Honeywords Generation Method for Passwords based on User Behaviours

K. Vinitha[1] Dr. R. Manicka Chezian[2]
[1]Research Scholar [2]Associate Professor & Head of Department
[1,2]Department of Computer Science
[1,2]NGM College, Pollachi, India

Abstract— The Honeywords are selected deliberately, such that a cyber-attacker who steals a file of hashed passwords cannot be sure, if it is the real password or a Honeywords for any account. Moreover, entering with a Honeywords to login will trigger an alarm notifying the administrator about a password file breach. At the expense of increasing the storage requirement by 24 times, the authors introduce a simple and effective solution to the detection of password file disclosure events. In this study, it scrutinizes the Honeywords system and highlight possible weak points. Also, it suggest an alternative approach that selects the Honeywords from existing user information, a generic password list, dictionary attack, and by shuffling the characters. Four sets of Honeywords are added to the system that resembles the real passwords, thereby achieving an extremely flat Honeywords generation method. To measure the human behaviors in relation to trying to crack the password, a tested engaged with by 820 people was created to determine the appropriate words for the traditional and proposed methods.

Key words: Cyber-Attacker, Detection Password, Generic Password, Shuffling Characters, Real Password

I. INTRODUCTION

A password is the popular authentication technique being used today despite many newer ones, such as biometric based techniques and dual factor authentication. Users tend to use small simple passwords and for this reason as well as their somewhat universal use, they are vulnerable to being compromised. Hence, it has become important to make progress in combating cracking techniques. Since these are becoming increasingly sophisticated, has become a salient issue. Intruders are increasingly eavesdropping on communication between legitimate users and servers as well as masquerading as authorized users or remote servers so as to be able to steal sensitive information. A good password has to have two features: a user can remember it and it is difficult to guess. Unfortunately, these two work against each other such that a password that is easy to remember is generally short and hence, easy to guess. Moreover, most people choose to use a single password for multiple accounts, because one is easy to remember. Invariably, people have a hierarchy of passwords, for example, they do not use the same password for email as they do for their bank account, in particular, because the bank requires more stringent security. The idea behind honey words is to create a relation between the real password and decoy hashed passwords, such that for every user the latter look like real passwords. The honey words are these decoys. An attacker can recognize the presence of honey words in a password file, as it is very unusual to have multiple passwords for a single user account. However, even if the attacker can crack multiple passwords associated with a user, he or she does not know which Honeywords are, and which the real ones are.

II. RELATED WORK

Hashing the plaintext or password is a one-way function, which makes it hard to find the required password. However, rainbow tables, which are massive tables filled with hash values and can be used to find a required password, whereby a hacker employs them to find the password by reversing the hashing function. Despite of a rainbow table taking up a lot of storage when holding it, attackers can usually crack the password in a shorter amount of time than when applying the brute force technique. Most existing biometric template protection schemes (BTPS) do not offer as strong security as cryptographic tools. Moreover, they are unable to determine whether or not a probe template has been downloaded the database by an imposter or an authentic user. Consequently, the “Honeywords” idea was proposed to detect the cracking of hashed password databases. In particular, an extra layer of protection is needed with biometric feature schemes, as these have been shown to be flawed. A honey template protection scheme relating to faces has been proposed and evaluated as representingan improvement on existing schemes user’s real password can be distinguished among Honeywords for each user by using a secure server called a “honey checker”, which triggers an alarm when a honey word is used .

III. PASSWORD ATTACKS

Password attacks include different character combinations being tried until a match with the correct password is found. There are several types of password attacks, some of the most important of which are described next.

A. Brute Force Attacks

In this type of attack, all the possible combinations of the password are applied to break it. It can also be applied to crack encrypted passwords wherever they are saved in the form of encrypted text.

B. Dictionary Attack

A dictionary attack is applied to verification data by trying every word in the dictionary. This kind of attack is targeted at sites with a high probability of success, such as those with weak passwords or with only a few key combination numbers. This attack is faster than an attack of brute force and is more successful when a weak, public or short password is used.

