Mini Project Report on

Study of Password Generator and Password Recommender System

Submitted in partial fulfilment of the requirement for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE & ENGINEERING

Submitted by:

Student Name:	**University Roll No.:**
Kartickey Aggarwal	**2018867**

Under the Mentorship of

Dr. Jay Bhatnagar
Associate Professor

Department of Computer Science and Engineering
Graphic Era (Deemed to be University)
Dehradun, Uttarakhand

July-2023

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the project report entitled **"Study of Password Generator and Password Recommender System"** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering of the Graphic Era (Deemed to be University), Dehradun shall be carried out by the under the mentorship of **Dr. Jay Bhatnagar, Associate Professor**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Student Name:	University Roll No.:
Kartickey Aggarwal	2018867

Table of Contents

Chapter 1

Introduction

1.1 Introduction

Passwords are essential for everyone's online security, especially in the covid era when e-commerce has grown rapidly and exposed users to more risks. However, many people choose passwords that are easy to remember but also easy to guess by hackers. There are various types of cyber-attacks that can break passwords, such as brute-force, dictionary attack, man in the middle, etc. Therefore, password strength is an important factor to consider when creating or selecting a password. Password strength refers to how hard it is to crack a password. A common way to enforce password strength is to check it in real-time when users create or change their passwords. This practice was recommended by password security researchers since the early 1990's.

Although there are many websites that can generate strong passwords for users, they may not be very memorable or user-friendly. There has been a lot of research on password generation and password strength estimation, and there are different methods to measure the strength of a password. The strength of a password is usually estimated by simulating how an attacker would try to break it.

Some of the methods are: -

- Attack Based Approach: - This approach uses information about the user and the most common passwords to try to guess the password. It usually uses a brute force strategy. One way to objectively quantify the strength of a password is to attack it until it is broken. Most of the existing works that use this approach are based on a single-attack strategy [1].

- Heuristic Based Approach: - This is one of the most popular methods to crack passwords. It uses a method called LUDS [counts of lower-case upper-case digits and symbols]. Most of the password policies in practice are based on this concept [2].
- In this project, we tested passwords against multiple approaches and methods to crack them, while also checking their usability for the users. The ultimate goal of a password is to protect the user's personal data while being easy to use. The paper will discuss the methodology used in generating and estimating passwords and how it ensures that they are user-friendly and hard to crack [3].

Chapter 2

Literature Survey

1.2 Surveys:

[1] A New Multimodal Approach for Password Strength Estimation. Part I: Theory and Algorithms

—After more than two decades of research in the field of password strength estimation, one clear conclusion may be drawn: No password strength metric by itself is better than all other metrics for every possible password. Building upon this certainty and also taking advantage of the knowledge gained in the area of information fusion, in the present work we propose a novel multimodal strength metric that combines several imperfect individual metrics to benefit from their strong points in order to overcome many of their weaknesses. The final multimodal metric comprises different modules based both on heuristics and statistics which, after their fusion, succeed to provide in real time a realistic and reliable feedback regarding the "guessability" of passwords. The validation protocol and the test results are presented and discussed in a companion paper.

[2] Designing password policies for strength and usability

Password-composition policies are the result of service providers becoming increasingly concerned about the security of online accounts. These policies restrict the space of user-created passwords to preclude easily guessed passwords and thus make passwords more difficult for attackers to guess. However, many users struggle to create and recall their passwords under strict password-composition policies, for example, ones that require passwords to have at least eight characters with multiple character classes and a dictionary check. Recent research showed that a promising alternative was to focus policy requirements

on password length instead of on complexity. In this work, we examine 15 password policies, many focusing on length requirements. In doing so, we contribute the first thorough examination of policies requiring longer passwords. We conducted two online studies with over 20,000 participants, and collected both usability and password-strength data. Our findings indicate that password strength and password usability are not necessarily inversely correlated: policies that lead to stronger passwords do not always reduce usability. We identify policies that are both more usable and more secure than commonly used policies that emphasize complexity rather than length requirements. We also provide practical recommendations for service providers who want their users to have strong yet usable passwords.

