Design and Evaluation of a Game-based Authentication Mechanism: Find That Animal

Author: Yang Liu
University of York

Supervisor: Dr. Alistair Edwards
University of York

Tuesday, 15 September 2015

20629 words as counted by Pages word count function over 60 pages of the main contents of this report
ABSTRACT 6
ACKNOWLEDGEMENTS 7
STATEMENT OF ETHICS 8
1. INTRODUCTION 9
2. LITERATURE REVIEW 10
   2.1 IDENTIFICATION AND AUTHENTICATION 10
   2.2 PASSWORD AUTHENTICATION 11
      2.2.1 Current Use of Passwords 11
      2.2.2 Attacks and Security Issues of Passwords
         Passwords cracking 12
         Examples of dictionaries: 12
         Lack of Enforcement for Creating Robust Passwords 13
   2.3 HUMAN MEMORY 14
      2.3.1 Short-term Memory 14
         Multicomponent model 15
      2.3.2 Long-term Memory 16
   2.4 GRAPHICAL PASSWORDS 17
      2.4.1 Password Retrieval 18
      2.4.2 Graphic Authentication Mechanisms 18
         Recall based mechanisms 18
         Cued-recall based mechanisms 19
         Recognition based mechanisms 20
   2.5 GAME DESIGN 20
      2.5.1 Iterative Design 20
      2.5.2 The Pyramid of Game Design 21
   2.6 THEORY OF FUN 23
      2.6.1 Lazzaro’s 4 Keys to Emotions 23
3. DESIGN 24
   3.1 GATHER REQUIREMENT 24
   3.2 EVALUATE GAMES 26
   3.3 ITERATE IDEAS 29
      3.3.1 Stealing shoes 30
      3.3.2 Tree game 31
      3.3.3 Tangram and Chinese characters 33
         Tangram 33
         Chinese characters 35
4. PROTOTYPE 36
4.1 IMPLEMENTATION TECHNOLOGY 36
  4.1.1 Coding languages 37
  4.1.2 Tools 37

4.2 IMPLEMENTATION PROCESSES 37
  4.2.1 Scenario-based Design 37
    Persona 38
    Scenario 38
    Claim analysis 39
    Site map 40
  4.2.2 Selection of pictures 41

4.3 PROTOTYPE 1.0 42
  4.3.1 Anatomy of the game 42
    FtA in login 42
    FtA in practice 43
  4.3.2 Code Implementation 43

4.4 PROTOTYPE 2.0 44
  Changes of rules 44
  Self-adjusted buffer for blocking time 45
  Usability issues improved 45
  Screenshots 46

5. TESTING 51

5.1 EXPERIMENT 1 51
  5.1.1 Participants 51
  5.1.2 Procedures 51
    Send email reminder for registration 51
    Clicking the first email reminder 52
    Send email reminders for login 52
    Clicking subsequent email reminders 52
    Device 52
    Grouping 52
  5.1.3 Result 53
    Rule out data 53
    Login distribution 54
    Error Analysis 54
    Time Analysis 56
    Conclusion 57

5.2 EXPERIMENT 2 58
  5.2.1 Participants 58
  5.2.2 Procedures 59
    Devices 59
Test the prototype of FtA 59
Finish the questionnaire 59

5.1.3 Result 59
  Ease of use 59
  Fun 62
  Conclusion 63

6. DISCUSSION 64

6.1 ISSUES OF FTA 64
  6.1.1 Selection of pictures 64
    Preference related effect 64
    Danger of choosing pictures in order 64
  6.1.2 Remembering secret pictures 65
    Practice is not enough 65
  6.1.3 Relying heavily on vision 65

6.2 VULNERABILITY EVALUATION 66
  Guessability 66
  Observability 67
  Recordability 67
  Memorability 67

7. CONCLUSION 68

8. FURTHER WORK 69

REFERENCE 70
Abstract

Various security policies are available to provide access control of different levels to protect confidential information and ensure the access for authorized people. With the advancement of technology, traditional authentication mechanism is widely implemented to plenty of secured systems. This authentication mechanism often requires a user ID and a textual password. While as the increased threaten of password attacks through the networks, people are encouraged or forced to use good passwords. However, a good password is often long, complex, and hard to remember. And it is also challenging human’s brain to manage different passwords across different secured systems.

As researches have proven that we have a better memory for remembering visual information, graphic authentication mechanism is proposed to facilitate the remembering of passwords. Different models have been proposed, but graphical passwords are still not in use widely. It may because the usability of graphical passwords is not improved so much than that of textual passwords, but the improvement is at a cost of login time.

This project explores a novel idea to implement a prototype which combines a graphical authentication mechanism with a simple game. Funny animal pictures are used in the game as passwords. The results of a usability testing show that it takes about 5 to 20 seconds to login through the game, and people think it is fun. But people still prefer textual passwords.
Acknowledgements

I would like to offer my sincerest thanks to my supervisor Dr. Alistair Edwards, for his guidance and assistance in carrying out this project. I must also acknowledge the support of Mr. Alex Cooper. Finally, I would like to thank my fellow students at the university for taking part in the experiments.
Statement of Ethics

Edwards [17] states that the following basic ethical principles should be applied in any project:
• Do not harm
• Informed consent
• Confidentiality of data

This project incorporated two experiments that test the overall usability of a game-based authentication mechanism. Participants were informed that taking part in the testing of the prototype (FtA) is voluntary and of no risk to them, and a reward is provided. They may quit and withdraw at any point of the experiment and their accounts and records will be deleted on the database.

The outline and instructions of the experiment is integrated in the prototype in a form of web pages. The participation is made by clicking a ‘Proceed’ button to give consent. And the consent is granted also by clicking button to proceed the experiment.

The procedures of the two experiment were nearly the same. Both experiments asked participants to follow the instructions on FtA and register an account with 3 secret pictures. Participants then practiced the game and login to the system. In experiment 1, participants were asked to test the prototype for 7 days. They would receive email reminders with a link to the prototype every morning during the experiment. At the last day, participants were asked to finish a questionnaire by clicking an additional link on the email reminder. In experiment 2, participants used the same prototype and followed the instructions to complete the procedures of registration, practice, and login. After the experiment, they were asked to finish a questionnaire. All the data created during the use of the prototype are kept on a database in the university server, and can only be accessed by typing two password authentication approaches. The result of the questionnaire is kept in Qualtrics, and can be accessed through university account.

During experiment 1, email reminders should be sent using ‘Bcc’ to prevent the disclosure of email addresses to other receivers in the mailing list. Reminders were send not using this method at first. These participants soon received apologies, and they were noticed they could withdraw their consent at that point.
1. INTRODUCTION

Password use is the most prevalent authentication method because it can be easily reused across multiple systems and applications without any additional devices. Typically, a user identifier is used together with a password to verify a unique user by matching the password with the registry information stored beforehand in the underlying system. To ensure the security of password authentication, not only is a secure transport channel necessary, but also the encoding of the password (e.g. using hash functions like MD4 and SHA) is needed. However, apart from the encryption and protection of data during transportation through networks or the security levels of hardware, passwords themselves can be a weak link to be cracked. Referring to Mitnick and Simon [1]: “A company may have purchased the best security technologies that money can buy, trained their people so well that they lock up all their secrets before going home at night, and hired building guards from the best security firm in the business. The company is still totally vulnerable... The human factor is truly security’s weakest link”.

To combat with the increasing number of password attacks, organizations are now implementing stricter rules to create good passwords. As it is suggested that ‘good’ passwords should always be complex and long, the memory load for remembering such passwords is increased. In addition, infrequently used passwords are forgotten repeatedly, which also decreases the usability of traditional textual passwords. By having different passwords for different systems, our brain is challenged to remember all of them. There is a general agreement that it is hard to manage many textual passwords for different services. To ease the problem of human inability to deal with a large number of passwords, the graphical password is proposed. The first model of a graphical password is called a locimetric system, first proposed by Blonder, which requires the person to verify their identity by clicking on positions within a picture. Other models are searchmetric (picking a picture from a grid of images), and drawmetric (drawing a secret pattern) [26].

Different models of graphical authentication are proposed, but it is hardly to connect the use experience of fun. At the same time, people seem increasingly happy to fill in any odd patches of time playing simple games [22]. Why not implement a graphical authentication in combination with a simple game?

This project aims to explore a novel idea to solve the usability issues of textual passwords. It investigates the idea of having to authenticate by a special pattern in a simple game. A prototype, named ‘find that animal’, is implemented, and is tested in two experiments. A vulnerability evaluation is conducted at last.
2. LITERATURE REVIEW

2.1 Identification and Authentication

Identity establishment aims to create a profile that can prove the entity is authentic, and the process of authentication verifies the entity's identity in order to rule out unauthorised access to confidential information. In general, the identity profile of an entity is encapsulated in an account, and the process of authentication takes place on every login attempt to access a controlled computing system. To establish a user identity, each entity should be associated with an authentication credential that is shared only between the entity and the computing system. There are three methods of achieving this:

- Presenting something the user knows: A password, a personal identification number (PIN), or a pass-phrase are the most common schemes in this category. Character-based passwords are used on most computing systems as they are easy to implement and easy to be verified. To make the password more secure, many restrictions on passwords, such as length, syntax, and alphabet, are required. The increased security level of passwords also makes it challenging for users to remember, especially when they have many different passwords for different services or applications. Due to the difficulty of remembering and managing complex passwords, users may use the same password across a number of computing systems. In the event of the password being disclosed on one system, the information kept on other systems is at great risk of disclosure as well.

- Presenting something that the user has: This authentication scheme relies on a device, known as a token, to store the credentials of identity verification, like a USB stick or card reader used when generating a PIN to make bank transactions.

- Presenting something that the user is: This scheme authenticates a user by biometrical traits that reliably distinguish users such as fingerprints, hand geometry, hand signature, eye shape, voice, and face recognition. This scheme is much more secure but will cost more money to implement. In addition, the verification of biometrical traits other than fingerprints may have inaccuracies due to issues with related technologies.
2.2 Password Authentication

2.2.1 Current Use of Passwords

When people enter the virtual world of the internet and access a variety of services, they cannot protect their confidential information using actual locks and keys, so they use passwords instead. Many researchers have investigated the habits and behaviours of people using passwords in different groups and countries.

According to Schneier’s study of 34,000 MySpace accounts in 2006, 65% of all passwords were shorter or equal to 8 characters and the most frequently used passwords were ‘password1’, ‘abc123’, ‘myspace1’, and ‘password’. [2] The security of a password is not only related to its length, but is also related to how many times the password is renewed. According to the results of Hoonakker’s password survey from a human factor perspective, on average each user had nine password-protected accounts, and each password was renewed 7 times a year. 18% of the respondents used the same password always, and 50% reused sometimes. [3] The results of a different password survey conducted on Haley’s blog showed that 44% of respondents owned more than 20 password-protected accounts, 63% used the same password always, and 45% reused old passwords on their accounts. [4] In Kumar’s study of students and staff in the University at New Delhi, 60% of respondents had more than 6 password-protected accounts, 79% reused old passwords, and only 30% used different passwords for different accounts. [5] Similarly, in a survey about the password habits of American consumers, 54% of respondents had less than 5 passwords, 44% used the same passwords for more than a year, and 61% reused old passwords. [6] Finally, in the extended results of a Norwegian password security survey, Helkala indicated that a person had 25 passwords on average, and 74% of people reused passwords. [45] Helkala also found that more than half of respondents believed some accounts, like bank accounts or email accounts, need to be well-protected by stronger passwords, but only 31% always used stronger passwords to protect these accounts. In addition, the general understanding of what constitutes a “good password” is one which has “a mixture of letters and digits”.

