Armchair Authentication

Karen Renaud Joseph Maguire

Department of Computing Science
University of Glasgow
Glasgow, G12 8QQ
{karen, joseph}@dcs.gla.ac.uk

ABSTRACT

Alphanumeric authentication, by means of a secret, is not only a powerful mechanism, in theory, but prevails over all its competitors in practice. However, it is clearly inadequate in a world where increasing numbers of systems and services require people to authenticate in a shared space, while being actively observed. This new reality places pressure on a password mechanism never intended for use in such a context. Asterisks may obfuscate alphanumeric characters on entry but popular systems, e.g. Apple iPhone and Nintendo Wii, regularly require users to use an on-screen keyboard for character input. This may not be a real concern within the context of secluded space but inadvertly reveals a secret within shared space. Such a secret has an economic cost in terms of replacement, recall and revenue, all of which affect the financial return of the offending systems and services.

In this paper, we present and evaluate a graphical authentication mechanism, Tetrad, which appears to have the potential to address these specific concerns.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Graphical user interfaces*; K.6.5 [**Computing Milieux**]: Security and Protection—*Authentication*

General Terms

Design, Economics, Experimentation, Human Factors, Security

Keywords

context, shared space, secluded space, graphical authentication mechanism

1. INTRODUCTION

When access to an electronic resource is restricted, legitimate users need to prove that they are entitled to access it. They do this by proffering an identity, and the identity is verified by means of an authentication step. This ensures accountability of individuals using restricted resources. Authentication is most commonly established by means of a shared secret password.

© The Author 2009.
Published by the British Computer Society

The sheer number of secrets, passwords and PINs individuals are expected to remember is placing them under undue pressure and they are responding by behaving insecurely: writing down or sharing 'secrets' or using personal details [11].

However, when authentication *is* necessary, arguably, individuals seem to prefer the convenience and anonymity of passwords [14] and often resist more time-consuming or privacy-violating authentication mechanisms [15]. Furthermore, they are endlessly innovative in finding ways of easing the memory burden imposed by endless password requests [13, 20]. Conventional approaches not only include writing passwords down but reusing the same one with several systems or generating an inconsequential "don't care" password for use with systems not having negative side-effects, if compromised [11]. However, individuals tend to put some effort into choosing their passwords for systems which hold sensitive or financial details. They will also be careful about entering these passwords in such a way that they will not be observed.

Password entry is routinely obfuscated so that casual observers cannot easily see what is being typed, which reassures end-users. Unfortunately, Tari *et al.* [23] found that when users were required to type long and obscure passwords their attempts were more easily observed by shoulder surfers than when they were typing in an easy and familiar word. Unfortunately, the passwords we choose to protect our bank account and tax records are likely to have exactly these characteristics, so our efforts to be "secure" actually backfire. Even so, most people are fairly confident that observers cannot guess their password with any degree of accuracy [26], even though this confidence is probably misplaced [23].

However, with the advent of devices such as the Nintendo Wii and televisions that can be used to browse the Web, it becomes clear that passwords are completely unsuitable. Many of these devices come without physical keyboards, so the user has to enter the password using an on-screen variant. This means that they have to navigate to the applicable key and activate it. The slowness of the process makes observation not merely possible, but inevitable.

Consider the situation where Mike suggests purchasing a movie from a service such as iTunes for the family to watch. Everyone is sitting comfortably in the living room with him, when he accesses the service, which then asks him to enter his password, so that the movie can be charged to his credit card. Mike is in a quandary. If he enters the password he is actually letting the rest of his family in on his secret. If he chooses not to enter his password he is effectively telling them that he does not trust them. Mike has the choice of either being uncomfortable with the disclosure of his password or having offended various family members. It's likely that Mike will avoid the situation altogether and refrain from suggesting purchasing the movie. Thus, the service loses custom simply because no one considered the context of authentication and its impact on Mike.

We propose a better way of managing authentication in a *shared space*. Section 2 discusses different authentication contexts, the economics of the problem and introduces Tetrad, an observation-resistant authentication mechanism. Section 3 outlines the design of Tetrad and Section 4 explains how it was evaluated. Section 5 gives the results of our evaluation, which are discussed in Section 6. Lastly, our concluding remarks and thoughts on future work are presented in Section 7.

2. CONTEXT

Authentication is most often achieved by means of a shared secret. By definition any disclosed secret is no longer secret and thus can no longer serve as an authenticator. Hence authentication context — the physical environment within which the user is authenticating — must be taken into consideration when designing the authentication mechanism. Two different user contexts need to be considered:

Shared Space : We are not alone, and we are aware that individuals or devices could be potential threats. Tan and Czerwinski [21] point out that large displays have a serious impact on privacy and argue that a solution to this should be sought.

Secluded Space : We are alone and do not need to be concerned that others are observing our actions. Most authentication mechanisms in use today implicitly assume this context by requiring the user to provide their entire secret. The only concession to possible observation is obfuscation of the entered text, and even that has been abandoned on the popular iPhones, which briefly display the entered text so as to enhance usability for the user.

These contexts are not necessarily mutually exclusive and users will behave differently in each, no matter how unaware they are of security issues. Customers in the United Kingdom have to enter a PIN when using their credit cards to purchase goods in a card-present transaction. After a number of fraud cases banks are now advising customers routinely to shield their PIN entry. The fact that the banks had officially to issue such advice confirms that many people simply do not understand the security threats they are vulnerable to [26].

However, even the least privacy conscious user will not want private or sensitive information displayed on a large screen for everyone to see in a shared space. We therefore have to accommodate the many people who want to be able to share movies, music and photos with others but do not want to give away their authentication secrets in the process. What is needed is a way for people to *prove* knowledge of a secret without revealing the secret to an active observer. It is hard to envisage doing this with an on-screen keyboard so we should facilitate it using a mouse or Wiimote, as the case may be.

The following section will discuss the economics of this shared space problem, Section 2.2 briefly discusses work done by other researchers in this area and Section 2.3 outlines our proposed solution to this problem.

2.1 Economic Cost

The implementation of an authentication mechanism when a system is being developed is relatively cost free if the ubiquitous password is used. However, there are a number of related costs, carried by the organisation, which are not always considered:

1. *Password Replacement* — Replacing a password can be an expensive business if it is done securely [5]. This cost is directly proportional to the number of registered users. It makes sense for companies to find new ways of enhancing the memorability and security of their authentication keys so as to minimise this cost.

