2016-05-18

Graphical One-Time Password (GOTPass): A usability evaluation

Alsaiari, H

10.1080/19393555.2016.1179374
Information Security Journal: A Global Perspective
Taylor and Francis

Graphical One-Time Password (GOTPass): A Usability Evaluation

Alsaiari H, Papadaki M, Dowland P, Furnell S

Centre for Security Communication and Network Research, School of Computing Electronics and Mathematics, Plymouth University, Plymouth, United Kingdom
email: info@cscan.org

Abstract

Complying with a security policy often requires users to create long and complex passwords to protect their accounts. However, remembering such passwords appears difficult for many and may lead to insecure practices, such as choosing weak passwords or writing them down. In addition, they are vulnerable to various types of attacks, such as shoulder surfing, replay and keylogger attacks (Gupta et al., 2012). One-Time Passwords (OTPs) aim to overcome such problems (Gupta et al., 2012); however, most implemented OTP techniques require special hardware, which not only adds cost, but there are also issues regarding its availability (Brostoff, Inglesant, & Sasse, 2010). In contrast, the use of graphical passwords is an alternative authentication mechanism which is designed to aid memorability and ease of use, often forming part of a multi-factor authentication process. This paper is a complementary to the earlier work that introduced and evaluated the security of the new hybrid user-authentication approach: Graphical One-Time Password (GOTPass) (Alsaiari et al., 2015). The scheme aims to combine the usability of recognition-based and draw-based graphical passwords with the security of OTP. The paper presents the results of an empirical user study that investigates the usability features of the proposed approach, as well as pre-test and post-test questionnaires. The experiment was conducted during three separate sessions, which took place over five weeks, to measure the efficiency, effectiveness, memorability and user satisfaction of the new scheme. The results showed that users were able to easily create and enter their credentials as well as remember them over time. Participants carried out a total of 1,302 login attempts with a 93% success rate and an average login time of 24.5 seconds.

Keywords: authentication, knowledge-based authentication, graphical passwords, One-Time Password, usable security

1. Introduction

In general, the task of recognising a displayed item has been demonstrated to be easier for people rather than relying on their memory to recall the same information without any assistance (Nielsen, 1994). Furthermore, a classic cognitive science experiment showed that humans have a strong memory ability for images (Standing, Conezio, & Haber, 1970). Thus, recognition-based techniques are an interesting branch of graphical passwords, which involve identifying a set of user-selected images among other decoy images. This technique has been proposed as a usable alternative to textual passwords, since it includes many useful features, such as ease of memorisation, simple use as well as providing a reasonable security level (Khot, Kumaraguru, & Srinathan, 2012). With respect to security, the password space is an important factor for a robust authentication scheme. Generally, most recognition-based schemes suffer from a small password space, whereas many recall-based schemes can offer a much larger password space. Therefore, the proposed scheme employs both techniques to gain the best out of each. An Android unlock pattern (a recall-based (draw-based) technique) is implemented as a point-of-entry defence for the main recognition-based (choice-based) technique.

One of the authentication mechanisms to withstand many of the traditional textual password security issues is the One-Time Password (OTP). The nature of this technique makes it appropriate to secure various financial services and online payments, since OTP generates a password that is valid for a single use which then expires. Thus, this paper proposes an authentication scheme that makes use of a graphical password to generate an OTP. It is

envisaged that the proposed mechanism could form a lower cost and more readily available alternative to token reader devices that are often used in online banking.

The rest of this paper is organised as follows: Section 2 briefly introduces relevant existing schemes. In Section 3, the GOTPass approach is described. Section 4 provides a detailed usability evaluation, as well as an overview of the security evaluation. Section 5 discusses the outcomes of the scheme's evaluation followed by the conclusions in Section 6.

2. Related Work

(Komanduri & Hutchings, 2008) implemented a picture password system with the ability to produce a memorable, high-entropy password. The proposed system consists of 80 unrepeated pictures, and each one is labelled with a character. Each participant is assigned with a unique arrangement of eight items known as the 'home grid', which they need to recognise to fulfil future authentication requirements. Pictures are always placed in a fixed location within the home grid with the same correspondent keyboard key. In this system, a dual input ability is enabled by using either the keyboard or an on-screen mouse cursor. Furthermore, another initiative was launched to accept an unordered input, thus allowing the selection of the correct images in any order. According to the study, a successful authentication system could benefit from this unordered recall.

(Gao et al., 2009) (Wang et al., 2010) innovated a solution based on a challenge-response protocol to protect graphical passwords against spyware attacks by utilising a Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA). The new authentication scheme is a combination of graphical password and a textual CAPTCHA that is assigned and embedded into each displayed image. To register, users need to choose and

remember a number of pass-images as their password. In order to authenticate, users are required to pass two steps. First is the image recognition step, where they need to look for their pass-images among other decoy images. The second step involves solving and typing in the assigned CAPTCHA string that appears below each pass-image in a certain way. The improved technique of this scheme uses a predefined random length as an alternative to the usual uniform length. As such, the user predetermines the position and the number of characters. Consequently, users need to select and memorise the letter positions (pass-positions) of each pass-image (e.g. the letters in the 1st, 3rd and 7th positions).

(De Angeli et al., 2002) (De Angeli et al., 2003) presented an innovative concept for user authentication called Visual Identification Protocol (VIP), which is based on the idea of replacing conventional PIN numbers with pictures. An authentication attempt is successful when the user correctly selects the images that are part of their portfolio among other decoys within the display panel. There are three variations of the VIP scheme, one of which is the advanced scheme (VIP3), which assigns a portfolio of eight pictures to each user. At every login attempt, a 4x4 challenge set is presented to the user, containing four random portfolio pictures together with an additional 12 distractors. To authenticate, users have to identify their pre-set images among the 16 images shown on the interface in any sequence.

(Van Oorschot and Wan, 2009) came up with a new scheme called TwoStep. The new scheme is a hybrid user-authentication scheme which utilises traditional text passwords and recognition-based graphical passwords. In the first step, users will still use a text-based password as normal, but the second step involves entering a graphical password. Users need to register a number of images as their graphical password components which are set over a particular number of rounds. Once this has been done, an index number is assigned to each image. The login screen will display, at random, the images along with their index numbers.

A selection panel is located at the lower part of the screen, which contains all of the index numbers in ascending order. To authenticate, the user needs to identify the image and select the corresponding index number from the selection panel. TwoStep has the advantage of the user being able to enter the graphical password part by clicking a mouse, which reduces the possibility of keylogging attacks.

