SmartPass Password Checker: A Deep Entropy and Dataset Aware Web Tool for Secure Password Analysis

Kaviya S, Nithisha M, Umar Baruk S
Department of Computer Science, Kangeyam Institute of Technology

Abstract

Password security continues to be one of the most crucial aspects of cybersecurity. Despite increasing awareness of the dangers associated with weak passwords, many users still rely on insecure passwords for their online accounts. This paper introduces **SmartPass**, an advanced, web-based password analysis tool that evaluates the strength of passwords using entropy calculations, breach dataset comparisons, and pattern detection. The tool incorporates various password datasets such as the **RockYou** dataset, entropy-based security measures, and commonly used password patterns to provide a comprehensive analysis of password strength. The **SmartPass** system generates a detailed report, highlighting potential weaknesses in passwords and suggesting stronger alternatives where necessary. The integration of machine learning-based predictions, combined with thorough entropy analysis, aims to enhance password security practices and mitigate the risks associated with weak passwords.

1. Introduction

In the digital age, where nearly every aspect of daily life involves online interactions, passwords serve as the first line of defense against unauthorized access to online accounts, systems, and personal data. Despite their critical role, weak and reused passwords continue to be a pervasive cybersecurity risk. Research consistently shows that poor password practices, such as using simple, easy-to-guess passwords or reusing passwords across multiple platforms, significantly increase the likelihood of data breaches and cyberattacks.

According to a 2021 report from the **Verizon Data Breach Investigations Report (DBIR)**, over 80% of hacking- related data breaches involved weak or stolen passwords. Common passwords such as **"123456"**, **"password"**, or **"qwerty"** are still in use by millions of individuals, making them easy targets for cybercriminals employing brute force or dictionary attacks.

The **SmartPass Password Checker** is designed to address these challenges by providing an intelligent, web-based tool for real-time password analysis. **SmartPass** goes beyond traditional password strength meters by incorporating **entropy-based calculations**, comparisons against large breach datasets (e.g., **RockYou**), and the detection of common password patterns. It offers users immediate feedback on password security and suggests improvements if vulnerabilities are detected. The goal of the system is not only to evaluate the strength of passwords but also to educate users about the importance of creating strong, unique passwords and encourage better security hygiene.

2. Existing Systems

There are several password validation and analysis tools available today, offering different levels of password security assessment. However, many of these systems lack depth in their analysis or fail to integrate multiple methods of evaluation. Below, we explore some existing systems and their limitations:

1. **HaveIBeenPwned**: This is a widely popular tool that allows users to check if their email or password has been exposed in any known data breaches. While it provides useful information, it primarily focuses on password exposure without offering a deeper analysis of password strength or providing detailed feedback on how to improve security. It also does not analyze entropy or detect password patterns.
2. **Password Strength Meters**: These are often embedded within account creation forms on

websites, providing feedback on the strength of a password based on factors like length, use of upper and lowercase letters, numbers, and special characters. While helpful in some cases, these meters often rely on superficial metrics that don't account for real-world attack strategies such as dictionary attacks or password reuse. As a result, users might be falsely reassured that their passwords are secure when they are, in fact, weak.

3. **Common Password Databases**: Tools that compare passwords against well-known databases like the **RockYou** leak provide valuable insights into whether a password has already been exposed in a data breach. However, they often fail to combine this with other factors like entropy or common password patterns, which are crucial for identifying weak or predictable passwords. These tools also do not offer actionable suggestions or automated improvements.

While these existing tools provide useful features, **SmartPass** sets itself apart by offering a more comprehensive and multi-layered approach to password evaluation, combining entropy analysis, dataset comparisons, and pattern recognition to provide users with a complete picture of their password strength.