2. *Compromising other passwords* — Individuals often use the same password on a number of systems [11]. If an attacker observes one password, leaked through poor implementation choices, he or she will be able to compromise others. A system or service can not risk being branded as insecure. Such connotations will lead to avoidance by potential customers who fear other services or systems, e.g. financial, may be at risk.

3. *Losing Revenue from purchases* — People who do not feel comfortable authenticating will simply not purchase content from the service, e.g. buy a movie. Futhermore, if Mike were able securely to authenticate in shared space and successfully purchase a movie, it would be a great advertisement. Individuals initially viewed as potential attackers could become potential customers.

4. *Losing Revenue from failed purchases* — Individuals who forget their password will be unable to purchase content. This is of specific significance to services such as iTunes, who are likely to have customers with varied purchasing habits. While some customers make frequent purchases, others while buy weekly, monthly or annually in the case of birthdays and holidays. Infrequently used passwords are the ones most easily forgotten. Not only will the service incur a replacement password cost, they also continue to lose revenue until the individual is able, once again, to authenticate.

Therefore, although alphanumerical authentication may have a comparatively low-cost implementation, this is intertwined with often overlooked overhead costs, as outlined above. The reality is that many businesses simply view this overall expense as the *cost of doing business*. In fairness, this is not an altogether unwise view from the perspective of authentication as a necessary evil.

Alphanumerical authentication requires little to no introduction or even training. The knowledge to interact with a keyboard has existed for well over a century and the concept of secret-words has existed even longer. The notion of shifting customers from a physical store to an online one, coupled with its own nuances and training, can be considered problematic enough without introducing a new authentication mechanism.

However, an alternative perspective is to view unconventional authentication mechanisms as an *investment*. Although implementation and associated overhead costs may be an initial expense there is great reward in the long-term. A mechanism which utilises memorable elements as secrets combats the expense of replacement and decay, i.e. (1) and (4) respectively. Furthermore, a suitably constructed interaction approach protects the authentication secret and eases user concerns, i.e. (2) and (3) respectively.

2.2 Related Work

Researchers have proposed a number of different ways of alleviating these problems. Tan *et al.* [22] propose a spy-resistant onscreen keyboard specifically designed for kiosks. Unfortunately, their users were somewhat uncomfortable using the keyboard. Hoanca and Mock [12] propose the use of eye tracking systems to determine the direction of the user's gaze. Whilst this is an innovative solution, its need for eye tracking hardware is likely severely to limit its applicability. Roth *et al.* [24] propose the use of a game that the user plays in order to demonstrate knowledge of their secret password/PIN without actually divulging it. Users liked the increased security of the game but did not like the extra effort involved to authenticate, typical of users who make cost-benefit analyses and tend to reject things that take too much effort and/or time [19].

A mechanism which makes use of a pointer, rather than the keyboard, is the graphical authentication mechanism. These mechanisms rely on the user to remember a set of secret images, and rely on the fact that these will be remembered better than an alphanumeric password. Various researchers have worked on alternative authentication mechanisms which are more easily remembered than passwords [8, 3, 16, 10, 25]. Unfortunately all the mechanisms require users to click directly on the secret images and therefore these are unsuitable for any shared space authentication.

There are a number of ways to alleviate the problems related to shoulder surfing of such image-based passwords. The simplest is to display only a randomly chosen subset of available targets at each authentication attempt. This means that multiple authentication attempts must be observed before the observer is able to gain knowledge of all the secret pictures.

One could also use different targets each time. For example, the Dynahand system [18] generates new PINs each time the user authenticates, which means that a casual observer has less chance of gaining access to the user's account later because what is being tested (the handwriting) is relatively obscure and less easily cracked than a straightforward set of pictures.

The mechanism we're advocating in this paper is called *Tetrad*, a *minimum disclosure* searchmetric mechanism. Using this mechanism, the user proves knowledge of the target pictures without identifying them directly.

Limited disclosure searchmetric mechanisms foil shoulder surfing and key-logging software, since they rely on the use of arrow keys or a mouse to manipulate sets of pictures. Most limited disclosure searchmetric mechanisms have some redundancy so that the observer is not able to deduce the key from casual observation but has either to observe a number of authentications or carry out an error-prone deduction of the key based on a few observations. The v.Crypt system from Bharosa[1], illustrated in Figure 1, requires the user to use arrow keys to line up a shape on the bottom row with an alphanumeric key on the top row or to rotate a dial to line up letters in the same way as a combination lock is operated. This is done for as many letters and numbers as there are in the key.

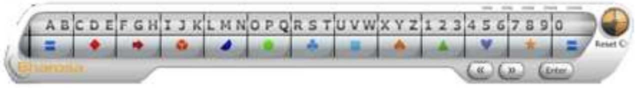
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z 1 2 3 4 5 6 7 8 9 0
Reset
Bharosa
Enter

Figure 1: Limited Disclosure Searchmetric Authentication

Another example of this kind of mechanism is the convex hull click scheme proposed by Weidenbeck *et al.* [27]. The user is assigned a number of pass-icons, which are displayed on the screen along with a number of other icons. He or she mentally constructs a convex hull using the pass-icons as vertices and then clicks inside the hull. They reported positive results but the mental effort required does seem significant.

2.3 Proposed solution

We propose *Tetrad*, a minimum disclosure searchmetric graphical authentication mechanism. A set of images are randomly positioned in grid format. Contained within the set are the user's target images, i.e. those images which constitute the image-based password.

A successful authentication attempt requires the user to re-position columns and rows of images within the grid. The user does not move or select *individual* images; he or she re-positions subsets of the images, i.e. rows and columns. The goal is to align the target images either horizontally, vertically or diagonally. Because the user is moving rows or columns at a time, it is hard for an observer to see exactly which pictures are the focus of the movement. The introduction of an element of redundancy provides the obfuscation which protects the user.

3. TETRAD

3.1 Web Prototype

We built a prototype of Tetrad for execution through a web browser using Javascript to accomplish interaction. Buttons or 'arrow-keys' were positioned at each column and row edge, i.e. to the right of a row would be a 'right arrow-key'. The user would click the arrow-key which would, in turn, execute a Javascript to re-position images. Each click represents one movement in that direction, e.g. click a right arrow-key and all images within the respective row moves one position right within the grid with the rightmost image wrapping around to the left. The appearance of an early version can be seen in Figure 3.

Informal testing revealed mixed reactions to the mechanism. We expected this to some extent: experimental authentication mechanisms often evoke connotations of unusable and cryptic methods in the mind of a user.

