Leveraging large language model to protect against targeted password guessing

USP 2025 GR 1

Table of contents

01 Introduction

02 Research Question

03 Results

04 Data Collection

05 System Architecture

06 Conclusion

Introduction

Context: Increasing target password-guessing sophistication.

Problem: Traditional passwords are vulnerable to targeted guessing attacks due to similarities among various passwords for the same user

Objective: Explore if LLM-generated passwords are more resilient and memorable than user-created ones.

Research Questions

1. **Guessability of LLM-generated vs. User-generated passwords**: Can LLMs reduce guessability by avoiding common patterns?

2. **Impact of Prior Passwords**: Does prior password knowledge influence LLM-generated password guessability?

Study Design

- **Data Collection**: Realistic dataset to simulate potential attacker knowledge (4.2 billion entries from leaked dataset, used 1000 for our study).

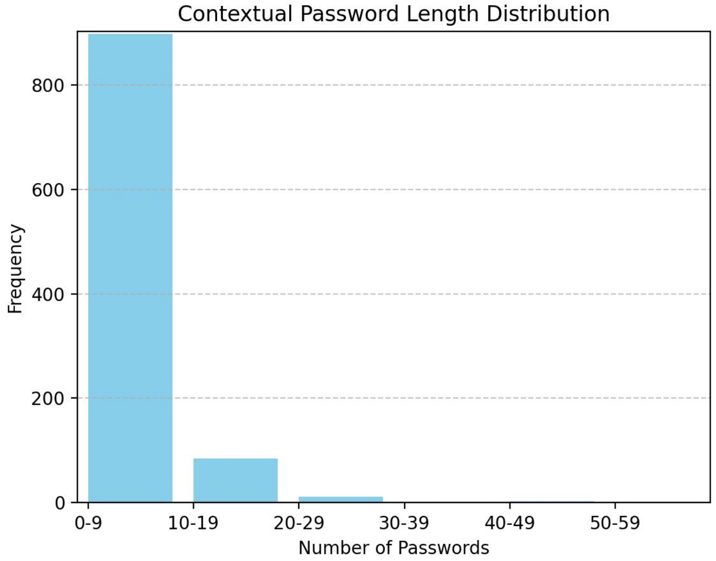
- **Password Generation**: Generate LLM-based passwords (used **llama3.2:3b**).

- **Guessability Analysis**: Use **Targuess**, to predict guessability of llm-generated password

Data Collection Architecture

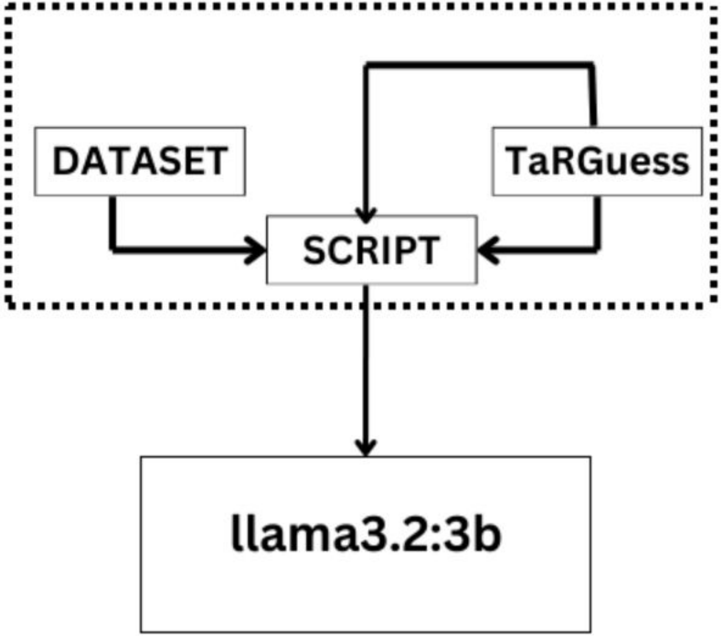
The dataset used in this research consists of 4.2 billion context-based passwords, collected from publicly available resources and password leak repositories. These sources include open datasets, password leaks, and breach disclosures that were accessed through ethical and legal means.

We used 1000 entries for our study.



System Architecture

- **DataSet**: Used by the main program
- **llama3.2:3b**: Deployed on server, used to generate context based password
- **TarGuess**: Guesses the password depending on the context (datapoint entry)
- **SCRIPT**: Controls the flow



Results

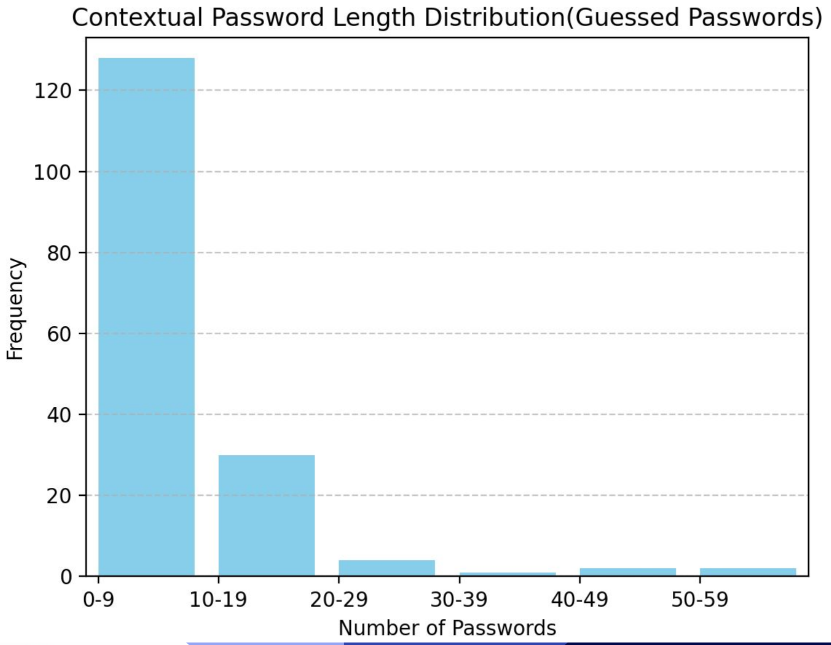
Out of the 1000 passwords generated by LLama 3.2:3b, 187 passwords were successfully guessed by Targuess. This performance is summarized as follows:

Passwords Generated = 1000

Passwords Guessed by Targuess = 187

Guessing Accuracy = 187

1000 × 100 = 18.7%



Conclusion

In this study, we explored the use of LLama 3.2:3b for generating passwords and the effectiveness of Targuess in guessing them. The results demonstrated that while LLama 3.2:3b can generate a significant number of passwords, the guessability of these passwords remains a concern, as a significant fraction (187 out of 1000) were guessed by Targuess.

This indicates that we need to improve the llm model to generate better passwords that are easy to remember and tough to guess.

Thank You!