

Question-Based Group Authentication

**Ann Nosseir &
Richard Connor**

Department of Computer and Information
Sciences, University of Strathclyde
Livingstone Tower, 26 Richmond St.,
Glasgow, G1 1EB, UK
ann.nosseir@cis.strath.ac.uk

Karen Renaud

Department of Computing Science,
University of Glasgow
17 Lilybank Gardens, Glasgow,
G12 8RZ, UK
karen@dcs.gla.ac.uk

ABSTRACT

There are various situations where a distinction needs to be made between group members and outsiders. For example, to protect students in chat groups from unpleasant incidents caused by intruders; or to provide access to common domains such as computer labs. In some of these situations the implications of unauthorized access are negligible. Thus, using an expensive authentication technique, in terms of equipment and maintenance, or requiring significant effort from the user, is wasteful and unjustified. Passwords are the cheapest access control mechanism but have memorability issues. As a result, various alternatives have been proposed. These solutions are often either insecure or expensive in terms of data collection and maintenance. In this paper we present a solution that is less costly since it is built on the data produced by user-system interactions. The mechanism relies on a dynamic (and unpredictable) shared secret. We report on our investigation into differentiating between group members and outsiders by means of their *group* characteristics. We also present an original analytical framework to facilitate the automatic generation of questions from group characteristics. Finally, we introduce a prototype of the mechanism.

Author Keywords

Human factors, Security, Web-based services, Social and Behavioral Sciences, Computers and Society

ACM Classification Keywords

H.1.2 Human factors, H.2.0 Security, H.3.5 Web-based services, J.4 Social and Behavioral Sciences, K.4 Computers and Society, Miscellaneous

MOTIVATION

Various settings require us to distinguish between group members and outsiders. For instance, there are plenty of

examples of unpleasant incidents on the net, where a stranger pretends to be a friend, or a group member, and reaches out to a child without the parents' knowledge, and causes trouble (BBC NEWS September, 24th 2003, 06:54 GMT).

Users often need to build trust relationships in order to conduct business over the web. Distinguishing group members from outsiders can reduce the risk associated with doing business to an acceptable level in some cases. Basu and Muylle (Basu and Muylle, 2003) outlined situations where affiliation with a group (for example, a corporate purchasing unit) may be more important than the individual's identity. They argue that this leads to a fourth category, on top of the three traditional authentication categories, of (1) *something the buyer knows*; (2) *something the buyer has*; and (3) *something the buyer is*, (Adams and Sasse, 1999; Adams et al., 1997; Anderson, 2001; Gollman, 2003) which is (4) *something the buyer does*. In this category users are authenticated based on their group characteristics, such as a group user-ID or the firm's digital certificate.

Group access is even more crucial in the medical context. There are certain domains that are accessible only to doctors and nurses but not to other staff or patients.

For the above-mentioned examples, the consequences of an outsider intruding are dangerous.

In other contexts the implications of unauthorized access are quite small. For instance, there are groups which are formulated in a physical environment, such as classmates, research groups or social groups. These groups establish mailing lists to publish upcoming events and group web pages providing classified information. Often group members meet via chat services to discuss issues or merely to socialize. The group members often wish to restrict access to bona fide group members.

However, outsider access will cause almost negligible damage. Hence, employing a heavyweight (in terms of memory or required hardware input device) access control mechanism technique is untenable. The traditional mechanisms are either expensive or not mobile. Biometrics and token-based authentication are costly in terms of installation, equipment and maintenance. Furthermore, tokens can be stolen or mislaid. Knowledge-based mechanisms such as passwords are the cheapest but they have memorability problems (Jianxin et al., 2004).

OzCHI'06, November 22-24, 2006, Sydney, Australia.

Copyright the author(s) and CHISIG

Additional copies are available at the ACM Digital Library (<http://portal.acm.org/dl.cfm>) or ordered from the CHISIG secretary (secretary@chisig.org)

OZCHI 2006 Proceedings ISBN: x-xxxxx-xxx-x

Users can not remember a random set of alphanumeric strings and this leads to insecure behaviour such as writing down passwords.

