Secure and Usable Bio-Passwords based on Confidence Interval

Aeyoung Kim[1], Geunshik Han[2], and Seung-Hyun Seo[3]*
[1]National Institute for Mathematical Sciences, Yuseong-gu, Daejeon 34047 Korea
aeyoung@nims.re.kr
[2]Hanshin University, Osan-si, Gyeonggi-do 18101 Korea
gshan@hs.ac.kr
[3]Korea University, Jochiwon, Sejong-si 30019, Korea
crypto77@korea.ac.kr

Abstract

The most popular user-authentication method is the password. Many authentication systems try to enhance their security by enforcing a strong password policy, and by using the password as the first factor, something you know, with the second factor being something you have. However, a strong password policy and a multi-factor authentication system can make it harder for a user to remember the password and login in. In this paper a bio-password-based scheme is proposed as a unique authentication method, which uses biometrics and confidence interval sets to enhance the security of the log-in process and make it easier as well. The method offers a user-friendly solution for creating and registering strong passwords without the user having to memorize them. Here we also show the results of our experiments which demonstrate the efficiency of this method and how it can be used to protect against a variety of malicious attacks.

Keywords: Key Extraction, Confidence Interval, Non-Cognitive Password, Bio-Cryptosystem

1 Introduction

A password is the most commonly used authentication solution to protect important or personal information from unauthorized access, but they can be quite vulnerable to attack and inefficient for users as well. The individual user's password composition and security awareness are key to the security of the password. Although its importance is emphasized, many users prefer to compose a short and simple password and to use the same password in many accounts, because of the ease of remembering. Such user tendencies often have serious consequences for their information and systems. An easy-to-remember password is often easy to guess and can leave the users account open to a plethora of attack techniques, e.g., a trial-and-error method such as brute force attack, an exhaustive-search method in a pre-arranged word listing such as dictionary attack, a direct observation method such as shoulder surfing, etc. [9] [17] [21]. In order to avoid the weaknesses of conventional log-in password systems which invite these types of attacks, many organizations use a password meter, enforce a password policy, or carry out a multi-factor authentication system. Many researchers have proposed alternatives to conventional passwords. A password meter measures the password strength by calculating its entropy or checking its vulnerability to a dictionary attack. This helps users create more secure passwords, but does not strictly require them. Thus, users still tend to choose their easy-to-remember passwords. A password policy helps to improve the password strength by forcing a user to create passwords that are more difficult to guess. However, a strict policy that satisfies the guidelines can lead to user frustration with a difficult-to-remember password.

Journal of Internet Services and Information Security (JISIS), volume: 7, number: 1 (February 2017), pp. 14-27

*Corresponding author: Department of Mathematics, Korea University, Sejong Campus 2511, Sejong-ro, Sejong City 30019, Korea, Tel: +82-44-860-1316

Because policies are different for each system, it is difficult to apply the same password. Thus, remembering one's passwords can be difficult. The difficult-to-remember passwords are usually stored on a note, smartphone, or PC. These user behaviors lead to another security problem. The recent multi-factor authentication combines the first factor ”something you know” (e.g., password, question and answer) with a second factor “something you have” (e.g., identification card, smartphone, credit card with IC (integrated circuit), or USIM (Universal Subscriber Identity Module)) and a third factor “something you are” (e.g., fingerprint, face, iris). Traditional one-factor-based authentication systems, like password, are easily discovered. To increase the difficulty of compromising a software system, thereby improving its level of security, many consumer-electronic-related systems commonly use the first and second factors. The usability in these system is low. On the other hand, the security is higher because a user must pass more steps and hold something, e.g., an ID card, security card, smart phone, etc., to log in.

Various password alternatives have been tried, such as noisy passwords, reformation-based passwords, geographic location-based passwords, time signature-based password, graphical passwords, and virtual passwords [7] [19] [24] [2] [13] [10]. Despite various attempts, these methods are still unwieldy alternatives to a strong password, because they also suffer from the dilemma of security and usability. Moreover, all four of these attempts to improve security or usability still find it difficult to balance the security and usability, protect against illegal sharing of authentication solutions, and guarantee uniqueness. Using a biometric as the third factor, e.g., a fingerprint or face, is the best method to improve these limitations and simultaneously enhance security and efficiency. While this approach is highly attractive, it is not easy to obtain high key stability to bind a password, since each time a sensor acquires an image, the image will most likely be somewhat different. To address this issue, we propose a simple bio-password method using eigenface-based confidence intervals. This will allow the creation and use of strong passwords, which will not need to be memorized by the user. This approach also permits for the creation of strong passwords, while side-stepping the need for a password policy, while likewise avoiding the need for memorization or even input. Finally, this method also protects against the many different forms of attack and weaknesses noted previously in this paper.

This paper is an extended version of our ICTC 2013 paper[12]. We are making it available in order to have more detailed descriptions and reasonable evaluation in comparison to its short version. The contributions of this paper are to describe a technique for generating the strong and easy-to-use password in more detail[12]. Obviously what one would want in password systems is a ideal system that uses a strong but easy-to-use password for each user and is safe for a variety of attacks. The strong password has a high combinatin rate, long length, and uniqueness. The high combination rate means that the letters, numbers, and special characters are uniformly mixed. The long length means a length more than the length defined by the password policy, and the uniqueness means that the same password among a user's passwords does not exist. Nevertheless, this strong password should be easy to use. The basic idea of our scheme is to extend the biometric-based confidence interval set introduced in [11] and to map unique information from a website or app, such as a domain, to the expanded set. The biometric-based confidence interval set is the result of a scheme for extracting stable values from ustable values based on noises data such as biometrics [11].

