An experimental comparison of secret-based user authentication technologies

I.Irakleous, S.M.Furnell, P.S.Dowland and M.Papadaki

ABSTRACT

The paper presents a comparative study of software-based user authentication techniques, contrasting the use of traditional password and Personal Identifier Numbers against alternative methods involving question and answer responses and graphical representation. All methods share the common basis of some secret knowledge and rely upon the users ability to recall it in order to achieve authentication. An experimental trial is described, along with the results based upon 27 participants. The alternative methods are assessed in terms of practical effectiveness (in this context relating to the participant's ability to authenticate themselves a significant time after initial use of the methods), as well as the perceived levels of user friendliness and security that they provide. The investigation concludes that while passwords and PIN approaches garner good ratings on the basis of their existing familiarity to the participants, other methods based upon image recall and cognitive questions also achieved sufficiently positive results to suggest them as viable alternatives in certain contexts.

KEYWORDS

User authentication, Passwords, PINs, Question and Answer passwords, Graphical. passwords.

INTRODUCTION

One of the main aims of IT security is to ensure the availability of systems, while at the same time protecting them against unauthorised access, destruction and misuse. To this end, systems must control access to keep intruders and masqueraders out and permit access to the legitimate users. Although a variety of alternative techniques have been developed, using token-based authentication approaches (such as smart cards) or biometric solutions (such as fingerprint and facial recognition) (Sherman 1992), the most common methods of authentication in current systems are based upon secret-knowledge approaches such as passwords and Personal Identification Numbers (PINs). Despite their popularity, however, these methods are typically characterised as providing weak authentication, due mainly to the vulnerabilities introduced by end users (Jobusch and Oldehoeft 1989). Common problems include the fact that many users forget their passwords, and compromise their protection by sharing them with other people. As such, it is appropriate to consider alternative methods of authentication that may overcome (or

at least reduce) these problems, without introducing unnecessary complexity from the user perspective. Potential approaches here include altering the basis of the techniques away from purely recall-based approaches (which is the case with standard PINs and passwords), towards methods that rely upon less demanding concepts, such as recognition and provision of personal information.

A number of prior works have been conducted to enhance login security, whilst still retaining secret knowledge as the foundation of the approach. However, these approaches have typically been researched and evaluated on an individual basis, and it is desirable to evaluate them in a comparative study in order to obtain a more informed view of which, if any, are likely to represent acceptable alternatives. This paper presents the results of such an investigation, and begins by providing more details about the alternative methods. The discussion then proceeds to describe an experimental procedure by which five methods were evaluated in a practical trial. The results of this exercise are then presented, leading to consideration of the implications for practical systems.

BACKGROUND

The predominance of password-based methods can largely be explained by the fact that they are conceptually simple for both systems designers and end users, and can provide effective protection if they are used correctly. However, users themselves often compromise the protection provided. Previous investigations have revealed a variety of problems, and typically include the fact that passwords are often: badly selected (and therefore more easily guessed or cracked), forgotten, written down, shared with colleagues, infrequently changed, and kept the same across multiple systems (Klein 1990; Kessler 1996). The work of Klein, for example, found that 25% of the passwords (from a total sample of 15,000) were cracked after 12 months of exhaustive testing, with the help of a number of dictionaries including foreign words. More significant though is the fact that 21% of the passwords (more than 3,000) were cracked in the first week, and 2.7% of them were cracked in the first 15 minutes.

If the password approach is to be replaced or supplemented, then alternative means of authentication are clearly required. Surveys have shown that fundamentally different approaches, such as using biometrics (authenticating users based upon their physiological or behavioural characteristics) or token based approaches (magnetic cards, smart cards) are not readily accepted by the user community, who for various reasons express a strong preference for the methods they already know (Furnell et al. 2000). In addition, the financial cost associated with the introduction and maintenance of these other approaches will often preclude their adoption in many environments. For this reason, other approaches based upon secret knowledge, which do not incur any additional expenditure on hardware technologies, are considered desirable.

Previous research has highlighted the potential of question and answer based approaches, in which the user is asked to answer a series of questions, with correct answers leading to successful authentication. Clearly, such questions must require answers that are suitably

distinctive to the legitimate user, in order to prevent everyone having similar answers or their responses being too easy to discover or guess. Such questions may be based upon cognitive or associative information (Haga and Zvran 1991), as described later in the paper. The use of such questions has the potential advantage of using easily memorable (but nonetheless secret) information, but can involve a rather lengthy exchange between the user and the system in order to gain acceptance.

The solutions discussed so far have all been of a textual nature. However, given the transition to graphical user interfaces that has occurred during the last two decades, it is perhaps unsurprising that graphical authentication approaches have also arisen. For example, Blonder (1996) patented a graphical password in which the user can select a number of areas in a picture as a password. The weakness of this technique was that the user had to recall the location and the order of the regions. In another alternative, proposed by Jermyn et al (1999), the 'pasword' method was realised as a simple picture drawn on a grid. Other variations include the recognition of previously seen images, with an example being the D ej a Vu system (Dhamija and Perrig 2000).