In Where You See is What You Enter (WYSWYE) (a scheme proposed by (Khot et al., 2012)), two variations of the proposed approach were implemented: Horizontal Reduce (HR) and Dual Reduce (DR). Although they are different in terms of the challenge grid size and the process of identifying and mapping the image pattern, the underlying strategy stays the same.

In the registration stage of the DR scheme, users are presented with a set of 28 images and required to create a password containing four images. During the login time, the scheme generates two side-by-side grids; the challenge grid contains random images, four of which correspond to the password. The user is expected to interact with the second grid only, the response grid, which is smaller in size; it is initially empty and is used for input entry purposes. In order to map between the different size grids, the user must reduce the bigger challenge grid to the size of the response grid. This is done by a mental elimination of the rows and columns that do not contain any of the password images from the challenge grid. Login is achieved by locating the password image positions inside the reduced challenge grid and by subsequently using the response grid to map them accurately.

(Ku et al., 2012) (Ku et al., 2013) proposed a solution to generate a graphical one time password (GOTP) for financial services using smartphones. The password creation is based on selecting an image portfolio that consists of four rounds that form a story – to act as a recall assistant. Each authentication round displays images on a 4x9 grid frame in the correct order.

The respective alphanumeric OTP code is shown at the top-left corner of the screen, and the user needs to memorise this for the next round. The final (fifth) round is the password input step, which contains a random layout display of 12 buttons to allow the user to enter the memorised four OTP texts that match the image portfolio. The study showed that the average registration time was quite fast, with positive results that evaluated the recall interference, authentication time and recall convenience.

However, the GOTP approach still requires the user to memorise an alphanumeric code obtained by identifying the pass-images over several rounds and then entering the code in the final round. That, in turn, may require memory recall from the user, resulting in usability issues. In addition, GOTP is designed for smartphone platform that can be used as an out-of-band channel for authentication, which is carried out away from the browser. In other words, there is a need for an additional device (smartphone) to be present in order to use the GOTP scheme; however, this is not always an issue for many users nowadays. Furthermore, the length of the OTP code generated by GOTP is short compared to other similar schemes which provide twice as long OTP codes (e.g. Picture Password and Gao's CAPTCHA). Therefore, the demand for an enhanced authentication mechanism that utilises the advantages of such schemes (e.g. one-time password and the use of separate means for data entry) and overcomes their limitations (e.g. the need for extra devices, burdening memory with codes to remember, short codes and static pass-images) has emerged.

3. The GOTPass Scheme

Having considered the contributions of the prior works, this section proceeds to propose the basis of an alternative approach that seeks to address the perceived shortcomings. As described in (Alsaiari et al., 2015), the proposed scheme is a hybrid multi-level authentication mechanism

called **G**raphical **O**ne **T**ime **Pass**word (GOTPass). The overall objectives of the proposed scheme are presented next, followed by details of the operational approach.

3.1. Objectives

The objective of this scheme is to enhance the usability features of the existing graphical authentication system by developing a new multi-graphical password technique that fulfils most of the usability requirements. The main usability characteristics that the GOTPass authentication system aims to satisfy can be highlighted as follows.

The first requirement is the ability to create a new password using a simple process and a minimal amount of steps. Second, the password should be easy to remember, so a user is not overwhelmed by a raft of complex secrets that they have to memorise. Third, it should be a simple to use scheme that is reliable (an unreliable system may result in denial of access). Fourth, it should be efficient to use, and the registration and login time should be acceptably short. Fifth, there should be nothing to carry, which means that a user should not rely on auxiliary devices (e.g. tokens) to perform the authentication task, excluding devices that users usually carry around at all times, such as mobile phones. Finally, it should be easy to recover, allowing users to regain the ability to login in case the authentication credentials are forgotten.

The key technical advantages of the proposed scheme considered to be:

- Combination of multiple authentication mechanisms (graphical password and OTP).
- Combination of multiple graphical password categories (recall-based [draw] and recognition-based [choice]).
- System-assigned themes with user-chosen images.
- Various GOTPass input formats (code locations).

One of the significant features of an image-based authentication technique is the ease of recall, which is something that a conventional text-based password lacks. Thus, this has motivated us to investigate and develop an enhanced graphical authentication mechanism. However, most recognition-based graphical password schemes are vulnerable to observation attacks (e.g. shoulder surfing), due to their very nature of being visible to surrounding people. Therefore, we employed a user-friendly graphical technique (unlock pattern) that acts as a front-line defender before the recognition-based technique. This is in line with the results of an earlier field study carried out over 21 days which confirmed that users were in favour of the pattern mechanism despite the repeated errors they made (Von Zezschwitz et al., 2013). According to (Chiang and Chiasson, 2013), the Android screen unlock technique is the most well-known deployed graphical password. Finally, the system's security is strengthened by the implementation of the OTP technique. Table 1 summarises the rationale behind the selection of these various authentication techniques.

	Authentication technique	**Rationale of selection**
1	**Pattern unlock**	Protect the main image-based scheme User-friendly and familiar
2	**Image recognition**	Easy to remember Easy to use
3	**OTP input format**	Provide robust security

Table 1: Rationale behind the selection of various authentication techniques

3.2. Approach

GOTPass scheme combines graphical and one-time passwords. In addition, various graphical password methods have been merged to form a new mix of recall- and recognition-based techniques. The final component of GOTPass involves the determination of input formats, or, in other words, the location of the associated codes. More precisely, the method will be established by solving the lock pattern (draw-based), followed by identifying pass-images

(image recognition) and the last step will be to enter the corresponding OTP code according to the pre-chosen format (knowledge-based).

The process flow for the enrolment and authentication phases is summarised in Table 2, which defines the requirements and procedures for each phase as well as showing the authentication classifications of each part.

General process flow	**Registration phase**	**Authentication phase**
Secret knowledge (username)	Select a unique username	Enter the correct username
Pattern unlock Graphical password (recall-based, draw-based)	- 4x4 pattern grid will be displayed - The user needs to draw a pattern in any preferred shape	Unlock the pattern grid by redrawing the pre-chosen pattern
Image recognition Graphical password (recognition-based, choice-based)	- The system will assign four random themes for the user - A panel of images from each of the assigned themes will be presented and the user will make his/her own selection	The system displays a 4x4 panel of images containing two random pass-images out of the four previously chosen pass-images, plus 14 other decoy images The user needs to identify the two pass-images
One-Time Password Formation of the final password entry	- Since the edge side of each row and column of the panel will be assigned four random digits, the user can choose from two available security level options: basic or advanced. Each level has two different GOTPass input format combinations and the system will randomly assign one to the user	Enter the associated GOTPass code with each image in the same previously chosen format and in the correct order

Table 2. Process flow for the enrolment and authentication phases

3.3. Enrolment

The registration stage involves three main phases. First, the user needs to choose a unique username and draw any shape on a 4x4 unlock pattern. Second, the system will automatically assign four random themes for each user, one after another. The user needs to select one pass-image from each of the given themes (a total of four altogether). Finally, the position of the pass-images in the grid will be used to indicate a code that needs to be entered using the keypad/keyboard, which is referred to as the GOTPass input format. These codes are located on the top or left-hand axis of each pass-image. There are two security level options for the user to choose from: basic or advanced. At the basic security level, the numeric codes for both pass-images are taken from the same axis, whereas the numeric codes in the advanced level are taken from a different axis for each pass-image. The system assigned input format is clearly presented to the user with an illustration example (e.g. top axis for the 1st pass-image + left-hand axis for the 2nd pass-image).