Researchers have proposed cognitive passwords as an alternative (Zviran and Haga, 1990). They use cue questions when the initial password is forgotten. They are more memorable and provide users with better mobility but they have the associated high cost of enrolment data collections and are more predictable than random passwords. In previous work, we tested a cognitive question-answer model using electronic personal histories in order to alleviate the costly data collection at enrolment. The studies showed that electronic personal history had potential for authentication purposes (Nosseir et al., 2005).

In this paper, we examine the extension of the traditional question-based model to determine whether it can be used to discriminate between legitimate group members and impostors by using the group rather than personal knowledge. In other words, our scheme authenticates users using *tacit* rather than factual knowledge. Tacit knowledge is difficult to record but easy to recall. Members of a group have this kind of knowledge about their group's characteristics and behaviour because of their shared history but outsiders simply cannot easily obtain the knowledge without being a member of the group for a certain time period.

This scheme is unlike other cognitive passwords since the expensive data collection step is unnecessary: group data is already available since it can be derived from the history of the group. The aim of this paper is to investigate two issues: Firstly; to explore the possibility of automatically extracting group characteristics from the shared group information; and secondly to determine the possibility of using this derived data to effectively differentiate between group members and outsiders. We examine these issues with both a physical *and* a virtual group.

GROUP CHARACTERISTICS

People are motivated to join groups to promote a positive self-image since the maintenance of relationships is a basic human need (Huitt, 2004). Tajfel and Turner (Tajfel, 1978; Thorne, 2003; Turner, 1985) argue that this self-image is made up of both personal and social identity. People appear to think of the particular group membership as being preferable to any alternative out-group. This social identity affects individuals' beliefs, behavior, values, and attitudes toward their in-group. In order to enhance this "notion of positive self-image distinctiveness", (Deaux, 1996; Hogg and Terry, 2000) group members emphasize real and imagined differences so as to highlight their group distinctiveness and thus preserve the uniqueness of their own social identity (Hogg and Terry, 2000; Streeter and Gillespie, 1992; Thorne, 2003). Accordingly, groups differ and build up their boundaries from out-groups by developing their own distinctive group characteristics (Tajfel, 1978; Thorne,

2003; Turner, 1985; Deaux, 1996; Hogg and Terry, 2000; Streeter and Gillespie, 1992).

FIRST EXPERIMENT

Problem Statement

In this pilot experiment, we investigated the possibility of differentiating between group members and outsiders based on group characteristic questions. We made two assumptions. Firstly, we assumed that members of the group should be able to identify their group knowledge whereas an outsider – anybody outside this group – would find it hard to identify group knowledge. Secondly, we assume that this group knowledge accurately reflects group characteristics.

Experiment Design

Two groups participated in this experiment. Each group had 40 participants: 36 students and 4 of their lecturers and tutors. We asked each group to assume a scenario in a chat room where more than one student, lecturer or tutor can sign in. We asked students to formulate questions related to their group activities, events or any other common knowledge or information they share which could be used to distinguish between themselves and intruders. In each group, we split students up into six sub-groups (see figure 1). We wanted to identify particular questions that any bona-fide member of the group could easily answer but that would be impossible for outsiders to answer correctly. A student would come up with a question, and then ask one of their lecturers or tutors the question; finally they asked another group the same question. We followed this procedure starting from group (A) and ending at group (F). We examined 12 questions.

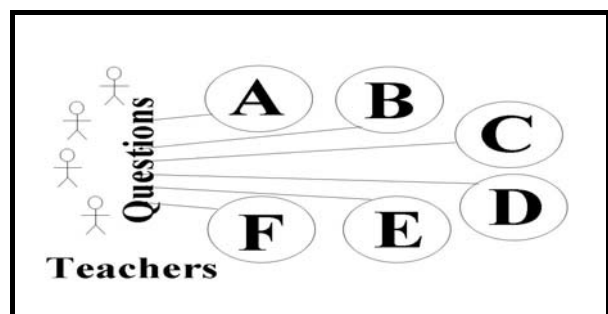


Figure 1. Typical group setting

Results

The results were the same for all groups. Students were able to answer all questions correctly. Lecturers and tutors were able to answer only 25% of the questions correctly, (3 questions out of 12). When we tested the first group's questions with the second group we duplicated the results. We used *sensitivity* (probability measure of true positives) and *specificity* (true negatives) analysis to classify these results.

