How Viable is Password Cracking in Digital Forensic Investigation? Analyzing the Guessability of over 3.9 Billion Real-World Accounts

By:
Aikaterini Kanta (University College Dublin and European Commission, Joint Research Centre, Sein Coray (University of Basel), Iwen Coisel (European Commission, Joint Research Centre), and Mark Scanlon (University College Dublin)

From the proceedings of
The Digital Forensic Research Conference
DFRWS USA 2021
July 12-15, 2021

How Viable is Password Cracking in Digital Forensic Investigation? Analyzing the Guessability of Over 3.9 Billion Real-World Accounts

DFRWS USA 2021

AIKATERINI KANTA, SEIN CORAY, IWEN COISEL AND MARK SCANLON

UCD DUBLIN

European Commission
JOINT RESEARCH CENTRE

UCD Forensics and
Security Research Group

Contribution of this work

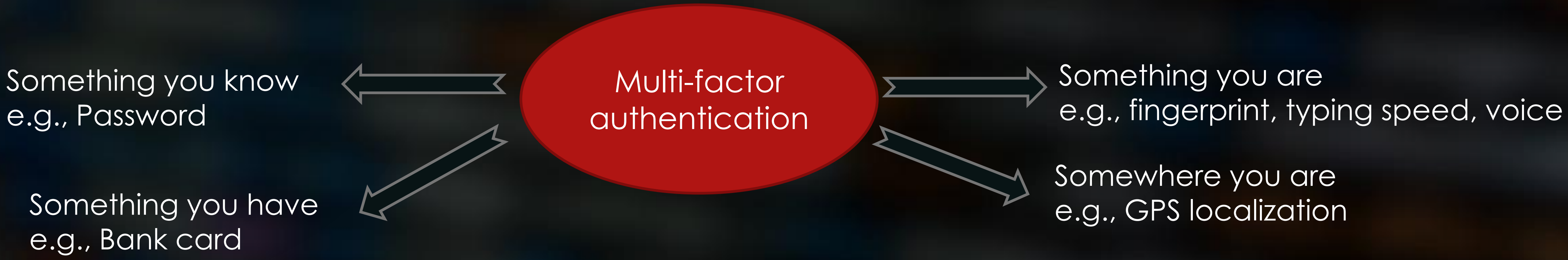
- The largest and most comprehensive analysis of real-world passwords conducted to date.
- An analysis of the passwords' pattern of construction after splitting the passwords into meaningful component fragments.
- A look into the most common semantic classes and what they mean for users' password construction.
- An analysis on the strength and crackability of the passwords.

All About Authentication

The average number of passwords users need to remember:
27 in one study to 191 in another

Usage

Type



Multi-factor authentication

Something you know
e.g., Password

Something you have
e.g., Bank card

Something you are
e.g., fingerprint, typing speed, voice

Somewhere you are
e.g., GPS localization

Password Cracking Attacks

Traditionally
Brute force attacks… guaranteed to work
Dictionary Attacks + mangling rules

More recently
Machine learning and AI techniques

US National Institute of Standards and Technology
recommendations
NIST 2013: LUDS-8 (updated in NIST 2017)

In practice:
NVIDIA 3080 GPU
54 10^9 passwords/s
600 euro
password 8 characters
2 days

f16QL~!>5mX#9dgj”+2

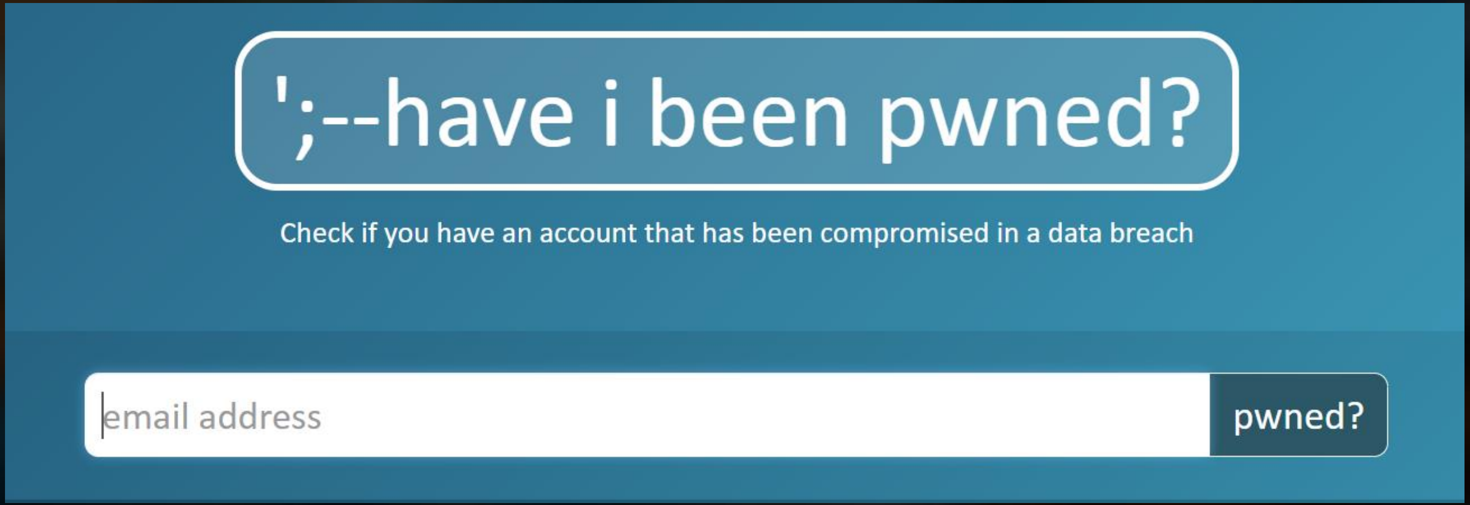
Age of universe: **13.8×10^9 years**
Time to crack this password: **22×10^{18} years**