3.4. Authentication

The system will prompt the registered user for their username and display an on-screen pattern lock (Figure 1), which requires the user to redraw the predefined unlock pattern shape by connecting nodes to re-form the correct pattern shape.

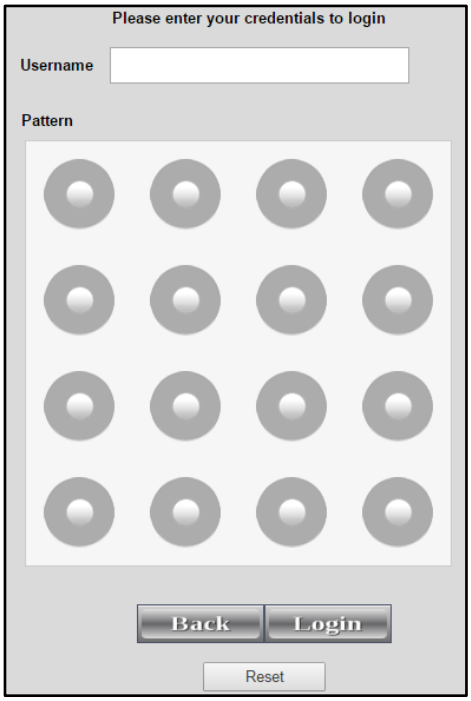

Figure 1: GOTPass unlock pattern step

If the preceding step is correct, the system will display a fresh (4x4) image panel, as illustrated in Figure 2, containing two random pass-images out of the four previously chosen pass-images, six distractor images that are associated with the pass-images (three for each) and another eight random decoy images. The system generates new OTP codes and fills the panel edges (axis) of each row and column (only the locations that are occupied by the correct pass-images will contain the correct GOTPass codes). To complete the authentication process, the user must first identify the password images among others in the panel (this is done mentally, there is no need to touch/click on the images). From the grid axis, the user needs to locate and enter the codes associated with each pass-image (these should be entered in the correct format, as previously assigned and shown in the registration phase). It is necessary to select the pass-images and, thereafter, the associated codes in the correct order depending on which pass-image appears first. Once the system ensures that all of the information that has been provided is correct, then the user is successfully authenticated and granted access.

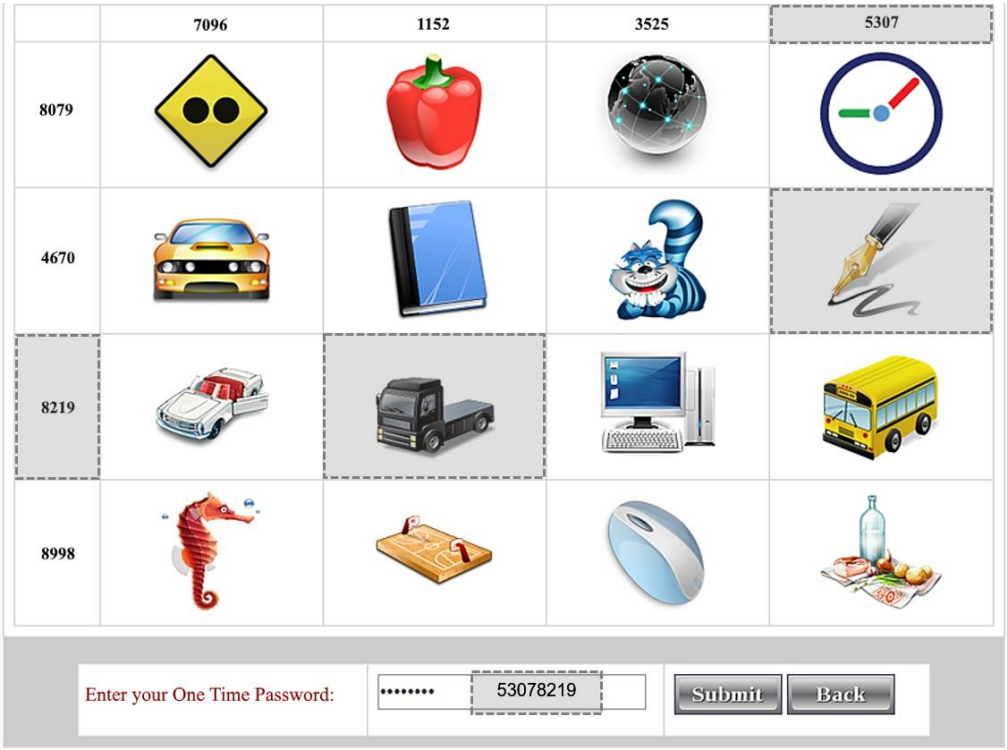

Figure 2: GOTPass image recognition and OTP code entry
Assuming security level option 3 is in use (top axis for the first pass-image + left axis for the second pass-image)

4. Evaluation

The study conducted by (Biddle, Chiasson, & van Oorschot, 2011) stated that the consistency of the published research within the domain of graphical authentication is almost absent, which complicates the task of reproducing results or comparing schemes. Many graphical password system proposals have an inadequate evaluation of either security or usability, or even both. The lack of an accepted usability standard in this area of research is a result of the missing coordination work between researchers, which led to the use of different evaluation criteria for nearly every system proposal. Furthermore, (Bonneau et al., 2012) realised that the original publications on such schemes have included optimistic and incomplete ratings. Therefore, standard evaluation methods and measurements are required to carry out a reasonable comparison against other works.

A proper framework is required to evaluate the design of a successful authentication mechanism against several aspects of security and usability (De Angeli et al., 2005). Hence, a collection of evaluation criteria and guidelines has been carefully identified by exploring the characteristics and methods of the existing graphical authentication schemes alongside the review of the available evaluation studies. However, it should be noted that fulfilling all the requirements of security and usability in a single authentication scheme is unlikely to be achievable (Schaub et al., 2013).

