8-3-2006

A Design and Analysis of Graphical Password

Xiaoyuan Suo

A DESIGN AND ANALYSIS OF GRAPHICAL PASSWORD

by

Xiaoyuan Suo

Under the Direction of Ying Zhu

ABSTRACT

The most common computer authentication method is to use alphanumerical usernames and passwords. This method has been shown to have significant drawbacks. For example, users tend to pick passwords that can be easily guessed. On the other hand, if a password is hard to guess, then it is often hard to remember. To address this problem, some researchers have developed authentication methods that use pictures as passwords. In this paper, I conduct a comprehensive survey of the existing graphical password techniques. I classify these techniques into two categories: recognition-based and recall-based approaches. I discuss the strengths and limitations of each method and point out the future research directions in this area.

I also developed three new techniques against the common problem exists in the present graphical password techniques. In this thesis, the scheme of each new technique will be proposed; the advantages of each technique will be discussed; and the future work will be anticipated.

INDEX WORDS: Graphical password, Authentication, Security, Password space, Survey

A DESIGN AND ANALYSIS OF GRAPHICAL PASSWORD

By

XIAOYUAN SUO

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2006

A DESIGN AND ANALYSIS OF GRAPHICAL PASSWORD

by

XIAOYUAN SUO

Major Professor: Ying Zhu
Committee: Scott Owen
Raheem Beyah

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
August 2006

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER1: INTRODUCTION

Human factors are often considered the weakest link in a computer security system. Patrick, et al. [1] point out that there are three major areas where human-computer interaction is important: authentication, security operations, and developing secure systems. Here we focus on the authentication problem.

Current authentication methods can be divided into three main areas:

- Token based authentication
- Biometric based authentication
- Knowledge based authentication
    - Text based authentication
    - Picture based authentication

Token based techniques, such as key cards, bank cards and smart cards are widely used. Many token-based authentication systems also use knowledge based techniques to enhance security. For example, ATM cards are generally used together with a PIN number.

Biometric based authentication techniques, such as fingerprints, iris scan, or facial recognition, are not yet widely adopted. The major drawback of this approach is that such systems can be expensive, and the identification process can be slow and often unreliable. However, this type of technique provides the highest level of security.

Knowledge based techniques are the most widely used authentication techniques and include both text-based and picture-based passwords. The picture-based techniques can be further divided into two categories: recognition-based and recall-based graphical techniques. Using recognition-based techniques, a user is presented with a set of images

and the user passes the authentication by recognizing and identifying the images he or she selected during the registration stage. Using recall-based techniques, a user is asked to reproduce something that he or she created or selected earlier during the registration stage.

The most common knowledge based authentication method is for a user to submit a user name and a text password. The vulnerabilities of this method have been well known. One of the main problems is the difficulty of remembering passwords. Studies have shown that users tend to pick short passwords or passwords that are easy to remember [2]. Unfortunately, these passwords can also be easily guessed or broken. According to a recent Computerworld news article, the security team at a large company ran a network password cracker and within 30 seconds, they identified about 80% of the passwords [3]. On the other hand, passwords that are hard to guess or break are often hard to remember. Thus a large portion of customer service calls are related to one's forgetting his or her password. Studies showed that since user can only remember a limited number of passwords, they tend to write them down or will use the same passwords for different accounts [4, 5].

Recently security researchers have detected a rise in the spread of Keylogger [6], a spyware built to capture login names and passwords and to send them to the attackers. Text-based passwords are particularly vulnerable to such attacks.

To address the problems with traditional username-password authentication, alternative authentication methods, such as biometrics [3, 7], have been used. In this paper, however, we will focus on another non-traditional authentication method: using pictures as passwords.

The primary goal of improving the current user authentication technology is to make the method secure yet easier for the user. Graphical password schemes have been proposed as a possible alternative to text-based schemes, motivated particularly by the fact that humans can remember pictures better than text. Psychological studies have shown that people can remember pictures better than text [8]. Pictures are generally easier to be remembered or recognized than text, especially photos, which are even easier to be remembered than random pictures.

It has also been suggested that graphical passwords may be hard to guess or broken by brute force search. If the number of possible pictures is sufficiently large, the possible password space of a graphical password scheme may exceed that of text-based schemes and thus presumably offer better resistance to dictionary attacks. Because of these (presumed) advantages, there is a growing interest in graphical password. In addition to workstation and web log-in applications, graphical passwords have also been applied to ATM machines and mobile devices.

In this thesis, I will first conduct a survey of the existing graphical password techniques. I will discuss the strengths and limitations of each method and also point out future research directions in this area. In conducting this survey, I want to answer the following questions:

- Are graphical passwords as secure as text passwords?
- What are the major design and implementation issues for graphical passwords?

I will then propose three different new techniques against the commonly seeing problems in graphical password area. RAF, or recall a formation, will allow the user to

choose from a set of images to be placed on a 8×8 grid; if both the formation and images are correctly placed, the user will be authenticated. The second algorithm is a neural network based approach. It authenticates the user by user's daily mouse motion. The third method is a shoulder surfing resistant passpoint; it overcomes shoulder surfing problem the passpoint scheme has. Details are as follows:

CHAPTER2: A SURVEY OF GRAPHICAL PASSWORDS

2.1 Recognition based techniques

In recognition based techniques, users are given a set of pictures and they pick and memorize some of them. During authentication, the users need to recognize and identify the pictures they have picked earlier.

Dhamija and Perrig [4] proposed an graphical authentication scheme based on Hash Visualization technique [9]. In their system, user will be asked to select certain number of images from a set of random pictures generated by a program (figure 1). Later, user will be required to identify the pre-selected images to be authenticated. The results showed that 90% of all participants succeeded in the authentication using their technique, while only 70% succeeded using text-based passwords and PINS. The average log-in time, however, is longer than the traditional approach, but has a much smaller failure rate. A drawback is that the server needs to store a large amount of pictures which may have to be transferred over the network, delaying the authentication process. Another weakness of this system is that the server needs to store the seeds of the portfolio images of each user in plain text. Interface-wise, the process of selecting a picture from picture database can be tedious and time consuming for the user.

In Akula and Devisetty's algorithm [10], the system displays a set of images to the user and the user would then select the correct pass-image. The basic scheme is similar to the technique proposed by Dhamija and Perrig [4]. The difference is that this technique uses the hash function SHA-1, which produces a 20 byte output. This makes the authentication secure and requires less memory. However, an image file still occupies more space than text even after hashing. The authors suggested a possible future

improvement by providing the persistent storage and this could be deployed on the Internet, cell phones and PDA's.

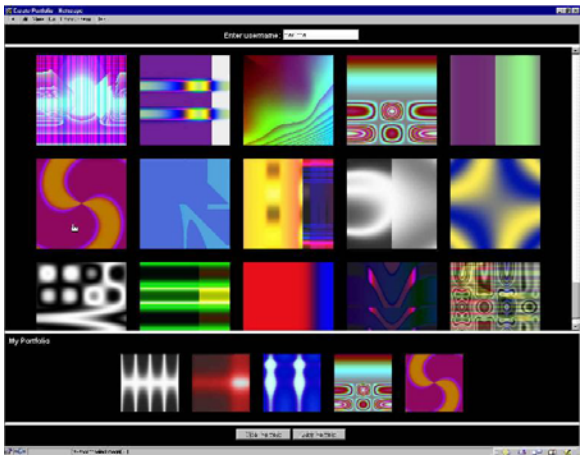

Figure 11. Random arts used by Dhamija and Perrig [4]

Weinshall and Kirkpatrick [11] identified a wide range of human memory phenomena as potential certificates of identity. They sketched several authentication schemes, such as picture recognition, object recognition, and pseudo word recognition, and conducted a number of user studies. In the picture recognition study, a user is trained to recognize a large set of images (100 – 200 images) selected from a database of 20,000 images. After one to three months, users in their study were able to recognize over 90% of the images in the training set. This study showed that pictures are the most effective among the three schemes tested. Pseudo codes can also be used, but require proper setting and training.

Sobrado and Birget [12] developed a graphical password technique that deals with shoulder-surfing problem. In the first scheme, the system will display a number of pass-objects (pre-selected by user) among many other objects. To be authenticated, a user needs to recognize pass-objects and click inside the convex hull formed by all the pass-objects (figure 2). In order to make the password hard to guess, Sobrado and Birget

suggested using 1000 objects, which making the display very crowded and the objects almost indistinguishable. On the other hand, using fewer objects may lead to a smaller password space, since the resulting convex hull can be large. In their second algorithm, a user moves a frame (and the objects within it) until the pass object on the frame lines up with the other two pass-objects. The authors also suggest repeating the process for a few more times to minimize the likelihood of logging in by randomly clicking or rotating. The main drawback of these algorithms is that the log in process can be slow.

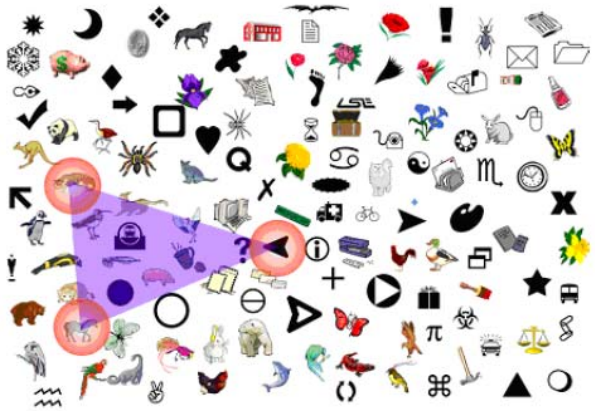

Figure 12. A shoulder-surfing resistant graphical password scheme. (Source: Sobrado and Birget [12])

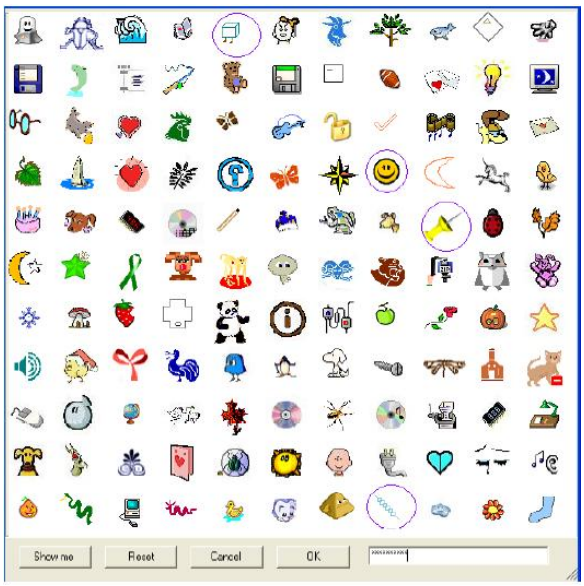

Figure 13. Another shoulder surfing resistant scheme developed by Hong, et al. [13]. The pass-string is 99dc8151up

Man, et al. [14] proposed another shoulder-surfing resistant algorithm. In this algorithm, a user selects a number of pictures as pass-objects. Each pass-object has several variants and each variant is assigned a unique code. During authentication, the user is challenged with several scenes. Each scene contains several pass-objects (each in the form of a randomly chosen variant) and many decoy-objects. The user has to type in a string with the unique codes corresponding to the pass-object variants present in the scene as well as a code indicating the relative location of the pass-objects in reference to a pair of eyes. The argument is that it is very hard to crack this kind of password even if the whole authentication process is recorded on video because where is no mouse click to give away the pass-object information. However, this method still requires users to memorize the alphanumeric code for each pass-object variant. For example, if there are 4 pictures each with 4 variants, then each user has to memorize 16 codes. Although the pass-objects provide some cues for recalling the codes, it is still quite inconvenient.

