OrangeCon 2024 Amsterdam

How to crack billions of passwords?

Jeroen van Beek

Scattered Secrets

Who am I

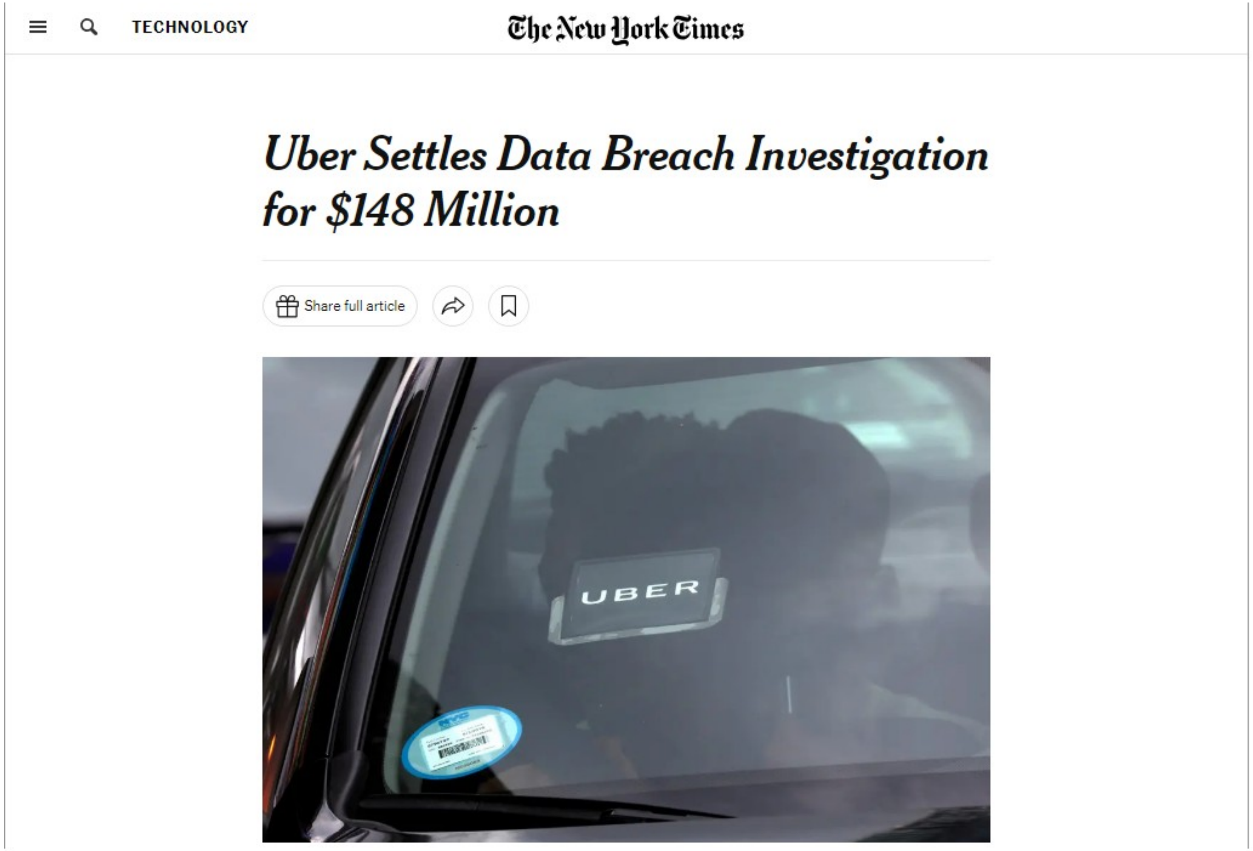
- Hardware evaluator @research lab
- Lecturer “offensive technologies” @university
- Penetration tester, consultant @self employed (was: Big Four)
- Co-founder @Scattered Secrets

Taylor Swift, what happened?

Taylor Swift, what happened?

- Taylor Swift ticketing handled by Ticketmaster
- Ticketmaster uses Snowflake (*“that powers the AI Data Cloud”*)
- Snowflake got hacked:
  - No mandatory Multi Factor Authentication for admins
  - Mandiant: *"[..]at least 79.7% of the accounts leveraged by the threat actor in this campaign had prior credential exposure"*
  - Data of 560 million Ticketmaster users affected
  - Besides Ticketmaster, seemingly 165+ other organizations affected
    - This story will continue for many months or years
      - Probably too much work for the hackers to ransom and extort them all at once

Uber, what happened?