Sounds like a good choice, right?
Depends! Data breaches and password reuse are a serious threat

Have I Been Pwned (HIBP)

Created by web security expert Troy Hunt to:

- Highlight the seriousness of data breaches
- Serve as a blacklist of passwords
- Help victims know their accounts have been compromised



';--have i been pwned?

Check if you have an account that has been compromised in a data breach

https://haveibeenpwned.com/

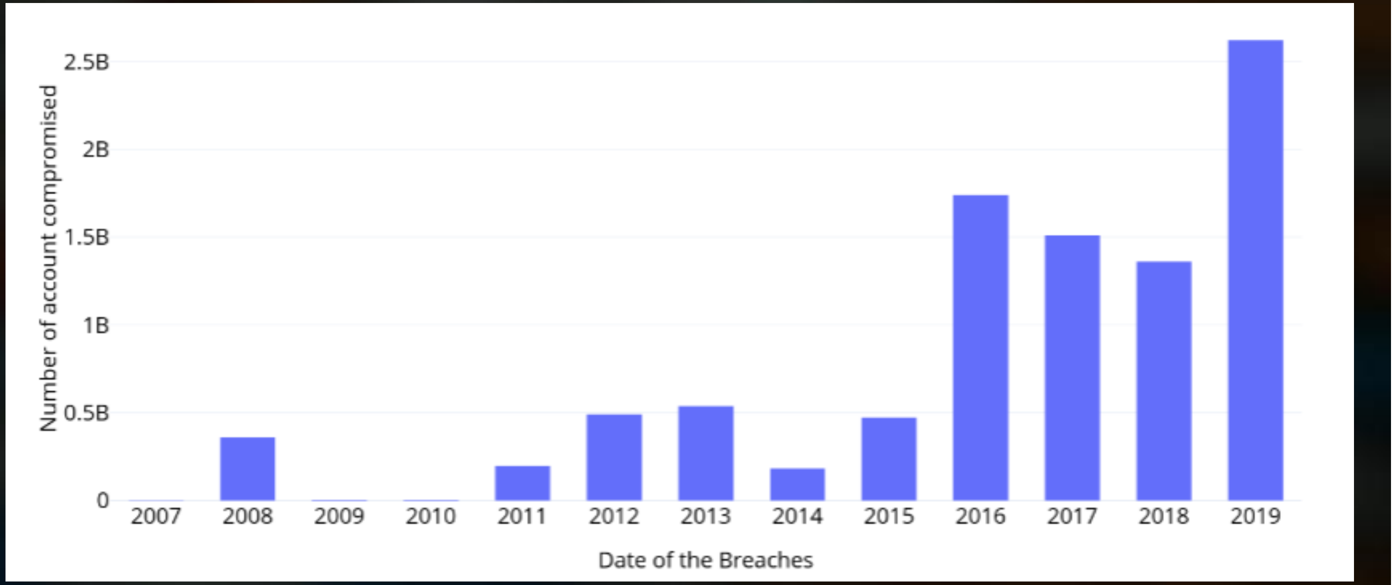
Analysis of Leaked Passwords

- Have I Been Pwned list
- 3.9 Billion Real-World Accounts
- Plaintext passwords from: hashes.org
- Statistics on:
  - Length
  - Makeup
  - Strength
- Fragment Analysis (Óðinn)
- Classification using WordNet®