Hong, et al. [13] later extended this approach to allow user to assign their own codes to pass-object variants. Figure 3 shows the log-in screen of this graphical password scheme. However, this method still forces user to memorize many text strings and therefore suffer from the many drawbacks of text-based passwords.

Figure 14. An example of Passfaces (source: www.realuser.com)

"Passface" is a technique developed by Real User Corporation [15]. The basic idea is as follows. The user will be asked to choose four images of human faces from a face database as their future password. In the authentication stage, the user sees a grid of nine faces, consisting of one face previously chosen by the user and eight decoy faces (figure 4). The user recognizes and clicks anywhere on the known face. This procedure is repeated for several rounds. The user is authenticated if he/she correctly identifies the four faces. The technique is based on the assumption that people can recall human faces easier than other pictures. User studies by Valentine [16, 17] have shown that Passfaces are very memorable over long intervals. Comparative studies conducted by Brostoff and Sasse [18] showed that Passfaces had only a third of the login failure rate of text-based passwords, despite with about a third the frequency of use. Their study also showed that Passface-based log–in process took longer than text passwords and therefore were used less frequently by users. Although the preliminary user studies have shown some promising results for the Passface technique, the effectiveness of this method is still

uncertain. Davis, et al. [19] studied the graphical passwords created using Passface technique and found obvious patterns among these passwords. For example, most users tend to choose faces of people from the same race. In their study, female faces were preferred by both male and female users. Better looking faces were more likely to be chosen. All of these make the Passface password quite predictable. This problem may be alleviated by arbitrarily assigning faces to users, but doing so would make it hard for people to remember the password.

Jansen et al. [20-22] proposed graphical password mechanism for mobile devices. During enrollment stage, a user selects a theme (e.g. sea, cat, etc.) which consists of thumbnail photos and then registers a sequence of images as a password (figure 5). During the authentication, the user must enter the registered images in the correct sequence. After a successful authentication, the user may change the password, selecting a new sequence, or possibly change the theme. One drawback of this technique is that while the amount of thumbnail image is limited to 30, the password space is small. Each thumbnail image is assigned a numerical value, and the sequence of selection will essentially generate a numerical password. The result showed that the image sequence length was generally shorter than the textural password length. To address this problem, two pictures can be combined to compose a new alphabet element, thus expands the image alphabet size.

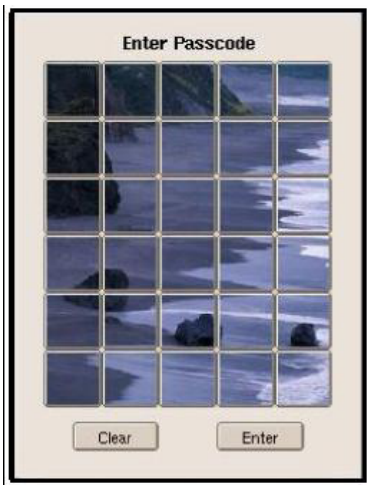

Figure 15. A graphical password scheme proposed by Jansen, et al. [20]

Takada and Koike discussed a similar graphical password technique for mobile devices. This technique allows users to use their favorite image for authentication [23]. The users first register their favorite images (pass-images) with the server. During authentication, a user has to go through several rounds of verification. At each round, the user either selects a pass-image among several decoy-images or chooses nothing if no pass-image is present. The program would authorize a user only if all verifications are successful. Allowing users to register their own images makes it easier for user to remember their pass-images. A notification mechanism is also implemented to notify users when new images are registered in order to prevent unauthorized image registration. This method does not necessarily make it a more secure authentication method than text-based password. As shown in the studies by Davis [19], users' choices of picture passwords are often predictable. Allowing users to use their own pictures would make the password even more predictable, especially if the attacker is familiar with the user.

2.2Recall based techniques

In this section we discuss two types of picture password techniques: reproducing a drawing and repeating a selection.

2.2.1Reproduce a drawing

Jermyn, et al. [24] proposed a technique, called “Draw - a - secret (DAS)”, which allows user to draw their unique password (figure 6). A user is asked to draw a simple picture on a 2D grid. The coordinates of the grids occupied by the picture are stored in the order of the drawing. During authentication, the user is asked to re-draw the picture. If the drawing touches the same grids in the same sequence, then the user is authenticated. Jermyn, et al. suggested that given reasonable-length passwords in a 5 X 5 grid, the full password space of DAS is larger than that of the full text password space.

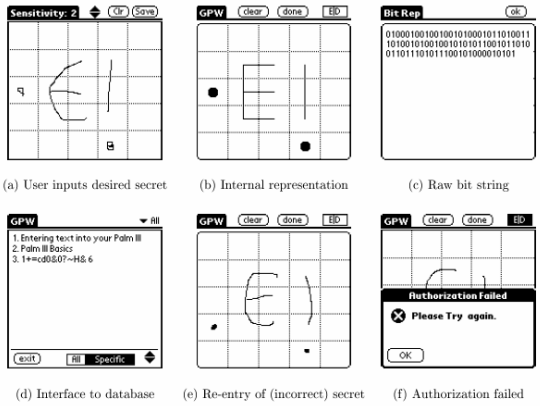

Figure 16. Draw-a-Secret (DAS) technique proposed by Jermyn, et al. [24]

Thorpe and van Oorschot [25] analyzed the memorable password space of the graphical password scheme by Jermyn et al. [24]. They introduced the concept of graphical dictionaries and studied the possibility of a brute-force attack using such dictionaries. They defined a length parameter for the DAS type graphical passwords and showed that DAS passwords of length 8 or larger on a 5 x 5 grid may be less susceptible

to dictionary attack than textual passwords. They also showed that the space of mirror symmetric graphical passwords is significantly smaller than the full DAS password space. Since people recall symmetric images better than asymmetric images, it is expected that a significant fraction of users will choose mirror symmetric passwords. If so, then the security of the DAS scheme may be substantially lower than originally believed. This problem can be resolved by using longer passwords. Thorpe and van Oorschot showed that the size of the space of mirror symmetric passwords of length about L + 5 exceeds that of the full password space for corresponding length L <= 14 on a 5 x 5 grid.

Thorpe and van Oorschot [26] further studied the impact of password length and stroke-count as a complexity property of DAS scheme. Their study showed that stroke-count has the largest impact on the DAS password space -- The size of DAS password space decreases significantly with fewer strokes for a fixed password length. The length of DAS password also has a significant impact but the impact is not as strong as the stroke-count. To improve the security, Thorpe and van Oorschot proposed a "Grid Selection" technique. Selection grid is an initially large, fine grained grid from which the user selects a *drawing grid*, a rectangular region to zoom in on, in which they may enter their password (figure 7). This would significantly increase the DAS password space.

Goldberg et al. [27] did a user study in which they used on a technique called "Passdoodle". This is a graphical password comprised of handwritten designs or text, usually drawn with a stylus onto a touch sensitive screen. Their study concluded that users were able to remember complete doodle images as accurately as alphanumeric passwords. The user studies also showed that people are less likely to recall the order in

which they drew a DAS password. The work nevertheless provided useful information in terms of graphical password as a possible alternative for text password. However, the user study was done using paper prototype instead of computer programs, verifications were done by human rather than computer. Therefore the accuracy of this study is still uncertain.

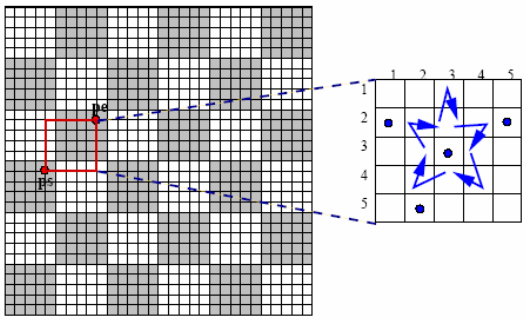

Figure 17. Grid selection: user selects a drawing grid. (Source: Thorpe and van Oorschot [28])

Nali and Thorpe [29] conducted further analysis of the "Draw-A-Secret (DAS)" scheme [24]. In their study, users were asked to draw a DAS password on paper in order to determine if there are predictable characteristics in the graphical passwords that people choose. The study did not find any predictability in the start and end points for DAS password strokes, but found that certain symmetries (e.g. crosses and rectangles), letters, and numbers were common. This study showed that users choose graphical passwords with predictable characteristics, particularly those proposed as "memorable". If this study is indicative of the population, the probability in which some of these characteristics occur would reduce the entropy of the DAS password space. However, this user study only asked the users to draw a memorable password, but did not do any recall-test on whether or not the passwords were really memorable.

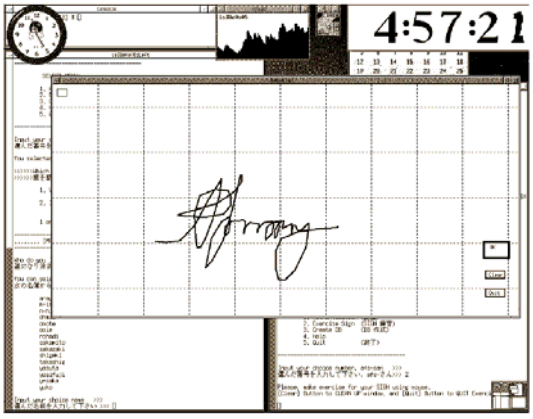

Figure 18. A signature is drawn by mouse. (Source Syukri, et al. [30])

Syukri, et al. [30] proposes a system where authentication is conducted by having user drawing their signature using mouse (figure 8). Their technique included two stages, registration and verification. During the registration stage: user will first be asked to draw their signature with mouse, and then the system will extract the signature area and either enlarge or scale-down signatures, rotates if needed, (also known as normalizing). The information will later be saved into the database. The verification stage first takes the user input, and does the normalization again, and then extracts the parameters of the signature. After that, the system conducts verification using geometric average means and a dynamic update of database. According to the paper, the rate of successful verification was satisfying. The biggest advantage of this approach is that there is no need to memorize one's signature and signatures are hard to fake. However, not everybody is familiar with using mouse as a writing device; the signature can therefore be hard to drawn. One possible solution to this problem would be to use a pen-like input device, but such devices are not widely used, and adding new hardware to the current system can be expensive. We believe such technique is more useful to small devices such as PDA.

2.2.2 Repeat a sequence of actions

In this group of authentication algorithms, a user is asked to repeat a sequence of actions originally conducted by the user during the registration stage.