To prepare an appropriate evaluation plan, a review of studies carried out by similar graphical password techniques was conducted. As Table 3 illustrates, almost all schemes carried out in-lab studies. Most schemes were performed over several sessions with various time intervals. The maximum number of sessions used was three and the minimum number was one. With regard to the number of trials, two schemes allowed 10 authentication attempts. The number of participants ranged between 10 and 61. Essential evaluation elements, such as effectiveness, efficiency, memorability and user satisfaction, were the components of most of the conducted studies. In addition, at the end of the table, a summary of the GOTPass scheme study is included to enable an easy basis for comparison.

Scheme	Type of study	Sessions	Trials	Participants	Evaluation elements
Komanduri Picture Passwords (Komanduri et al., 2008)	In-lab and any location	- Day 1 in-lab - Day 2 any location - Day 9 in-lab	Eight complete correct inputs	- 23 participants - Only 15 participants received picture-based passwords	Effectiveness, efficiency and memorability
TwoStep (van Oorschot et al., 2009)	No user study	Future work: lab/field studies	–	–	–
WYSWYE Dual-Reduce	Controlled lab	One login session	Three login attempts	- 24 participants. - None of them knew about GP	Accuracy, efficiency, learnability and user satisfaction

(DR) (Khot et al., 2012)					
VIP (De Angeli et al., 2002)	Controlled lab	Two login sessions: first day and after one week	10 authentication attempts – with three incorrect attempts	61 participants	Effectiveness, efficiency and user satisfaction
GOTP (Ku et al., 2012)	In-lab	–	–	10–20 participants with prior knowledge of use	Password creation time, login time, recall convenience and recall disturbance
Gao CAPTCHA (Wang et al., 2010)	In-lab	Three login sessions: day one, one week later and one month later	- Test 1 (day 1): 10 times, - Test 2 (one week) - Test 3 (one month): three times	36 participants unfamiliar with the scheme	Login success %, login time and memorability
GOTPass	In-lab	Three login sessions: day one, one week later and one month later	Allowed: maximum10 login attempts for each session. Required: only 5 correct logins	81 participants	Effectiveness, efficiency, user satisfaction and memorability

Table 3: Summary of the graphical password technique studies

4.1. GOTPass Usability

A successful authentication system should maintain a balance between usability and security. System usability is an essential design aspect that should not be compromised for security (and vice versa). The GOTPass proposal contains some interesting usability design features (Table 4), such as the use of image themes that prompt users to remember password images. Although the system prohibits users from using their own images, to protect against a guessing attack by a familiar person and help reduce the impact of users tendency to choose predictable images, they are allowed to choose preferred images from a specified theme, which adds flexibility to the system as well as freedom of choice for the user. One of the GOTPass goals is to have a

reasonable level of memorability so users manage to remember their pass-images easily. However, there is no use of mnemonics to assist users in remembering their passwords, since the proposed scheme uses multiple authentication mechanisms which makes applying such a feature on each mechanism both difficult and pointless.

	Usability features				
	System-assigned Themes	**User-provided images**	**User-selected images**	**Memorability**	**Mnemonic**
GOTPass	✓	×	✓	✓	×

Table 4: GOTPass usability features

4.1.1. Experiment Design and Implementation

The GOTPass prototype was developed as a web-based application using Microsoft Visual Studio 2013 – C# and SQL Server 2012 as the Database Management System. The prototype application was hosted on a laptop with a 15.6" screen display set at a resolution of 1366x768 pixels and running Windows 8.1.

A user study was conducted that involved three separate trial sessions on the first day of the study, one week later and after one month. A within-subjects design method was used in which the same users participated in all experimental tasks – that is, repeated measures are taken from the same people. Participants performed two main assignments: firstly to enrol and authenticate for several times over specific time intervals and secondly to act as observers to try and capture the experimenter's login password using various attacking techniques. This study is a longitudinal testing method, since several observations of the same subjects were conducted over a period of time.

Experiments to evaluate the usability and security of the GOTPass approach were conducted in a controlled lab environment, as all users were required to be physically present and use the same computer to perform the study tasks. For study purposes, the implemented scheme

generated some significant activity logs in such a way that it stores timestamps, login status (successful, failed) as well as details of the duration of each session. In addition, results of the responses to the pre-test and post-test questionnaires were also collected. Only the research investigator and the participant were allowed in the lab, to avoid any possible disruption and observe any usability or security issues, as well as record the participant's comments. Nevertheless, attention was paid to the session duration, in which we tried to remain focused on the experiment and discouraged any side conversations during the trials, unless participants chose to talk.

Given the longitudinal nature of the study, and the necessity for those involved to remain available for each stage of the work, the participants were sourced from the local staff/student community at the authors' university, and recruited via several methods: including word-of-mouth, student portals, emails and posters. Participation did not require any specific level of computing ability. Each participant received reasonable compensation for their participation, payable upon the completion of the study at the end of the third session. As for the session duration, the allocated time for each session never exceeded 30 minutes.

The experiment was conducted over five weeks and involved 81 participants (63 male, 18 female) who attended all three separate sessions. Most participants were university staff and students, with a mix of educational levels ranging from undergraduate and postgraduate. Most participants were aged between 18 and 39 years. Fifty percent of participants reported an intermediate level of computer experience, yet 17% indicated a basic level. Almost all participants indicated that they knew about at least one type of graphical technique. Draw-based graphical passwords were most familiar to the users, followed by recognition-based passwords, whereas only a few respondents had prior knowledge of the click-based technique.

4.1.2. User Study Procedure

Below is the series of tasks the users were required to perform at each session.

A. Initialisation session – Day one

The first session started with a brief introductory overview of the procedure, participants' rights as well as an explanation about the system functionalities and the process of enrolment and authentication. An instruction manual 'guide booklet' and video demo that describes the registration and login sequential steps were made available as training materials.

After gaining the required understanding of the system and how it works, participants started the registration phase, where they created a new account.

Once the users were registered, they filled out a short online pre-test questionnaire on demographic and authentication experience. This acted as a separator role between phases to distract the user's attention away from the registration process to aid a better evaluation of memorability during the next phase. This is similar to the Mental Rotation Tasks (MRTs) procedure, which aims to clear the participants' working memory.

The final task of the first session was the login phase, where participants were required to login (maximum 10 total attempts) under the following conditions:

- Total of five correct authentication attempts > successfully completed this session.
- Total of five incorrect attempts > receive the guide booklet or play the video demo, then try again.

Participants were instructed to avoid clicking on the pass-images, instead they were encouraged to mentally locate the images and map them to the right axis of the OTP code.

