Analysis of Minimum and Maximum Character Bounds of Password Lengths of Globally Ranked Websites

Dr. Jatinderkumar R. Saini
Director I/C & Associate Professor
Narmada College of Computer Application,
Bharuch, Gujarat, India
Email: saini_expert@yahoo.com

----------**ABSTRACT**----------

Most of the websites today require a password for logging in to the website. The current paper presents a detailed study and analysis of length of passwords based on both minimum as well as the maximum bounds of the length of password. The paper deploys nearly two dozen of the top websites in the global popularity rankings. It has been found that the average minimum password length bound is 6.48 while the average maximum password length bound is 51.41. There are also popular websites listed in the paper with minimum password length of 1 (e.g. Wikipedia) and maximum password length of more than 20000 characters (e.g. Twitter and Facebook). The paper laments the lack of uniformity in the web community about the length of passwords. On the sidelines of these results, the paper also deliberates on various other aspects of password lengths of websites.

Keywords - **Characters, Maximum Bound, Minimum Bound, Password, Password Length, Website.**

I. INTRODUCTION

In In today's world highly influenced by online operations, there are many websites which require the user to create an online account in order to be able to access the services of the website. The services provided by the website may range from providing simple search facility, online shopping, chatting, social networking to accessing online technical assistance. There are also instances when the same services are also provided without asking the user to create an online account. But in such a case either the service is limited or the period of availability of the service is limited.

It has become a common practice on various websites to ask the user to first create an account and then access the services. This is beneficial to website administrators in two ways. Firstly, it lets the website administrators to collect website visit records and hence perform mining on the data in order to enhance the provision of services. Secondly, it lets the website administrators to design an algorithm that provides a more customizable and personalized experience to the visitors, more so if they are visiting the website again. As the user visits the website and optionally performs some transactions, the security of the user is of key concern for all stakeholders involved. The security of the user is also of key importance as the online operations may involve monetary transactions also. Besides, the typical pieces of information required during the registration process of creating an account on any website also form a part of the private data of the user. In all such cases, the password is largely responsible for assuring the security of the users' financial as well as non-financial data on the website. This paper deliberates in details about the length of passwords required by websites while creating an online account.

The remainder of the paper is divided into four sections respectively for related literature review, methodology, results and findings, finally followed by conclusion.

II. RELATED LITERATURE REVIEW

Lamport [1] in his historical work has presented a method of user password authentication which is secure even if an intruder can read the system's data, and can tamper with or eavesdrop on the communication between the user and the system. The method of Lamport assumes a secure one-way encryption function and can be implemented with a microcomputer in the user's terminal. Wu et al. [2] have discussed the vulnerability of passwords to various types of attacks including the dictionary attacks and impersonation attacks. They discussed the security for a simple and efficient three-party password-based authenticated key exchange protocol. Por [3] has addressed a newly discovered security threat named Frequency of Occurrence Analysis (FOA) attack in searchmetics password authentication scheme. A countermeasure technique that utilizes Metaheuristic Randomisation Algorithm (MRA) is proposed by Por to address the FOA attack. Por has presented the algorithm and an offline FOA attack simulation tool is developed to verify the effectiveness of the proposed method. In addition, a shoulder surfing testing is conducted to evaluate the effectiveness of the proposed method in terms of mitigating shoulder surfing attack. The experiment results of Por show that MRA is able to prevent

FOA and mitigate shoulder surfing attacks. Moreover, Por has proved that the proposed method is able to provide larger password space compared to the scheme used for benchmarking.