AN EXPERIMENTAL STUDY OF ALTERNATIVE METHODS

In order to enable a comparative study of alternative authentication methods, an experimental trial was devised incorporating five secret-knowledge based techniques. The methods selected were PINs and passwords (familiar methods, included to provide a baseline for reference), alongside two question and answer methods (using cognitive and associative questions respectively), and a graphical technique using an image-based PIN (hereafter termed ImagePIN). The study sought to assess the practical effectiveness of the techniques, as well as friendliness and the perceived level of security from the user's perspective.

The effectiveness was gauged by means of a practical trial, using specially designed profiling and authentication systems to present the various techniques to a series of participants. Opinions relating to the friendliness and security of the methods were then obtained using a written questionnaire – completed by participants after they had participated in an authentication phase and witnessed their own performance using each technique. The construction of the experimental tools and the follow-up questionnaire are described in the subsections that follow.

The Profiler

The Profiler required each participant to identify him/herself and then provide appropriate responses for each of the methods under test. The profiling procedure for each of the methods is summarised below.

– **Passwords and PINs**

The implementation of these methods was fairly standard, with each participant being asked to supply a 4 digit PIN and a password of at least 8 characters. Participants were requested not to select a password or PIN that they already used on other systems, as the aim of the exercise was to assess their ability to recall new details, and thereby put these more familiar methods on an equal footing with the other techniques when it came to assessing ease of information recall. Nonetheless, as later results will indicate, some participants did not follow this guideline.

– **Cognitive questions**

Participants were asked to provide answers to a series of twenty questions, each requiring factual or opinion-based answers. The questions requested information that was personal to the participant, and would therefore be difficult for a potential masquerader to guess in an operational scenario. The questions used are listed in Table I.

What is your mother's maiden name?
Where were you born?
What is your favourite colour?
What was the name of your best friend at school?
What is your favourite music?
What is your favourite food?
What was the name of your first pet?
Which primary school did you go to?
What is your favourite sport?
Where was your first house?
What make was your family's first car?
How old were you when you had your first kiss?
What is your favourite film?
Where was the first place you remember going on holiday?
What was your favourite subject at school?
What is the most important part of your body?
What is your favourite type of animal?
What is the name of your favourite relation?
How many cousins do you have?
What is your favourite shape?

Table I : Cognitive questions

Even in cases where the participants might not have had a genuine answer (e.g. they may never have had a pet), it was expected that they would still be able to

provide a response that could later be reproduced if prompted to answer that question.

– **Associative questions**

Participants were then asked to provide word association based responses to a set of twenty keywords. The keywords are listed in Table II, and were carefully chosen to ensure that a number of different responses were theoretically possible in each case

Blue	House	Table	Computer	Friend
Peace	Glass	Marriage	Sea	Love
Cat	Music	Fire	Seven	Video
Father	Food	Remote	Fast	Door

Table II : Associative keywords

– **ImagePIN**

The user had to select five images from a number of icons, by clicking on them with the mouse. Later authentication would work by the user reselecting the same images in the correct sequence.

The user interface of the profiling system is illustrated in Figure 1.

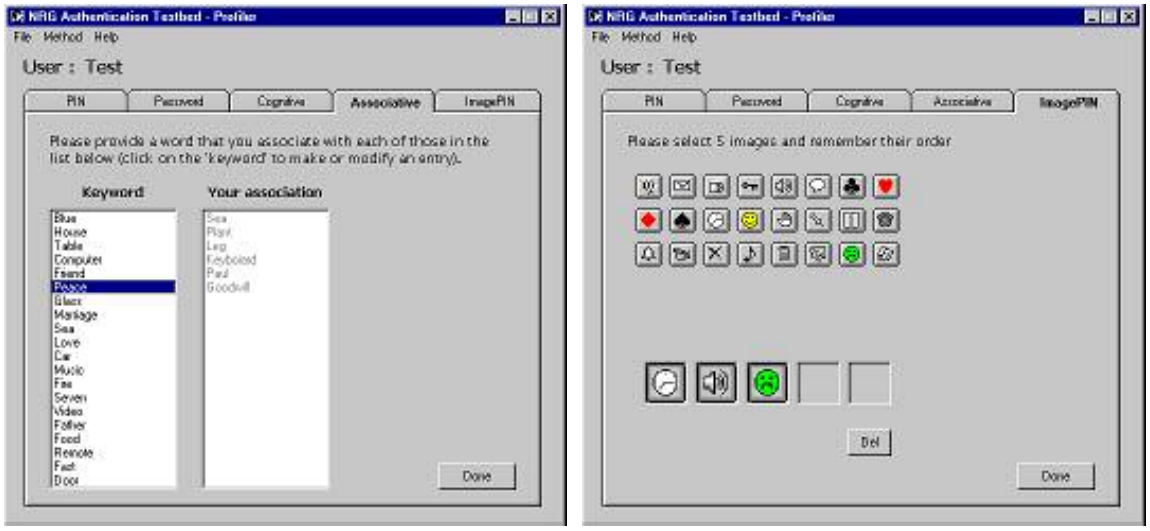

Figure 1 : Profiler system (showing Associative Questions and ImagePIN screens)

After the profile had been created, a short training exercise was performed using the second program, the Authenticator, in order to familiarize the users with how the later authentication test would work. After this, it was up to the participants to attempt to remember the details they had provided in order to perform the later authentication tests.

