Maximilian Golla[†‡], Miranda Wei[†⋆], Juliette Hainline[†], Hayley Szymanek[†], Annika Braun[†], Annika Hildebrandt[†], Blair Christensen[†], David Langenberg[†], Blase Ur[†]
† University of Chicago, ∗ Carnegie Mellon University,
‡ Max Planck Institute for Security and Privacy, ⋆ University of Washington

Abstract

Credential-guessing attacks often exploit passwords that were reused across a user's online accounts. To learn how organizations can better protect users, we retrospectively analyzed our university's vulnerability to credential-guessing attacks across twenty years. Given a list of university usernames, we searched for matches in both data breaches from hundreds of websites and a dozen large compilations of breaches. After cracking hashed passwords and tweaking guesses, we successfully guessed passwords for 32.0% of accounts matched to a university email address in a data breach, as well as 6.5% of accounts where the username (but not necessarily the domain) matched. Many of these accounts remained vulnerable for years after the breached data was leaked, and passwords found verbatim in breaches were nearly four times as likely to have been exploited (i.e., suspicious account activity was observed) than tweaked guesses. Over 70 different data breaches and various username-matching strategies bootstrapped correct guesses. In surveys of 40 users whose passwords we guessed, many users were unaware of the risks to their university account or that their credentials had been breached. This analysis of password reuse at our university provides pragmatic advice for organizations to protect accounts.

1 Introduction

Despite their disadvantages, passwords remain widely used for authentication [7]. Organizations must protect against large-scale attacks on users' passwords. An adversary may leverage **reused passwords**—when the same individual picks similar or identical passwords for different services [10, 80] to cope with having to remember numerous passwords [16]. If any one of these services suffers a data breach, attackers typically try to log into another service with the same email address alongside a password that is either the same as the leaked password, or tweaked in small ways. Such credential-stuffing attacks are this paper's focus. Additionally, attackers

curity23/presentation/nisenoff-retrospective st frequently chosen

The ability to conduct attacks that exploit reused password has increased as hundreds of websites have had their password databases stolen and leaked over the last decade [34]. We term the breach of a single service an **individual service breach**. In recent years, hackers have also packaged credentials from many different services into **breach compilations** containing hundreds of millions or even billions of credentials [24].

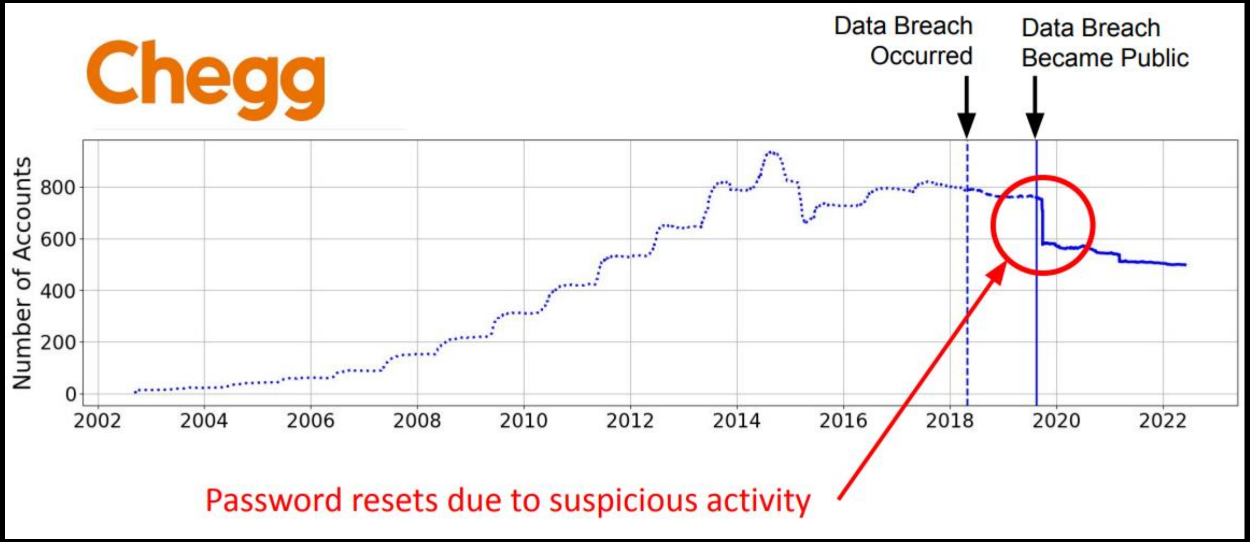
To protect an organization against attacks exploiting common passwords, system administrators can institute straightforward blocklists [25, 70]. Protecting an organization from reused passwords, however, is far more complex. A vulnerable password is specific to one user based on their credentials on other sites at any past or future time. Furthermore, prospective attackers often have far more information than system administrators. Attackers may know about a successful breach that system administrators may not hear about for years, or ever. Further, attackers may pool resources to crack hashes and reveal the plaintext needed for an attack, while the system administrator may be left only with uncracked hashes [9].

In recent years, researchers and practitioners have developed compromised-credential-checking tools to try to defend users. For instance, Chrome [73], Firefox [55], and Safari [12] notify users if their passwords appear in a data breach. The Have I Been Pwned (**HIBP**) service [32], itself integrated with 1Password [13], enables users to check for their appearance in a data breach. Supporting these efforts, academic work has proposed protocols that underpin compromised-credential-checking tools [40, 43, 44, 59, 82, 83] and sought to improve the usability of data breach notifications [22, 31, 53, 79, 90, 92].

Despite prior work, many questions remain for system administrators trying to protect their organizations from attacks exploiting reused passwords. For what amount of time are accounts vulnerable? Out of hundreds of data breaches, how important is it to account for them all? Should defenders devote resources to trying to crack hashes to protect users? Is it sufficient to look for matching email addresses, or should they also search for matching usernames? How often do attackers appear to have exploited reused passwords, and what factors make them more likely to have done so?

Results

- Guessed 32% of passwords in historical DB by leveraging reuse
    - As compared to 6.5% without considering reuse
    - 35.5% of valid guesses were for current passwords
- Of those guessed by reuse
    - 54.7% were verbatim reuse, vs. 45.3% based on tweaks
- Vulnerability is real
    - Some historical observed exploits coincided with data breaches
    - Passwords were vulnerable for long after a breach (median of 5 years)

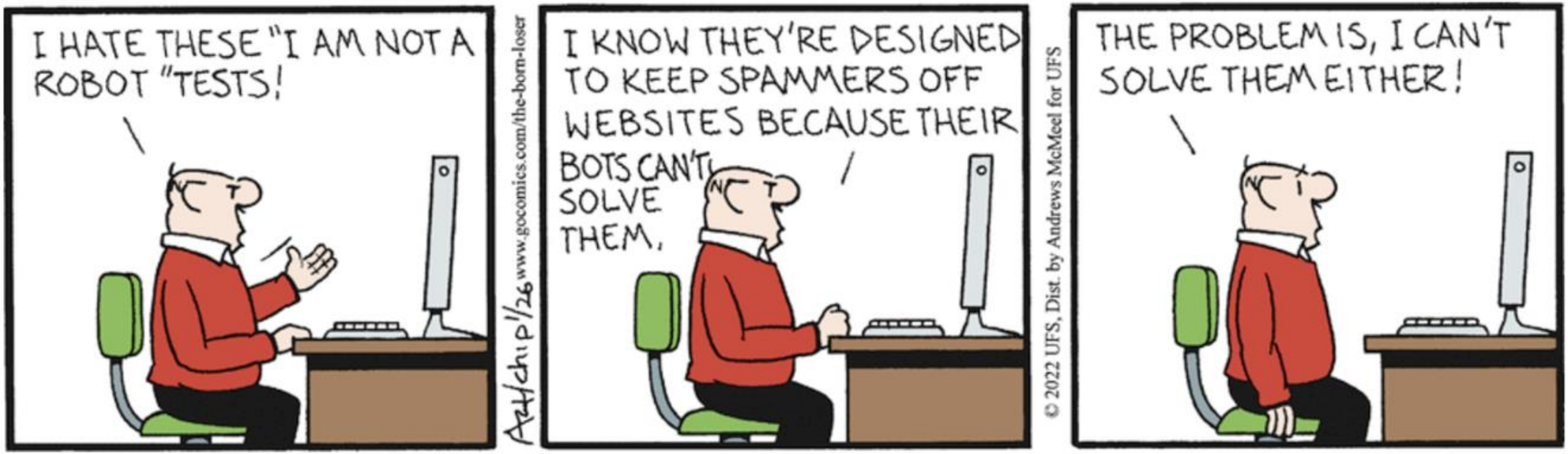


The human "threat"

- Malicious humans
- Clueless humans
- Unmotivated humans
- Humans constrained by human limitations

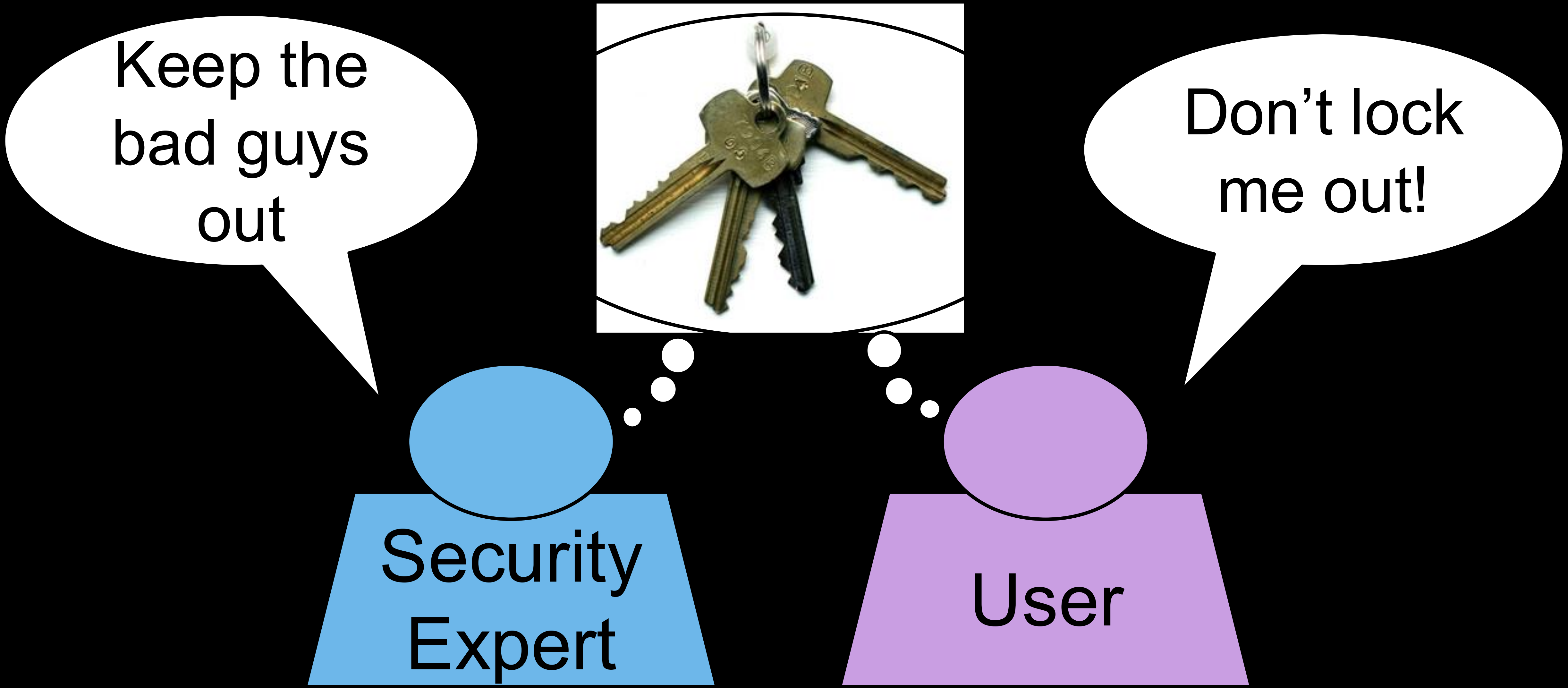
Thanks to Lorrie Cranor for this and some of the following slides

The Born Loser by Art and Chip Sansom for January 26, 2022


I HATE THESE "I AM NOT A ROBOT" TESTS!
I KNOW THEY'RE DESIGNED TO KEEP SPAMMERS OFF WEBSITES BECAUSE THEIR BOTS CAN'T SOLVE THEM.
THE PROBLEM IS, I CAN'T SOLVE THEM EITHER!
Art/chip 1/26
www.gocomics.com/the-born-loser
© 2022 UFS, Dist. by Andrews McMeel for UFS

Security is a secondary task

Concerns may not be aligned



Keep the bad guys out

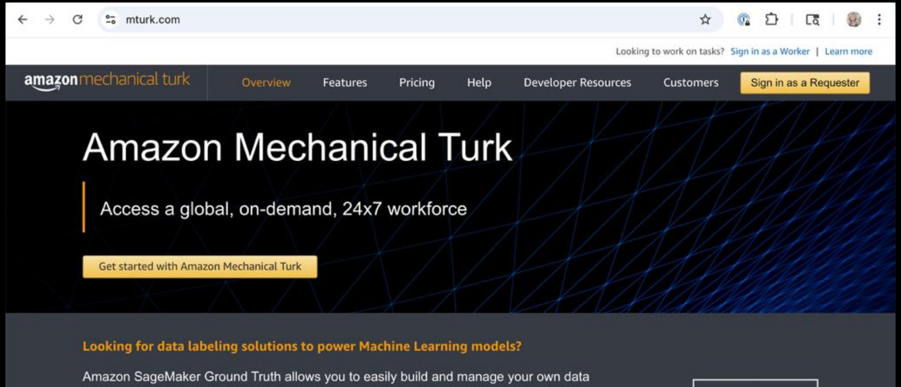
Don’t lock me out!