Blonder [31] designed a graphical password scheme in which a password is created by having the user click on several locations on an image. During authentication, the user must click on the approximate areas of those locations. The image can assist users to recall their passwords and therefore this method is considered more convenient than unassisted recall (as in text-based password). Passlogix [32] has developed a graphical password system based on this idea. In their implementation (figure 9), users must click on various items in the image in the correct sequence in order to be authenticated. Invisible boundaries are defined for each item in order to detect whether an item is clicked by mouse. A similar technique has been developed by *sfr* [33]. It was reported that Microsoft had also developed a similar graphical password technique where users are required to click on pre-selected areas of an image in a designated sequence [34]. But details of this technique have not been available.

The "PassPoint" system by Wiedenbeck, et al. [35-37] extended Blonder's idea by eliminating the predefined boundaries and allowing arbitrary images to be used. As a result, a user can click on any place on an image (as opposed to some pre-defined areas) to create a password. A tolerance around each chosen pixel is calculated. In order to be authenticated, the user must click within the tolerance of their chosen pixels and also in the correct sequence (figure 10). This technique is based on the discretization method proposed by Birget, et al. [38]. Because any picture can be used and because a picture may contain hundreds to thousands of memorable points, the possible password space is

quite large. Wiedenbeck, et al. conducted a user study [37], in which one group of participants were asked to use alphanumerical password, while the other group was asked to use the graphical password. The result showed that graphical password took fewer attempts for the user than alphanumerical passwords. However, graphical password users had more difficulties learning the password, and took more time to input their passwords than the alphanumerical users.

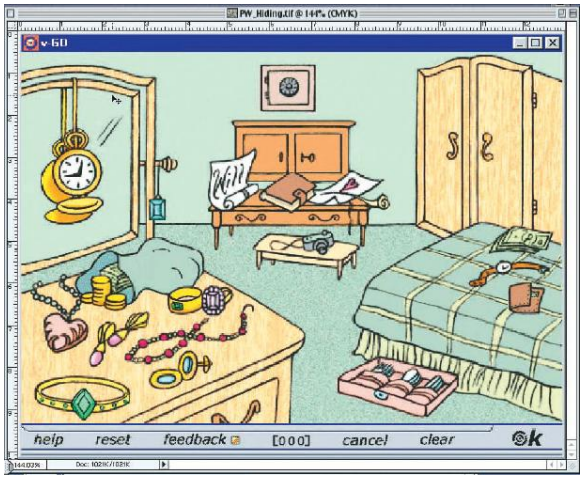

Figure 19. A recall-based technique developed by Passlogix. (Source: Paulson [34])

Figure 20. An image used in Passpoint system. (Source: Wiedenbeck, et al. [35])

Later Wiedenbeck, et al. [36] also conducted a user study to evaluate the effect of tolerance of clicking during the re-authenticating stage, and the effect of image choice in

the system. The result showed that memory accuracy for the graphical password is strongly reduced after using smaller tolerance for the user clicked points, but the choices of images do not make a significant difference. The result showed that the system works for a large variety of images

Passlogix [32] has also developed several graphical password techniques based on repeating a sequence of actions. For example, its v-Go includes a graphical password scheme where users can mix up a virtual cocktail and use the combination of ingredients as password. Other password options include picking a hard at cards or putting together a "meal" in the virtual kitchen. However, this technique only provides a limited password space and there is no easy way to prevent people from picking poor passwords (for example, a full house of cards).

Adrian Perrig was also reported to be working on a system (called Map Authentication) that was based on navigating through a virtual world [34]. In this system, users can build their own virtual world. The authentication is carried out by having users navigate to a site that is randomly chosen each time they log on. However, the details of this system are not available.

CHAPTER3: ANALYSIS OF GRAPHICAL PASSWORD

After a survey of existing graphical password techniques, we try to answer the questions we proposed at the end of section 1.

3.1 A taxonomy for graphical password

Techniques	**Usability**		**Security issues**	
	Authentication process	Memorability	*Password space*	*Possible attack methods*
Text-based password	Type in password, can be very fast	Depends on the password. Long and random passwords are hard to remember	94^K (there are 94 printable characters excluding SPACE, N is the length of the password).The actual password space is usually much smaller.	Dictionary attack, brute force search, guess, spyware, shoulder surfing, etc.
Perrig and Song [9]	Pick several pictures out of many choices. Takes longer to create than text password	Limited user study showed that more people remembered pictures than text-based passwords	N!/K!(N-K)! (N is the total number of pictures; K is the number of pictures in the graphical password)	Brute force search, guess, shoulder-surfing
Sobrado and Birget [12]	Click within an area bounded by pre-registered picture objects, can be very fast	Can be hard to remember when large numbers of objects are involved.	N!/K!(N-K)! (N is the total number of picture objects; K is the number of pre-registered objects)	Brute force search, guess
Man, et al. [14] Hong, et al. [13]	Type in the code of pre-registered picture objects; can be very fast	Users have to memorize both picture objects and their codes. More difficult than text-based password	Same as the text based password	Brute force search, spyware
Passface [15]	Recognize and pick the pre-registered pictures; takes longer than text-based password	Faces are easier to remember, but the choices are still predictable	N^K (K is the number of rounds of authentication, N is the total number of pictures at each round)	Dictionary attack, brute force search, guess, shoulder surfing
Jansen et al. [20-22]	User register a sequence of images; slower than text-based password	Pictures are organized according to different themes to help users remember	N^K (N is the total number of pictures, K is the number of pictures in the graphical password. N is small due the size limit of mobile devices)	Brute force search, guess, shoulder surfing
Takada and Koike [23]	Recognize and click on the pre-registered images; slower than text-based password. Slower than text-based password	Users can use their favorite images; easy to remember than system assigned pictures	(N+1)^K (K is the number of rounds of authentication, N is the total number of pictures at each round)	Brute force search, guess, shoulder surfing

Table 1. Comparison of major graphical password techniques

Jermyn, et al. [24], Thorpe and van Oorschot [25-26]	Users draw something on a 2D grid	Depends on what users draw. User studies showed the drawing sequence is hard to remember	Password space is larger than text based password. But the size of DAS password space decreases significantly with fewer strokes for a fixed password length	Dictionary attack, shoulder surfing
Syukri, et al. [30]	Draw signatures using mouse. Need a reliable signature recognition program.	Very easy to remember, but hard to recognize	Infinite password space	Guess, dictionary attack, shoulder surfing
Goldberg et al. [27]	Draw something with a stylus onto a touch sensitive screen	Depends on what users draw	Infinite password space	Guess, dictionary attack, shoulder surfing
Blonder [31], Passlogix [32], [33], [34] , Wiedenbeck, et al. [35-37]	Click on several pre-registered locations of a picture in the right sequence.	Can be hard to remember	N^K (N is the number of pixels or smallest units of a picture, K is the number of locations to be clicked on)	Guess, brute force search, shoulder surfing

3.2 Major factors in evaluating graphical passwords

There is still no clear answer to this question. Many user studies in our survey have confirmed that people can recall graphical password more reliably than text-based password over a long period of time. This seems to be the main advantage of graphical passwords. Some graphical password techniques have been shown to provide a password space similar to or larger than that of text-based passwords [24, 25, 38]. Although some research exists in the field, very little research has been done to study the actual difficulty of cracking graphical passwords. There is little study on the possible techniques for breaking graphical passwords. As a result, there is still no concrete evidence to prove whether graphical password in general is more or less secure than text-based password. This question has to be answered on a case by case basis, depending on specific algorithms and implementations. Recognition based graphical passwords tend to have smaller password spaces than the recall based methods, and therefore seem more vulnerable to attacks. However, there has been no study to compare the level of security between recognition-based methods and recall-based methods. In addition, studies on the

Passface technique have shown that people often choose weak and predictable graphical passwords [19], a serious problem typically associated with text-based passwords. Nali and Thorpe's study [29] revealed similar predictability among the graphical passwords created with DAS technique [24]. Much more research efforts are needed to understand the nature of graphical passwords created by real world users. This information is very important for evaluating the level of security of graphical passwords.

There are many aspects of security issues, brutal force attack (large password space can solve this problem), shoulder surfing problems (there are existing works discussed in section three which take care of it), dictionary attack, key log (A Keylogger is a program that runs in the background, recording all the keystrokes. Once keystrokes are logged, they are hidden in the machine for later retrieval, or shipped raw to the attacker. The attacker then peruses them carefully in the hopes of either finding passwords, or possibly other useful information that could be used to compromise the system or be used in a social engineering attack), easy to guess (we need to prevent people from choosing simple and weak password, log in spoofing (most of the existing graphical password techniques can typically prevent this problem).

Typical graphical password schemes require large system storage, even with the usage of existing picture hash techniques, the storage is still larger than text based password. it is probably impossible to require graphical password to occupy as little database as the text-based password, but it is certainly necessary to minimizing the current database.

Encryption and transferring over the internet are another two issues that remain un-discussed among all the works we encountered. Traditional text-based password can

be encrypted into a string while transferring, but if pictures are encrypted into a string as well, it then reveals no advantage against text based password. So the question remains to be how to encode the graphical password in reality.

Very little research has been done to study the difficulty of cracking graphical passwords. Because graphical passwords are not widely used in practice, there is no report on real cases of breaking graphical passwords. Here we briefly exam some of the possible techniques for breaking graphical passwords and try to do a comparison with text-based passwords.

3.3 What are the major design and implementation challenges of graphical passwords?

The main design issue for recognition based techniques is how to make it easier for users to remember and recognize the images. A number of techniques have been proposed, such as grouping images by theme, using human face images, or allowing users to register their own images.

The major design issue for recall-based methods is the reliability and accuracy of user input recognition. This is a typical pattern recognition problem. In this type of methods, the tolerances of error have to be set carefully – overly high tolerances may lead to lots of false positives while overly low tolerances may lead to lots of false negatives.

In the above section, we have briefly examined the security issues with graphical passwords.

Maintaining a large password space is a major design issue for both recognition-based and recall-based methods. A large password space is necessary to defend against guess-based attacks. For recognition-based methods, one solution is to have several

rounds of verifications. But this will make the log-in process longer and tedious. Another solution is to deploy large number of decoy-images. Some proposed methods involve hundreds of decoy-images. This would also slow down the log-in process. In addition, this solution is not suitable for mobile devices due to very limited user interface space. For "reproduce-a-drawing" methods, possible solutions include maintaining a large canvas, reducing the tolerance of error, and requiring users to draw complex pictures. However, this may result in sophisticated and perhaps overly sensitive recognition programs that generate lots of false negatives. For "repeat-a-sequence" methods, the solution is to use a highly detailed image and provide large number of potential click points. Users are also required to click on many points in order to generate a long password. The drawback, however, is that users may have difficulty to memorize the long sequence of clicks.

Shoulder-surfing resistance is an important design consideration. At this point, only a few recognition-based techniques are designed to resist shoulder-surfing. None of the recall-based based techniques are considered should-surfing resistant.

Graphical passwords require much more storage space than text based passwords. Tens of thousands of pictures have to be maintained in a centralized database. Network transfer delay is also a concern for graphical passwords, especially for recognition-based techniques in which hundreds of pictures may need to be displayed for each round of verification.

3.4 Security factors

3.4.1 Brute force search

The main defense against brute force search is to have a sufficiently large password space. Text-based passwords have a password space of 94^N, where N is the length of the password, 94 is the number of printable characters excluding SPACE. Some graphical password techniques have been shown to provide a password space similar to or larger than that of text-based passwords [24 - 27, 30, 38]. Recognition based graphical passwords tend to have smaller password spaces than the recall based methods.