The Authenticator

The authentication tests took place one month after the initial profiling, with the aim of assessing whether the participants were able to adequately recall the information that they had previously provided during profiling and thereby authenticate themselves successfully. The interface of this system was very similar to that of the Profiler, and two aspects are illustrated in Figure 2.

In the case of the PIN, Password and ImagePIN methods, the participant was directly asked to provide the same information as originally profiled. For the cognitive and associative methods, however, they were asked to answer five randomly selected questions out of the twenty that had been profiled in each case. This was considered to represent a good simulation of how such question and answer authentication techniques would be implemented in practice.

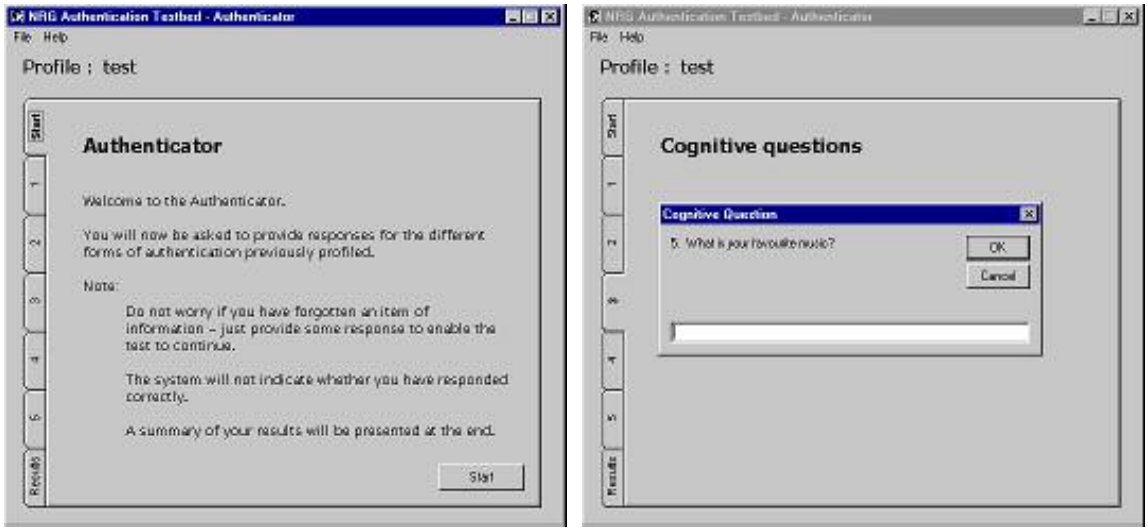

Figure 2 : Authenticator system (showing Welcome and Cognitive Question screens)

Participant questionnaire

Following the authentication test, all the participants were asked to complete a questionnaire, in order to determine their regular exposure to user authentication methods in other contexts and to assess their views about the different methods under trial. The following key elements of information were collected:

- the number of different passwords they use;
- the frequency with which they use PINs and passwords;
- whether or not they use the same password in their applications;
- the composition of their password(s) (e.g. letters, numbers, symbols)
- ranking the trialed methods according to the perceived user friendliness, level of security, and overall preference.

A total of 27 participants were involved in the profiling and subsequent authentication testing, and the results of the study are described in the next section.

EXPERIMENTAL RESULTS

The results presented here encompass the effectiveness of the techniques (in terms of user recall) that was observed in the practical trial, as well as the participant's subsequent opinions in relation to the methods. It should be noted that, in the discussions and graphs that follow, the percentage figures have been rounded to whole numbers.

In order to gauge their current exposure to authentication techniques, the participants were asked how many different passwords they have to remember and how often they use them. Only 4% of the participants had just a single password or PIN to remember, whilst 59% claimed to have up to five, 22% claimed to have between five and ten, and 15% claimed to have in excess of this number. In terms of their frequency of use, 85% claimed daily usage, while 11% indicated once every two days, and 4% claimed three to five times a week. No one indicated that they used PINs or passwords on a less frequent basis than this. It can be concluded from these findings that, although the overall sample of users was small, the participants all had considerable experience of using traditional authentication methods and were therefore suitably qualified to participate and comment on this study.

