

Independent Project Proposal

by Nino Aguilar

Submission date: 20-Oct-2025 10:13AM (UTC+0100)

Submission ID: 265716823

File name: 49114_Nino_Aguilar_Independent_Project_Proposal_695044_180413121.docx (127.28K)

Word count: 2110

Character count: 12533



INDEPENDENT PROJECT PROPOSAL

Assessing Strengths of Password: A Python-Based
Simulation of Password Cracking Technique

Nino Aguilar (2305326)
2305326@student.uwtsd.ac.uk

Table of Contents

Introduction:	2
Problem Statement:	2
Aim:	2
Objectives:	3
Expected Outcomes:.....	3
Literature Review:	4
Project Design:	5
Workflow:	5
Data Collection:.....	6
Ethical, Legal and Limitation Issues:	6
Resources and Planning:	7
Project Timeline:	8
REFERENCES	Error! Bookmark not defined.

Introduction:

Cybersecurity is essential for the protection of systems, online services and data especially in this day and age where communication, financial transactions, and data storage are being heavily relied upon online. This rapidly growing digital economy has many upsides, one of which is convenience and accessibility, as the digital era has made almost anything accessible from almost anywhere whether that's enabling financial systems to process more efficiently or accessing services and products a lot quicker than ever before [1]. However, with the positives also comes the negatives such as an increased rate of cyber threats and attacks as well as data privacy being breached which can cost the global economy up to \$10.5 trillion annually by the year 2025 [2]. This is due to cybercriminals exploiting security systems that are not strong enough to ensure the protection of their data. Password-based authentication is one of the most used methods to authorize and confirm a person's identity. However, it is also the most common vulnerability as over 74% of data breaches relate to weak and unsecure passwords and the lack of good habits when it comes to password security such as reusing passwords across multiple accounts [3].

Problem Statement:

Although it is publicly known that cyber-attacks can happen to anyone, it is still a shame that a large number of people still practice bad habits of reusing the same password across various services as well as not using strong enough passwords to enhance security and safety. This reasoning could be due to the fact that users place convenience over security. This issue seems to stem, not from users knowing that threats can happen because they already are aware, but from users just not understanding how easy it truly can be to compromise a weak and unsecure password and how quickly it can happen to them. This shows that there is a gap between the awareness and understanding of password security.

Aim:

Therefore, the aim of this project is to bridge that gap by developing a Python-based simulation that helps people truly visualize how weak passwords lead to vulnerability, spreading practical insights on why and how people should prioritize security over convenience.

Objectives:

1. Research password security standards, policies and cracking methods to explore vulnerabilities and typical behaviors of users that lead to the use of weak passwords.
2. Design and develop a Python-based simulation to perform brute force and dictionary attacks on passwords to showcase the differences between password strengths.
3. Evaluate numerical data such as time and number of attempts needed to compromise passwords of various lengths and complexity.
4. Analyse and interpret the results providing recommendations to promote awareness and understanding of password security and why it's important.

Expected Outcomes:

1. Python-based simulation designed to crack passwords.
2. Dataset and visual graphs showcasing results tested on the relationship between password complexity and cracking time.
3. Report on the findings and recommendations to improve best practices for password security.

Literature Review:

“Memorized secrets” commonly known as passwords or PINS, are the most commonly used authentication method alongside other growing methods such as multi-factor authentication. NIST discusses the importance of the complexity and secrecy that a memorized secret should have as it can aid in preventing passwords from being compromised [4]. Although adding complexity can help, there are also weaknesses within that aspect too, especially in user-generated password schemes as they seem to encourage bad habits when it comes to password security such as using the same passwords across multiple services and using predictable password strategies such as character substitution e.g. using a 4 instead of an A or a 0 instead of an o [5]. This demonstrates the vulnerability of human error within password security. Cybercriminals are aware of these common vulnerabilities and utilize this opportunity for an attack.

A common way cybercriminals compromise weak and unsecure passwords is through using automated techniques such as a brute-force attack where it uses an automated software to guess usernames and passwords through trial and error until the right username and passwords has been authenticated. Another method would be dictionary attacks; this is where software, again, is used to try a list of common words, phrases and combinations to gain unauthorized access. These password attacks being automated means their success of unauthorized access is dependent on factors which can be measurable such as password entropy, length and the reuse of credentials. For example, a low entropy and shorter password which makes it easier to guess decreases the time and effort for an attacker to compromise a password and vice versa [6]. This allows for an appropriate choice of attack methods to analyze the time and effort to crack passwords for this project.

Furthermore, there are also multiple studies that imply that human behavior has become the main weakness of password security. A survey by Google & Ipsos clearly shows that 65% of Americans reuse the same passwords for different accounts online and a staggering 52% include personal info such as names and birthdate information into their passwords. Another interesting statistic taken from that survey reports that one in three people have shared their password information with others or even have access to someone else's passwords, which are not their own [7]. NordPass also reports that easy passwords to guess just like “123456” have been ranked the most used but also one of the quickest passwords to crack within a matter of under 1 second [8]. This further suggests

that bad habits within password security stems from the human nature to choose convenience over security which is why it is important to visually educate users on how and why to improve password security habits.