C. Phishing Attack

This is where an attacker attempts to retrieve legitimate users’ confidential and sensitive credentials fraudulently by mimicking electronic communications from a trustworthy or public organization in an automated fashion. The aim of

phishing is to steal sensitive information, such as online banking passwords and credit card information from Internet users [18]. These attacks use a combination of social engineering and technical spoofing techniques that persuade users into giving away sensitive information that the attacker then uses to make a financial profit.

D. Password Guessing Attack

In this attack, the adversary steals the file of the password from the main server, and also obtains plaintext passwords by reversing the hash values detected.

IV. HONEY WORDS GENERATION METHODS

In this section, some of the Honeywords generation methods are discussed.

A. Chaffing-by-Tweaking

This method involves tweaking the real password by selecting the character positions that will be tweaked to produce the honey words, so the user password will be the seed of the generator algorithm. The same type of character will be selected: letters are replaced by letters, digits by digits, and special characters by special characters. For instance, when t=3 and the last characters have been selected for tweaking, the method for the generator algorithm is Gen (k,t). While another approach called "*chaffing-by-tweaking-digits*" is carried out by tweaking the last positions that contain digits. For instance, if the last algorithm has been used, then for the password *42hungry* and , the Honeywords*12hungry* and *58hungry* may be generated.

B. Chaffing-with-a-Password Model

In this technique, the generator algorithm takes the password from the user, and then a probabilistic model of the original passwords is relied upon to generate the Honeywords. To give an example of applying this technique, known as *modeling syntax*, the model is divides the real password into character sets. For example, the password *mice3blind* is decomposed as four-letters + one-digit + five-letters (L4+D1+L5) and is replaced with the same structure, such as *gold5rings*.

C. Hybrid Method

This method involves combining of the strength of different honey word generation methods, e.g. chaffing-with a-password model and chaffing-by-tweaking-digits. For instance, let the original password be apple1903, then the Honeywordsangel2562 and happy9137 might be produced as seeds to chaffing-by-tweaking-digits.

V. ANALYSING THE FLATNESS IN THE NEW HONEYWORDS

The Honeywords created in this first group are associated with personal questions will most probably lead to personal answers. In this case, six answers, which are either in letter or digit form and the letters and digits, are then randomly mixed to produce five Honeywords. The high level of association of these Honeywords with user real answers will make it difficult for the adversary to identify which one is the real password, i.e. this increases the flatness. In contrast, the traditional methods do not take into consideration whether there is a personal password, because all the Honeywords are generated by tweaking some letters or digits in the real password. Dictionary attacks are commonly used to break passwords, but in the proposed method they are used to generate the Honeywords. Such an attack involves most of the passwords that have been created by users around theworld, by using an algorithm based on English language rules to make the search in this dictionary to find Honeywords very close to the original password. In addition, a minority of users have a strong password, whereby they select some letters randomly and create a meaningless one. However, most users in this group still select letters or digits from their names and/or personal dates.

VI. RESULTS & DISCUSSION

It is a difficult to measure how people are thinking when they are creating a password, because it depends on unpredictable user behavior. The scenario involved dividing the passwords into three groups: good, personal, and generic. Then, the participants were provided with the, and ask to nominate words that could be passwords, this column being titled "nomination". The idea behind this step was to ascertain how many people would nominate the real password among the honeywords, and how many words they would choose amongst which they believed the password would be found. Having chosen their words, they were asked to identify the single one that they thought was the real password and if they got it wrong then Intrusion Detection System IDS would trigger attempted intrusion, but if successful access was granted. The first type, namely the good password, was strong, being created with random letters, digits, and special characters. The results showed that this type of password is very strong, as most people who participated in the tested experiment did not choose it among the honeywords, The second type of password is the personal password, which was created based on information relating to the users. The tested revealed that the new method is better than the traditional methods. Finally, with the same scenario, the third type of password.