TECHNOLOGY

The New York Times

Uber Settles Data Breach Investigation for $148 Million

Share full article

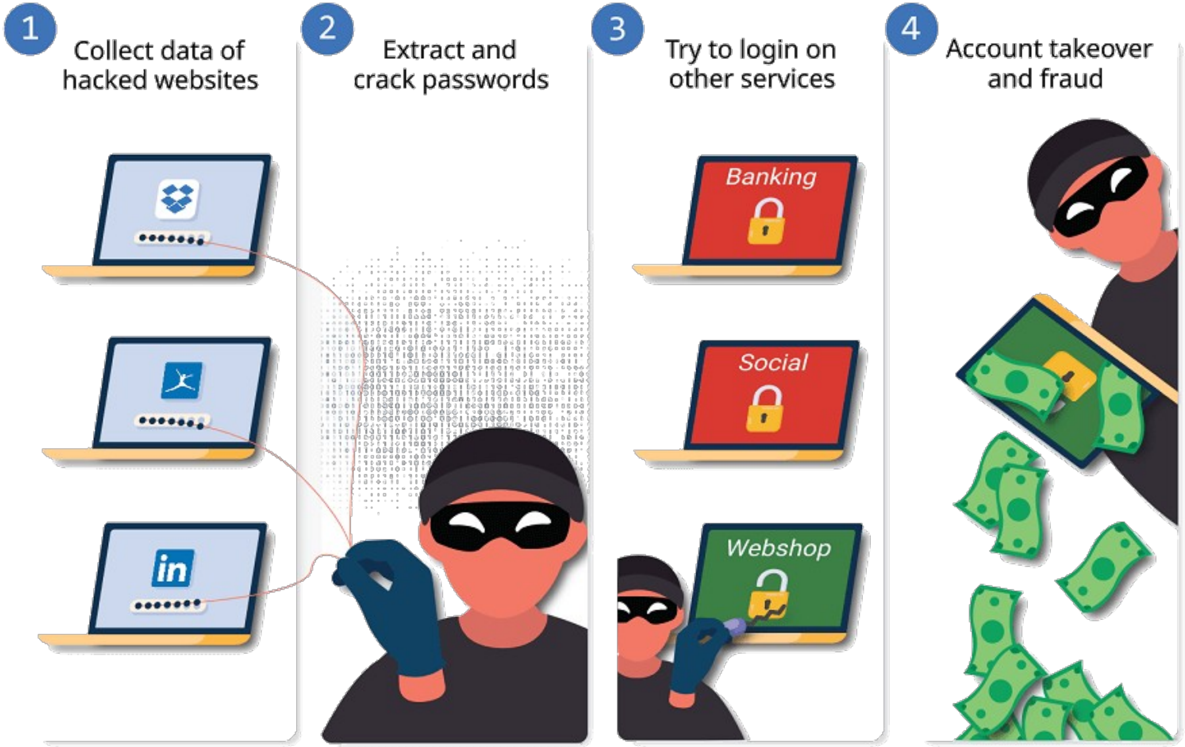
Uber, what happened?

- Uber source code stored private in GitHub repository
  - No mandatory Multi Factor Authentication for developers
- Developer re-used password on other services:
  - Other service got hacked, password cracked
- GitHub account take over:
  - Storage credentials in Uber source code (AWS S3)
  - Data of 57 million Uber users and 600 thousand drivers stolen
- Bonus:
  - Uber did not report the breach & paid hackers $100,000 to keep quiet