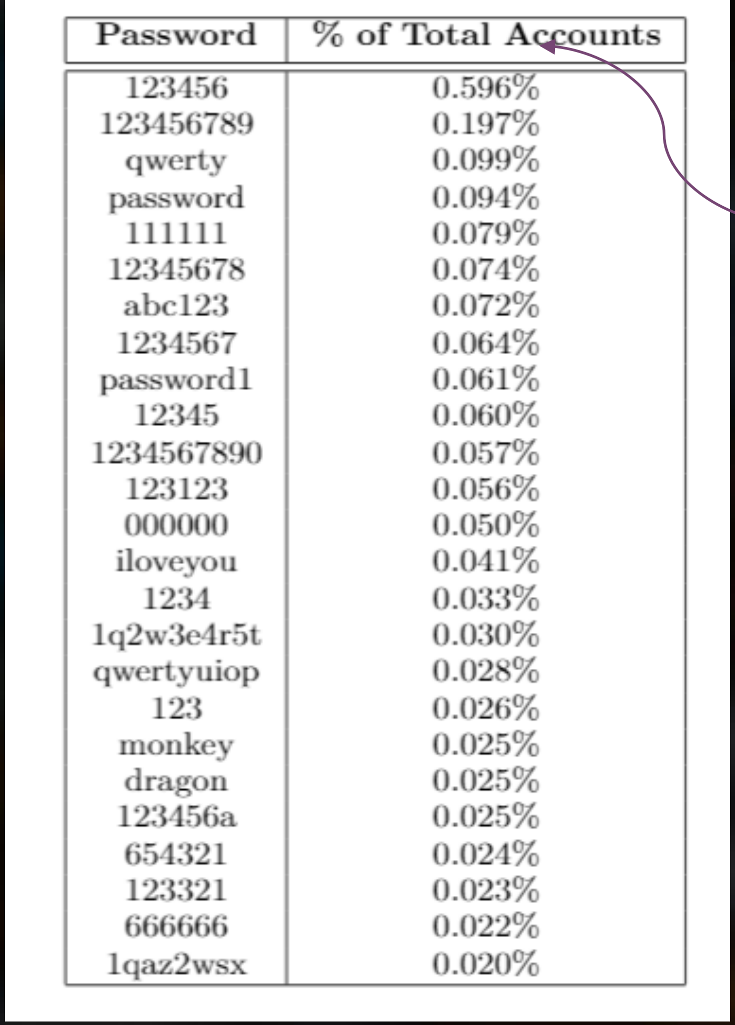
WordNet: https://wordnet.princeton.edu/

Have I Been Pwned Analysis 1

Number of Breached Accounts listed in Have I Been Pwned



Top 25 passwords in HIBP

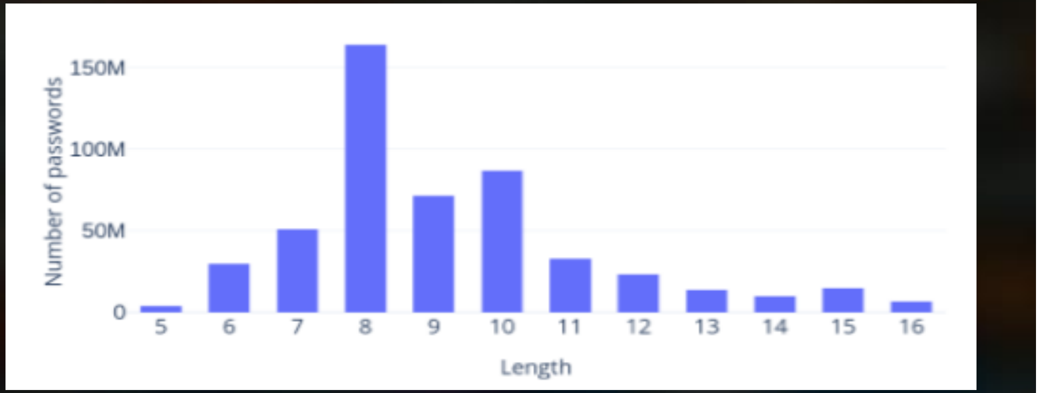


Password	% of Total Accounts
123456	0.596%
123456789	0.197%
qwerty	0.099%
password	0.094%
111111	0.079%
12345678	0.074%
abc123	0.072%
1234567	0.064%
password1	0.061%
12345	0.060%
1234567890	0.057%
123123	0.056%
000000	0.050%
iloveyou	0.041%
1234	0.033%
1q2w3e4r5t	0.030%
qwertyuiop	0.028%
123	0.026%
monkey	0.025%
dragon	0.025%
123456a	0.025%
654321	0.024%
123321	0.023%
666666	0.022%
1qaz2wsx	0.020%

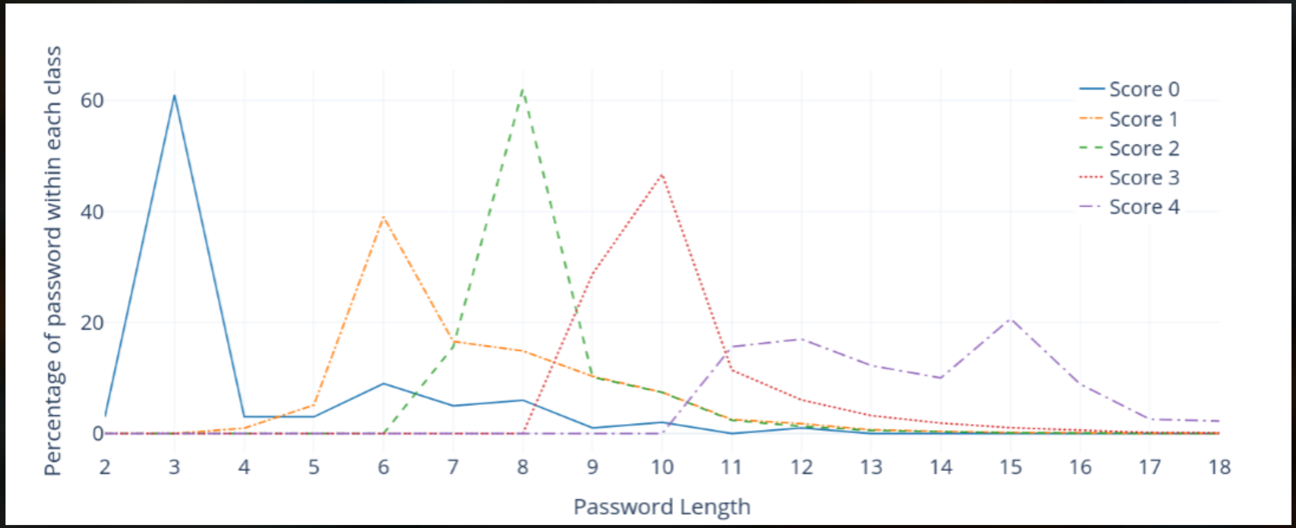
That's more than 23 million accounts!