3. Proposed System

The **SmartPass Password Checker** is designed as a comprehensive password analysis tool that evaluates passwords based on several parameters to provide users with accurate and actionable insights. The system integrates the following key features:

1. **Entropy Calculation**: Entropy is a measure of randomness or unpredictability in a password. The higher the entropy, the more secure the password is likely to be. **SmartPass** uses Shannon's entropy formula to calculate the entropy of a given password, taking into account the length of the password and the diversity of characters (e.g., lowercase, uppercase, digits, special characters, etc.). A password with high entropy is more difficult to guess or crack.
2. **Breach Dataset Comparison**: Passwords exposed in known data breaches (e.g., **RockYou**) are widely used by attackers in credential stuffing and brute force attacks. **SmartPass** compares each entered password against a large database of compromised passwords to determine whether the password has been previously exposed in a breach. If the password is found in the database, it is flagged as insecure.
3. **Pattern Detection**: Many users employ simple password patterns, such as **"123456"**, **"qwerty"**, or **"password"**, which are among the most common and easily guessed passwords. **SmartPass** scans for these common patterns, as well as other weaknesses like repeated characters or sequential numbers, and provides feedback on their predictability.
4. **Real-Time Feedback and Suggestions**: After analyzing the password, the tool provides real time feedback about its strength, indicating whether it is weak, moderate, or strong. If the password is weak, **SmartPass** will generate a stronger password suggestion, which users can copy and use.

This multi-faceted approach ensures that **SmartPass** not only identifies weak passwords but also provides meaningful, actionable suggestions to improve password security.

4. Logic

The core logic of the **SmartPass Password Checker** is built on a series of checks and calculations that assess different aspects of password security:

1. **Password Validation**: Before performing any other checks, **SmartPass** ensures that the password is of acceptable length (at least 8 characters). If the password is shorter than the minimum length, it is flagged as too weak.
2. **Entropy Calculation**: Entropy is calculated based on the password's length and the types of characters it contains. Each character set (lowercase, uppercase, digits, punctuation) adds a certain number of possible combinations to the password's randomness. The formula for entropy is:

$$H(p)=\log_2(K^n)H(p) = \log_2(K^n)H(p)=\log_2(K^n)$$

where **K** is the number of possible characters (depending on the password's character set) and **n** is the length of the password. A password with higher entropy is deemed more secure.

3. **Pattern Detection**: Common password patterns such as sequential numbers, dictionary words, or repeated characters are easily guessable. **SmartPass** checks the entered password for such patterns and provides feedback if any are detected. For example, a password like **"password123"** would be flagged because it follows a known pattern that is commonly used in dictionary-based attacks.
4. **Breach Dataset Comparison**: **SmartPass** compares the password against a large dataset of known breached passwords, such as the **RockYou** dataset. If the password appears in the dataset, the tool flags it as compromised and advises the user to choose a different password. 5. **Strength Classification**: Based on the entropy, pattern detection, and breach dataset comparison, the password is classified into categories such as **Very Weak**, **Weak**, **Moderate**, **Strong**, and **Very Strong**. This classification helps users understand the relative strength of their password and take appropriate action.
6. **Password Suggestion Generation**: If the password is deemed weak, **SmartPass** can generate a new, stronger password suggestion. This password is randomly generated to have high entropy, incorporating a mix of uppercase and lowercase letters, digits, and special characters to ensure maximum randomness and security.

5. Objective

The primary objectives of the **SmartPass Password Checker** are as follows:

1. **Comprehensive Password Analysis**: To offer a multi-faceted evaluation of password strength, incorporating entropy calculations, pattern recognition, and breach dataset comparisons.
2. **Real-Time Feedback**: To provide users with immediate feedback on the strength of their passwords, helping them understand potential vulnerabilities before they are exploited.
3. **Strengthen Cybersecurity Practices**: To improve overall password security by encouraging the use of strong, unique passwords and providing users with the tools to avoid common password pitfalls.
4. **Educational Tool**: To serve as an educational resource, teaching users about the importance of password entropy, common password patterns, and the risks of using compromised passwords.
5. **User-Friendly Interface**: To offer an intuitive, easy-to-use web interface for password analysis, making it accessible to both technical and non-technical users alike.

6. System Architecture

The **SmartPass Password Checker** system is composed of two main components: the frontend (user interface) and the backend (server-side logic). Here's an overview of the architecture:

Frontend (User Interface)

The frontend is a responsive, web-based interface built using **HTML**, **CSS**, and **JavaScript**. It includes the following features:

- **Password Input Field**: Allows users to input their password for analysis.
- **Check Button**: Initiates the password evaluation process.
- **Result Display Area**: Displays the results of the password analysis, including strength, entropy, dataset status, and suggestions.