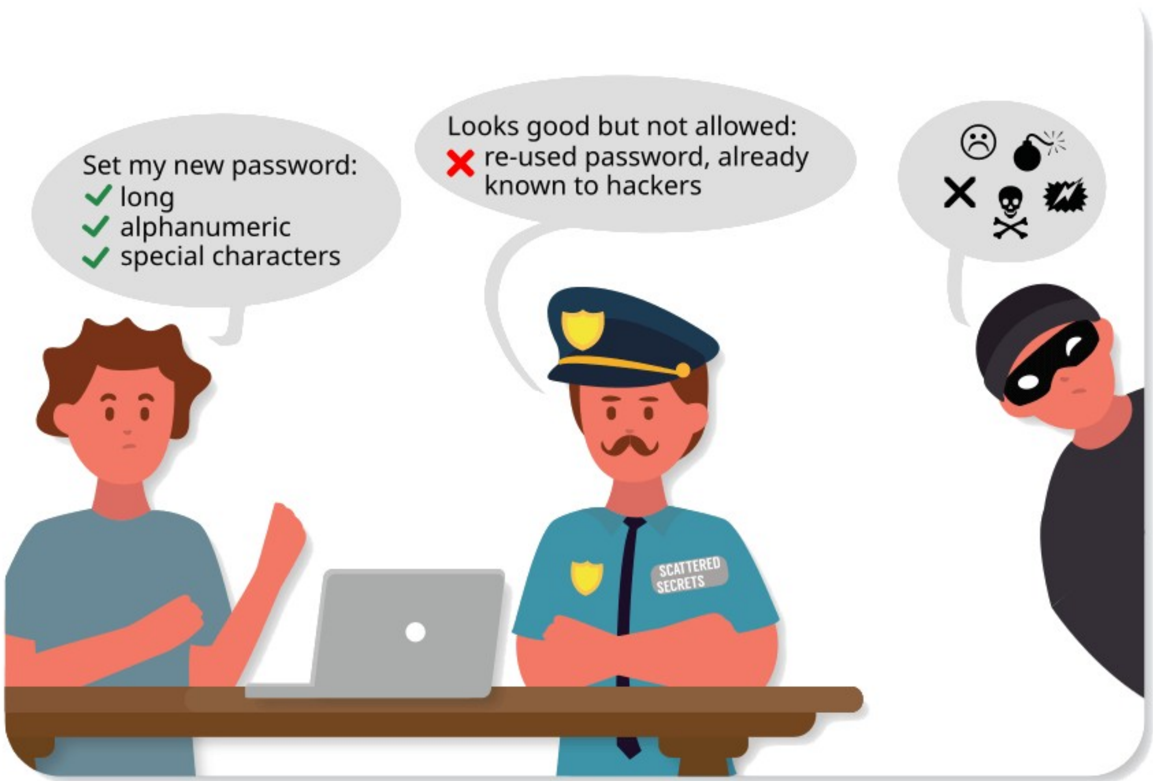
Account takeover is the new 0day(?)

- Timeline:
  - Exploit operating system flaws
  - Exploit other software flaws
  - Exploit application code
  - Exploit account takeovers 🡐 **we are here**
- Verizon: *“credential abuse is the big thing to focus on”*
- Huge increase in amount of available breach data in the last years
  - Bulk sets with hundreds of millions of lines are quite common
- Huge decrease in costs and complexity for an attacker

What do cybercriminals do?


1
Collect data of hacked websites
2
Extract and crack passwords
3
Try to login on other services
4
Account takeover and fraud
in
Banking
Social
Webshop

What does Scattered Secrets do?

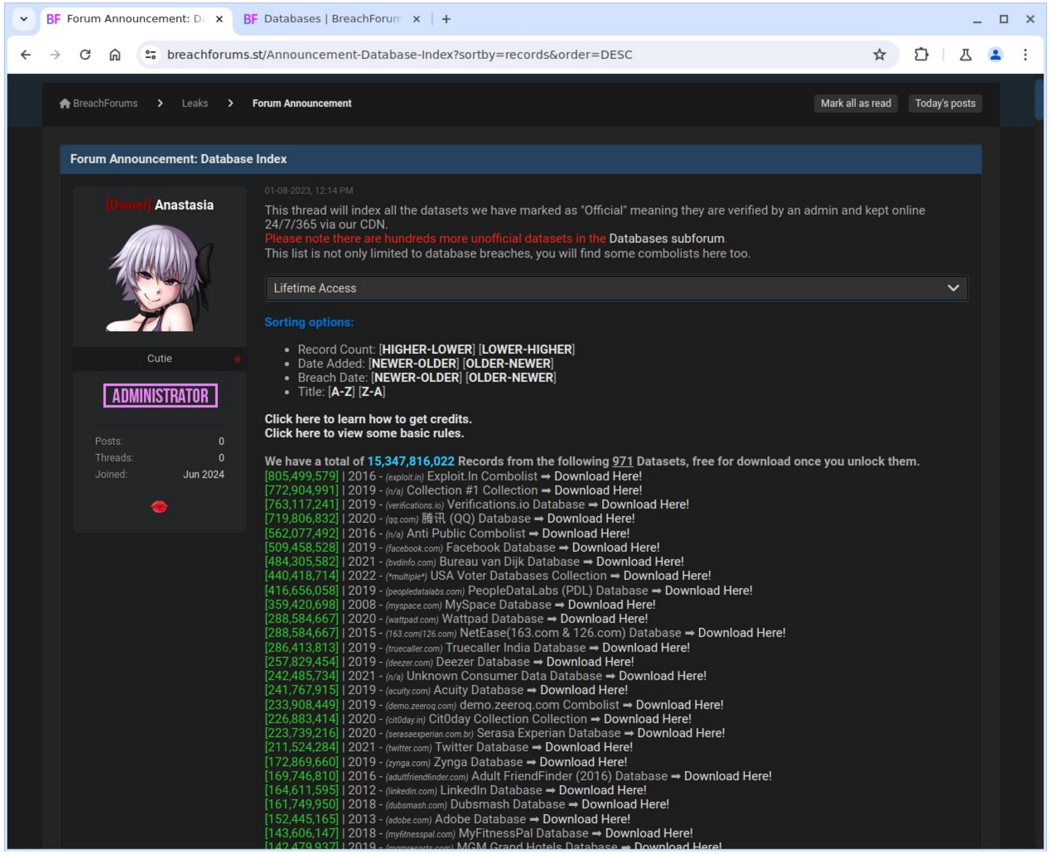


Aaaah, like HIBP / Apple / Microsoft?