In a similar work, Saini and Desai [4] have presented a detailed structural analysis of username segment in e-mail Addresses of MBA Institutes of Gujarat State of India. Email has become a key mechanism of electronic communication for professional organizations that like to communicate with their subjects online and are slowly shifting to paper-less office. Their paper focuses specifically on academic institutions offering MBA courses in Gujarat state and attempts for textual analysis of the usernames of the institutional e-mail addresses. They have found that the institutions tend to design the username segment of their e-mail addresses by choosing words or combination of words from specific categories. Their paper also highlights the use of special characters, digits and random words in designing the usernames. On the sidelines, their paper lists the style of employing department names and designations for the design process. In another work, Saini and Desai [5] have presented a textual analysis of digits used for designing yahoo-group identifiers. Their paper presents digit-wise statistics and a detailed analysis of the usage of digits based on the study of nearly 1400 unique Yahoo group identifiers containing digits. The results show that the digit 0 contributes one-fourth of the total usage of digits, whereas the three digits 0, 1 and 2 collectively contribute more than 50% of the total digit usage. A number of areas that influence the selection of a digit or digits by the users in the design process have also been identified. They claim their attempt to be first of its kind in studying online user behavior based on the usage of digits for designing a unique online identifier.

Saini and Desai [6] have also presented a classification of character usage in unique addresses employed for accessing Yahoo! Groups service too. A tremendous increase in the use of internet for online communication like message sending is witnessed worldwide. Yahoo! Inc. provides one such service in the form of Yahoo! Groups. Each such group is identified and accessed using a unique group address. Their paper presents an analysis of nearly 5000 Yahoo-group addresses. It presents a classification of characters employed by users in designing these addresses into 5 major sets. Their results show that around 90% characters used for designing the Yahoo-group addresses are alphabets whereas the remaining l0% constitute from the domain of digits and special characters. The paper also elaborates on the divisional values of these proportions highlighting the user's preference for selecting a particular character. Saini and Desai [7] have also presented an analysis of usage of four-digit year number in designing Yahoo-group identifiers. For Yahoo! groups, each such group is uniquely identified using a group address or Yahoo-group identifier. Their work provides statistics on usage of different characters as well as on usage of digits for design of Yahoo-group identifiers. Their work specifically analyzes the usage of digits for year number usage in 'yyyy' format in design of more than 300 yahoo-group identifiers by the users of online world. They have found that Yahoo-group identifiers containing year value constitute 36.27% of the total usage of digits for designing Yahoo-group identifiers. Their paper also elaborates on the various reasons behind using year number in designing the Yahoo group identifier as well as the trend of usage of year number in this design process.

Choi and Un [8] have presented an anti-forensic technique by which password can be recovered only by the user's genuine fingerprint. Zhou et al. [9] have presented a method that is more resistant to linkage attacks and combines password to biometric systems for increasing performance. In a similar work, Islam et al. [10] have presented a mechanism of biometric template protection using watermarking with hidden password protection. Tan and Teo [11] have discussed various phishing attacks and the limitations of Secure Socket Layer (SSL). They have proposed an ID-based SSL protocol to counter the weaknesses of SSL. Henry and Luo [12] have presented a method that uses common password for protection of multiple accounts managed by the user. The web-based implementation of their method assures that even if the common password is leaked, the security of other accounts will not be compromised. They claim that their method is more convenient and secure for protection of multiple accounts compared to the traditional counterparts. Wang et al. [13] have presented an architecture that provides a trusted framework for users with roaming capability. Their method is based on Universal Serial Bus (USB) key and is provided in the form of a tool called IDKeeper. Shahid and Qadeer [14] have presented with techniques of protecting the passwords as well as to try to eliminate the problem of memorizing it. They have also tried to eliminate the dictionary attacks and brute-force attacks as well as evolve a password scheme with some graphic technique which can be used in highly secured applications. Hussain [15] has discussed the role of password in protecting the cloud computing environment. He has discussed the role of One Time Password (OTP) as well as One Day Password (ODP) in providing a secure mechanism to access the cloud environment. Joshuva et al. [16] have discussed the usage of graphical passwords as an alternative to the conventional alphanumeric passwords.