Results imply almost perfect authentication (1, 0.75); there is evidence that group members recall or recognize group knowledge infallibly but outsiders don't.

	Group Members	Outsider
Correct	Sensitivity =1	0.25
Wrong	0	Specificity =0.75

Table 1. Sensitivity and specificity

Analysis

We will now present an analysis of the questions by mapping them to known group characteristics. Firstly, question topics reflected group *solidity*; this is evidenced by the fact that there are more than ten questions about a recent successful hockey competition. Shared information is also evident. A particularly *active member* was easily identified, with questions such as “what is Jane’s nickname” or “which pet does she have?” They also share common *cues or signals*. They have a hand drawn doodle depicting a lecturer with a moustache. The group *shares activities*. Questions depicted trips and birthday parties. They also have shared *norms, ideas and opinions*; for instance, they have formulated common preferences for a list items such as magazines, night clubs, shops, singers, drama series, TV shows and films. With these results in hand we moved forward to investigate the viability of using this technique with virtual groups

SECOND EXPERIMENT (PILOT STUDY)

Virtual Groups

People participate in online groups for work, education and leisure. Virtual group members experience the same sense of community as face-to-face communities. Blanchard and Markus (Blanchard and Markus, 2004) carried out empirical research on virtual communities and using the Sense Of Community (SOC) conceptual model. They found evidence of typical membership behaviour; member influence, in terms of enforcing and challenging norms; mutual support among members and shared emotional connections (Blanchard, 2004).

This research suggests that, in many ways, online groups are as “real” as any other groups. Zorn (Zorn, 2005) argues that the computer mediated communication (CMC) group interactions are still different from to face-to-face interaction, however, because members use a distinctly different context and medium. Interaction is purely via written communication and this precludes other cues such as gestures, tone of voice, etc. Furthermore written communication demands more time and effort from the communicator which may well cause him/her to more concise.

This is probably why virtual group members have developed their own vocabularies and abbreviations, by using short phrases. For example: ASAP (As Soon As Possible), CUTL8R (See You Later), A/S/L? (What is your Age, Sex and Location?). There is a whole glossary of abbreviations¹. They also use group-specific graphical

¹abbreviations.virtualsplat.com/category/chat-abbreviation.asp

conventions such as emoticons and abbreviations to convey emotions (Turner, 1985).

The medium lacks the immediate feedback that characterizes synchronous communication. This means that context cannot be assumed and the communicator needs to be more precise about their text than they would be when speaking.

Finally, the medium exhibits distinctly different types of behaviors. Burnett (Burnett, 2000) characterizes *interactive behaviors* and *non-interactive behaviours*. The former imply active participation demonstrated by means of written messages and responses. *Non-interactive behaviors*, commonly referred to as “lurking”, implies passive participation as a reader.

In designing and implementing a group authentication mechanism we need to ensure that all kinds of group members involved in the different categories of groups are catered for. The following section describes an experiment carried out with non-virtual groups in order to determine the viability of such a mechanism.

Experiment Design

In this experiment, a virtual group with a group mail list was studied in order to determine differences between group members and outsiders. We sent an email asking participants to send us a list of questions and answers about the group. Six members actively participated and sent us about ten questions each. We proceeded to test these questions. Each member answered all other members’ questions.

Results

All participants answered all questions correctly. We then asked them to identify questions that their family members or close friends would be able to answer. Three participants identified one or two questions their family members would be likely to answer correctly. This means others can answer or guess only 20% correctly. This result is similar to the previous study. It is almost perfect authentication (1, 0.80) using the Sensitivity and Specificity analysis.

Analysis

Mapping the questions with the group characteristics, we found a relationship between the question topics and the group characteristics. First, *group dynamic and leadership* is clearly reflected in the question topics. For example, all participants have listed at least one question about the moderators, also about *active members’ activities and roles* e.g., who organized the York trip? Secondly, the group recognized the importance of *group objectives*. A few questions were explicitly asking about the group objectives and activities. This also demonstrates cohesiveness. Thirdly, the *group values and norms* were clear to all members; they formulated a question about the group values? Fourthly, there are at least 3 questions about the *group activities*. For example,

which trip had the biggest number of non-participants? Or, what was the major event organized by the group in London? Fifthly, *attitudes and behavior* were clear in questions. All of them had at least two questions about the funniest person or the most interesting events or trips or preferred restaurants. In summary, both experiments demonstrated the evidence of the feasibility of distinguishing between group members and outsiders based on questions generated by the group members. In the following section, we present the group question-based model and a framework to automate this idea.