It is more difficult to carry out a brute force attack against graphical passwords than text-based passwords. The attack programs need to automatically generate accurate mouse motion to imitate human input, which is particularly difficult for recall based graphical passwords. Overall, we believe a graphical password is less vulnerable to brute force attacks than a text-based password.

3.4.2 Dictionary attacks

Since recognition based graphical passwords involve mouse input instead of keyboard input, it will be impractical to carry out dictionary attacks against this type of graphical passwords. For some recall based graphical passwords [24][30], it is possible to use a dictionary attack but an automated dictionary attack will be much more complex than a text based dictionary attack. More research is needed in this area. Overall, we believe graphical passwords are less vulnerable to dictionary attacks than text-based passwords.

3.4.3 Guessing

Unfortunately, it seems that graphical passwords are often predictable, a serious problem typically associated with text-based passwords. For example, studies on the Passface technique have shown that people often choose weak and predictable graphical passwords [19]. Nali and Thorpe's study [29] revealed similar predictability among the graphical passwords created with the DAS technique [24]. More research efforts are needed to understand the nature of graphical passwords created by real world users.

3.4.4 Spyware

Except for a few exceptions [13][14], key logging or key listening spyware can not be used to break graphical passwords. It is not clear whether "mouse tracking" spyware will be an effective tool against graphical passwords. However, mouse motion alone is not enough to break graphical passwords. Such information has to be correlated with application information, such as window position and size, as well as timing information.

3.4.5 Shoulder surfing

Like text based passwords, most of the graphical passwords are vulnerable to shoulder surfing. At this point, only a few recognition-based techniques are designed to resist shoulder-surfing [13][14]. None of the recall-based based techniques are considered should-surfing resistant.

3.4.6 Social engineering

Comparing to text based password, it is less convenient for a user to give away graphical passwords to another person. For example, it is very difficult to give away

graphical passwords over the phone. Setting up a phishing web site to obtain graphical passwords would be more time consuming.

Overall, we believe it is more difficult to break graphical passwords using the traditional attack methods like brute force search, dictionary attack, and spyware. There is a need for more in-depth research that investigates possible attack methods against graphical passwords.

3.5 Usability

Finally, the usability of graphical passwords has to be addressed. A major complaint among the users of graphical passwords is that the password registration and log-in process take too long, especially in recognition-based approaches. For example, during the registration stage, a user has to pick images from a large set of selections. During authentication stage, a user has to scan lots of images to identify a few pass-images. For example, in Dhamija and Perrig's system [4], users have to scan through at least 25 images. Users may find this process long and tedious.

One of the main arguments for graphical passwords is that pictures are easier to remember than text strings. Preliminary user studies presented in some research papers seem to support this. However, current user studies are still very limited, involving only a small number of users. We still do not have convincing evidence demonstrating that graphical passwords are easier to remember than text based passwords.

A major complaint among the users of graphical passwords is that the password registration and log-in process take too long, especially in recognition-based approaches. For example, during the registration stage, a user has to pick images from a large set of selections. During authentication stage, a user has to scan many images to identify a few

pass-images. Users may find this process long and tedious. Because of this and also because most users are not familiar with the graphical passwords, they often find graphical passwords less convenient than text based passwords.

3.6 Reliability

The major design issue for recall-based methods is the reliability and accuracy of user input recognition. In this type of method, the error tolerances have to be set carefully – overly high tolerances may lead to many false positives while overly low tolerances may lead to many false negatives. In addition, the more error tolerant the program, the more vulnerable it is to attacks.

3.7 Storage and communication

Graphical passwords require much more storage space than text based passwords. Tens of thousands of pictures may have to be maintained in a centralized database. Network transfer delay is also a concern for graphical passwords, especially for recognition-based techniques in which a large number of pictures may need to be displayed for each round of verification.

CHAPTER4: ANALYSIS OF THE PASSWORD SPACE

In this section, we analyze the two main categories of the graphical password techniques from several perspectives. We focus mainly on security and usability but also take into consideration system and communication issues. For security, we focus on password space and the strength of the password. For usability, we focus on the easiness of registration and easiness of authentication.

4.1 Recognition based techniques

Table 2: variable representations for recognition based techniques

Data Definition	**variable**
Number of pictures in each page	n
Number of scene/rounds of authentication	s
Distraction image in each scene	d
Password length	l

Security. The password space of the recognition based techniques largely depends on the size of the content. Most recognition based techniques do not consider the order of the selection. They often involve many rounds of authentication with users going through several pages of images [5, 8].

So the password space for recognition based technique is a function of total number of pictures:

$$password\_space = f(s \times n)$$

The chances of creating weak password are high in recognition based password. The work by Davis, et al. [9] found obvious patterns among the PassFace password [5]. For example, most users tend to choose faces of people from the same race.

Random art [4] could be one solution toward weak password; in which user has no familiarity to any of the picture password. However doing so may decrease the usability by making the password hard to remember.

Usability. Content (pictures), spatial layout of the content, and input devices are all important factors that influence usability. For example, users' favorite pictures tend to be easier to remember but also easier to be guessed by attackers. Too many distraction pictures tend to slow down the authentication process. Several existing techniques are proven to have usability due to the crowded content arrangement [10].

Other issues. Overly large storage requirement is a significant issue for recognition based techniques, since the size of a typical picture is much larger than the equivalent text. In order to achieve the larger password space, thousands of pictures need to be stored at one time. Sending large number of pictures over the network is also a problem for low speed networks.

4.2 Recall based techniques

Security. It is considerably difficult to calculate the password space of a recall based technique, since there are many variations in recall based techniques. Here we demonstrate a general mathematical model. Some of the most important elements that determine the password space of a recall based technique are listed in table 2.

Table 3: variable representations for recall based techniques

Data Definition	**variable**
Number of pixels in each scene	n
Sequence of password	s
Password length	l

The maximum password space that a recall based technique can have is extremely large, since certain techniques requires the user not only have the proper shape of drawing or clicking, but also the sequence of drawing [6]. If the drawing allows the same pixel be chosen multiple times, the password space for a specific password length *l* is:

$$\max = n^l$$

For *n* pixels on the scene, the total password space is defined as:

$$\sum_{l=1}^{n} n^l$$

If a password scheme does not allow the drawing to pass the same pixel multiple times, or if it requires mouse clicking to match the pre-registered sequence, the password space is the smallest:

$$\min = n$$

In general, the password space for recall based approach is:

$$n \leq space \leq \sum_{l=1}^{n} n^l$$

Usability. A major complaint in recall based graphical passwords is that it is difficult to draw shapes with mouse. Most users are not familiar with using mouse as a drawing tool. However, on mobile devices, stylus pen is a good choice for such techniques.

CHAPTER5: RECALL A FORMATION—RAF

5.1 RAF methodology

As mentioned above, one of the most important factors about a secure graphical password is determined by its password space. Typically, the larger the password space, the more secure the graphical password is. We have therefore developed a new technique: RAF, or a recall a formation. This technique belongs to the recall-based techniques. Using java, we have created a simple interface, in which there are two 8×8 tables: data table and input table. Data table contains the possible choices of icons, there are 4 different themes available for the data table; they are "singleObject", in which there is only one kind of icons, a star (figure 11);

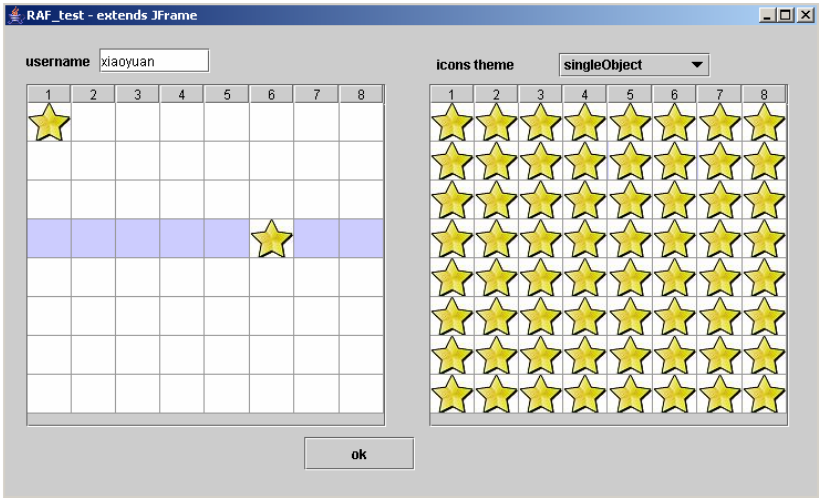

Figure11, RAF interface, user will be asked to choose their desired icon from the right hand menu, and place them into the left hand side table. This is the single Object theme, in which there is only one type of icons.

"animal", in which there are 10 distinctive icons (figure 12);

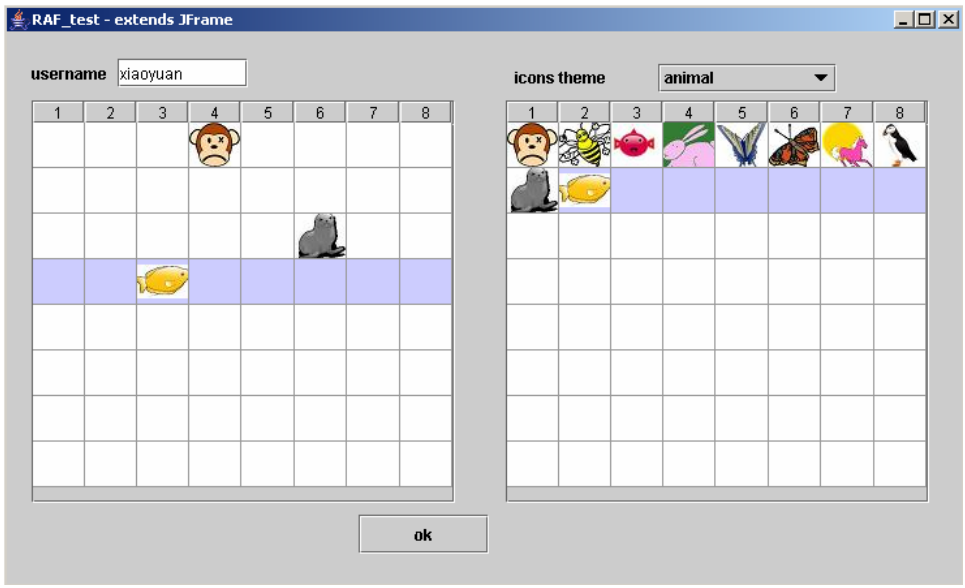

Figure12, this is the animal theme, in this theme, only 10 different kinds of objects are represented in the data table.