Kato and Klyuev [17] have argued that most of the users don't know important elements for password protection. In their paper, they have analyzed the tendencies in the password protection by surveying more than 250 students of a University. A questionnaire consisting of 15 questions about length of a password, strategy to create a new password etc. was given to the students for analyzing their understanding of creating a strong password. Hwang et al. [18] have presented with a work that creates flexibility into the RSA and the split-private-key RSA cryptography. In

the split-private-key RSA, the user can select, at his/her discretion, a password to derive the first of two private key portions. Their paper provides an extension that further allows the user to change the password without resorting to a regeneration of the public/private key pair. Additionally, validation of password input uses neither a hash value of the password nor other derivatives derived from just the password. They advocate that the cryptosystems created accordingly strengthen the password protection. Hirano et al. [19] have proposed a novel password protection mechanism that they call T-PIM or Trusted Password Input Method. Their system employs a hypervisor to isolate a trusted domain. They have also discussed security and usability of their proposed method in addition to discussion of attacks and comparison with conventional methods against data stealing malware. Yim [20] has proposed a method to prevent passwords from being hacked through sniffing of hardware protocols. The author proposed effective solution to keyboard vulnerability by providing a method that mixes noise with the password in order to make it in-feasible for the hacker to make out the exact phrase or text of the password.

III. METHODOLOGY

As a first step towards the analysis of password strengths and security of websites, a list of websites was identified for which the testing could be applied. The analysis was done for 23 websites listed in Table 1. It is noteworthy that out of the 23 websites listed in Table 1, many have been listed in the Alexa's list of top and most popular websites. Table 1 also lists the global Alexa Popularity Rank [21] of each website. The websites in Table 1 are presented according to the alphabetical listing order of the website name. It is noteworthy to mention here that Alexa does not provide any assistance regarding the password lengths of the websites. Alexa has been used only for the global popularity ranking of the websites. Further, only randomly selected few websites are chosen for study owing to three reasons, viz. popularity of website, possibility of creating an account on website and in-directly non-repeatability of same website. In-directly non-repeatability of website intends to avoid website like YouTube, which is ranked 3rd on Alexa popularity ranking list but which in-directly uses Google account, while the Google account has already been considered for discussion in the present paper.

In order to find the password length bounds, that is minimum and maximum number of characters allowed for creating the password, two approaches were used. As a first approach, the respective website was visited and the tool-tip or help listed adjoining the password field provided the description about the minimum and maximum number of characters allowed for the password field. The second approach was followed for cased where the website did not directly provide the assistance or description regarding the upper and lower bounds allowed for the length of the password. As a second approach, it was tried to create a user account on the website. The websites variously provide 'Sign Up', 'Register' and 'Create Account' options – all for creation of account on the website through username and password. The account creation process at many websites also asks for other information pieces as part of the account creation process. During the second approach, an intentional error was introduced while creating the password. This activated the website to alert the user about the length of the password in addition to other website specific criteria for password creation. An example of such an intentional error was to try creating the password initially with only a single character.

Table 1: Password Length Bounds and Alexa Popularity Rank of Websites

Sr. No.	Website	Minimum Characters	Maximum Characters	Alexa Popularity Rank [21]
1	Adobe	6	100	64
2	Amazon	6	128	8
3	Apple	8	32	48
4	ebay	6	64	27
5	Evernote	6	64	412
6	Facebook	6	20001+	2
7	Flipkart	4	20001+	101
8	Google	8	100	1
9	Hotmail	8	16	13125
10	IEEE	8	64	3178
11	Linkedin	6	16	10
12	Lycos	6	20	5372
13	Oracle	8	80	483
14	Outlook	8	16	625
15	PayPal	8	20	34
16	Pinterest	6	20001+	23
17	Rediff	6	12	309
18	Springer	6	60	2757
19	stackoverflow	8	20001+	50
20	Twitter	6	20001+	7
21	Wikipedia	1	20001+	6
22	WordPress	6	50	25
23	Yahoo	8	32	4

Making out the minimum number of characters for the password field on the website was an easy task compared to making out the maximum number of characters for the password field. For most of the websites, neither of the two listed approaches provided a clear indication of the maximum number of characters allowed for the website. This was truer to make out specifically for those websites which allow a large number of characters for populating the password field on the website. The last resort was to keep

on increasing the password length while creating an account as long as the website does not restrict it. For most of the websites, an upper limit was hence found. For six websites, namely Facebook, Flipkart, Pinterest, Stackoverflow, Twitter and Wikipedia, even though initially upto 100 characters were tried, the maximum limit was not found. This was followed by trying for 1000 and 15000 characters. These websites still allowed the creation of accounts. As a result, it was decided to try for 20001 characters as a final value to freeze upon. This was decided to avoid creation of otherwise not-to-be-used accounts. Further, instead of using 20000 characters, 20001 characters were used to emphasize that even though as part of the current research work, a stopping condition has been decided, but the websites are still allowing more than 20000 characters. The '+' sign suffixed to value of 20001 in Table 1 indicates that the website still allows more characters and that this is not the final value.