Have I Been Pwned Analysis 2

Distribution of password length in HIBP



Distribution of password length per class (zxcvbn)

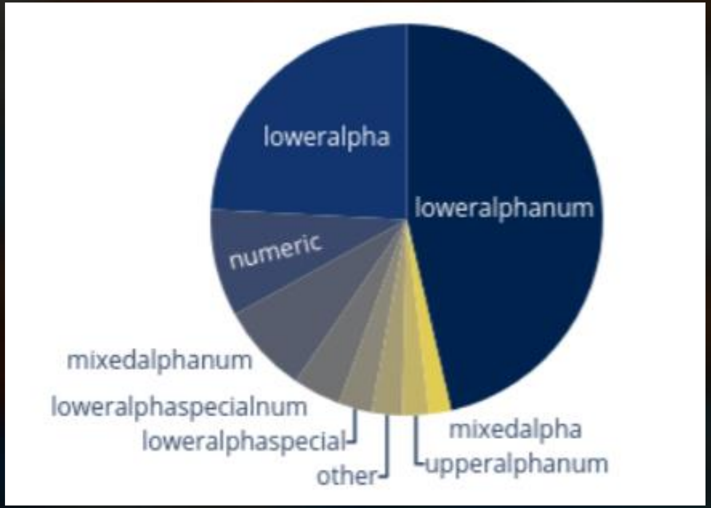


Password Strength (zxcvbn)

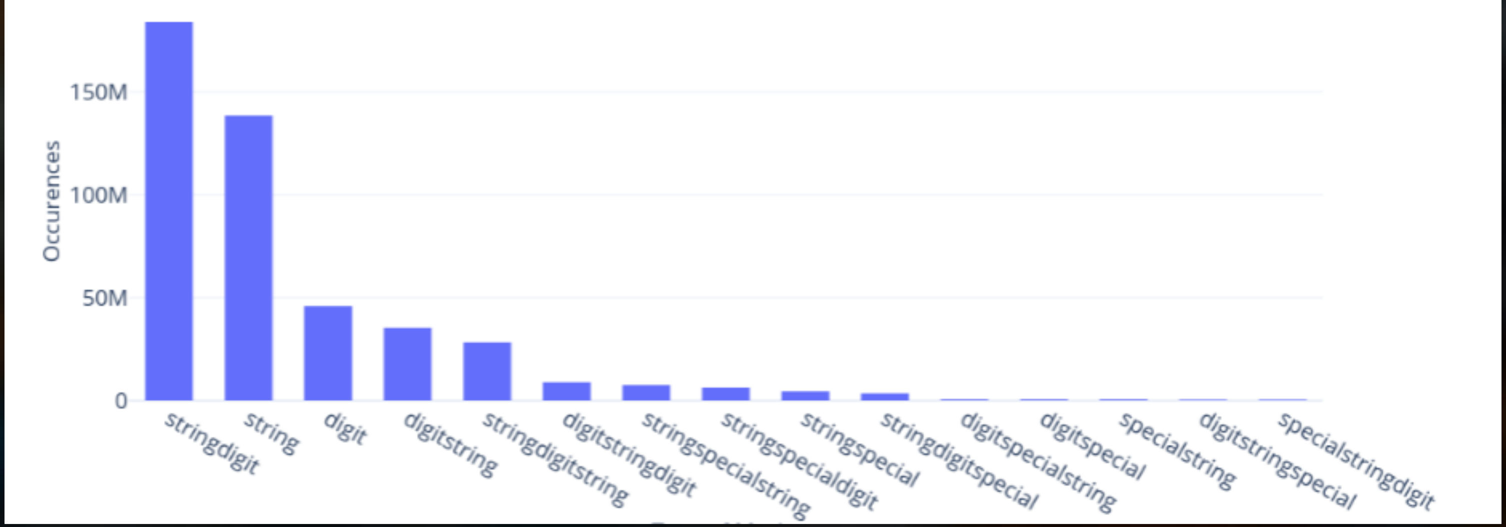
Score	Percentage
0	0.04%
1	14.7%
2	47.3%
3	26%
4	12%

Example: 15 digits passwords - 11% of passwords in class 4
MD5: ~12 hours
BCRYPT: ~650 years
considering a 3080 NVIDIA GPU