Additionally, the passwords generated by the proposed scheme also are evaluated and compared with the hacked passowrds that were actually used and the collecteded passwords by anonymous surveys. In the short version, we mainly focused on introducing our idea, and experimented with creating a few sample passwords with just a few face images. In contrast, we used AR Face DB[14] to generate the strong passwords and the 14 million passwords in the RockYou list[4] and the collected 245 passwords[15] collected from 109 users.

The organization of this paper is as follows: In Section 2, we will describe issues for strength and usability of passwords and confidence intervals with eigenfaces for extracting stable values from unstable data. In Section 3 we will present an outline of our suggested approach along with the algorithms which are needed in its implementation. Section 4 shows face-based passwords and the results of comparison with other passwords. Finally, Section 5 presents some concluding remarks.

2 Related Work

2.1 Pasword Strength and Usability

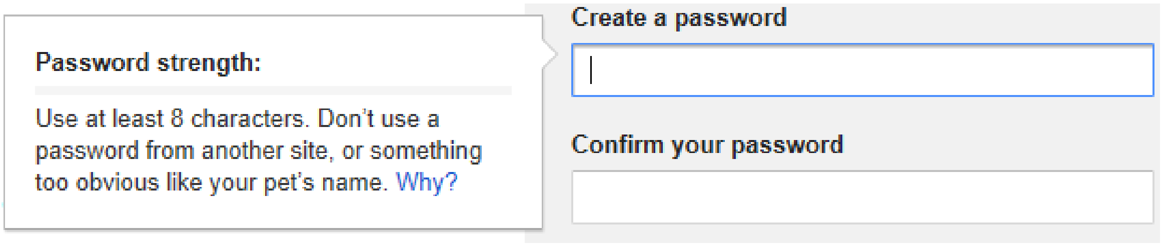
Several previous studies have shown the weakness of passwords against many type of attacks [18] [23] [22] [3]. Especially, Weir et al. attempted to gauge the security provided by conventional password creation rules with the RockYou set which contained over 32 million passwords [21], [23].

Overall, user-generated passwords with few restrictions can easily be guessed and are susceptible to dictionary or brute force attacks. The use of stronger passwords is helpful to prevent such attacks or an additional authentication solution to protect our information from unauthorized access. Stronger passwords can be created by using methods such as password meters, password policies, and alternative passwords. However, most of these efforts are rather uncomfortable.

Some sites use password meters to force users to generate strong passwords [20], [8]. Early methods of evaluating password strength were quite simple. For example, given length N and the number of alphabet C, the strength is calculated as $Nlog2C$ bits [5]. This strength is a result of the length and character composition.

More recently, password evaluation involves the measuring of password input by noting the password length. Certain attributes are preferred, such as the presence of uppercase and lowercase letters, numbers, symbols, etc., while other features are avoided, like character repetition, letters placed in alphabetical sequence, or the use of only letters and/or numbers.

Password meters indicate the strength of passwords and help users generate stronger ones, as shown in Fig 1. Although feedback with a password measure is provided, many meters to measure password strength have limited accuracy because their rules are over simplified and users still tend to create easy-to-remember passwords. Moreover, such meters couldn't work to check re-use. Therefore, increasing the password strength is limited.



Password strength:

Use at least 8 characters. Don't use a password from another site, or something too obvious like your pet's name. Why?

Figure 1: Example of a password meter

Unlike a password meter, enforced password policies with strong rules are more useful to obtain a strong password than password meters. The principles of password security are included, such as [1].

- Longer and more complex password
- Uniqueness (one password per one website or account)
- Increasing change regimes (change password once a month)

- Reduction in the allowed input error rates (three times to input error)

The use of longer and more complex password is the first step toward a secure password. The OWASP (Open Web Application Security Project) has recommended best practices for password length (a minimum of eight and a maximum of 160 characters) and complexity (combination of uppercase letters, lowercase letters, numbers, punctuation marks, mathematical and other conventional symbols) [19]. On the other hand, with this recommended length, the average length of representative password lists is about 7 characters [23]. Password uniqueness is also an important attribute to consider when creating a new password which does not reuse in anywhere. However, it is difficult to satisfy uniqueness because there is a memorization problem that requires the user to have to remember a number of passwords that are as many as the user's accounts. The guidelines of the GIACP (Global Information Assurance Certification Paper) have suggested that a user cannot reuse any of their previous five passwords [6]. According to the experimental results of Poornachandran et al., among many studies on the uniqueness or reuse of passwords, about 59 percentage of users reuse their passwords for multiple accounts and about 35 percentage of users of major SNS websites, including Twitter, Facebook, and Gmail use extensive passwords [16]. Liao and Yu revealed password repetitive patterns and the way these patterns would lead to reconsideration of a password security policy [25].

In addition, users are required to change their password once a month or once every three months in systems with increased change regimes. Furthermore, users are required to exactly input such longer, more complex, frequently changed, and unique passwords. Users who exceed the allowed input error rates have to prove their own id or right to access with another independent authentication process. The user then has to generate a new password again. Ideal passwords based on these enforced password policies are not only more difficult to guess but also harder to remember. Many users may write them down on a notepad or store them in an electronic device such as a smart phone.

Most users tend to select the easier way rather than the more difficult one. They often create short and simple passwords, reuse them for multiple accounts, use them for a long time, and write them down on a note. This behavior causes other vulnerabilities for password-based authentication systems. This situation is known as dilemma or trade-off between security and usability of passwords. Some researchers have overcome such trade-offs by proposing generation, management, or support methods to ensure a secure and usable password; there are noisy passwords, reformation-based passwords, geographic location-based passwords, time signature-based password, graphical passwords, and virtual passwords [7] [19] [24] [2] [13] [10]. However, these attempts have not yet been implemented because most of them are impracticable or not secure enough.