Unfortunately, our early prototype only reinforced this perception.

There were almost as many 'arrow-keys' as there were images, resulting in a cluttered and confusing mess. The interaction was not intuitive, although a 'right arrow-key' may logically communicate the concept of re-positioning images to the right, it failed visually to communicate this interaction. The re-positioning of images was not animated: they simply appeared in their new grid position, at the blink of an eye. Individuals would frequently click 'arrow-keys' to extract meaning, citing they only did so because they were the only objects on-screen which were not generic images.

[1] http://www.bharosa.com

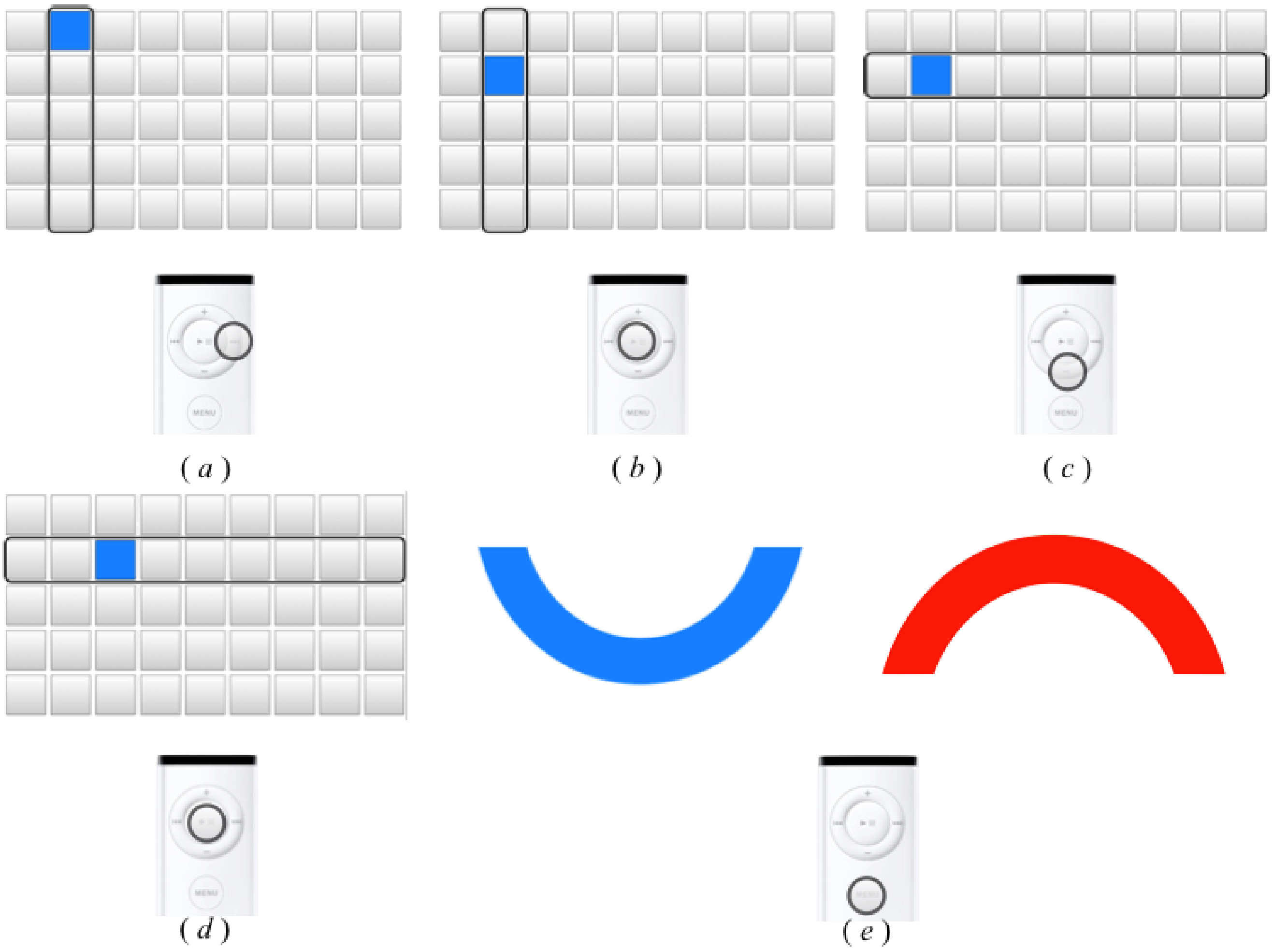

Figure 2: Interaction example

Indeed, many individuals clicked their images directly, expecting a response from the system. This not only undermined the main purpose of Tetrad, i.e. resilience to casual observation, but highlighted the interaction flaws.

The interaction initially seemed simple and could scale to a range of *shared space* devices. However, our initial implementation had 28 'arrow-keys'. This vast number of buttons is not only difficult to navigate using the remote-control of a large-scale display but consumes precious space on small-screen touch-based devices. Furthermore, we cannot have 28 separate voice commands or haptic responses for use in other implementations.

We later experimented with 14 buttons which would allow for circular movement in one direction only, deeming this acceptable since users often simply clicked the same 'arrow-key' to achieve movement, than both. However, 14 buttons was still excessive, and did not resolve other concerns.

Lastly, the generic images themselves also failed adequately to facilitate lightweight recognition [17].

3.2 Shared Space Prototype

In developing our *shared space* prototype we needed to start with a fresh perspective and to create a somewhat generic approach which could not only scale between shared space devices, such as televisions and mobile phones, but also address the concerns of our earlier web prototype.

Furthermore, we needed to investigate and select a suitable image-type for Tetrad that would facilitate lightweight recognition.

3.2.1 Design

We developed our shared space prototype using Objective-C for use with Apple's OS X. Apple's operating system is utilised across their entire range of devices, in one variant or another, i.e. iMac, MacBook, Apple TV, iPod touch and iPhone. We felt this a wise investment as it gave us scope to trial future prototypes across numerous devices.

Three main concerns to be addressed in our shared space prototype were: (1) visual communication of image movement, (2) exposure of target images and (3) interaction required to re-position images.