Although the results of the different surveys vary, in general users tend to have more than 5 passwords and reuse old passwords at a high rate, but still some of them may not change their passwords over a long period. Our brain has a limited ability to remember a number of different passwords, and so people will forget the passwords that are not used frequently. However, using the same passwords across
different CISs will increase the risk of password disclosure if security of one of the computing systems is compromised.

### 2.2.2 Attacks and Security Issues of Passwords

**Passwords cracking**

The power of password authentication can be largely reduced by a weak password which is short and easy to be guessed. The main attacks against passwords authentication includes dictionary attack, brute-force attack, and rainbow table attack. The first two techniques are used for password cracking by checking different combinations of words, characters, and numbers, and the rainbow table attack is used to reverse hash functions. In particular, a dictionary attack is more aggressive than other password cracking techniques if given enough dictionaries of passwords and a proper cracking tool like John the Ripper. Dictionaries refer to lists of weak passwords that have been leaked in the past few years.

**Examples of dictionaries:**

*Top500.* This dictionary came from “Top 500 Worst Passwords of All Time” [7] which was then developed to Top 10000 passwords in 2011. The dictionary includes passwords like 123456, password, qwerty, and master.

*John the Ripper (JtR).* This tool is a popular password cracker with a dictionary of 3,456 passwords [8]. This dictionary also includes simple words like those in Top500, but some of them are little more complex like trustno1.

*C&A.* Cain & Abel is a password cracker like JtR which has a dictionary with 306,706 words [9]. The words in this dictionary are long lower case words like constantness.

*RY5.* [RockYou.com](http://RockYou.com) is a gaming website that was subject to an SQL injection attack in 2009, resulting in the leak of 32.6 million cleartext user passwords. This dictionary has 14.3 million unique passwords and has become one of the largest real user-generated password databases in existence today.

*YA.* This dictionary consists of cleartext passwords of 450,000 Yahoo! users due to an SQL injection attack in July 2012. The number of unique passwords is 342,541
Lack of Enforcement for Creating Robust Passwords

Although guidelines for creating secure passwords are proposed and suggested, the policies for creating passwords in different websites vary from one to another. As for the restriction of passwords, the minimum length of six characters is still very common and there is not enough strict and clear instruction for creating strong passwords [10]. Passwords on such websites are vulnerable to password cracking if a proper dictionary is used in accordance with that specific password policy. For example, LinkedIn was attacked in June 2012 and as a result millions of hashed passwords were released and partially cracked independently. The accounts on LinkedIn just required passwords with a minimum length of 6 characters, and so the leaked passwords were a useful, natural, and real-user generated dictionary for cracking passwords under similar policies. In general, users may think that some accounts are less important than others like bank accounts, and they may knowingly choose relatively weak passwords. [11] Accounts on LinkedIn contained personal information related to work experience which might have caused further losses if this information was disclosed. Numerous researchers have suggested many guidelines for strong passwords, but users often just follow the enforcement of the websites even though they are aware of these guidelines. Websites should leverage the security level of the content or services they have and implement appropriate enforcement for creating passwords. However, if the policy is too clear or specific in confining the range of the passwords, it may in turn reduce the difficulty in cracking the passwords.

Users can deviate from best practice in many ways, such as: always using only one password to access different systems; using a short password (equal to or less than 8 characters); not using a complex combination of capital letters, lower-case letters, and special characters; reusing old passwords; writing down their passwords [12]; saving passwords in an electronic file without protection[13]; and sharing passwords with others. [14] To encourage users to use strong passwords during their creation, password checkers have been deployed as password-meters on many websites. The checkers measure the strength of a given password usually by means of regarding the length and character-set complexity, or detecting weak patterns such as common words, repetitions, and keyboard spatial sequence. Xavier and Mannan evaluated a large-scale of password strength meters on many mainstream websites and indicated the weaknesses and inconsistencies of these meters, for example, weak passwords being labeled as strong passwords. [15]
However, even a strong password cannot ensure a high security level, because many other less frequently considered types of attacks can deprive the user of all the advantages. These attacks include phishing, social engineering, key logging, shoulder surfing, etc. [16]

2.3 Human Memory

Atkinson and Shiffrin proposed an influential theory of human memory in 1968, as shown in Figure 2.1. [17] In this model, the input information is first processed by a parallel series of temporary sensory memory systems, including iconic and echoic memory processes [18]. And then the information flows into the short-term store that not only feeds information into or out of the long-term memory store, but also serves as a global workspace that selects and operates strategies, and rehearses information. Atkinson and Shiffrin also found that verbal items could be encoded into long-term memory by rote rehearsal.

![Figure 2.1 Model of Human Memory](image)

2.3.1 Short-term Memory

Short-term memory refers to a temporary store of a small amount of material over periods of a few seconds. And working memory is an extension of short-term memory, which refers to a memory system that underpins our capacity to “keep things in mind” when performing complex tasks [17].
Multicomponent model

The initial working memory model proposed by Baddeley and Hitch consists of three components: the phonological loop, the visuo-spatial sketchpad, and the central executive [19].

In this model, the phonological loop refers to a subsystem of working memory that holds sequences of acoustic or speech-based items. It is assumed to have a limited short-term store and an articulatory rehearsal process. Verbal items can be registered into the store as memory traces, and they decay within a few seconds. But the memory traces can be refreshed by repeating the item in a vocal or subvocal rehearsal that belongs to the articulatory process.

The other subsystem of working memory is visuo-spatial sketchpad that temporarily holds visual and spatial information. Visual STM is suggested to have a limited capacity as well according to Phillips (1974) and Luck and Vogel (1997) [20, 21]. But in contrast, on Conezio and Haber’s study in 1970, they proved that the capacity of visual LTM appeared to be extremely large [22]. In addition, a series of studies also suggests that people can retain many details of pictures rather than a broad semantic gist of a large number of stimuli [23], and it also indicates that visual LTM can store a huge amount of complexity.
2.3.2 Long-term Memory

In general, there are two types of long-term memory: declarative memory and non-declarative memory. The declarative memory refers to facts and events, and non-declarative memory refers to skills and habits that are hard to describe. The declarative memory again has two components: episodic memory that links to events, and semantic memory that links to facts. Evidence suggested that the learning of verbal material depended on the depth of processing: the greater the process has been, the better the subsequent memory. A greater process of verbal information benefits from the retention of the meaning. This process involves semantic coding that generates a richer and more elaborate code, which in turn is more readily retrieved.

Another feature that promotes the remembering of verbal material is visual imagery. In general, it is easier to remember words that are concrete and more imageable than abstract ones. Allan Paivio studied this effect of imageability and explained in a dual-coding hypothesis that imagable words like the name of concrete objects can be encoded in terms of both the verbal meaning and the visual appearance. Therefore, the words can be retrieved in both the verbal and non-verbal routes from memory. If one route decays, the retrieval can still recover the information from the other route. In a dual-coding mechanism, he assumed that two separate systems processed verbal information and nonverbal information.

![Figure 2.3 Dual-coding scheme](image)
individually. He called verbal entities like speech and text as “logogens”, and non-verbal entities and mental images as “imagens” [24].

There are two types of connections between logogens and imagens in the system. Within a single verbal or non-verbal system, the items have associative connections with each other, forming a verbal-verbal connection or a non-verbal-non-verbal connection. For example, the word school may associate with words like teacher, book, boredom, etc. Between the verbal and non-verbal systems, logogens and imagens have referential connections that enable the operations like imaging to words or images to words, and words to images [25].

2.4 Graphical Passwords

The biggest motivation for graphical authentication mechanisms is that our brain has a much stronger memory for graphic information than it does for textual information. Initially, Paivio described a “picture superiority effect”, and suggested that pictures were encoded using a dual coding mechanism consisting of verbal representation (logogen, word generator) and imaginal representation (imagen, image generator) [27]. Compared with textual passwords, graphic passwords can be recalled and activated either in a verbal route or a non-verbal route, but textual passwords can only be reached in a verbal route. The route to passwords stored in long-term memory decays. If the route is forgotten, the passwords cannot be retrieved and accessed. But if the retrieval item is visual, there will potentially be multiple routes to access it, and the retrieval failure of one route does not render the item unreachable.

Many graphical authentication mechanisms are proposed, but very few of them are used in practice, e.g. Windows 8’s picture password and Android’s lock-screen pattern. Renaud et al. classified some current graphical authentication mechanisms into 3 categories according to the different password retrieval methods, and evaluated the vulnerabilities of these mechanisms.
### 2.4.1 Password Retrieval

#### Table 2.1 Methods for Password Retrieval

<table>
<thead>
<tr>
<th>Retrieval</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall</td>
<td>Information is retrieved from memory. E.g. textual passwords, and drawmetric graphical passwords that need to draw a secret pattern.</td>
</tr>
<tr>
<td>Cued Recall</td>
<td>Information is retrieved from memory when cued. E.g. locimetric graphical passwords that need to click a position on a picture.</td>
</tr>
<tr>
<td>Recognition</td>
<td>Information is presented and the individual is able to identify the correct item. E.g. searchmetric graphical passwords that need to pick a picture from a collection of distractors.</td>
</tr>
</tbody>
</table>

### 2.4.2 Graphic Authentication Mechanisms

**Recall based mechanisms**

**Draw-a-secret (DAS)** provides authentication to individuals by drawing a secret pattern on a regular grid of size $G \times G$ [28]. For example, if a user draws a pattern, a sequence of coordinate pairs will be generated as $(2,2)$, $(3,2)$, $(3,3)$, $(2,3)$, $(2,2)$, $(2,1)$, $(5,5)$ where $(5,5)$ represents a “pen up” indicator — a stroke. If a new part of the pattern is added after a stroke, then the sequence of coordinate pairs will be appended ending with $(5,5)$. However, if two different drawings cross the same sequence of grids with strokes occurring in the same places, they will have an equivalent encoding.

This approach is meant to be memorable because it relies on visual, lexical, and kinaesthetic memory [29]. As shown in the figure above, DAS relies on the user remembering a visual pattern of a drawing, but not a sequence of characters. However, a real concern proposed by Thorpe and van Oorschott [30] is that users’ drawing habits could largely reduce the password-space of DAS. They conducted an informal user study on 16 participants and found that 45% of the secret drawings were symmetrical, 56% were centred, and 80% used fewer than 3 strokes. Although a complex drawing with many strokes seems to be much more secure, it may increase the memory load and the error rate of repeating the drawing. The lock-screen pattern of Android is the latest version of a drawmetric mechanism, but it is still vulnerable to shoulder attacks and observations of smudges on the touchscreen [31]. Lin et al. proposed a Qualitative DAS (QDAS) to solve the observation problem...
[32], and Chakrabarti et al. proposed another method to solve this problem by rotating the canvas [33].

griDsure is another graphical authentication relying on knowledge of a secret pattern [34]. This mechanism requires a user to create a secret pattern on a 5*5 grid. Each cell of the 5*5 grid generates random numbers from 0 to 9 for each authentication attempt, and the random numbers are not unique to a cell. However, Bond claimed that this approach can be cracked using only two forged authentication grids [35].