[3] Encouraging users to improve password security and memorability.

Security issues in text-based password authentication are rarely caused by technical issues, but rather by the limitations of human memory, and human perceptions together with their consequential responses. This study introduces a new user-friendly guideline approach to password creation, including persuasive messages that motivate and influence users to select more secure and memorable text passwords without overburdening their memory. From a broad understanding of human factors-caused security problems, we offer a reliable solution by encouraging users to create their own formula to compose passwords. A study has been conducted to evaluate the efficiency of the proposed password guidelines. Its results suggest that the password creation methods and persuasive message provided to users convinced them to create cryptographically strong and memorable passwords. Participants were divided into two groups in the study. The participants in the experimental group who were given several password creation methods along with a persuasive message created more secure and memorable passwords than the participants in the

control group who were asked to comply with the usual strict password creation rules. The study also suggests that our password creation methods are much more efficient than strict password policy rules. The security and usability evaluation of the proposed password guideline showed that simple improvements such as adding persuasive text to the usual password guidelines consisting of several password restriction rules make significant changes to the strength and memorability of passwords. The proposed password guidelines are a low-cost solution to the problem of improving the security and usability of text-based passwords.

Chapter 3

Methodology

The algorithm generates passwords based on the information provided by the user like their first name, last name, date of birth, and father's name. After the generation each password is tested by the algorithm that how strong each password is, and a score is provided to the password according to its strength tested by the algorithm.

- First the algorithm checks the length of the password, whether it's 6 characters long or more than that.
- Then the algorithm checks if the password contains any commonly used words that are generally used by the majority of users while creating a password and a score is provided according to the test.
- Then the algorithm checks whether the password contains the information of the user without encrypting it in some way or how much the password is predictable based on the information provided by the user.
- Then the password is checked against brute-force attacks i.e., how much time would it take to crack the password.
- After the brute-force attacks, the randomness of the password is checked, checking how predictable can the password be based on how many times a character is repeating in the password.
- Each step provides a relative score to the password and based on the score the password is rated as Easy, Moderate or Hard.

Survey

Survey Design:

The survey was designed with two identical forms, each consisting of 30 questions. The questions in both forms included passwords, and the participants were asked to select one of the three options (Easy, Moderate, or Hard) based on the perceived difficulty of each password. The forms were distributed to a diverse group of participants.