Security Expert

User

Password security: Perception vs. reality

- Password predictability in 2016 (and today?) is high. Maybe:
  - Users unwittingly select predictable passwords because they misunderstand what makes a password guessable
  - Users prioritize memorability, or some other feature, over predictability
- Let’s test these hypotheses experimentally!
  - Give participants a small set of technical exercises and free-response questions
  - Analyze results (e.g., with password cracking tools)

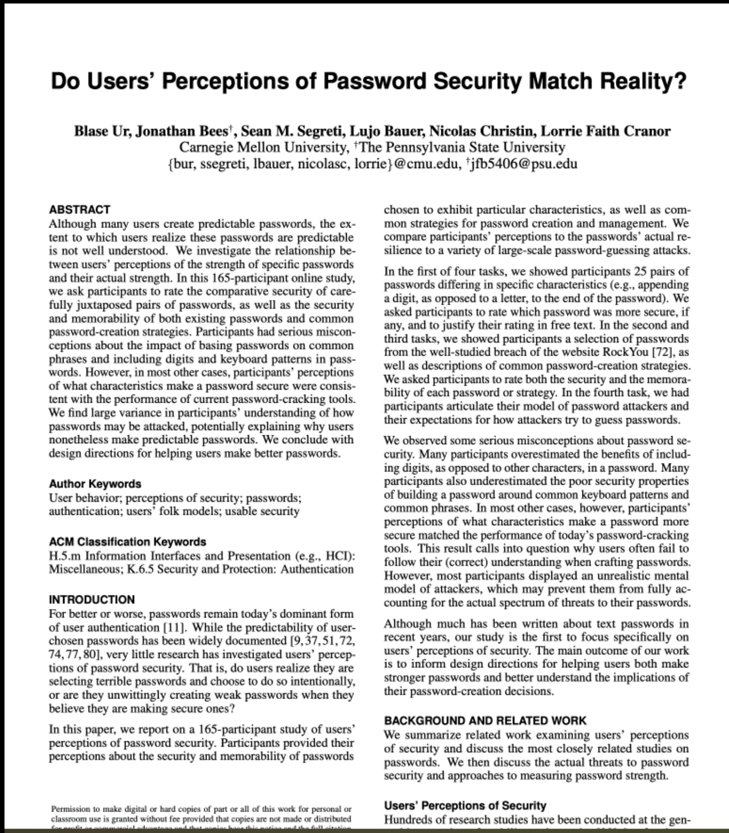


Amazon Mechanical Turk

Access a global, on-demand, 24x7 workforce

Looking for data labeling solutions to power Machine Learning models?

Amazon SageMaker Ground Truth allows you to easily build and manage your own data



Do Users’ Perceptions of Password Security Match Reality?

Blase Ur, Jonathan Bees†, Sean M. Segreti, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor
Carnegie Mellon University, †The Pennsylvania State University
{bur, ssegreti, lbauer, nicolasc, lorrie}@cmu.edu, †jfb5406@psu.edu

ABSTRACT

Although many users create predictable passwords, the extent to which users realize these passwords are predictable is not well understood. We investigate the relationship between users’ perceptions of the strength of specific passwords and their actual strength. In this 165-participant online study, we ask participants to rate the comparative security of carefully juxtaposed pairs of passwords, as well as the security and memorability of both existing passwords and common password-creation strategies. Participants had serious misconceptions about the impact of basing passwords on common phrases and including digits and keyboard patterns in passwords. However, in most other cases, participants’ perceptions of what characteristics make a password secure were consistent with the performance of current password-cracking tools. We find large variance in participants’ understanding of how passwords may be attacked, potentially explaining why users nonetheless make predictable passwords. We conclude with design directions for helping users make better passwords.

Author Keywords

User behavior; perceptions of security; passwords; authentication; users’ folk models; usable security

ACM Classification Keywords

H.5.m Information Interfaces and Presentation (e.g., HCI): Miscellaneous; K.6.5 Security and Protection: Authentication

INTRODUCTION

For better or worse, passwords remain today’s dominant form of user authentication [11]. While the predictability of user-chosen passwords has been widely documented [9,37,51,72, 74,77,80], very little research has investigated users’ perceptions of password security. That is, do users realize they are selecting terrible passwords and choose to do so intentionally, or are they unwittingly creating weak passwords when they believe they are making secure ones?

In this paper, we report on a 165-participant study of users’ perceptions of password security. Participants provided their perceptions about the security and memorability of passwords chosen to exhibit particular characteristics, as well as common strategies for password creation and management. We compare participants’ perceptions to the passwords’ actual resilience to a variety of large-scale password-guessing attacks.

In the first of four tasks, we showed participants 25 pairs of passwords differing in specific characteristics (e.g., appending a digit, as opposed to a letter, to the end of the password). We asked participants to rate which password was more secure, if any, and to justify their rating in free text. In the second and third tasks, we showed participants a selection of passwords from the well-studied breach of the website RockYou [72], as well as descriptions of common password-creation strategies. We asked participants to rate both the security and the memorability of each password or strategy. In the fourth task, we had participants articulate their model of password attackers and their expectations for how attackers try to guess passwords.

We observed some serious misconceptions about password security. Many participants overestimated the benefits of including digits, as opposed to other characters, in a password. Many participants also underestimated the poor security properties of building a password around common keyboard patterns and common phrases. In most other cases, however, participants’ perceptions of what characteristics make a password more secure matched the performance of today’s password-cracking tools. This result calls into question why users often fail to follow their (correct) understanding when crafting passwords. However, most participants displayed an unrealistic mental model of attackers, which may prevent them from fully accounting for the actual spectrum of threats to their passwords.

Although much has been written about text passwords in recent years, our study is the first to focus specifically on users’ perceptions of security. The main outcome of our work is to inform design directions for helping users both make stronger passwords and better understand the implications of their password-creation decisions.

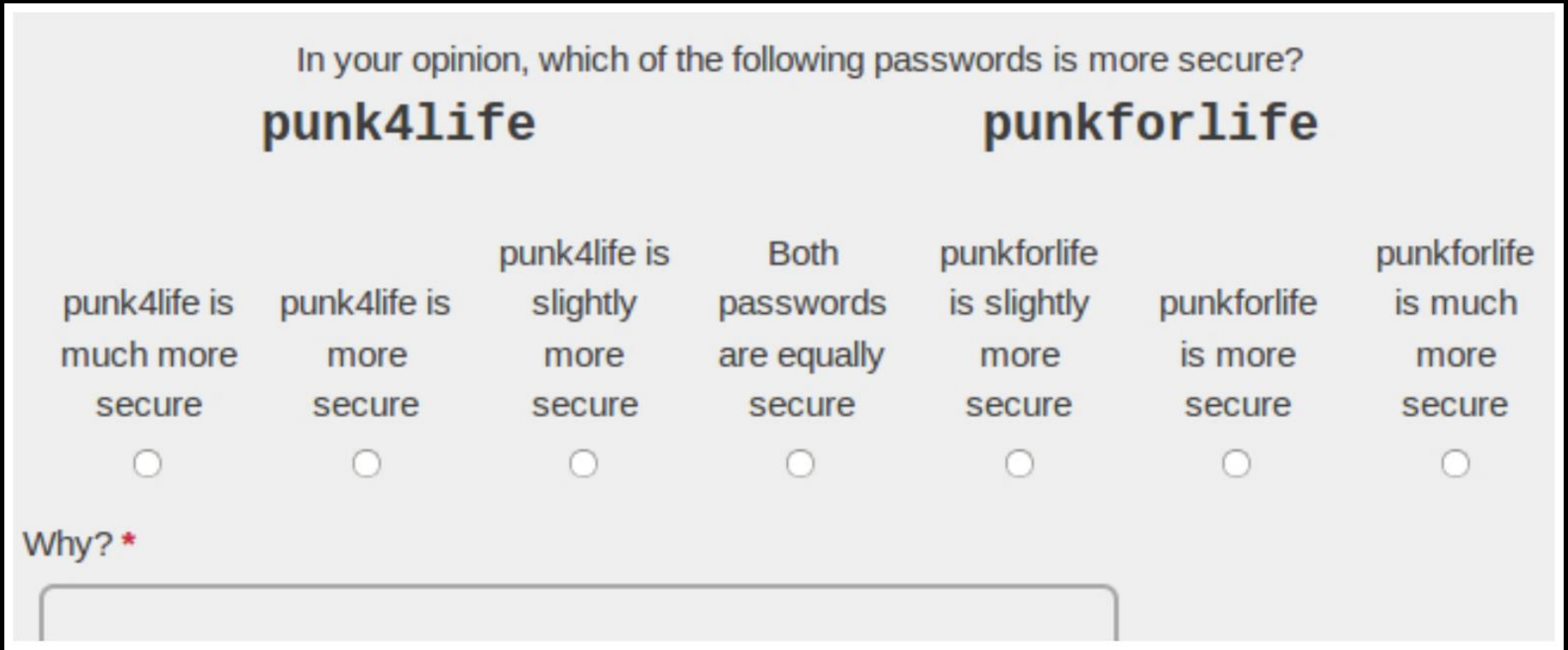
BACKGROUND AND RELATED WORK

We summarize related work examining users’ perceptions of security and discuss the most closely related studies on passwords. We then discuss the actual threats to password security and approaches to measuring password strength.

Users’ Perceptions of Security

Hundreds of research studies have been conducted at the gen-

Example question type: Comparison



In your opinion, which of the following passwords is more secure?

punk4life **punkforlife**

punk4life is much more secure	punk4life is more secure	punk4life is slightly more secure	Both passwords are equally secure	punkforlife is slightly more secure	punkforlife is more secure	punkforlife is much more secure
○	○	○	○	○	○	○

Why? *

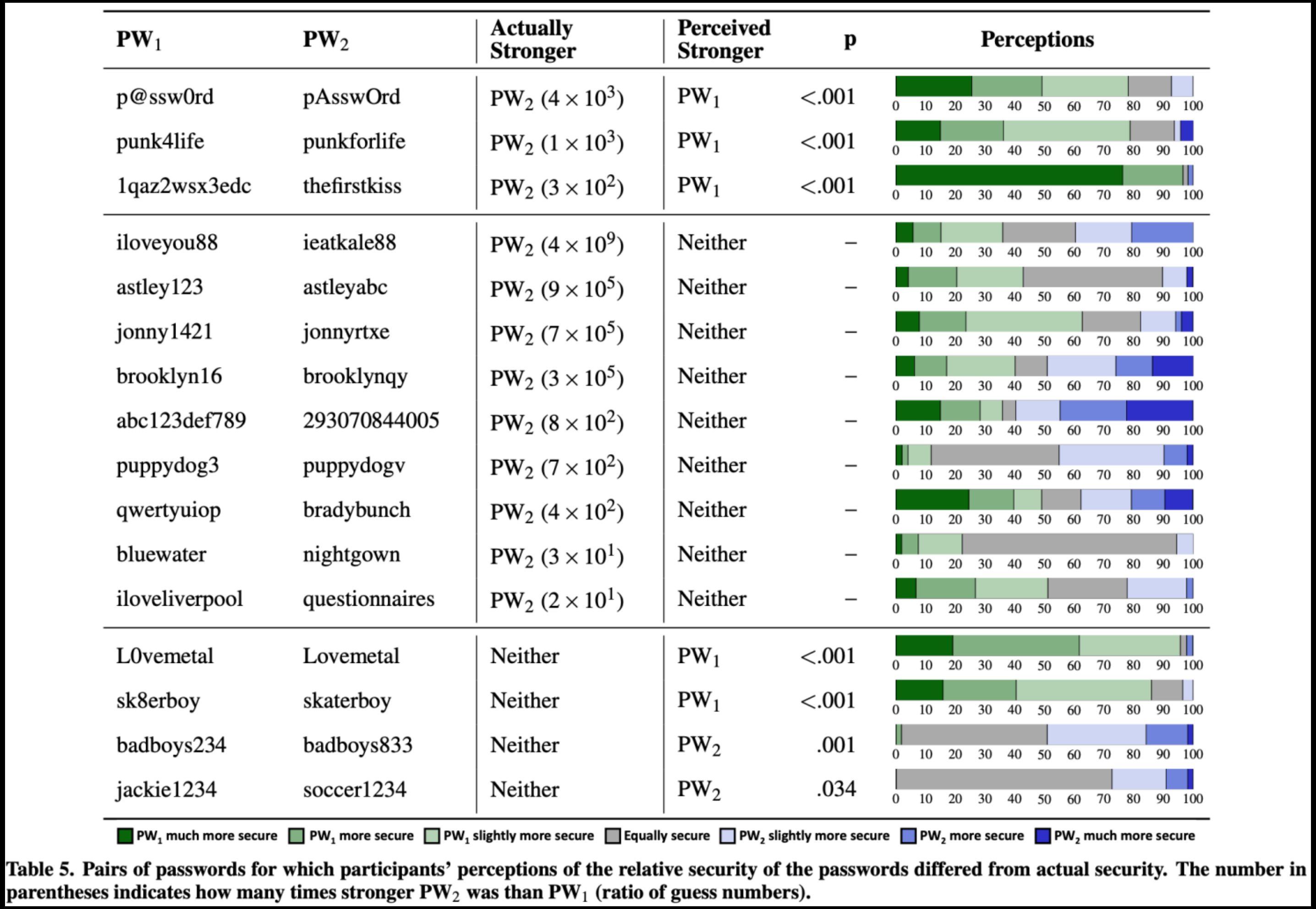
25 kinds of difference, users saw one of three password pairs per category (25 x 3 = 75 total pairs)
Password pairs selected randomly/systematically from RockYou dataset (real-world PW DB that was leaked)
26th category is an **attention check**: Show the same password
Presented in randomized order, and randomized left/right position

Results: Comparison

16 pairs (21%) were not consistent

Misconceptions:

- adding digits inherently makes a password more secure than using only letters
- substituting digits or symbols for letters makes a password more secure
- overestimated the security of keyboard patterns
- misjudged the popularity of particular words and phrases



PW_1	PW_2	Actually Stronger	Perceived Stronger	p	Perceptions
p@ssw0rd	pAsswOrd	PW_2 (4×10^3)	PW_1	<.001	
punk4life	punkforlife	PW_2 (1×10^3)	PW_1	<.001	
1qaz2wsx3edc	thefirstkiss	PW_2 (3×10^2)	PW_1	<.001	
iloveyou88	ieatkale88	PW_2 (4×10^9)	Neither	–	
astley123	astleyabc	PW_2 (9×10^5)	Neither	–	
jonny1421	jonnyrtxe	PW_2 (7×10^5)	Neither	–	
brooklyn16	brooklynqy	PW_2 (3×10^5)	Neither	–	
abc123def789	293070844005	PW_2 (8×10^2)	Neither	–	
puppydog3	puppydogv	PW_2 (7×10^2)	Neither	–	
qwertyuiop	bradybunch	PW_2 (4×10^2)	Neither	–	
bluewater	nightgown	PW_2 (3×10^1)	Neither	–	
iloveliverpool	questionnaires	PW_2 (2×10^1)	Neither	–	
L0vemetal	Lovemetal	Neither	PW_1	<.001	
sk8erboy	skaterboy	Neither	PW_1	<.001	
badboys234	badboys833	Neither	PW_2	.001	
jackie1234	soccer1234	Neither	PW_2	.034	

🟩 PW_1 much more secure 🟩 PW_1 more secure 🟩 PW_1 slightly more secure ⬜ Equally secure 🟦 PW_2 slightly more secure 🟦 PW_2 more secure 🟦 PW_2 much more secure

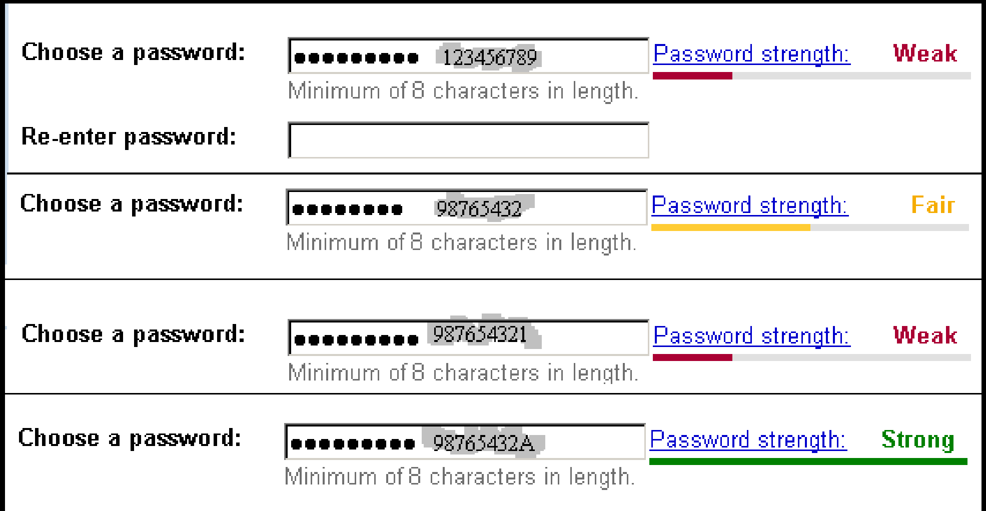
Table 5. Pairs of passwords for which participants' perceptions of the relative security of the passwords differed from actual security. The number in parentheses indicates how many times stronger PW_2 was than PW_1 (ratio of guess numbers).