2.2 Biometric Stability by Using Confidence Interval Analysis

Overall, the best way to design strong authentication involves combining one of the previous factors with the third, ”something you are”. Security and efficacy are most readily acquired through the use of this type of multi-factor approach founded on the incorporation of biometrics. However, though this approach holds much potential, currently, biometrics is not easily adapted for use in present cryptography-oriented two-factor systems. It is especially challenging to maintain stability of key features in biometrics as the visual data acquired by sensors will always be a little different. Fortunately, Eigenface-based confidence interval sets can help to integrate biometrics into traditional passwords systems. The use of such sets as predictors of recognition performance has already been proposed [11]. Then can improve the stability of data received from biometric input. The concept of confidence interval set *CIS* generation is founded on a T-test which uses varied visual data of the same face.

The algorithm 1 is for generation of *CIS* to get the stable values from noise and unstable data or signal. This generation algorithm includes methods for calculating a principal component pc_i based on the Principle Component Analysis (PCA), taking a test value t, and generating a confidence interval ci_i. Given $t, \bar{x}, s, u, v$, and c_i; set $B_{(j)} = pc_{i(j)}$, where $1 < i < m$ and $2 < j < n$ for the input; and set $CIS = \{ci_i\}_{i=1}^{m}$ for the output. The algorithm produces CIS as an output using the parameters listed above. The formal description of this algorithm 1[11] is as follows:

Algorithm 1: PRE_TCI_GENERATION

```
$CIS, \bar{x}_i, s_i, u_i, v_i \leftarrow 0$ ;
$t \leftarrow t_{0.01}(n-1)$ ;
for $i \leftarrow 1$ to $m$ do
    $\bar{x}_i \leftarrow \frac{1}{n}\sum_{j=1}^{n} pc_{i(j)}$ ;
    $s_i \leftarrow \frac{1}{n}\sum_{j=1}^{n} {pc_{i(j)}}^2 - {\bar{x}_i}^2$ ;
    $u_i \leftarrow \bar{x}_i - t \cdot s_i \cdot n^{-\frac{1}{2}}$ ;
    $v_i \leftarrow \bar{x}_i + t \cdot s_i \cdot n^{-\frac{1}{2}}$ ;
    $ci_i \leftarrow (u_i, v_i)$ ;
    $CIS \leftarrow CIS \cup ci_i$ ;
```

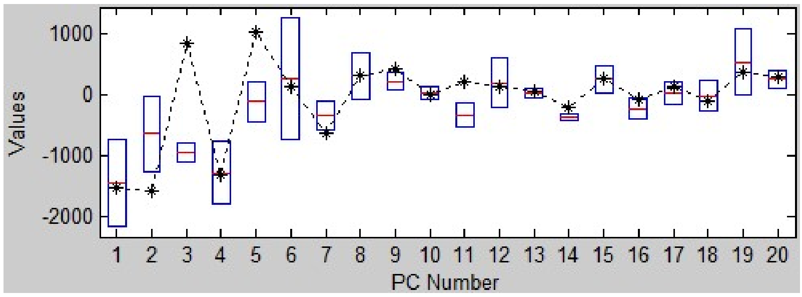


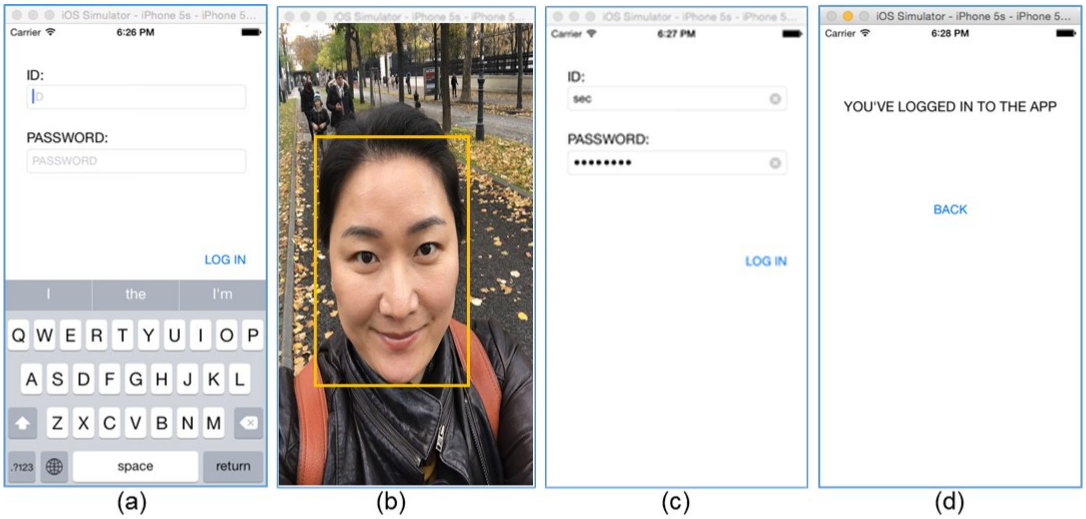
Figure 2: Example of ci_i in CI and pc_i^B in B

The CIS is stored securely such as private key and is used to compare with m principal components of a new acquired face as the input data. Given CIS and new face B, we check whether pc_i^B in B is included in $ci_i = (u_i, v_i)$ of CIS as shown in Fig. 2; the result ir_i is shown in Eq. (1). The sum of ir_i for m is the inclusion rate, IR. The IR is used for a recognition or prediction in the previous work, but it is used to get a permission to generate a bio-password with the CIS as described in the next section.

$$ir_i = 1,\ if(pc_i^B \leq u_i) \wedge (pc_i^B \geq v_i) \quad (i = 1, ..., m) \tag{1}$$

3 Bio-Password by Using a Map based on Confidence Interval

This section presents a new password generation algorithm for face-based user authentication systems. The proposed bio-password scenario presented in Fig.3 is the same as the scenario in general password systems with the exception that this bio-password scenario uses a picture instead of a typed input. The algorithm converts the picture into a password, which is automatically entered into the system. The algorithm uses information, which identifies each website like the domain, and confidence interval set based on eigenfaces. Previous studies usually extracted features from the eigenfaces by using complex extraction methods for an effective matching process. Unlike previous studies, this paper uses the difference between the lower and upper bounds of the CI based on the eigenface, which is a simpler approach compared to the complex feature extraction process.