The practical evaluation began by examining the participant's performance in relation to the password and PIN methods. The results indicated that 70% of the participants had succeeded in authenticating themselves using passwords, and a similar proportion (67%) were successful using the PIN based method. Although these results initially appear very encouraging from the perspective of the participants being able to accurately recall the details after an absence of a month, the results of the accompanying survey revealed that a significant number of people had not followed the request to use different passwords and PINs than the ones normally used in other applications. In fact, only 56% used different passwords and 41% used different PINs. Within these subgroups, the authentication success was markedly lower - 53% of them succeeded in the password test and only 36% in the PIN version. By contrast, within the subgroups that used the same details as in other systems, 92% of them succeeded with passwords and 87% succeeded with PINs, so these figures can be considered to have artificially inflated the overall results.

In the cognitive and associative question tests, the participants were presented with a random selection of five questions out of the twenty that they were profiled for. Authentication was judged to be successful if all five questions were answered correctly. With the cognitive questions, a success rate of 59% was observed, whilst a number of further participants did succeed in answering a proportion of the questions presented to them. The distribution of correct answers in the cognitive test is shown in Figure 3.

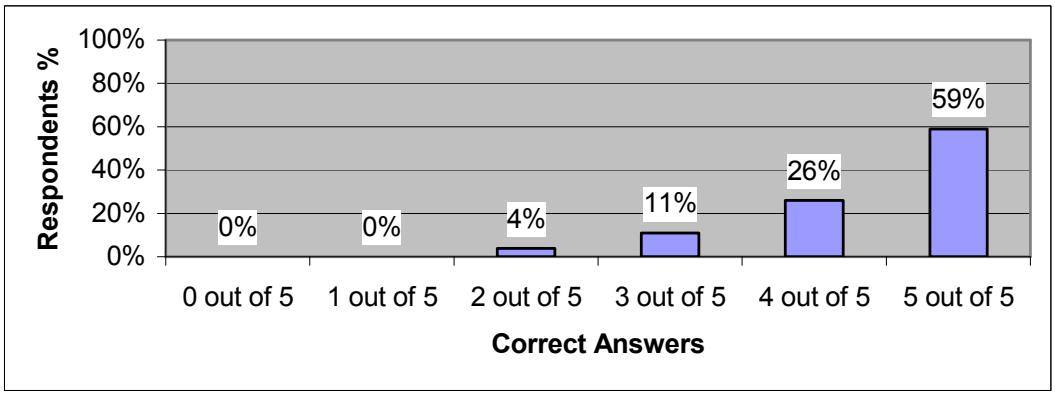

Figure 3 : Distribution of correct answers in cognitive questions

With the associative questions, the success rate was significantly lower. Only 4% (equivalent to one person) managed to correctly answer all five questions and the distribution of correct answers across five random questions is shown in Figure 4.

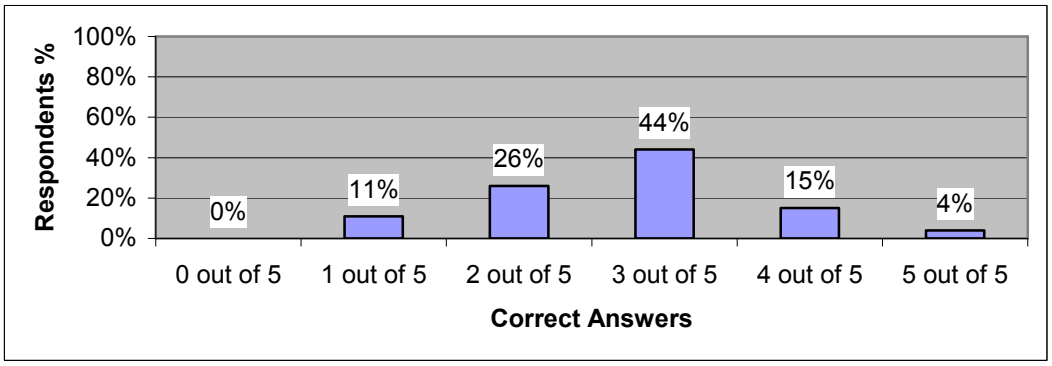

Figure 4 : Distribution of correct answers in associative questions

These results suggest that the associative question method is extremely problematic in relation to the correct recall of the information, and that participants are inconsistent in the words that they most readily associate with the keyword prompts. A further problem observed in the results of this study was that many participants chose the same associations for certain keywords, suggesting that the method could be easily targeted for masquerade attacks if used in practice. Table III summarises the cases in which the same associations were chosen for each keyword. The highest frequency of duplication was 44%, in which respondents had chosen the word "control" as the associative response to the keyword "remote".