Other results

- 79% of comparison pairs were correct!
    - Capitalizing the middle of words is better than capitalizing just the beginning
    - Putting digits and symbols in the middle is better than at the end
    - Dictionary words are better than common first names
    - …
- Users' incomplete understanding of the scale of potential attacks seems to be a root cause of bad passwords
    - 1/3 of participants: secure if can withstand several dozen guesses
    - Others: password must withstand quadrillions of guesses or more

Password strength meter

- Gives user feedback on the **strength of the password**
  - Intended to **measure guessability**
  - Research shows that these **can work**, but the design must be **stringent**
    - Ur et al, “How does your password measure up? The effect of strength meters on password creation”, Proc. USENIX Security Symposium, 2012.
  - Some password requirements now debunked – people use odd characters in a predictable way!



Choose a password: ●●●●●●●●● 123456789 Password strength: Weak
Minimum of 8 characters in length.
Re-enter password:

Choose a password: ●●●●●●●● 98765432 Password strength: Fair
Minimum of 8 characters in length.

Choose a password: ●●●●●●●●● 987654321 Password strength: Weak
Minimum of 8 characters in length.

Choose a password: ●●●●●●●●● 98765432A Password strength: Strong
Minimum of 8 characters in length.

password → P@ssw0rd!

Password manager

1Password

LastPass

RoboForm

KEEPER

- A password manager (PM) stores a **database of passwords, indexed by site**
  - Encrypted by **a single, master password** chosen (and remembered) by the user, used as a key
  - **PM can generate complicated per-site passwords**
    - Hard to guess, hard to remember, but the latter doesn’t matter!
- **Benefits**
  - Only a single password for user to remember
  - User’s password at any given site is hard to guess
  - Compromise of password at one site does not permit immediate compromise at other sites
- But:
  - Must still **protect** and **remember strong master password**

Impact of password managers on security

- RQ: Impact of password managers on password strength and reuse?
- Data collected: Survey on Mechanical Turk, then data collection from some of these about password management using a Chrome plugin
    - Plugin used for four days
- Data analysis: Correlate factors against measures of password strength and occurrences of reuse
    - Factors include how exactly password managers were used or not used
    - Allows relating to baselines, identifying key aspects of success

Better managed than memorized?
Studying the Impact of Managers on Password Strength and Reuse

Sanam Ghorbani Lyastani
CISPA, Saarland University

Michael Schilling
Saarland University

Sascha Fahl
Ruhr-University Bochum

Michael Backes
CISPA Helmholtz Center i.G.

Sven Bugiel
CISPA Helmholtz Center i.G.

Abstract

Despite their well-known security problems, passwords are still the incumbent authentication method for virtually all online services. To remedy the situation, users are very often referred to password managers as a solution to the password reuse and weakness problems. However, to date the actual impact of password managers on password strength and reuse has not been studied systematically.

We provide the first large-scale study of the password managers' influence on users' real-life passwords. By combining qualitative data on users' password creation and management strategies, collected from 476 participants of an online survey, with quantitative data (incl. password metrics and entry methods) collected in situ with a browser plugin from 170 users, we were able to gain a more complete picture of the factors that influence our participants' password strength and reuse. Our approach allows us to quantify for the first time that password managers indeed influence the password security, however, whether this influence is beneficial or aggravating existing problems depends on the users' strategies and how well the manager supports the users' password management right from the time of password creation. Given our results, we think research should further investigate how managers can better support users' password strategies in order to improve password security as well as stop aggravating the existing problems.

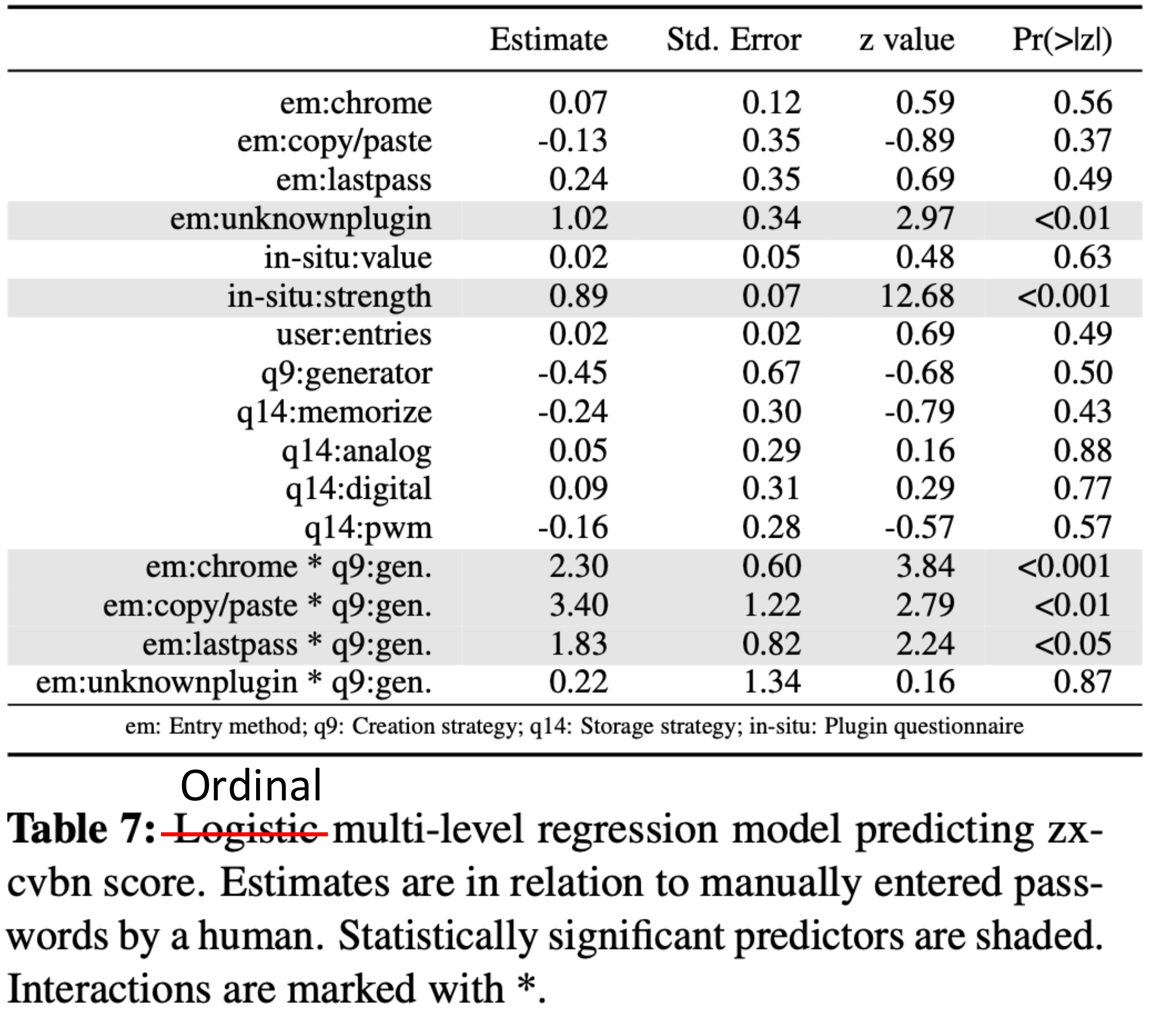
1 Introduction

For several decades passwords prevail as the default authentication scheme for virtually all online services [44, 11, 30]. At the same time, research has again and again demonstrated that passwords perform extremely poor in terms of security [48]. For instance, various attacks exploit that humans fail to create strong passwords themselves [10, 19, 45, 31, 34]. Even worse, there is an observable trend towards an increasing number of online services that users register to. This increasing number of required passwords in combination with the limited human capacity to remember passwords leads to the bad practice of re-using passwords across accounts [26, 51, 16, 66].

In the past, different solutions have been implemented to help users creating stronger passwords, such as password meters and policies, which are also still subject of active research [41, 54, 17, 45, 68]. Among the most often recommended solutions [28, 59, 53, 62, 56] to these problems for end-users is technical support in the form of password management software. Those password managers come built-in to our browsers, as a browser plugin, or as separate applications. Password managers are being recommended as a solution because they fulfill important usability and security aspects at the same time: They store all the users' passwords so the users do not have to memorize them; they can also help users entering their passwords by automatically filling them into log-in forms; and they can also offer help in creating unique, random passwords. By today, there are several examples of third party password managers that fit this description, such as Lastpass [5], 1Password [1], and even seemingly unrelated security software, such as anti-virus [4] solutions.

Unfortunately, it has not been sufficiently studied in the past whether password managers fulfill their promise and indeed have a positive influence on password security or not? To break this question down, we are interested in *1) whether password managers actually store strong passwords that are likely auto-generated by, for instance, password generators, or if they really are just storage where users save their self-made, likely weak passwords?* Further, we are interested whether *2) users, despite using password managers, still reuse passwords across different websites or if do they use the managers' support to maintain a large set of unique passwords for every distinct service?* Prior works [66, 51] that studied password reuse and strength in situ have also considered password managers as factors, but did not find an influence by managers and could not conclusively answer those questions.

Modeling strength



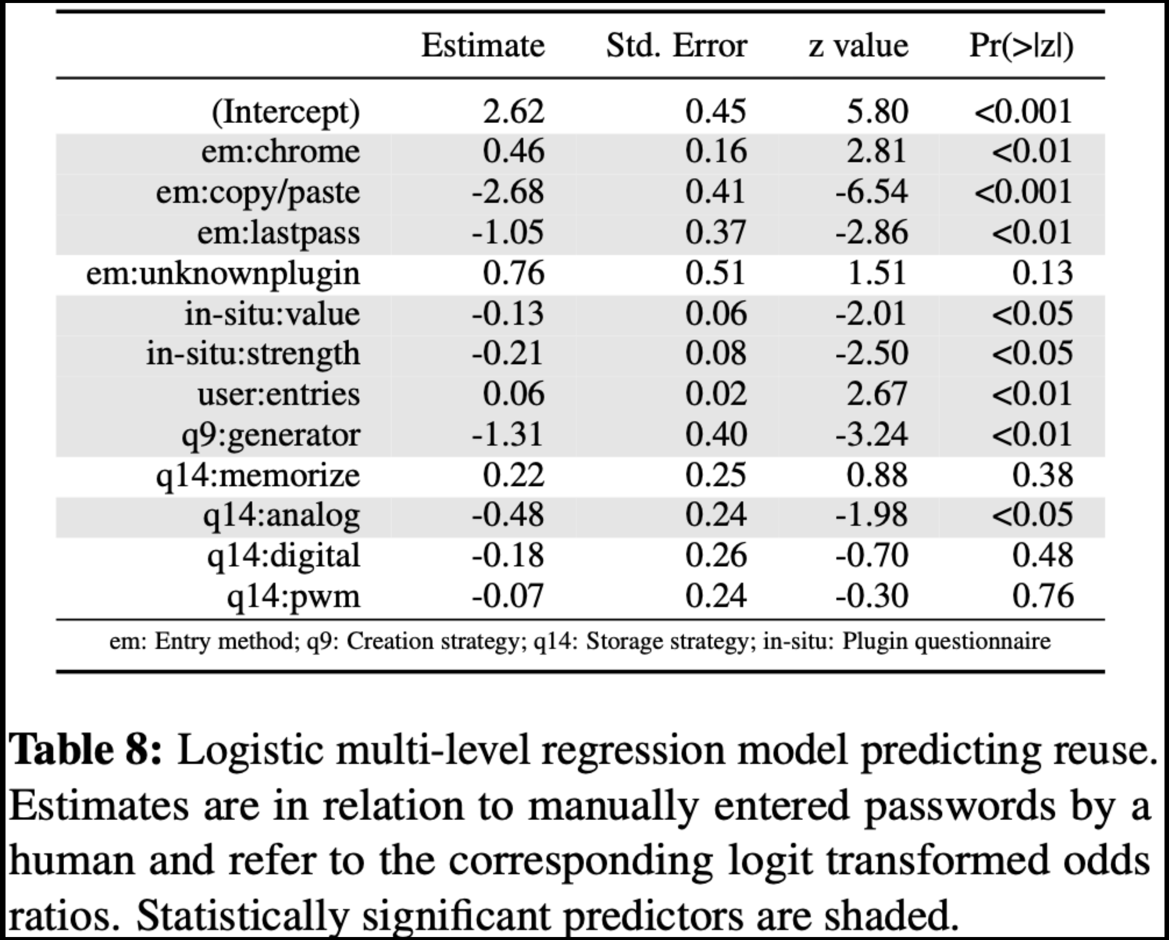
	Estimate	Std. Error	z value	Pr(>|z|)
em:chrome	0.07	0.12	0.59	0.56
em:copy/paste	-0.13	0.35	-0.89	0.37
em:lastpass	0.24	0.35	0.69	0.49
em:unknownplugin	1.02	0.34	2.97	<0.01
in-situ:value	0.02	0.05	0.48	0.63
in-situ:strength	0.89	0.07	12.68	<0.001
user:entries	0.02	0.02	0.69	0.49
q9:generator	-0.45	0.67	-0.68	0.50
q14:memorize	-0.24	0.30	-0.79	0.43
q14:analog	0.05	0.29	0.16	0.88
q14:digital	0.09	0.31	0.29	0.77
q14:pwm	-0.16	0.28	-0.57	0.57
em:chrome * q9:gen.	2.30	0.60	3.84	<0.001
em:copy/paste * q9:gen.	3.40	1.22	2.79	<0.01
em:lastpass * q9:gen.	1.83	0.82	2.24	<0.05
em:unknownplugin * q9:gen.	0.22	1.34	0.16	0.87

em: Entry method; q9: Creation strategy; q14: Storage strategy; in-situ: Plugin questionnaire

Table 7: Ordinal ~~Logistic~~ multi-level regression model predicting zxcvbn score. Estimates are in relation to manually entered passwords by a human. Statistically significant predictors are shaded. Interactions are marked with *.

