Nosseir, A. and Connor, R. and Dunlop, M.D. (2005) *Internet authentication based on personal history - a feasibility test.* In: Proceedings of Customer Focused Mobile Services Workshop at WWW2005. ACM Press. ISBN 1-59593-046-9

Internet Authentication Based on Personal History – A Feasibility Test

Ann Nosseir
Computer and Information Sciences
University of Strathclyde
26 Richmond St.
G1 1XH
Ann.Nosseir@cis.strath.ac.uk

Richard Connor
Computer and Information Sciences
University of Strathclyde
26 Richmond St.
G1 1XH
Richard.Connor@cis.strath.ac.uk

Mark Dunlop
Computer and Information Sciences
University of Strathclyde
26 Richmond St.
G1 1XH
Mark.Dunlop@cis.strath.ac.uk

ABSTRACT

On the Internet, there is an uneasy tension between the security and usability of authentication mechanisms. An easy three-part classification is: "something you know" (e.g. password); "something you hold" (e.g. device holding digital certificate), and "who you are" (e.g. biometric assessment) [9]. Each of these has well-known problems; passwords are written down, guessable, or forgotten; devices are lost or stolen, and biometric assays alienate users.

We have investigated a novel strategy of querying the user based on their personal history (a "Rip van Winkle" approach.) The sum of this information is large and well-known only to the individual. The volume is too large for impostors to learn; our observation is that, in the emerging environment, it is possible to collate and automatically query such information as an authentication test.

We report a proof of concept study based on the automatic generation of questions from electronic "calendar" information. While users were, surprisingly, unable to answer randomly generated questions any better than impostors, if questions are categorized according to appropriate psychological parameters then significant results can be obtained. We thus demonstrate the potential viability of this concept.

General Terms

Experimentation, Security, Human Factors, Standardization, and Verification.

Keywords

Internet security, password, human memory, user studies, security usability Identity theft, personal electronic data, user mobility,

1. INTRODUCTION

Users require secure authentication over the Internet with minimum personal effort. In the Internet café culture, they require to access sensitive personal information and transaction facilities on a “drop-in” basis from public machines. Furthermore users require to interact securely with many different organizations; performing separate authentication activity for each aggravates the security problem, and is furthermore resented by users. The state-of-the-art is the use of trusted third party for single sign-on via password, which gives an acceptable compromise.

Figure 1 presents a general framework of the information flow between a user, third trusted party and other web sites.