YOU'VE LOGGED IN TO THE APP

Figure 3: Scenario for face-based password usage

The proposed scenario with the algorithm gives the user some certain benefits in relating to security and usability as follows:

- the user does not need to remember passwords.
- the user does not need to type passwords into the password box.
- the user does not need to enter the same password multiple times.
- the user does not need to worry about creating a strong password that conforms to the password policy.

For the accurate functioning of the bio-password authentication system, the password system must extract a fixed value, which is the registered password, from the registered user's face. Generally, the values extracted from biometrics, including those of the face, are not stable. In order to extract a sufficiently stable value to use as a password, we designed a password generation algorithm using a CI. The proposed bio-password generation scheme consists of three steps as shown in 4: acquiring permission for access to the CI set (step 1), generating a specific code for the account (step 2), and generating a bio-password (step 3). Prior to the execution of these steps, the user must have a CI set. If a CI set is not available, create a new CI set with the user's face by means of algorithm 1 mentioned previously in Sec. 2 [11]. Although there are various methods for storing the password, this paper assumes that the generated CI set is safely stored in the user's smart device. The saved CI set works like a master key or security certificate in the proposed scheme.

The first step is similar to an access control step and involves getting permission from the user for using the CI. After acquiring the license, the CI-based information for generating a password is obtained in the next step. In order to obtain this information, the captured image of the face is used to generate the eigenfaces $F = \{pc_1, ..., pc_m\}$ for the number of principle components, m, which equals the size of the CI set $CIS = \{ci_1, ..., ci_m\}$, where $ci_i = (u_i, v_i)$. When the i-th element, pc_i, is included in ci_i such that $(u_i \leq r_i)$ and $(r_i \leq v_i)$, the inclusion rate R increases by one. The value of R indicates how similar the registered face from the CI set is to the face F in the photo provided for authentication. If R is greater

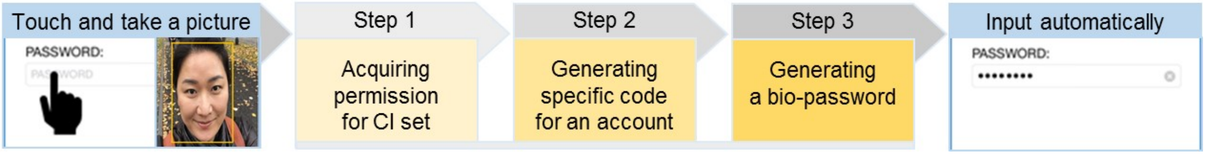

Touch and take a picture
PASSWORD:
PASSWORD
Step 1
Acquiring permission for CI set
Step 2
Generating specific code for an account
Step 3
Generating a bio-password
Input automatically
PASSWORD:
••••••••

Figure 4: Steps for the proposed bio-password generation

than the threshold value T for permission acquisition, the final output of this step can be obtained from the CI-based information as shown in Fig. 5. The *CIS* includeing w_i can be represented by a cube map with three coordinates, (u,v,w) as shown in Fig. 6, and it consists of characters with 26 uppercase letters, 26 lowercase letters, 10 numbers, and 32 symbols in ASCII.

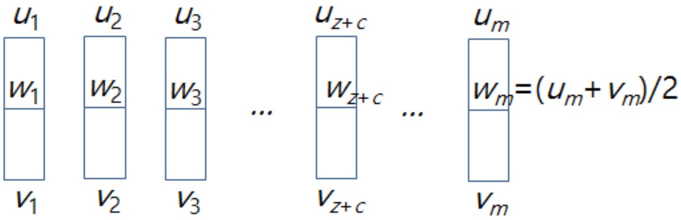


Figure 5: Composition using *CIS*

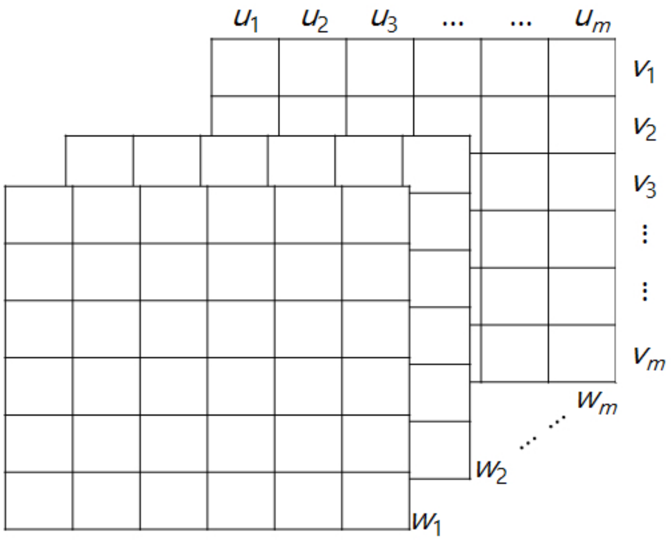


Figure 6: The cube map represented by *CIS*

The second step in the authentication process is the generation of unique information from the site to allow login. Each website or app has unique information such as domain, IP, and identification number, which can distinguish it from other websites and apps. This information can be used as the input value in the second step. The output in the form of (x, y, z) can be used in combination with the output of the first step. In the proposed method, we used the domain with XOR, ADD, separation, and merge. The constituent characters $\{c_1,...,c_n\}$of a domain are separated into three parts

