An Improvement In Password Protection Overriding Trade-Off Between Security And Usability

S. Vinothkumar, S.Varadhaganapathy, R.Shanthakumari

Abstract: To ensure data security and to maintain privacy over the internet, authentication is used. The password based on text is the most commonly used verification form eventhough it is not favorable. The user tends to choose small password in the entire password space provided by the website. People use personal information as their password for easy memorization. They have an inherent trade-off between usability and security: while powerful passwords are hard for unauthorized person to guess, they are on the other side also hard for the user to recollect. The leaked dataset is used for research to evaluate the user's personal information residing in the passwords. Individual-PCFG method cracks the password much faster by generating personalized guesses. To guard the user from this type of attacks, we use distortion function. The distortion function transforms the user-chosen password into encrypted form that reduces the continuation of personal information residing in the passwords which gives the attacker a tough job to guess the password.

Index Terms: Private information, Depth, Password guess, distortion function

1 INTRODUCTION

Authentication is a process of distinguishing an individual, typically dependent on a username and a secret key. The procedure of a director giving rights and providing a way towards checking client account creates consent for access to the assets as an approval. The benefits and inclinations passed for the approved record rely upon the client's authorizations, which are either put away locally or on the verification server. [1]. Authentication is a mechanism introduced to validate user identity based on the user's data comparison during validation and the existing data stored at the time of registration. Authentication types include password-based authentication, authentication based on biometrics, etc. The authentication based on biometrics includes physical biometrics consisting of facial scanning, finer scanning, retina scanning, iris scanning. The biometrics of actions involves speech scanning, fingerprint scanning. Although biometric scanning guarantees better results, text-based password authentication dominates and other authentication methods may still be irreplaceable. Authentication is commonly performed in personal and public networks in the form of user IDs and passwords. The login credentials' knowledge is used to ensure the client is legitimate. First of all, every person registers using his or her own code. The user must remember the code and use the same for every subsequent use. With the increase in the wide variety of internet-enabled devices, system authentication is essential to allow comfortable communication in domestic automation and different networked environments.

- *Mr.S.Vinothkumar is an Assistant professor of Information Technology in Kongu Engineering College, Erode, Tamilnadu, India,.PH-9003683789. E-mail: vinoths@kongu.ac.in*
- *Dr.S.Vardhaganapathy is a professor of Information Technology in Kongu Engineering College, Erode, Tamilnadu, India .PH-9443034110.E-mail: svg@kongu.ac.in*
- *Ms.R.Shanthakumari is an Assistant professor of Information Technology in Kongu Engineering College, Erode, Tamilnadu, India.PH-9443416490.E-mail: rsk_shan@kongu.ac.in*

In the internet of things scenario, that's more and more turning into a reality, almost any conceivable entity or object may be made addressable and able to exchange records over a network. It is essential to realize that every access factor is a capability intrusion point. Every networked tool needs sturdy machine authentication and additionally, in spite of their generally constrained activities, those gadgets need to be configured in such a way that restricted permissions get access as well, to restrict what can be performed even though they may be breached. To prevent unauthorized access to valuable resources and data, existing developed organizations rely on passwords. Organizations are especially reliant on the protection of passwords. If a malicious party is aware of a single user's password, a whole device can end up being compromised. Research has found that if that user is forced to memorize too many passwords too often, a user may become less effective in memorizing passwords. There are two ways to compromise an account. Next, every account is subject to a given number of brute force attacks on a daily basis. The longer and more complicated the account's code, the less likely it is in such an attempt to become compromised. Third, if an account's password is written, it is exposed to an increased risk of being compromised regularly. [2]. Numerous specialists have proposed diverse validation systems; however no option can bring every one of the advantages of passwords without acquainting additional weights with clients. Many times user used to give password which will be easily memorized or related to the user's personal information, but at the same time it would be benefit for the hacker who can easily hack the password by using the personal information of the user.