In addressing the first concern we looked at how others had dealt with visually communicating re-positioning of content. Minimis-

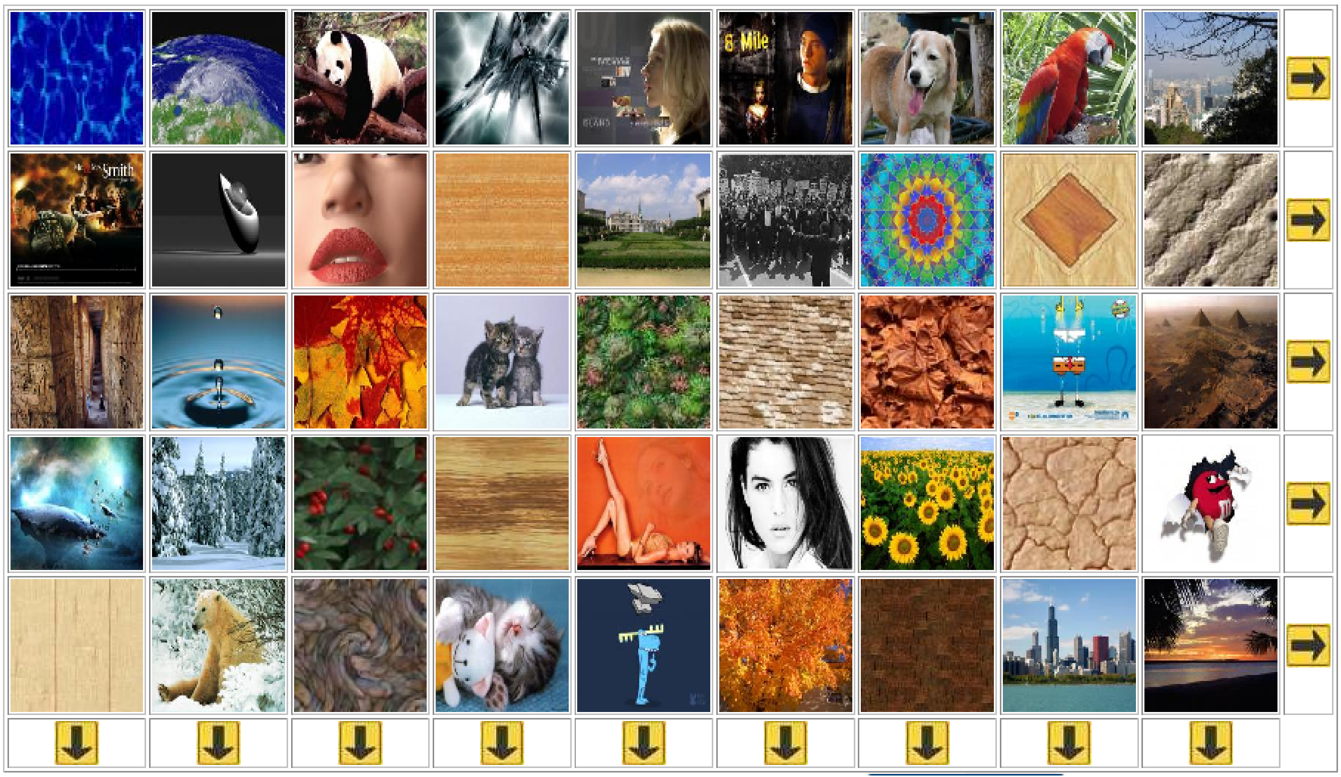

Figure 3: Web Prototype

ing a window is such an example. Historically, windows would simply disappear when minimised and appear in another location, e.g. the task bar. This could be confusing to new users as it's unclear where the window disappeared to, and how to retrieve it. Apple's solution to this concern was to animate the window, shrinking it from its current location to its new location, i.e. the 'Genie effect'.

Animation, often seen as a frill, serves the purpose in this case of visually communicating to the user the location of the now minimised window. A user consequently knows exactly where the window now resides.

We made extensive use of Apple's Core Animation framework within our new prototype. We animated images moving, rather than disappearing and reappearing. The intention behind this decision is to reduce the number of 'exploratory' button pushes by the user, often used to discover how the images moved within the grid.

Our next concern was the interaction required to re-position images. Even 14 buttons was still unrealistic. In addressing this concern we looked at how television manufacturers had tackled the problem of navigating complex programming guides for hundreds of channels. The solution in almost all cases was a directional-pad complimented with a 'select' or 'OK' button on the television remote control.

We then revisited Tetrad's interaction and attempted to map the function of 14 buttons to just five. Our solution was a horizontal and vertical 'selection bar'. Users navigate the grid of images using these bars, rather than a free flowing cursor. The horizontal bar would map to the 'up' and 'down' buttons of the directional-pad while the vertical bar would map to the 'left' and 'right'. Obviously, both bars cannot be on-screen at the same time, so they fade-in and -out in response to directional-pad movements, e.g. while the horizontal bar is on-screen the user can move it up and down — if a user presses 'right' the horizontal bar fades out and the vertical bar fades in. Naturally, the navigation bars remember their position when fading-in and -out, aiding entry.

Once an individual has navigated to the column or row they wish to manipulate, they press the selection button. If a row is highlighted, and the select button is pressed, a circular movement to the right occurs, e.g. all images move right one space, with the last image becoming the first. Similarly, if a column is selected and the selection button is pressed, a circular movement downwards occurs, each image moving down one position, with the last image moving to the top of the column.

This approach also addresses our last concern: exposure of target images, as users control only the two selection bars. This means they are able to highlight individual rows and columns but *not* individual images — preventing inadvertent disclosure.

Figure 2 illustrates interaction within the shared space prototype. Let us assume that the highlighted square represents an image we wish to reposition, its intended location being one column along and one row down. The following steps are required:

(a) *Select column* - Using the 'right' navigation button we move the vertical selection bar to highlight the second column.

(b) *Move image downwards* - Using the selection button we activate the highlighted column, moving all images within the column downwards, with the last image becoming the first.

(c) *Select row* - Using the 'down' navigation button we move the horizontal selection bar downwards. The vertical selection bar fades out.

(d) *Move image right* - Using the selection button we activate the highlighted row, moving all images within the row to the right, with the last image becoming the first.

(e) *Submission* - Using the submission button, i.e. 'Menu' on the Apple remote, we can submit our efforts for assessment.

   i. *Success* - if successful, a 'smile' is displayed on-screen indicating that access has been granted to the service or system.

   ii. *Failure* - otherwise, a 'frown' is displayed on-screen indicating that access has been denied and that another attempt can be made.

These simplified facial gestures are generated and animated using the images within Tetrad. The images are repositioned and filled blue for a successful entry, red if otherwise. This approach allows us to extend the accessibility and simplicity of feedback while avoiding language.

This sequence of steps represents merely one way of repositioning an image; several other paths could be utilised.