Cued-recall based mechanisms

Background DAS (BDAS) is proposed by Dunphy and Yan to solve the problems of forgetting secret drawings of DAS, especially the sequence of strokes. [36] A background image is presented together with a grid of cells, cueing the secret drawings and strokes. However, users still tend to choose weak passwords. [37]

PassPoints, proposed by Wiedenbeck [38], requires a user to choose five click points on an image which can be selected from a library or provided by the user. The sequence of the click points acts like the secret key. Although each click point has a tolerance radius of clicking accuracy, the password-space is not reduced because an image can theoretically provide so many clickable points that can ignore the tolerance space. However, the password-space can be reduced by “hot-spots” which refers to popular click points, and patterns of choosing points. Researchers believe that the sequence of points is predictable. [39] In addition, Dirik et al. developed a model to discover the popular regions for points, and proved that they were capable of extracting 70% to 80% of the points. [40]

Cued Click Points (CCP) is an authentication approach based on PassPoints [41]. It is developed to release the problems of “hot-spots” and popular positions inherited from PassPoints. During the authentication, an individual needs to click somewhere on a series of images with one point on each, instead of several click points on one image.

Persuasive Cued Click Points (PCCP) is a developed version of CCP which aims further to release the problem of popular positions on each image [42]. A feature called persuasive viewport is added to the original CCP to restrict the range of a click point. In the registration phase, a random viewport is generated which confines the area of selectable points. The viewport is emphasized clearly by lightness contrast,
and users can only choose a point from within that area, yet they can shuffle to change the position of that point.

**Recognition based mechanisms**

**Deja Vu**, proposed by Dhamija and Perrig, requires a user to choose 5 abstract images as secret keys and then to pick them from a set of images during the authentication process [54]. The abstract images are generated by a mathematical formula, so that the system just stores the initial seeds for the formula but not the actual images. Dhamija and Perrig compared the performance of Deja Vu and PINs, and they found Deja Vu was easier to use but at the expense of time and security. In addition, they reported that the image-type is related to the security of the password. For example, users tend to select the same images when using semantic images such as photographic scenes.

**PassFaces** [43], a commercial graphical authentication mechanism, assigns a user 3 face images as secret keys. During the authentication process, a grid of 9 faces is presented, and one of the faces is from the secret keys. The user needs to identify that face among 8 distractors.

**Convex Hull Click (CHC)** is a recognition-based graphical authentication system proposed by Wiedenbeck et al. [44]. The secret keys are 5 icons or pass-icons. During the login process, there are several challenges that need users to click within a convex hull mentally created by their pass-icons. A convex hull is formed by at least 3 points (pass-icons), and the lines are not visible but formed in mind. Users need to identify a convex hull and click within it to go to the next round. Wiedenbeck found that this approach was lengthy, but it could resist observation, and the process is meant to delight and excite users to keep them interested in this approach.

---

**2.5 Game Design**

**2.5.1 Iterative Design**

Iterative design is a play-based process. Emphasizing play testing and prototyping, iterative design is a method by which design decisions are made based on the experience of playing a game while it is in development. In an iterative methodology, a rough version of the game is rapidly prototyped as early in the
design process as possible. The prototype should focus only on the interaction and not on the aesthetic trappings at this stage. This prototype is played, evaluated, adjusted, and played again, allowing the designer or design team to make decisions on the successive iterations or versions of the game. Iterative design is a cyclical process that alternates between prototyping, play testing, evaluation, and refinement.

### 2.5.2 The Pyramid of Game Design

Game design can be divided into three levels like a pyramid. The higher levels provide game experience based on the lower levels.

On the lowest level, Components are the basic objects or structures that provide interaction (operation and feedback) with users.

In the middle, Mechanics are the various actions, behaviors and control mechanisms afforded to the player within a game context. Together with the game’s Components (levels, assets and so on), the mechanics support overall gameplay Dynamics.
At the highest level, the Dynamics contribute to the overall game experience in aesthetic terms. This level can provide “fun”, and can prompt users to keep playing a game.

The aesthetic part is the most important part of a game, and if the game is not attractive at this stage, it is likely to be a boring game. Aesthetics are bound tightly to the Dynamics of a game, as the Dynamics usually engage users to play by providing some engagement and motivation. This engagement can be through learning and practicing the game, and the motivation can be the exploration of the game story (or journey), or achieving certain rewards.

A game itself consists of Components contributing to the rich experience of gameplay, and the pleasure and fun come from the whole structure of the game, including its rules and systems. The counterparts to rules and systems are Mechanics and Dynamics, which build a basic structure and create various interactions in a game. The ultimate goal of a game is to evoke positive aesthetics and thus to deliver fun.

A password game should be fast and simple so that it does not take much time. Games like Zombies & Plants and Fruit Ninja are fun because they are strategic and highly controlled. To some extent, these two games exercise people’s brains and hands whilst testing people’s ability, and people can confirm this themselves during the games. That’s probably why some people cannot stop playing such games, and others drop the games after a short time. In a short game time (5-10 seconds), it is hard to develop a password game involving strategies or skills. Finding and recognizing may be suitable Mechanics for password games.
2.6 Theory of Fun

2.6.1 Lazzaro’s 4 Keys to Emotions

Lazzaro developed 4 keys that can boost emotions when playing games [47]. These 4 keys are not defined as 4 types of fun but 4 aspects that lead to different types of enjoyment and pleasure.

**Hard Fun: Emotions form meaningful challenges, strategies, and puzzles.**

Players may play a game because it is hard, and the achievement of overcoming an obstacle creates a positive emotion that is related to fun and pleasure. Challenges, and testing one's skills, may boost players to develop and apply creative strategies, or push them to practise their operational skills, e.g. action games or shooting games. At the end, the feedback on progress, rewards, or a winning performance in the game lead to a sense of accomplishment for players.

Players who enjoy hard fun may use a game to prove how good they are and test their strategies. A game usually has different levels where the difficulty of the challenge matches the skills of novices and advanced players alike. The transition of difficulty prevents novice users from feeling stressed and bored.

**Easy Fun: Grab attention with ambiguity, incompleteness, and detail.**

Players focus on experiencing the environment created by the game and then immerse themselves in the game world with ambiguity, incompleteness, and a focus on the detail in the game. Players may describe their “easy fun” as being like exploring new worlds with intriguing people, excitement and adventure, wanting to figure out this world, seeing what happens, liking the sound, feeling like the user and the character are one, etc.

**Altered States: Generate emotion with perception, thought, behavior, and other people**

Players get fun and pleasure through altered states such as excitement and relief during and after playing a game. Their enjoyment is produced by internal experiences in reaction to visceral, behavioral, cognitive, and social properties in games. Players may describe this experience as being like clearing their mind by clearing a level, feeling better about themselves, avoiding boredom, being better at something that matters, etc. Examples of game that can alter states are Collapse, Crosswords, Halo, GTA, Civilization, etc.

**People factor: Create opportunities for player competition, cooperation, performance, and spectacle**

In multiplayer games, some players may think the most attractive factor is the interaction with other players. Competition, teamwork, and communication promote enjoyment and make the game experience meaningful. In addition, players have the opportunity to gain social bonding and personal recognition whilst playing
with others. Players who enjoy the people factor may say things such as “it is the people that are addictive, not the game”; “I want to invite my friends”; or “I don’t play, but it is fun to watch”. Examples of this kind of game are EverQuest, Halo, GTA, etc.

Lazzaro’s theory considers fun from the players’ aspect, and explores the emotion generated through gameplay. In contrast, LeBlanc’s theory considers this more from the game designers’ aspect, which directs the user on how to accomplish different types of fun during the process of implementation [48]. This theory is related to the pyramid of game design, and the 8 kinds of fun are the delivery methods of the aesthetics created by gameplay. To figure out what fun is, Koster considers Lazzaro’s theory and explains it like this [49]:

<table>
<thead>
<tr>
<th>Fun</th>
<th>is the act of mastering a problem mentally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aesthetic appreciation</td>
<td>is enjoyable, but is not always fun.</td>
</tr>
<tr>
<td>Visual reactions</td>
<td>are generally physical in nature and relate to physical mastery of a problem.</td>
</tr>
<tr>
<td>Social status manoeuvres</td>
<td>of many sorts are intrinsic to our self-image and our standing in community.</td>
</tr>
</tbody>
</table>

### 3. DESIGN

#### 3.1 Gather Requirement

This project aims to develop a game that performs like a password, authenticating users in login process. In general, a password authentication mechanism requires users to create an account profile with at least an username and a secret password. The password should be confidential to other people to ensure the power of the authentication. Every time users login to the system, they need to input their passwords. The authentication is fulfilled by verifying whether the password is right. Therefore, a weak password can seriously undermine the security level of authentication. The basic process of an authentication mechanism consists of register and login, and the basic elements are an username and a password. To create a game to substitute a textual password, the game need to contain a secret pattern that authenticates users to login to the system, but the secret pattern needs
to decrease the memory load to remember it compared with a textual password, and the overall operation of inputting the secret pattern need to be short and fun.

The prototype of a game authentication mechanism requires two basic functions:

- **Register**
  - Users need to create an account to identify themselves into the system.
  - The account consists of an username and a secret pattern.
  - The secret pattern authenticates users based on something they know.
  - The secret pattern is not hard to remember, but aims to ease the memory burden of remembering.
  - The secret pattern needs to be hard to describe so it will be hard to guess, e.g. intruders cannot guess at the secret pattern by analyzing their personal life form any social applications.

- **Login**
  - Users need to input their username and play a game to login to the system.
  - Users need to identify their secret pattern by playing the game.
  - The rules of the game is clear and simple to decrease extra burden of users’ memory.
  - The process of playing the game is quick and short.
  - The game protects the secret pattern from shoulder surfing by other people.
  - The game has a mistake tolerance mechanism to avoid annoying users if they make mistakes accidentally.
  - The game aims to be fun in gameplay.
3.2 Evaluate Games

Considering the password game needs to be simple and quick, the game itself needs to be light and loaded quick as well. Therefore, the implementation of the game is confined to small-sized games with simple operations. To implement a prototype of the password game, it needs to determine 3 key elements, including rule, control, and secret pattern. The ideas firstly come from evaluating popular games, shown in Table 5.

In flappy bird, players feel fun to control the bird to go through obstacles and get higher points. During the gameplay, the game tests players’ coordination of finger and brain, and they can practice their skill to get better performance and in turns get better ability of coordination. In this game, the extrinsic motivation for players is to get higher points, and the intrinsic motivation is to practice their ability of reaction and coordination of their brains and fingers. The result of the game not only gives a feedback of the gameplay, but also reflects the performance of their ability. For most players, this game is too challenging to play so they may stop playing over a short time. But for professional players, they like to play it and achieve hard fun of challenging. For a password game, it is better to have a simple control and game context, but it cannot be too hard to play.

In Fruit Ninja, the fun comes from cutting multiple fruits to get higher points without cutting a boom. The control is different from flappy bird, but the operation is as simple. This game also tests and practices the ability of human brains. Players need to cut as more fruits as possible by one slide, yet they need to be attention about booms. The higher points players get, the better skills they have, which reflects they get better ability of coordination and reaction. In this game, the 1-minute time limit plays an important role of gameplay to form the basic rules and give time pressure. But one minute is too long for a login process, so the password game is not suitable to be implemented like this kind of game.

