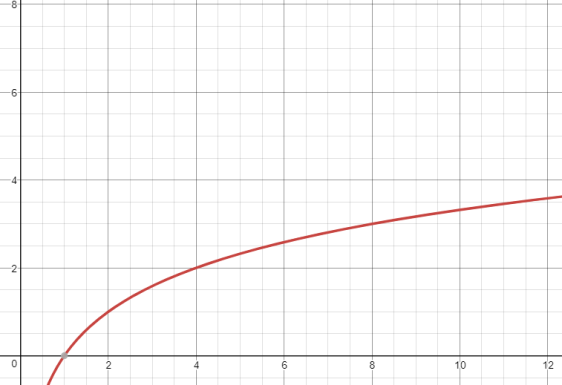
Password Security, Dictionary Attacks, and Entropy:

Password security is becoming an increasingly prevalent factor in modern society. Malicious attackers seeking to exploit vulnerabilities in security systems will target the weakest area; this is most often the user. Taking advantage of poorly designed and common passwords, cracking has evolved from straight forward brute force enumeration to targeted attacks utilising dictionaries and compromised user account databases.

Security vulnerabilities arise when passwords are stored in plaintext format. If a database is breached, the data has no inherent protection and can be readily exploited. As outlined in the GDPR, companies storing sensitive information have a responsibility to ensure its protection. Resultantly, passwords should have their cryptographic hashes computed using a SHA (Secure Hash Algorithm) and stored in place of the password. Hashes have an appearance of randomness and are computed using a linear function, leaving them infeasible to invert **(Dr M. Pound, 2017)**. Despite this, it is still crucial to use a strong password; attackers have been known to precompute the hashes of common keys, usually with the aid of ‘Rainbow Tables’ **(Digital Guide, 2018).**

With reference to dictionary attacks, many user-created passwords are inadequately designed and are therefore attacks are designed with human predictability in mind. Current generation GPU architectures such as Pascal (NVIDIA) and Polaris (AMD) suited to multi-core arithmetic pipelining (extreme parallelism) have enabled highly efficient cracking - further emphasising the importance of creating resistant passwords **(O Afonin, 2016).**

*Logarithmic relationship between entropy and character set size:*

*Password Entropy*

The strength of a password is a function of its length and complexity **(US-CERT, 2009)**, resulting in a measure of entropy (a sense of its difficulty to crack). Higher entropy generally translates to a stronger password in real terms, with the bits of entropy increasing as a logarithmic function of the size of the password’s character set (x) and linearly with length (n), given by the formula: **(M Rouse, 2014).**

*Size of character set*

Many users create passwords consisting of long strings of related words, as their length proves as an effective measure against simplistic brute force attempts. Character based entropy does not provide a good indication of a password’s resilience against dictionary attacks. We can model passwords containing known words as being composed of a very large character set, with each known word in a dictionary corresponding to a character (an example of this is the DiceWare word list). There are many more words than characters in the English language; but a logarithmic function of a character set yields diminishing returns as the set increases. As a result, against dictionary based attacks, a short password containing seemingly random characters can prove more effective than a long password of common words **(RIT, 2011)**. Theoretically, this would be the best solution; however, the difficulty to remember a password increases with its complexity. To solve this problem, a user could install and make use of a reputable password manager and generator. Managers usually only require the memorisation of a single ‘master password’ to decrypt a local vault. Generation is also handled by the manager itself, removing the need for the user to create their own and potentially flawed passwords. For this reason, I believe that it not only the safest solution, but also a relatively user friendly answer to our problem.

Behavioural Tracking and Data Controllers:

The gathering and processing of users’ online data has become a lucrative business over the past decade. With the rise of Big Data, mining to search for trends and patterns in aggregated user data allows for targeted marketing and predictive analytics. This task is divided between data controllers (those who determine the purpose for which data is collected) and data processors (those who process the data on behalf of the controller). In many cases, there is a large network of controllers and processors, and since data is shared, each node in this network poses its own security risk. As a result, a data breach can compromise user information stored in all affiliated companies across said network.