Have I Been Pwned Analysis 3



Types of masks in HIBP



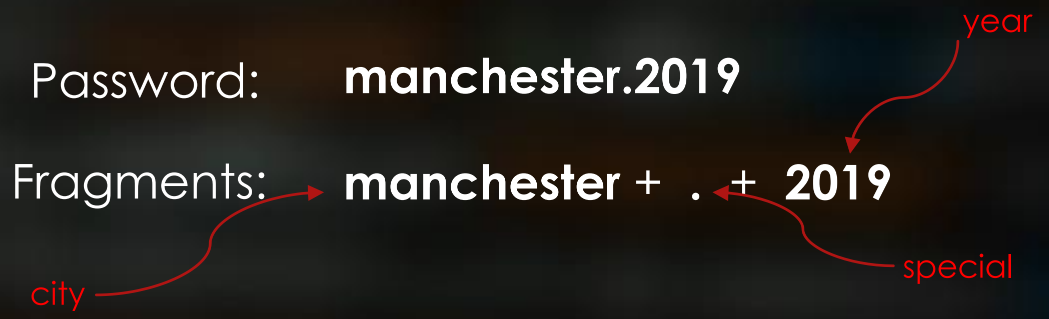
Examples:
Lower alpha: password
Mixedalphanum:passWoRD12
Upperalphaspecialnum: PASSWORD1!
All: !pa$$1worD

Character categories

- String
  - Lower alpha
  - Upper alpha
  - Mixed alpha
- Numeric
- Special

ex. Digitstringdigit: 123p123 or 1PASSWORd1

Have I Been Pwned Advanced Analysis 1



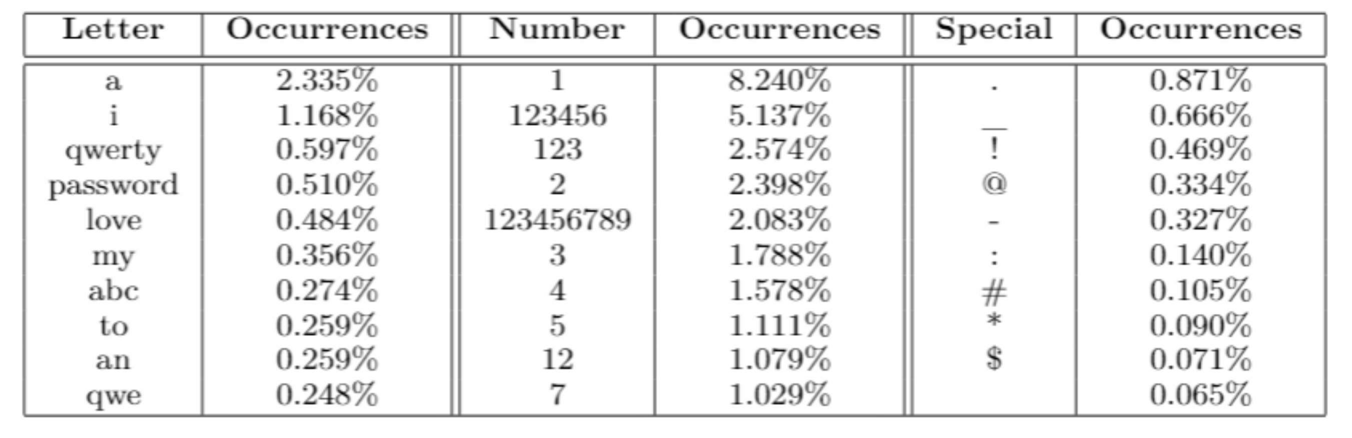
Password: **manchester.2019**

Fragments: **manchester** + **.** + **2019**

Number of fragments per category

letters	1,074,196,225
numbers	439,727,373
special	61,336,778
total	**1,575,290,376**