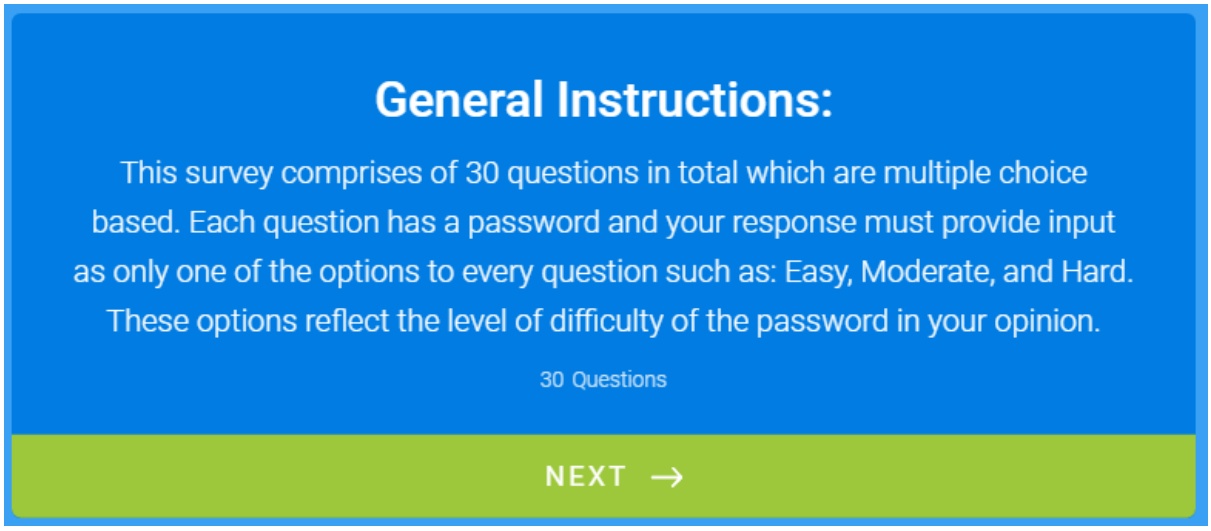

Fig. 3.1: General Instructions window

Tabular Form of both forms:

Password	Easy	Hard
32sId19		
9109ksD~		
aG~1>K1u		
gAu1805		
03+03gJk		
RhS@+26a		
96sUm96		
27saR]96		
9}NA9hk(		
1999hRi		
[0909Chh		
089!;hhC		
aAr2697		
'01naN01		
{9ir1~NN		
20kUn05		
23@00nCk		
]Cnk%523		
1997dEv		
S$}S61vt		
s-S'0326		
nIk3333		
VN02:53h		
3)lN/3iS		
99aMi24		
03Si{k@N		
3390"Svn		
aKa1902		
02aTa-03		
k~]An3T9		
0799aVi		
19,vaB19		

@BVhj9{4		
20aNk01		
01)21Psa		
.h9V9vA>		
28aRj95		
28:19Mah		
[2M7*aHr		
1919rOh		
aNr04.26		
>RNn]6d9		
vIk2220		
20-20Kdv		
‚KoD8(2a		
0306aBh		
11#Cka20		
!akC~111		
06aRj19		
04aDa06@		
2-D2]Vak		
05aMi05		
Pga0117		
aTa%A!19		
1920aDi		
04&!2aaT		
1920*aaT		
aRj0919		
17+Kl@yA		
Ska~2626		