Indeed, such redundancy has the potential to offer flexibility to the user, who, if feeling under threat, could take less obvious routes to reposition images. Furthermore, individuals could perform 'trick-moves', repositioning images not required for authentication to confuse onlookers.

3.2.2 Image Type

This experiment was not intended to test memorability of different kinds of images and so we decided to choose the most memorable image type so that any observed effects would be easier to attribute to the nature of the mechanism than to the efficacy of the image type used. Very few alternative authentication mechanisms are being used in real-life settings, since the majority remain within the experimental setting. Two exceptions are Handwing [16] and Passfaces [3]. The former uses doodles and the latter faces. In this experiment we decided to make use of faces because Passfaces is the more mature of the two, and has demonstrated its efficacy in the long term.

Humans are good at recognising previously seen faces. Bruce [4] explains that humans posses a face vocabulary similar to the lexical vocabulary that supports speech. If a face has been seen before, it becomes part of the person's vocabulary and will be recognised. Desimone *et al.* [7] studied Macaque monkeys and found that neurons in the inferotemporal cortex of a Macaque responded exclusively to faces, both of monkeys and humans, but since a weaker response was observed for human faces, this suggests a species preference. Yamane *et al.* [28] confirmed these findings. Moreover, facial memory is remarkably durable. Bährick *et al.* [2] found that people recognised the faces of their peers 90% of the time, even after periods as long as 48 years. Facial memory is clearly stronger than memory for other image types and should prove more memorable and be particularly suitable for image-based authentication.

4. EVALUATION

To evaluate Tetrad, we needed to test two aspects: how easy Tetrad was to use, and how easy it was for observers to identify the secret images if they watched someone else authenticating using Tetrad. The first aspect assesses the usability and the second the security of the mechanism. We therefore asked participants to engage in three tasks, using a within-subject design.

The three tasks are outlined below:

1. *Authenticating with Tetrad.*
   Although participants were familiar with alphanumeric authentication mechanisms and their mechanics and processes, it's unlikely they would be familiar with image-based authentication. Therefore, the first task asked participants to authenticate using Tetrad. This task assessed the usability of the mechanism and also prepared participants for the second task. We were also able to estimate the cognitive workload of the authentication task.

2. *Observing Authentication*
   This task asked the participant to determine the secret key being entered by another user. The participant viewed two videos, of equal length, one showing an unknown individual making an alphanumeric authentication attempt with an on-screen keyboard and the other showing the same individual authenticating using Tetrad. The *independent variable* is the authentication mechanism while the *dependent variable* is the success or failure of the participant determining the password entered. The *experimental hypothesis* is that Tetrad will be more resistant to casual observation than alphanumeric authentication.

3. *Questionnaire*
   The last task asked participants to provide additional information based on participants' thoughts and concerns regarding authentication in shared spaces.

4.1 Subjects

Eleven participants were recruited: 6 females and 5 males. Their ages ranged from 20 to 70 and included various backgrounds and professions, e.g. student, retired, professional etc.

4.2 Apparatus & Materials

The system used was an Apple MacBook, Model: MB062LL/A, with 2GB RAM. The MacBook's accompanying Apple Remote was used for interaction.

Tetrad required two image sets, one for the first task, and one for the second task. A total of 90 face images, 45 for each set, were extracted from the University of Massachusetts LFW database[2].

The videos used in Task 2 were captured using Screenium 1.0 in advanced using our MacBook. The first video required the Nintendo Wii to be connected to our MacBook using Elegato EyeTV Hybrid. The output from the Nintendo Wii was viewed using Elegato EyeTV 3. The captured videos were played full-screen during the trial using QuickTime 7 Pro.

Finally, participants were provided with pens and a handout to complete which included instructions for each task, questions regarding