IV. Results and Findings

Based on the analysis of around two dozen globally most popular websites, it is found that there is no consistency and common agreement as far as the minimum and maximum number of characters required for creating a password are concerned. The reason for choosing the most popular websites was to emphasize that these websites being visited by maximum number of persons on earth ought to have in place highest level of security standards. It is found that all websites decide their minimum and maximum number of characters to be allowed in the password field according to their own wish. Of the considered 23 websites, 13 websites are in the list of top 50 websites visited on earth while 18 websites are among the top 500 websites visited worldwide. It is specifically noteworthy to mention here that there are trillions of websites the world over.

Table 2: Number of Websites Corresponding to Unique Minimum Password Length

Sr. No.	Minimum Password Length	No. of Websites
1	1	1
2	4	1
3	6	12
4	8	9
Total	-	23

The number of websites corresponding to unique minimum password lengths are presented in tabular format in Table 2. The same data are presented graphically in Fig. 1.

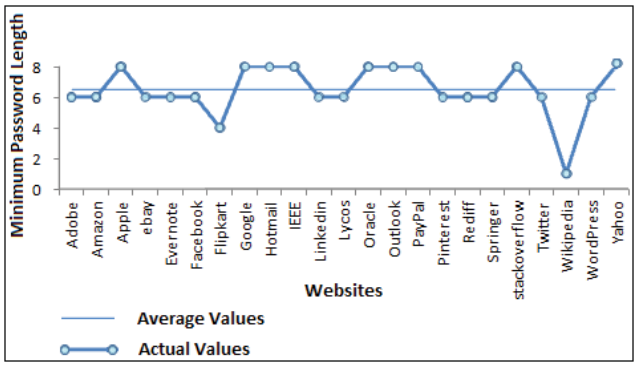

Fig. 1: Comparison of Average and Minimum Number of Password Characters of Websites

It has been found that 12 websites prefer to keep their password's minimum length to 6 while 8 being the second most preferred figure for 9 websites for minimum number of characters required for populating the password field. The average value of minimum number of characters required for creating the password by the websites is 6.48. There are 14 websites whose minimum password length is below average while there are 9 websites whose minimum password length is above average. The trend indicates that it would be preferable for the users to have a password of more than 6 characters.

Table 3: Number of Websites Corresponding to Range of Maximum Password Length

Sr. No.	Maximum Password Length Range	No. of Websites
1	<=10	0
2	11-20	6
3	21-50	3
4	51-100	7
5	101-150	1
6	>=151	6
Total	-	23

Similar to the analysis of minimum number of characters required for password field, an analysis of maximum number of characters required for password field was also done. But instead of using the unique minimum value of password length, a range of maximum password length was chosen. This was done to normalize the variation among maximum password length values. Accordingly, the 6 value ranges were decided and have been presented in Table 3. Excepting the case of outlier extreme of 20001 characters, Table 3 indicates that on an average, the maximum password length preferred and allowed by websites is either between the range of 51 to 100 or between the range of 11 to 20.