- No:
  - HaveIBeenPwnd: email without password ➔ false positives
    - Pwnd Passwords: password without email, see Apple
  - Apple: password without context ➔ limited risk
    - Alert if a grocery store in Australia is using the same password
      - To takeover an account, an attacker needs the account name too
  - Microsoft: password blacklisting ➔ different scope
    - *"global banned password list"* and optionally a *"custom banned password list"*
      - Protects against password spraying, not against credential stuffing
- What we provide: email + password pairs ➔ actionable information

How to get a quality dataset?

- Step 1: collect raw data
  - The amount of available data is overwhelming
  - Basic task
- Step 2: extract credentials
  - cut, grep, sed, awk, jq etc.
  - Basic task
- Step 3: analysis & cracking
  - Here you can make the difference
  - But what kind of data is out there?



BF Forum Announcement: D
BF Databases | BreachForum
breachforums.st/Announcement-Database-Index?sortby=records&order=DESC
BreachForums > Leaks > Forum Announcement
Mark all as read
Today's posts

Forum Announcement: Database Index

[Owner] Anastasia

Cutie

ADMINISTRATOR

Posts: 0
Threads: 0
Joined: Jun 2024

01-08-2023, 12:14 PM

This thread will index all the datasets we have marked as "Official" meaning they are verified by an admin and kept online 24/7/365 via our CDN.
Please note there are hundreds more unofficial datasets in the Databases subforum.
This list is not only limited to database breaches, you will find some combolists here too.

Lifetime Access

Sorting options:

- Record Count: [HIGHER-LOWER] [LOWER-HIGHER]
- Date Added: [NEWER-OLDER] [OLDER-NEWER]
- Breach Date: [NEWER-OLDER] [OLDER-NEWER]
- Title: [A-Z] [Z-A]

Click here to learn how to get credits.
Click here to view some basic rules.

We have a total of 15,347,816,022 Records from the following 971 Datasets, free for download once you unlock them.
[805,499,579] | 2016 - (exploit.in) Exploit.In Combolist ➡ Download Here!
[772,904,991] | 2019 - (n/a) Collection #1 Collection ➡ Download Here!
[763,117,241] | 2019 - (verifications.io) Verifications.io Database ➡ Download Here!
[719,806,832] | 2020 - (qq.com) 腾讯 (QQ) Database ➡ Download Here!
[562,077,492] | 2016 - (n/a) Anti Public Combolist ➡ Download Here!
[509,458,528] | 2019 - (facebook.com) Facebook Database ➡ Download Here!
[484,305,582] | 2021 - (bvdinfo.com) Bureau van Dijk Database ➡ Download Here!
[440,418,714] | 2022 - (*multiple*) USA Voter Databases Collection ➡ Download Here!
[416,656,058] | 2019 - (peopledatalabs.com) PeopleDataLabs (PDL) Database ➡ Download Here!
[359,420,698] | 2008 - (myspace.com) MySpace Database ➡ Download Here!
[288,584,667] | 2020 - (wattpad.com) Wattpad Database ➡ Download Here!
[288,584,667] | 2015 - (163.com|126.com) NetEase(163.com & 126.com) Database ➡ Download Here!
[286,413,813] | 2019 - (truecaller.com) Truecaller India Database ➡ Download Here!
[257,829,454] | 2019 - (deezer.com) Deezer Database ➡ Download Here!
[242,485,734] | 2021 - (n/a) Unknown Consumer Data Database ➡ Download Here!
[241,767,915] | 2019 - (acuity.com) Acuity Database ➡ Download Here!
[233,908,449] | 2019 - (demo.zeeroq.com) demo.zeeroq.com Combolist ➡ Download Here!
[226,883,414] | 2020 - (cit0day.in) Cit0day Collection Collection ➡ Download Here!
[223,739,216] | 2020 - (serasaexperian.com.br) Serasa Experian Database ➡ Download Here!
[211,524,284] | 2021 - (twitter.com) Twitter Database ➡ Download Here!
[172,869,660] | 2019 - (zynga.com) Zynga Database ➡ Download Here!
[169,746,810] | 2016 - (adultfriendfinder.com) Adult FriendFinder (2016) Database ➡ Download Here!
[164,611,595] | 2012 - (linkedin.com) LinkedIn Database ➡ Download Here!
[161,749,950] | 2018 - (dubsmash.com) Dubsmash Database ➡ Download Here!
[152,445,165] | 2013 - (adobe.com) Adobe Database ➡ Download Here!
[143,606,147] | 2018 - (myfitnesspal.com) MyFitnessPal Database ➡ Download Here!
[142,479,937] | 2019 - (mgmresorts.com) MGM Grand Hotels Database ➡ Download Here!