Keyword	Frequent word associations
Blue	Sky (41%), Sea (15%)
House	Big (15%)
Table	Food (22%)
Computer	Work (11%), Game (7%), Internet (7%)
Peace	War (15%)

Glass	Wine(22%), Broken (11%)
Sea	Blue (11%)
Love	Hate (11%), Marriage (7%)
Music	Rock (15%), Dance (7%)
Fire	Red (11%), Alarm (11%), Engine (7%)
Seven	Film (15%), Seven (7%), Days (7%)
Video	Games (11%), Movie (11%), Tape (7%)
Father	Mother (19%), Names (15%)
Remote	Control (44%)
Fast	Food (22%), Car (19%)
Door	Key (11%), Open (11%), Closed (7%)

Table III : High frequency associative responses

For the final technique, the ImagePIN, the participants had to recall their graphical PIN by reselecting the original icons in the correct order, with 63% being successfully authenticated. Even though the implementation of the method offered the participants the opportunity to undermine the security by selecting the same icon five times, only two participants actually did this.

Figure 5 summarizes the overall results of the authentication tests, indicating the percentage of respondents who would have been successfully authenticated using each of the methods.

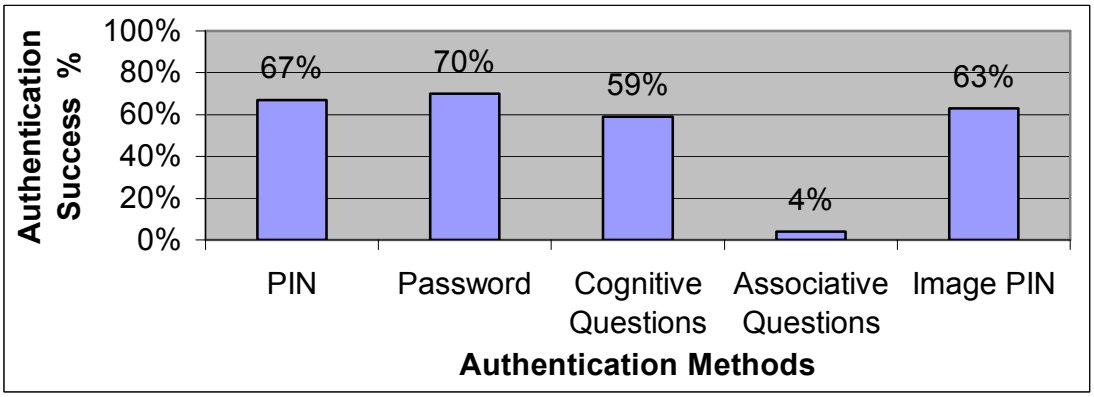

Figure 5 : Authentication methods success

Having experienced the techniques and witnessed their own performance, the participants were asked to rate the approaches on the basis of user-friendliness, security, and overall preference.

In terms of user-friendliness, participants were asked to assess the methods on a five-point scale, progressing from ‘easy’ to ‘hard’. The best outright indicator of preference in this case was where methods were ranked as ‘easy’. In this context, passwords were ranked first, receiving 48%, followed by the PIN method with 44%. The third position was shared by the cognitive question and ImagePIN methods, with 22% respectively.

Last was the associative method with only 4%. Taking a wider view, and considering the total percentages for which methods were rated 'medium' or above, the password was still favourite, with 96%, followed by the PIN with 93%, cognitive questions with 81%, the ImagePIN with 59%, and associative questions with 48%. Looking from this viewpoint serves to place some separation between the cognitive and ImagePIN methods, and shows that more people tended to express concern over the friendliness of the latter technique. The full results are presented in Figure 6.