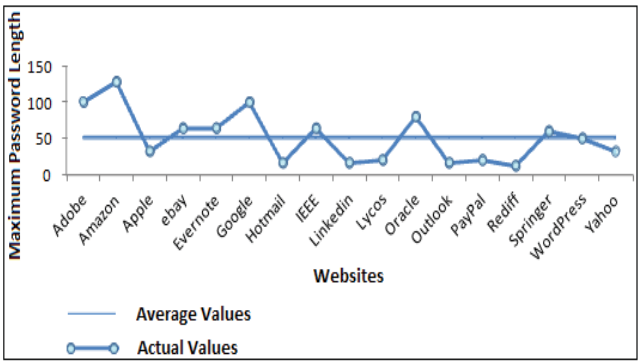

Fig. 2: Comparison of Average and Maximum Number of Password Characters of Websites

Further, in order to plot a graph from the values of Table 1, the extreme outlier values were ignored and the remaining values were plotted. The resultant graph is shown in Fig. 2. The average value of remaining websites was found to be 51.41. Fig. 2 indicates that there are 8 websites which have maximum password length more than the average value while the remaining 9 websites have maximum password length less than the average value. The trend indicates that the websites have an inclination to keep 50 or higher value as the upper limit of password length.

V. Conclusion

An analysis of maximum and minimum character bounds of most popular globally ranked websites by Alexa was done in the present work. For this purpose, 23 websites of which 18 were among the world's top 500 visited websites were analyzed. It is concluded that the average of minimum bound and maximum bound for populating the password field of websites is respectively 6.48 and 51.41, respectively. There are also websites which allow more than 20000 characters as the password length. It is concluded that there is no consistency or common understanding on fixing the minimum and maximum bound on password length. Almost 50% each of the websites were found to have their minimum password length below and above the average value. The same result was also found for maximum bound of password length.

To the best of author's knowledge this is the first attempt to formally study and analyze the variation in minimum and maximum bounds of password lengths across websites. The paper is not at all intended to promote or discourage the usage of a specific website. The paper also does not intend to deprecate any website owing to the minimum or maximum length of password of the website. This is solely an academic research meant to analyze the in-consistency and lack of common consensus and understanding among the web community on deciding the uniform values for minimum and maximum bounds of lengths of passwords. The presented results are best reported on the websites under study and the reader's discretion is expected while deciding the length of password on any website. Also, password length is just one criterion in deciding the strength of the password while other criteria like sequence, type and placement of characters, to name a few, also play a significant role in deciding password strength.

References

[1] Lamport L., Password authentication with insecure communication, published in the *Communications of the ACM, 24* (11), Nov. 1981, pp. 770-772 , doi: 10.1145/358790.358797

[2] Wu S., Chen K. and Zhu Y., Enhancements of A Three-Party Password-Based Authenticated Key Exchange Protocol, published in *The International Arab Journal of Information Technology*, *10*(3), May 2013, pp. 215-221

[3] Por L. Y., Frequency of Occurrence Analysis Attack and Its Countermeasure, published in *The International Arab Journal of Information Technology*, *10*(2), May 2013, pp. 189-197

[4] Saini J. R. and Desai A. A., Structural Analysis of Username Segment in e-mail Addresses of MBA Institutes of Gujarat State of India, published in *International Journal of Human and Social Sciences*, ISSN: 1307-6892, 5(6), October 2010, pp. 356-360

[5] Saini J. R. and Desai A. A., A Textual Analysis of Digits Used for Designing Yahoo-group Identifiers, published in *The IUP Journal of Information Technology*, ISSN: 0973-2896, 6(2), June 2010, pp. 34-42

[6] Saini J. R. and Desai A. A., A Classification of Character Usage in Unique Addresses Employed for Accessing Yahoo! Groups Service, published in *Karpagam Journal of Computer Science*, ISSN: 0973-2926, 12(1), January 2011, pp. 233-240

[7] Saini J. R. and Desai A. A., An Analysis of Usage of Four-digit Year Number in Designing Yahoo-group Identifiers, published in *ADIT Journal of Engineering*, ISSN: 0973-3663, 8(1), September 2011, pp. 22-27

[8] Choi W. Y. and Sung K. U., Anti-forensic approach for password protection using fuzzy fingerprint vault, published in the *proceedings of IEEE 7th International Conference on Computing and Convergence Technology (ICCCT)*, 2012, pp. 643-646

[9] Zhou X., Opel A., Merkle J., Korte U. and Busch C., Enhanced template protection with passwords for fingerprint recognition, published in the *proceedings of IEEE Third International Workshop on Security and Communication Networks (IWSCN)*, 2011, DOI: 10.1109/IWSCN.2011.6827719, PP. 67-74