system sends him a One Time Password (OTP) by email and (9) he gets access by the OTP. (10) If the client decides to resign or the group master decides to expel him, the system deactivates his username. This mechanism enables new users to access the group website until they contribute and know the group history. Additionally, it enables group master to expel resigned members. See figure 2

Automating these questions requires identifying and analyzing stored group-related data. Virtual communities communicating via text communications such as chat or email lists generate communications patterns and text

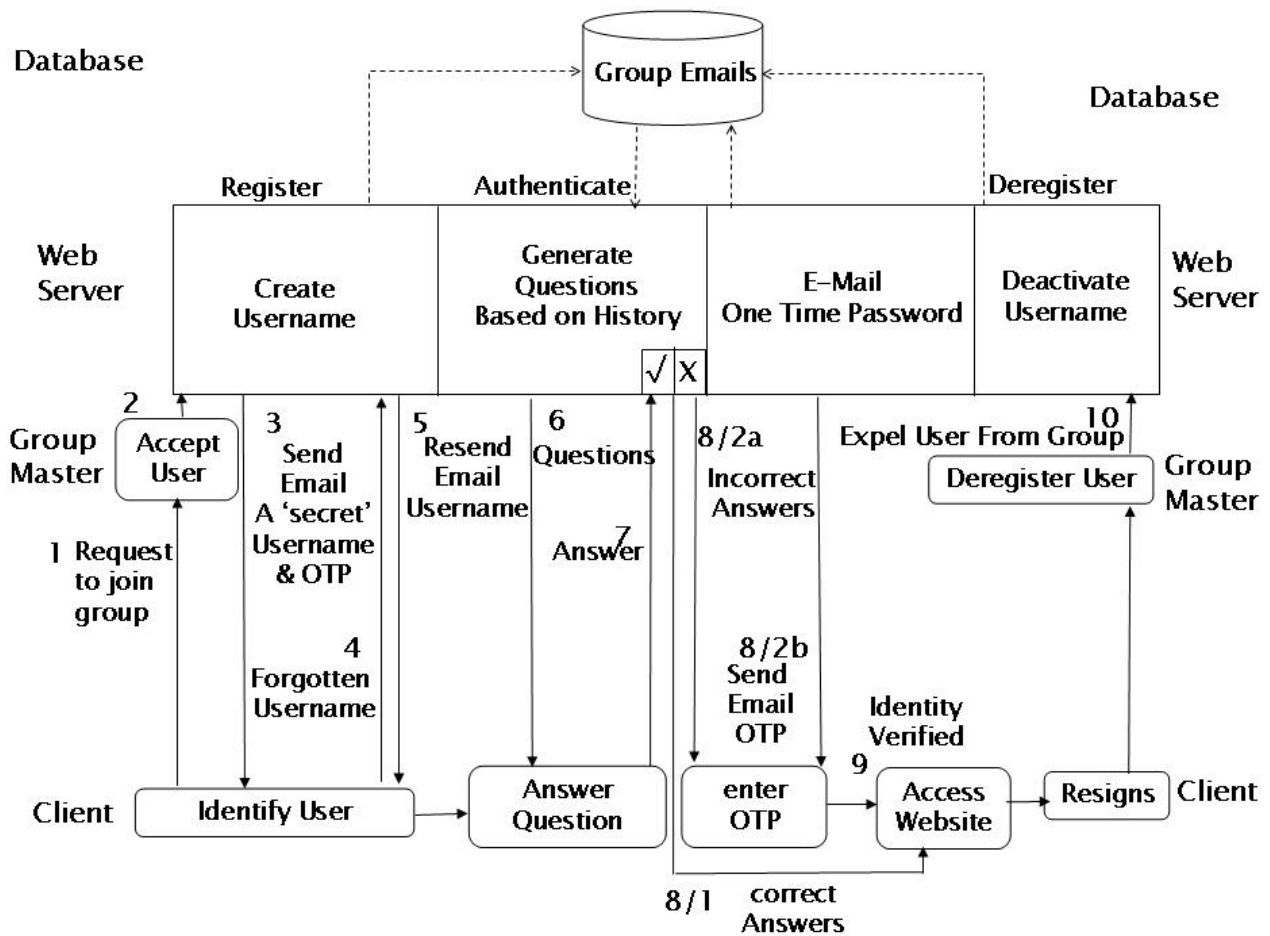


Figure 2. Group-question based model

GROUP QUESTION- BASED MODEL

Our Group-Question Based Model is a web-based mechanism that identifies group members from outsiders through the following steps: Initially, (1) a client requests joining the group and (2) if the group master accepts this request, (3) the system creates a username and sends it to the client. (4) If the client forgets his username; he can send an email to the system and (5) the system sends him back his username. (6) To access the group web site, the system sends questions generated from the group email database, next (7) the client answers the questions. (8/1) If the answer is correct, (9) he accesses the group website; otherwise (8/2a) if the answer is incorrect, (8/2b) the

content that can be mined. To analyze this information, and extract authentication questions, we can use a variety of techniques.

Streeter and Gillespie (Streeter and Gillespie, 1992) have developed a technique called SNA, which analyses the relationships between actors in a network. This technique can be used to understand how virtual communities communicate. The technique views social relationships in terms of *nodes* and *ties* where nodes are the individual actors within the networks, and ties are the relationships