[2] http://vis-www.cs.umass.edu/lfw/

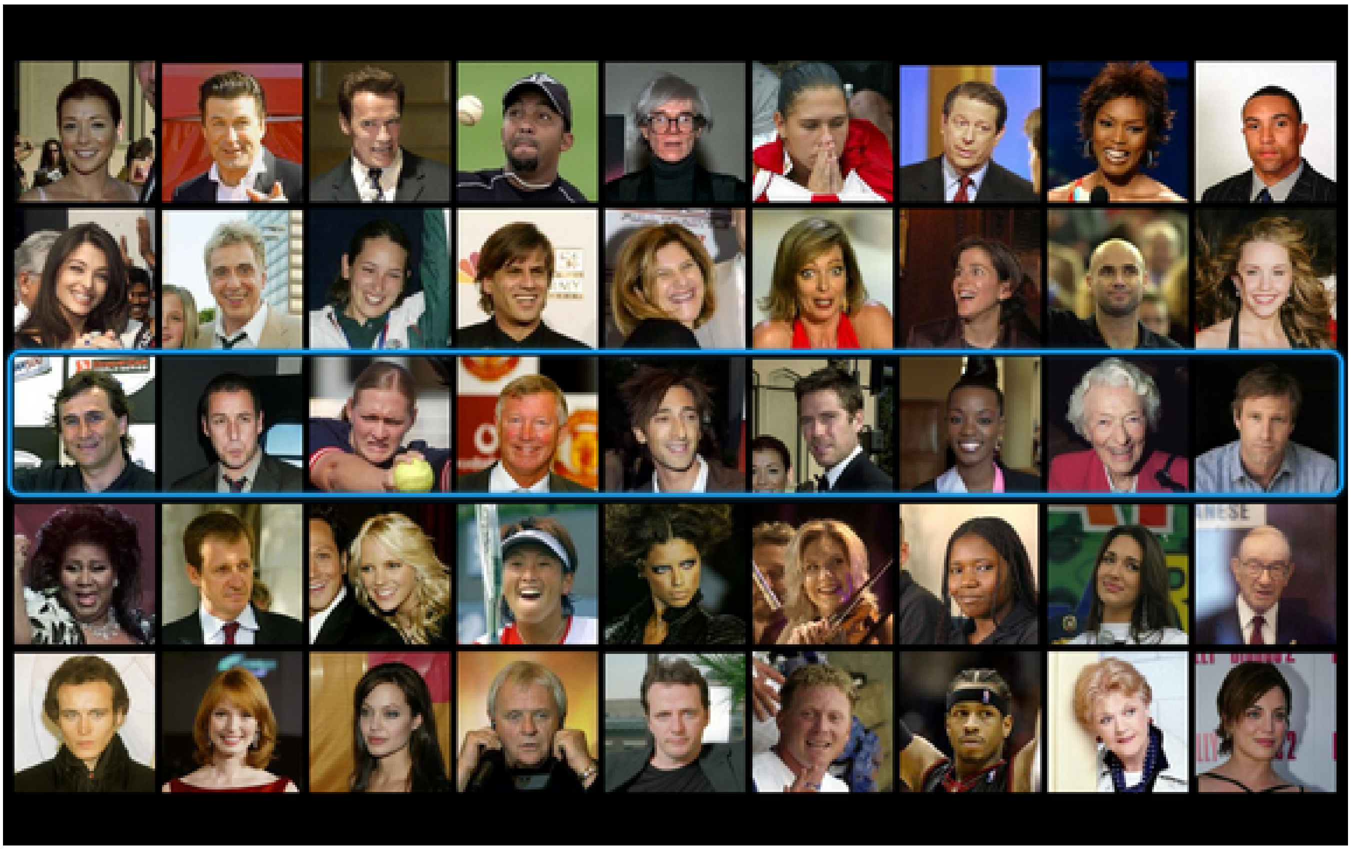

Figure 4: Shared Space Prototype

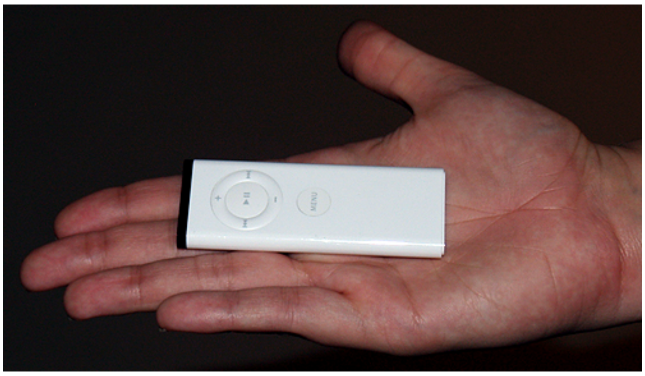

Figure 5: Apple Remote

the experiment, cognitive workload assessments and a brief one-page questionnaire.

4.3 Procedure

Participants were requested first to read the cover-page of our handout, which outlined the nature of the trial, estimated time to complete, three tasks which we expected participants to complete and our contact details should they have any queries. Lastly, participants indicated consent by signature before the experiment commenced.

Task 1 introduced our image-based authentication mechanism, Tetrad, to the participants and explained the concepts necessary to make a successful authentication attempt. Four images, which represented an image-based password, were printed as part of the instructions. Participants were advised there was no time-limit and that they did not need to memorise any of the images.

Upon completion of an authentication attempt, participants were requested to complete two evaluation procedures which examined cognitive workload. We used NASA-Task Load Index or NASA-TLX[3]. Participants first completed weighting then magnitude ratings for each sub-scale.

Task 2 instructed participants to watch two videos, of equal length. The purpose of viewing the videos was to extract the password entered by an unknown individual. In the case of alphanumeric authentication individuals wrote down the characters in their recalled position, e.g. if the password entered was ‘east’, a response of ‘seat’ would result in all characters being correctly identified but only one with the correct position, ‘t’. Similarly, for the image-based password, participants were requested to select four images from the image-set printed in the handout, as well as identifying that image’s position within the password. In both cases, participants were asked to rate their confidence on a scale of 0 to 100, i.e. how confident they felt about their estimations.

The second video had an additional question, which was for the participant to guess the alignment and position of the secret set of images within the image set when the person had completed moving all images around to authenticate. Participants indicated this on

[3] http://humansystems.arc.nasa.gov/groups/TLX/

a generic template of Tetrad's layout and rated their confidence in their estimation.

Upon completion of Task 2, participants were requested to complete two evaluation procedures which examined cognitive workload for extracting the image-based password. We used NASA-TLX thus participants first completed weighting then magnitude ratings for each sub-scale.

Lastly, participants were invited to complete a short questionnaire, i.e. Task 3.

5. RESULTS

5.1 Task 1

Task 1 was completed by all 11 participants, with every attempt being successful. Although time and memorability were not a consideration during this experiment, anecdotal evidence suggests that faces were memorable. Furthermore we were aware that the time it took to authenticate varied between participants. The evaluation procedure for the first task was completed by all 11 participants, which generated a cognitive workload score. If you are unfamiliar with NASA-TLX please visit their website.

In the rest of this paper, the TLX terminology used is as follows: Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Effort (E), Performance (P) and Frustration (F).

Table 1 shows the mean, median, minimum and maximum weighted ratings for Task 1.

	Workload Score
Mean	61
Median	60
Min	15
Max	91

Table 1: Workload Score for Task 1

The mean weighted rating or workload score for Task 1 is 61. The factors and their respective weights which contribute to the workload score can be seen in Figure 6. The factor contributing the most to the workload score is *Effort* with an approximate mean weight of 4. The median, minimum and maximum for *Effort* is 4, 2 and 5 respectively. While the factor contributing the least to the workload score is *Physical Demand* with an approximate mean weight of 1. The median, minimum and maximum for *Physical Demand* is 1, 0 and 4 respectively.

The factor with the highest rating is *Temporal Demand* at approximately 65. The median, minimum and maximum for *Temporal Demand* is 70, 20 and 100 respectively. The lowest rating is *Physical Demand* at approximately 15. The median, minimum and maximum for *Physical Demand* is 10, 0 and 50, respectively.

5.2 Task 2

All 11 participants attempted Task 2. After watching the first video all participants successfully extracted the characters within the alphanumeric password and their positions. The mean confidence rating was 95 out of 100, with 75 being the minimum and 100 the maximum confidence rating.

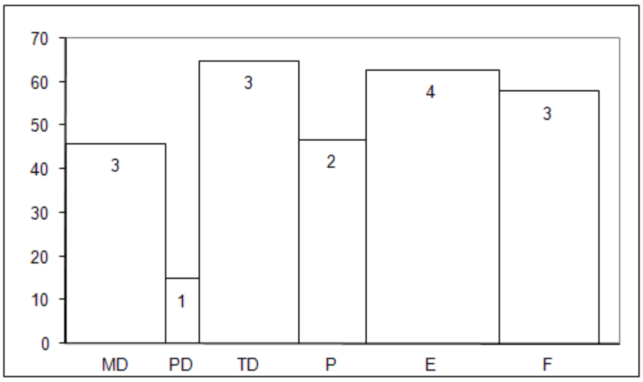

Figure 6: Mean Weighted Ratings for Contributing Factors for Task 1

However, after watching the second video, all participants failed to extract any of the images contained within the image-based password. Furthermore, 45% of participants identified at least 4 incorrect images, the mean being 2 images, with 27% of participants not identifying any images. If we remove these participants, the mean increases to approximately 3 images.

