Poster: The last step of password strength evaluation

Yan Shao
University of International Relations
y_shao@uir.edu.cn

Xin Xin
University of International Relations
grace_xin77@163.com

Hong Di*
University of International Relations
di_hong@163.com

Abstract—Passwords are typically the first line of protection in a security system, yet they are also the most vulnerable. The key to ensuring security is to use strong passwords. The security systems encourage users to create strong passwords through password strength estimation. There is no doubt that the existing password strength estimate methods are already very accurate for the identification of the weak passwords. However, among the passwords that meet their strong password condition, there are still some misjudged passwords. In this poster, we propose Character Distance Strong Password Checker (CDSPC). Specifically, Consecutive Lead Character Distance (CLCD) and Average Adjacent Character Distance (AACD) are used in CDSPC. CLCD is the sum of distance between all characters and the first character in the password. And AACD is the average value of every two adjacent characters distance in the password. The types of characters contained in the password and the lengths of password are recognized by CLCD, while the permutations of characters in the password are recognized by AACD. In the experiment, CDSPC was able to distinguish a misjudged strong password that had been evaluated as strong password using the password strength evaluation methods LPSE and ZXCVBN.

I. INTRODUCTION

Passwords are essential for safeguarding information property. Text password occupies an irreplaceable position in the Internet because of its convenience and low cost [1]. As the Internet grows in popularity, so does the number of weak passwords. There have been three primary ways to limit the spread of weak passwords in recent years. The first is for the system to create user passwords that are not only lengthy enough but also random enough [2], [3]. In general, users do not need to memorize such complicated passwords but rely on third-party password storage services. Users would face new authentication and trust issues as a result of a third-party involvement. The second type of password rules are set by the web administrator [4]. The length and character type of passwords are typically governed by the regulations, in order to register properly, the users must adhere to them. This coercive approach does improve password strength, but previous research indicated that users can only fulfill the regulations’ minimal requirements, thus it does not play a positive role. The third is to use a password strength evaluation method on the registration interface [5] that can provide real-time feedback on password strength in the form of text [strong, medium, weak] or a colored bar. Users can only register if their passwords satisfy the specified strength requirement. Previous researches [6] have shown that when users are aware of the strength of passwords, they will actively use various characters to construct stronger passwords.

The existing password strength evaluation methods will evaluate the passwords as unreasonable strength [7]. Obviously, judging weak passwords as strong is more dangerous. The Internet suffers greatly as a result of the incorrect classification of weak password intensity levels. For the same input password, different password strength evaluation methods will produce different output strengths, leaving users perplexed while creating passwords for several websites [8]. The existing password strength evaluation methods, according to Wang et al. [9], are ineffective in detecting weak passwords. Research by Ur et al. [10] showed that password strength may be significantly enhanced only if a password strength meter is provided with accurate strength.

In order to increase the accuracy of the password strength as much as possible, there are two password strength evaluation methods: one is ZXCVBN [11] which from industry, the other is LPSE [12] which from academia. Although the above-mentioned password strength evaluation methods have high accuracy, there are still some misjudgment in the strong passwords. We present Character Distance Strong Password Checker (CDSPC) to identify this part of misjudged passwords. Specifically, CDSPC calculates the character distance of the password to determine whether its strength is evaluated correctly. Two character distances are utilized in particular: Consecutive Lead Character Distance (CLCD) and Average Adjacent Character Distance(AACD). In the experiment of calculating the character distance of strong passwords, two password strength evaluation methods from different datasets are selected, the experimental results show that the proposed method is effective in identifying misjudged passwords.

II. METHODOLOGY

A. Character Distance

In order to distinguish the type of character effectively, we need to redesign the value of the character, which is called pivot value. As shown in Table.I, since the continuous length of lowercase and uppercase letters is 26, the minimum difference between \([\xi_1, \xi_2, \cdots, \xi_{26}]\) and \([\xi_1, \xi_2, \cdots, \xi_{26}]\) should be at least 26 in order to reflect the change of character types. Similarly, the minimum difference between \([s_1, s_2, \cdots, s_{10}]\) and \([\xi_1, \xi_2, \cdots, \xi_{26}]\) should also be at least 26. Obviously, the minimum difference between \([\xi_1, \xi_2, \cdots, \xi_{26}]\) and \(\eta\) should be at least the number of symbols.

* Hong Di is corresponding author.
With the pivot value of characters, for any two characters \( \alpha \) and \( \beta \), the character distance is renewed to:

\[
distance(\alpha, \beta) = |PV(\alpha) - PV(\beta)|,
\]

in which \( PV(\chi) \) represents the pivot value of character \( \chi \).

**B. Two Indicators Based on Character Distance**

The factors that affect the strength of a password are not only the characters which compose the password, but also the permutation of the characters. Specifically, for a password composed of multiple characters, it is necessary to consider the length of the password, the type of internal characters, the proportion of the same type characters, and the permutation of characters.