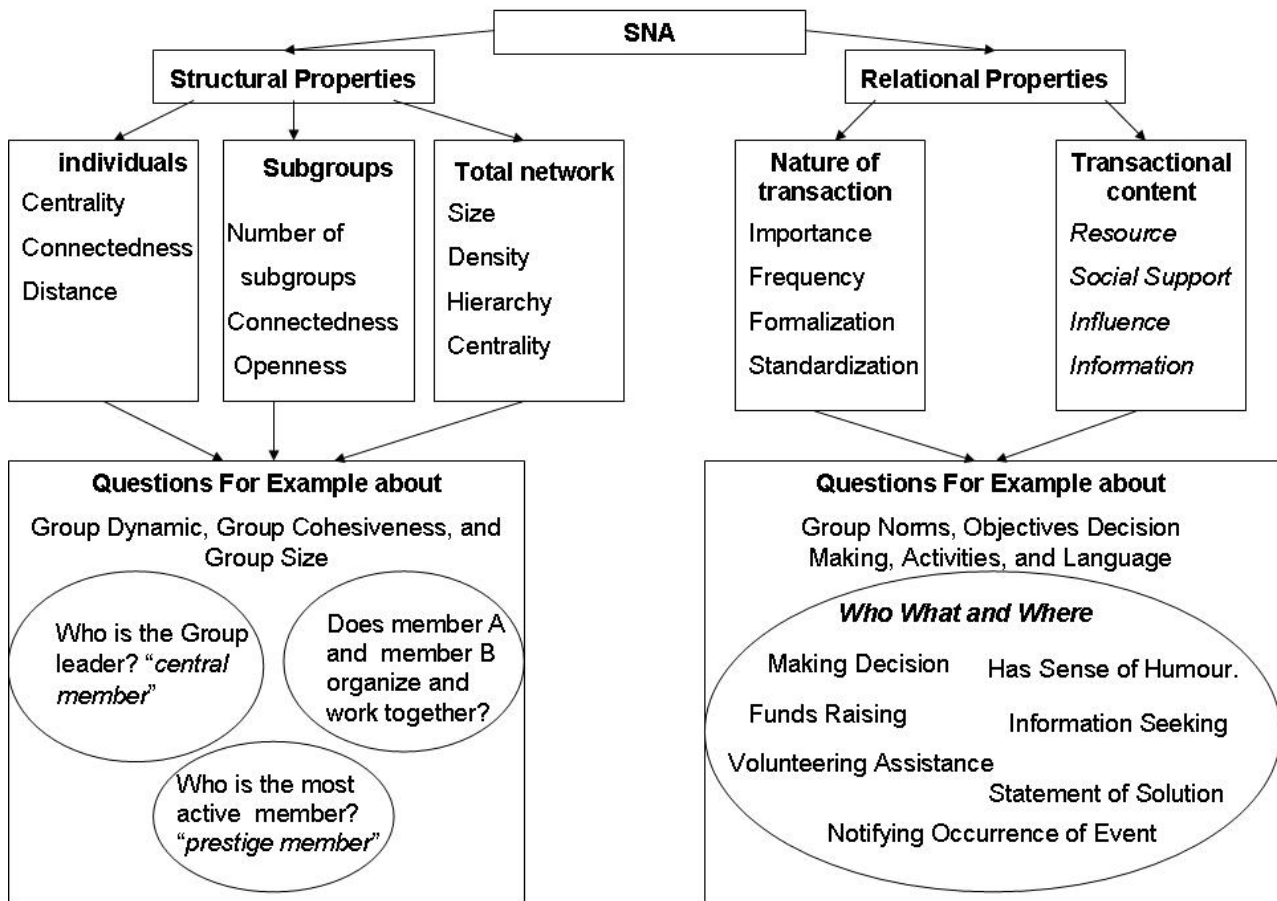


Figure 3. Analytical framework

between the actors. SNA maps and measures relationships and information flows (Hanneman and Riddle, 2005; Krebs, 2002; Walther, 2005; Zorn, 2005); e.g. who is the group leader? Or what was the last event? SNA is in some sense a network x-ray. SNA does this by measuring structural and relational properties of a network.

Structural properties describe the ways in which people form social networks and seeks to identify bona fide members of the network, and particular subgroups. Qureshi (Qureshi, 1995) argues that SNA makes it possible to identify central or prestigious actors in a network as well. A *central actor* will typically have extensive involvement in many relationships. *Prestigious actors* have extensive involvements in relations directed to them and are tend to influence the behavior of others in the network. The identification of these actors is helpful in terms of group authentication because it is often only group members who can infallibly identify such actors.

Relational properties refer to members' relationships with each other. This is analyzed by considering flows and exchanges between actors. These might be purely informational or offering social support or exerting influence. Using these exchanges, Qureshi (Qureshi, 1995) exploited *text content* information to determine the initiator and passive participant in conversations. Boucauvalas (Boucauvalas, 2003) also used *text content*

information to develop an emotion extraction engine based on word tagging and analysis of sentences.

Using these properties, the system can generate various types of questions asking about group norms, objectives, decision making, activities, and language. For example, it could generate a questions based on who made a particular decision, what the decision was and who the detractors were.