[10] Islam M. R., Sayeed M. S. and Samraj A., Biometric template protection using watermarking with hidden password encryption,

published in the *proceedings of IEEE International Symposium on Information Technology, 2008*, ITSim 2008, 1, DOI: 10.1109/ITSIM.2008.4631572, pp. 1-8

[11] Tan C. H. and Teo J.C.M., Protection AgainstWeb-based Password Phishing, published in the *proceedings of IEEE Fourth International Conference on Information Technology, 2007*, ITNG '07, DOI: 10.1109/ITNG.2007.162, pp. 754-759

[12] Henry P. and Luo H., A common password method for protection of multiple accounts, published in the *proceedings on 14th IEEE Conference on Personal, Indoor and Mobile Radio Communications*, 2003. PIMRC 2003, 3, DOI: 10.1109/PIMRC.2003.1259242, pp. 2749-2754

[13] Wang X., Zhen Han Z. and Zhang D., IDKeeper: A Web Password Manager with Roaming Capability Based on USB Key published in the *proceedings of IEEE International Conference on Industrial Control and Electronics Engineering (ICICEE)*, 2012, DOI: 10.1109/ICICEE.2012.326, pp. 1228-1231

[14] Shahid M. and Qadeer M. A., Novel scheme for securing passwords, published in the *proceedings of 3rd IEEE International Conference on Digital Ecosystems and Technologies*, 2009. DEST '09, DOI: 10.1109/DEST.2009.5276738, pp. 223-227

[15] Hussain S., Access Control in Cloud Computing Environment, published in *International Journal of Advanced Networking and Applications*, ISSN: 0975-0290, 5(4), 2014, 2011-2014

[16] Joshuva M., Rani T. S. and John S. M., Implementing CHC to Counter Shoulder Surfing Attack in PassPoint – Style Graphical Passwords, published in *International Journal of Advanced Networking and Applications*, ISSN: 0975-0290, 2(6), 2011, 906-910

[17] Kato K. and Klyuev V., Strong passwords: Practical issues, published in the *proceedings of IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS)*, 2013, 2, DOI: 10.1109/IDAACS.2013.6662997, pp. 608-613

[18] Hwang J., Liao G. and Hsu Y., RSA with User Chosen and Changeable Password, published in *the proceedings of IEEE International Conference on Information Science and Applications (ICISA)*, 2011, DOI: 10.1109/ICISA.2011.5772312, pp. 1-2

[19] Hirano M., Umeda T., Okuda T., Kawai E. and Yamaguchi S., T-PIM: Trusted Password Input Method against Data Stealing Malware, published in the *proceedings of IEEE Sixth International Conference on Information Technology: New Generations*, 2009, ITNG '09, DOI: 10.1109/ITNG.2009.35, pp. 429-434

[20] Yim K., A New Noise Mingling Approach to Protect the Authentication Password, published in the *proceedings of IEEE International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)*, 2010, DOI: 10.1109/CISIS.2010.185, pp. 839-842

[21] Alexa Website, Available Online: www.alexa.com

Dr. Jatinderkumar R. Saini is Ph.D. from VNSGU, Surat. He secured First Rank in all three years of MCA and has been awarded Gold Medals for this. Besides being University Topper, he is IBM Certified Database Associate (DB2) as well as IBM Certified Associate Developer (RAD). Associated with more than 50 countries, he has been the Member of Program Committee for more than 50 International Conferences (including those by IEEE) and Editorial Board Member or Reviewer for more than 30 International Journals (including many those with Thomson Reuters Impact Factor). He has more than 55 research paper publications and nearly 20 presentations in reputed International and National Conferences and Journals. He is member of ISTE, IETE, ISG and CSI. Currently he is working as Associate Professor and Director I/C at Narmada College of Computer Application, Bharuch, Gujarat, India. He is also Director (Information Technology) at Gujarat Technological University, Ahmedabad (GTU)'s A-B Innovation Sankul.