2 LITERATURE REVIEW

Vaishnavi Mahajan, Pratima Nikam, Kalyani Padmawar, Nihar Ranjan [4] have used a simple distortion function for the user to enter password that would mix with a dictionary of passwords providing a strong password to the end user. The algorithm takes all personal details and password as inputs and provides strong password as output. It collects all the details and password and finds correlation between the information collected and password using Probability Context Free Grammar with a semantic rich algorithm. If the password consists of user's private data as a weak password then the end user has to create a new password. If the password is

strong, it will be mixed with some more complicated password chosen from the dictionary and a strong password will be mailed to the user. Ming Lei, Yang Xiao, Susan V. Vrbsky, Chung-Chih Li, and Li Li Liu have suggested a digital password definition that contains a little human computation to protect online user passwords. Based on the fact that a database has more data than any adversary does, user-determined linear generation functions are implemented to secure passwords for users. It is demonstrated that the proposed scheme is protecting against attacks on phishing, key logger, and shoulder surfing. H. Luo and P. Henry [5] have designed a technique that requires a user to remember only one password, called a common password, to access any of his/her accounts. Every account, called a special password, is secured by another different password. It is created by a one-manner hash work of an account-precise arbitrary wide assortment that is spared in an encryption shaped at the account server or an intermediary where the encryption key is extricated from the same ancient password. It overrides a convenient but dangerous method of using one or more passwords to secure multiple accounts. This guarantees that cracking a single password would not expose the standard password as well as any other similar password. This method uses a database server to store account identifiers and encrypted random numbers for every user and provide a web page containing a password calculator written in Javascript. Using a web browser, the user will create a common password for any process that is in need of password authentication. The existing system considers either the password security or the user memorability. This paper provides the nutshell of both security and usability. Based on the result, the distortion function is proposed to balance both security and usability. The distortion function can be implemented in the client side as per the client requirements using an individual web page in order that the user has to access the account anywhere at anytime. The distortion function transforms the user created original password that may contain personal information into strong password that deduces the correlation between personal information and passwords.

3 IMPLEMENTATION OF DISTORTION FUNCTION

A public password dataset, which contains private information from a Chinese website, had been utilized. It is called a 12306 dataset because all passwords are from a website www.12306.cn, the official website of China's online train ticket reservation system. The exact number of users on the 12306 website is not available; however, we conclude at least tens of millions of registered users in the system as it is the official website for the entire Chinese railway system. The database contains more than 110,000 Chinese passwords that are plaintext passwords, as well as several forms of personal information, such as the name of a user and the unique identification number issued by the government. The information in this database is considered reliable, as the website needs a real ID number to register and people need to provide valid personal information to book a ticket. Various forms of personal information are also included in 12306. Personal information is tabulated in TABLE1. The ID number issued by the government is a unique 18-digit number that includes the personal information of the owner intrinsically. Specifically, digits 1-6 represent the place of birth, digits 7-14 represent the date of birth, and digit 17 represents the sex.

TABLE 1
Personal Information

Type	*Representation*
Name	*User's Chinese Name*
Email address	*User's Registered Email address*
Cell phone	*User's Registered Cell Phone Number*
Account Name	*The username used to log in to the system*
ID Number	*Government issued ID number*

Using the Chinese data set, it was found that 657 guesses were generated by PCFG and 459 guesses were created by Individual-PCFG for 100 records taken in the test set. The distortion functions can be proposed to protect weak passwords that include personal information, as they increase the security of passwords by significantly reducing the correlation between user passwords and personal data. Distortion functions can mitigate the problem of including personal data in user passwords without losing the password chosen by the user significantly [3]. To evaluate the involvement of personal information in creating an individual password in an accurate and systematic fashion, a novel metric called Depth is introduced. Depth value ranges from 0 to 1. A greater range is correlated with a stronger association. Depth − 0 Business means that no personal information is included in a password, and Depth − 1 means that the entire password matches perfectly with one type personal data. Though Depth is mainly used for measuring an individual password, the average Depth also reflects the degree of correlation in a set of passwords. The Depth algorithm process is provided in figure 1.