Plain passwords

In the old days (early 1960s), passwords were stored in plain text only

- Trouble:
  - Password file breached means all accounts breached
    - Including randomly generated quality passwords
    - Perfect for seeding dictionaries
  - Plain storage still occasionally used nowadays!
- Fixes:
  - ~~Obfuscation~~
  - 6th Edition Unix (1974): basic password hashing using crypt(3)
    - Actually an encryption algorithm (DES) used for hashing
    - Maximum password length = 8

Basic hashes (MD-* SHA-*)

Introduction of basic plain hashing (early 1990s)

- Trouble:
  - Same password results in same hash: hashes can be pre-calculated
    - A big lookup table with hash ↔ password
  - Efficient and effective cracking introduced
    - Crack (1991), Jack (1993), John the Ripper (1996)
- Fix: salting, not *hash(password)* but *salt + hash(password + salt)*:
  - N times more compute required for N hashes, no more pre-calculation

Basic hashes with salts

Introduction of salted basic algorithms (mid 1990s)

- Trouble:
  - Herd immunity only, no protection against targeted attacks
    - Cracking a single salted hash is just as fast as cracking a plain basic hash
  - Moore's law goes password cracking: CPU power doubles about every 2 years
    - Algorithms unchanged ➔ easier to crack passwords over time
- Fix: *key stretching*, use multiple rounds to slow a cracker down:
  - MD5crypt (1994): 1,000x MD-5 (and a salt), later examples: SHA*crypt

Basic hashing with salts and iterations

Introduction of salted and key stretched basic algorithms (mid 1990s)

- Trouble:
  - New technology introduced for cracking: Graphic Processing Units
    - General purpose programming with CUDA (2007), OpenCL (2009)
    - 50 to 100x speed-up for most hashing algorithms
  - Moore's law also works for GPUs
    - Algorithms unchanged → easier to crack passwords over time
- Fix: hardware-specific slowdowns & password specific algorithms
  - Use more fast memory than available on the cracking platform
  - For example bcrypt (1999): does not work efficiently for GPUs, even today
    - Number of iterations can be configured as well

Hashing with salts, iterations and H/W limitations

Introduction of H/W slowdowns & password specific algos (late 1990s+)

- Trouble:
  - New technology introduced for cracking: Field Programmable Gate Arrays
    - Specific FPGAs supported for specific algorithms (2019)
    - 50 to 100x speed-up for specific algorithms when released (including bcrypt)
  - Moore's law still works for this and all earlier technology
    - Algorithms unchanged → at some time tech will overcome the H/W slowdowns
- Fix: fully configurable hardware-specific slowdowns:
  - Scrypt (2009), Argon2 (2015): ← **developers, you want one of those**
    - Configurable number of iterations, configurable resource usage
      - Update parameters over time, to beat the latest and greatest cracking platform

Effective & efficient cracking

- Password cracking is an exponential problem:
  - Brute force 10M hashes, mixed-case alphanumeric with special characters:
    - Basic hash: length ~8-9
    - Salted basic hash: length ~5
    - Key stretched salted basic hash: length ~3-4
    - Key stretched salted hash with H/W limitations: length ~1-2 ← **more and more bcrypt seen**
    - N extra length requires ~100^N times more computing power
      - Just “buy more computers” is not a practical solution
- Think different:
  - Use tricks
  - Use the best and most powerful platform for each specific job
  - Data analysis on cracked passwords to improve cracking rules over time