$C_x = \{c_1,...,c_{k1}\}, C_y = \{c_1,...,c_{k2}\}, C_z = \{c_1,...,c_{k3}\}$, where n is $k1+k2+k3$ and C_i is the i-th character of the target domain represented in ASCII. Following this reconstruction of the domain, (x,y,z) is formularized by 2. However, this reconstruction is not limited to this formula.

$$(x,y,z) = (\sum_{j=1}^{k1} c_{3j-2}, \sum_{j=1}^{k2} c_{3j-1}, \sum_{j=1}^{k3} c_{3j}) \mod m \tag{2}$$

The third step with the algorithm 2 is for generating a bio-password using the *CIS* from the result of the first step and (x,y,z) from the result of the second step. Additionally, given the password length n and change count c, (x,y,z) is applied to this map as shown in Fig.7 to determine the target bio-password. The password length n depends on the password policies of the target websites. Finally, the bio-password consists of algorithmic operations in the algorithm 2 as shown in Fig.8. This password represented on the map in Fig. 6 is shown in Fig.9. The password obtained from the bio-password extraction algorithm in the third step is automatically applied to the password box.

Algorithm 2: BIO-PASSWORD_EXTRACTION

```
$biopwd[] \leftarrow 0$ ;
$w \leftarrow (u_{z+c} + vz + c)/2$ ;
for $i \leftarrow 1$ to $n$ do
    $biopwd[i] \leftarrow u_{(x+i-1) \mod m} \oplus v_y \oplus w$ ;
    $biopwd[i] \leftarrow biopwd[i] \mod (p-q)$ ;
    $biopwd[i] \leftarrow biopwd[i] + q$ ;
```

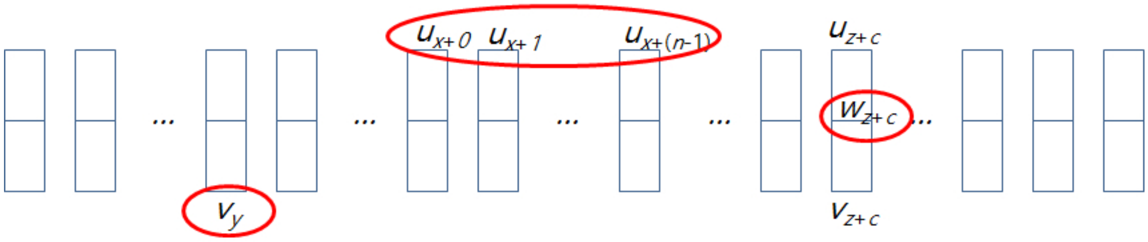


Figure 7: Example of the usage (x,y,z) on the *CIS*

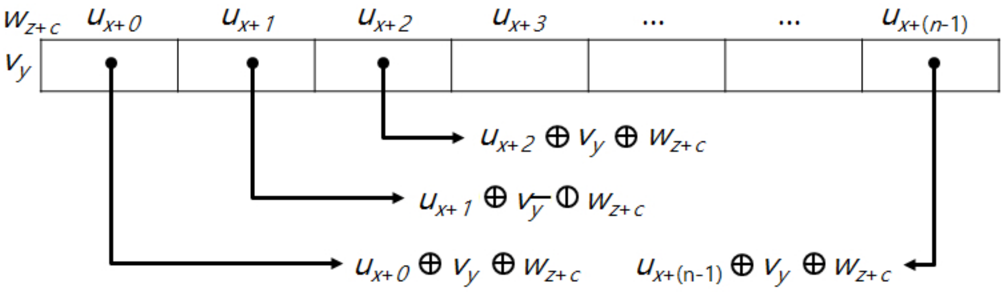

w_{z+c}
u_{x+0}
u_{x+1}
u_{x+2}
u_{x+3}
...
...
$u_{x+(n-1)}$
v_y
$u_{x+2} \oplus v_y \oplus w_{z+c}$
$u_{x+1} \oplus v_y \oplus w_{z+c}$
$u_{x+0} \oplus v_y \oplus w_{z+c}$
$u_{x+(n-1)} \oplus v_y \oplus w_{z+c}$