In Angry Bird, players get fun from making strategies to kill enemies. The strategy is to find a good angle to project the bird and break the enemies’ nest. It is not necessary to kill all the enemies to complete a level, but in that way players can get all the 3 stars. The difficulty increases in higher levels, but it is more difficult to get 3 stars for later levels. The intrinsic motivation to play this game is to practice problem solving, and the extrinsic motivation for some players is to get all 3 stars for all the levels. However, it is hard to make a secret pattern in this game as the
outcome of projecting a bird is unpredictable. For example, players cannot hit a specific spot by projecting a bird every time even with an error tolerance. Therefore, it is hard to distinguish whether the error is made by the authorized users or intruders. But the idea of levels is acceptable in a password game.

In Candy Crush, the fun comes from making choices of crushing candies to get more points. The trick is that players need to choose from one or more possible choices to get special sweets to earn more points within a time limit or step limit. The limit constrains the gameplay and constructs an important part of the game experience that allows players to make strategies to optimize their choices. However, there are only 6 kind of basic candies, and it takes at least 3 candies in a row or column to crush them. It is hard to make a secret pattern in this game because it will relate too many objects and still have limited possible combinations. For example, if the secret pattern is to remove 3 groups of yellow candies, there are $6 \times 6 \times 6 = 216$ possible combinations for intruders to guess. But the removal of one group of candies affects 3 or more candies, if there is no possible combination for the target candies, players need to remove other candies to wait for the target combination, which means they spend more time. The security level can be increased by removing more groups of candies, e.g. 5 groups of yellow candies, indicating that players may waste more time on waiting for the right combination. Still 1 minute is too long for a login process, while in a few seconds players may not be able to complete a secret pattern in a game like this. Yet the idea of step limit can be used to constrain the gameplay.

Peak is an application with a set of small games aiming to improve the ability of human brains. It is fun because it is meaningful for players to exercise their brains and improve their mental skills. The games require players to spend mental efforts to solve problems or identify targets, and the difficulty is adjusted according to players' performance. Although the games are more like tasks, players can get fun from achieving higher scores during the gameplay, and they can also feel better if they break previous records which reflects their mental skills are improved. Also, in the games with moderate difficulty, players can get into flows of solving problems and get simple pleasure of making progress. For a login process, it is better to be quick and simple, so the idea of a fast-paced remembering and picking mechanism can be borrowed. Additionally, it is better to create a meaningful play that can practice players' skills, so players can get fun from achieving better performance.
### Table 3.1 Evaluation of Games

<table>
<thead>
<tr>
<th>Games</th>
<th>Rules</th>
<th>Controls</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flappy Bird</td>
<td>- Players control a bird to fly through tubes.</td>
<td>- Players click on screen to control the bird.</td>
<td>- The game practices players’ reaction of rich stimuli.</td>
</tr>
<tr>
<td></td>
<td>- Score is counted by the number of passed tubes.</td>
<td>- The bird drops down by gravity during the game.</td>
<td>- The game is challenging and practicing players’ skills of coordinating mind and fingers.</td>
</tr>
<tr>
<td></td>
<td>- Difficulty is increased by adding more clustered tubes and enemies.</td>
<td>- The bird flies higher counteracting the gravity by one click on the screen.</td>
<td>- The control of bird with gravity is challenging that boost hard fun.</td>
</tr>
<tr>
<td></td>
<td>- No time pressure.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Need skill.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fruit Ninja</td>
<td>- Players slide the touch screen to cut out fruits.</td>
<td>- Players slide fingers as knives to cut fruits.</td>
<td>- The game also practices players’ reaction of rich stimuli as flappy bird.</td>
</tr>
<tr>
<td></td>
<td>- Score is counted by the number of cut fruits.</td>
<td>- Multiple fruits can be cut by one slide and generate more points.</td>
<td>- The challenge is to achieve more scores within 1 minute which practice players’ skill to cut multiple fruits in one slide.</td>
</tr>
<tr>
<td></td>
<td>- Difficulty is increased by booms which result in game-over.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Time is limited.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Need skill.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angry Bird</td>
<td>- Players control several birds as shells by breaking physical structures (enemies’ nest).</td>
<td>- Players launch a bird like a shell to kill enemies by demolishing enemies’ nest.</td>
<td>- The game practices players’ strategy to break a physical structure and kill enemies.</td>
</tr>
<tr>
<td></td>
<td>- Score is counted by the number of killed enemies.</td>
<td>- Players can adjust the angle of the launch of a bird.</td>
<td>- The challenge is to find a good strategy to break the enemies’ nest and kill more enemies.</td>
</tr>
<tr>
<td></td>
<td>- Difficulty is increased by adding more enemies.</td>
<td></td>
<td>- The more enemies are killed, the more stars are earned.</td>
</tr>
<tr>
<td></td>
<td>- No time pressure.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Need strategies.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candy Crush</td>
<td>- Players clear candies to earn points.</td>
<td>- Players need to move 3 same candies together to clear them out.</td>
<td>- The game practices players’ strategy to get a certain amount of points.</td>
</tr>
<tr>
<td></td>
<td>- Score is counted by the number of cleared candies.</td>
<td>- Players need to spend more effort to clear out barriers.</td>
<td>- The challenge is to clear candies that can earn more points.</td>
</tr>
<tr>
<td></td>
<td>- Difficulty is increase by earning more points and adding barriers.</td>
<td>- Players can create special candy to earn more points.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Time (or step) is limited.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Need strategies.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To sum up, four ideas can be borrowed from popular games to implement the password game:
1. **Game constrain**: Use a step limit or a round limit instead of time limit because the login process needs to be quick.
2. **Game control**: Use simple operation to play the game like quick picking of targets.
3. **Outcome**: Create meaningful play to practice players’ skills.
4. **Game type**: Implement a quick and simple task-like game instead of a strategy game involving complex objects. Because the complexity of affected objects regarding to the secret pattern may hinder the process of the verification, and it is necessary to increase the number of distractors to increase the difficulty of guessing the targets.

### 3.3 Iterate Ideas
After reviewing popular games, some ideas are generated to implement the password game for the login process. At the beginning, the idea is to implement games with background stories that create an extrinsic meaning of gameplay. Because the main purpose of the game is to login to the system, serving as a extrinsic motivation for users to play the game, it not necessary to boost intrinsic motivation to immerse into the game. But at the last, the idea tends to implement a task-like game that exercises skills and prompts intrinsic motivation to play.
3.3.1 Stealing shoes

The first idea is to implement a game relying on keyboards. The story of the game is based on a cartoon, Usavich [33]. A main character, Kirenenko, who wears a prison uniform with red strips, number 04, is very dangerous when he becomes angry and berserk. It is fun to watch the cartoon where Kirenenko goes crazy and beat enemies heavily, especially to see his angry face. He collects sneakers and loves them a lot so that he will become crazy if other people touch his sneakers. So the story is to steal Kerenenko’s sneaker into a well and make him angry.

Rules:
- Users register an account with at least 3 pictures, and they need to remember the order of the pictures.
- During the login process, the pictures are generated randomly together with random key indicators (e.g. up-down-left-right).
- Users need to find out their secret pictures and type the key indicators in sequence of the determined order.
- There is a status bar to indicate a progress. When users complete typing the indicators under one picture, the color of the status bar will become green, and the sneaker will approach closer to the well.
- When users complete typing the indicators under 3 pictures in a right way, the sneaker is stolen and Kirenenko is angry. Then users can login to the system.
- There is an error tolerance in case users make mistakes unexpectedly.
• Advantages:
1. It is easier to identify pictures than to recall textual passwords.
2. The game can practice fingers.
3. It is hard to be observed by people nearby to type the indicators on a keyboard.
4. It is cannot be cracked using leaked password dictionary.

• Disadvantages:
1. Users cannot get fun from stealing the sneaker or making Kirenenko angry. For one thing, users may not know Usavich so that they can neither understand the background story nor feel fun to see Kirenenko’s angry face.
2. Although users need to type a long sequence of indicators, the number of possible situations is not secure enough as intruders just need to guess 3 pictures in some order. The number of all possible situations is $10 \times 9 \times 8 = 720$.

This idea cannot be used because many people are not familiar with this background story so that they will not think it is fun. Also the number of possible situations is not enough to secure the authentication. But the idea of using picture cues can be implemented.

3.3.2 Tree game

Figure 3.2 Tree game
The second idea is inspired by gambling and electronic pet. When people play a fruit machine, they have a chance to win a round yet the possibility is low. But if the reward is attractive, people keep playing it. The pleasure comes from getting reward of a win-state result as well as achieving a small possibility event. So it is feasible to construct a game mechanism that combines components of possibility and reward. When raising an electronic pet, it is necessary to feed the pet regularly and this function is usually performed by clicking a “feed” button. Considering the idea of creating new elements in games like “Little Alchemy”, the act of clicking a “feed” button can be replaced by clicking many elements to create an effect [18].

The story in this game is to fertilize a tree. The tree is like an electronic pet that needs players to “feed” it, and it starts from a seed and grows up according to its level. The fertilizer is mixed by more than 3 ingredients, and the special recipe acts like a secret pattern that authenticate an user. Different recipes generate different ranges of points that contribute to the value of levels, and every time people mix their ingredients to fertilize their trees, it generates different points in a range according to their performance. For example, if a user mixes his recipe within 3 seconds, it will add more points from 0 to 10 to the basic points of this recipe. Additionally, three characters— width, height, and number—can be used to show more detailed status of a tree, and they are triggered randomly.

Rules
- Users need to register an account and choose more than 3 elements as their secret pattern. In other words, users need to choose some special elements to create a special recipe to fertilize their trees.
- The elements are divided into many categories and related to routine life, e.g. Chopin’s nocturne, KitCat, Titanic movie.
- During a login process, users need to pick out the elements to create their recipes and click “mix” to fertilize the tree.
- Creating a right recipe generates some points contributing to levels. And the figure of the tree will change according to the level status. For example, there is only soil in level 0, and a small bud in level 1.
- Different recipes generate different ranges of points.
- A random value is added according to special achievement of the performance (e.g. fertilizing within 3 seconds, continuous right recipes).
- Width, height, and number are changed randomly every time.
- There is an error tolerance machines to occasional mistakes.

Advantages
1. Raising an electronic tree may interest more people.
2. People can make odd recipes of the fertilizer, e.g. Chopin’s nocturne, lemon juice, and salt.
3. People can get easy fun of creating recipes of the fertilizer, and exploring a more powerful recipe.
4. So many possible combinations increase the security level.

Disadvantages
1. It is complex to implement this game.
2. It is very subjective to create the library of basic elements.
3. Too many random values may confuse users and users cannot get a clear feedback from the status.
4. Marking the chosen elements and clicking mix later also facilitate shoulder surfing.

This idea increases the number of possible combinations of a secret pattern, but it is hard to implement. To build the mechanism of generating different ranges of points in different recipes, it is necessary to create an algorithm of generating points towards different combinations. The algorithm together with the creation of the library of basic elements are subjective. In terms of the operation of this game, it is easy to be observed if the elements are marked obviously. The pictures of the elements need to be clear but with limited size, so it is necessary to have textual tags to indicate the items. In this game, the exploration of recipes plays a more important role than the operation skills and making strategies, so the art and look of this game is crucial to create a good aesthetic experience.

### 3.3.3 Tangram and Chinese characters

**Tangram**

![Tangram Image](image.png)

Figure 3.3 Tangram
To extend the idea of tree game but remove the complex algorithm of generating points, the idea of this game also authenticates users by clicking the right pictures. Tangram is a traditional Chinese dissection puzzle consisting of 7 flat shapes, and the 7 pieces of shapes can form many different patterns. Users need to choose 3 patterns as their secret keys, and they need to pick out their target pattern to login. In contrast to tree game, this game is more dynamic that users need to click the target within a time limit, otherwise the pictures are blocked in a brief time. During the login process, a set of 12 tangram patterns are generated randomly, and two target patterns are from users’ secret keys. Users can login to the system if they pick out the two targets. This game is much simpler than tree game; it does not have a background story, but focuses on completing a small task.