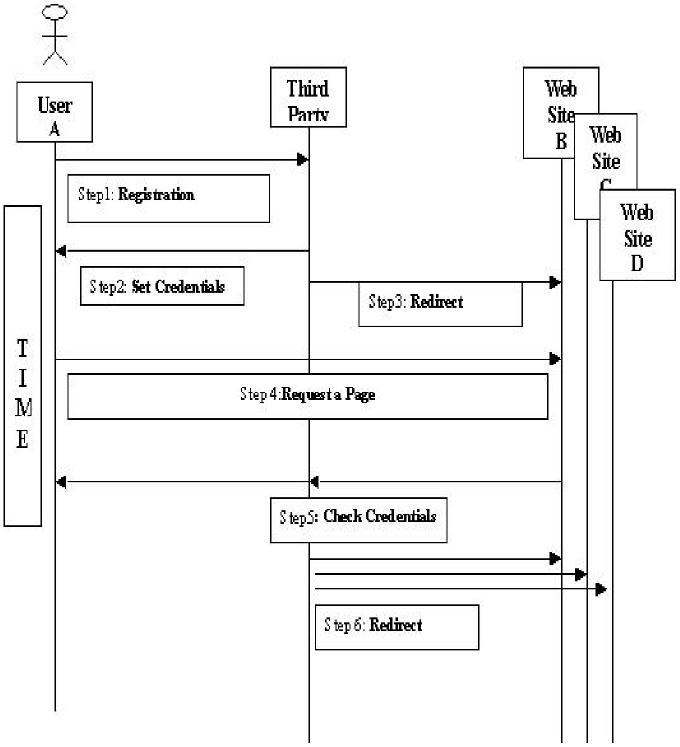


Figure (1) Authentication using trusted third party

There are two major authentication stages: the initial authentication stage and the authentication at instant time stage [8]. (i) Users, in the initial authentication stage, register through some appropriate process with the trusted third party (TTP), at which point a shared secret is created. (ii) When users attempt to access a secure site, they are redirected to the TTP, which performs an authentication process with the user and reports the result back to the originating site. If the TTP site is able to keep "session" information, then the TTP can confirm authentication with other sensitive sites during the session without further user interaction.

A TTP can use any authentication mechanism; clear contenders are passwords, digital certificates, or biometric assay. Certificates offer users a far higher level of electronic security, but are not memorable and therefore require to be kept on a device, which is restrictive and open to physical loss. Biometric assay is, still, unacceptable to users and carries a relatively high opt-in cost, although of course these factors may change over time. We thus contend the majority requirement for TTP authentication is information the user knows, which is fully portable and lowest entry cost

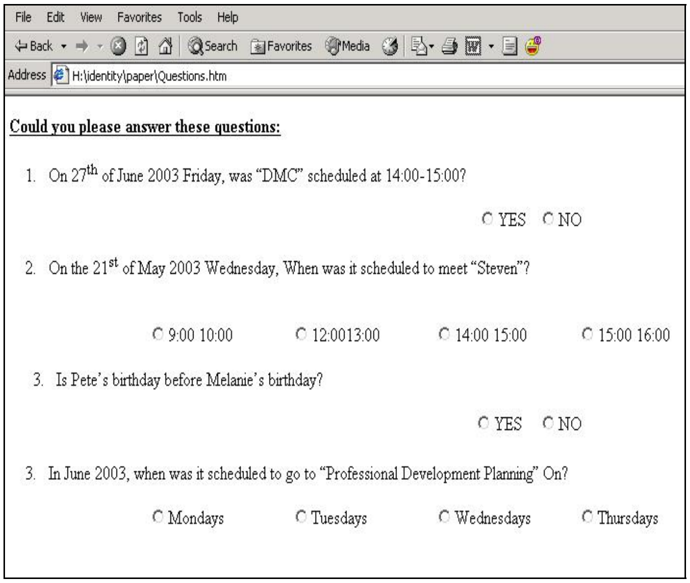
Passwords are not without human problems however. Even small passwords are difficult to remember. Yan [7] addressed this problem and remarked that there are trade offs between good non-guessable passwords and the limitations of the human memory. It is hard for users to remember random passwords, but others are guessable; although passwords provide users with mobility, they can be stolen, guessed, or cracked. Zviran [12] presented cognitive passwords as a method to overcoming this. Cognitive passwords are 'question based' according to either fact or opinion, and are shown to be more memorable than alphanumeric strings. Significant others could not remember or guess more that 27%. Cartwright [3] presented challenge response questions that authenticate customers. The questions are generated from different resources: consumer credit, vehicle ownership, property ownership, and reference files.

Dhamija and Perrig [6] and Brostoff and Sasse [2] presented an approach to improve the security of these systems using passwords. The approach relies on recognition-based, rather than recall-based authentication. The former researchers developed a web-based prototype that allows users to create image portfolios and to authenticate themselves to the system later by selecting their portfolios from a challenge set. Brostoff and Sasse presented passfaces as an alternative to password. Passfaces are a combination of different human faces. They mentioned that users better recognize passfaces than passwords.

2. QUESTION BASED-MODEL

Our research presents a solution that relies on electronic personal history. The underlying security relies not on a secret, but a volume of information known in detail only to the authentic user and too large for an impostor to learn. As much pertinent history is already recorded electronically, it is feasible to generate sparse questions from a very large data set, with each authentication attempt generating a different set of questions. Such information may be continually and incrementally released to the TTP and used as the basis of authentication, rather than the sharing of an explicit secret. We report here results of a pilot study based on a prototype of this concept.