Figure 8: Example of bio-password composition in the *CIS*

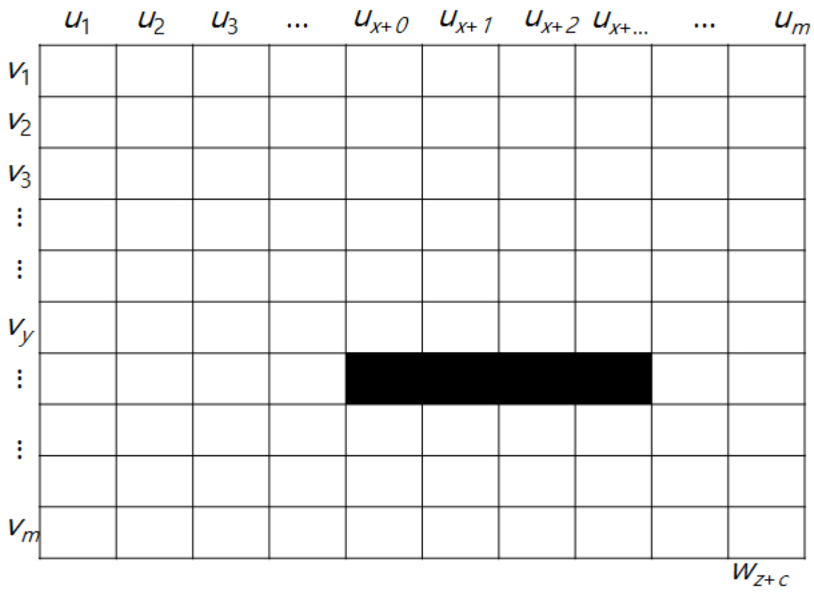


Figure 9: Example of the bio-password on the *CIS*-based map

Table 1 shows examples of bio-passwords of lengths 8, 10, and 12 characters for eight websites. If needed, we can generate a longer password by using the bio-password extraction algorithm and taking a photograph. The number of passwords that can be created in this process is m^3 for a face.

Table 1: Example of Passwords by Extracted Bio-Key

Website	Bio-Password Length		
	8	10	12
facebook	+”$ˆ0Mwc	+”^{0}MwcV	+”^{0}MwcV/K
naver	7xk‘X#5/	7xk‘X#5/n’	7xk‘X#5/n’0u
google	F907xk‘X	F907xk‘X#5	F907xk‘X#5/n
sbs	^{0}MwcV	^{0}MwcV/K	^{0}MwcV/K%:
nate	>cQFrr*y	>cQFrr*yTv	>cQFrr*yTvA[
microsoft	$u>cQFrr	$u>cQFrr*y	$u>cQFrr*yTv
shinhan	Sv4/.HQi	Sv4/.HQi#*	Sv4/.HQi#*&F
nonghyup	v5A*C.Rg	v5A*C.Rg,N	v5A*C.Rg,NE”

4 Experimental Results

We evaluated the strength scoring provided by a password creation policy against the bio-passwords and real life password-datasets. We assessed the effectiveness and security for our scheme by implementing the algorithm in MATLAB and used AR Face DB as grayscale face picutures [14]. The face-based *CIS* was calculated with reference to the algorithm *PRE_TCI_GENERATED*, which becomes an effective tool for extracting stable information from unstable data or noised data like biometrics [11]. The strength scores in this test were compared by collecting the RockYou password list (Dec. 2009), the Facebook password list (Sep. 2010), and surveyed password list (Dec. 2015) as real life datasets. The RockYou list, which was users's passwords in RockYou.com, is a large-scale password list that has been hacked and released [21]. For the experiment, the number of available passwords by the list is about 14×10^6 passwords. The Facebook password list was leaked or stolen from the website, and then it was also posted for generating and testing passwords for a good reason. The surveyed password list that contained 181 passwords was collected for this experiment from a college freshman who usually create their accounts and passwords in 2 or 3 years. The test module for evaluating strength of passwords was implemented by applying a policy in www.passwordmeter.com. There are two indicators to represent the strength: a score and complexity. The score is added or subtracted as to whether or not the given condition in the policy is satisfied. The complexity is determined by the calculatd score such as below 20 is 'Very Weak', 21-40 is 'Weak', 41-60 is 'Good', 61-80 is 'Strong', and above is 'Very Strong'.

The test datasets were divided into G1 to G4 as follows.

- The bio password list by using G1
- The RockYou password list by using G2
- The Facebook password list by using G3
- he Surveyed password list by using G4

Table 2 presents password information in terms of how many uppercase letters, lowercase letters, numbers, and symbols are contained in each password list on average. The average of the length is almost the same with 9.98 in G1, 8.75 in G2, 9.35 in G3, and 9.66 in G4. On the other hand, the component ratio differs. The percentage that contains uppercase letters and symbols are 4.35, 1.25 in G2, 1.74, 1.75 in G3, and 5.89, 5.15 in G4. The passwords in these lists are overly weighted in term of the lowercase letters and numbers. However, the percentage that contains an uppercase letter and symbol in G1 are 21.53 and 41.21, respectively. Uppercase letters are almost the same as the percentage of lowercase letters and numbers. Considering the number of symbols in the proposed scheme is 32, which exceeds the number of uppercase or lowercase letters, the percentage that contains a symbol in G1 is reasonable in comparison with other lists. These results show the composition of the passwords by the proposed scheme is the best because of impartiality.

Table 2: Password information from each password list

Password lists	**G1**	**G2**	**G3**	**G4**
Number of passwords	12,508	14,342,307	2,493	181
Average of length	9.98	8.75	9.35	9.66
% that contain uppercase letter	21.53	4.35	1.74	5.89
% that contain lowercase letter	20.98	61.04	67.30	49.14
% that contain number	16.29	33.36	29.20	39.82
% that contain uppercase letter	41.21	1.25	1.75	5.15

Table 3 lists the password strength score for different lengths. Table 3 and Fig.10 show that the longer the password length, the higher the score. The score 86.40 of 8 characters in G1 is larger than the score 42.65, 37.90, and 68.22 of 12 characters in G2, G3, and G4, respectively. As can be seen from the table, the passwords in G1 using the proposed scheme shows the best performance when the average strength score is 95.31 and the complexity is 'Very Strong'.