"computer", in which there are 24 distinctive icons (figure 13);

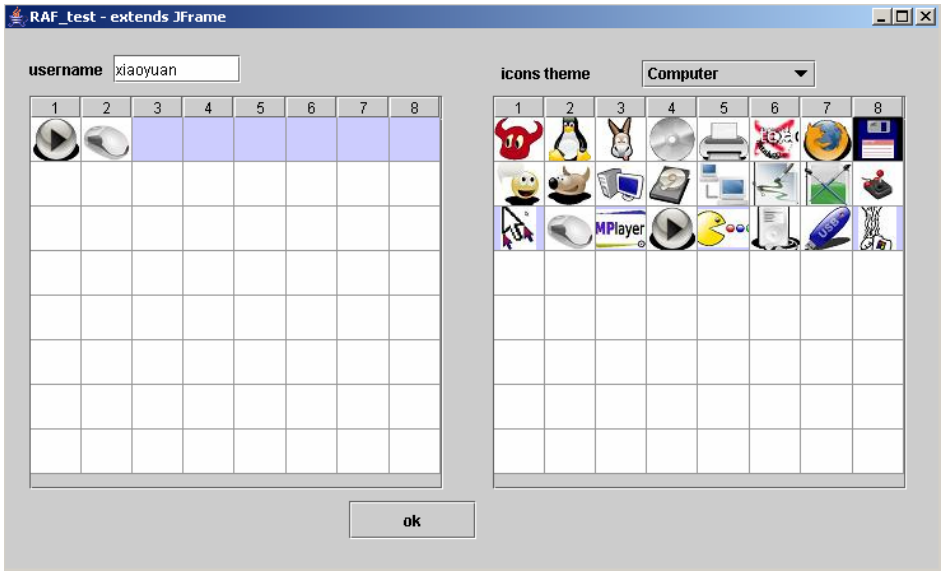

Figure13, this shows the computer theme, in which there are 24 different kinds of object in the data table.

and "simpleObject", in which there are 64 distinctive choices (figure 14).

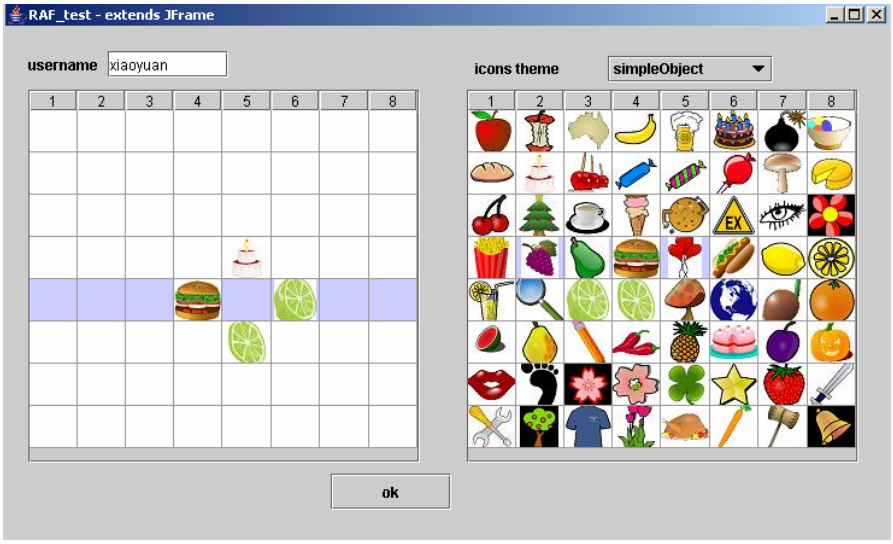

Figure14, this is the simpleObject theme, in which there are multiple kind of objects, but each of them are commonly seen, and very distinctive from each other.

Different theme had different amount of icons, since we would like to study how the simplicity of the data table will affect the effectiveness of RAF. User can choose any of these themes upon their preference; the data table will show the icons correspondingly.

RAF has two stages, registration and authentication. For the registration stage, user can choose icons from the data table, and place them into the desired cells of the input table. The icons in the input table and their formation will be served as the graphical password for users' future authentication purposes. The data, the contents, inside each cell of the input table, will be saved into our database with the user's name as the actual graphical password.

For the authentication stage, user will be given the same interface, but required to place the exact icon into the exact table cell. Then the system will determine whether or not the user has been authenticated.

5.2 User studies and analysis

We now compare our technique with the existing technique, DAS, recall from section two, as well as the traditional text-based password. So far we have taken an 8×8 input table, and a 8×8 data table. User has a choice of leaving one of the input table cells empty. Therefore, there are 65 different choices for each cell. In total we have 64^{65} choices. RAF has benefit for theme choices, since the user can ideally choose any icon upon their preference; this makes our theme choices to be infinite, thus, the total amount of choices of RAF is

$$64^{65} \times \infty$$

Of course this is the extreme case, in reality; no one will use all 64 objects. However, no one will fill all the grids in DAS either. Suppose we use 8 objects (which is more reasonable), and then the password space is $64^8 \times \mathit{themeChoices}$. Now consider text-based password. Suppose we have 83 printable characters, and then the password space of an 8-character password is $8^{83} = 9.046 \times 10^{74}$.

User study: we have further conducted a user study among 30 users, users chosen are mostly 20-30 years old, and they are either fellow student at GSU or close friends. We required the users to use our RAF system to create a graphical password, at the same time wrote down a simple text password, with no more than 64 digits. The users were asked to recall their password the next day also using our system. Among the 30 users we have chosen, only 11 people remembered their exact formation and the exact icons they have chosen. 15 others remembered more than half of the icons they inputted the day before. Among those 15 people, 9 people remembered their icons exactly, but not the formation. However, 100% of the users remembered their exact text-based password,

among those text-based password, 29 are easy-to-remember words or numbers, for instance their names and birthday dates.

3 people reported that they need to write the RAF formation on a piece of paper to assist them for future log in purposes. One person reported that she only inputted one object into the input table. Almost 90% of the users reported they prefer the single-objected theme to the multi-objected theme, since for a single-objected theme, they can remember the formation better, rather than the icons, and therefore the "all-star" theme was clearly the most popular theme choice. 80% of the user chose no more than 4 different objects in the whole input table, and the icons chosen were the simplest objects, for instance, simplest stars, circular objects, or their favorite item.

The user studies further proved that the choice of icons is the most critical issue when it comes to memorablity. Typical users preferred fewer icons especially fewer varieties of icons for their memory purposes, this proves that we need to unify the themes of icons. We have also tried an all faces theme at the beginning of the research, motivated by the research done by "realuser" corporation [15], however, users reported that the interface appears to be very confusing and it's difficult to distinguish the faces from one another.

CHAPTER6: AUTHENTICATION WITH MOUSE MOTION--A NEURAL NETWORK BASED APPROACH

6.1 Introduction

In chapter of my thesis, we will be proposing a new algorithm of authentication, in which both biometrical and graphics are combined. The motivation behind this technique is based on the observation that many computer users have a habit of moving around their mouse while waiting for a certain task; for instance, waiting for the pictures to download. We have therefore developed this technique to help us to study the pattern in each individual user's mouse motion; and further make such patter available for authentication purposes. This technique currently falls into the recall-a-formation category; it is one of the reproduce a drawing techniques.

In this technique, the user will be asked to move the mouse based on their intuition. Then the processes repeats and corresponding information such as the mouse motion will be recorded. Back-propagation learning algorithm is employed in this case to predict the next-value, then after the reconstruction of the mouse motion, the actual value and the predicted value are compared and if the error rate is within the tolerance rate, the user is authenticated. The algorithm is implemented using matlab and openGL; we later conducted a user study, the algorithm is proven to be accurate and has better usability than the existing similar techniques in a longer term.

6.2 Algorithm and simulation result

Here we divide our techniques into five stages: registration, prediction, reconstruction, comparison, and authentication. Each stage requires a different technique and uses a different programming language.

The user may start and terminated the program in their convenience. Registration Stage will record the users' mouse motion, and record the corresponding x and y values into the database. Prediction stage, takes the existing datasets as the input, will predict a set of datasets, that is, the predicted mouse motion. Then the corresponding predicted data and registered data will be compared, and a difference value will be recorded; an average of the difference values will be calculated and set as the tolerance rate. In order to be identified, the user will need to move their mouse within the tolerance rate compare with the predicted corresponding mouse motion.

Below are the detailed algorithm and simulation result.

6.2.1 Registration

This stage requires mouse tracking and motion recording techniques. In this stage, the user is asked to move the mouse based on their intuition; our program, using openGL, will record the user's mouse motion concurrently, please refer to figure 3, a screen shot taken while the user's mouse motion is being recorded, for illustrational purposes. The user may feel free to terminate the program at any time. Each point the mouse has passed will be saved into our database for future use. The mouse motion will continue being recorded on a daily or more frequent basis according to the user's will, the data will append to the database.

Figure 15. User mouse motion being recorded

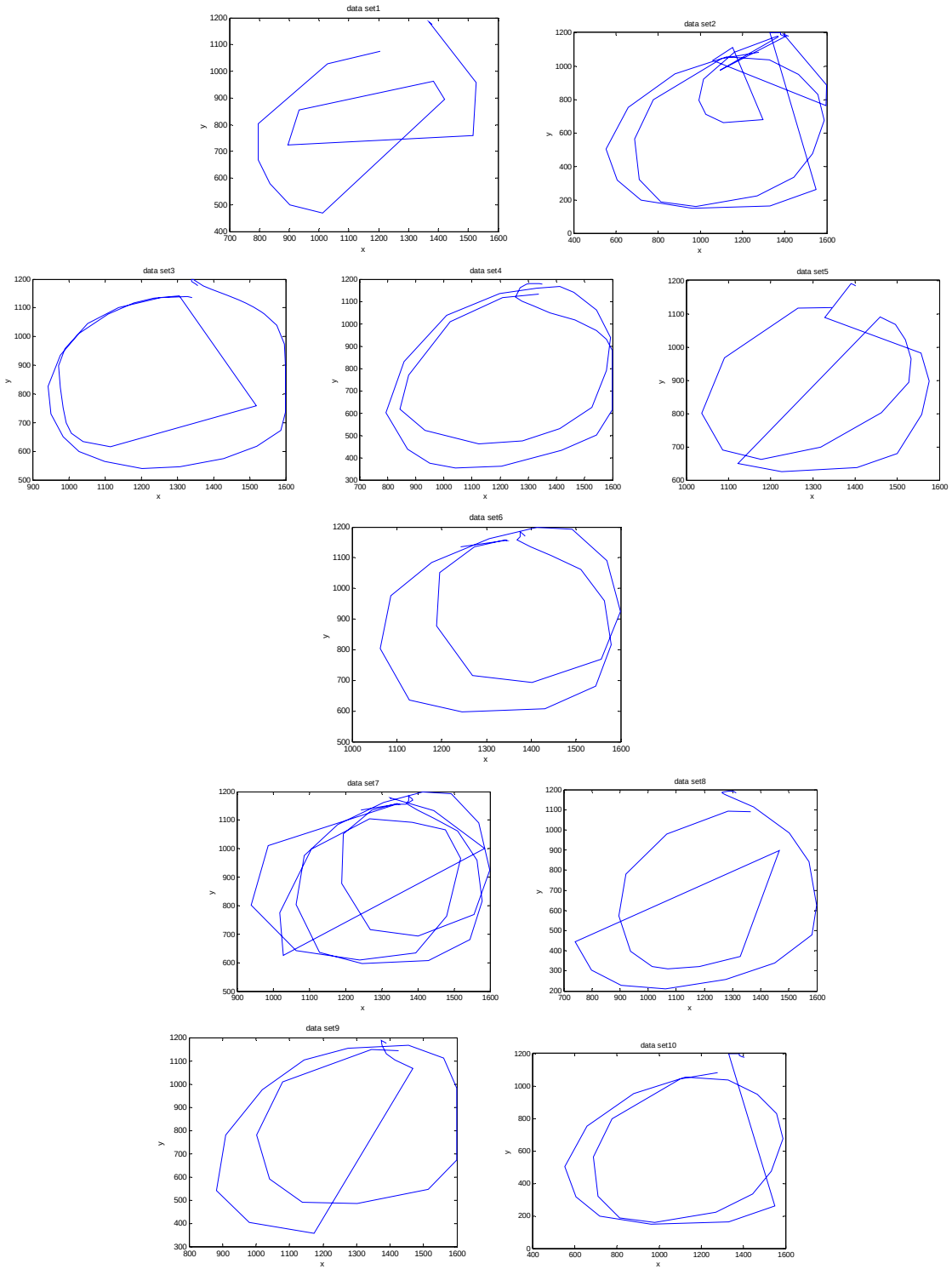

Figure 16. 10 dataset taking by the user within 10 different days

6.2.2 Prediction

Using back-propagation learning algorithm, we can consequently predict the next day's value. The back propagation learning algorithm takes the first n days of data sets as its input value, and generates an output value, or the predicted values. The predicted values will be saved in the same format as the registration values, provided for future purposes.