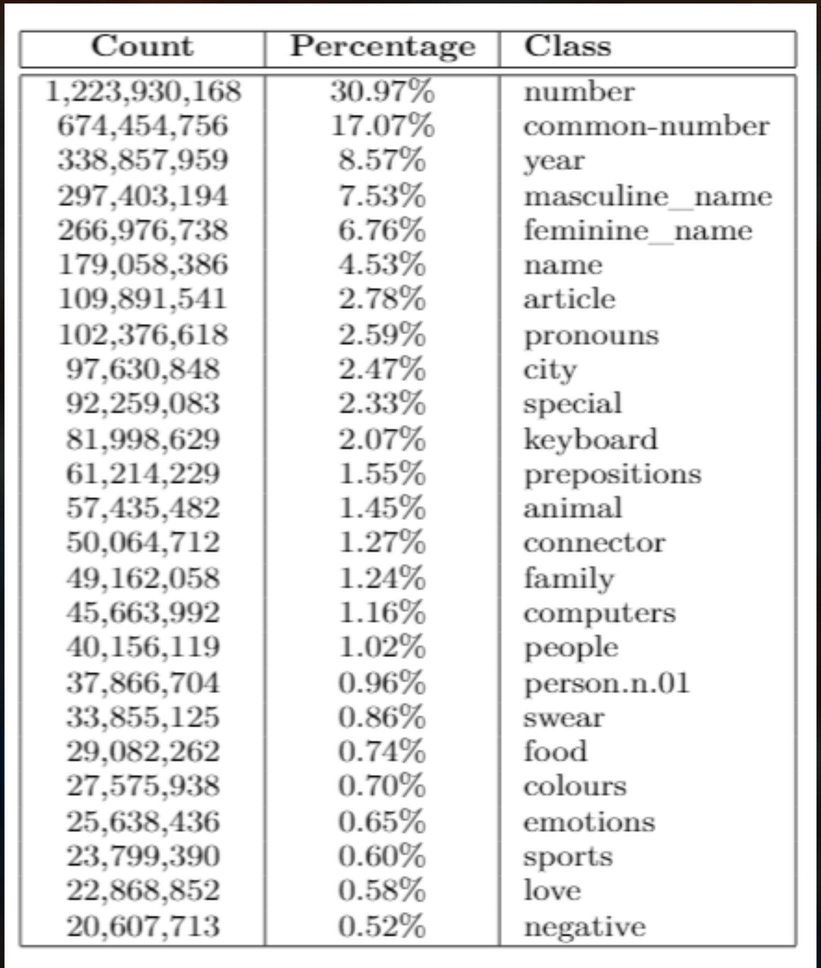
10 most popular fragments per category



Letter	Occurrences	Number	Occurrences	Special	Occurrences
a	2.335%	1	8.240%	.	0.871%
i	1.168%	123456	5.137%	_	0.666%
qwerty	0.597%	123	2.574%	!	0.469%
password	0.510%	2	2.398%	@	0.334%
love	0.484%	123456789	2.083%	-	0.327%
my	0.356%	3	1.788%	:	0.140%
abc	0.274%	4	1.578%	#	0.105%
to	0.259%	5	1.111%	*	0.090%
an	0.259%	12	1.079%	$	0.071%
qwe	0.248%	7	1.029%		0.065%

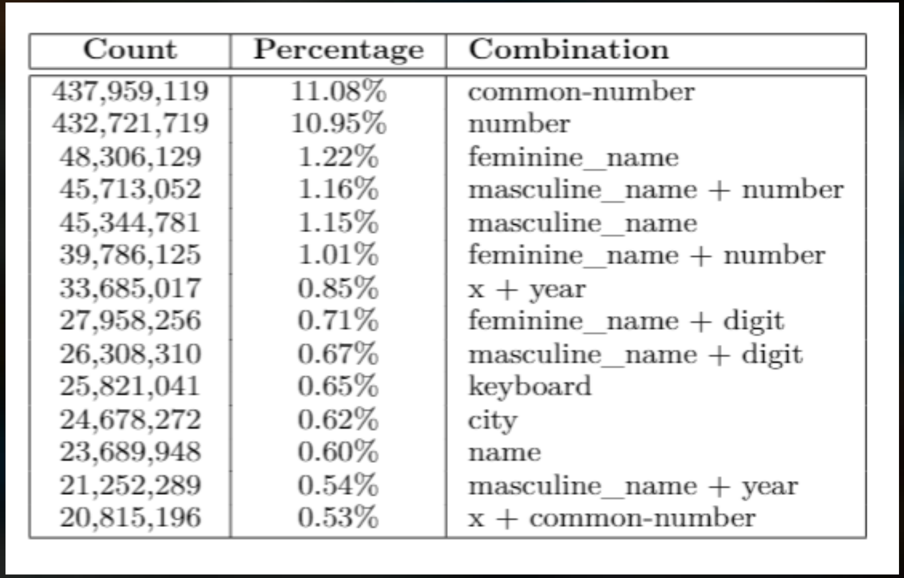
Have I Been Pwned Advanced Analysis 2

Most common fragment categories



Count	Percentage	Class
1,223,930,168	30.97%	number
674,454,756	17.07%	common-number
338,857,959	8.57%	year
297,403,194	7.53%	masculine_name
266,976,738	6.76%	feminine_name
179,058,386	4.53%	name
109,891,541	2.78%	article
102,376,618	2.59%	pronouns
97,630,848	2.47%	city
92,259,083	2.33%	special
81,998,629	2.07%	keyboard
61,214,229	1.55%	prepositions
57,435,482	1.45%	animal
50,064,712	1.27%	connector
49,162,058	1.24%	family
45,663,992	1.16%	computers
40,156,119	1.02%	people
37,866,704	0.96%	person.n.01
33,855,125	0.86%	swear
29,082,262	0.74%	food
27,575,938	0.70%	colours
25,638,436	0.65%	emotions
23,799,390	0.60%	sports
22,868,852	0.58%	love
20,607,713	0.52%	negative