Rules
- Users register an account with an username and 3 tangram patterns as their keys.
- During the login process, 12 tangram patterns are generated randomly, and two of them are users’ secret keys.
- After a brief time, all the pictures are blocked.
- Users need to pick out the two target tangram patterns before the pictures disappear.
- Users can login to the system by clicking the two targets, and each successful login can get 10 points.
- It is better to develop a features of unblocking titles to make the point meaningful.
- There is an error tolerance of 3 chances to try.

Advantages:
1. The game is simple and easy to operate.
2. The game is a bit challenging.
3. The simplicity of tangram patterns creates aesthetic value.
4. The choice of secret patterns is not related to personal information or common words. So it increases the difficulty to guess the secret patterns.
5. The pictures and colors make up the unavailability of the art design of the game.
6. To block the pictures after a brief time helps prevent from observation and shoulder surfing.

Disadvantages:
1. It may too challenging to pick out 2 target patterns within a brief time.
2. It may not be easy for users to remember 3 tangram patterns as they are similar to some extent, and are not very concrete.
3. The number of tangram patterns is limited, so to make more patterns, different colors and a turn to some angle can be used to enlarge the library. But in turns, it may increase the challenge of picking out the targets from many similar patterns.

4. People can observe two patterns in one round.

**Chinese characters**

An alternative idea of this game is to change the pictures of tangram patterns to Chinese characters as they are just like drawing patterns. In a separate page, it is necessary to make detailed explanation of every characters to build a memory trace of verbal encoding based on meaning. The rules of this game are same to tangram game, but after the pictures disappear, the grid of pictures rotates 90 degrees to increase the difficulty.

**Advantage:**
1. People can learn at least 2 Chinese characters. It may interest people who like Chinese.
2. It is easy to identify a character in a relatively small size, so in a same space, the number of characters is larger than the number of tangram patterns. Therefore, there are more possible situations to guess.
3. The rotation makes it more challenging and more like a game.
4. It may be hard to guess secret characters of an user because people may choose them relying on either the meaning or the look. And it may be hard to crack the secret characters according to any leaked passwords dictionaries or personal information, because it is a foreign language.

5. To block the characters and rotate the grid helps prevent from observation and shoulder surfing.

Disadvantage:

1. This method is recommended to people whose mother language is not Chinese.
2. The library of characters is subjective.
3. It may be too challenging to pick out 2 target characters under the condition of blocking and rotation.
4. People may have a preference to choose several characters that are simpler or much more complex, so they can pick them out easily from distractors. But the preference reduce the difficulty to guess and crack the secret characters.

To sum up, the idea of tangram game is used to develop a prototype of game authentication mechanism, but to balance the challenge of picking out target pictures, the pictures should be more distinct and concrete to facilitate remembering. In terms of the challenge of picking out two target pictures within a limited time, it can be changed to choose one target a round to decrease the difficulty and avoid the exposure of observing attacks.

4. PROTOTYPE

This section will explain the development of a game-based authentication mechanism. The prototype, named Find that Animal (FTA), is developed from the idea of the tree game. Also included are the types of languages and tools that are used to implement the prototype. Two versions of the prototype were created during the implementation phase. In particular, the second version is used for testing FTA, integrated with experiment instructions and a concert form (page). This version improved the usability of the previous version for the purposes of testing.

4.1 Implementation technology

This project implemented the prototype according to the web design process. The technologies used are closely related to web development. The documents of the prototype are cached on the university server under the student account. Also, the
data generated from participants is kept on a database that extends from the university server.

### 4.1.1 Coding languages

To implement the web-based prototype, five types of languages are involved. On the client side, html and CSS are used together to build the basic structure of the web pages, arranging the content, layout, and style of the components. Javascript is used to develop the actions of the web pages, including client side validation of forms, the feedback action of clicking, the blocking of pictures, etc. On the server side, all the html documents are saved as php documents on the student account of the university server. Php is used to implement the complex functions of retrieving and sending data to the database on the server, and also the function of generating a grid of random pictures. In particular, SQL commands together with php language are used to manipulate the database on the server side.

### 4.1.2 Tools

During the coding process, Bracket is used to write the coding documents, and Google Chrome browser is used to check and debug the web pages. The coding work complies with the compatibility problems across different web browsers according to the guidelines on W3school [51]. To support the data transactions of registration and login processes, a database is created on the university server under the student account, and it can be manipulated remotely by terminal.

### 4.2 Implementation Processes

#### 4.2.1 Scenario-based Design

This project aims to test the game-based authentication mechanism so that it may not be easy to identify target users. However, as the idea of game-based authentication mechanism is a new concept and still at an exploratory stage, and as the participants for testing the prototype are likely to be young people from around the campus, it is better to consider the implementation of web pages more from the
perspective of young people who are capable of using new applications on computers, rather than of people who cannot use computers well. In addition, the main purpose of the implementation is to build the game as a prototype, and to test whether this game can work well, instead of building web pages of services for target users. To help develop the basic functions of the prototype and test the usability problems, it is necessary to create a persona to help walk through the prototype and evaluate it.

**Persona**

Alex is an undergraduate majoring in chemistry at the University of York. He is a professional user of computers and smart devices. When he feels tired of studying, he likes to play games on his phone to kill time. Although he is not a fan of huge social games such as “War of Warcraft”, he sometimes plays smaller single-player games such as “Need for Speed”. He has registered many accounts for more than ten websites, but has forgotten his registration details for some of them that are not often used. To solve this problem, he just registers on new websites using his Google plus account. He rarely changes his passwords unless there is a notification of irregular login indicating security attacks.

**Scenario**

Alex received a request email regarding the testing of a prototype of a game-based authentication mechanism. The email introduced the instructions to use the prototype with a link to the website. Alex clicked the link and he was directed to a login page requiring a username. He noticed a link for registration under the login box and clicked it. On the registration, he filled in the information of username and email address. In particular, he chose three pictures as the keys for his account. After he finished the registration, he clicked the “Finish” button to proceed. He then saw the regular login page with a game. Although he knew the rules of the game from the instructions, he still made a mistake. He logged in successfully to the system when he picked out the right picture. He thought this testing was finished so he closed the website.
## Claim Analysis

### Integrate instructions on the pages

+ Facilitate a fluent flow when using the websites
+ Avoid confusion over procedures and operations
- Long sentences may take up too much space
- Inappropriate words may lead to confusion

### Give a training section of the game after registration

+ Let users practice the game
+ Help remember the secret pictures
+ Avoid mistakes on the game resulting from a lack of knowledge
- Users need to click a link to go to the game in the regular login process
- Users may not know they need to get out of the practice page

### Highlight the login status

+ Give a clear indication of whether users have logged in to the system successfully
+ Big titles attract users’ attention
- Users may not notice other content on the page

### Give clear feedback of registration and login

+ Validate data on the registration form
+ Indicate repeated usernames on registration
+ Indicate invalid username on login
- Need to connect to a powerful database
- Need to use PHP language to implement server-side functions

### Give a setting page for secret pictures

+ Check the secret pictures
+ Change secret pictures
- Easy to be observed

### Give a clear feedback on the game

+ Tell users whether they pick the right target
- The type of feedback is limited
- May not be supported in different web browsers
As shown in this Figure, 6 web pages construct the prototype.

- **Login page**: The login page contains a login box that requires a username. A verification of the username is processed if users proceed to login. New users need to register an account beforehand. An existing user can just type a valid username and proceed, and then automatically go to the game page that performs the regular login process.

- **Register page**: To create an account on the register page, users need to fill out a username and email address twice respectively. Users need to choose three different pictures as their secret keys. Apart from the validation of the registration form, it is necessary to check whether the username is taken. When a user has filled out all of the data in the right way and has clicked the “continue” button, the page will be redirected to the practice page.

- **Practice page**: On the practice page, users can practice the game but there is no punishment if they click a wrong picture. They can click a link to access the game page to perform the regular login.

- **Game page**: The game page serves as though typing and validating a textual password. When users pick out one of their secret pictures from a grid of

---

**Figure 4.1 Site map for the first version of FtA**
distractors, the page will be redirected to the finish page, which is deemed to show that the user has successfully logged in to the system.

- **Finish page:** This page can only be accessed through the regular login process, whereby users complete a game validation. It indicates that users are logged on to the system, so that they can then manipulate the account information. Users can also go to the practice page if they want to play the game.
- **Setting page:** This page contains secret pictures.

### 4.2.2 Selection of pictures

One issue for the implementation of the prototype is the selection of pictures. This prototype chose funny animal pictures as the basic components because people are likely to generate positive emotions on seeing this kind of picture. Therefore, the positive feelings caused by seeing these pictures boost fun and pleasure during the gameplay, which makes up for the lack of art and aesthetic design in the game. In addition, it is easier to remember concrete pictures as such pictures can be encoded in a visual code as well as a verbal code. For example, if a user chooses Figure 4.2 as a key, apart from remembering the details of the picture, he may also remember it by tagging keywords like cat, big face, and big round eyes, etc. Every time he plays the game, he can recall his key by keywords or a piece of imagery on his mind. The key is in the grid of pictures, and the recognition of the key among distractors is more common than pure recall.

The pictures are collected from Google by searching for the term “funny animal”. The basic requirement for collecting a picture into the library is that the content of the picture is clear; it is easy to identify what kind of animal it is, and what they are doing. The displayed size of a picture is limited because if a picture is too large, the number of pictures in a grid may decrease to fit the screen, which will decrease the possible choices. Equally, if a picture is too small, it will be hard to recognize the picture and therefore to pick out the target secret picture. On the prototype, the size of a picture is 120px×120px, which makes a grid of 16 pictures fit the screen without having to scroll down.

The pictures are collected randomly and share common characteristics, such as a big animal face on which you can clearly see the expression, or a gesture or movement of an animal. The type of animal is not specifically considered. In the collection of 40 pictures that build the library seen on registration, there are different kinds of animal such as dogs or cats, different individual pictures of each kind of animal, and also different pictures of an individual animal. The order of the pictures on the library is not designed on purpose.
4.3 Prototype 1.0

The first version of the prototype was implemented according to the site map. During and after the implementation, changes were made to improve the usability for the purpose of testing. Most of these changes were made to optimize the procedure so the appearance of the web pages did not change too much. On this section, the focus is on introducing procedures and problems of the prototype, and the screenshots of the web pages will be provided in the next section.

4.3.1 Anatomy of the game

The prototype is named Find that animal (FtA) as it uses animal pictures as secret keys, and users need to pick out their target animal during the game. The idea comes from Tangram, so the basic mechanism of picking and blocking is the same.

FtA in login

- Components:
  1. The basic component that constructs the game mechanism is the grid of 16 pictures generated randomly in the center of the computer screen. One of these pictures is the user’s secret picture, and this picture is also randomly chosen from the user’s three secret pictures. For example, users may encounter the same secret picture for three rounds, or they may meet three different pictures for three rounds. In addition, the pictures in the grid are clickable. Once a picture is clicked, the browser will handle the client-side checking and will play different sounds to indicate right or wrong. At the same time, the server will also check whether the picture is the target secret picture and process the data transactions with the database. After the quick checking of the selected picture, the page will redirect to the other web pages.
  2. The grid of pictures will be blocked after 5 seconds.
  3. A reward board is used to show how many challenges have been achieved. The reward is calculated according to how many times users achieve a specific challenge. For example, users can get a star if they login to the system at the first attempt ten times.
• Dynamics:
  1. There are three rounds in the game. Users have three chances to pick out the target picture. Once they pick the right one, they can login to the system successfully. If a user picks out wrong pictures for all three rounds, he or she will be logged out.
  2. Users pick a picture by clicking. They have a limited time to find their target pictures among the distractors before the pictures disappear.