Our program, using the back-propagation learning algorithm (Pandya, et al. [15]), takes the data file, the learning rate η, the momentum constant α, which has the effect of smoothing the error surface in weight space by filtering out high-frequency variations, and the training error as the parameters to train the neural networks. We then begin the process by normalizing the weight of each dataset, that is, each data should divided by the largest number; the datasets are therefore all normalized to be between 0 and 1, and they are ready to be used for the program.

The predictions can starts no earlier than the 4^{th} dataset, since we need at least 3 inputs datasets to train the neural networks. The figure below is the difference between the predicted and real x, y values. In the case of dataset 4, the average difference between the predicted x values and real values is 13.7%, while the average difference for y values is only 6% (figure 17 and figure 18).

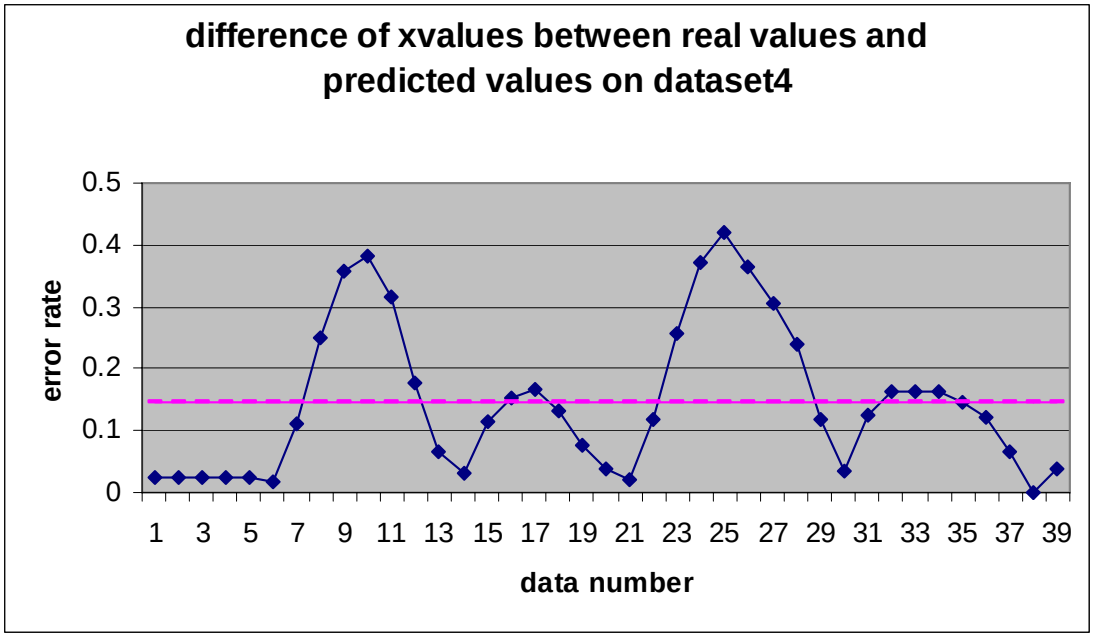

Figure 17. Difference in X values between predicted and real data set 4

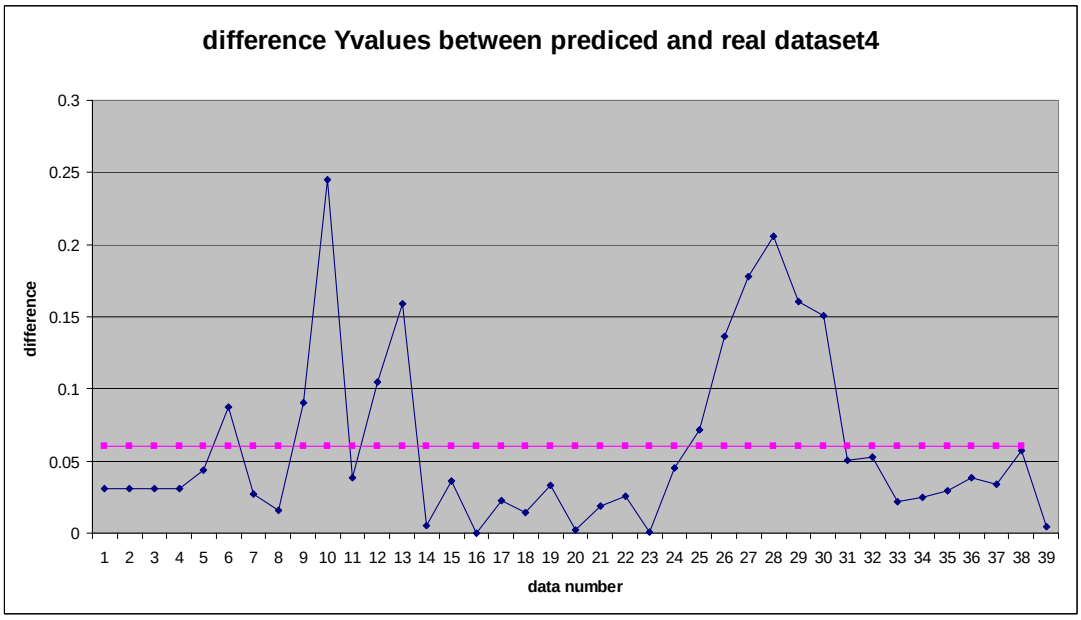

Figure 18. Difference in Y values between predicted and real data set 4

After calculating the average

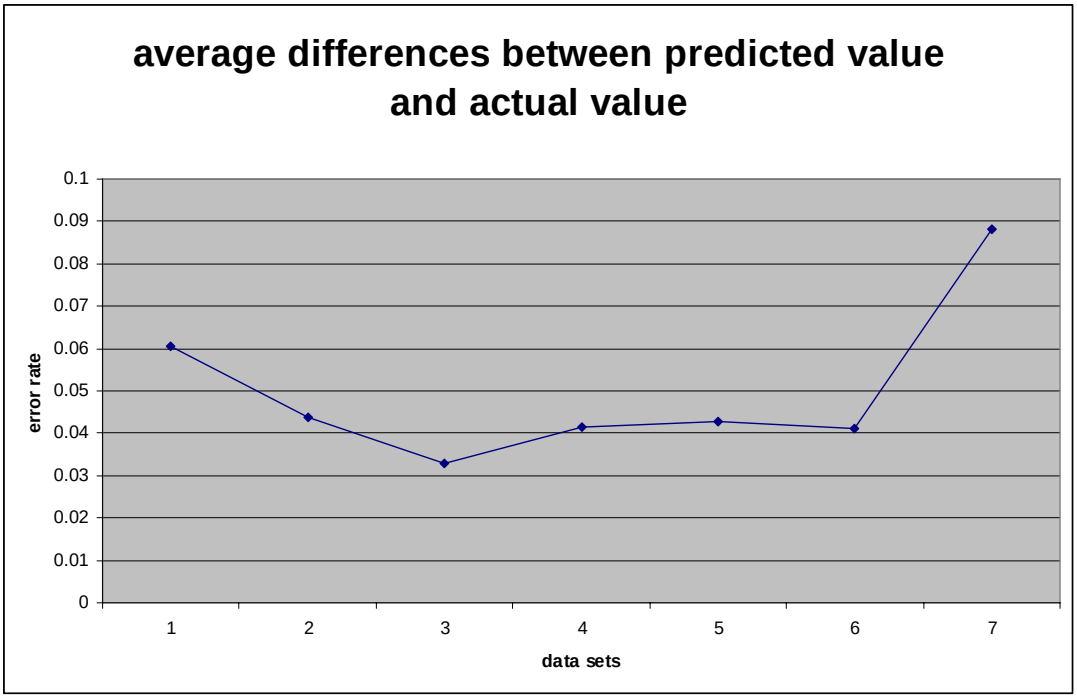

Figure 19. Average differences between predicted value and actual value

6.2.3 Reconstruction

Reconstruction stage is done by matlab; matlab program first reads from both the registration database and prediction database, and re-construct a 2-dimentional image by connecting the point, or the x, y values from top to bottom; the user's pre-registered mouse motion is then generated as a image. The images are saved for comparisons during the next step. Figure 4 showed the total 10 datasets of reconstruction, that is, the actual mouse motion the user made.

6.2.4 Comparison

There are several steps involved in this stage, two main algorithms have been employed: the first one I call it Suo's Algorithm, originally developed by Xiaoyuan Suo, or the first author. The algorithm normalizes the images; make them comparable to each other in terms of scale and location. The algorithm involves two major steps: centralizing

and scalizing.

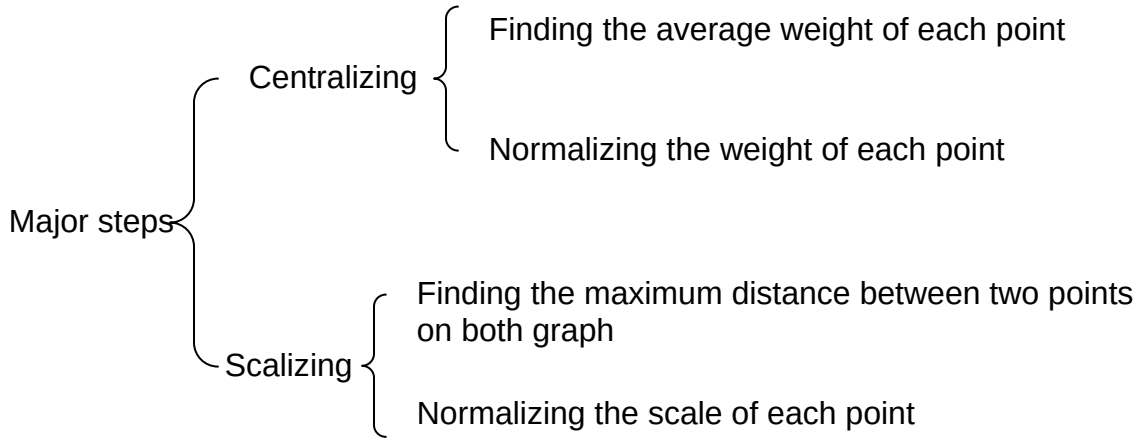

Figure 20. Major steps involved

Centralizing involves two steps: find the average weight and normalizing the weight of each point. In this section, we will be using dataset1 and dataset2 as our two comparable images for demonstration purposes; dataset1 is represented in red, while dataset2 is in blue.