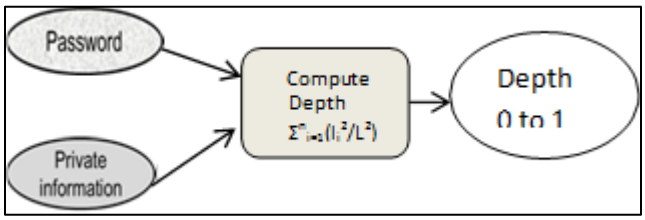

Figure 1 *Depth Evaluation*

In terms of strings, a window method is implemented which takes word and private information as input. A dynamic-sized window sliding from the beginning to the top of the term is retained to conduct the computation. The window's initial width is two. If the window process matches a certain type of personal data, the width of the window will expand by one. Instead we tend to try again to suit the private information with the new phase. If a match is found, the width of the window will be increased until a match occurs. The window resets to the initial size at this point and begins to move from wherever the game happens. Whereas,

because the term is used to record the duration of each matched word process, the array known as tag array has a similar length. After windowing through the entire word sequence, the tag array is used to illustrate the meaning of Depth, which is adding squares of the matched word step length separated by the word length square. Mathematically,

$$\text{DEPTH} = \sum_{i=1} n(1^2/L^2) \text{---------- Equation 1}$$

Where n indicates the number of matched password segments, li indicates the length of the corresponding matched password segment; L is the length of the password [3].The best distortion function can be judged from the result such that the depth value must be nearest to 0.

3.1 DISTORTION FUNCTION

Distortion function is a method of changing the plaintext password chosen by the users into secured form that disposes the user-chosen password. Even a simple distortion function can mitigate the correlation between personal information and password. This function must be implemented in the client side in order to balance both the security and usability. This can be implemented in several forms like creating a webpage with distortion function deployed in the backend database, software integrated with the distortion function.

DISTORTION FUNCTION 1

This function calculates the middle position m from the length of the password. At the beginning of the password, aa is added and after including the values of 0 to m-1 positions in the original password, add da and then include the values in the position of m to l and finally generate two random numbers and include at the end of the original password. For example, sasirekha will be transformed into aasasirdaekha51.The depth value should be calculated for individual password and the average can be used for comparison to find the optimized function. The flow is given in the figure 2. When the user password is passed on through the distortion function 1, the encrypted passwords are obtained. The depth value is calculated for each password and the average is taken as 0.0061.The depth value is subjected to small variation due to the use of random

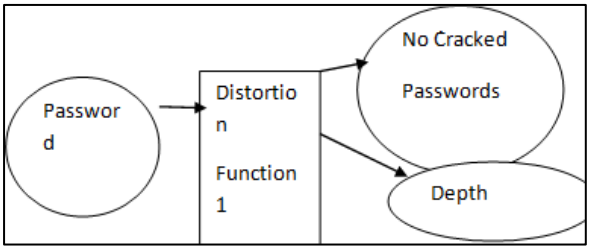

Figure 2.*Distortion Function 1*

DISTORTION FUNCTION 2

This function calculates the length of the password. Generate two random numbers and append at the end of the original password. Then append two random characters such that the random numbers are replaced by corresponding characters based on ASCII values .For example, sasirekha will be transformed into sasirekha95ux.The depth value should be calculated for individual password and the average can be used for comparison to find the optimized function. The flow is given in the figure 3. When the user password is passed on through the distortion function 2, the encrypted passwords are obtained. The depth value is calculated for each password and the average is taken as 0.0194.The depth value is subjected to small variation due to the use of random number generation. Apply Individual-PCFG to the test set which consists of 100 passwords, no password has been cracked. When compared with the function 1, function 2 is not secure since it does not split the personal information in the password but appends two digits and two characters.