Tricks

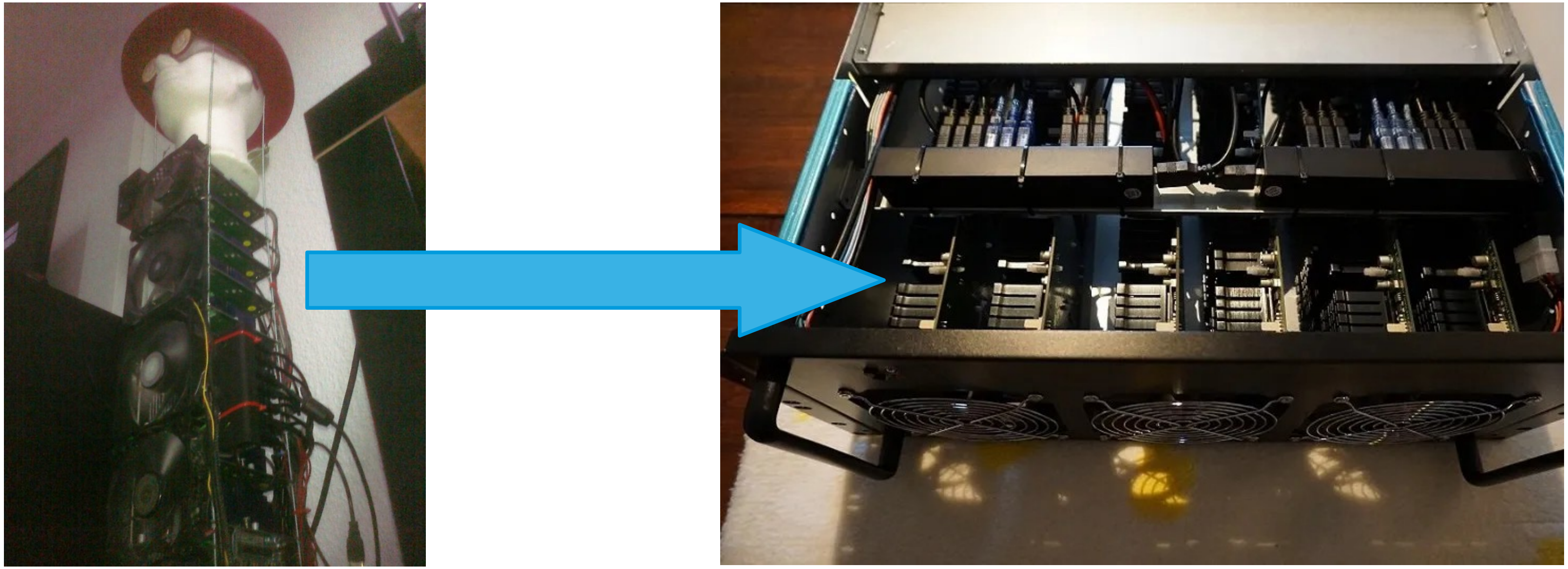
- Applications and code changes over time:
  - Website updates from MD-5 to bcrypt with workfactor 12
    - From extremely fast to extremely slow
  - Reverse engineer leaked code:
    - Database still stores case-insensitive legacy hashes for upgraded accounts
      - For upgrade paths and backward compatibility
      - Reset procedures still uses mechanism based on legacy algorithm
    - Crack legacy fast hashes, use results as dictionary for new slow hashes
- Great success:
  - Using MD-5 based loginkey to recover a significant percentage of bcrypts
  - That's over a 100 million times speed-up!

Best cracking platforms

- Basic hashes: GPUs
  - With or without salts and key stretching
  - Bigger is better:
    - Consumer grade hardware is most cost effective
- Huge sets of basic hashes / complex rules: CPUs
  - GPUs haven't got enough VRAM memory
  - Use high core count devices with AVX-512 support
- Hashes with hardware slowdowns: FPGAs
  - More speed for less power
  - Exotic breed, find supported devices and built it yourself

Best cracking platforms, ct.

- From zero to (data centre) hero:
  - Each box contains 72 FPGAs @~585 Watts
    - Bcrypt performance per box: ~14 RTX-4090s @~6,500 Watts



Results & observations

- Over 7 billion email + password pairs recovered
  - On average 1 to 2 million a day, especially a lot of unique bcrypt content
- There's a lot of fake password news and disinformation out there
- A typical organization:
  - Over ~1,337 accounts: account takeovers just works (≥1 accounts)
  - At scale: 5-15% of accounts can be breached
  - No (effective) protection against credential stuffing
- Old passwords matter
  - Still new passwords recovered from the 2012 LinkedIn breach in 2024

Results & observations, ct.