First we need to find the average weight of the data set. The path consists with x and y values, suppose there are n sets of x and y values, so the central value would be average of the x and y values respectively, that is,

For x values, $x_c = \frac{1}{n}\sum_{i=1}^{n} x_i$

For y values: $y_c = \frac{1}{n}\sum_{i=1}^{n} y_i$

Secondly, we need to normalize the weight of each point regarding to the central value, that is,

For the new $x_i \longrightarrow x_i - x_c$;

For the new $y_i \longrightarrow y_i - y_c$

Therefore the comparable images are on the same center as figure 9 showed.

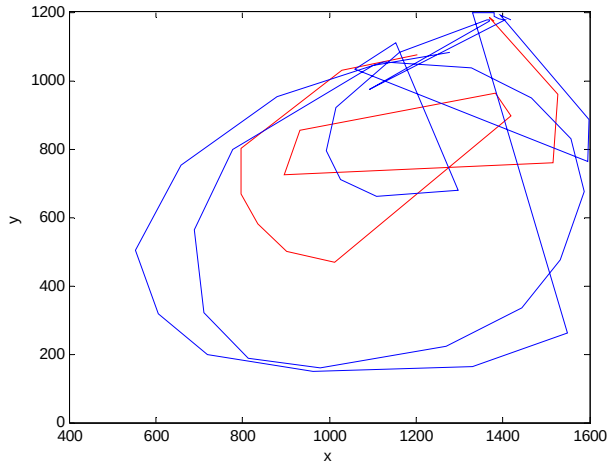

Figure 21. The two different images are on the same center.

The purpose of normalizing the comparable images is to make sure they are on the same scale. The reason as follows, the user might draw a larger scale image but the same shape, but we should identify the two images to be the same no matter of the scale. In order to normalize the weight or each point, we first need to find 4 sets of the extreme values, that is, the maximum and minimum x values on both of the comparable image and their corresponding y value; the maximum and minimum y values on both of the comparable image and their corresponding x value. We first find the distance between the maximum values on each graph, and then use each point to divide such distance. We use equations for better illustration purposes:

For x values,

$$\left.\begin{aligned} x_{\max} &= \max(x_i) \\ x_{\min} &= \min(x_i) \end{aligned}\right\} \quad y_{\max,} y_{\min}$$

$$dist1 = \sqrt{(x_{\max} - x_{\min})^2 + (y_{\max} - y_{\min})^2}$$

For y values,

$$\left.\begin{aligned} y_{\max} &= \max(y_i) \\ y_{\min} &= \min(y_i) \end{aligned}\right\} \quad x_{\max,} x_{\min}$$

$$dist2 = \sqrt{(x_{\max} - x_{\min})^2 + (y_{\max} - y_{\min})^2}$$

Then we find the bigger value between dist1 and dist2; we define that value as our largest distance between any two points on the image, or *dist* . So for each of the x and y values on the image,

$$newvalue = \frac{x_i}{dist}$$

$$newvalue = \frac{y_i}{dist}$$

We therefore re-plot the two images; the result is on figure

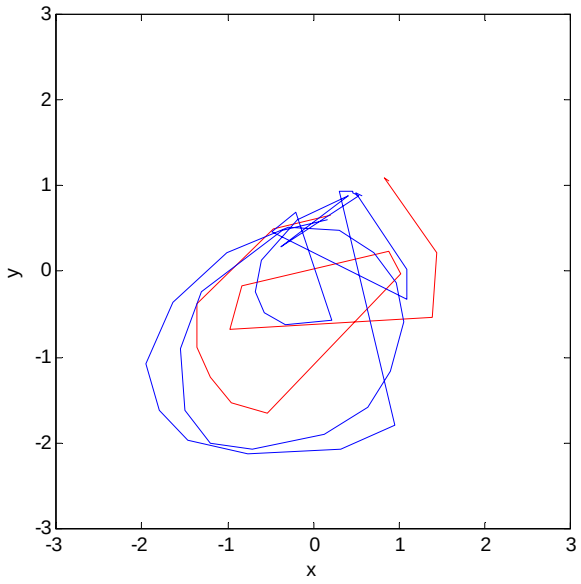

Figure 22. The two different images are now on the same center and with the same scale.

At this moment, the images are still incomparable, main reason is that the two images may have different amount of points; therefore we need to add more values on the curves in order to make them smoother.

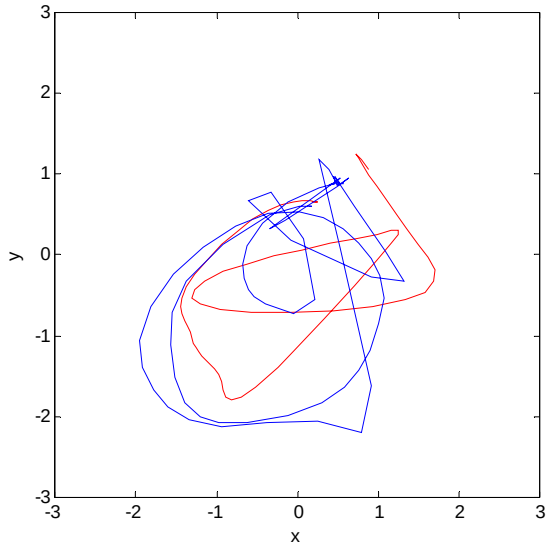

Figure 23. After applying the smooth function

Now the two datasets have same amount of points and they are comparable!

6.2.5 Authentication

From the previous step, we calculated that the difference between the two datasets was 1.3870. The figure below is a graphical illustration of all the differences among 10 different datasets and the predicted values.

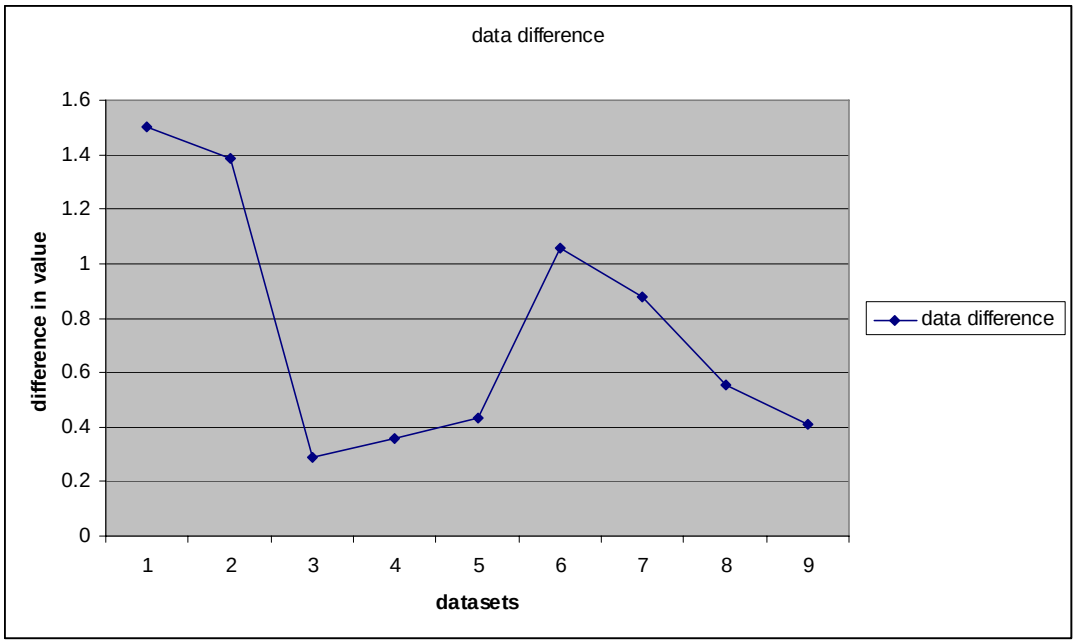

Figure 24. Datasets differences.

Notice that the differences among first 3 datasets are slightly higher than the others, which mostly are less than 1.2.

Our technique authenticates a pre-registered user by allowing user to move the mouse motion within the data difference range. In the example above, we can see that the average difference is around 0.8; therefore, for the next possible data set, or data set No. 11, we set the tolerance to be 0.8, that is, only if the user's mouse motion re-performance is no more different than 0.8 compare with the previous performances, the user can be authenticated. The user in this case was authenticated on the 5th dataset, which has an error rate clearly less than the previous ones.

6.3 Analysis

This technique is a novel approach in terms of using user's own behavior as the authentication method. In this paper, we have demonstrated that using user's mouse motion as an authentication method is plausible, yet easy to implement, and the

simulation showed that the user started to be authenticated during the second time of authentication.

The project can be further extended as a person identification program for computer security and internet services. To be differentiated from the traditional authentication method, which typically requires a user name and a password (text, graphical, finger prints. etc) and authenticates the user right away, our program may allow the user to do their daily activities based on their will, and then identifies the user based on the comparison between their activity and the predicted activity.

Adding more users' biometric parameters and factors into the registration process would certainly be an improvement, examples such as: recording the users' typing patter alone with the mouse motion, or the eye fixation during different activities.

Nevertheless, biometrics as an authentication technique is still at its very early stage, there is still much more work to do.

CHAPTER7: A SHOULDER SURFING RESISTANT PASSPOINT

Figure 25, passpoint technique

7.1 Introduction

Passpoint technique (figure 24), as we can see from the chapter one, has a large password space, and user studies proved its usability. However, the major problem with this technique is that it is not shoulder surfing resistant, which is one of the main hazards in graphical password. This technique is also not mouse tracking resistant; one could easily reproduce the password by recording the user's mouse motion.

7.2 Methodology

We therefore developed a mouse tracking and shoulder surfing resistant passpoint; it differs the traditional passpoint by requiring the user to select from a set of decoyed images. The idea of using decoyed image was inspired by passface technique. In our case, the decoyed image's idea is different from passface; therefore instead of using several different pictures, we use the same pictures, but each time we emphasis on the different possible area (figure 25).

Figure 26. The image is blurred except the decoyed area. If the user's passpoint is within the decoyed area, the user may click on "Y"; "N" otherwise

Since there is no mouse motion required during the authentication stage, mouse tracking devices will certainly be useless. User will be asked to input either "Y" or "N" for decoyed areas, that is, "Y" if the decoyed area was among the pre selected passpoints and "N" for otherwise. Since the decoyed area can change every time randomly, there is no use of recording the keyboard input either.

The process repeats for several rounds for security purposes. In order to save the user's time, each decoyed area will stay for 5 seconds, there will be no more than 10 different decoyed areas shown each time. The total process will therefore take no more than 1 minute each time. Compare with traditional text-based password approach, the process is longer than the traditional approach.

The password space remains to be the same as passpoint in this case; any pixels on a picture can be used and any pictures can be used. The authentication stage doesn't affect the password space either, since the decoyed area can be anywhere on a given picture.