B. Follow-up session (short-term memorability experiment) – One week later

After a week of non-use, participants returned to the lab where they were asked to repeat the login task.

C. Final session (long-term memorability experiment) – One month later

The third and final session took place one month after the first session. The first task was again to login using the created account with the same rules and conditions as the first and second trials.

Finally, each participant received an online post-test questionnaire to assess their impression of the GOTPass system, as well as find out their opinion on it.

4.1.3. Usability Study Results

As defined by ISO 9241-11 (International Organization for Standardization., 1998), effectiveness, efficiency and satisfaction are the main components of usability in a particular context. However, there are no absolute measures of usability (Bangor et al., 2008). Nevertheless, major usability features from ISO and previous studies were extracted to build a usability evaluation criteria for the new graphical password system. This paper reports the quantitative results for all usability components except user satisfaction, which reports qualitative results from the surveys regarding the user perceptions.

i. Efficiency

Usability elements	Measurements	Assessment type	Assessment method
Average entry time for registration/ authentication	$Av\ (R) = \frac{\text{Sum (successful\_registration\_times)}}{\text{number\_of\_successful\_registrations}}$ $Av\ (L) = \frac{\text{Sum (successful\_login\_times)}}{\text{number\_of\_successful\_logins}}$	Objective/ quantitative	Experiment/ user trial

Table 5: Efficiency evaluation elements

Table 5 describes the details of the measurements used to calculate the efficiency of the proposed scheme. As anticipated, creating a GOTPass account took relatively long time since registering for GOTPass includes typing a username, drawing a pattern, clicking the 'Register Pattern' button, initial thinking time (image viewing), selecting four pass-images, choosing the security level and, finally, clicking the 'Submit' button. As shown in Table 6, the average registration time was 134 seconds. It is worth mentioning that participants were totally new to the system and, while they created their accounts, spent quite a lot of time talking and asking questions about the prototype, trying to start discussions about several aspects, such as the potential advantages and disadvantages of the system and the way it was implemented. Although the registration time was relatively high, it was considered generally acceptable for most participants, as indicated in the post-test questionnaire result, where 80% of the users stated that they managed to complete the required tasks quickly. In contrast, only one participant disagreed with this statement.

	Total attempts	Total time	Average	SD	Minimum	Maximum
Registration	81	10,833	134	36.5	59	254

Table 6: Registration entry time details (in seconds)

In the analysis of the time it took to enter the correct submission, the average was 24.5 seconds, as presented in Table 7. The long input time was also expected in the login phase, since the login task involves a number of keystroke and mouse activities. In addition, the time taken to mentally locate the correct pass-images and their associated codes is also considered to be a

significant factor that increased the login time. There was a slight variation in the average login time between trials: 23.6, 25.5 and 24.3 seconds respectively.

	Total attempts	Success	Total time	Average	SD	Minimum	Maximum
Login	1,302	1,215	29,754	24.5	11	8	83

Table 7: Entry time details for successful authentication (in seconds)

ii. Effectiveness

Usability elements	Measurements	Assessment type	Assessment method
Login success rate	$SR(L) = \frac{\text{number\_of\_successful\_logins}}{\text{number\_of\_total\_logins}}$	Objective/ quantitative	Experiment/ user trial

Table 8: Effectiveness evaluation elements

The details of the measurements used to calculate the effectiveness of the proposed scheme can be seen in Table 8. The study looked at the proportion of all successful login attempts across all trials to calculate the success rate of the proposed system. In total, data from 1,302 login attempts carried out by all participants were analysed. Table 9 provides details of the success and failure rates for the authentication phase over the three trial sessions. The results show a relatively high success rate, as over 93% of the attempts were successful. Although the first trial was preceded by MRTs, to distract the users after the registration task and free up their working memory, this did not have any clear impact on the success rate of the first trial in particular. In the final session (Trial 3), there seems to be some associations of the GOTPass in the participants' memory, as the number of incorrect inputs was lower than in Trial 2.

	Total attempts	Successful		Failed	
Trial 1	429	405	94.4%	24	5.6%
Trial 2	438	405	92.5%	33	7.5%
Trial 3	435	405	93.1%	30	6.9%
Total	1,302	1,215	93.3%	87	6.7%

Table 9: Login success and failure rates

Interestingly, the study showed that none of the users were completely unable to login within the given number of attempts. Approximately 40% of the participants managed to complete their login tasks without error. Moreover, since many systems limit the number of consecutive incorrect attempts a user is allowed to make, we introduced this measure to determine the highest number of repeated failed attempts. The results show that only one user failed to login, with three consecutive incorrect login attempts, and seven others failed for two logins. In addition, only one participant was responsible for the maximum non-consecutive failed attempts by a user (five attempts), as shown in Figure 3 below.

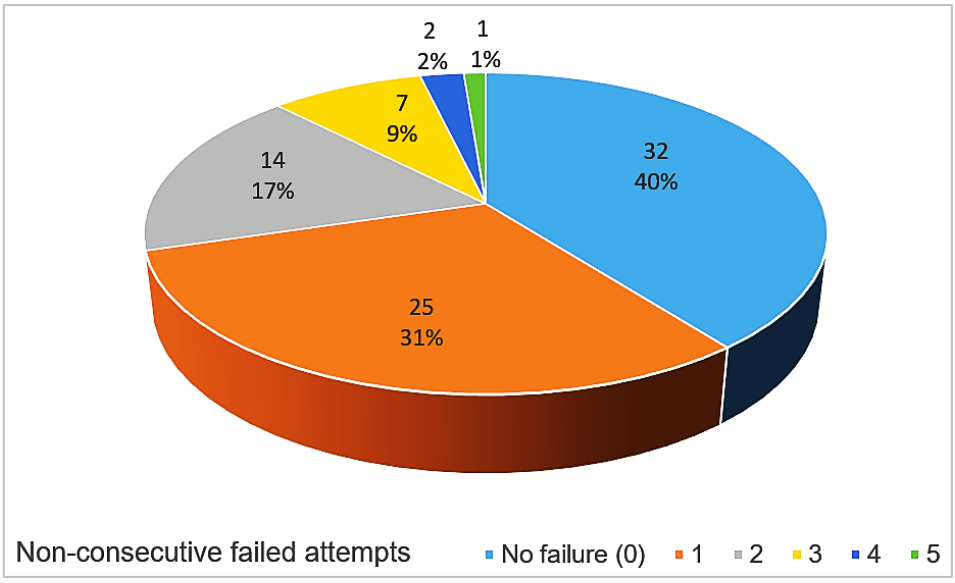

Figure 3: Number of users and their non-consecutive failed attempts

One of the observations from the trials highlighted that almost all failures occurred within the recognition part of the authentication process, more precisely the wrong codes or inputting codes in the wrong order, since the majority of the participants claimed that they were sure they recognised their pass-images correctly but might have entered them in the incorrect order or made a typographical mistake.

iii. Memorability

Usability elements	Measurements	Assessment type	Assessment method
Memorability over time intervals Short (one week), Extended (one month)	Matched at first attempt Matched within three login attempts	Objective/ quantitative	Experiment/ user trial

Table 10: Memorability evaluation elements

Table 10 shows the details of the measurements used to calculate the memorability of the proposed scheme. Participants carried out a memorability experiment twice. The first took place after one week of non-use (Trial 2) and the second was one month later (Trial 3). The results showed that all users managed to login successfully to their GOTPass accounts, but the number of attempts to do so varied. There was no lockout event since all consecutive incorrect attempts were three or fewer.