• Aesthetics:
  1. It aims to boost positive emotion when seeing the funny animal pictures.
  2. It builds the challenge of picking out a target from distractors in a limited time.
  3. It requires users to concentrate their attention to find the target.

FtA in practice

To let users practice the game and remember their secret pictures better, slight changes are made on the rules of play:
- The performance data such as time and selected picture will not be stored in the database.
- Users cannot play another round until they pick out the target picture.
- Users can play as many rounds as they want.

Figure 4.2 A cat picture in FtA

4.3.2 Code Implementation

During the implementation of the coding process, lots of problems were encountered in connecting to the database. One important lesson I’ve learnt from a connection failure of a table is that it is important to ensure the table names are
correct. Because of this simple mistake, time was wasted trying to solve this problem.

There are two methods of presenting pictures. On the registration page, all of the pictures are in a form of a ‘checkbox’, so that users can click three items on the submission of the registration form. On the game practice page and the regular login page, the pictures in the grid are in the form of ‘\'<input type="image">’\’, which submits the game record to the server once a picture is clicked, and the page will then be redirected to the next round.

One problem is in relation to the playing of a sound. The sound is played by a client-side (javascript) function at the same time as a picture is clicked. The sound is played to indicate whether the user has picked the right picture, but a redirection may interrupt the sound after the process of submission. If it takes longer to process the submission of the game record, the page is likely to have enough time to play the sound in full. However, if the process is too quick, the page is likely to play only a small part of the sound. This would result in two different sounds being played, and would confuse users.

Another important problem is that there are compatibility issues on Firefox and Safari. FtA can work well on the Google Chrome browser. On Firefox, it seems that it does not support the session variables used to implement a login state. On Safari, the loading of the game seems to be longer than on other web browsers. Additionally, it cannot play sound. As these compatibility issues cannot be fixed because of my limited programming skill, participants are advised to use the Chrome web browser to login to FtA.

4.4 Prototype 2.0

Changes of rules

The dynamic part has changed due to an improvement of the complexity of FtA. The maximum number of rounds of the game is increased to 5. Users must pick out their target pictures right for 3 rounds to login. The maximum number of error tolerance is 2, which indicates users can still login to the system successfully by clicking two wrong pictures. However, users will be logged out if they pick wrong pictures for 3 time. In this way, the repeated rounds cannot only increase the difficulty of guessing, but also increase the challenge of the game, which can in turns elicit a sense of achievement. FtA in practice is the same one in the login
process. So a new round will start no matter whether users clicking a right picture. The difference between FtA in practice and FtA is there is no punishment of clicking a wrong picture in the training section.

**Self-adjusted buffer for blocking time**

The game is a bit static although a reward board is suggested in prototype 1.0. To improve the dynamics between the game and users, a self-adjusted buffer is implemented into the game. It is used to adjust the blocking time of the pictures according to users’ performance. The default blocking time for pictures is also 5 seconds, and the default value of the buffer is 0. The buffer will be increased by 1 if users pick out their secret pictures once and decreased by 1 if they pick a wrong picture. The range of the buffer is from minus 4 to 3 (-4 to 3). And the blocking time is the result of a simple subtraction.

\[
\text{blocking time} = \text{default blocking time} - \text{buffer value}
\]

The range of blocking time is from 2 to 9 seconds, which tries to balance the difficulty for different users.

**Usability issues improved**

Several palaces have been changed for the purpose of testing. In particular, the prototype is built into a much linear series of procedures. Then participants can just follow the instructions and think less.

Changes:
1. Take out links on top. So it won’t confuse participants; e.g. “Log Out | Introduction | Practice | Setting | System | Regular game”
2. Shorten the words on the instructions, unless it is important.
3. Separate long paragraphs into different pages.
4. Make sure users know they need to use FtA on Chrome browser.
5. Make participants know beforehand they can still click the squares after the pictures are gone. And it is the only way to proceed.
6. Require users to register by email address only.
7. Use very different sounds, so it can still give different feedbacks if the browser corrupts the play of the sound.
8. Add a big, clear button on practice page to notice users they need to continue to a next step.
9. Broaden out the space between pictures on the grid, so it will not so difficult to play the game.
## Screenshots

<table>
<thead>
<tr>
<th>Page name</th>
<th>Screenshots</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instruction page 1:</strong></td>
<td></td>
</tr>
<tr>
<td>- introduce the outline of the experiment</td>
<td><img src="image" alt="Fun Security:" /></td>
</tr>
<tr>
<td>- note participants need to use Chrome browser</td>
<td><img src="image" alt="Introduction" /></td>
</tr>
<tr>
<td>- click the button to proceed</td>
<td><img src="image" alt="Outline briefing" /></td>
</tr>
<tr>
<td><strong>Instruction page 2:</strong></td>
<td></td>
</tr>
<tr>
<td>- ask consent for participation</td>
<td><img src="image" alt="Consent" /></td>
</tr>
<tr>
<td>- click the button to proceed</td>
<td><img src="image" alt="Registration" /></td>
</tr>
<tr>
<td><strong>Instruction page 3:</strong></td>
<td></td>
</tr>
<tr>
<td>- introduce the procedures on registration</td>
<td><img src="image" alt="Registration" /></td>
</tr>
<tr>
<td>- click the button to register</td>
<td><img src="image" alt="Registration" /></td>
</tr>
<tr>
<td><strong>Page name</strong></td>
<td><strong>Screenshots</strong></td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Register page:</td>
<td><img src="image1" alt="Register Page Screenshot" /></td>
</tr>
<tr>
<td>- require an email address and 3 pictures to register an account</td>
<td></td>
</tr>
<tr>
<td>- click continue to practice the game</td>
<td></td>
</tr>
<tr>
<td>- 40 pictures in the library in a random order</td>
<td></td>
</tr>
<tr>
<td>Instruction page 4:</td>
<td><img src="image2" alt="Instruction Page 4 Screenshot" /></td>
</tr>
<tr>
<td>- indicate an account is created successfully</td>
<td></td>
</tr>
<tr>
<td>- introduce the operation of the game</td>
<td></td>
</tr>
<tr>
<td>- click the button to practice</td>
<td></td>
</tr>
<tr>
<td>Practice page:</td>
<td><img src="image3" alt="Practice Page Screenshot" /></td>
</tr>
<tr>
<td>- practice the game without punishment</td>
<td></td>
</tr>
<tr>
<td>- click Done to the regular login process</td>
<td></td>
</tr>
</tbody>
</table>

**Success message:**

Successfully registered! ^_^

**Go to practice:**

You will see a set of 16 pictures. Among them, there will be one of the pictures you chose. You will only see them for a short time, though. Then you must click on the square where you saw your picture. The speed of hiding images will be adjusted according to your performance.

When you click the picture, there is a sound indicating whether it is the right picture. If you pick a wrong one, you will hear a lower pitch than the sound indicating a right one. If you’re not sure where the right one is, just click any square and it will lead you to the next round. So please do not refresh or go back.

Please go to the regular login after the practice. On the regular game, you need to pick the right image for 3 times, and then you can enter the system. But if you pick the wrong images for 3 times at first, you will fail to login.

Please click **Done** to continue.
<table>
<thead>
<tr>
<th><strong>Page name</strong></th>
<th><strong>Screenshots</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction on Practice:</td>
<td><img src="image1.png" alt="Instruction Screenshot" /></td>
</tr>
<tr>
<td>- a choice to see detailed instructions</td>
<td><strong>Instruction:</strong> You will see a set of 18 pictures. In among them will be one of the pictures you choose. You will only see them for a short time, though, and then you must click on the square where you saw your picture. The speed of holding images will be adjusted according to your performance. When you click the picture, there is a sound indicating whether it is the right picture. If you pick a wrong one, you will hear a lower pitch than the sound indicating a right one. If you’re not sure where the right one is, don’t be anxious. Just click any square and it will lead you to a new round. So please do not refresh or go back. Please go to the regular game to log out after the practice. On the regular game, you need to pick the right image for 5 times, and then you can enter the system. But if you pick the wrong images for 3 times at once, you will have to log in again.</td>
</tr>
<tr>
<td>Game to login:</td>
<td><img src="image2.png" alt="Game to login Screenshot" /></td>
</tr>
<tr>
<td>- a maximum of 5 rounds to pick out the target pictures</td>
<td><strong>Game to login:</strong> If you click the right secret picture for 3 times, a reward board to indicate status in the game will be displayed.</td>
</tr>
<tr>
<td>- login to the system if a user click the right secret picture for 3 times</td>
<td></td>
</tr>
</tbody>
</table>
### Page name

<table>
<thead>
<tr>
<th>Finish page:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- logged in to the system successfully</td>
</tr>
<tr>
<td>- click the button to log out</td>
</tr>
<tr>
<td>- choice to check secret pictures</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Check secret pictures:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 3 out of 6 pictures are the secret keys</td>
</tr>
<tr>
<td>- it is expected to evaluate a natural memory decay of animal pictures, so that the choice is on the bottom.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Example of messages:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- login failure message</td>
</tr>
<tr>
<td>- end-of-test message</td>
</tr>
</tbody>
</table>

### Screenshots

**Finish page:**
- Congratulations! ^o^
- One last step: please click log out to finish
- Log out

**Check secret pictures:**
- Congratulations! ^o^
- One last step: please check secret pictures

**Example of messages:**
- Sad! T_T
- Fail to enter the system. Please login and try again.

**Goodbye! See you next time!**
You will receive an email in approximately 24 hours asking you to log in to the system again - and then again every day for a week. Now you can definitely close the window!
<table>
<thead>
<tr>
<th><strong>Page name</strong></th>
<th><strong>Screenshots</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Login page:</td>
<td><img src="image" alt="Login page screenshot" /></td>
</tr>
<tr>
<td>- enter email address to enter the game</td>
<td></td>
</tr>
<tr>
<td>- no links on this page to enter other pages like instruction pages</td>
<td></td>
</tr>
<tr>
<td>- confine a linear procedure</td>
<td></td>
</tr>
</tbody>
</table>
5. TESTING

5.1 Experiment 1

As the instructions are integrated into prototype 2.0, the prototype (FtA) is ready to be tested. To conduct this experiment, participants were first asked to give their email addresses, so that later they could receive email reminders with a short introduction and the link to FtA. The reminders lasted for a week, and participants just needed to check their emails and click a link. This experiment aims to evaluate the usability of FtA according to participants’ performance of using it for a week. In addition, a questionnaire is deployed to explore participants’ feelings and experience towards FtA, and to gather demographic information as well.

5.1.1 Participants

To conduct this experiment, more than forty people were invited face-to-face or through emails, and twenty seven participants responded to register accounts on FtA after the first email reminder. Of these, there were twenty two valid accounts that had at least one login record created in the database. These participants were invited from the University of York, and nearly all of them were young students. As half of the participants who registered a valid account did not do the questionnaire sent out with the last email reminder, it has not been possible to provide more detailed demographic information about the participants at this stage.