- Creation strategy + EM significant together, but not individually
  - So: using a PM only leads to significant improvement in password strength when users also employ supporting techniques for password creation
- Self-reported password strength was a significant predictor of the measured password strength

Modeling reuse



	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.62	0.45	5.80	<0.001
em:chrome	0.46	0.16	2.81	<0.01
em:copy/paste	-2.68	0.41	-6.54	<0.001
em:lastpass	-1.05	0.37	-2.86	<0.01
em:unknownplugin	0.76	0.51	1.51	0.13
in-situ:value	-0.13	0.06	-2.01	<0.05
in-situ:strength	-0.21	0.08	-2.50	<0.05
user:entries	0.06	0.02	2.67	<0.01
q9:generator	-1.31	0.40	-3.24	<0.01
q14:memorize	0.22	0.25	0.88	0.38
q14:analog	-0.48	0.24	-1.98	<0.05
q14:digital	-0.18	0.26	-0.70	0.48
q14:pwm	-0.07	0.24	-0.30	0.76

em: Entry method; q9: Creation strategy; q14: Storage strategy; in-situ: Plugin questionnaire

Table 8: Logistic multi-level regression model predicting reuse. Estimates are in relation to manually entered passwords by a human and refer to the corresponding logit transformed odds ratios. Statistically significant predictors are shaded.

- Reuse was significantly influenced by the entry method of the password
  - Odds for reuse were 2.85 times *lower* by LastPass, 14.29 times lower if C&P
  - odds for reuse were 1.65 times *higher* by Chrome auto-fill
- Creation by alg: odds of non-reuse 3.70 times higher

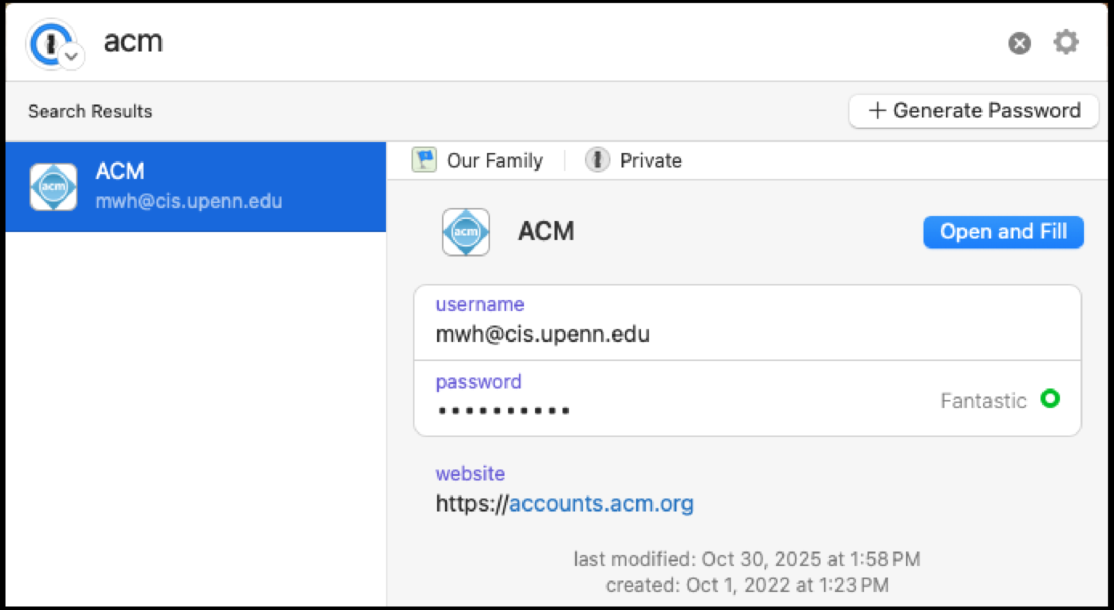
Summary of PM study results

- Users that rely on **technical support for password creation** had both **stronger and more unique passwords**
    - Even if entered through other channels than a manager

- **Chrome's auto-fill option increased password reuse** – more than 80% of Chrome auto-filled passwords were reused
    - Chrome at the time did not have password generation enabled by default

Better together

- Password manager
  - One security decision, not many
- Password meter
  - Users can explore ramifications of various choices by visualizing quality and reasoning of password
  - Do not permit poor choices (or reduce the chances of them) by enforcing a minimum score
- Best: **Let PM generate password**
  - We’ll see some studies that show this later



ACM
mwh@cis.upenn.edu

ACM

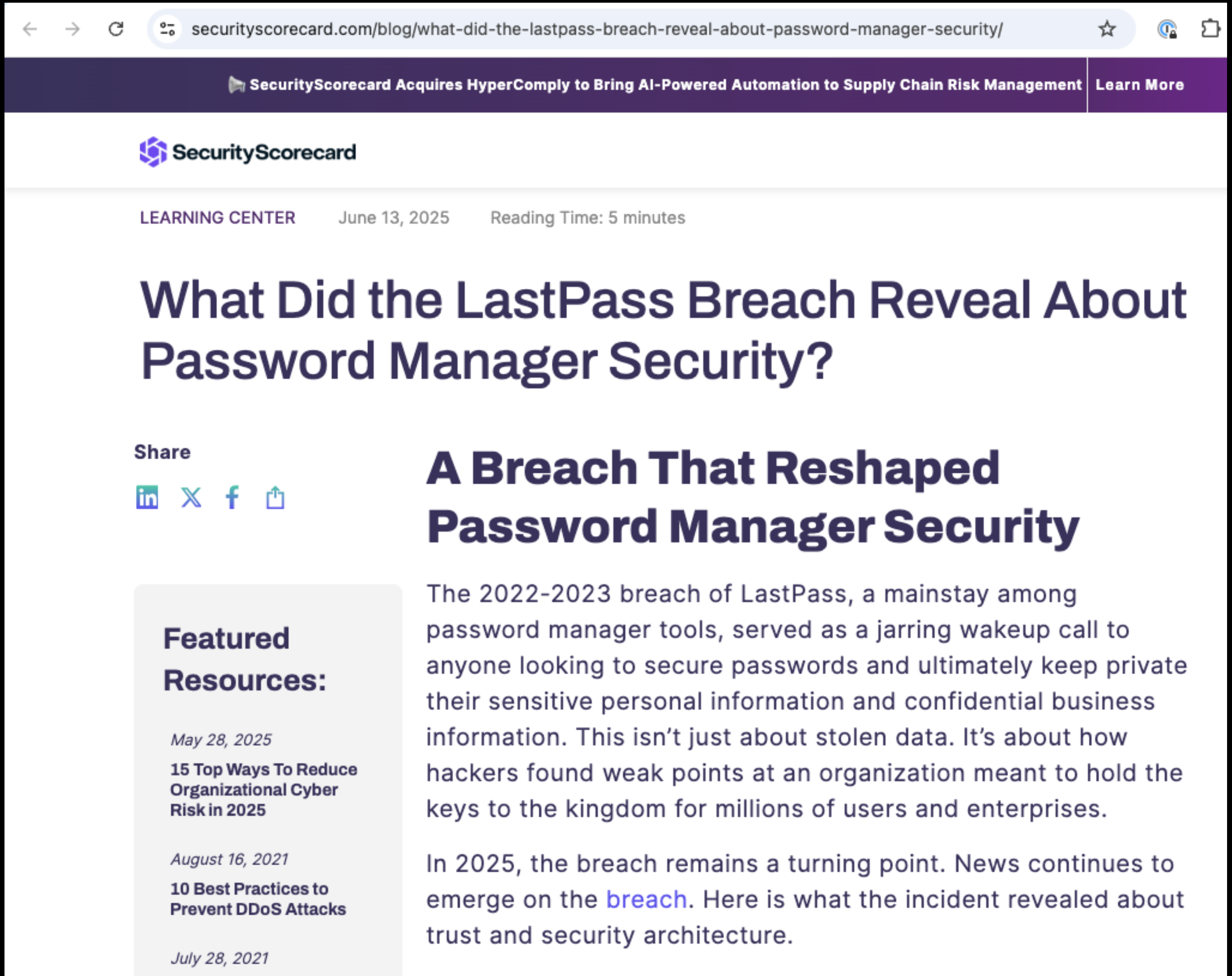
username
mwh@cis.upenn.edu

password
•••••••••• Fantastic ○

website
https://accounts.acm.org

last modified: Oct 30, 2025 at 1:58 PM
created: Oct 1, 2022 at 1:23 PM

PM Risk: Central point of failure?



securityscorecard.com/blog/what-did-the-lastpass-breach-reveal-about-password-manager-security/

SecurityScorecard Acquires HyperComply to Bring AI-Powered Automation to Supply Chain Risk Management | Learn More

SecurityScorecard

LEARNING CENTER June 13, 2025 Reading Time: 5 minutes

What Did the LastPass Breach Reveal About Password Manager Security?

Share

A Breach That Reshaped Password Manager Security

The 2022-2023 breach of LastPass, a mainstay among password manager tools, served as a jarring wakeup call to anyone looking to secure passwords and ultimately keep private their sensitive personal information and confidential business information. This isn't just about stolen data. It's about how hackers found weak points at an organization meant to hold the keys to the kingdom for millions of users and enterprises.

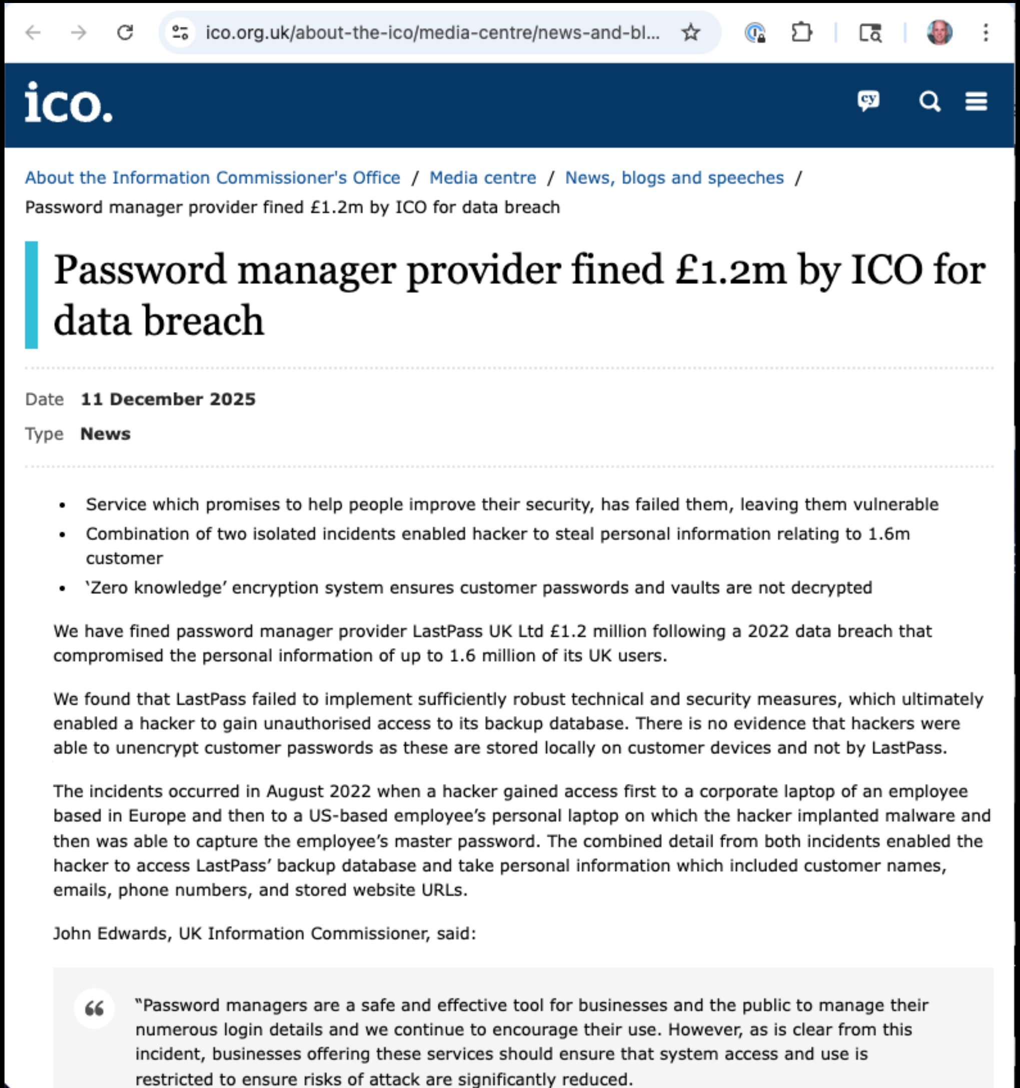
In 2025, the breach remains a turning point. News continues to emerge on the breach. Here is what the incident revealed about trust and security architecture.

Featured Resources:

May 28, 2025
15 Top Ways To Reduce Organizational Cyber Risk in 2025

August 16, 2021
10 Best Practices to Prevent DDoS Attacks

July 28, 2021



ico.org.uk/about-the-ico/media-centre/news-and-bl...

ico.

About the Information Commissioner's Office / Media centre / News, blogs and speeches / Password manager provider fined £1.2m by ICO for data breach

Password manager provider fined £1.2m by ICO for data breach

Date **11 December 2025**

Type **News**

- Service which promises to help people improve their security, has failed them, leaving them vulnerable
- Combination of two isolated incidents enabled hacker to steal personal information relating to 1.6m customer
- 'Zero knowledge' encryption system ensures customer passwords and vaults are not decrypted

We have fined password manager provider LastPass UK Ltd £1.2 million following a 2022 data breach that compromised the personal information of up to 1.6 million of its UK users.

We found that LastPass failed to implement sufficiently robust technical and security measures, which ultimately enabled a hacker to gain unauthorised access to its backup database. There is no evidence that hackers were able to unencrypt customer passwords as these are stored locally on customer devices and not by LastPass.