	Trial 2						**Trial 3**					
Attempt sequence	**1**st	**2**nd	**3**rd	**4**th	**5**th	**6**th	**1**st	**2**nd	**3**rd	**4**th	**5**th	**6**th
Failure frequency	12	6	6	4	3	2	15	3	5	4	2	1
Total	**33**						**30**					

Table 11: Details of the frequency of the failed attempts based on trials and attempts

Table 11 illustrates the number of failed login attempts in each sequence. It can be inferred from the table that 85% of the participants in Trial 2 managed to login successfully on their first attempt. In addition, the number of failed attempts seems to reduce over time. One month later, in Trial 3, when participants tried to re-enter their GOTPass secrets, only 19% were unable to correctly login at the first attempt. However, during all trials almost all users logged in successfully within three attempts, which shows an encouraging outcome from a password recall perspective.

iv. User Satisfaction

Usability elements	Measurements	Assessment type	Assessment method
Overall satisfaction (simplicity, ease of use, understandability and perception of using GOTPass)	Satisfied Neutral Unsatisfied (7-point Likert scale/ multiple choice)	Subjective/ qualitative	Questionnaire/ attitude scale

Table 12: User satisfaction evaluation elements

The details of the measurements used to analyse the level of user satisfaction of the proposed scheme is shown in Table 12. User satisfaction was measured through a post-test questionnaire, which was given to the users at the end of their final study session. The aim was to discover the users' feelings towards the perceived aspects of usability and security of the proposed system. Most measurements were carried out using a 7-point Likert scale, ranging from 1 (strongly agree) to 7 (strongly disagree), whereas some others used multiple-choice measurements. All 81 participants of the user study took part in the survey. The results indicate that 86% of the respondents agreed that learning how to use the system and how to create a GOTPass account was simple, with the remaining 14% showing an average response. Almost 91% of the participants stated that this authentication method would become easier and quicker to use with practice. The vast majority of the participants (98.7%) stated that they would be confident using the GOTPass system. Ninety-four per cent of the participants thought that the GOTPass system could be used for sensitive web authentication. The overall level of user satisfaction with the GOTPass system was very high, as 98% were in support of the idea. Note that the results of all responses were mostly in the positive half of the scale, which, in turn, reflects positive outcomes towards a prospective solution.

4.2. GOTPass Security

Of particular interest to our work is the security aspect, which was evaluated in detail in a parallel work (Alsaiari et al., 2015). In brief, the key points from the preliminary results are also presented in this paper. Two types of security evaluation were conducted, the first, 'theoretical', was based on assessment criteria and the second, 'empirical', was where several attacks were simulated and tested.

The security experiment involved 81 participants who were divided into three groups based on the assigned security attack experiment. Simulations of three security attacks were prepared (guessing, intersection and shoulder-surfing attacks) to evaluate the proposed system's capability to resist such attacks. Participants were asked to act as attackers to try and steal a victim's credentials. Overall, the analysis of the security evaluation showed that GOTPass had a high resistance against common graphical password attacks. The results showed that only 3.3% of the 690 login attempts succeeded in compromising the system.

5. Discussion

Compared with other graphical password techniques that are similar in nature, such as (Khot et al., 2012) (Komanduri et al., 2008) (Gao et al., 2009), GOTPass has both advantages and disadvantages. At first glance, many users thought it might be too complex; however, learning and practising the system created an opposite impression, as the majority found it easy to use and adoptable.

The long account creation time is a disadvantage of the system, but, at the same time, it is worth mentioning that GOTPass is a multi-level authentication approach which employs several graphical password techniques into a single robust mechanism. That, in turn, might justify the extended time taken to create user accounts. In order to register, users need to complete

multiple steps: username selection, unlock pattern drawing, multi-round pass-images selection, and, finally, choose the security level along with the input format. In addition, these factors have an obvious impact on the complexity of the registration process. However, although it seems complex and takes time, the user study shows that, overall, users were satisfied – there were no complaints about the duration of the registration process or the level of difficulty. Furthermore, the GOTPass scheme provides strong resistance against various common security attacks, which is one of the primary objectives of this system.

Although the combination of several security methods may yield a higher level of security, it may also affect the usability of the system. However, that is not the case with the GOTPass scheme, as it aims to keep a reasonable balance between security and usability and avoid any trade-off. According to the results of the user study, there is no evidence of a negative impact on usability as a result of combining multiple security methods. Additionally, reporting a high success rate even after a period of time, as well as the users' positive perception regarding the simplicity of the system, prove that multi-security levels do not hamper the usability of GOTPass.

Focusing more on one of the chained steps and neglecting the others by choosing weak passwords should not be a major issue, as the success of breaking one of the authentication steps will not compromise the entire credentials. In addition, the employment of the implicit feedback technique plays an important role in hiding which step is actually incorrect. In this way, it is difficult for an attacker to find out whether the strong or the weak step is wrong. In other words, GOTPass works as a package where each part or feature complements the other.

Comparing the login time of GOTPass to other graphical schemes (see Figure 4) shows that the login time still appears to be sensible. As mentioned earlier, a significant reason that influences the performance time of an authentication scheme is the involvement of multiple steps, which also justifies the longer time taken to register and login to GOTPass. However,

GOTPass is still comparable to other two-step approaches, and is even superior within its category (three-step).