In order to take these factors into account, two indicators are proposed in this poster. One is CLCD, the other is AACD, the definition of them are as follows:

For CLCD, the password length with the same character type is calculated by the sum operation. For example, \( \text{Xbox360PS2021NS} \) and \( \text{Xbox360PS2021NSOLED} \) are composed of the same character type, the CLCD value of the former is significantly smaller than that of the latter. Another advantage of CLCD is that it is sensitive to the proportion of the same characters in passwords. For example, if there are lowercase \( \text{sourGRAP S} \) and \( \text{sourgraps} \), due to the large gaps between the pivot values of different character types, CLCD can calculate the proportion of lowercase letters and uppercase letters in the former, while knowing there are only lowercase letters in the latter.

For AACD, the former character type changes more often than the latter. Therefore, the AACD value of the former is higher than that of the latter.

**C. Checking Strong Password**

CLCD and AACD evaluate password strength from different perspectives. Therefore, to evaluate the strength of a given password accurately, it is necessary to calculate not only the two character distances, but also the threshold of two proposed indicators. It is worth mentioning that the threshold is different for different datasets. We analyze the distribution of characters in different datasets with different password strength evaluation methods to determine the threshold. Finally, whether the strength level of the strong password has changed is determined by judging the threshold of two character distances.

**ACKNOWLEDGMENTS**

This work was specially supported by National Natural Science Foundation of China (62102113), Fundamental Research Funds for the Central Universities (32620211T25), and Undergraduate Academic Support Program of the University of International Relations (3262021SYJ006, 3262021SYJ001).

**REFERENCES**


The last step of password strength evaluation

Yan Shao  
University of International Relations  
y_shao@uir.edu.cn

Xin Xin  
University of International Relations  
grace_xin77@163.com

Hong Di  
University of International Relations  
di_hong@163.com

The motivation of this poster is to identify the part of the misjudged passwords that hide in the strong passwords. In this poster, we propose Character Distance Strong Password Checker (CDSPC). Specifically, Consecutive Lead Character Distance (CLCD) and Average Adjacent Character Distance (Aacd) are introduced in CDSPC.

Research Problem

- Among the passwords which are identified as strong passwords, some of them may not satisfy the strong evaluation condition.
- Medium strength passwords or weak strength passwords that are misjudged as strong passwords will bring inestimable harm to the cyberspace or computer systems.

Our Approach

1. Evaluation Indicators

\[
P(i) = \frac{\sum_{i} P[i]}{\sum_{i} P[i]}, \quad \text{Aacd} = \frac{\sum_{i} |P[i] - P[i-1]|}{\sum_{i} P[i]}.
\]

\[\text{CLCD} = \sum_{i} |P[i] - P[0]|, \quad \text{Aacd} = \frac{\sum_{i} |P[i] - P[i-1]|}{\sum_{i} P[i]}.
\]

Aacd and CLCD reflect the comprehensive quality of passwords.

2. Determining Pivot Value

The number of passwords contained in each dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>hotmail</td>
<td>12,050</td>
</tr>
<tr>
<td>myspace</td>
<td>12,060</td>
</tr>
<tr>
<td>redc</td>
<td>12,050</td>
</tr>
<tr>
<td>yahoo</td>
<td>12,060</td>
</tr>
<tr>
<td>total</td>
<td>12,050</td>
</tr>
<tr>
<td>Native speakers of English</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>12,050</td>
</tr>
<tr>
<td>Native speakers of Chinese</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>12,050</td>
</tr>
</tbody>
</table>

Over the past few years, real password datasets have been leaked for various reasons. The leaked password datasets provide a great aid to password cracking technology. We choose the password datasets leaked in recent years to determine pivot value of password strength evaluation method proposed in this paper.

3. Checking Strong Password

CLCD and Aacd values evaluate password strength from different aspects. Therefore, to evaluate the strength of a given password accurately, it is necessary to calculate not only the distance between two characters, but also the threshold of two proposed indicators. Combining with the existing password strength evaluation method, a good password is determined through the pivot value to determine whether it is evaluated accurately.

Real Dataset Experimental Results

In order to demonstrate CDSPC's ability of identifying misjudged passwords, we use the password cracking tool Next-gen PCC to analyze the misjudged passwords patterns. By analyzing the top 2 patterns and probabilities of misjudged passwords from LPS and ZXCVMN, we find that the top 2 patterns and probabilities are shown in the left table.

Conclusion

- Our proposed method takes 0.05s time to calculate the character distance.
- CDSPC can be deployed on both client and server because of its lightweight feature.
- Experimental results show that our CDSPC can identify misjudged passwords effectively.
- Combining with the state-of-the-art password strength evaluation methods, the CDSPC can improve the accuracy of password strength evaluation results.

This work was specially supported by National Natural Science Foundation of China (62102133, Fundamental Research Funds for the Central Universities (3262021I125), and Undergraduate Academic Support Program of the University of International Relations (3262021S1006, 3262021I1001).