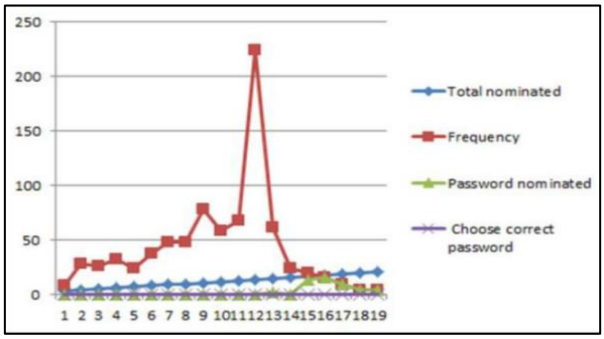

Fig. 1: The Results of the Proposed Method when a strong Password Was Applied

Fig 1.Illustrates the results of the strong password for the new method, with the total nominated representing how many words the participants chose, while the frequency is how many people selected a particular amount of passwords.

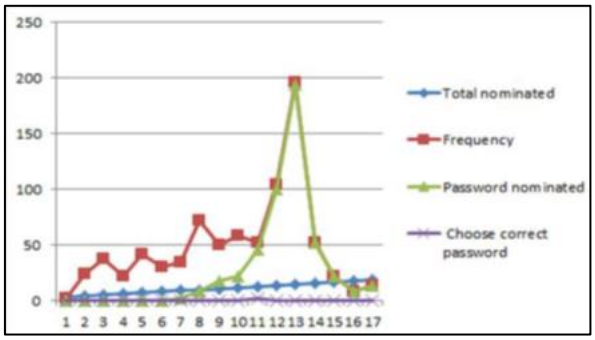

Fig. 2: The Tested Results when the real Password Contained Personal Information

Fig.2 shows the results of the proposed method when the real password is the personal information type and clearly, the number of people who nominated the password amongst their choices increased, being 502 out of 820.

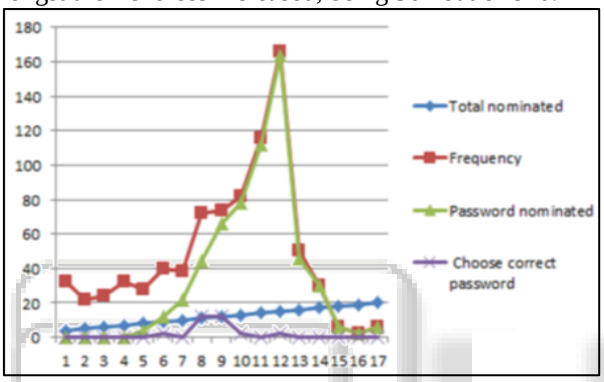

Fig. 3: The tested results for the proposed method when the real password is generic

Fig.3 illustrates the new method when a generic password was the real password and the results show that this type provides the worst outcomes of the three, but the new method still gives better results than with the traditional one. Specifically, the total number who chose this password was 630 out of 820, and it was guessed correctly 21 times.

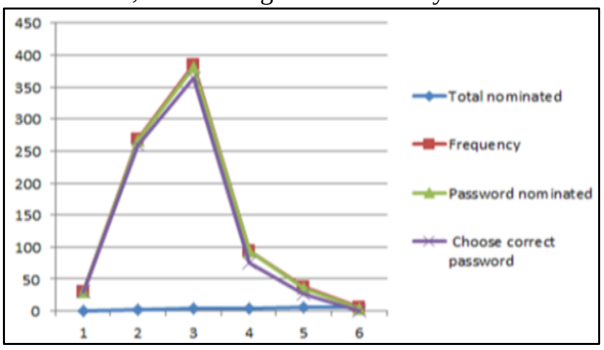

Fig. 4: The tested results for the traditional method of "Chaffing-by-Tweaking".

Fig.4 showing the outcomes when Chaffing-by-Tweaking was applied in the tested, it is clear that the number of participants guessing the real password was very high, standing at 794 times out of 820, whilst the number who nominated was 812. Moreover, most people nominated just one or two words out of 25 in and no one nominated more than six, which suggests that many were confident they from the beginning which was the correct password.