Table 3.1: Passwords table as were in the forms

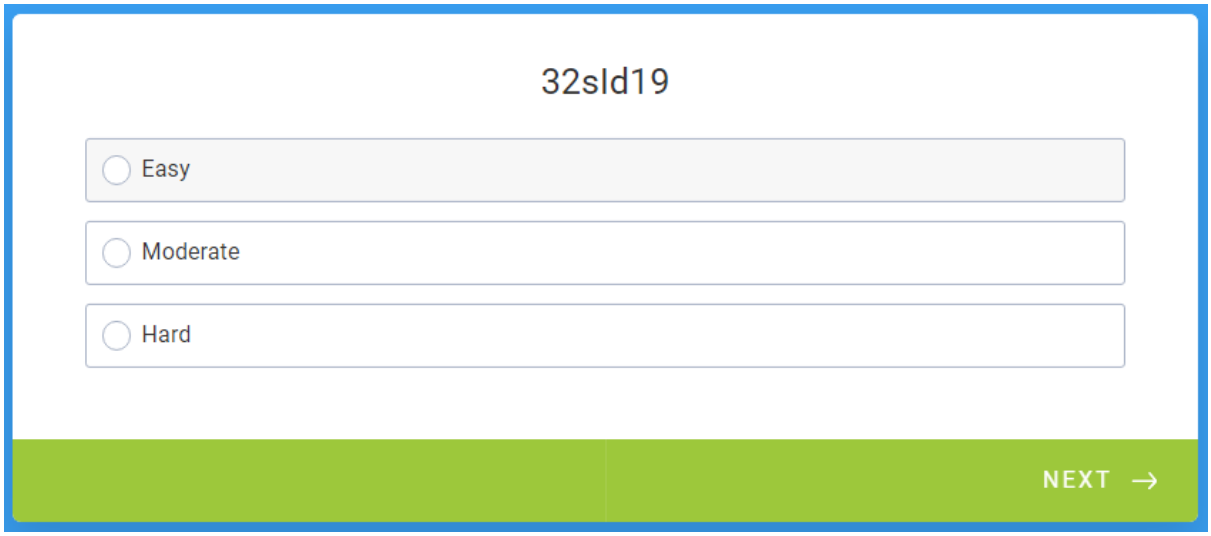

Fig. 3.2: Block Diagram of the process

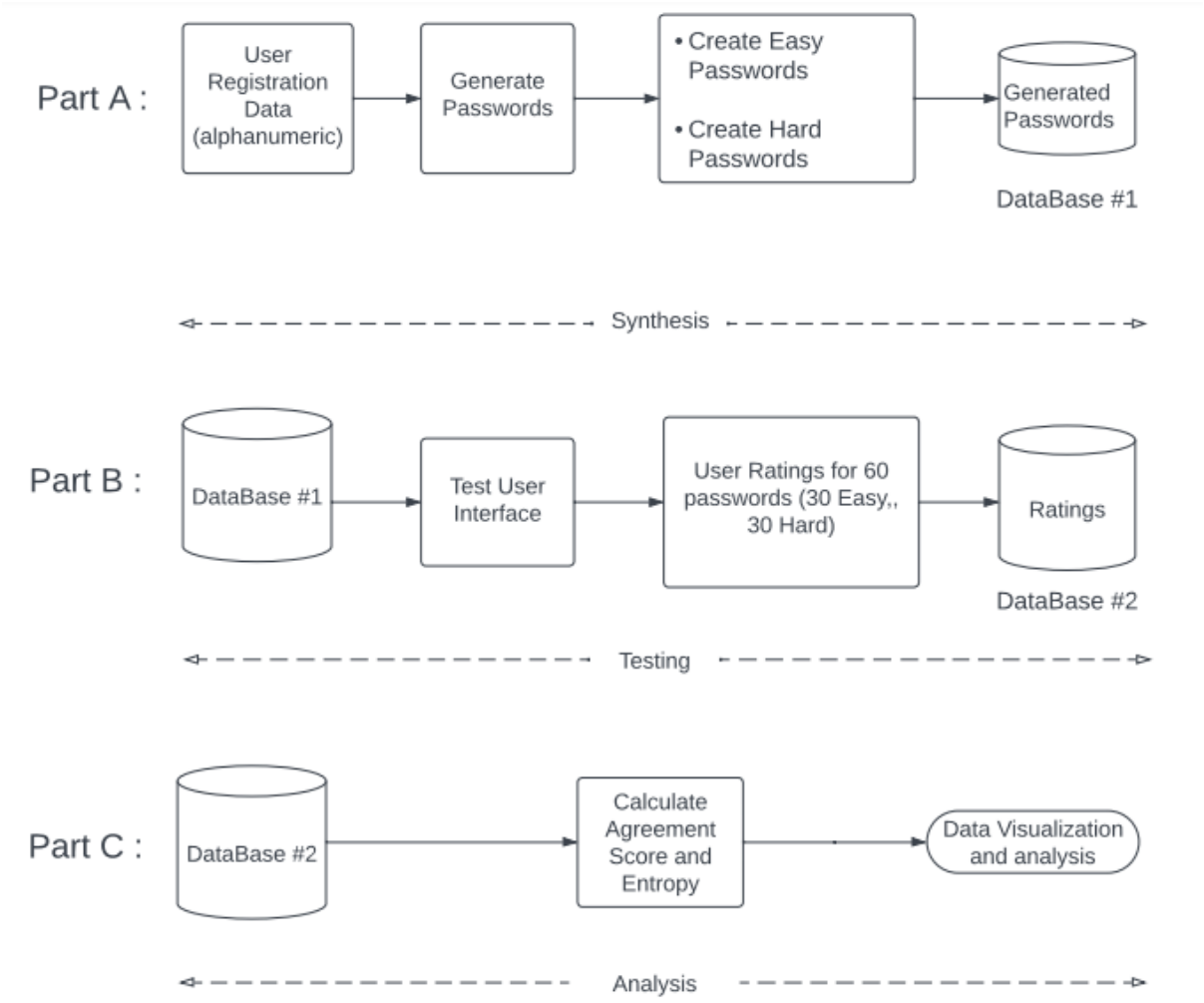

Fig. 3.3: Block Diagram of the process

Chapter 4

Result and Discussion

Graphs:

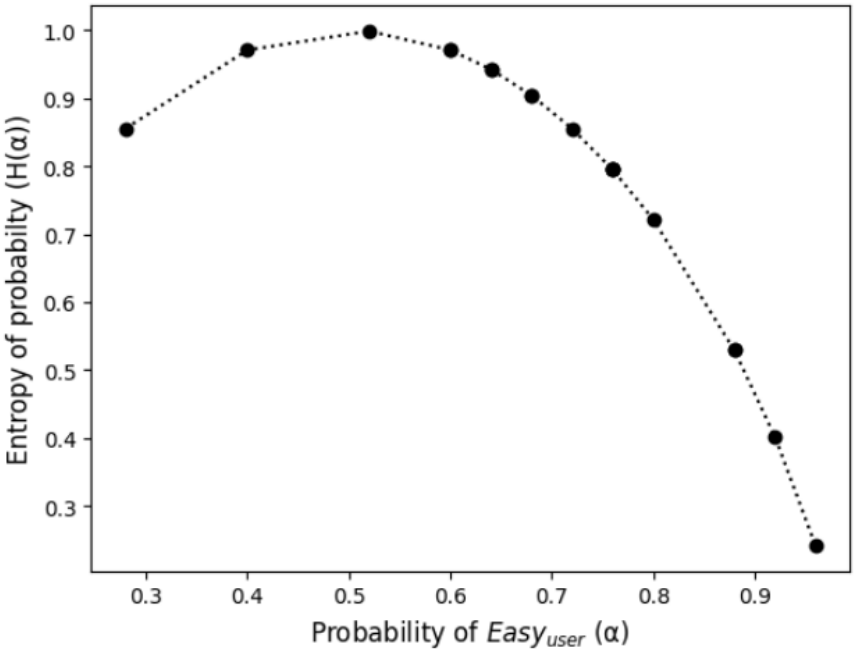

Fig. 4.1: Entropy-probability curve of 'Easy' passwords

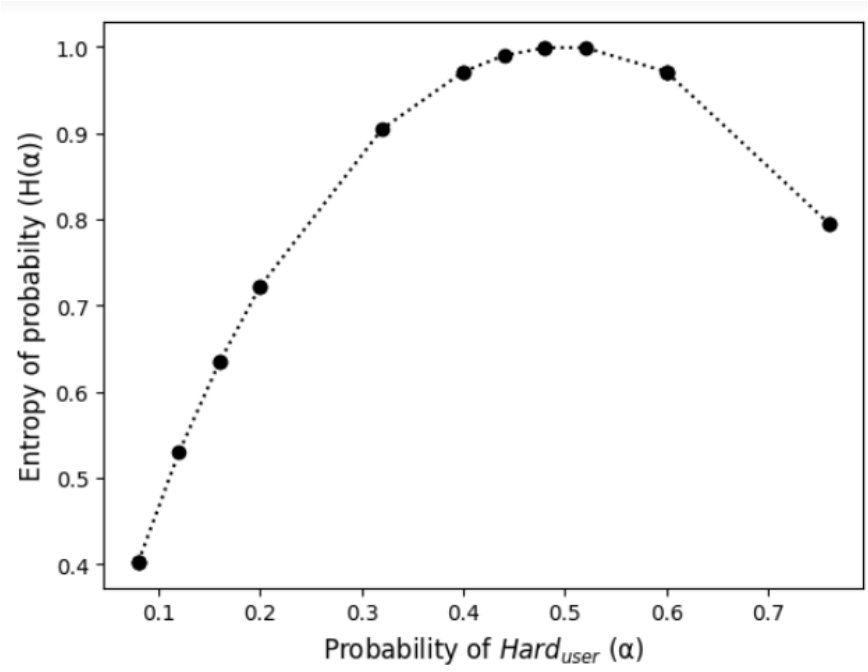

Fig. 4.2: Entropy-probability curve of ‘Hard’ passwords

Each point on curve represent a password, there are 20 such points (some are overlapping due to exact same values).