The incidents occurred in August 2022 when a hacker gained access first to a corporate laptop of an employee based in Europe and then to a US-based employee's personal laptop on which the hacker implanted malware and then was able to capture the employee's master password. The combined detail from both incidents enabled the hacker to access LastPass' backup database and take personal information which included customer names, emails, phone numbers, and stored website URLs.

John Edwards, UK Information Commissioner, said:

> "Password managers are a safe and effective tool for businesses and the public to manage their numerous login details and we continue to encourage their use. However, as is clear from this incident, businesses offering these services should ensure that system access and use is restricted to ensure risks of attack are significantly reduced.

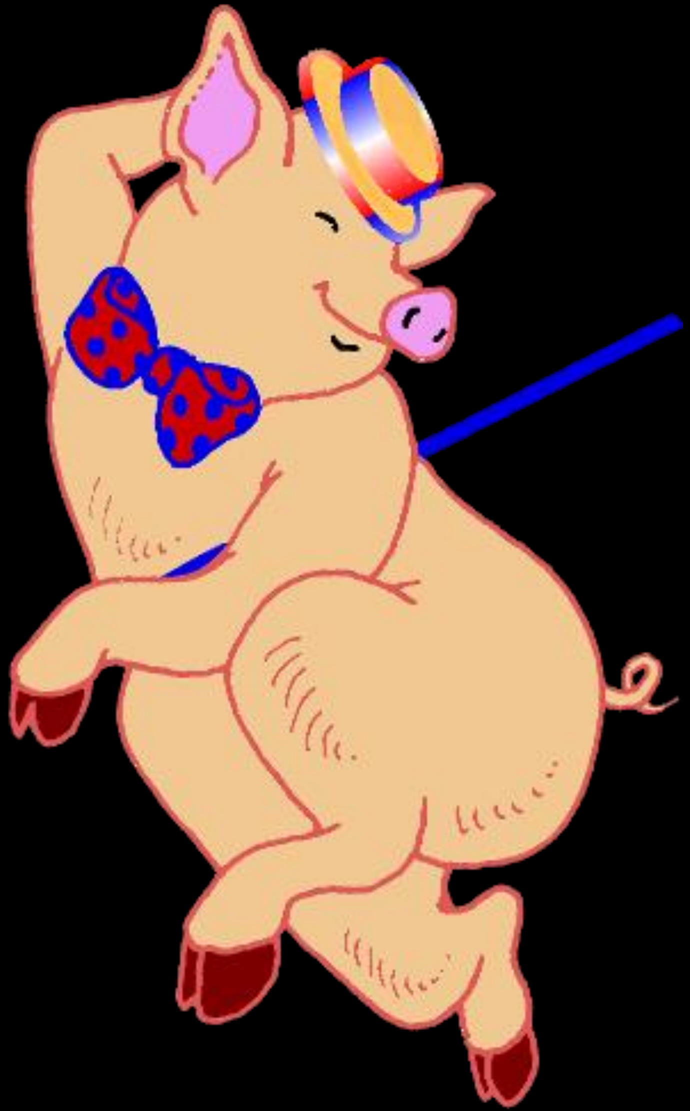
Understand humans in the loop

- Do they know they are supposed to be doing something?
- Do they understand what they are supposed to do?
- Do they know how to do it?
- Are they motivated to do it?
- Are they capable of doing it?
- Will they actually do it?

Remember: Convenience always wins

Given a choice between dancing pigs and security, users will pick dancing pigs every time.

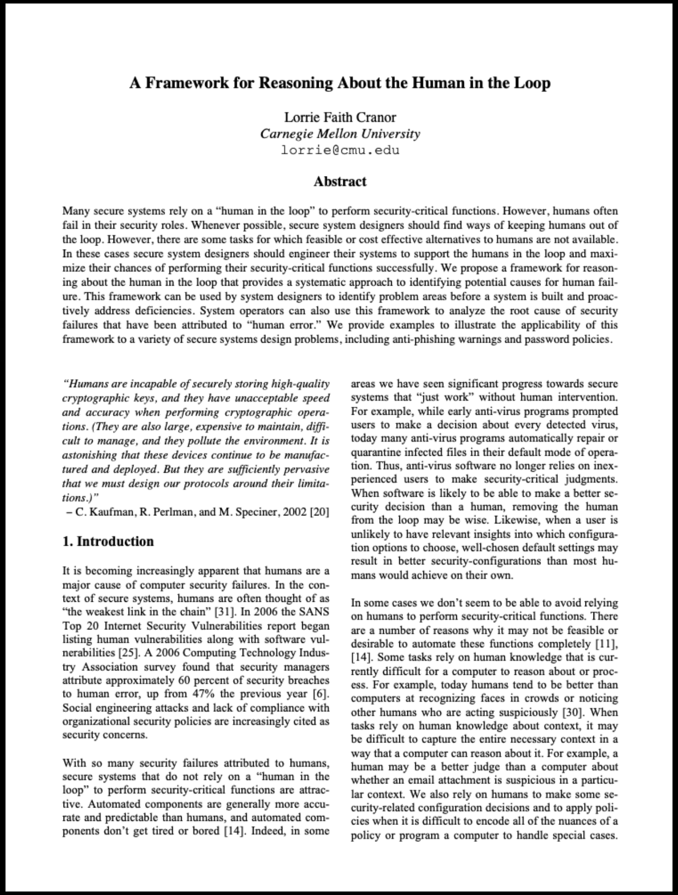
– Ed Felton



Human-in-the-loop framework

- Based on Communication-Human Information Processing Model (C-HIP) from Warnings Science
- Models human interaction with secure systems
- Can help identify human threats

We will look at this a bit later



A Framework for Reasoning About the Human in the Loop

Lorrie Faith Cranor
Carnegie Mellon University
lorrie@cmu.edu

Abstract

Many secure systems rely on a "human in the loop" to perform security-critical functions. However, humans often fail in their security roles. Whenever possible, secure system designers should find ways of keeping humans out of the loop. However, there are some tasks for which feasible or cost effective alternatives to humans are not available. In these cases secure system designers should engineer their systems to support the humans in the loop and maximize their chances of performing their security-critical functions successfully. We propose a framework for reasoning about the human in the loop that provides a systematic approach to identifying potential causes for human failure. This framework can be used by system designers to identify problem areas before a system is built and proactively address deficiencies. System operators can also use this framework to analyze the root cause of security failures that have been attributed to "human error." We provide examples to illustrate the applicability of this framework to a variety of secure systems design problems, including anti-phishing warnings and password policies.

"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, difficult 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.)"
– C. Kaufman, R. Perlman, and M. Speciner, 2002 [20]

1. Introduction

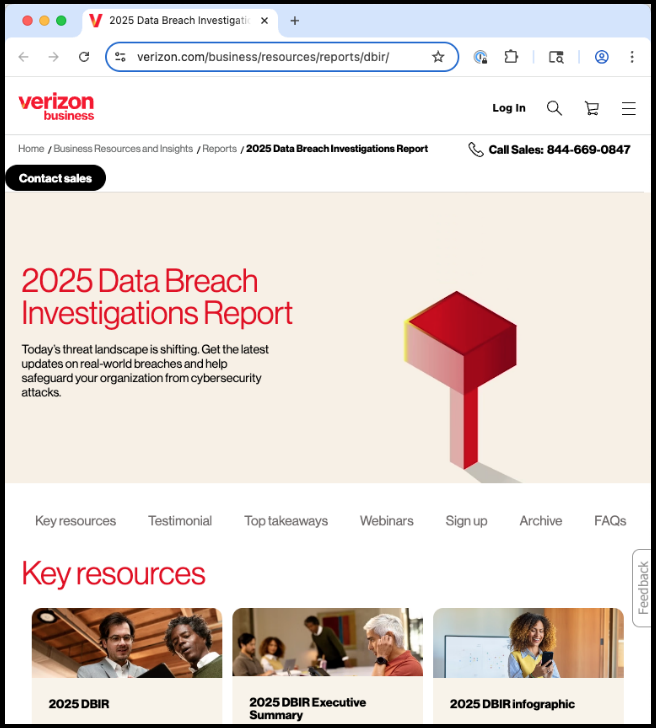
It is becoming increasingly apparent that humans are a major cause of computer security failures. In the context of secure systems, humans are often thought of as "the weakest link in the chain" [31]. In 2006 the SANS Top 20 Internet Security Vulnerabilities report began listing human vulnerabilities along with software vulnerabilities [25]. A 2006 Computing Technology Industry Association survey found that security managers attribute approximately 60 percent of security breaches to human error, up from 47% the previous year [6]. Social engineering attacks and lack of compliance with organizational security policies are increasingly cited as security concerns.

With so many security failures attributed to humans, secure systems that do not rely on a "human in the loop" to perform security-critical functions are attractive. Automated components are generally more accurate and predictable than humans, and automated components don't get tired or bored [14]. Indeed, in some areas we have seen significant progress towards secure systems that "just work" without human intervention. For example, while early anti-virus programs prompted users to make a decision about every detected virus, today many anti-virus programs automatically repair or quarantine infected files in their default mode of operation. Thus, anti-virus software no longer relies on inexperienced users to make security-critical judgments. When software is likely to be able to make a better security decision than a human, removing the human from the loop may be wise. Likewise, when a user is unlikely to have relevant insights into which configuration options to choose, well-chosen default settings may result in better security-configurations than most humans would achieve on their own.

In some cases we don't seem to be able to avoid relying on humans to perform security-critical functions. There are a number of reasons why it may not be feasible or desirable to automate these functions completely [11], [14]. Some tasks rely on human knowledge that is currently difficult for a computer to reason about or process. For example, today humans tend to be better than computers at recognizing faces in crowds or noticing other humans who are acting suspiciously [30]. When tasks rely on human knowledge about context, it may be difficult to capture the entire necessary context in a way that a computer can reason about it. For example, a human may be a better judge than a computer about whether an email attachment is suspicious in a particular context. We also rely on humans to make some security-related configuration decisions and to apply policies when it is difficult to encode all of the nuances of a policy or program a computer to handle special cases.

L. Cranor. A Framework for Reasoning About the Human In the Loop. Usability, Psychology and Security 2008. http://www.usenix.org/events/upsec08/tech/full_papers/cranor/cranor.pdf

Top attack vectors



2025 Data Breach Investigati

verizon.com/business/resources/reports/dbir/

verizon business

Log In

Home / Business Resources and Insights / Reports / 2025 Data Breach Investigations Report

Call Sales: 844-669-0847

Contact sales

2025 Data Breach Investigations Report

Today's threat landscape is shifting. Get the latest updates on real-world breaches and help safeguard your organization from cybersecurity attacks.

Key resources Testimonial Top takeaways Webinars Sign up Archive FAQs

Key resources

2025 DBIR

2025 DBIR Executive Summary

2025 DBIR infographic

Feedback

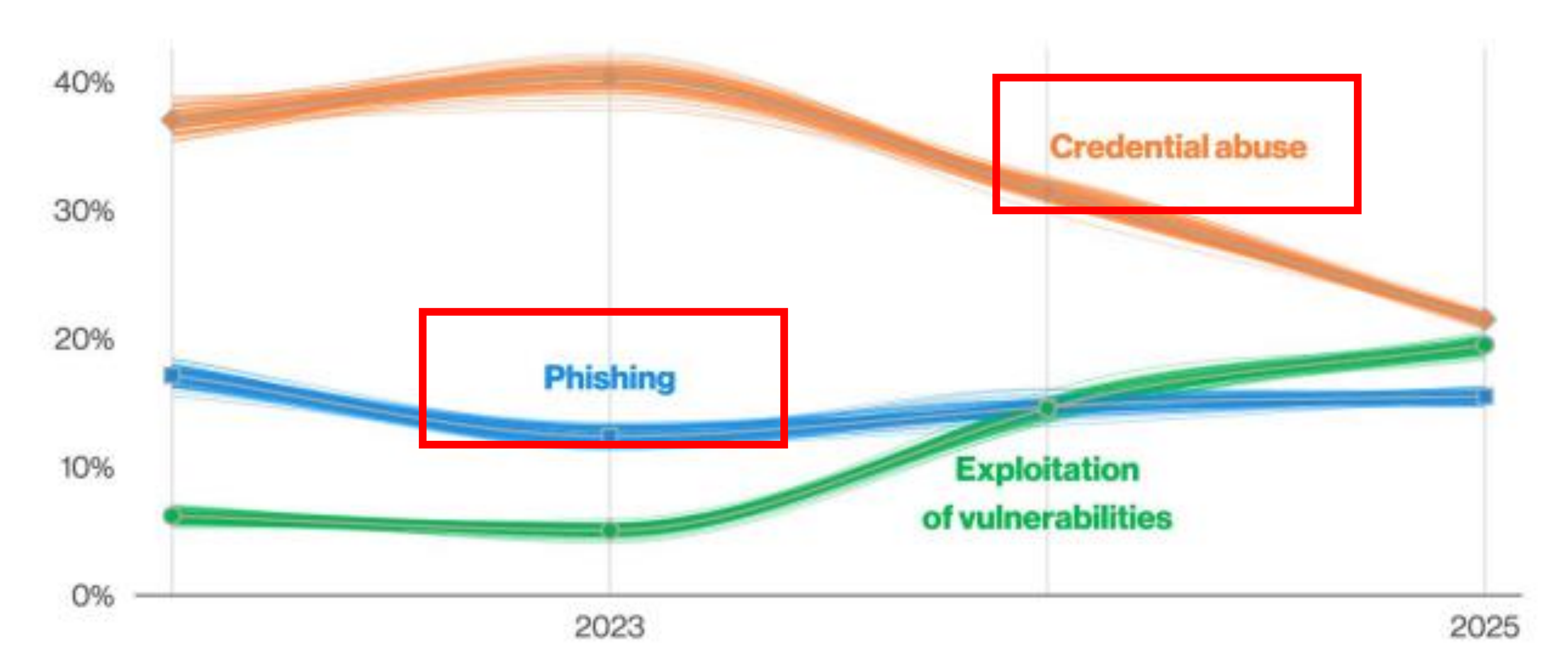
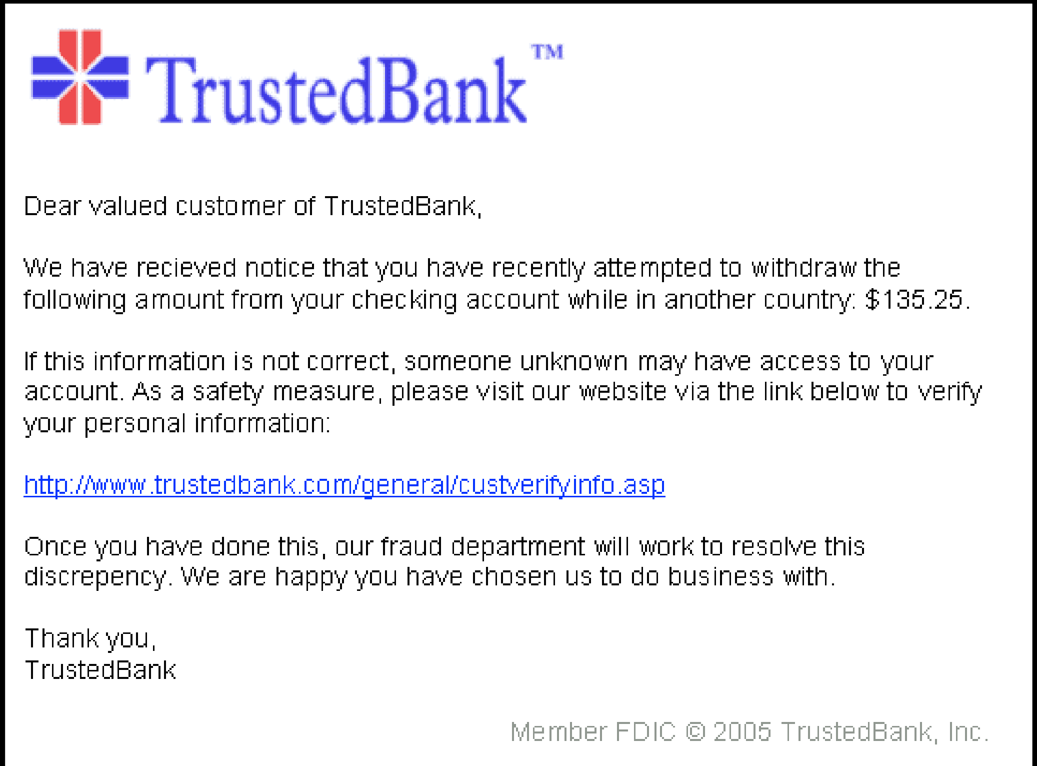