It's clear that the research shows that everyone uses passwords daily and that it is the primary form of authentication, however, due to human behavior, leading to predictable outcomes, this becomes an extreme vulnerability for the user and an easy catch for cybercriminals. Although governments and organizations such as NCSC and NIST actively spread awareness and run campaigns to promote practicing good habits, showcasing easy and straightforward tips to improve password security [9], this still remains to be a problem and will continue to carry out if nothing concrete is being done to solve this. It appears that awareness is being spread through generalized recommendations and tips; however, it seems as though it is not enough. Therefore, this project aims to go that extra mile to create a Python-based simulation, to not just give recommendations but to also visually display the ease or difficulty it truly is to crack passwords, motivating users to change and improve their habits when it comes to password security.

Project Design:

This project is designed to create an experimental and simulation environment to analyze how password length, complexity, and entropy can affect how easy or how hard it is for a password to be compromised. The two attack methods I will be using for the demonstration will be the brute-force attack, and the dictionary attack and will be incorporated into Python to ensure it is a well-controlled and safe environment to carry out these attacks without having to expose real user data.

Workflow:

- Create a dataset of passwords in different strength groups containing (weak, moderate, strong).
- Hash each password to simulate real authentication processes.
- Run a brute-force attack that generates every possible password combination.
- Run a dictionary attack using a list of common passwords and phrases.
- Record data such as cracking time, number of attempts, and password length.

- Store results in CSV format using Pandas to analyze results.
- Visualize the findings with Matplotlib to compare time-to-crack across password categories.

Data Collection:

- This project involves no human participants; all data will be generated synthetically.
- The Python-based simulation automatically collects key data points such as cracking time, number of attempts made, length of password and the group complexity from each password test.
- This data is then collected to analyze the vulnerability of the different types of passwords.
- This data will be used to calculate the mean, median, and variance of the password cracking performance to analyze the results.
- The Matplotlib library will be used to create graphs and charts to give us a visual representation of the final analyzation.

Ethical, Legal and Limitation Issues:

- This project will be conducted in a controlled and offline environment to ensure that no real systems / personal info is at risk.
- Only fake passwords will be used, excluding personal data / leaked password lists.
- This project will comply with the ethical and legal standards set by the UK cyber security law by making sure that there is no unauthorized access or data misuse involved.
- Tools and scripts used in this project are only for educational purposes only and will also add safety measures to prevent any malicious use of intent.
- Hardware performance – as personal computers may not have the same processing power as professional hacking tools, shorter password lengths will be used to keep this project more practical.

Resources and Planning:

To carry out this project effectively, the following resources will be required:

- **Hardware:**
 - Personal laptop or desktop computer with sufficient processing power to run Python-based simulations.
 - Stable internet connection
- **Software:**
 - Python (latest version)
 - Libraries: hashlib, bcrypt, itertools, string, pandas, and matplotlib for simulation, hashing, data handling, and visualization.
 - IDE: Pycharm
 - Microsoft Office Suite — used for documentation, data organization, and report writing.
- **Data Storage and Management:**
 - OneDrive — for secure file storage and version control, in line with the UWTSd Research Data Management Policy.
 - Local encrypted storage — as a secondary backup for project files and datasets.

Project Timeline:

Phase / Date Range	Key Tasks & Activities	Expected Outcome
21 Oct – 3 Nov 2025	Review supervisor feedback on proposal. Finalize project scope and objectives. Set up Python environment in PyCharm and install required libraries (hashlib, bcrypt, itertools, string, pandas, matplotlib).	Environment ready and project plan finalized.
4 – 24 Nov 2025	Conduct in-depth literature review on password security, authentication standards (NIST/NCSC), hashing algorithms, and password attack techniques.	Completed research foundation for report background section.
25 Nov – 15 Dec 2025	Develop Python modules for password generation and hashing. Test simple password sets and log initial results.	Basic password creation and encryption modules functional.
16 Dec 2025 – 5 Jan 2026	Implement brute-force and dictionary attack simulations. Begin collecting early test data for short passwords.	Working attack simulation with initial sample results.
6 – 26 Jan 2026	Run structured experiments testing different password lengths, complexities, and hash types. Collect cracking times and performance metrics.	Full dataset collected and ready for analysis.
27 Jan – 16 Feb 2026	Analyse dataset using pandas to calculate mean, median, and variance. Use matplotlib to visualize findings through graphs and charts.	Analytical results and visuals are completed.
17 Feb – 9 Mar 2026	Write and refine sections of report covering methodology, design, and analysis. Begin discussion and conclusion drafts.	Draft report with integrated data and analysis.