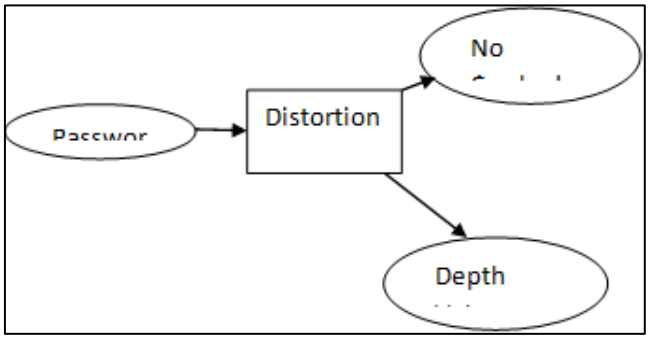

Figure 3. *Distortion Function 2*

DISTORTION FUNCTION 3

This function calculates the length of the password. Check the value in the last position of the password. If it is a number, generate two random numbers and append the at the end of the original password else generate three random characters such that the random numbers are replaced by corresponding characters based on ascii values and append at the end of the original password else. For example, sasirekha will be transformed into sasirekhausq and sasirekha9 will be transformed into sasirekha940.The depth value should be calculated for individual password and the average can be used for comparison to find the optimized function. The flow is given in the figure 4. When the user password is passed on through the distortion function 3, the encrypted passwords are obtained. The depth value is calculated for each password and the average is taken as 0.0296.The depth value is subjected to small variation due to the use of random number generation. Apply Individual- PCFG to the test set which consists of 100 passwords, no password has been cracked. When compared with the function 1 and 2,3 is not secure since it does not split the personal information in the password but appends three digits or characters.

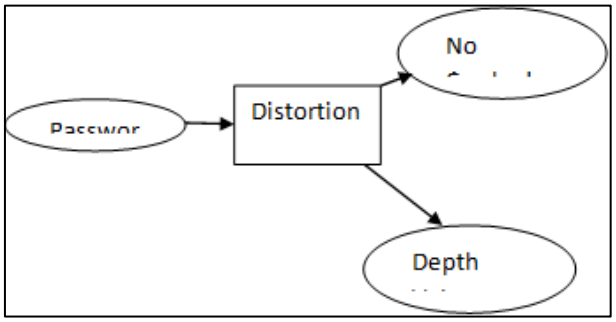

Figure 4. *Distortion Function 3*

DISTORTION FUNCTION 4

This function generates a random character and inserts it in every odd position of the password. For example, sasirekha will be transformed into syawsgiyrqeakehtau. The depth value should be calculated for individual password and the average can be used for comparison to find the optimized function. The flow is given in the figure 5. When the user password is passed on through the distortion function 4, the encrypted passwords are obtained. The depth value is calculated for each password and the average is taken as 0.0033.The depth value is subjected to small variation due to the use of random number generation. Apply Individual-PCFG to the test set which consists of 100 passwords, no password has been cracked. When compared with the function 1,2 and 3 , function 4 is secure with less depth value since it splits each position of the password.

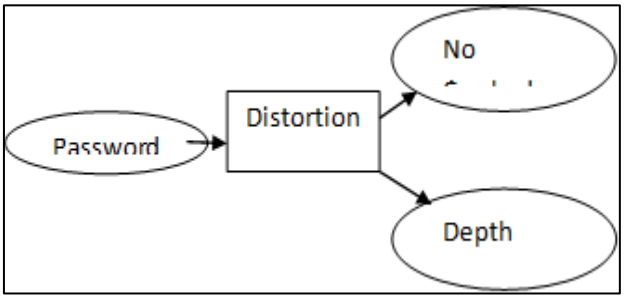

Figure5. *Distortion Function 4*

DISTORTION FUNCTION 5

This function sets distance value d as 1 and shifts d positions in the alphabet with mod 26 if it a character and shifts d positions with mod 10 if it is a number. For example, sasirekha will be transformed into tbtjsflib. The depth value should be calculated for individual password and the average can be used for comparison to find the optimized function. The flow is given in the figure 6. When the user password is passed on through the distortion function 5, the encrypted passwords are obtained. The depth value is calculated for each password and the average is taken as 0.0117.Apply Individual-PCFG to the test set which consists of 100 passwords, no password has been cracked. When compared with the function 1,2, 3 and 4, function 4 is secure with less depth value since it splits each position of the password.