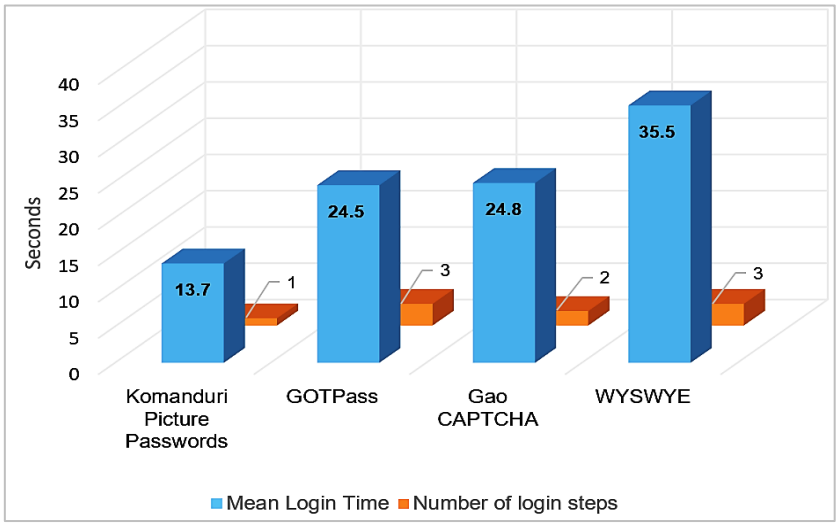

Figure 4: Comparison of the mean login time and number of steps to login

In terms of comparing GOTPass with its closest scheme, GOTP, a direct comparison is not straightforward, given that the evaluation data for GOTP are limited to post-test survey responses and not experimental data (Ku et al., 2012). Nonetheless, a brief comparison between the two schemes is presented next. The data of our survey had to be adjusted from a 7-point Likert scale to a 5-point Likert scale to enable a direct comparison. In order to gain comparable results, the response values of the relevant questions were converted by using the following method (IBM Support, 2015):

1. L_i = Multiply the response value by its frequency (e.g. 7-point Likert scale × number of selected times).
2. S = Sum, the total of all points ($L_7 + \ldots + L_1$).
3. P = Divide S by the number of participants (S ÷ 81) {the mean value in a 7-point Likert scale}.
4. Q = Divide P by 7 (P ÷ 7) {the value in the range between 0 and 1}.
5. R = Multiply Q by the new Likert point number (Q × 5) {the mean value in a 5-point Likert scale}, the value of R represents the original result but using a 5-point Likert scale.

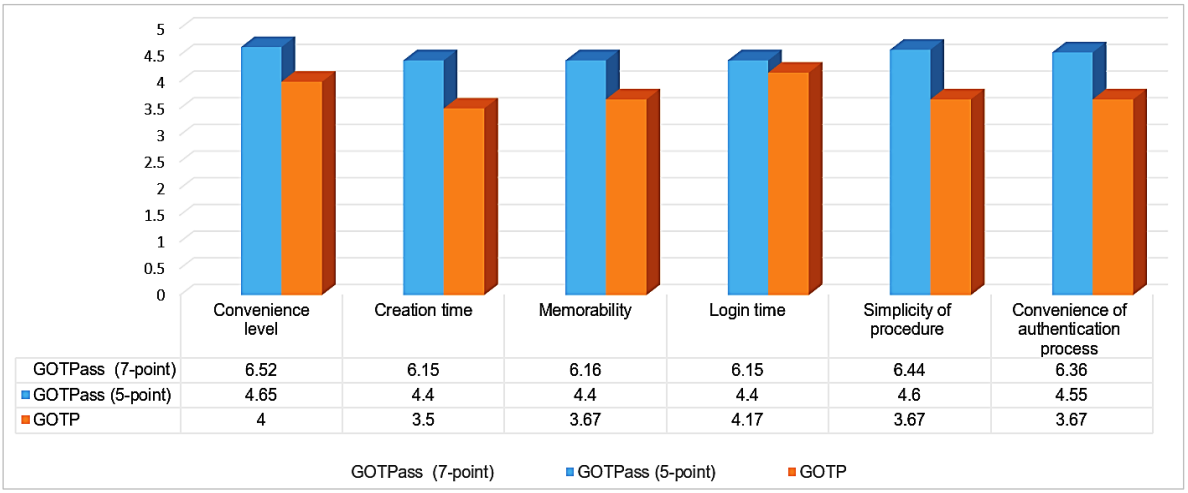

Figure 5: Comparison summary of GOTP and GOTPass

Figure 5 highlights the differences based on the available evaluation data of the GOTP scheme. It demonstrates that GOTPass has a major advantage of having a larger number of participants, which increases the accuracy and reliability of the result. Although GOTP scored highly regarding the level of memorability, GOTPass showed even better results, which satisfies one of the main requirements of any prospective alternative authentication system. In relation to that, ease of use is another important feature, and GOTPass achieved a higher result than that of GOTP. However, across all comparison parameters GOTPass has performed very well, with over four out of five in all aspects.

In addition, the GOTP scheme requires the user to memorise four alphanumeric codes obtained by identifying the pass-images over four rounds. That, in turn, would require memory recall from the user, posing possible usability issues. In contrast, the GOTPass scheme does not involve the memorisation of codes, since they are visible on a single screen. In addition, GOTP is designed for smartphone platform that can be used as an out-of-band channel authentication, which is usually carried out away from the browser, whereas GOTPass utilises an in-session authentication system using the existing browser. In other words, there is no need for additional devices, such as a token or mobile phone, to use the GOTPass scheme. Regarding the length of the OTP code, GOTP submits a four-character-long code while GOTPass requires an eight-

character code. Themes and images used in GOTP are static and unchangeable, but in GOTPass they are dynamic and shuffling. The letters and numbers in the top corner of each GOTP image are barely readable on a mobile phone screen (Figure 6), which can be considered to be a major usability drawback of the system.

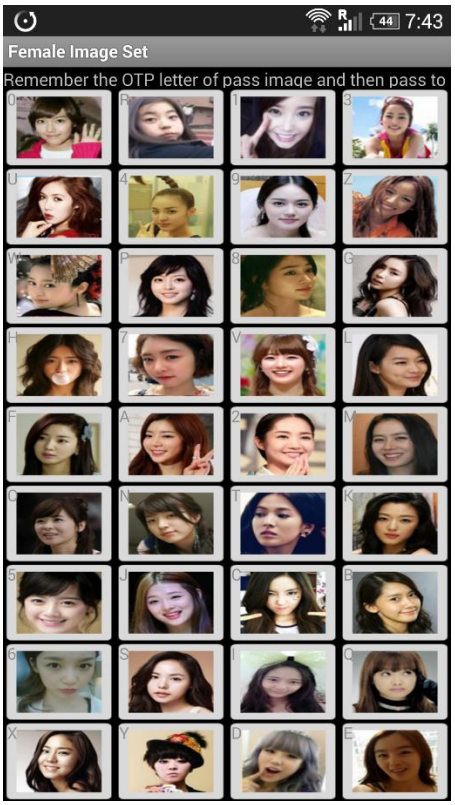

Figure 6: A screenshot of the GOTP login screen

6. Conclusions and Future Research

This paper has presented a usable mechanism to help authenticate users by using combined graphical password techniques along with an OTP. The main contribution is the introduction of draw-based and recognition-based graphical methods with the employment of an OTP to resist many of the common security threats without sacrificing the ease of use. Initially, the results of the experiments indicated that the scheme has an acceptable level of efficiency and effectiveness as well as a high level of user satisfaction. Moreover, the study showed that GOTPass has the potential to succeed and contribute towards the adoption of graphical password technologies. Further research is recommended that should concentrate on