Recent legislation introduced in the EU’s GDPR requires data controllers to be transparent with their customers with respect to sharing data amongst third party processors **(Art. 12 GDPR).** Websites now need to ask for permission to utilise otherwise non-essential cookies for advertising and tracking purposes; pushing a notification that is clearly laid out and giving an option to decline **(Cookiebot, 2018)**. Advertising and marketing services in particular collate data from vast numbers of users, with the intent to construct profiles based on their interests and characteristics (such as age and ethnicity) **(C. Casteluccia, 2012)**. It is not feasible for an individual to keep track of all the companies holding their behavioural data, meaning that if any are compromised, a leak is unlikely to be brought to the user’s attention.

By working alongside online advertising agencies, the behavioural tracking industry generates an estimated $31 billion per annum **(C Sorensen, 2012)**. Through the eyes of some, these are ill-gotten gains. The industry also poses the argument that users receive more relevant adverts tailored to their viewing habits. There is a strong incentive for ad networks to place more relevant banners to maximise profit, due to the fact that the network collects a small payment from the corresponding advertiser for each click. Conversely, online anonymity advocates warn that tracking takes place out of the view of the public and is carried out with insufficient regulation. With collected data being used for corporate gain, it creates severe privacy concerns by allowing a few institutions to gather and concentrate a large amount of information about internet users. In my opinion, the worst case scenario is to transition into a surveillance Internet, where our online and physical activities are collected and correlated together.

Ultimately, one must take a certain degree of responsibility for their own online privacy. You should strike a balance between a more tailored user experience and your own online anonymity. Personally, I make use of Ghostery, a tracker blocker browser extension to prevent invasive cookies tracing my online activity. Additionally, I make use of the ad-block Plus extension, which not only prevents obtrusive advertising, but also provides a supplementary layer of privacy against data profiling due to their ‘Acceptable Ads without third-party tracking’ feature, encouraging a culture of responsible web-based marketing.

To conclude, I believe that we will only see real change in privacy rights with additional legislation brought in to oversee and regulate the personal data and analytics industry.

Jordan Taylor

Bibliography:

Article title: Choosing and Protecting Passwords

Website title: Us-cert.gov

URL: <https://www.us-cert.gov/ncas/tips/ST04-002>

---

Article title: Password Entropy

Website title: whatis.techtarget.com

Author: Margaret Rouse

URL: <https://whatis.techtarget.com/definition/password-entropy>

---

Article title: How to calculate password entropy.

Website title: RIT Fundamentals of Computer Security Class Blog

Author: Anonymous / alias: mlb9252

URL: <https://ritcyberselfdefense.wordpress.com/2011/09/24/how-to-calculate-password-entropy/>

---

Article title: NVIDIA Pascal: a Great Password Cracking Tool

Website title: ElcomSoft

Author: Oleg Afonin

URL: <https://blog.elcomsoft.com/2016/07/nvidia-pascal-a-great-password-cracking-tool/>

---

Article title: Weak passwords banned in California from 2020

Website title: BBC News Technology

Author: Anonymous

URL: <https://www.bbc.co.uk/news/technology-45757528>

---

Article title: SHA: Secure Hashing Algorithm

Website title: YouTube

Author: Computerphile – Dr M. Pound

URL: <https://www.youtube.com/watch?v=DMtFhACPnTY>

---

Article title: Rainbow Tables

Website title: Digital Guide - IONOS

Author: Anonymous

URL: <https://www.ionos.com/digitalguide/server/security/rainbow-tables/>

---

Article title: Cookie Consent

Website title: Cookiebot

Author: Anonymous

URL: <https://www.cookiebot.com/en/cookie-consent/>

---

Article title: Behavioural Tracking: You’re being stalked across the web

Website title: Macleans

Author: Chris Sorensen

URL: <https://www.macleans.ca/economy/business/behavioural-tracking-youre-being-stalked-across-the-web/>

---

Article title: Behavioural Tracking on the Internet: A Technical Perspective

Author: Claude Casteluccia

URL: <https://www.researchgate.net/publication/287252970_Behavioural_Tracking_on_the_Internet_A_Technical_Perspective>