Figure (2) shows the prototype system in use. The TTP has access to a large volume of personal data, and randomly selects questions during authentication. Our study uses multiple choice questions for ease of generation and to allow statistically analysis; while these can clearly be guessed and are not suitable for real-world authentication, we are at present only attempting to assess the feasibility of such a mechanism in terms of statistical analysis of results and user acceptability; a real implementation would clearly require questions that are harder to guess.


File Edit View Favorites Tools Help
Back
Search
Favorites
Media
Address H:\identity\paper\Questions.htm

Could you please answer these questions:

1. On 27th of June 2003 Friday, was "DMC" scheduled at 14:00-15:00?

○ YES ○ NO

2. On the 21st of May 2003 Wednesday, When was it scheduled to meet "Steven"?

○ 9:00 10:00 ○ 12:0013:00 ○ 14:00 15:00 ○ 15:00 16:00

3. Is Pete's birthday before Melanie's birthday?

○ YES ○ NO

3. In June 2003, when was it scheduled to go to "Professional Development Planning" On?

○ Mondays ○ Tuesdays ○ Wednesdays ○ Thursdays

Figure (2) Question form generated from a calendar

3. PILOT STUDY

In the pilot study we have run two experiments. The data, in both experiments, included the electronic calendar data of staff in a UK university department. The electronic calendars were either in PalmOS or WindowsCE format.

3.1. Problem Statement

To assess the possibility of automatically generating questions based on recorded electronic history that a user can answer better than others who know the user well.

3.1.1. First Experiment

This experiment was designed as a low-cost feasibility test of the concept, and was not designed to extract meaningful statistics. Our sample size was six calendars, chosen from a set of colleagues well known to each other; and we generated five different true/false questions. Each subject was asked to attempt to answer all 30 questions. Events were selected randomly from all events entered in the PDA diary over the fortnight immediately preceding the test date; for each event, a true/false correct result was selected randomly, and then a question was generated according to a simple set of rules based on the time, day and date upon which it occurred. There was no attempt to apply any "intelligence" to understanding the nature of each diary entry

We show sensitivity/specificity analysis of the results. Briefly, sensitivity is the probability measure of true positive, while specificity measures true negative.

Sensitivity is the probability that a test (or symptom, or sign) is positive given the disease/condition being present. It is the true positive. Sensitivity = TP/TP+FN.

Specificity the probability that a test (or symptom, or sign) is negative or absent given that the disease/condition is not present. It is the true negative. Specificity= TN/TN+FP [5]

Thus a result of (1,1) implies perfect authentication, while for example (1,0) implies a test that everyone can pass. For single true/false questions, the best possible outcome is (1,0.5). see table (1).

Table (1) Sensitivity and Specificity

	Genuine	***Impostor***
Correct	Sensitivity = TP/TP+FN TP (True Positive)	FP (False Positive)
Wrong	FN (False Negative)	TN (True Negative) TN/TN+FP
Total	TP+FN =1	TN+FP =1

3.1.2. Results

We expected users to remember their calendar data with a high sensitivity. However the result obtained was 0.53 with a range of 0.8 i.e. (±0.40). Interestingly, the high range is caused by highly variable performance among subjects, with between one and five questions being answered correctly.

The specificity is consistent with our expectations of a random distribution, giving us some confidence that the questions were not answerable by impostors. However the fact that not everyone was able to answer their own questions better than randomly caused us to examine more deeply the patterns of questions that were and were not correctly answered with respect to the literature governing human memory performance.

Table (2) Sensitivity and Specificity (First Experiment)

	Genuine	*Impostor*
Correct	0.53 ± 0.40	0.5
Wrong	0.47	0.5 ± 0.20

These results were unexpected; we had expected a far better recall of recent calendar events. Intuitive analysis of the individual questions however suggested the marked differences among individual performance could have been caused not by their memory per se, but by different behaviour patterns as to the type of event entered in their calendars. How does the human remember and what are parameters that can affect the human memory? What are the types of questions that are difficult for others to answer?