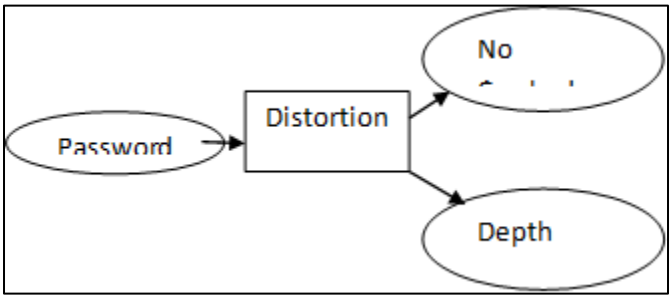

Figure 6.*Distortion function 5*

DISTORTION FUNCTION 6

This function inserts exclamatory symbol at all the odd positions. For example, sasirekha will be transformed into s!a!s!i!r!e!k!h!a!.The depth value should be calculated for individual password and the average can be used for comparison to find the optimized function. The flow is given in the figure 7. When the user password is passed on through the distortion function 6, the encrypted passwords are obtained. The depth value is calculated for each password and the average is taken as 0.0012.Apply Individual-PCFG to the test set which consists of 100 passwords, no password has been cracked. When compared with the function 1,2,3,4 ,5 and 6, function 6 is secure with lowest depth value since it splits each position of the password by adding least preferred symbol.

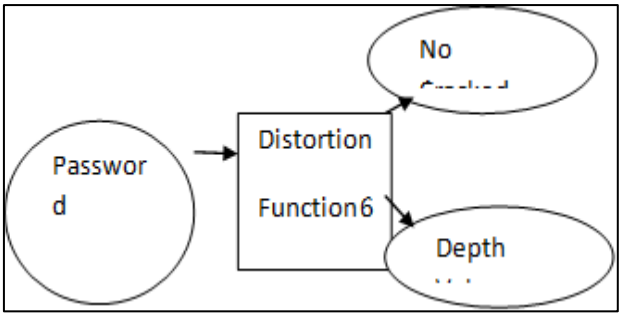

Figure 7.*Distortion function 6*

4 PERFORMANCE EVALUATION

The implementation of distortion function shows that it reduces the relationship of private information and the secret word. Even though all the proposed functions mitigates the depth value, function 6 which adds the exclamatory symbol at each odd position results in the lowest depth value 0.0012 i.e nearer to 0.Hence, function 6 is the optimized solution that can be preferred by the users to transform their personal information included password to ensure security of the password. The comparison graph is given in the figure 8.

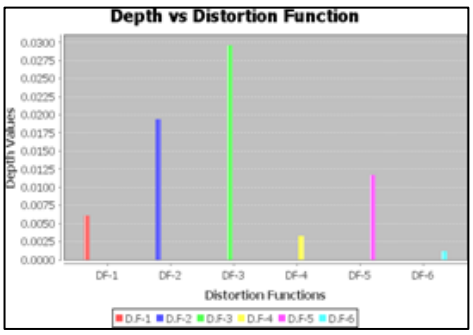

***Figure 8**.Comparison of distortion function's performance*

5 REALTIME SCENARIO

In real time environment, the distortion function has to be chosen based on the depth value and by considering the norms provided by all the accounts. The best suggestion is to implement it in the client side since the user can have the full rights to change the distortion function according to the security measures need by them. The flow is given in the figure 8. First, the webpage has to be created for each individual. Apply distortion function in the backend database while storing the credentials using admin option. Whenever the particular user wishes to open an account, he/she has to open the webpage and provide the login id and the password chosen by the user. Then the data will be compared with the database and password in the database is decrypted and checked with the password provided by the user at an instant time. If both the password matches, the encrypted password will be provided to the user that shows a trade-off between security and usability. The user can copy the password and use it for actual account sign in.