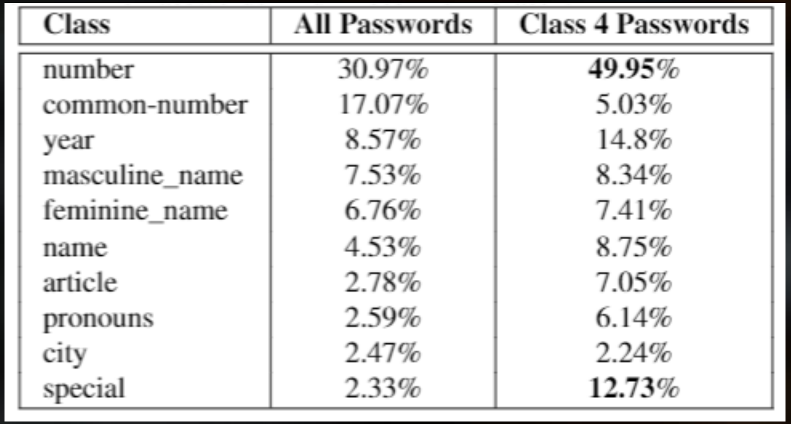
Most frequent fragment combinations



Count	Percentage	Combination
437,959,119	11.08%	common-number
432,721,719	10.95%	number
48,306,129	1.22%	feminine_name
45,713,052	1.16%	masculine_name + number
45,344,781	1.15%	masculine_name
39,786,125	1.01%	feminine_name + number
33,685,017	0.85%	x + year
27,958,256	0.71%	feminine_name + digit
26,308,310	0.67%	masculine_name + digit
25,821,041	0.65%	keyboard
24,678,272	0.62%	city
23,689,948	0.60%	name
21,252,289	0.54%	masculine_name + year
20,815,196	0.53%	x + common-number

Comparison of most frequent fragment categories between all passwords and class 4

HIBP (average): 2.1 fragments per password
Class 4: 4.4 fragments per password



Class	**All Passwords**	**Class 4 Passwords**
number	30.97%	**49.95%**
common-number	17.07%	5.03%
year	8.57%	14.8%
masculine_name	7.53%	8.34%
feminine_name	6.76%	7.41%
name	4.53%	8.75%
article	2.78%	7.05%
pronouns	2.59%	6.14%
city	2.47%	2.24%
special	2.33%	**12.73%**

Context in Passwords

Demographic

- Male/female
- English/non-English speaking
- Age range
- Profession

Personal Information

- Male/female names
- Birthdates
- City names
- Pet names

User Interests

- Animals
- Food
- Swear words
- Emotions
- IT knowledge

Leveraging context

Digital Life of suspect

Local Device Information

Previous Passwords

Online Presence

PASSWORD
password
Pass12word
PASSWORD12
pasSWord12
PASS12word
PASSWORD!
pa$$word

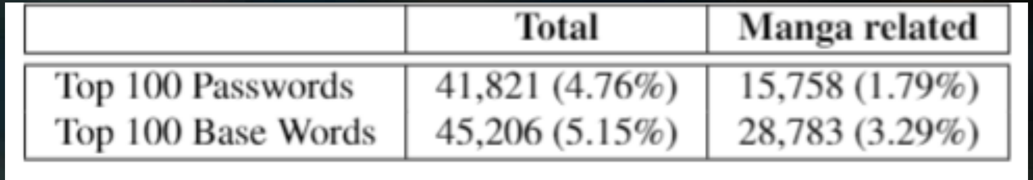
Mangling Rules

Smarter Dictionary List

A preliminary Analysis

Password leak: mangatraders.com
881,468 entries or
618,237 unique passwords

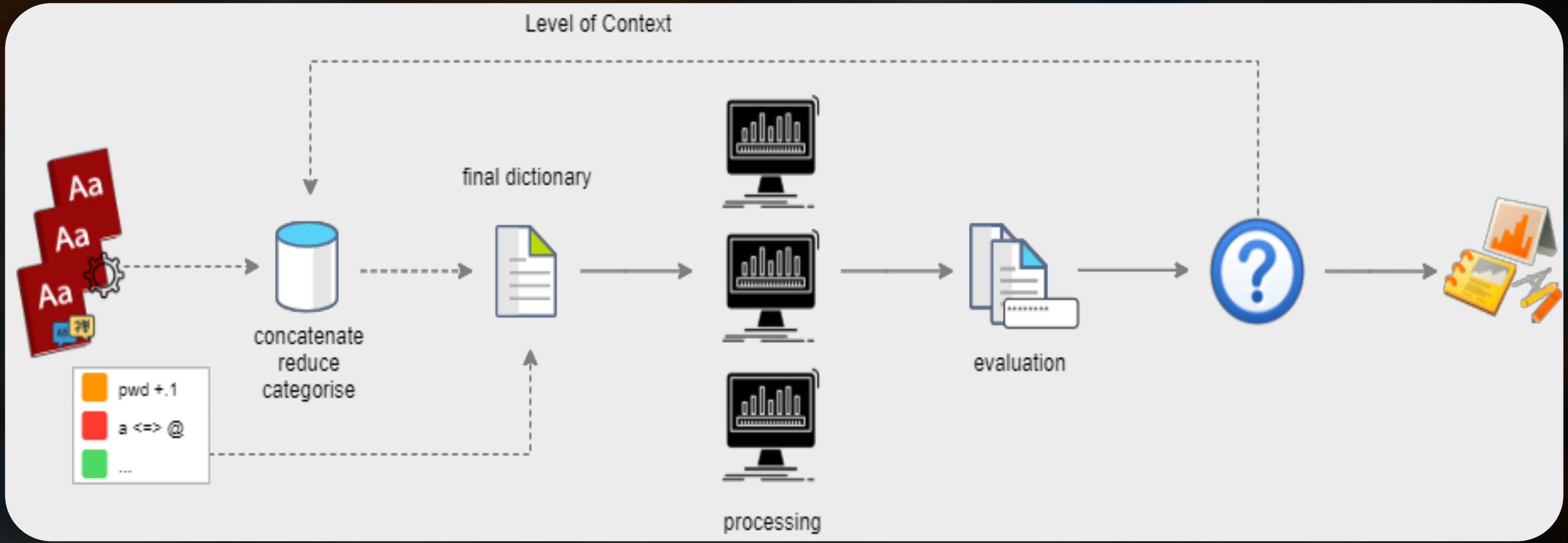
Manga Related Passwords in mangatraders.com