Despite of 'Very Strong', the use of the proposed bio-password such as those in G1 is very convenient. There is no typing and remembering user's password of the target site. When the user simply touches the password input box and takes a face with the camera in smart device stored *CIS*, the password

Table 3: Strength score with different length

Password lists	**G1**	**G2**	**G3**	**G4**
less than 7	62.07	29.18	22.57	40.36
8	86.40	35.39	31.54	49.03
9	94.26	38.06	31.80	52.97
10	97.97	44.64	40.59	60.19
11	99.34	42.10	42.83	58.37
12	99.81	42.65	37.90	68.22
over 13	99.90	50.94	37.22	87.00
Average	95.31	36.83	31.96	57.03
Complexity	Very Strong	Weak	Weak	Good

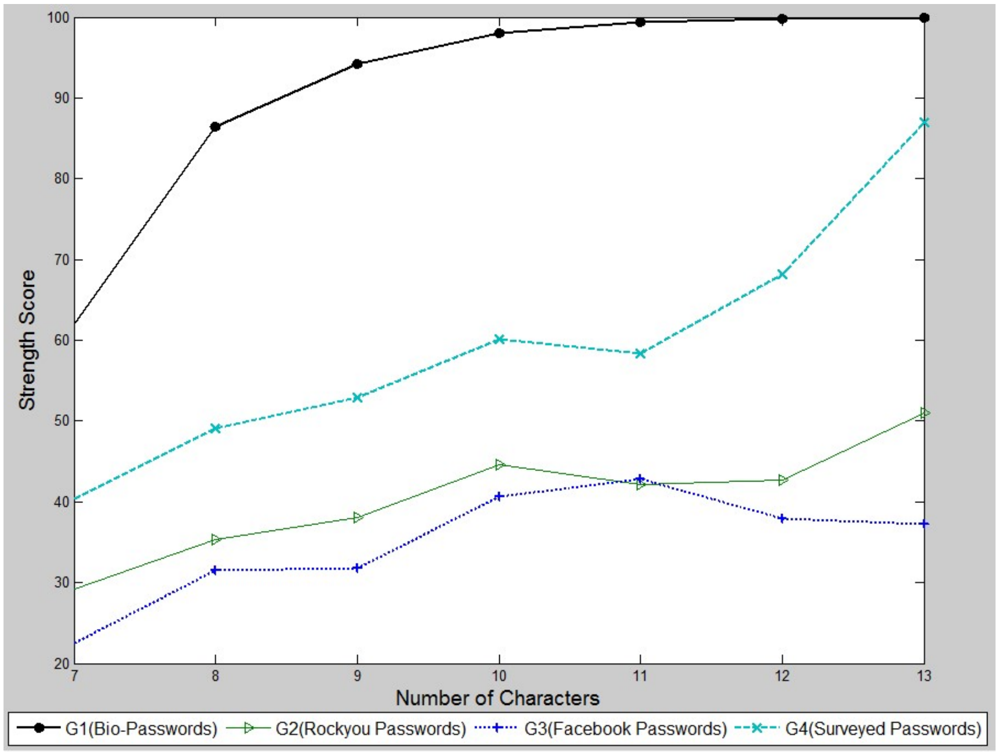


Figure 10: Strength score on each password group with Different Length

of the target site is generated and input by the proposed scheme. After generating it, the used face image is deleted. Moreover, the user can create and use passwords as many as registered sites or accounts. Given user's password $pwd = \{pw_1, ..., pw_l\}$, $pw_i \neq pw_j$ with our proposed scheme. This means there is no repetitive use of passwords. Resolving the problem of repetitive use of passwords improves the security of the password system incomparably.

5 Conclusion

Ginven the results in Sec. 4, the trade-off between security and usablity of passwords, and the recent interest into the next generation authentication, the bio-password-based password system is worth investigating more. Futhermore, today's smart devices have a variety of built-in biometric sensors, as good face and fingerprint sensors in them are common for user authentication and iris sensors have also begun to be equipted. We designed and implemented the bio-password scheme which uses face and confidence interval sets to improve the security and usability of the log-in process. The proposed scheme offers user-friendly solution to generating strong and easy-to-use passwords. The results of our experiments show that the average strength is 'Very Strong' while other password lists are 'Weak' or 'Good'. The bio-passwords with 'Very Strong' strength are also robust against a dictionalry attack, a brute force attack, and attacks by any other password guessing or observation methods.

Acknowledgments

The authors thank Jiyoung Ahn for assistance with test of password strength and Kyunghyun Min for survey of passwords. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT Future Planning (NRF-2014R1A1A1005490).

References

[1] A. Adams, M. A. Sasse, and P. Lunt. Making passwords secure and usable. In *People and Computers XII - Proc. of the 7th International Conference on Human-Computer Interaction (HCI'97), San Francisco, California, USA*, volume 1, pages 1–19. Springer London, August 1997.

[2] K. Alghathbar and H. A. Mahmoud. Noisy password scheme: A new one time password system. In *Proc. of the 22th Canadian Conference on Electrical and Computer Engineering (CCECE'09), St.John's NL, Cananda*, pages 841–846. IEEE, May 2009.

[3] J. Bonneau. The science of guessing: analyzing an anonymized corpus of 70 million passwords. In *Proc. of the 33th IEEE Symposium on Security and Privacy (SP'12), San Francisco, USA*, pages 538–552. IEEE, May 2012.

[4] C. Braz and J.-M. Robert. Security and usability: the case of the user authentication methods. In *Proc. of the 18th Conference on l'Interaction Homme-Machine (IHM'06), Montreal, Canada*, pages 199–203. ACM, April 2006.

[5] W. E. Burr, D. F. Dodson, E. M. Newton, R. A. Perlner, W. T. Polk, S. Gupta, and E. A. Nabbus. Electronic authentication guideline. Technical Report SP 800-63-1, National Institute of Standards & Technology, December 2011.

[6] L. Danielle. Global information assurance certification paper: Introduction to dsniff. *SANS Institute, USA*.

[7] S. Egelman. My profile is my password, verify me!: the privacy/convenience tradeoff of facebook connect. In *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'13), Paris, France*, pages 2369–2378. ACM, April 2013.