- The grass is greener on the other side:
  - Cyber criminals often perform test-runs before the full attack starts
    - Too few results ➔ try an easier victim
    - Outsmart your competitors, even basic measures can help
- A good password policy can work counterproductive:
  - Alphanumeric, mixed-case, ≥12 length, organization-specific blacklist
  - The ***number*** of leaked candidate passwords goes down, but the ***quality*** of candidates goes up drastically ➔ higher efficiency & stealthier

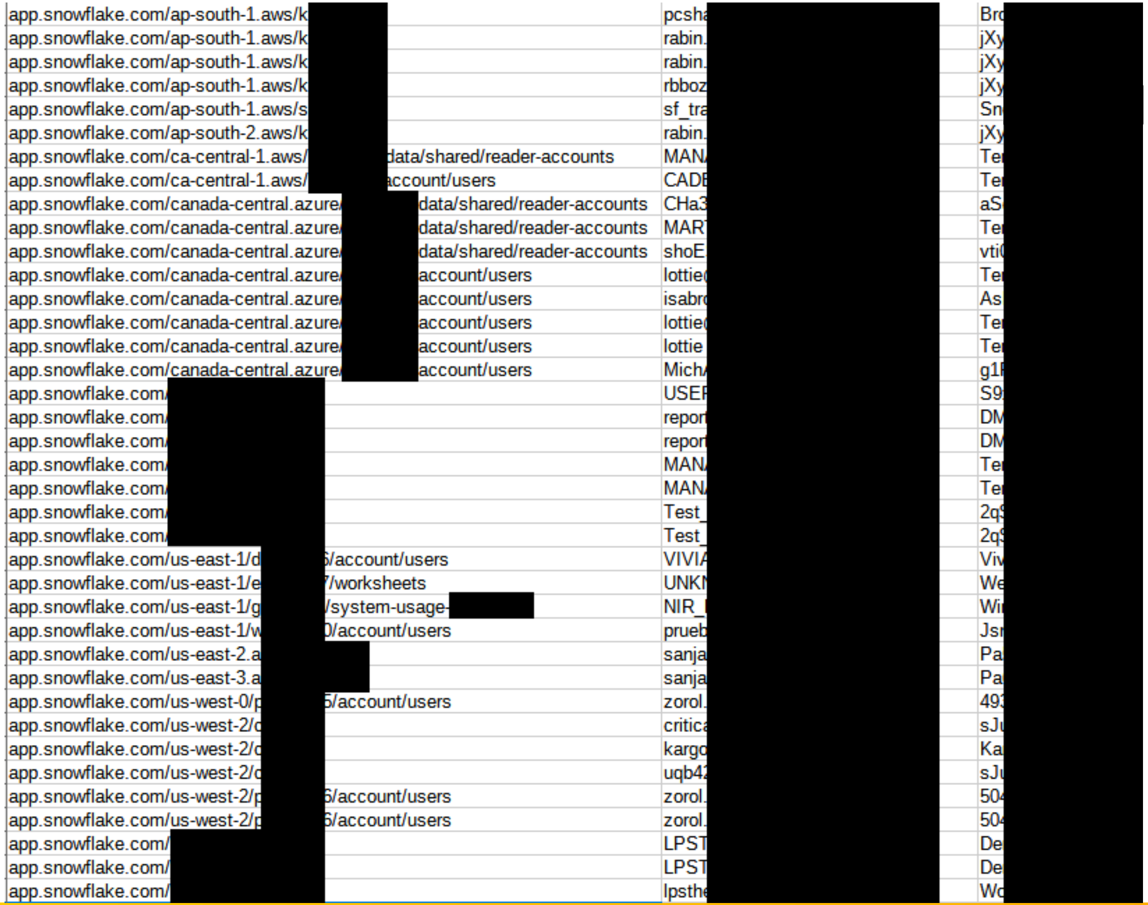
Fake news

- Hookers.nl got hacked:
  - Statement: *“no passwords stolen”*
  - Reality: cracked 57% of ~300k passwords in three days
    - This was a quick burn-in test for one of our then new boxes

No basic protection

- Hacked because of missing basic cyber hygiene? Say this:
  - *“Very sophisticated attack, never seen before”*
  - *“State sponsored threat actor”*
- Reality:
  - A school kid with $5 and a gaming PC can launch a successful attack
    - If you can run Call of Duty, you can brute force an eight position basic hash in days

No basic protection, ct.