Figure 16. Known inital access vectors over time in non-Error, non-Misuse breaches (n in 2025 dataset=9,891)

Source: 2025 Verizon Data Breach Investigations Report

Phishing

- **User is tricked** into thinking that a **site** or **e-mail** is legitimate
  - Real logos, plausible circumstances
  - But actually: It is a **scam**
- May **stoke fear**, sense of **urgency**
  - Engages lizard brain!
- User persuaded to install malware or **perform other harmful actions**
  - By **clicking a link** or **engaging in a communication**
  - Or by **opening an attachment**



TrustedBank™

Dear valued customer of TrustedBank,

We have recieved notice that you have recently attempted to withdraw the following amount from your checking account while in another country: $135.25.

If this information is not correct, someone unknown may have access to your account. As a safety measure, please visit our website via the link below to verify your personal information:

http://www.trustedbank.com/general/custverifyinfo.asp

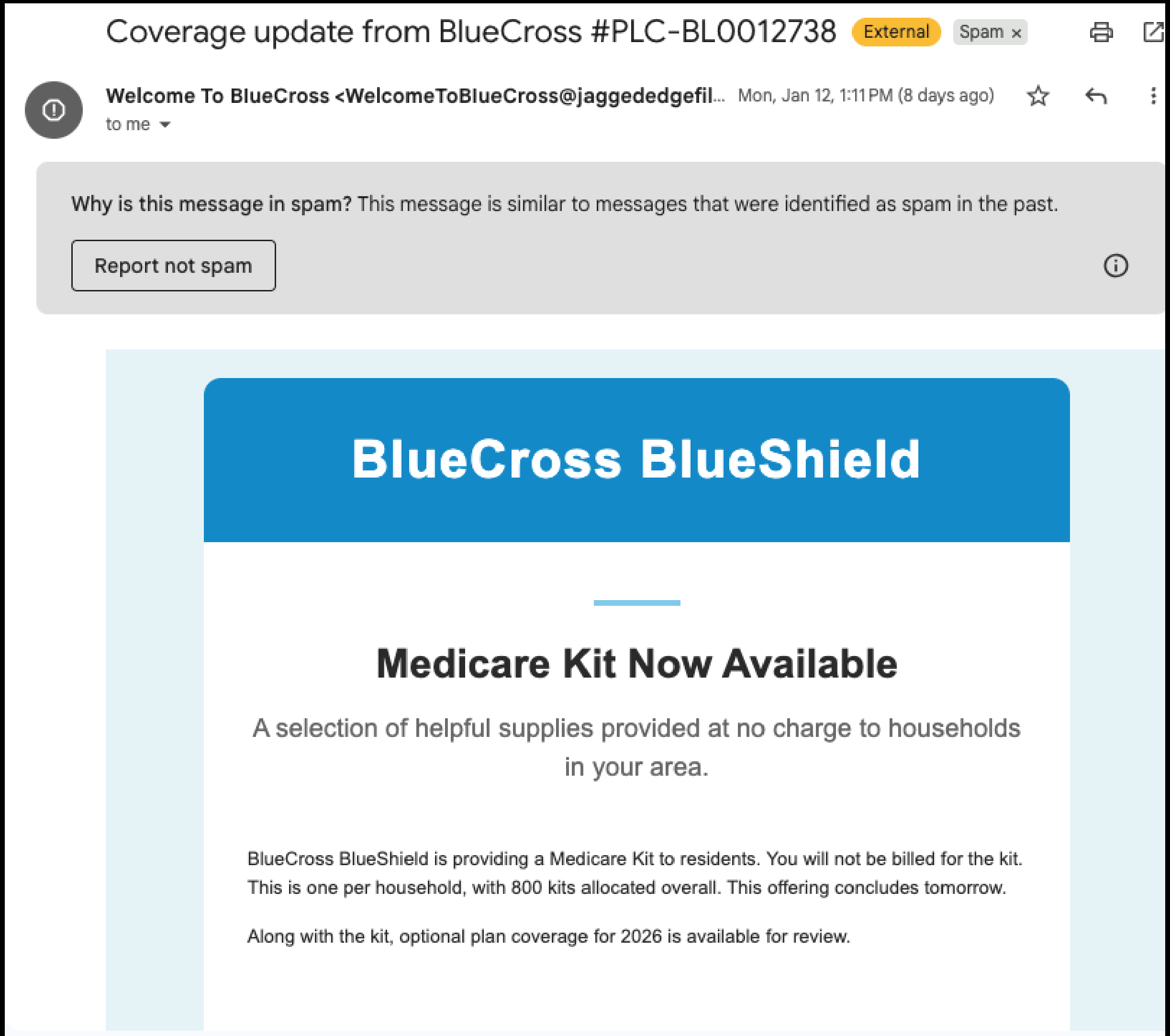
Once you have done this, our fraud department will work to resolve this discrepency. We are happy you have chosen us to do business with.

Thank you,
TrustedBank

Member FDIC © 2005 TrustedBank, Inc.

Phishing mitigation: Automated detection

Machine learning:
Looks similar to other phishing emails



Coverage update from BlueCross #PLC-BL0012738 External Spam ×

Welcome To BlueCross <WelcomeToBlueCross@jaggededgefil... Mon, Jan 12, 1:11 PM (8 days ago)
to me

Why is this message in spam? This message is similar to messages that were identified as spam in the past.

BlueCross BlueShield

Medicare Kit Now Available

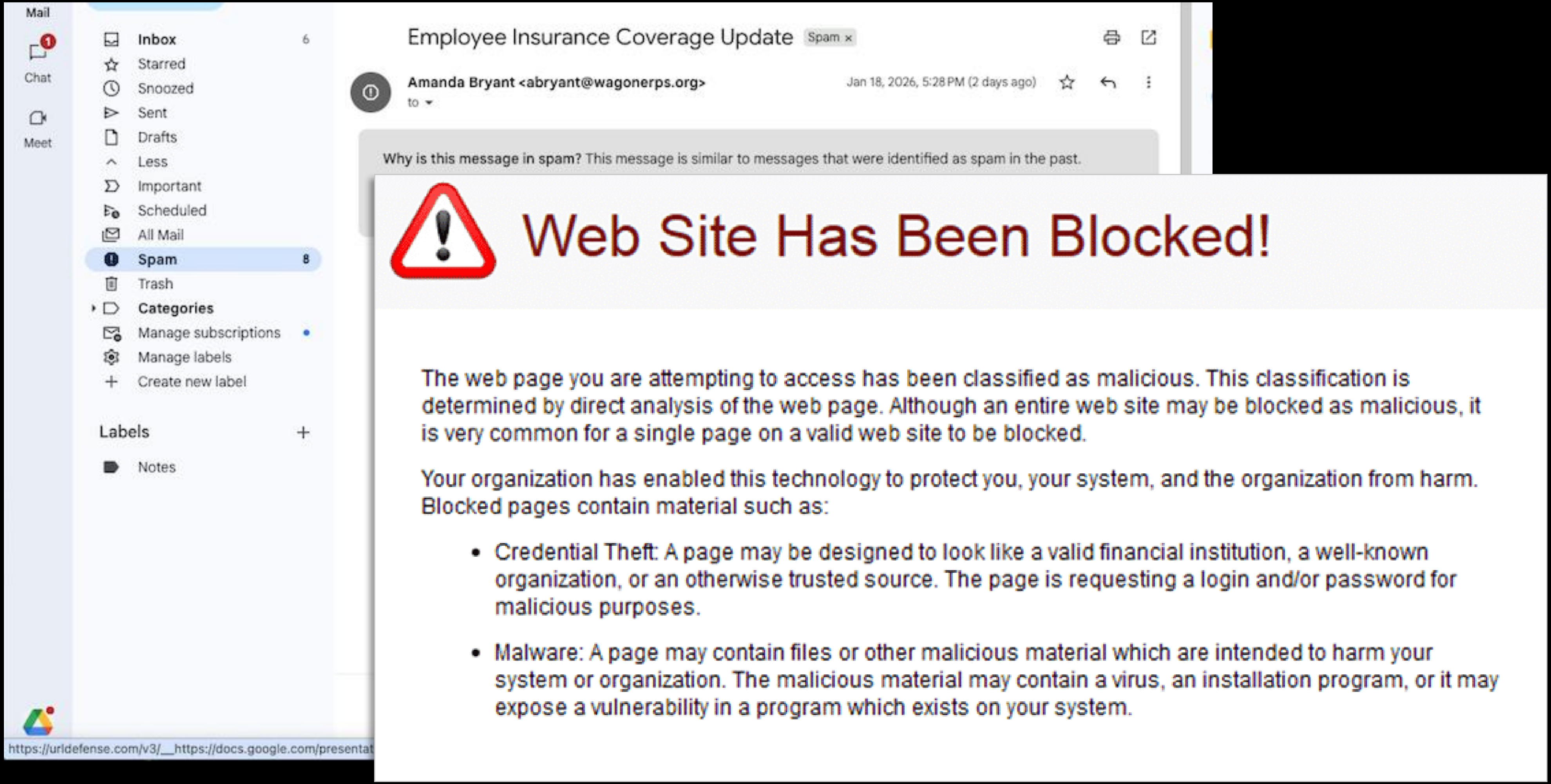
A selection of helpful supplies provided at no charge to households in your area.

BlueCross BlueShield is providing a Medicare Kit to residents. You will not be billed for the kit. This is one per household, with 800 kits allocated overall. This offering concludes tomorrow.

Along with the kit, optional plan coverage for 2026 is available for review.

Phishing mitigation: Link-based protection

Wraps link with redirect to security gatewau; will block clicks on bad links



Employee Insurance Coverage Update

Amanda Bryant <abryant@wagonerps.org>

Jan 18, 2026, 5:28 PM (2 days ago)

Why is this message in spam? This message is similar to messages that were identified as spam in the past.

https://urldefense.com/v3/__https://docs.google.com/presentat

Web Site Has Been Blocked!

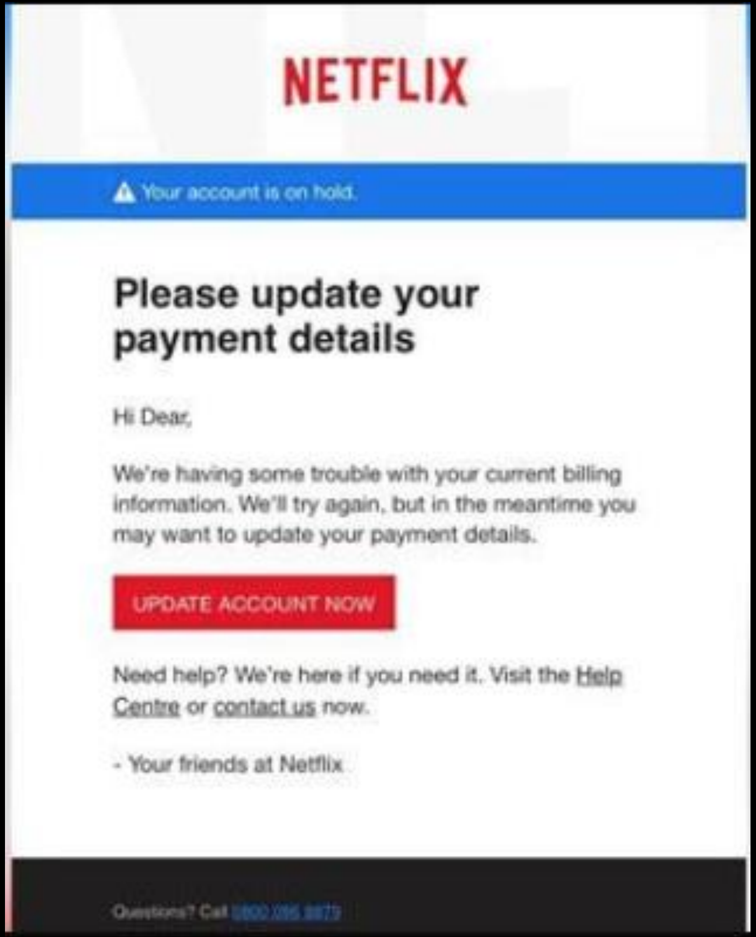
The web page you are attempting to access has been classified as malicious. This classification is determined by direct analysis of the web page. Although an entire web site may be blocked as malicious, it is very common for a single page on a valid web site to be blocked.

Your organization has enabled this technology to protect you, your system, and the organization from harm. Blocked pages contain material such as:

- Credential Theft: A page may be designed to look like a valid financial institution, a well-known organization, or an otherwise trusted source. The page is requesting a login and/or password for malicious purposes.
- Malware: A page may contain files or other malicious material which are intended to harm your system or organization. The malicious material may contain a virus, an installation program, or it may expose a vulnerability in a program which exists on your system.