Participants did identify similar images, with two images in particular being identified by 45% and 27% of participants, respectively. If we remove those individuals who made no attempt to identify any images these values increases to 62.5% and 37.5% respectively.

The mean confidence rating from participants, regarding their identification of images, was approximately 27 out of 100. One of the participants identified only one image (incorrectly) but their confidence rating was 100, confident that above all else the single image they had identified was part of the image-based password. If we remove this outlier, the mean confidence rating drops to approximately 16 out of 100.

Participants were asked an additional question for the second video, which was the alignment of the image-based password. The alignment used within the video was diagonal but none of the participants identified this alignment, 27.2% could not identify the alignment, 27.2% identified horizontal as the alignment while the majority of participants, 45.4%, identified vertical as the alignment.

Participants were asked to rate their confidence in the alignment they had identified, on a scale of 0 to 100, the mean confidence rating was approximately 28.

The participants were asked to complete an evaluation procedure, which assessed workload, for the second video. Table 2 shows the mean, median, minimum and maximum weighted rating for the second video.

	Workload Score
Mean	74
Median	76
Min	23
Max	98

Table 2: Workload Score for Task 2, Video 2

The mean weighted rating or workload score for identifying the password is 74. The factors and their respective weights which contribute to the workload score can be seen in Figure 7. The factors contributing the most to the workload score are *Mental Demand* and *Effort*, with an approximate mean weight of 4.

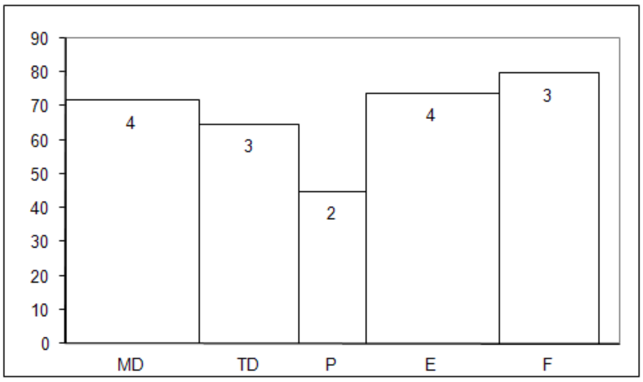

Figure 7: Mean Weighted Ratings for Contributing Factors for Task 2

The factor contributing the least to the workload score is *Physical Demand*, with an approximate mean weight of 0.

The factor with the highest rating is *Frustration* at approximately 80. The median, minimum and maximum for *Frustration* is 90, 10 and 100 respectively. The factor with the lowest rating is *Physical Demand* at approximately 10. The median, minimum and maximum for *Physical Demand* is 5, 0 and 45 respectively.

5.3 Task 3

Lastly, the answers to the questionnaire reveal that 54.5% of participants have purchased on-demand content through their television using their remote control. When asked if they would authenticate when not alone, 90.9% of participants said they would authenticate in front of others, with 70% ranking family as the least threatening and strangers the most.

6. DISCUSSION

We set out to assess the usability and security of the shared space Tetrad prototype. The first task assessed the usability of the mechanism. Our participants clearly put some effort into authenticating with Tetrad, but at least did not find it physically demanding. Fewer than half indicated that it was mentally demanding. This means that the workload score is less than than optimal, and could be improved. However, all participants managed to authenticate successfully in what was their first use of Tetrad, which is encouraging.

Task 2 assessed the observability of Tetrad. We asked participants to attempt to record the secret after watching someone enter either their alphanumeric password or their image-based password. It was no suprise that they all correctly observed the alphanumeric password. However, we did not expect that no one would be able to pick out at least one of the images involved in the secret images used in the image-based password. In terms of vulnerability to observation it would appear that Tetrad is as strong as we had hoped.

However, when considering that several individuals mistakenly identified the same faces, it could be that an individuals choice was influenced by attractiveness or race [6]. Thus, Tetrad's interaction redundancy, assumed to increase security, could itself prove *redundant* due to image-type and/or secret-creation. This could be tackled in numerous ways. Whether any such approaches could curb the inherent problems in using faces for authentication is another question.

It is interesting to note that the workload score for users attempting to uncover the image-based password was higher than that of the workload score for Task 1, indicating that authenticating with Tetrad requires less effort than observing someone else authenticating with Tetrad with a view to extracting their authentication secret. However, further investigation will be required to determine strength outside the realm of shared space.

With respect to security, Tetrad appears to meet the needs of shared space authentication in terms of resisting observation. However, resisting observation is not the same as immunity, and such a claim could only be assessed through longitudinal field studies coupled with varied interaction times and closely monitoring 'attacker' attempts to extract authentication secrets in a laboratory setting. The longer it takes an individual to authenticate using an alphanumerical password, the greater the probability of that password being compromised. It is entirely plausible that this is also true of Tetrad and we need to assess its strength with lengthier authentication attempts.

Furthermore, the extra effort perceived by our participants needs to be addressed if we are to convince people to use such a mechanism. Even though people complain about passwords, the undeniable fact is that they are very convenient when authentication is required [14] and people will always minimise their cognitive effort if at all possible [9, 15].

In strengthening the appeal and credibility of Tetrad we need to compare and contrast it to competing graphical authentication mechanisms. This will require us first to finalise our prototype, carefully considering our procedures for secret-creation and identification. Moreover, we need to contemplate the services and systems suitable to Tetrad, as we certainly do not advocate Tetrad as a one-size-fits-all solution. This product can then evaluated using traditional metrics, such as a theoretical security assessment and longitudinal usability assessments in terms of accessibility, login-time and memorability with an increased number of participants.

It is well established that people are the weak link when it comes to security [1, 19]. They make clear judgements about costs and benefits. If the cost of authenticating securely is balanced against their risk perception, even if it is inaccurately low, and they might well prefer not to use a mechanism such as Tetrad but rather to accept the risk of traditional mechanisms [26].

However, the idea of buying online content, using a device such as a Wii, iPhone or Apple TV, is relatively novel. Perhaps, as people start using these devices in shared spaces, the issues we have envisaged will come to the fore and companies will start looking for mechanisms akin to Tetrad to mitigate the threats of shared space authentication.