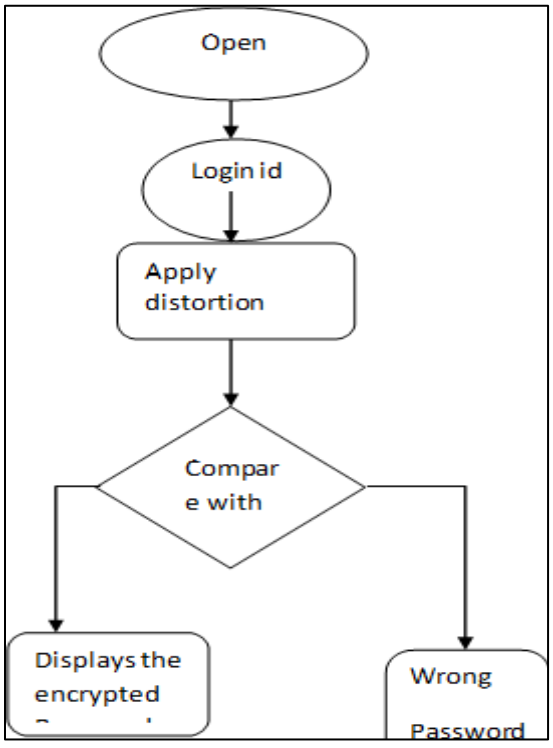

***Figure 9.** The flow of real time implementation*

6 CONCLUSION

We have demonstrated that the use of distortion function mitigates the relationship between private information and secret word. This makes the attacker harder to guess and crack the passwords as done already. This ensures the security without putting the user in the hectic situation when the user really feels hard to remember the random password. Thus the enhancement of password protection was done based on trade-off between security and usability.

7 FUTURE WORK

The distortion functions can be integrated with the server and the unique key can be provided to the user at a time of account creation and change of password to reduce the burden of user to own a webpage and opening it every time whenever the user logs into the account. The extra layer of security can be provided by applying hash algorithm to the encrypted password.

8 REFERENCES

[1] Richard Shay , Elisa Bertino, A Comprehensive Simulation Tool for the Analysis of Password Policies‖, International Journal of Information Security; Volume 8,Number 4; August 2009

[2] P.Sasirekha,S.Varadhaganapathy,J.Premalatha,

[3] C.Visali,,Improvement on Traditional LDS Structures to Crack the Passwords Based on Individual's Information‖,International Journal for Modern Trends in Science and Technology, Vol. 03, Issue 12,December 2017, pp.39-45.

[4] Vaishnavi Mahajan Pratima Nikam Kalyani Padmawar Nihar Ranjan, Secure Of Web-Accounts Using Personal-PCFG, International Education and Research Journal, E-ISSN No : 2454-9916 | Volume: 3 | Issue : 12 | Dec 2017.

[5] H. Luo, P. Henry, –A common password method for protection of multiple accounts‖, Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings, January 2004.

[6] Yue Li, Haining Wang, and Kun Sun, Personal Information in Passwords and Its Security Implications, IEEE Transactions on Information Forensics and Security, vol. 12, no. 10, October 2017.

[7] R.Veras, C. Collins, and J. Thorpe, On the semantic pat- terns of passwords and their security impact, in Proceedings. NDSS, Feb. 2014.

[8] Kuo, C., Romanosky, S., Cranor, L.F. Human selection of mnemonic phrase-based passwords. In: SOUPS '06: Proceedings of the second symposium on Usable privacy and security,pp. 67–78. ACM Press, New York, NY, USA (2006).

[9] S. Gaw and E. W. Felten, Password Management Strategies for Online Account, Proceedings of the 2nd Symposium on Usable Privacy and Security, ser. SOUPS '06, 2006, pp. 44–55.

[10] B. Ur, P. G. Kelley, S. Komanduri, J. Lee, M. Maass, M.L. Mazurek,0 T. Passaro, R. Shay, T. Vidas, L. Bauer,N. Christin, and L. F. Cranor, —How does your password measure up? The effect of strength meters

on password creation‖, Proceedings of the 21st USENIX conference on Security symposium, ser. USENIX-SS ’12. USENIX Association, 2012, pp. 65–80.

[11] R. Shay, S. Komanduri, P. G. Kelley, P. G. Leon, M. L. Mazurek, L. Bauer, N.Christin, and L. F. Cranor,—Encountering Stronger Password Requirements: User Attitudes and Behaviors‖, Proceedings of the 6th Symposium on Usable Privacy and Security, ser. SOUPS ’10,2010, pp. 1-20J. Bonneau, C. Herley, P. C.Van Oorschot, and F.Stajano, — The quest to replace passwords: A framework for comparative evaluation of web authentication schemes,‖ in IEEE Security & Privacy, 2012.