Mitigation by training: Spot “bad smells”

- Text content phishy
  - The email has a generic greeting
  - The email says your account is on hold because of a billing problem
  - The email invites you to click on a link to update your payment details
- Link URL (hover over it) not to netflix.com
  - Netflix.com.ru not the same thing!
  - Email might be similarly bogus
  - .bit.ly links suspicious



NETFLIX

⚠ Your account is on hold.

Please update your payment details

Hi Dear,

We're having some trouble with your current billing information. We'll try again, but in the meantime you may want to update your payment details.

Need help? We're here if you need it. Visit the Help Centre or contact us now.

- Your friends at Netflix

Questions? Call 0800 095 8879

https://consumer.ftc.gov/articles/how-recognize-avoid-phishing-scams

Spear phishing

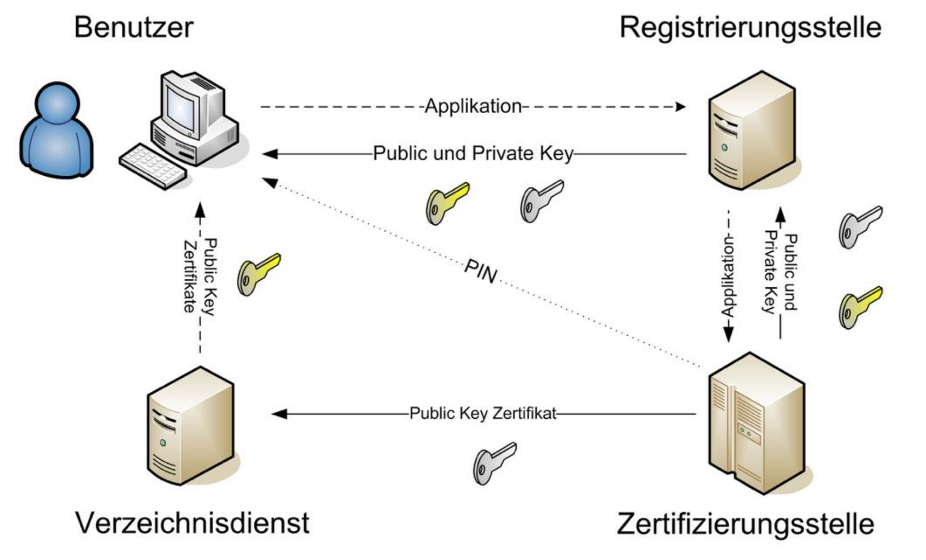
- Email is highly customized to the recipient
  - Think: Email sent to CEO, CIO, etc.
  - Contextually valid, appears to be from someone they know, etc.
  - Expensive to produce, but the hope for the attacker is a big payoff
- Several of the previous mitigations won’t work
  - No generic greeting, topic, etc.
  - Won’t look like prior phishing attempts
- More onus on the human to spot the deception

No perfect defense

- **Failure: Site or e-mail not (really) authenticated**
  - Internet e-mail and web protocols **not originally designed for remote authentication**
  - Solution is **hard to deploy**
    - Use hard-to-fake notions of identity, like **public key cryptography**. But which system? How to upgrade gradually?

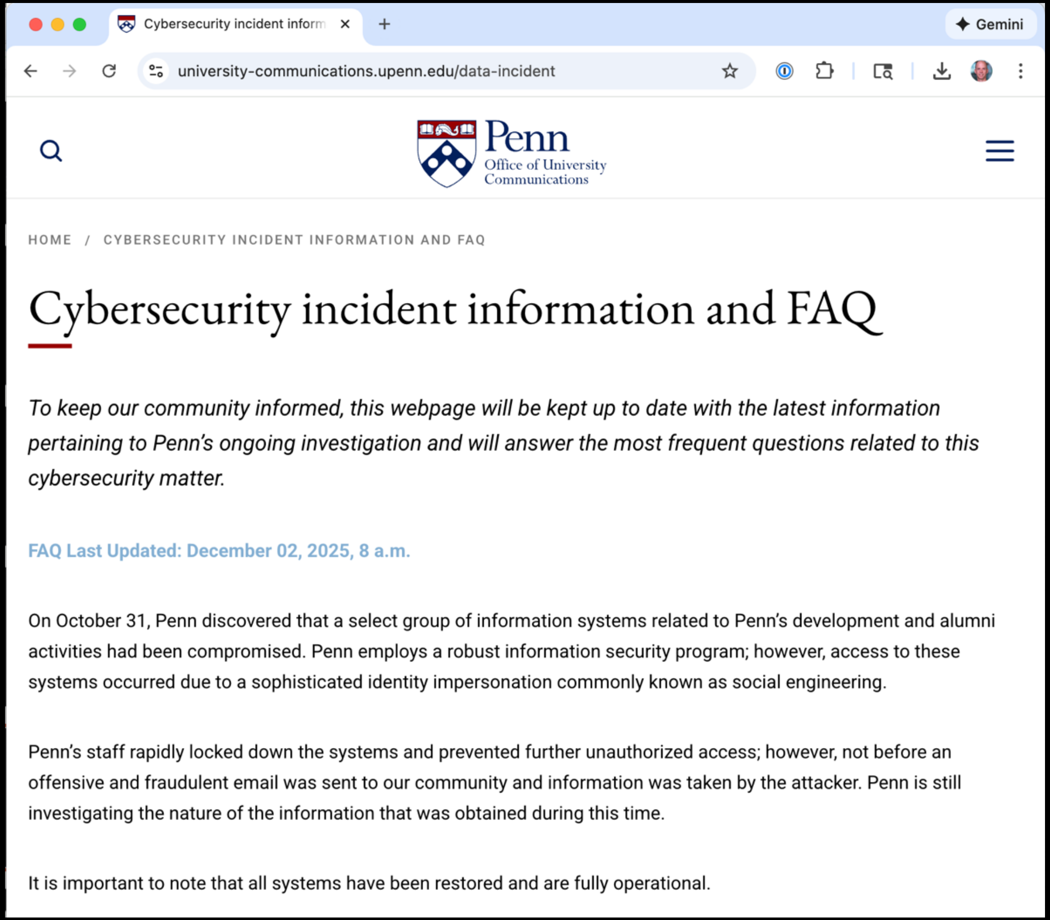
GPGTOOLS

THE ENIGMAIL PROJECT
OPENPGP EMAIL SECURITY FOR MOZILLA APPLICATIONS



Social engineering

- Social engineering is a type of psychological attack
  - Misleads target into doing something the attacker wants
  - Phishing is a kind of social engineering attack
- Vectors: Phone call, text message, etc.
- Signs
  - They are creating a sense of urgency
  - They are asking for information they should already have
  - The situation is too good to be true
- Defense: contact claimed source directly (bank, gov't agency, etc.)


Cybersecurity incident inform
Gemini
university-communications.upenn.edu/data-incident
Penn
Office of University Communications
HOME / CYBERSECURITY INCIDENT INFORMATION AND FAQ

Cybersecurity incident information and FAQ

To keep our community informed, this webpage will be kept up to date with the latest information pertaining to Penn's ongoing investigation and will answer the most frequent questions related to this cybersecurity matter.

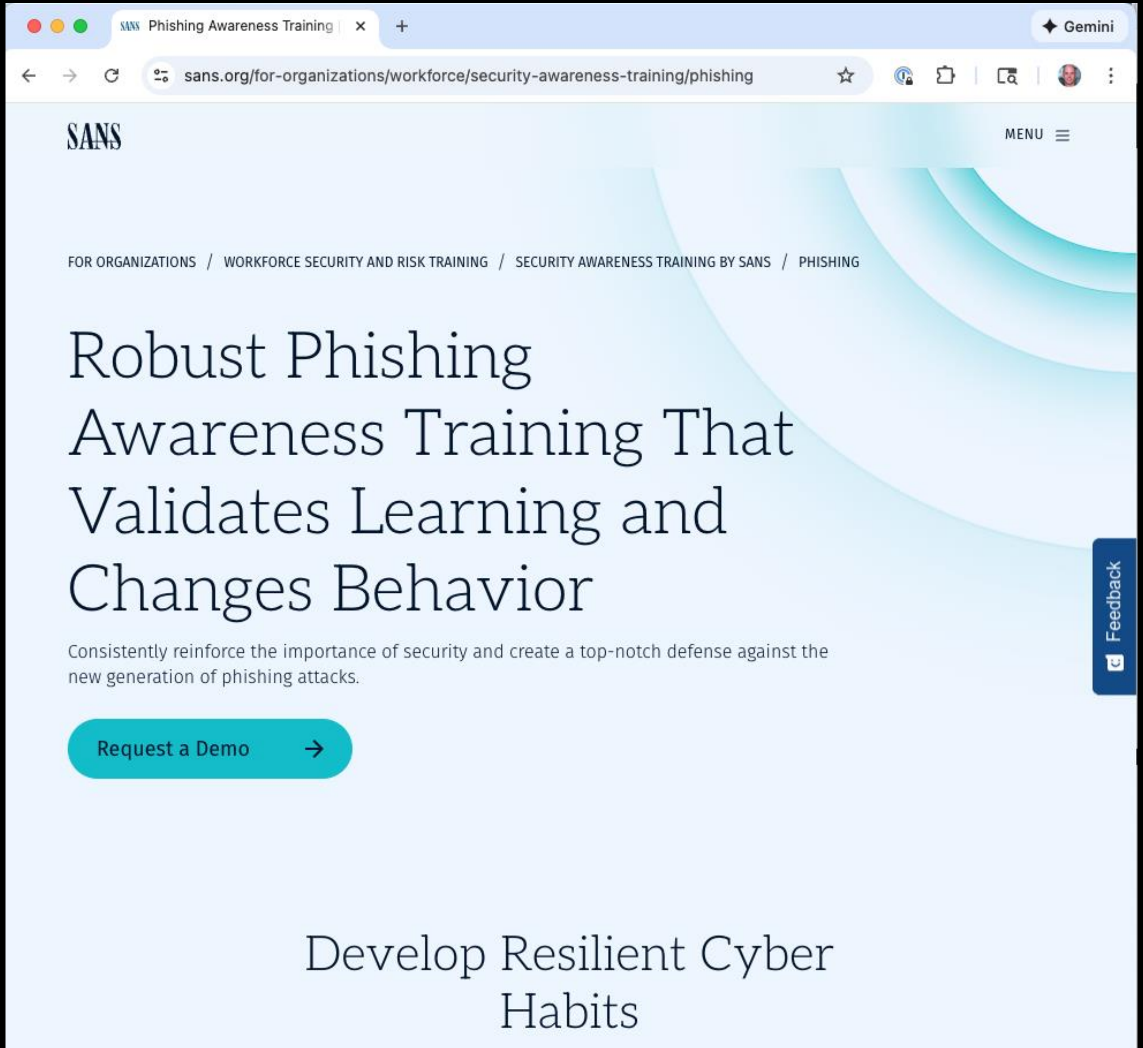
FAQ Last Updated: December 02, 2025, 8 a.m.

On October 31, Penn discovered that a select group of information systems related to Penn's development and alumni activities had been compromised. Penn employs a robust information security program; however, access to these systems occurred due to a sophisticated identity impersonation commonly known as social engineering.

Penn's staff rapidly locked down the systems and prevented further unauthorized access; however, not before an offensive and fraudulent email was sent to our community and information was taken by the attacker. Penn is still investigating the nature of the information that was obtained during this time.

It is important to note that all systems have been restored and are fully operational.

“Protect yourself” training: Does it work?



Phishing Awareness Training
Gemini
sans.org/for-organizations/workforce/security-awareness-training/phishing
SANS
MENU
Feedback

FOR ORGANIZATIONS / WORKFORCE SECURITY AND RISK TRAINING / SECURITY AWARENESS TRAINING BY SANS / PHISHING

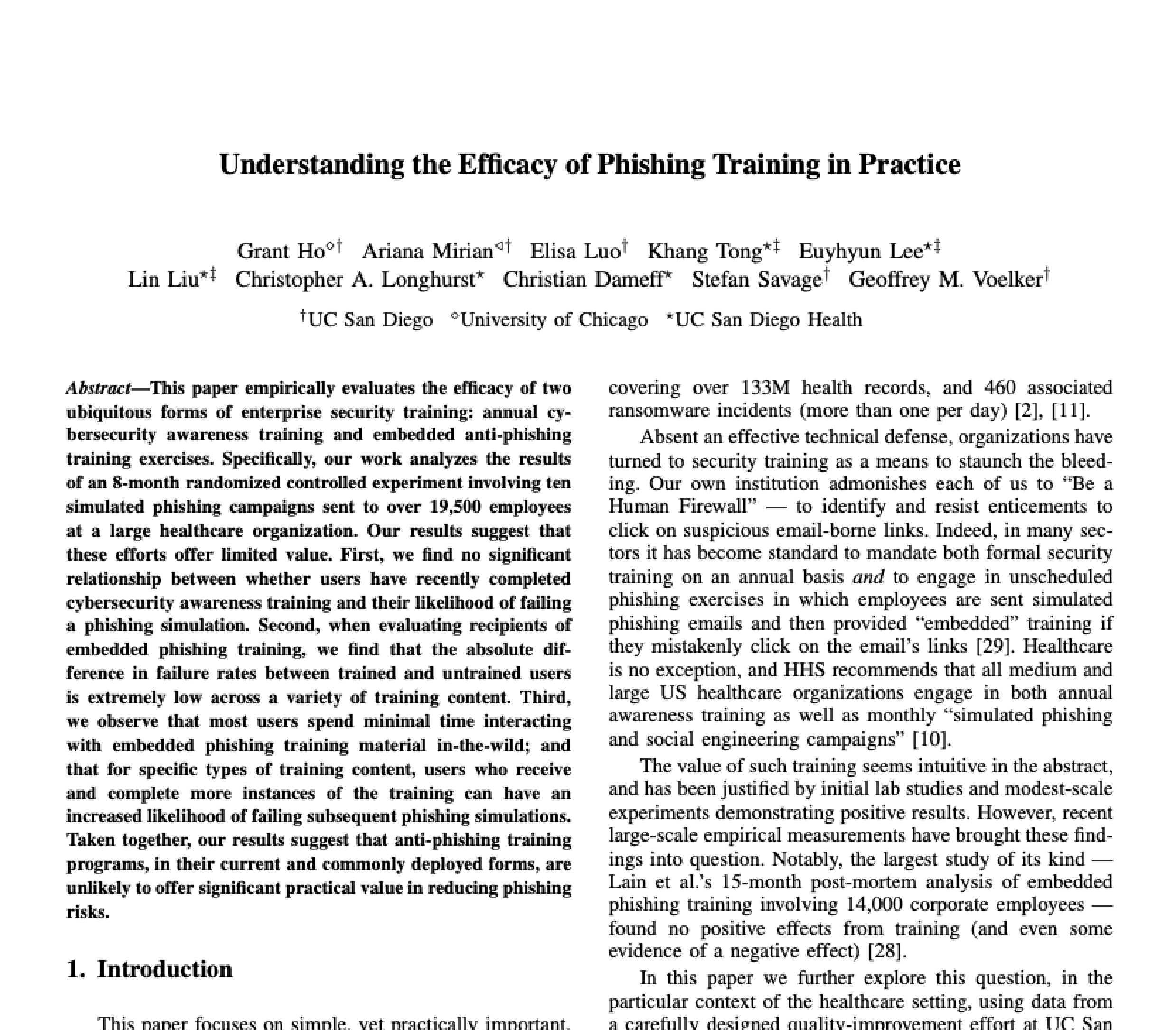
Robust Phishing Awareness Training That Validates Learning and Changes Behavior

Consistently reinforce the importance of security and create a top-notch defense against the new generation of phishing attacks.