It clear that to investigate electronic groups, for instance, we can use a combination of structural theories, techniques, and social network analysis to build an understanding of a group and its characteristics. These characteristics enable the generation of authenticating questions. Figure 3 demonstrates a framework for achieving this.

IMPLEMENTATION

Based on this framework, we developed a prototype to demonstrate the feasibility of automating this scheme. Our system is a web group authentication question-based system. It poses three questions based on the group's history to authenticate users:

- (1) the first question is about the most active member,
- (2) the second is about the most recent topic discussed amongst the group,

(3) the third question asks users to identify a picture that was taken of a recent group activity.

To generate these questions, our prototype retrieves the recent few weeks' group emails. To simplify accessing the group emails, we cached them. In creating the first question, our prototype extracts all senders' names and checks the most frequent one and prints the question, by using a template "Who is the most active member" and sets the most frequent name with 4 random names (not necessarily of group members) in multiple choice options. Second, in generating the second question, the system extracts the most recent subject title. It uses a template to print the question "which message topic has recently been sent?" The system displays a multiple choice option of the most recent topic and other random topics. Third, the system retrieves one of the group images and uses other semantically-similar images as distractors. To generate the question, the system uses a template "Which is the group picture?" and presents a set of images: one of which is the group picture. Finally, the system directs users to the group web page only if all the correct answers are identified.