Tetrad's first evaluation was promising, but we hasten to add that it is a first step in the journey towards what we hope will be an acceptable, secure authentication mechanism for shared spaces, which will be as convenient as possible.

7. CONCLUSION

In this paper we introduced an authentication mechanism called Tetrad, which can be used to authenticate in full view of other people without the secret authentication key being compromised. Our evaluation showed that the system was indeed resistant to casual observation, specifically when contrasted and compared to alphanumeric authentication in a similar setting. Furthermore, it has addressed the main issues we identified with the Web prototype.

We believe Tetrad has a future and we intend embarking on a more extensive evaluation, with a larger number of evaluators, the better to isolate the usability issues with a view to further developing Tetrad. We are convinced that mechanisms such as Tetrad are essential in today's world where we will increasingly have to authenticate in public places.

8. REFERENCES

[1] ADAMS, A., AND SASSE, M. A. Users are not the enemy: Why users compromise security mechanisms and how to take remedial measures. *Communications of the ACM 42*, 12 (December 1999), 40–46.

[2] BÄHRICK, H. P., BÄHRICK, P. O., AND WITTLINGER, R. P. Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General 104*, 1 (1975), 54–75.

[3] BROSTOFF, S., AND SASSE, A. Are passfaces more usable than passwords? A field trial investigation. In *People and Computers XIV - Usability or Else! Proceedings of HCI 2000* (2000), S. McDonald, Ed., Springer, pp. 405–424.

[4] BRUCE, V. Recognizing faces. *Philosophical Transactions of the Royal Society of London, Series B, Biological Scienes 302*, 1110 (1983), 423–436.

[5] CLAESSENS, J., DEM, V., DE COCK, D., PRENEEL, B., AND VANDEWALLE, J. On the security of today's online electronic banking systems. *Computers & Security 21*, 3 (2002), 253–265.

[6] DAVIS, D., MONROSE, M., AND REITER, M. On user choice in graphical password schemes. In *13th USENIX Security Symposium* (2004). `http://www.cs.jhu.edu/~fabian`. Accessed Sept 2006.

[7] DESIMONE, R., ALBRIGHT, T. D., GROSS, C. G., AND BRUCE, C. Stimulus-selective properties of inferior temporal neurons in the Macaque. *Journal of Neuroscience 4*, 8 (1984), 2051–2062.

[8] DHAMIJA, R., AND PERRIG, A. Déjà vu: A user study using images for authentication. In *Proceedings of USENIX Security Symposium* (Denver, Colorado, August 2000), pp. 45–58.

[9] FISKE, S. T., AND TAYLOR, S. E. *Social Cognition*. Random House, 1984.

[10] FURNELL, S., PAPADOPOULOS, I., AND DOWLAND, P. A long-term trial of alternative user authentication technologies. *Information Management & Computer Security 12*, 2 (2004), 178–190.

[11] GAW, S., AND FELTEN, E. W. Password management strategies for online accounts. In *SOUPS '06: Proceedings of the second symposium on Usable privacy and security* (New York, NY, USA, 2006), ACM Press, pp. 44–55.

[12] HOANCA, B., AND MOCK, K. Secure graphical password system for high traffic areas. In *ETRA* (San Diego, 27-29 March 2006), p. 35.

[13] IVES, B., WALSH, K. R., AND SCHNEIDER, H. The domino effect of password reuse. *Commun. ACM 47*, 4 (2004), 75–78.

[14] MORRIS, R., AND THOMPSON, K. Password security: A case history. *Communications of the ACM 22*, 11 (1979), 594–597.

[15] O'GORMAN, L. Comparing passwords, tokens, and biometrics for user authentication. *Proceedings of the IEEE 91*, 12 (dec 2003), 2019–2040.

[16] RENAUD, K. A visuo-biometric authenticaton mechanism for older users. In *Proc British HCI 2005. Sept 5-9, Edinburgh* (2005), pp. 167–182.

[17] RENAUD, K. On user involvement in production of images used in visual authentication. *Journal of Visual Languages and Computing 20*, 1 (2009), 1–15.

[18] RENAUD, K., AND OLSEN, E. Dynahand: Observation-resistant recognition-based web authentication. *IEEE Technology and Society. Special Issue on Usable Security and Privacy. 26*, 2 (2007), 22–31.

[19] SASSE, A. Computer security: Anatomy of a disaster, and a plan for recovery. In *Workshop on Human-Computer Interaction and Security Systems* (Fort Lauderdale, Florida, Apr. 2003), ACM.

[20] SASSE, A., AND FLECHAIS, I. Usable security: Why do we need it? how do we get it? In *Security and Usability*, L. Cranor and S. Garfinkel, Eds. O'Reilly, 2005, ch. 2.

[21] TAN, D. S., AND CZERWINSKI, M. Information voyeurism: Social impact of physically large displays on information privacy. In *Proceedings CHI* (Fort Lauderdale, Florida, 5-10 April 2003), pp. 728–9.

[22] TAN, D. S., KEYANI, P., AND CZERWINSKI, M. Spy-resistant keyboard. In *OZCHI* (Canberra, Australia, 23-25 November 2005), pp. 1–10.

[23] TARI, F., OZOK, A. A., AND HOLDEN, S. H. A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In *Symposium on Usable Privacy and Security* (Pittsburgh, July 12-14 2006), pp. 56–66.

[24] V ROTH, K. R., AND FREIDINGER, R. A PIN-entry method resilient against shoulder surfing. In *CCS '04: Proceedings of the 11th ACM conference on Computer and communications security* (2004), ACM Press, pp. 236–245.

[25] WEINSHALL, D. Secure authentication schemes suitable for an associative memory. Tech. Rep. TR 2004-30, Hebrew University, Leibniz Center for Research in Computer Science, 2004.

[26] WEIRICH, D., AND SASSE, M. A. Pretty good persuasion: A first step towards effective password security for the real world. In *Proceedings of the New Security Paradigms Workshop 2001* (Cloudcroft, NM, 10-13 sep 2001), ACM Press, pp. 137–143. http://www.getrealsecurity.com/publications.htm.

[27] WIEDENBECK, S., WATERS, J., SOBRADO, L., AND BIRGET, J.-C. Design and evaluation of a shoulder-surfing resistant graphical password scheme. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces* (New York, NY, USA, 2006), ACM Press, pp. 177–184.

[28] YAMANE, S., KAJI, S., AND KAWANO, K. What facial features activate face neurons in the interferotemporal cortex of the monkey. *Experimental Brain Research 73*, 1 (1988), 209–214.