Request a Demo →

Develop Resilient Cyber Habits

Reading



Understanding the Efficacy of Phishing Training in Practice

Grant Ho[◇†] Ariana Mirian[◁†] Elisa Luo[†] Khang Tong[⋆‡] Euyhyun Lee[⋆‡]
Lin Liu[⋆‡] Christopher A. Longhurst[⋆] Christian Dameff[⋆] Stefan Savage[†] Geoffrey M. Voelker[†]

[†]UC San Diego [◇]University of Chicago [⋆]UC San Diego Health

***Abstract*—This paper empirically evaluates the efficacy of two ubiquitous forms of enterprise security training: annual cybersecurity awareness training and embedded anti-phishing training exercises. Specifically, our work analyzes the results of an 8-month randomized controlled experiment involving ten simulated phishing campaigns sent to over 19,500 employees at a large healthcare organization. Our results suggest that these efforts offer limited value. First, we find no significant relationship between whether users have recently completed cybersecurity awareness training and their likelihood of failing a phishing simulation. Second, when evaluating recipients of embedded phishing training, we find that the absolute difference in failure rates between trained and untrained users is extremely low across a variety of training content. Third, we observe that most users spend minimal time interacting with embedded phishing training material in-the-wild; and that for specific types of training content, users who receive and complete more instances of the training can have an increased likelihood of failing subsequent phishing simulations. Taken together, our results suggest that anti-phishing training programs, in their current and commonly deployed forms, are unlikely to offer significant practical value in reducing phishing risks.**

1. Introduction

This paper focuses on simple, yet practically important,

covering over 133M health records, and 460 associated ransomware incidents (more than one per day) [2], [11].

Absent an effective technical defense, organizations have turned to security training as a means to staunch the bleeding. Our own institution admonishes each of us to "Be a Human Firewall" — to identify and resist enticements to click on suspicious email-borne links. Indeed, in many sectors it has become standard to mandate both formal security training on an annual basis *and* to engage in unscheduled phishing exercises in which employees are sent simulated phishing emails and then provided "embedded" training if they mistakenly click on the email's links [29]. Healthcare is no exception, and HHS recommends that all medium and large US healthcare organizations engage in both annual awareness training as well as monthly "simulated phishing and social engineering campaigns" [10].

The value of such training seems intuitive in the abstract, and has been justified by initial lab studies and modest-scale experiments demonstrating positive results. However, recent large-scale empirical measurements have brought these findings into question. Notably, the largest study of its kind — Lain et al.'s 15-month post-mortem analysis of embedded phishing training involving 14,000 corporate employees — found no positive effects from training (and even some evidence of a negative effect) [28].

In this paper we further explore this question, in the particular context of the healthcare setting, using data from a carefully designed quality-improvement effort at UC San

Key research questions

- How well do common trainings protect against phishing attacks?
    - Annual security awareness training & Embedded phishing training

- Do different forms and styles of training have varying levels of anti-phishing protection?
    - Interactive vs. Static material, Generic vs. Contextualized content, etc.

- What underlying practical factors decrease training's efficacy?

These slides are based on ones provided by Grant Ho – thanks!

Study Design: Real-world RCT at UCSD Health

UCSD Health: existing embedded phishing training -> controlled experiment

- 19,500+ employees: anonymized statistics (IRB & QI approval)
- 8 months: 10 different phishing emails of varying sophistication
- Clicking on the embedded phishing link = “failed” simulation

Randomly partitioned all users into 5 training groups

- 1 control group: no training (“404 not found page”)
- 2 static training groups: generic vs. customized educational page
- 2 interactive training groups: generic & customized content

Statistical analysis of results w/ GLME models to control for confounders

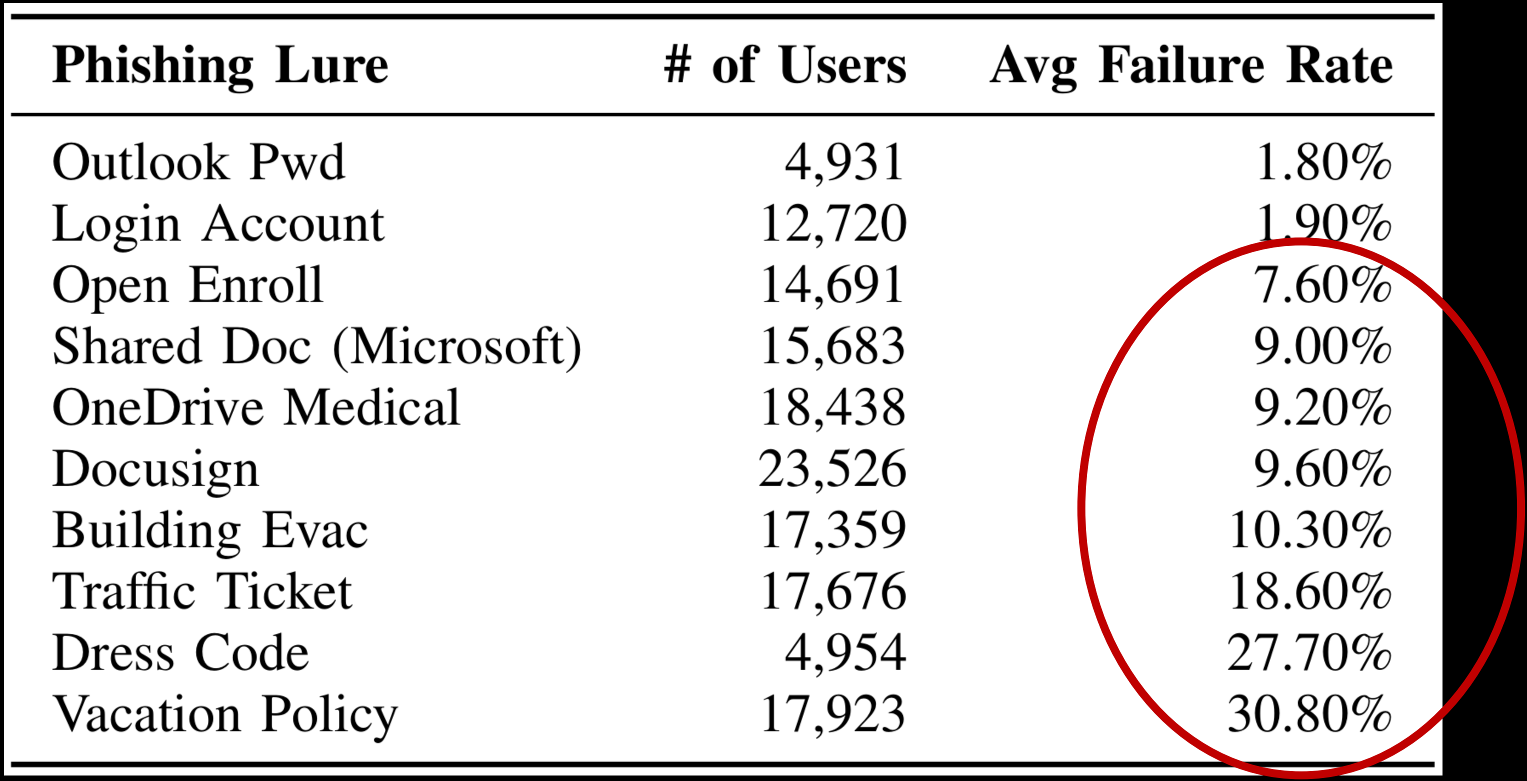
Result: Embedded Phishing Training has Minimal Benefit

Phishing Lure	Control	Generic Static	Context. Static	Generic Interactive	Context. Interactive
Login Account	3.44%	1.14%	1.27%	0.97%	1.13%
Outlook Pwd	1.62%	1.72%	2.41%	1.85%	1.52%
John Davis	9.56%	7.01%	6.38%	6.4%	7.45%
Docusign	11.06%	9.98%	10.2%	10.05%	9.75%
OneDrive Medical	9.89%	9.37%	8.54%	9.25%	9.16%
Open Enroll	9.02%	6.67%	6.68%	7.01%	6.76%
Vacation Policy	31.02%	30.58%	31.99%	30.58%	29.85%
Traffic Ticket	20.39%	20.07%	16.37%	17.25%	19.37%
Building Evac	11.67%	8.25%	8.4%	8.55%	9.32%
Dress Code	29.96%	27.01%	27.41%	26.98%	26.88%

Overall, training yielded only a 1.7% average improvement over control group

- See paper for additional analysis & statistical models

Phishing Lure Efficacy Varies & Far Outstrips Protection



Phishing Lure	# of Users	Avg Failure Rate
Outlook Pwd	4,931	1.80%
Login Account	12,720	1.90%
Open Enroll	14,691	7.60%
Shared Doc (Microsoft)	15,683	9.00%
OneDrive Medical	18,438	9.20%
Docusign	23,526	9.60%
Building Evac	17,359	10.30%
Traffic Ticket	17,676	18.60%
Dress Code	4,954	27.70%
Vacation Policy	17,923	30.80%

The phishing lure used (chosen by attacker) ultimately dictates the attack outcome:

- 1.7% improvement from training vs. 30.8% clickthrough rate from specific lures

Limited Engagement w/ Embedded Training

User statistics across training for all simulated phishing emails:

< 25%	complete & acknowledge the training
> 33%	immediately close the training website
> 75%	spend < 1 minute on the training website

Paper has many additional results & further analysis:

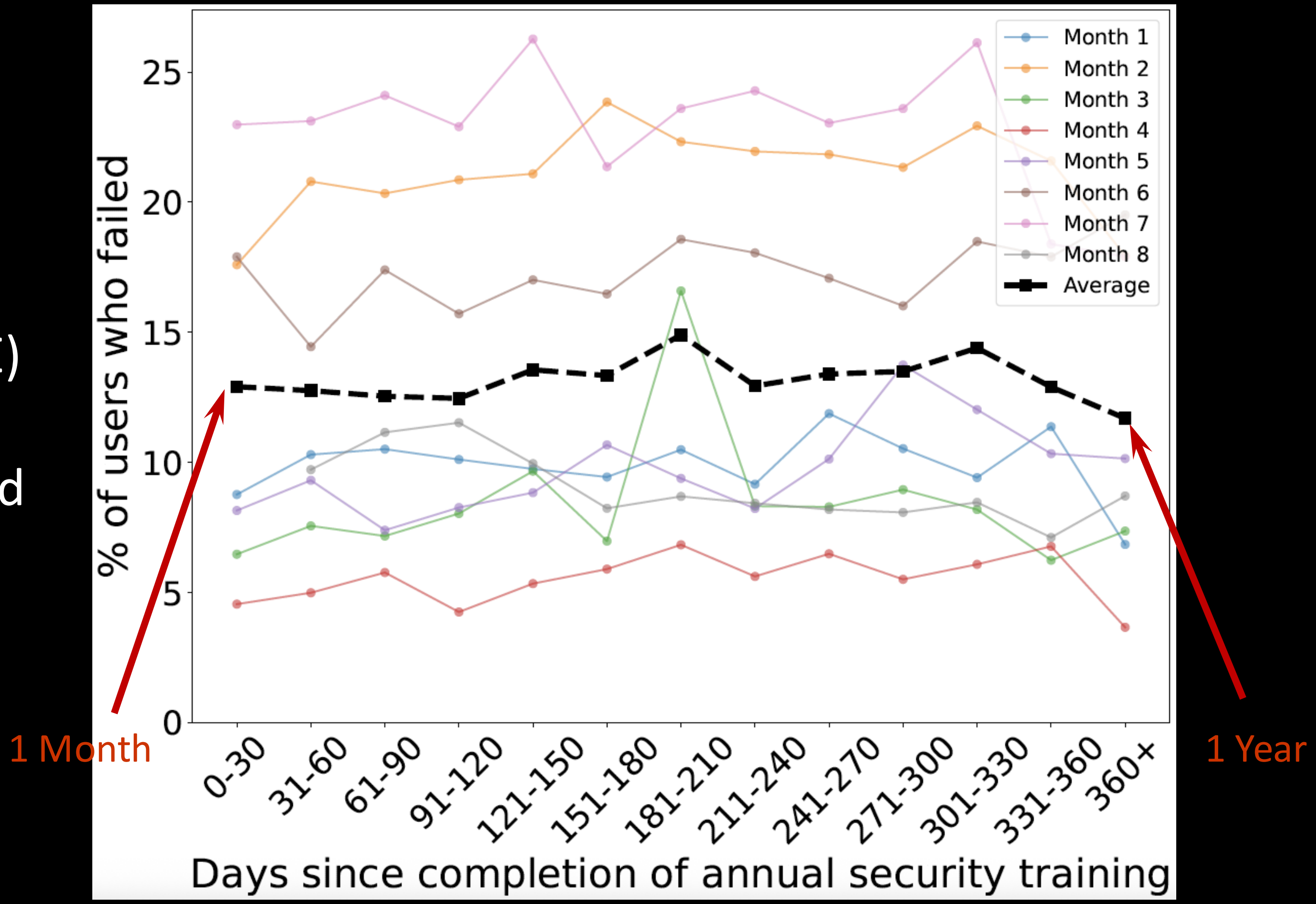
- Outcomes for users w/ substantial training engagement, premised on inefficient training delivery, impact of training material design, etc.

Annual Awareness Training has Minimal Benefit

Mandatory annual security awareness training made by a leading vendor (KnowBe4)

Data and statistical models (GLME) find ***no association*** between:

1. how long ago a user completed training and
2. their likelihood of failing a phishing simulation



Summary and reflection

- Minimal anti-phishing benefits from common security training
  - Embedded phishing training & Annual cybersecurity awareness training
  - 19.5k employees: 8-months of in-situ, randomized controlled experiments
- Study limitations: one institution, one type of phishing action
  - But results consistent with prior large-scale studies

- Realistically, phishing training will continue… so what should we do?
  - Whatever it is, we should evaluate it rigorously & empirically in-the-wild: go beyond marketing claims