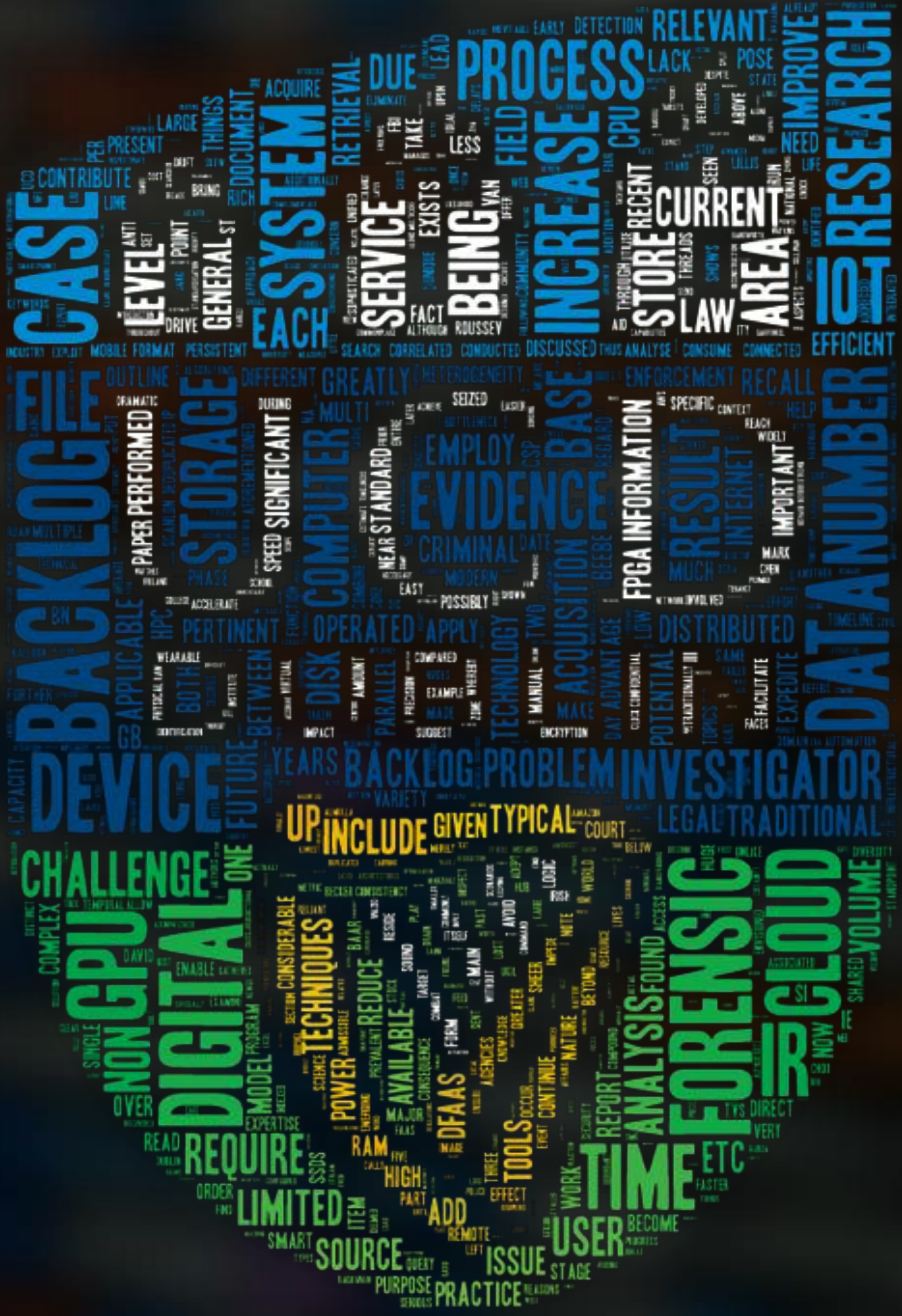
	Total	Manga related
Top 100 Passwords	41,821 (4.76%)	15,758 (1.79%)
Top 100 Base Words	45,206 (5.15%)	28,783 (3.29%)

Future Work

- PCWQ: A Framework for Evaluating Password Cracking Wordlist Quality



- Creation of custom dictionary lists



✉ Aikaterini.Kanta@ucdconnect.ie

Aikaterini.Kanta@ec.Europa.eu

www.ForensicsAndSecurity.com

@ForSecResearch

Joint Research Centre
JRC

UCD Forensics and
Security Research Group