10 – 30 Mar 2026	Proofread and edit report. Cross-check IEEE references and ensure formatting matches university standards.	The final report version is ready for supervisor review.
31 Mar – 20 Apr 2026	Prepare presentation materials, demo video (if applicable), and final poster visuals to summarize findings.	All presentation and poster assets are completed.
21 Apr – May 2026	Final review of all materials and submission preparations.	Project fully completed and ready for assessment.

References

- [1] McKinsey, “How COVID-19 has pushed companies over the technology tipping point—and transformed business forever,” *McKinsey & Company*, 2021.
<https://www.mckinsey.com/~media/McKinsey/Business%20Functions/Strategy%20and%20Corporate%20Finance/Our%20Insights/How%20COVID%2019%20has%20pushed%20companies%20over%20the%20technology%20tipping%20point%20and%20transformed%20business%20forever/How-COVID-19-has-pushed-companies-over-the%20technology%20tipping-point-final.pdf> (accessed Oct. 15, 2025).
- [2] S. Morgan, “Cybercrime To Cost The World \$10.5 Trillion Annually By 2025,” *Cybercrime Magazine*, Nov. 10, 2020. <https://cybersecurityventures.com/cybercrime-damage-costs-10-trillion-by-2025/> (accessed Oct. 15, 2025).
- [3] Ipsos, “PUBLIC POLL FINDINGS AND METHODOLOGY,” May 2022. Accessed: Oct. 15, 2025. [Online]. Available:
https://www.ipsos.com/sites/default/files/ct/news/documents/2022-05/Google%20Password%20Topline_5.5.2022.pdf
- [4] P. A. Grassi *et al.*, “Digital identity guidelines: authentication and lifecycle management,” *NIST Special Publication 800-63B*, Jun. 2017, doi:
<https://doi.org/10.6028/nist.sp.800-63b>.
- [5] “Password Guidance.” Accessed: Oct. 16, 2025. [Online]. Available:
https://assets.publishing.service.gov.uk/media/5a806bb9e5274a2e87db9b6a/Password_guidance_-_simplifying_your_approach.pdf
- [6] I. Šlekýtė, “Password entropy and how to calculate it | NordVPN,” *nordvpn.com*, Aug. 18, 2023. <https://nordvpn.com/blog/what-is-password-entropy/> (accessed Oct. 17, 2025).
- [7] “PUBLIC POLL FINDINGS AND METHODOLOGY A majority of Americans are concerned about the safety and privacy of their personal data,” May 2022. Accessed: Oct. 17, 2025. [Online]. Available: https://www.ipsos.com/sites/default/files/ct/news/documents/2022-05/Google%20Password%20Topline_5.5.2022.pdf
- [8] NordPass, “Most Common Passwords of 2020,” *nordpass.com*, 2024.
<https://nordpass.com/most-common-passwords-list/> (accessed Oct. 17, 2025).

[9] National Cyber Security Centre, “Cyber Aware,” www.ncsc.gov.uk, 2024.
<https://www.ncsc.gov.uk/cyberaware/home> (accessed Oct. 17, 2025).

Independent Project Proposal

GRADEMARK REPORT

FINAL GRADE

GENERAL COMMENTS

67 /100

PROJECT PROPOSAL – ACCA6020 INDEPENDENT PROJECT

**Student Nino
Aguilar
Mark:
67**

Intro / A&O (400 words)		
Project title; Identification of a specific problem to be investigated; Well-defined aims and SMART objectives; Explaining why this project is worth doing; Discussion of what the project needs to do in order to be a success; Description any deliverables that will be produced at the end of the project	20	15
Problem identified. Aim need to be improved, as it is not about developing the simulation tool but more about using python-based simulation.		
Objectives are clear but can do with		

refinement.		
Background to the Project / Literature Review (400 words)		
Description of how the work to be done by in the project relates to the work identified in the literature review; Discussion of any ethical issues raised by the project Identifying what has already been done within the field; Topic familiarity; Brief literature review using appropriate sources – 6-8 journal papers and book chapters rather than just websites; Selection of relevant models & concepts Lit review can be enhanced. I t would be good to use graphical representation to discuss the statistical trend from the survey outcomes. All website used, try to also use peer reviewed articles.	20	12
Project Design (500 words)		
Clear statement of which development methods are to be used and why they are appropriate; Clear statement of which tools are to be used and why they are appropriate; Clear description of how data collected will	30	20

be analysed This need to be more granular; a block diagram would enhance the design method. Tools to use stated. Data collection listed.		
Resources / Planning (200 words)		
List of any deliverables; Clear diagram showing both timing and resource use; Completed ethics form; Limited and relevant appendices This done satisfactory Ethics form completed.	15	10
Document Presentation		
Correct spelling, grammar and punctuation; Flow of narrative/fluency; Coherence; Word processed ; Correct	15	10

use of referencing with bibliography; Strict word limit		
Could do with formatting alignment. Introduction could do with paragraphing, etc.		
TOTAL	100	67

PAGE 1
PAGE 2
PAGE 3
PAGE 4
PAGE 5
PAGE 6
PAGE 7
PAGE 8
PAGE 9
PAGE 10
PAGE 11
PAGE 12