conducting a field study and improve registration and login times. Enlarging the sample of participants and running the user study for an extended period of time are suggested to allow more conclusive analysis of the data. It is also suggested to investigate the compatibility and effectiveness of the current design on different platforms, especially handheld devices. In terms of security, the resilience of the proposed scheme has been investigated in parallel with this study. In fact, the results of the earlier security experiment, involving three different attack simulations against GOTPass, were encouraging and complementary to this work.

7. References

Alsaiari, H., Papadaki, M., Dowland, P. & Furnell, S. (2015) 'Secure Graphical One Time Password (GOTPass): An Empirical Study'. *Information Security Journal: A Global Perspective*, 24 (4-6). pp 207-220.

Bangor, A., Kortum, P. T. & Miller, J. T. (2008) 'An Empirical Evaluation of the System Usability Scale'. *Intl. Journal of Human–Computer Interaction*, 24 (6). pp 574-594.

Biddle, R., Chiasson, S. & Van Oorschot, P. (2011) 'Graphical passwords: Learning from the first twelve years'. *ACM Computing Surveys*, 44 (4).

Bonneau, J., Herley, C., Van Oorschot, P. C. & Stajano, F. (2012) 'The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes', *Security and Privacy (SP), 2012 IEEE Symposium on*. IEEE, pp. 553-567.

Brostoff, S., Inglesant, P. & Sasse, M. A. (2010) 'Evaluating the Usability and Security of a Graphical One-Time PIN System', *Proceedings of the 24th BCS Interaction Specialist Group Conference*. British Computer Society, pp. 88-97.

Chiang, H.-Y. & Chiasson, S. (2013) 'Improving User Authentication on Mobile Devices: A Touchscreen Graphical Password', *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*. ACM, pp. 251-260.

De Angeli, A., Coutts, M., Coventry, L., Johnson, G. I., Cameron, D. & Fischer, M. H. (2002) 'VIP: A Visual Approach to User Authentication', *AVI '02 Proceedings of the Working Conference on Advanced Visual Interfaces*. New York, USA ACM, pp. 316-323.

De Angeli, A., Coventry, L., Johnson, G. & Coutts, M. (2003) 'Usability and user authentication: Pictorial passwords vs. pin'. in McCabe, P.T. (Ed.), *Contemporary Ergonomics*. Taylor & Francis, London, pp 253-258.

De Angeli, A., Coventry, L., Johnson, G. & Renaud, K. (2005) 'Is a Picture Really Worth a Thousand Words? Exploring the Feasibility of Graphical Authentication Systems'. *International Journal of Human-Computer Studies*, 63 (1–2). pp 128-152.

Gao, H., Liu, X., Wang, S. & Dai, R. (2009) 'A New Graphical Password Scheme Against Spyware by Using CAPTCHA', *Proc. Symposium On Usable Privacy and Security (SOUPS)*. pp. 15-17.

Gupta, S., Sahni, S., Sabbu, P., Varma, S. & Gangashetty, S. V. (2012) 'Passblot: A Highly Scalable Graphical one Time Password System'. *International Journal of Network Security & Its Applications (IJNSA)*, 4 (2).

IBM Support *Transforming different Likert scales to a common scale*. http://www-01.ibm.com/support/docview.wss?uid=swg21482329 (Accessed: June 2015).

International Organization for Standardization, ISO 9241-11 (1998). *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs): Part 11: Guidance on Usability.*

Khot, R. A., Kumaraguru, P. & Srinathan, K. (2012) 'WYSWYE: shoulder surfing defense for recognition based graphical passwords', *Proceedings of the 24th Australian Computer-Human Interaction Conference*. ACM, pp. 285-294.

Komanduri, S. & Hutchings, D. R. (2008) 'Order and Entropy in Picture Passwords', *Graphics Interface Conference 2008*. Ontario, Canada Canadian Information Processing Society, pp. 115-122.

Ku, Y., Choi, O., Kim, K., Shon, T., Hong, M., Yeh, H. & Kim, J.-H. (2012) 'Extended OTP Mechanism Based on Graphical Password Method'. In *Future Information Technology, Application, and Service*, vol. 1, pp 203-212. Springer Netherlands. doi:10.1007/978-94-007-4516-2_20

Ku, Y., Choi, O., Kim, K., Shon, T., Hong, M., Yeh, H. & Kim, J.-H. (2013) 'Two-factor Authentication System Based on Extended OTP Mechanism'. *International Journal of Computer Mathematics*, 90(12), pp 2515-2529, Taylor & Francis. doi:10.1080/00207160.2012.748901.

Nielsen, J. (1994) 'Usability Heuristics'. in Nielsen, J. (Ed.), *Usability Engineering.* AP PROFESSIONAL, London, pp 129-130.

Schaub, F., Walch, M., Könings, B. & Weber, M. (2013) 'Exploring the Design Space of Graphical Passwords on Smartphones', *Proceedings of the Ninth Symposium on Usable Privacy and Security*. ACM, pp. 11.

Standing, L., Conezio, J. & Haber, R. N. (1970) 'Perception and memory for pictures: Single-trial learning of 2500 visual stimuli'. *Psychonomic Science*, 19 (2). pp 73-74.

Van Oorschot, P. C. & Wan, T. (2009) 'TwoStep: An Authentication Method Combining Text and Graphical Passwords'. in Babin, G., Kropf, P. and Weiss, M. (eds.) *E-Technologies: Innovation in an Open World.* Springer Berlin Heidelberg, Ottawa, Canada, Vol. 26, pp. 233–239.

Von Zezschwitz, E., Dunphy, P. & De Luca, A. (2013) 'Patterns in the Wild: A Field Study of the Usability of Pattern and PIN-based Authentication on Mobile Devices', *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*. ACM, pp. 261-270.

Wang, L., Chang, X., Ren, Z., Gao, H., Liu, X. & Aickelin, U. (2010) 'Against Spyware Using CAPTCHA in Graphical Password Scheme', in *24th IEEE International Conference on Advanced Information Networking and Applications (AINA)*, 2010, pp. 760–767.