X-axis: α-probability of agreement score, i.e. how many users agree with program generated easy password,

Y-axis: H(α)-Entropy of probability

Formulae:

$$\alpha = \frac{\text{n(count of users which agree with program)}}{25\text{(total number of ratings per password)}}$$

$$H(\alpha) = -\alpha * \log_2 \alpha - (1 - \alpha) * \log_2 (1 - \alpha)$$

Venn-Diagrams:

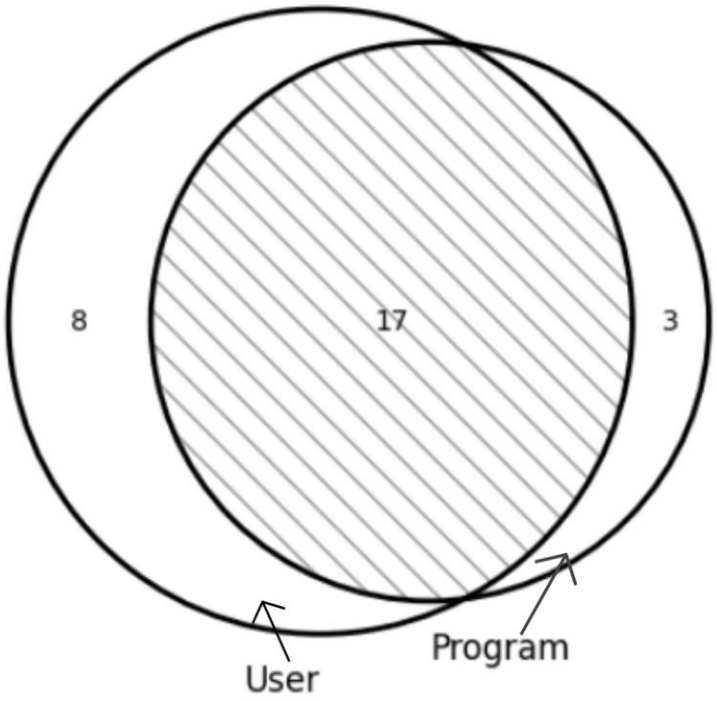

Fig. 4.3: Easy passwords venn-diagram

This shows how many easy passwords generated by program coincide with user marked easy passwords where,

n_u (user) = 25

n_p (program) = 20

n (user ∩ program) = 17

where, n stands for count.