app.snowflake.com/ap-south-1.aws/k[redacted]	pcsha[redacted]	Br[redacted]
app.snowflake.com/ap-south-1.aws/k[redacted]	rabin.[redacted]	jXy[redacted]
app.snowflake.com/ap-south-1.aws/k[redacted]	rabin.[redacted]	jXy[redacted]
app.snowflake.com/ap-south-1.aws/k[redacted]	rbboz[redacted]	jXy[redacted]
app.snowflake.com/ap-south-1.aws/s[redacted]	sf_tra[redacted]	Sn[redacted]
app.snowflake.com/ap-south-2.aws/k[redacted]	rabin.[redacted]	jXy[redacted]
app.snowflake.com/ca-central-1.aws/[redacted]data/shared/reader-accounts	MAN[redacted]	Te[redacted]
app.snowflake.com/ca-central-1.aws/[redacted]account/users	CADE[redacted]	Te[redacted]
app.snowflake.com/canada-central.azure/[redacted]data/shared/reader-accounts	CHa3[redacted]	aS[redacted]
app.snowflake.com/canada-central.azure/[redacted]data/shared/reader-accounts	MAR[redacted]	Te[redacted]
app.snowflake.com/canada-central.azure/[redacted]data/shared/reader-accounts	shoE[redacted]	vti[redacted]
app.snowflake.com/canada-central.azure/[redacted]account/users	lottie[redacted]	Te[redacted]
app.snowflake.com/canada-central.azure/[redacted]account/users	isabr[redacted]	As[redacted]
app.snowflake.com/canada-central.azure/[redacted]account/users	lottie[redacted]	Te[redacted]
app.snowflake.com/canada-central.azure/[redacted]account/users	lottie[redacted]	Te[redacted]
app.snowflake.com/canada-central.azure/[redacted]account/users	Mich[redacted]	g1[redacted]
app.snowflake.com/[redacted]	USER[redacted]	S9[redacted]
app.snowflake.com/[redacted]	report[redacted]	DM[redacted]
app.snowflake.com/[redacted]	report[redacted]	DM[redacted]
app.snowflake.com/[redacted]	MAN[redacted]	Te[redacted]
app.snowflake.com/[redacted]	MAN[redacted]	Te[redacted]
app.snowflake.com/[redacted]	Test_[redacted]	2q[redacted]
app.snowflake.com/[redacted]	Test_[redacted]	2q[redacted]
app.snowflake.com/us-east-1/d[redacted]6/account/users	VIVIA[redacted]	Viv[redacted]
app.snowflake.com/us-east-1/e[redacted]7/worksheets	UNKN[redacted]	We[redacted]
app.snowflake.com/us-east-1/g[redacted]/system-usage-[redacted]	NIR_[redacted]	Wi[redacted]
app.snowflake.com/us-east-1/w[redacted]0/account/users	prueb[redacted]	Js[redacted]
app.snowflake.com/us-east-2.a[redacted]	sanja[redacted]	Pa[redacted]
app.snowflake.com/us-east-3.a[redacted]	sanja[redacted]	Pa[redacted]
app.snowflake.com/us-west-0/p[redacted]5/account/users	zorol.[redacted]	493[redacted]
app.snowflake.com/us-west-2/c[redacted]	critica[redacted]	sJu[redacted]
app.snowflake.com/us-west-2/c[redacted]	kargo[redacted]	Ka[redacted]
app.snowflake.com/us-west-2/c[redacted]	uqb42[redacted]	sJu[redacted]
app.snowflake.com/us-west-2/p[redacted]6/account/users	zorol.[redacted]	504[redacted]
app.snowflake.com/us-west-2/p[redacted]6/account/users	zorol.[redacted]	504[redacted]
app.snowflake.com/[redacted]	LPST[redacted]	De[redacted]
app.snowflake.com/[redacted]	LPST[redacted]	De[redacted]
app.snowflake.com/[redacted]	lpsthe[redacted]	Wo[redacted]

Old passwords matter

- 2012 Dropbox password, still used for all her accounts:

..normaal zijn dat soort wachtwoorden een aantal jaren geleden gehackt, normaal hebben mensen die wel aangepast. Dat heb jij niet gedaan.

Uuhm, nee.

Old passwords matter, ct.

- Passwords can help you to follow persons over time:
  - Historical passwords can help to unravel the mechanisms used
  - Find personal accounts based on professional accounts and vice versa
- Let’s check politicians!
  - Some politicians use a high quality password, that is unique worldwide
  - Using this password you can identify their personal accounts:
    - john.doe@official.tld → johnd@gmail.com, jdoe@hotmail.com, jdburner@yahoo.com
  - Member of house of parliament: what did we find?

Old passwords matter, ct.

- pr0n :)

Thank you OrangeCon!

Questions?

Remarks?

Check your accounts for breached passwords:

https://scatteredsecrets.com/

Jeroen van Beek
jeroen@scatteredsecrets.com

Multi Factor Authentication is not a silver bullet

- MFA not enabled on ***all*** *internal* services ➔ you’re still vulnerable
- MFA not enabled on ***all*** *external* services ➔ you’re still vulnerable
  - Or is your staff not using sensitive information on third party platforms?
- Specific attacks are out there to successfully bypass MFA:
  - MiTM reverse proxies stealing session tokens
  - Keyloggers stealing session tokens *(“infostealers”)*
  - MFA fatigue attacks, and other types of social engineering
- However using MFA is better than not using it!