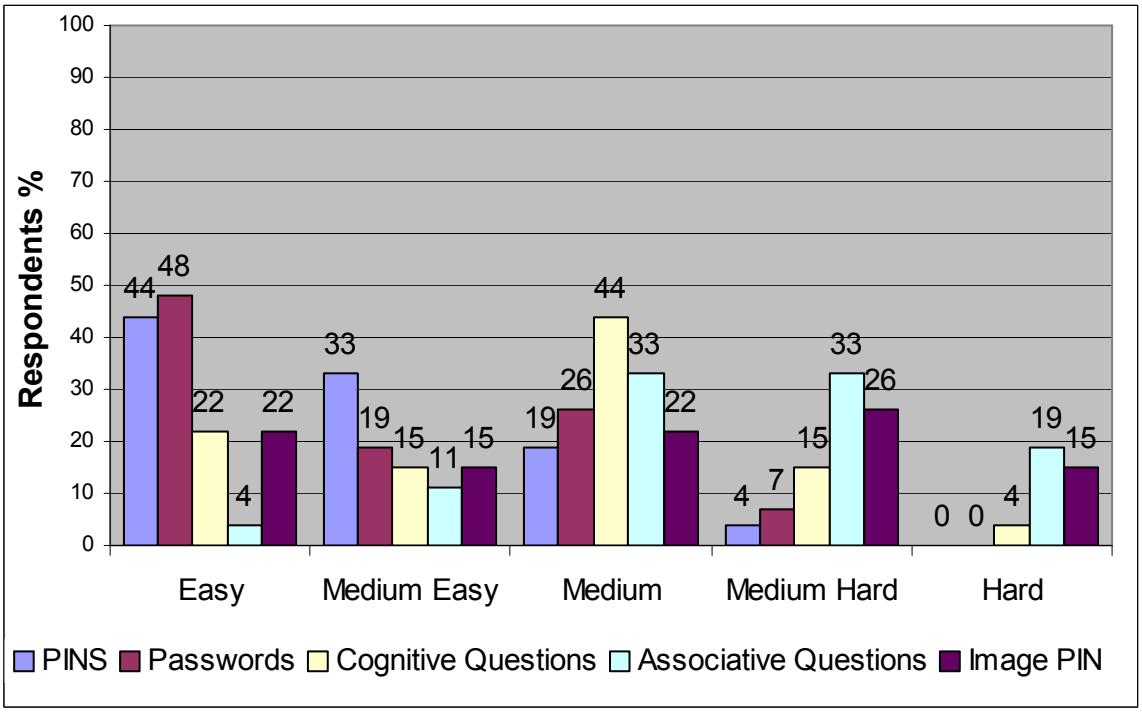

Figure 6 : Perceived user-friendliness

The second ranking addressed the perceived level of security. In this case, the password still fared well, with a combined total of 78% rating it to offer a 'medium' to 'high' level of protection. In this instance however, the popularity was also equalled by the cognitive and ImagePIN methods (and it can be noted that both of these methods actually exceed the results for passwords if only the 'high' and 'medium high' ratings are considered). Meanwhile, the PIN method attained 53%, and the associative approach was again ranked lowest, with 45% ranking it in the 'medium' to 'high' range. Figure 7 presents the perceived level of security for each authentication method.

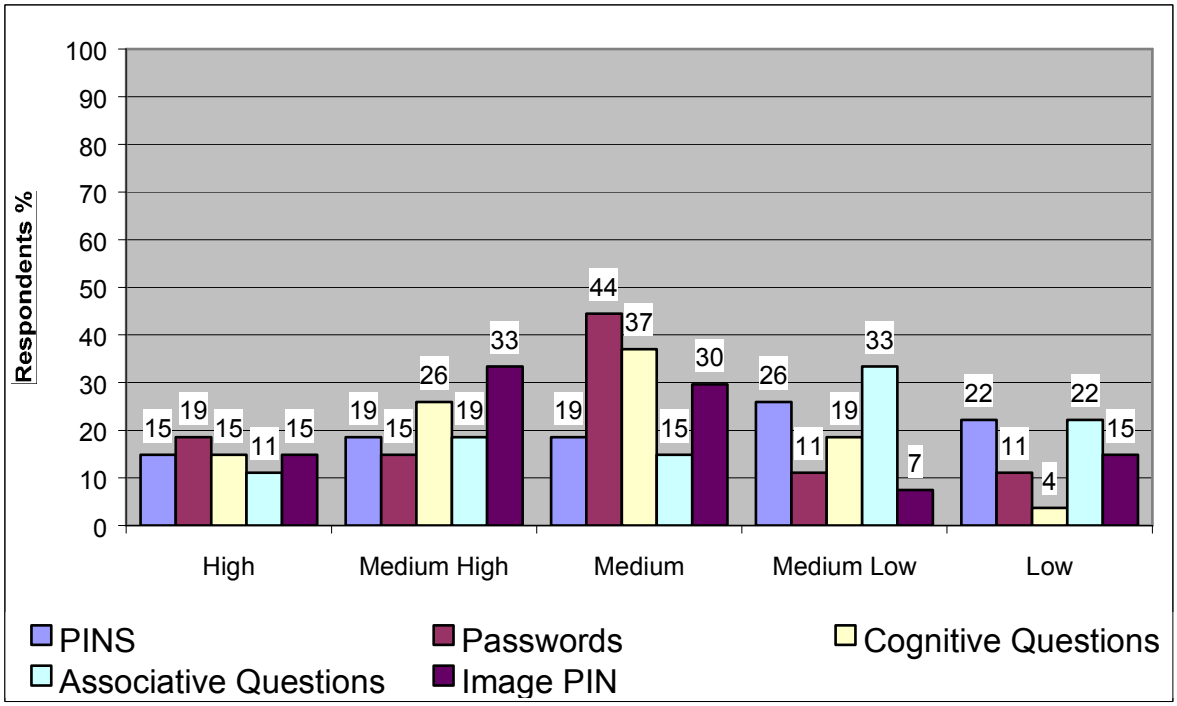

Figure 7 : Perceived security

The final question asked the participants to rank the methods according to their overall preference. Password method was again the most preferred form of authentication, with 44%, as shown in Figure 8. In the second place is PIN method with 22%, and third is the ImagePIN method with 19%. It is therefore clear that the more traditional and familiar methods of authentication are still the most readily accepted. However, if the rationale behind the alternative methods is accepted (i.e. that passwords and PINs are open to compromise), then it is relevant to give further consideration to the results and responses in the other categories.