3.2. Second Experiment

3.2.1. OVERVIEW

The human memory in psychology is classified as follow: sensor memory, working memory, and long-term memory. Long-term memory is divided into episodic, procedural and semantic memory.

In our research, we have focused more on the long-term memory and in particular the episodic memory which some researchers define it as autobiographical memory [1]. Human being can remember: recent events, repetitive events, and pleasant events better than any other events [4][10]. In this experiment, the questions are generated based on these memorable events. The calendar data are classified as recent, repetitive or pleasant events as follows. Matching the date of a calendar event against the current date classifies these events as recent events if it is within a month time. If the event is repeated once or twice within a month time, it is classified as repetitive event. At last, matching calendar events against a list of words e.g. birthdays, parties, and concerts etc. classifies events as pleasant events if these words are found in the calendar event.

The other parameters, which were observed while the participants were answering the questions, were “Difficult” and “Easy”. As mentioned earlier, in the first experiment, the questions are “True” and “False” questions. Participants preferred to choose “True” as an answer to the questions. They commented on the question, which its correct answer is “True” as an “Easy” question. Conversely, they commented on the question that its correct answer is “false” as a “difficult” question.

3.2.2. Experiment Design

The Sample size was nine calendars, from each of which we generated eight questions. Their types were six (true / false) using the parameters (recent, repetitive, and pleasant) in each type there was a (true / false) question, and two four-part multiple choice. This choice was driven by again keeping the experimental cost low, but attempting to show in a statistically significant manner than individuals could answer questions better than impostors; only a weak result, but necessary given the outcome of the first experiment.

3.2.3. Sensitivity and Specificity Results (Second Experiment) All Questions

In this experiment, we have done two sensitivity and specificity analysis, one for all questions types and another for the multiple-choice questions type. The results of the all types of questions came as follows: The sensitivity is 0.71 ±0.19 while the specificity is 0.57 ±0.10.

This means that the person can answer his calendar correctly if we consider the memorable events and others can’t recall or guess the answers correctly. See Table (3).

Table (3) Sensitivity and Specificity Second Experiment

Answer	*Genuine*	*Impostor*
Correct	0.71 ±0.19	0.45
Wrong	0.29	0.43 ±0.10
Total	1	1

The sensitivity of the multiple-choice questions is 0.75 ±0.25, and the specificity is 0.78 ±0.18. Although over very small numbers of tests, these results are already significant, and are included mainly to show the potential of generating much less guessable questions to maximise the specificity whilst maintaining the sensitivity.

Table (4) Sensitivity and Specificity (Multiple-Choice Questions)

answer	***Genuine***	***Impostor***
Correct	0.75 ±0.25	0.78
Wrong	0.25	0.22 ±0.18
Total	1	1

3.2.4. ROC GRAPH

"Receiver Operating Characteristic Analysis" ROC illustrates relations between the sensitivity and specificity of each question. The diagonal line on the graph from (0,0) in the lower left hand corner to (1,1) in the upper right hand corner reflects the characteristics of a test with no discriminating power. The closer the graph gets to the upper left hand corner (0, 1), the better the test is at discriminating between cases and non-cases [11]. All of the questions are above the curve except for question four which is type recent difficult. The curve is slightly under the accepted area. See Figure (4).