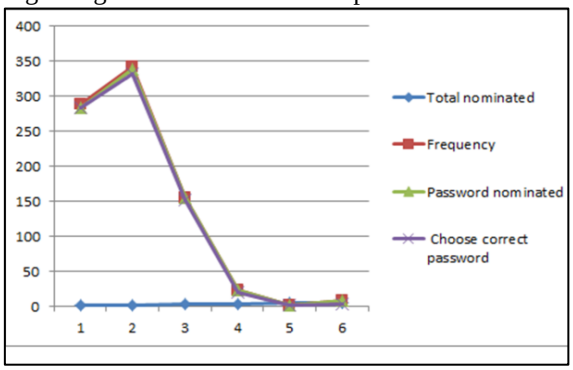

Fig. 5: The tested results for the traditional method of "Chaffing-by-Tweaking"

Fig.5 the results for the traditional method of Chaffing-by-Tweaking-Digits are shown. This method provides slightly better results than Chaffing-by-tweaking in that the password was guessed correctly 756 times out of a possible 820. To give an example of how the proposed method generates the honeywords, in TABLE I the password is "Ujemgzae91#e". Clearly, the first row contains honeywords generated based on personal information, while the second row has those created based on the worst passwords list. The rest of the table was generated by shuffle the letters and digits. A dictionary attack was not used in this table, because no word is similar this password.

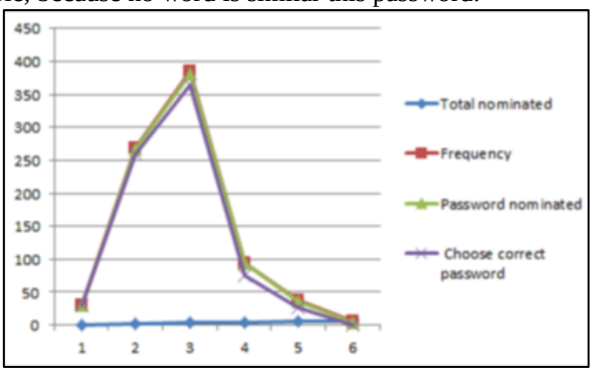

Fig. 6: The tested results for the traditional method of "Chaffing-by-Tweaking-Digits".

Prestol#70	Jordy$86	Steves@75	Mechaill$81	Anna^1945
Liverpool@2005	Football&1234	Password*1111	Music@6666	bond@007
Booboo&75	Love&2014	Mustang@16	Zme1qo@55req	Epalm#1999ks
Pufna*37xy	Msac^hs31	Neadjg_69	Vlpheo$10r	Kp#12zxme
Ltcbas!00j	Tg36$ewba	Ujemgzae91#e	Rpnq#fxg	Lsczyr&12

Table 1: Testbed with the New Method & A Good Password

StationRoad1960	Church2016	Morgan2010	Stevs1958	Andy2000
Alunaliceza	Andralice2004	Anasialice1977	Anaalice85	Hello131313
Nicholas123	Andrew 1212	Password222	Welcome777	Alice1974
ElArzd204	O9lefcm7ss	Oxsr15dox	Z7erpmc0	Enm12q

Movxg20w	Qica12r00	Hvagjr4193	Nlpqroo1870	Zaqu2w88

Table 2: Testbed with the New Method and A Generic Password

Table II illustrates an example when the tested was applied with the generic password, "password222", being drawn from the list of worst passwords. The honey words in the second row were generated based on a dictionary attack.

VII. CONCLUSIONS

In this paper, a new honey words generation method has been proposed. This method was developed to overcome the problems that exist with the traditional methods. The proposed method is based on personal information, dictionary attacks, the worst password list (generic passwords) and shuffling the characters. User behaviour is the underpinning principle the new method, because creation of the passwords differs from one user to another. Some limitations regarding the extant honey words methods were mentioned in this have been discussed and these have been overcome by the proposed method have been explained. A tested has been applied to obtain the results using 820 participants and these have shown that the new method is better than the traditional ones.