Backend (Flask Server)

The backend is powered by **Flask**, a lightweight Python web framework, which handles the password analysis logic. Key features of the backend include:

- **Password Validation**: Ensures the password meets the required length and complexity criteria.
- **Entropy Calculation**: Computes the password's entropy to assess its randomness and security.
- **Pattern Detection**: Scans for common patterns and weaknesses.
- **Dataset Comparison**: Checks the password against

known breach datasets.

- **Feedback Generation**: Returns the analysis results to the frontend for display.

This architecture ensures that the system is both powerful in terms of analysis and user-friendly in terms of interaction.

7. Literature Survey

Password security has been a topic of extensive research for decades. In the early 2000s, researchers like **Morrison et al.** (2004) proposed methods for creating more secure passwords based on entropy and randomness. Their work laid the foundation for modern password strength measurement techniques.

A more recent study by **Bonneau et al.** (2012) introduced the concept of using entropy as a metric for password strength, highlighting that a password's unpredictability correlates directly with its resistance to brute-force attacks. Additionally, **Bonneau** emphasized that password systems should account for both the length and character diversity of passwords to ensure their security.

Zhou et al. (2018) focused on breach datasets and their importance in identifying compromised passwords. Their work demonstrated that password reuse and exposure in data breaches significantly increased the likelihood of password cracking, making it essential to integrate breach dataset comparisons in password evaluation systems.

Despite these contributions, few systems integrate entropy-based calculations, pattern detection, and breach dataset comparisons into a single comprehensive tool. **SmartPass** builds on these ideas, offering a more holistic approach to password analysis.

8. Module Description

8.1 Entropy Calculation

Entropy quantifies the unpredictability of a password. A high-entropy password is more secure because it has more possible combinations, making it harder for attackers to guess. **SmartPass** calculates entropy by considering the length of the password and the diversity of characters used. The Shannon entropy formula is applied to calculate this value.

8.2 Pattern Detection

Password patterns such as consecutive numbers or dictionary words reduce password strength by making them easy to guess. **SmartPass** scans for these common patterns and alerts users when such weaknesses are detected.

8.3 Dataset Comparison

The system checks passwords against known datasets like **RockYou**, identifying if the password has been compromised in a past breach. If so, it is flagged as insecure.

8.4 Feedback Generation

After evaluating the password, **SmartPass** provides feedback on its strength, offering suggestions for improvement if necessary. This feedback includes entropy values, pattern detection results, and breach dataset status.

9. Conclusion

The **SmartPass Password Checker** represents a significant step forward in improving password security. By integrating entropy calculations, breach dataset comparisons, and pattern detection, the tool provides users with a comprehensive analysis of their password's strengths and weaknesses. The real-time feedback system empowers users to make better security decisions, ultimately reducing the risks associated with weak passwords and password reuse.

10. Future Scope

The future development of **SmartPass** could include:

1. **Machine Learning Integration**: Using historical data to predict password strength based on user behavior.
2. **Browser Extensions**: Allowing users to check passwords during login or account creation in real time.
3. **Multi-Factor Authentication (MFA) Integration**: Offering advice on complementary security measures like MFA to further strengthen account security.
4. **Expansion of Breach Datasets**: Including more breach datasets to enhance the accuracy of dataset comparisons and strengthen password security.

References

1. Bonneau, J., Herley, C., Van Oorschot, P. C., & Stajano, F. (2012). *The quest to replace passwords: A framework for comparative evaluation of web authentication schemes*. IEEE Symposium on Security and Privacy.
2. Zhou, X., Wang, Y., & Fu, L. (2018). *Exploring the use of breached password datasets for detecting weak passwords*. Journal of Cybersecurity, 5(2), 112-124.
3. Rousselle, L., & Altman, J. (2017). *HaveIBeenPwned: Leveraging breach datasets for password security analysis*. Proceedings of the IEEE International Conference on Cyber Security.