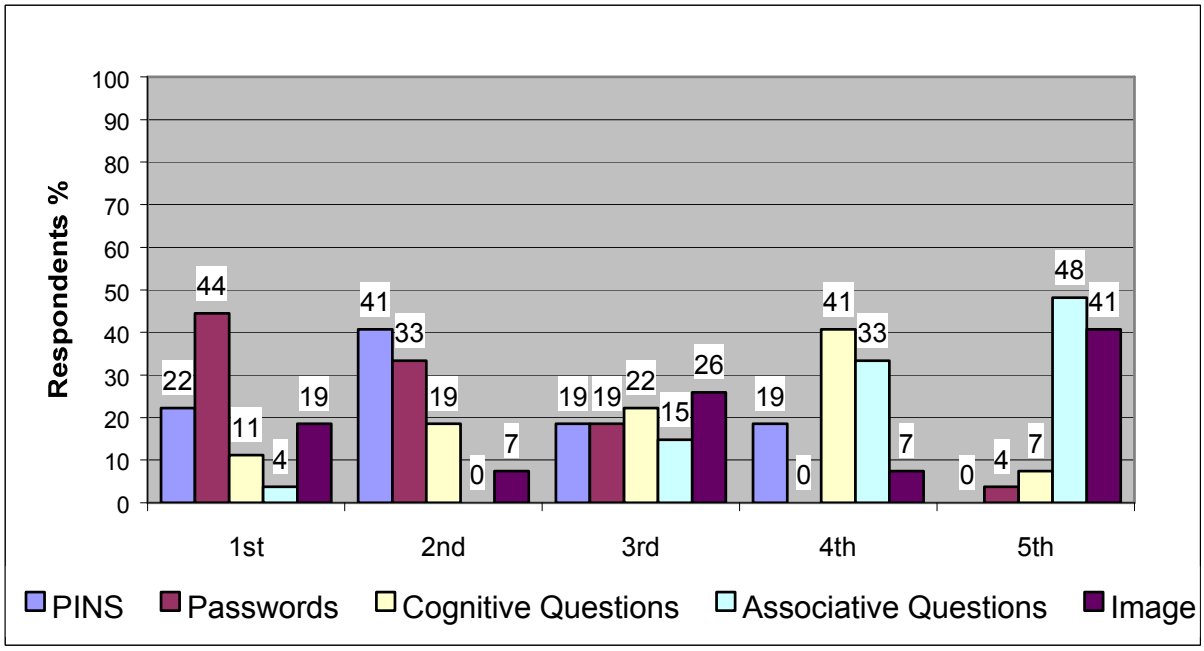

Figure 8 : Overall preference of trailed methods

DISCUSSION

Although people clearly prefer passwords and PINs, the other results obtained continue to suggest concerns about the level of security they actually provide. For example, analysis showed that 48% of the participants selected passwords that might be easily guessed or cracked (e.g. based upon dictionary words, variations of their name, or foreign words written in English characters). Only 38% of participants used an alphanumeric combination, and fewer still (4%) introduced other symbols into their passwords. These results increase the attractiveness of the other methods, which may be less vulnerable to such unintentional compromise.

The participant's performance in relation to the cognitive questions was relatively strong, with 59% successful authentication (interestingly, a previous study by Haga and Zviran (1991) reported better results, with 74%, for a broadly similar set of cognitive questions). Further points noted about the cognitive technique were the relatively time consuming nature of the profiling phase, in which the participants had to provide answers for all 20 questions. In addition, several participants expressed concern about the nature of the information that was requested, and were reluctant to provide genuine answers to the questions during the trial for fear that the information might be accidentally divulged. Particularly notable questions in this respect were in relation to mother's maiden name (a commonly used identity verification question in other contexts, such as bank accounts), place of birth, and age of first kiss. Overall, however, this method was ranked relatively high in terms of perceived user-friendliness and security.

The associative approach proved to be weak as an authentication method, with the performance of the participants (only 4% success) suggesting that it cannot deliver an

adequate level of effectiveness. It is considered that this poor performance can in part be explained by the fact that users still have to remember potentially abstract information (as opposed to the more recognition-oriented approaches of cognitive questions), placing more or less the same demand on their memory as the password method. In addition, the results raise questions over the level of security that the approach would provide – the fact that many participants chose the same word associations suggests that the method would be vulnerable to attackers attempting to guess the likely associations. At the very least, this requires that more care must be taken in the selection of the keywords, to ensure that none of them have obvious first-choice answers. It may again by observed that a previous study of the same basic method reported a far higher success rate, with an overall average 69% recall after a period of three months (Haga and Zviran 1991). It must be noted, however, that there was a significant difference in the experimental procedure in this case, as participants were asked to select their own keywords, as well as the appropriate associative responses.