5.1.2 Procedures

Send email reminder for registration

Before the experiment started, more than forty people were invited to test FtA. They gave their email addresses and agreed to receive email reminders. A first email reminder was sent to these participants. The email contained the link for registration. When participants registered an account with their email address and three secret pictures, this data was recorded onto the database. Participants registered at different times, so they were separated into groups to receive subsequent emails. All of the email reminders were sent in the morning.
Clicking the first email reminder

This link led participants to walk through a linear process in the order of reading instructions, granting consent, registration, practising the game, and login. The trick was that when participants practiced the game, they might not have noticed that they need to click “Done” to finish the training section and continue to test the game in the login process.

Send email reminders for login

The second email reminder contained a link for direct login. For participants who registered their accounts with valid email addresses, they would receive login reminders every day for a week. At the last login reminder, participants were asked to finish a questionnaire about the overall experience of using FtA.

Clicking subsequent email reminders

Participants were taken directly to the login page when they clicked the link. They needed to type their email address and continue to play the game where they need to pick out their secret pictures for three rounds to enter the system. After they logged into the system, they could access the practice page and check the instructions there. However, often they just clicked the “log out” button to finish one test. For the last email, the reminder contained a link to a questionnaire. After participants finished the login process, they needed to click the second link on the email to finish the questionnaire.

Device

To ensure that FtA has a good layout and all the functions work well, it is recommended to use Google Chrome as the browser on the computer.

Grouping

A problem which occurred during the experiment was that participants started to do the experiment on different days, even though they received the first email reminder on a same day. The solution to this is to check new accounts before sending reminders each morning, and to add the new user to the mailing list and start to count the days for each user. When collecting data, accounts can be grouped by their different starting date. Finally, it is necessary to order the dates of the records in each group, so that all of the data can be analyzed together with an indicator of sequence.
5.1.3 Result

Two tables were created during the implementation phase to save account information and login records. At the end of this experiment, 27 accounts and 154 login records had been created. Among the data, there were 23 valid accounts and 108 good login records that were used for further analysis.

Rule out data

In terms of accounts, four accounts were deemed invalid for the following reasons:
- No login record created: There is an indicator to adjust the blocking time of the game, and the value of the indicator will be changed if users play the game during the login process. If this indicator is null, it indicates there is no login record created in the database. So once account was ruled out for this reason.
- Not on mailing list: one account was ruled out because the email address was not on the mailing list. It created one good login record, which may result from not receiving subsequent email reminders.
- Exception of generating pictures: two accounts were ruled out because FtA did not work well, although they almost chose the right targets all the time. In the data of the login records, two types of variables indicate which secret picture is presented (target) and which picture is selected (selected). The value of these variables are numbers from 1 to 40, which are also the name of the pictures. However, the values for these two accounts were 0, indicating a white empty picture instead of a secret picture displayed in the grid. It can be inferred that they used a browser which did not support the code of FtA.

In terms of login records, records were removed for the following reasons:
- Exception of generating pictures: login records of two accounts were removed because of unexpected errors in not generating the target secret pictures.
- Incomplete login: The game verifies a user in at least three rounds. Records were removed because they contained data for less than two rounds, indicating an incomplete login or just a login attempt.
- Out of the experiment span: As the experiment lasted for 7 days, most of the participants would not use FtA anymore after that period. For
participants who responded late for the last email reminder, they may have logged in to the system after day 7. In particular, one participant kept using the prototype until day 18, and created all of their nine records after day 8. It is better to remove these records to keep a general and objective evaluation of the overall login performance.

**Login distribution**

As shown in the Figure 5.1, the login records last to 8 days. Most of the participants stopped logging in to the system on day 7, because the reminders were sent for 7 days only. There were four records created on day 8. It may be because the participants forgot to login the day before. If we add the four participants on day 8, the average number of records created in a day is 15.4286.

![Login records](image.png)

Figure 5.1 Distribution of login records in days

**Error Analysis**

To analyze how many errors participants made, data was collected carefully to examine the overall performance of using FtA. During the game, the tolerance level for errors is two. It means that participants can still login to the system successfully if they pick out one or two wrong target pictures during the game. However, if they make three mistakes, they will be logged out of the system and requested to login again. In Figures 5.2 and 5.3, the legends of 0 to 3 indicate the number of errors made by participants during the login process. As shown in Figure 5.2, twenty one participants logged in to the system without making an error; twelve participants had made one mistake; eight participants had made two mistakes; and nine participants had made three mistakes and therefore failed to login. The 108 login records
include the second login requested after the login failure. An interesting finding is that nearly all of the participants could login to the system without making a mistake, while nearly half of the participants had failed to login to the system. The reason is not clear at this stage.

The chart on Figure 5.3 converts the number of login records with different numbers of errors into a percentage. There are 48.15% of records without an error, 28.7% with one error, 9.26% with 2 errors, and 13.83% with three errors. The result indicates most of the participants can pick out their target pictures in the game, and they can still login successfully to the system, despite having clicked on one or two wrong pictures. As for the small number of records showing login failure, this may result from one of two reasons: the game was too challenging for these participants; participants could not remember their secret pictures clearly. Another reason could be that this resulted from distraction in the surrounding environment. As the pictures disappear after a brief time, participants need to focus on seeing the pictures. So, if they are distracted or interrupted by people or things around them, they may miss a round in the game.
Time Analysis

To evaluate the performance of the game, one problem is the examining of the time spent in the login process. The data used to generate diagrams is collected from the records suggesting a successful login. Ninety three records were selected to compute the results. The average login time is 15.101 seconds, and the standard deviation is 8.77 seconds.

As shown in Figures 5.4 and 5.5, 57 records show that participants could login to the system successfully within 5 to 15 seconds. The login time of 33.33% of the records is between 5 to 10 seconds; the login time of 27.96% of the records is between 10 to 15 seconds; and the login time of 32.26% of the records is more than 15 seconds.
15 seconds. The login time of only one record is less than 5 seconds. This indicates that the fastest speed of picking out three target pictures is about 5 seconds.

![Distribution of login time](image)

Figure 5.6 Distribution of login time

Figure 5.6 shows the distribution of login time of the login records. The points of login time are more crowded under 20 seconds. In particular, the tendency line of the distribution slopes down to the right. It may indicate participants can use shorter time to finish finding their target pictures as they become much familiar with the game.

**Conclusion**

The overall performance of using FtA is examined from the aspects of errors and time used in playing the game. By examining the login records created on the database, we can see that nearly 90% of these records indicate a successful login, although some of these records indicate that the participants made one or two errors. Nearly all of the participants could complete the game to login to the system without making a mistake. This result may imply that the game is of moderate difficulty. However, nine participants experienced login failure. One implication of this result is that these participants may not have remembered their secret pictures clearly, and so they kept picking the wrong targets. Another implication is that the game is just too challenging for them, although the blocking time is self-adjusted according to each user’s performance. In addition, the difference in performance may be related to personal preference. For users who like animals, they may be more excited during the gameplay than those users who are not interested in funny animals. So they may be more focused and therefore perform better.

The general login time is about 5 to 20 seconds, which is quick for playing a game. Still, it takes up more time than typing a textual password. However, a strong password is usually long and has a certain kind of permutation which involves
different areas on the keyboard and is not easy to type quickly. Also, it takes more time if people type wrong passwords. Déjà Vu and CHC are both previous frameworks based on searchmetric models. The login time in Déjà Vu is 27 to 32 seconds, and the login time in CHC is up to 72 seconds, which includes the loading time of the application. [44] To compare the login time of FtA to those of Déjà Vu and CHC, the login time of 5 to 20 seconds is acceptable.

The analysis of this experiment is concerned with an overall examination of the performance of using FtA. It lacks analysis from an individual perspective, which could examine how many times a participant use FtA, and could then divide the individuals into groups according to their login date. In this way, the analysis of login time and errors could be more accurate regarding different groups of participants.

5.2 Experiment 2

As there were only twelve participants who finished the questionnaire in experiment 1, this experiment was conducted to gather more responses to the questionnaire. New participants were asked to test the same prototype only for just one time, and then they were asked to finish a questionnaire. The questionnaire is implemented by Qualtrics, of the university, and can be accessed through a link [52]. The questionnaire is also on the appendix.

5.2.1 Participants

Twenty one participants were invited to test FtA. These participants were mainly students in the university. So, there were thirty three participants in total who finished the questionnaire. Among these participants, twenty one participants were female and twelve participants were male. Twenty nine participants were between the ages of 19 and 29, two participants were between 30 and 39, one participant was between 50 and 59, and one participant was over 60. Thirty of the participants said that they use computers or the web several times a day, and three of the participants said that they use these once or twice a day.
5.2.2 Procedures

Devices

Some participants were invited face-on-face, so they did the following procedures on the Google Chrome web browser on a Mac computer. Other participants were invited by emails with a link to FtA and a link to the questionnaire.

Test the prototype of FtA

Participants just needed to click the link for FtA and follow the instructions. After they read the introduction and gave their consent for participation, they registered an account and practiced the game. They then clicked “Done” to quit the training section and play the game for login.

Finish the questionnaire

After the participants finished testing FtA, they needed to click a link and finish the questionnaire.

5.1.3 Result

There are three sections on the questionnaire: ease of use, fun, and demographic information.

Ease of use

Figure 5.7 Distribution of rating results
This section contains eight questions regarding using pictures as secret keys. Seven of them ask for a rating, ranging from positive opinion to negative opinion. The other question is a multiple choice question.

Figure 5.7 presents the rating result of the questions, and the score from 1 to 5 indicates a rating from positive opinion to negative opinion. Figure 5.8 merges the two positive ratings into positive and two negative ratings into negative.

1. When you first registered with the system, you had to pick three pictures. How hard did you find it to pick them?

The average rating is 2.36 with a standard deviation of 0.98. Although the number of neutral ratings is the largest, more participants tended to make positive ratings, which indicates most of the participants think it is easy or not hard to pick three pictures on registration.

2. Did you feel that it was a lot of effort to try to learn to remember your three pictures?

The average rating is 2.76 with a standard deviation of 1.35. The percentage values of positive ratings and negative ratings are both 39.39% of the total ratings. It indicates there is no strong preference in the ratings, and participants may think it was no effort or a bit effort to learn to remember the three pictures.
3. Were you at all anxious about not being able to remember your three pictures?

The average rating is 2.64 with a standard deviation of 1.34. The distribution of ratings for this problem is similar to question 2. The percentage of positive ratings is 48.48%, and of negative ratings it is 36.36%. Although the number of positive ratings is higher than that of negative ratings, it is not clear whether there is a tendency for giving a rating of not anxious about not being able to remember the three pictures.

4. When you came back to log on to the system was it hard to remember your pictures?

The average rating is 2.39 with a standard deviation of 1.10. The percentage of positive and neutral ratings are 54.55% and 24.24% respectively, so it indicates that most of the participants think it was not hard to remember secret pictures when they came back to login again.

5. The pictures were covered over after a short time. Did that make it much harder to pick your picture?

The average rating is 3.82 with a standard deviation of 1.22. The percentage of negative ratings is 69.70%, which indicates that most of the participants think it is a bit hard or very hard for them to pick out pictures under the mechanism of coverage.

6. If you picked the right pictures, you were logged on. Do you think you were successful at logging on?

The average rating is 2.55 with a standard deviation of 1.05. The percentage of negative ratings is 18.18%. This indicates that most of the participants think they were successful when they picked the right pictures.