[8] S. Egelman, A. Sotirakopoulos, I. Muslukhov, K. Beznosov, and C. Herley. Does my password go up to eleven?: the impact of password meters on password selection. In *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'13), Paris, France*, pages 2379–2388. ACM, April 2013.

[9] K. Fu, E. Sit, K. Smith, and N. Feamster. The dos and don'ts of client authentication on the web. In *Proc. of the 10th USENIX Security Symposium, Washington, USA*, pages 251–268. USENIX, August 2001.

[10] S. Furnell and N. Bär. Essential lessons still not learned? examining the password practices of end-users and service providers. In *Proc. of the 1st International Conference on Human Aspects of Information Security, Privacy, and Trust (HAS'13), Las Vegas, Nevada, USA*, volume 8030 of *Lecture Notes in Computer Science*, pages 217–225. Springer Berlin Heidelberg, July 2013.

[11] A. Kim and S.-H. Lee. A scheme for predicting recognition performance by using confidence intervals. *IEICE Electronics Express*, 9(3):133–139, February 2012.

[12] A. Kim, G. Park, Y. Shin, and S.-H. Lee. A non-memorization bio-password scheme with confidence interval sets. In *Proc. of the 4th International Conference on ICT Convergence (ICTC'13), Jeju Island, Korea*, pages 1102–1106. IEEE, October 2013.

[13] M. Lei, Y. Xiao, S. V. Vrbsky, C.-C. Li, and L. Liu. A virtual password scheme to protect passwords. In *Proc. of the 11th IEEE International Conference on Communications (ICC'08), Beijing, China*, pages 1536–1540. IEEE, May 2008.

[14] A. M. Martinez. The AR face database. Technical Report 24, Centre de Visió per Computador (CVC), June 1998.

[15] K. Min and G. Han. Design of a password management application for ensuring password strength consistency and controlling multiple-use of passwords. In *Proc. of the 12th International Conference on Multimedia Information Technology and Applications (MITA'16), Luang Prabang, Lao PDR*. The Korea Multimedia Society, July 2006.

[16] P. Poornachandran, M. Nithun, S. Pal, A. Ashok, and A. Ajayan. Password reuse behavior: How massive online data breaches impacts personal data in web. In *Innovations in Computer Science and Engineering*, volume 413 of *Advances in Intelligent Systems and Computing*, pages 199–210. Springer, Singapore, 2016.

[17] M. Raza, M. Iqbal, M. Sharif, and W. Haider. A survey of password attacks and comparative analysis on methods for secure authentication. *World Applied Sciences Journal*, 19(4):439–444, April 2012.

[18] W. C. Summers and E. Bosworth. Password policy: the good, the bad, and the ugly. In *Proc. of the 2004 Winter International Symposium on Information and Communication Technologies (WISICT'04), Cancun, Mexico*, pages 1–6. Trinity College Dublin, January 2004.

[19] J. Thorpe, B. MacRae, and A. Salehi-Abari. Usability and security evaluation of geopass: a geographic location-password scheme. In *Proc. of the 9th symposium on usable privacy and security (SOUPS'13), Newcastle, UK*, page 14. ACM, July 2013.

[20] B. Ur, P. G. Kelley, S. Komanduri, J. Lee, M. Maass, M. L. Mazurek, T. Passaro, R. Shay, T. Vidas, L. Bauer, et al. How does your password measure up? the effect of strength meters on password creation. In *Proc. of the 21st USENIX Security Symposium, Washington, USA*, pages 65–80, August 2012.

[21] A. Vance. If your password is 123456, just make it hackme. *The New York Times*, 20:A1–A1, January 2010.

[22] J. E. Weber, D. Guster, P. Safonov, and M. B. Schmidt. Weak password security: An empirical study. *Information Security Journal: A Global Perspective*, 17(1):45–54, March 2008.

[23] M. Weir, S. Aggarwal, M. Collins, and H. Stern. Testing metrics for password creation policies by attacking large sets of revealed passwords. In *Proc. of the 17th ACM conference on Computer and communications security (CCS'10), Chicago, USA*, pages 162–175. ACM, October 2010.

[24] K.-P. Yee and K. Sitaker. Passpet: convenient password management and phishing protection. In *Proc. of the 2nd symposium on Usable privacy and security (SOUPS'06), Pittsburgh, USA*, pages 32–43. ACM, July 2006.

[25] X. Yu and Q. Liao. User password repetitive patterns analysis and visualization. *Information & Computer Security*, 24(1):93–115, March 2016.

Author Biography

Aeyoung Kim received the B.Sc.(2000) in computer science and statistics from Hanshin University, and the M.Sc.(2003) and Ph.D.(2012) in computer science and engineering from Ewha Womans University. She works at National Istitute for Mathematical Sciences in Korea. Her research interests include bio-cryptographic algorithms, lightweight cryptographic algorithms, privacy of big data, authentication for IoT.

Geunshik Han received the B.S.(1984) in statistics from Korea University and the M.S.(1990) in statistics from Iowa state university and Ph.D.(1993) in statistics from Oklahoma state university. He is a professor at Hanshin University, Osan, Korea since 1994. His main research interests include pattern recognition, statistical computations and sampling designs.

Seung-Hyun Seo received her B.S.(2000), M.S.(2002), and Ph.D.(2006) from Ewha Womans University in Korea. She is an assistant professor at Korea University sejong campus since 2015. Before that, she was a post-doctoral researcher of Computer Science at Purdue University for 2 and half years, a senior researcher of KISA (Korea Internet and Security Agency) for 2 years and a researcher for 3 years in FSA (Financial Security Agency), Korea. Her main research interests include cryptography, IoT security, mobile security, secure cloud computing, and malicious code analysis.