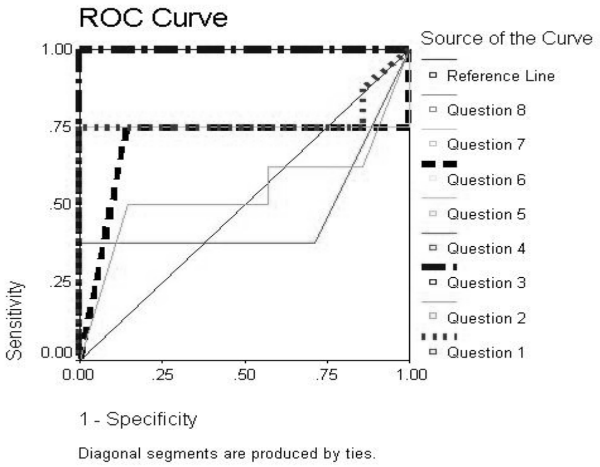


Figure (4) ROC Curve

Table (5) presents the results of the "Area under the Curve" AUC that varies from 1, perfect accurate, to 0.5. Almost all the questions are above 0.5[11], which mean the idea is feasible.

Table (5) Area Under the Curve (AUC)

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
0.777	0.545	1	0.464	0.75	0.696	0.75	0.75

Q1 Pleasant Easy, Q2 Pleasant Difficult, Q3 Recent Easy, Q4 Recent Difficult, Q5 Repeat Easy, Q6 Repeated Difficult, Q7 Multiple-Choice, Q8 Multiple Choice

4. CONCLUSIONS AND FURTHER WORK

Question based model of authentication presented a solution to some of the problems in identifying users over the Internet. The pilot study determined the feasibility of using personal electronic data to authenticate users. In the second experiment, because the types of questions were selected based on the memorable parameters, sensitivity was high enough to demonstrate the feasibility of the approach. The recent, repetitive, pleasant question types are better remembered and these types need further investigations in a bigger experiment. Additionally, using multiple choice improved specificity without impacting sensitivity. The ranking scale of correctness acknowledged the two factors, which are easy and difficult. The easy questions are on the top of the scale and difficult questions are on the bottom of the scale. This goes for both the genuine and impostor answers.

In the future, more investigation is required to gain more confidence in the results. This pilot study used only one format of the electronic data. The research can go further and use other electronic data such as data stored on mobile phone, GPS, PC, government or organizations database and, in the future with the smart environment application, there will a huge amount of stored electronic personal data. This large bulk of information can provide better security and, at the same time, will provide users with mobility because it is memorable.

4. REFERENCE

1. Baddeley, A. *Human Memory Theory and Practice.* New York Psychology Press, IN, NY, 1997.
2. Brostoff, S., Sasse, M. A., Are Passfaces More Usable than Passwords? A Field Trial Investigation', IN *Proc. HCI, Springer* September (2000), 405-424.
3. Cartwright, K. Information Sources and Metrics. *Conference and Expos/ trainings*. The KBA Symposium, Washington, IN, USA (2004)

4. Conway, M., *Autobiographical Memory: An introduction* Open Press University, IN Philadelphia,1990

5. Daly, L. Bourke, E., Geoffrey J. and Mc Gilvray, J. *Interpretation and Uses of Medical Statistics*, Oxford, IN, UK, 1991

6. Dhamija, R., and Perrig, A. Déjà Vu: A user study using images for authentication, *9th Usenix Security Symposium*, August, 2000.

7. Jianxin, Y., Blackwell, A., Anderson, R., Grant, Password memorability and security empirical results Cambridge, *IEEE Security & Privacy* September, (2004)., 26-31.

8. Mitchell, C.J. *Security for Mobility*. The Institution of Electrical Engineers, IN, London, 2004.

9. Stajano, F. *Security for Ubiquitous Computing*. WILEY: University of Cambridge, IN, UK, 2002.

10. Wagenaar W. My memory: A study of autobiographical memory over six years. *Cognitive Psychology, 18,* (1986)., 225-252

11. Witten, I., Frank, E., *Data Mining Practice Machine Learning Tools and Techniques with Java Implementations.* Morgan Kaufmann Publisher, IN London, 2000

12. Zviran, M., and Haga, W. Cognitive passwords: The key to easy access control. *Computer Security*, 9, (1990), 723-736.