7. As you will be aware, for security on password-based systems, it is a good idea to change your password frequently. If you were using a picture-based system like this, do you think it would be easy to change your three pictures regularly (say once per month), or hard to remember them?
The average rating is 3.85 with a standard deviation of 1.16. The percentage of negative ratings is 72.73%. It indicates most of the participants think it would be hard or very hard for them to remember the three pictures if they were to change their secret pictures once a month.

8. How do you choose your 3 secret images?

This problem explores how participants choose their secret images. Three participants chose pictures of one character; seven participants chose pictures of one kind of animal, e.g. all dog pictures; nineteen participants chose pictures of at least one different animal like two dog pictures and one cat picture; six participants chose pictures related to their life; and twenty-one participants chose pictures which tap into their emotions. The result indicates that participants tend to choose pictures of at least one different animal, and choose pictures which elicit a kind of emotion. The implication of the result is that participants choose distinct pictures to facilitate the picking in the game, and they choose emotion-related pictures to facilitate remembering.

![Figure 5.9 Distribution of choices](image)

Fun

Three questions were asked in the questionnaire to examine whether participants enjoy using FtA, and whether they would like to use this kind of authentication mechanism.

1. Did logging in to the system feel like playing a game?
2. Did you think it was fun?
3. The game was intended to be a replacement for passwords. If you had the choice, would you prefer to use the game or passwords?

The average rating for question 1 is 2.58 with a standard deviation of 1.44; for question 2 it is 2.75 with a standard deviation of 1.36; and for question 3 it is 3.67 with a standard deviation of 1.50. As shown in Figures 5.10 and 5.11, nineteen participants (57.58%) thought the login process was like playing a game. Over half of the participants (51.52%) thought it was fun. However, still over half of the participants (51.52%) would prefer to use textual passwords.

![Figure 5.10 Distribution of rating results](image1)

![Figure 5.11 Percentage distribution of rating result](image2)

**Conclusion**

The first section of the questionnaire is concerned with the usability of the prototype. An overall answer to the problems indicates that it is easy to pick three pictures as secret keys; it is not hard to remember the keys; the coverage of pictures during the game is challenging; it is likely to elicit a sense of achievement when picking out the target pictures; and it is hard for most of the participants to change their secret pictures regularly. As twenty one of the participants who answered this questionnaire had used FtA only once before they completed the questionnaire, these participants did not experience memory decay of the pictures, and so they contributed to the positive ratings seen in question 4. Thus, the result cannot tell clearly whether it is easy to use funny animal pictures as secret keys.

The second section of the questionnaire asked about users’ experience of playing the game. Most of the participants thought it was like a game more than just work, and it was fun to play, yet they still preferred textual passwords. This result seems to be
robust for an overall examination among new users and old users, but it is not clear whether new users would get bored with playing FtA over time.

6. DISCUSSION

6.1 Issues of FtA

6.1.1 Selection of pictures

Preference related effect

One reason to choose pictures of funny animals is to elicit positive emotions like pleasure and amusement, even in small amounts. So, using cute, eye-catching artwork in a game will serve as a substitute to aesthetic values. However, the attraction of these pictures may have a different effect on different people’s minds, and then might elicit different amounts of positive emotions. For people who like animals and have pets themselves, they may feel more sympathy on seeing an animal which is like their pets. As a result, they will enjoy playing the game of “find that animal”. While for people who are not interested in these pictures for some reason, they may feel less sympathy on seeing the pictures, and in turn the funny pictures may not be able to boost an adequate amount of positive emotions to deliver fun. These people are likely to think FtA is like a task rather than a game. As this project only tested the game mechanism with funny animal pictures, further experiments need to be conducted to validate this implication.

Danger of choosing pictures in order

The pictures in the library seen on registration are listed randomly without a clear classification. When users register an account, they are asked to choose three impressive pictures as keys. Figure 6.1 shows the distribution of selected keys for all of the accounts in FtA, from pilot testing through to the experiments. There may be a preference for certain pictures, but the preference may result from the limited number of pictures in the library (40 pictures). However, it seems that the three keys are selected in a sequence: the first picture is selected from the top part, the second one is selected from the middle part, and the last one is selected from bottom part. Still, it is not clear whether this effect is due to the limit resources available in library.
6.1.2 Remembering secret pictures

Practice is not enough

It is suggested for users to choose impressive pictures as their keys on registration. The term “impressive” indicates that users pay more attention to a picture, and have a stronger memory of it. Although users go through a training session to help them remember their secret pictures, they may still forget one of the three secret pictures after a few days. One suggestion for remembering the pictures is to let users spend more time choosing their secret pictures. Instructions could be used, such as: “Describe the picture to yourself with two special characters”. As well as a semantic coding of a picture, serious observation of the details can help build a memory trace in the long-term memory with rich, complex stimuli, and in turns users can have a better memory of a picture [23].

6.1.3 Relying heavily on vision

FtA exercises the abilities of quick reactions and attention. The authentication in the game cannot be completed without using your eyes. As the pictures are covered after a short time, if users miss the chance to find the target picture, they can just guess. This game is not suitable for people who have lost their vision. Other people
who have eye problems, such as color blindness, may be able to use this kind of
game-based authentication mechanism, as long as they can distinguish the details
of the pictures. During the experiment, participants said that the coverage of the
pictures was not challenging, but the distraction of other pictures was. In FtA,
picking out a target picture before the coverage is a challenge in the game, and the
challenge is based on the limited ability of visual attention. However, it can also be a
barrier to people who are lacking in or are not interested in visual attention.

6.2 Vulnerability Evaluation

The vulnerability of graphical authentication can be classified into four
categories: guessability, observability, recordability, and memorability. These
four aspects of vulnerability can be used to evaluate the security level of a
given graphical authentication mechanism. [53]

<table>
<thead>
<tr>
<th>Vulnerability</th>
<th>Description</th>
<th>Attack Type</th>
<th>Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guessability</td>
<td>The level to be guessed and cracked by attackers.</td>
<td>Brute Force Attack</td>
<td>Offline</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dictionary Attack</td>
<td>On &amp; Offline</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shoulder Surfing</td>
<td>Human Observer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spyware</td>
<td>Technology</td>
</tr>
<tr>
<td>Observability</td>
<td>The level to be observed by attackers.</td>
<td>Social Engineering</td>
<td>Deception</td>
</tr>
<tr>
<td>Recordability</td>
<td>The level to be recorded.</td>
<td>Theft</td>
<td>Unsecured Record</td>
</tr>
<tr>
<td>Memorability</td>
<td>The level to be memorised.</td>
<td>Forgetting</td>
<td>Coping Techniques</td>
</tr>
</tbody>
</table>

Guessability

FtA uses pictures as keys to authenticate users. The secret pictures cannot be
cracked by a dictionary attack, because the picture library is a unique resource and
there are no cracked picture libraries to be used as a reference. However, the system
can be threatened by a brute force attack of repeated attempts.
In FtA, users can login to the system by picking out the target pictures three times, and they can make a maximum of two mistakes. In addition, the secret picture in each round is generated randomly from three pictures. For example, there is a 1 in 3 possibility of the same secret picture being generated in a subsequent round. In terms of picking out three target pictures in the first three rounds, the number of possible guessing combinations is $16 \times 16 \times 16 = 4096$, which indicates that it would require a maximum of 4,095 attempts to crack the secret pictures. However, the security level will be decreased dramatically if an intruder occasionally guesses a secret picture and keeps it in mind immediately, and if the same secret picture is generated on the two subsequent rounds, it only takes $16 \times 3 \times 3 = 144$ chances to login to the system successfully. So, the system will be at a great risk if one of the secret pictures is guessed or observed by intruders.

**Observability**

In FtA, the pictures are displayed on the computer screen, and users click a target picture to proceed. The secret pictures cannot be cracked by spyware, but they may be observed by surrounding people. The coverage of pictures not only serves as a challenge in the game but also help protect the user from shoulder attack.

**Recordability**

It is hard to record secret pictures compared with textual passwords. However, users can keep keywords to describe a picture. People may make different keywords and comments for the same picture. In addition, there are pictures of one kind of animal and of one individual animal. Even if an intruder gets a simple cue, they may not be able to enter the system. However, a simple cue will also decrease the difficulty of guessing because the current library is too small.

**Memorability**

Funny animal pictures are filled with rich stimuli. They can be remembered not only by semantic keywords such as “a smelling cat”, but also by the details of the pictures. So, funny animal pictures are supposed to facilitate remembering. According to the result of the questionnaire, the majority of participants did not think it was hard to remember the secret pictures when they logged in again. However, better instruction should be provided to help strengthen the memory of the pictures apart from in the training part of the game.
7. CONCLUSION

This project has attempted to explore the novel idea of a game-based authentication mechanism. It was hoped that such an authentication mechanism would not only provide security, but would also improve the usability of a password. A game idea based on graphical authentication was determined through iterate designs, and then this idea was implemented to a prototype named ‘Find that Animal’. The prototype was improved and integrated with instructions for an experiment in a second version. To test the usability of FtA, two experiments were conducted.

To summarise the first experiment, the results indicate that participants were able to login to the system without making mistakes in the game. Almost half of the participants had a login failure, which indicates that the game may be a bit challenging for these participants, or that these participants could not remember their secret pictures. Another implication is that these people could have been distracted from the game, which relies heavily on visual attention, and could then have missed the chance to pick out their target pictures, causing them to fail to login. The time spent in the game is about 5 to 20 seconds. It may take longer time than a textual password, but the time is shorter than in two other graphical authentication mechanisms. The login time in Déjà Vu is 27 to 32 seconds, and the login time in CHC is up to 72 seconds which includes the loading time of the application. In addition, there is a tendency towards using shorter time in the game to login. However, an experiment with a longer time span is needed. In addition, one danger in choosing secret pictures in FtA is that participants may choose the pictures from the top to the bottom of the screen, which decreases the difficulty of guessing.

To summarise the second experiment, the results indicate that most of the participants think FtA is fun and feel it is like a game, yet they still prefer to use textual passwords. It can be inferred that the custom of using textual passwords cannot be changed easily, and the development of game-based authentication mechanisms is promising because people can have fun and enjoy a sense of achievement in the game.

Upon conducting a vulnerability evaluation of the approach, it was concluded that it is not very secure against brute force attack. So, this authentication mechanism is not suitable for any system which requires a high level of security. However, this approach can be implemented as a supplement to other authentication mechanisms, to improve the overall security level. Then FtA will work like a PIN, but it is more dynamic. The implementation of the game mechanism can prevent
attack from shoulder surfing, as the pictures on the screen will be covered over after a short time, and the page will be refreshed soon after a click. It is hard to record secret pictures, yet people may record keywords. The memorability of funny animal pictures is supposed to be easy, because such pictures contain rich stimuli for remembering. The results of the questionnaire indicate people have a good memory of such pictures, however an experiment with a longer time span would be required to validate this result.

8. FURTHER WORK

Improvement can be made in the implementation of levels. Apart from the auto-adjusted mechanism of the blocking time in the game, different levels with different challenges could satisfy users’ different needs.

In addition, the picture library seen on registration needs an expansion so that people can have a better opportunity to choose pictures in which they are interested, and can have a better experience in gameplay.
REFERENCE


S. Wiedenbeck, J. Waters, L. Sobrado, and J. Birget, 2006. Design and


http://www.cs.northwestern.edu/~hunicke/pubs/MDA.pdf


[51] http://www.w3schools.com/