REFERENCES

[1] S. M. Gurav et al., "Graphical password authentication: Cloud securing scheme," in 2014 International Conference on Electronic Systems, Signal Processing and Computing Technologies, 2014, pp. 479-483.
[2] D. Mishra et al., "A secure password-based authentication and key agreement scheme using smart cards," Journal of Information Security and Applications, vol. 23, pp. 28-43, 8, 2015.
[3] S. Houshmand, S. Aggarwal and R. Flood, "Next Gen PCFG Password Cracking," IEEE Transactions on Information Forensics and Security, vol. 10, pp. 1776-1791, 2015.
[4] B. Gurung et al., "Enhanced virtual password authentication scheme resistant to shoulder surfing," in 2015 Second International Conference on Soft Computing and Machine Intelligence (ISCMI), 2015, pp. 134-139.
[5] Siqiong Fan, Zhenfu Cao and Xiaolei Dong, "Cryptanalysis and improvement of a smart card-based identity authentication scheme," in ICINS 2014 - 2014 International Conference on Information and Network Security, 2014, pp. 152-157.
[6] J. Ma et al., "A study of probabilistic password models," in 2014 IEEE Symposium on Security and Privacy, 2014, pp. 689-704.
[7] S. M. TaiabulHaque, M. Wright and S. Scielzo, "Hierarchy of users ′ web passwords: Perceptions, practices and susceptibilities," International Journal of Human-Computer Studies, vol. 72, pp. 860-874, 12, 2014.
[8] Dr. Ari Juels RSA, Professor Ronald L. Rivest MIT. Dr. Ari Juels RSA, Professor Ronald L. RivestMIT., "For Stronger Password Security, Try a Spoonful of Honeywords ," 2013.
[9] D. Vishwakarma and C. E. V. Madhavan, "Efficient dictionary for salted password analysis," in 2014 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2014, pp. 1-6.
[10] H. Kumar et al., "Rainbow table to crack password using MD5 hashing algorithm," in 2013 IEEE Conference on Information & Communication Technologies, 2013, pp. 433-439.
[11] H. M. Ying and N. Kunihiro, "Decryption of frequent password hashes in rainbow tables," in 2016 Fourth International Symposium on Computing and Networking (CANDAR), 2016, pp. 655-661.
[12] E. Martiri, B. Yang and C. Busch, "Protected honey face templates," in 2015 International Conference of the Biometrics Special Interest Group (BIOSIG), 2015, pp. 1-7.
[13] M. J. Bhole, "Honeywords: A New Approach for Enhancing Security," International Research Journal of Engineering and Technology (IRJET), vol. 02, pp. 1563, 2015.
[14] L. Catuogno, A. Castiglione and F. Palmieri, "A honeypot system with honeyword-driven fake interactive sessions," in 2015 International Conference on High Performance Computing & Simulation (HPCS), 2015, pp. 187-194.
[15] N. Chakraborty and S. Mondal, "A New Storage Optimized Honeyword Generation Approach for Enhancing Security and Usability," arXiv Preprint arXiv:1509.06094, 2015.
[16] Jesudoss and N. Subramaniam, "A SURVEY ON AUTHENTICATION ATTACKS AND COUNTERMEASURES IN A DISTRIBUTED ENVIRONMENT," .
[17] E. I. Tatli, "Cracking More Password Hashes With Patterns," IEEE Transactions on Information Forensics and Security, vol. 10, pp. 1656-1665, 2015.
[18] L. Wu, X. Du and J. Wu, "Effective Defense Schemes for Phishing Attacks on Mobile Computing Platforms," IEEE Transactions on Vehicular Technology, vol. 65, pp. 6678-6691, 2016.
[19] Uusitalo, J. M. Catot and R. Loureiro, "Phishing and countermeasures in spanish online banking," in Emerging Security Information, Systems and Technologies, 2009. SECURWARE '09. Third International Conference On, 2009, pp. 167-172.