The image will have major areas being blurred out except the decoyed area. If the user's pre-selected passpoint is within the decoyed area, the user may click on the mouse's left button; right button otherwise. After several rounds, if all of the decoyed images are correctly identified, the user will be authenticated.

7.3 Analysis:

According to the Chapter 2, we judge a graphical password based on two basic criteria, that is: security and usability. Security, according to the previous section, can be judged based on several important areas; they are brute force search, dictionary attacks, guessing spyware, shoulder surfing and social engineering. Brute force attacks can be prevented by a large password space.

Password space is considerably large in passpoint, since any pixels and any pictures can be used. Employment of decoyed area will not affect the password space in this case, since the decoyed area may move randomly, and as long as the passpoint is within the decoyed area, the user may click on "Y". Therefore due to the large password

space, it is hard to carry out brute force and dictionary attacks for this kind of technique. Guessing is also impossible in this case, since the decoyed area may change every time randomly, and there could be countless possible combinations for a password.

Since there is no mouse motion during the authentication stage, there is no use of recording the mouse motion. And since the decoyed area will change each time and there's only a choice of "Y" or "N", key logging or key listening will be useless also.

This technique prevents the shoulder surfing problem. Video taping or shoulder surfing will not be helpful in this case, since it is hard to tell exactly which pixel the user is referring to in the decoyed area and the decoyed area may change randomly. Typical graphical password can not be described in simple words; therefore, typical graphical password prevents the social engineering problem.

The method nevertheless inherits some shortcomings from the passface scheme, such as: the process of authentication may take too long and typical user would find it hard to use at the first sight.

CHAPTER8: CONCLUSION AND FUTURE WORK

The past decade has seen a growing interest in using graphical passwords as an alternative to the traditional text-based passwords. In this paper, we have conducted a comprehensive survey of existing graphical password techniques. The current graphical password techniques can be classified into two categories: recognition-based and recall-based techniques. A comparison of current graphical password techniques is presented in Table 1.

Although the main argument for graphical passwords is that people are better at memorizing graphical passwords than text-based passwords, the existing user studies are very limited and there is not yet convincing evidence to support this argument. Our preliminary analysis suggests that it is more difficult to break graphical passwords using the traditional attack methods such as brute force search, dictionary attack, or spyware. However, since there is not yet wide deployment of graphical password systems, the vulnerabilities of graphical passwords are still not fully understood.

The first technique: RAF, or recall a formation, is proven to have a larger password space, therefore considered to be safer than the existing graphical password techniques. One main drawback of RAF is that the user studies proved that RAF does not have good memorability. The user studies further proved that the choice of icons is the most critical issue when it comes to memorablity. Typical users preferred fewer icons especially fewer varieties of icons for their memory purposes, this proves that we need to unify the themes of icons. We have also tried an all faces theme at the beginning of the research, motivated by the research done by “realuser” corporation [15], however, users

reported that the interface appears to be very confusing and it's difficult to distinguish the faces from one another.

The second technique, a Neural Network Based mouse motion authentication method, is a novel approach in terms of using user's own behavior as the authentication method. In this paper, we have demonstrated that using user's mouse motion as an authentication method is plausible, yet easy to implement, and the simulation showed that the user started to be authenticated during the second time of authentication.

The project can be further extended as a person identification program for computer security and internet services. To be differentiated from the traditional authentication method, which typically requires a user name and a password (text, graphical, finger prints. etc) and authenticates the user right away, our program may allow the user to do their daily activities based on their will, and then identifies the user based on the comparison between their activity and the predicted activity.

Adding more users' biometric parameters and factors into the registration process would certainly be an improvement, examples such as: recording the users' typing patter alone with the mouse motion, or the eye fixation during different activities. Nevertheless, biometrics as an authentication technique is still at its very early stage, there is still much more work to do.

The third technique, a shoulder surfing resistance passpoint, overcomes many of the existing problems of the traditional passpoint. For instance, the new technique is mouse tracking resistance and shoulder surfing resistant, and it still retains the large password space the traditional passpoint technique has. Some of the possible drawbacks

of this technique may include: time consuming and difficult to use. However the security and system concerns have not been addressed. we consider it as our future work.

Overall, the current graphical password techniques are still immature. Much more research and user studies are needed for graphical password techniques to achieve higher level of maturity and usefulness.

References

[1] A. C. L. Andrew S. Patrick, Scott Flinn, "HCI and Security Systems," in CHI, Extended Abstracts (Workshops). Ft. Lauderdale, Florida, USA., 2003.

[2] A. Adams and M. A. Sasse, "Users are not the enemy: why users compromise computer security mechanisms and how to take remedial measures," *Communications of the ACM*, vol. 42, pp. 41-46, 1999.

[3] K. Gilhooly, "Biometrics: Getting Back to Business," in *Computerworld, May 09*, 2005.

[4] R. Dhamija and A. Perrig, "Deja Vu: A User Study Using Images for Authentication," in *Proceedings of 9th USENIX Security Symposium*, 2000.

[5] M. Kotadia, "Microsoft: Write down your passwords," in *ZDNet Australia, May 23*, 2005.

[6] A. Gilbert, "Phishing attacks take a new twist," in *CNET News.com, May 04*, 2005.

[7] A. Jain, L. Hong, and S. Pankanti, "Biometric identification," *Communications of the ACM*, vol. 33, pp. 168-176, 2000.

[8] R. N. Shepard, "Recognition memory for words, sentences, and pictures," *Journal of Verbal Learning and Verbal Behavior*, vol. 6, pp. 156-163, 1967.

[9] A. Perrig and D. Song, "Hash Visualization: A New Technique to Improve Real-World Security," in *Proceedings of the 1999 International Workshop on Cryptographic Techniques and E-Commerce*, 1999.

[10] S. Akula and V. Devisetty, "Image Based Registration and Authentication System," in *Proceedings of Midwest Instruction and Computing Symposium*, 2004.

[11] D. Weinshall and S. Kirkpatrick, "Passwords You'll Never Forget, but Can't Recall," in *Proceedings of Conference on Human Factors in Computing Systems (CHI)*. Vienna, Austria: ACM, 2004, pp. 1399-1402.

[12] L. Sobrado and J.-C. Birget, "Graphical passwords," *The Rutgers Scholar, An Electronic Bulletin for Undergraduate Research*, vol. 4, 2002.

[13] D. Hong, S. Man, B. Hawes, and M. Mathews, "A password scheme strongly resistant to spyware," in *Proceedings of International conference on security and management*. Las Vergas, NV, 2004.

[14] S. Man, D. Hong, and M. Mathews, "A shoulder-surfing resistant graphical password scheme," in *Proceedings of International conference on security and management*. Las Vegas, NV, 2003.

[15] RealUser, "www.realuser.com," last accessed in June 2005.

[16] T. Valentine, "An evaluation of the Passface personal authentication system," Technical Report, Goldsmiths College, University of London 1998.

[17] T. Valentine, "Memory for Passfaces after a Long Delay," Technical Report, Goldsmiths College, University of London 1999.

[18] S. Brostoff and M. A. Sasse, "Are Passfaces more usable than passwords: a field trial investigation," in *People and Computers XIV - Usability or Else: Proceedings of HCI*. Sunderland, UK: Springer-Verlag, 2000.

[19] D. Davis, F. Monrose, and M. K. Reiter, "On user choice in graphical password schemes," in *Proceedings of the 13th Usenix Security Symposium*. San Diego, CA, 2004.

[20] W. Jansen, "Authenticating Mobile Device Users Through Image Selection," in *Data Security*, 2004.

[21] W. Jansen, S. Gavrila, V. Korolev, R. Ayers, and R. Swanstrom, "Picture Password: A Visual Login Technique for Mobile Devices," National Institute of Standards and Technology Interagency Report NISTIR 7030, 2003.

[22] W. A. Jansen, "Authenticating Users on Handheld Devices," in *Proceedings of Canadian Information Technology Security Symposium*, 2003.

[23] T. Takada and H. Koike, "Awase-E: Image-based Authentication for Mobile Phones using User's Favorite Images," in *Human-Computer Interaction with Mobile Devices and Services*, vol. 2795 / 2003: Springer-Verlag GmbH, 2003, pp. pp. 347 - 351.

[24] I. Jermyn, A. Mayer, F. Monrose, M. K. Reiter, and A. D. Rubin, "The Design and Analysis of Graphical Passwords," in *Proceedings of the 8th USENIX Security Symposium*, 1999.

[25] J. Thorpe and P. C. v. Oorschot, "Graphical Dictionaries and the Memorable Space of Graphical Passwords," in *Proceedings of the 13th USENIX Security Symposium*. San Deigo, USA: USENIX, 2004.

[26] J. Thorpe and P. C. v. Oorschot, "Towards Secure Design Choices for Implementing Graphical Passwords," in *Proceedings of the 20th Annual Computer Security Applications Conference*. Tucson, Arizona, 2004.

[27] J. Goldberg, J. Hagman, and V. Sazawal, "Doodling Our Way to Better Authentication," presented at Proceedings of Human Factors in Computing Systems (CHI), Minneapolis, Minnesota, USA., 2002.

[28] J. Thorpe and P. C. v. Oorschot, "Towards Secure Design Choices for Implementing Graphical Passwords," in *20th Annual Computer Security Applications Conference (ACSAC)*. Tucson, USA.: IEEE, 2004.

[29] D. Nali and J. Thorpe, "Analyzing User Choice in Graphical Passwords," Technical Report, School of Information Technology and Engineering, University of Ottawa, Canada May 27 2004.

[30] A. F. Syukri, E. Okamoto, and M. Mambo, "A User Identification System Using Signature Written with Mouse," in *Third Australasian Conference on Information Security and Privacy (ACISP)*: Springer-Verlag Lecture Notes in Computer Science (1438), 1998, pp. 403-441.

[31] G. E. Blonder, "Graphical passwords," in *Lucent Technologies, Inc., Murray Hill, NJ*, U. S. Patent, Ed. United States, 1996.

[32] Passlogix, "www.passlogix.com," last accessed in June 2005.

[33] sfr, "www.viskey.com/tech.html," last accessed in June 2005.

[34] L. D. Paulson, "Taking a Graphical Approach to the Password," *Computer*, vol. 35, pp. 19, 2002.

[35] S. Wiedenbeck, J. Waters, J. C. Birget, A. Brodskiy, and N. Memon, "Authentication using graphical passwords: Basic results," in *Human-Computer Interaction International (HCII 2005)*. Las Vegas, NV, 2005.

[36] S. Wiedenbeck, J. Waters, J. C. Birget, A. Brodskiy, and N. Memon, "Authentication using graphical passwords: Effects of tolerance and image choice," in *Symposium on Usable Privacy and Security (SOUPS)*. Carnegie-Mellon University, Pittsburgh, 2005.

[37] S. Wiedenbeck, J. Waters, J. C. Birget, A. Brodskiy, and N. Memon, "PassPoints: Design and longitudinal evaluation of a graphical password system," *International Journal of Human Computer Studies*, to appear.

[38] J.-C. Birget, D. Hong, and N. Memon, "Robust discretization, with an application to graphical passwords," Cryptology ePrint archive 2003.