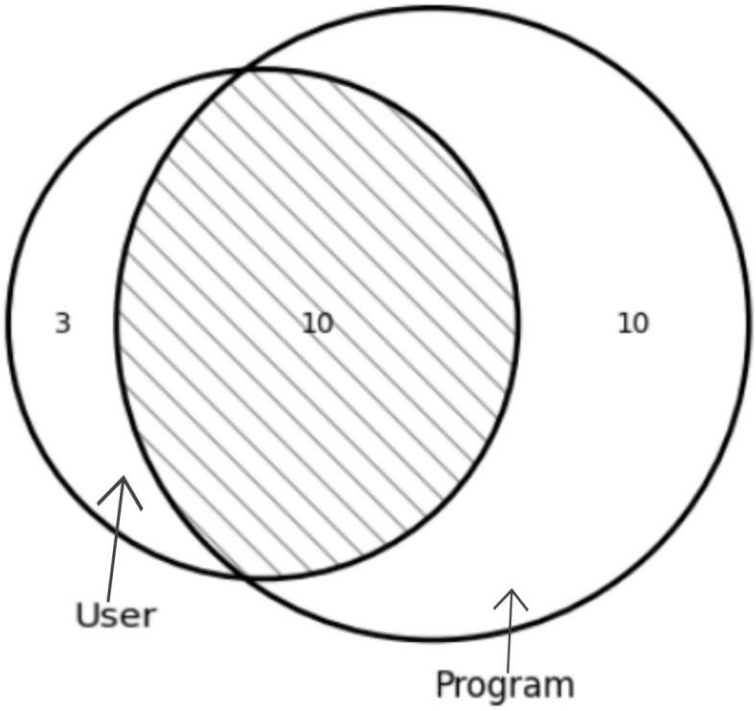

Fig. 4.4: Hard passwords venn-diagram

This shows how many hard passwords generated by program coincide with user marked hard passwords where,

n_u (user) = 13

n_p (program) = 20

n (user ∩ program) = 10

where, n stands for count.

Above diagrams represent the data collected in visual formats.

Results and Interpretation:

The results obtained from the analysis of the survey data and the Venn diagrams provided insights into the participants' perception of password difficulty compared to the machine model's classification. The overlapping area in the Venn diagrams represented the passwords where both the participants and the machine model agreed on the difficulty level. The non-overlapping areas indicated passwords where there were discrepancies in the classifications.

Implications and Recommendations:

The findings of the survey can have implications for password security practices and user education. If the participants' classifications align with the machine model's classification, it may indicate a reasonable consensus on password difficulty. However, if there are significant discrepancies, it suggests the need for further investigation and improvement in the machine model or user education regarding password security.

Limitations:

It is important to acknowledge certain limitations of the survey. The results may be influenced by the composition and characteristics of the participant sample. Additionally, the survey relied on subjective perceptions of password difficulty, which can vary among individuals. The machine model's classification may also have limitations, depending on the training data and algorithm used.

Chapter 5

Conclusion and Future Work

Conclusion:

The survey on password difficulty, conducted through two forms with 30 questions each, provided valuable data for comparing the classifications made by form fillers and a machine model. The analysis using Venn diagrams allowed for a clear visualization of the agreement and discrepancies between the two classifications. The survey findings can contribute to improving password security practices and enhancing user understanding of password difficulty.

The reverse Turing test is a promising new method for evaluating password recommender systems. It allows us to measure the system's ability to generate human-like passwords. Future Work will focus on refining the test and evaluating its effectiveness.

Future Work:

The algorithm to generate and recommend human like passwords can help improve security with helping to create passwords that are readily usable and are easy to remember. This can be used on the websites to improve their password recommendations while keeping security check. The current password recommendations provided by the websites are not user friendly and make passwords difficult to remember.

Personalization and User Preference: Allow users to provide preferences or requirements for their passwords, such as minimum length, character types to include/exclude, or specific words or phrases to incorporate. This customization can help users create passwords that align with their preferences and make them more memorable.

Adaptive Recommendations: Develop an adaptive system that learns from user behavior and preference overtime. This system can adjust its recommendations based on user feedback, password usage patterns, and evolving security best practices.

User Education and Awareness: provide educational resources and guidelines to user on creating strong yet memorable passwords. Help them understand the importance of password security and encourage the password hygiene practices, such as using unique passwords for different accounts and regularly updating passwords.

Usability Testing and User Feedback: Conduct usability testing and gather feedback from users to continuously improve the algorithm. Incorporate user insights and preferences into the recommendation system to ensure it meets the requirements of a wide range of users.

References: -

[1] J. Galbally, I. Coisel and I. Sanchez, "A New Multimodal Approach for Password Strength Estimation—Part I: Theory and Algorithms," in *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 12, pp. 2829-2844, Dec. 2017, doi: 10.1109/TIFS.2016.2636092.

[2] "Designing password policies for strength and usability," ACM Trans. on Infor. and Systems Security, vol. 18, pp. 13:1–13:34, 2016.

[3] Encouraging users to improve password security and memorability | SpringerLink