The ImagePIN approach demonstrated positive results in the authentication phase, with 63% success, placing it very close to the results observed for passwords. This result is partially explained by the findings from previous surveys, which have shown that people tend to have less difficulty in recognising previously seen pictures than they do in recalling passwords or phrases from the memory (Bensinger 2000; Sasse et al. 2001). In addition to its practical effectiveness, the ImagePIN scored well in terms of user acceptance, which bodes well for the rating that it might receive if users were given additional time to familiarise themselves with it. Another point worth noting is that the ImagePIN method as implemented for the study was rather crude, with a set of standard Windows icons having been used as the selection of available images. With more consideration given to the number and range of images available, it is likely that the perceived user friendliness of the approach could be further improved. Having said this, there was also a fairly high proportion of respondents who put it as their clear least favourite, whereas most of the other methods did not elicit such strong negative opinions.

Although some techniques suggested themselves as potential alternatives to standard passwords and PINs, it does not necessarily follow that they would make good replacement methods in all contexts. For example, the use of cognitive questions could potentially be too time consuming as a regular means of login authentication. The technique could, however, provide a good secondary level of authentication, which could be invoked in a number of scenarios (e.g. when a user tries to perform a sensitive activity, in response to a suspected masquerade attack, or simply at random intervals). Image based authentication techniques could be more easily implemented as an initial login technique, but their applicability would be limited to systems that are able to offer sufficient graphical displays. This would, currently, rule out devices such as mobile phones (where standard PIN methods currently predominate), but could still usefully include other PIN-based devices such as Personal Digital Assistants and Automated Teller Machines.

CONCLUSIONS

The paper has presented a comparative study of five user authentication techniques based upon secret knowledge. With the clear exception of the associative approach, the practical effectiveness of the techniques was closely comparable. However, in terms of the overall preference, the known and familiar methods of passwords and PINs were, perhaps unsurprisingly, favoured. Having said this, if the previous arguments and evidence regarding the weaknesses of these methods are accepted, then it may be reassuring to consider that the cognitive and ImagePIN methods are already comparably effective from a user recall perspective, and given further training and exposure these methods may gain greater acceptance.

Although the initial results are encouraging, two significant aspects were not addressed by the work to date. Firstly, the judgements relating to user-friendliness of the methods were based on a relatively brief level of exposure in the case of the question and answer approaches and the ImagePIN method. A longer-term trial is therefore required in which participants use the alternative methods in day-to-day operations, in place of their normal passwords or PINs. This will allow a more accurate impression to be gained regarding the perceived user-friendliness. The second aspect that requires attention is the level of protection that the new methods actually deliver when compared to the traditional approaches. The study described here did not attempt to assess the ability of participants to successfully masquerade as other users – although the duplication of responses that was observed for the associative questions would suggest that this would clearly be possible. As such, the methods need to be assessed in terms of their susceptibility to compromise by informed parties (e.g. those who know the person they are trying to impersonate, and may therefore be able to determine the correct cognitive and associative responses) and by simple guesswork. A further aspect that is worthy of investigation (in relation to the ImagePIN, or indeed other graphical approaches) is how well users are able to cope with multiple sets of login information. As earlier results confirmed, users today have to remember multiple PINs and passwords, and it is therefore relevant to know whether remembering multiple ImagePINs serves to simplify the issue or complicate it further. These aspects represent the focus of ongoing investigation by the authors.

REFERENCES

Bensinger, D. (1998), “Human Memory and the Graphical Password”, available at: http://www.passlogix.com/

Blonder, G. (1996), United States Patent, 1996, United States Patent 5559961.

Dhamija, R. and Perrig, A. (2000), “D ej a Vu: A User Study Using Images for Authentication”, SIMS/ CS, University of California Berkeley.

Furnell, S.M., Dowland, P.S., Illingworth, H.M. and Reynolds, P.L. (2000), "Authentication and Supervision: A survey of user attitudes", Computers & Security, vol. 19, no. 6, pp529-539.

Haga, W.J. and Zviran, M. (1991), "Question-and-Answer Passwords: An Empirical Evaluation", *Information systems*, vol. 16. no. 3. pp.335-343.

Jermyn, I., Mayer, A., Monrose, F., Reiter, M.K., and Rubin, A.D. (1999), "The design and analysis of graphical passwords". In *Proceedings of the 8th USENIX Security Symposium*, August 1999.

Jobusch, D.L. and Oldehoeft, A.E. (1989), "A Survey of Password Mechanisms : Part 1", Computers & Security, Vol. 8, No. 7: 587-604.

Kessler, G.C. (1996), "Passwords – Strengths and Weaknesses", January 1996, available at: http://www.garykessler.net/library/password.html.

Klein, D. (1990), "Foiling the cracker: A survey of, and improvements to, password security". In *Proceedings of the 2nd USENIX Security Workshop*, August 1990.

Sasse, M.A., Brostoff, S. and Weirich, D. (2001), "Transforming the 'weakest link'- a human/ computer interaction approach to usable and effective security", BT Technology Journal, vol. 19, no 3. pp122-131.

Sherman, R. (1992), "Biometrics Futures", Computers & Security, vol. 11, no. 2: 128-133.