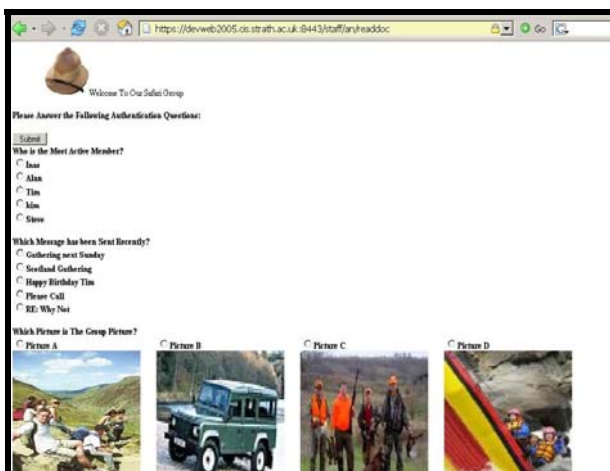


Figure 4. Our prototype interface

DISCUSSION

In this study, we investigated the idea of differentiating between group members and outsiders by means of their tacit knowledge of group characteristics. We have come up with a scheme which appears to have merit in this context. There are two questions that have to be answered before such a scheme can be at all viable. 1) How are new members authenticated? 2) How do we ensure that ex-members can no longer gain access to the system?

It is possible to make use of one-time passwords for new members. The member can keep requesting one-time passwords from the system until he/she is ready to move to the group authentication mechanism – once some knowledge of the group has been assimilated. The way to deal with ex-members is to make use of a secret username identifier. The user needs to use this to gain access to the

system, but it is never used apart from that and therefore remains secret. All one then has to do to expel members is to invalidate their username. The system will then no longer grant access even if the ex-member can answer authenticating questions.

CONCLUSION

In this study, we investigated the idea of differentiating between group members and outsiders by means of their tacit knowledge of group characteristics. We introduce this mobile and less expensive scheme for low risk situations and it can be used with other techniques as a second layer of authentication. In the future, more investigation is required to gain more confidence in the results with virtual groups using the suggested framework. Additionally, we also need to investigate other parameters that affect the authentication process. Finally, we need to explore various security issues to prevent hackers accessing any group information sources. This is to assure effectiveness of the authentication mechanism. Moreover, we need to seriously consider all privacy issues while running further experiments and avoid these situations like the current G-mail privacy problems² by requesting permission before storing or using emails and chat texts to derive authentication questions.

REFERENCES

- Adams, A., and Sasse, 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.
- Adams, A., Sasse, A., and Lunt, P. Making passwords secure and usable. In H. Thimbleby, B. O’Conaill, and P. Thomas, editors, *People & Computers XII, Proceedings of HCI’97(1997)*, pages 1–19.
- Anderson, R. *Security Engineering*, Wiley, N.Y. 2001,
- Anderson, R.E. Social impacts of computing: Codes of professional ethics. *Social Science Computing Review* 10, 2 (1992), 453-469.
- Basu, A. and Muylle, S. Authentication in e-commerce. *Commun. ACM Press* 46(12) (2003): 159-166.
- BBC NEWS September, 24th 2003, 06:54 GMT Document retrieved from Internet. Available at <http://news.bbc.co.uk/1/low/technology/3134342.stm>
- Blanchard, A., and Markus, M. The Experienced 'Sense' of a Virtual Community: Characteristics and Processes, *The Data Base for Advances in Information Systems*, 35, 1(2004.) 65-79.

² Google stores all e-mail even deleted e-mail for commercial purposes. This violates user privacy and lawmakers are issuing legislation to prevent that. http://www.usatoday.com/money/books/reviews/2005-11-13-google-book_x.htm

- BOUCOUVALAS, A. C., and Zhe, X. Text-to-Emotion Engine for Real Time Internet Communication. In *International Symposium on CSNDSP*, Staffordshire University, July 15-17, (2003) 164-168.
- Burnett, G. Information exchange in virtual communities: a typology, *Information Research*, Vol. 5 No. 4, July (2000) Available at: <http://informationr.net/ir/5-4/paper82.html>
- Deaux, K. "Social Identification," In Higgins, E. T., Kruglanski, A. W. (Eds.), *Social Psychology Handbook of Basic Principles*, New York and London: The Guilford Press, (1996) 777-798.
- Forsyth, D. R. *Group dynamics* (4th ed.). Pacific Grove, CA: Brooks/Cole. (2006).
- Gollman, D. *Computer Security*, John Wiley and Sons Ltd., N.Y. (2003).
- Hanneman R. and Riddle, M. *Introduction to social network methods*. (free introductory textbook on social network analysis). (2005), Available at <http://faculty.ucr.edu/hanneman/nettext>
- Hogg, M.A. and Terry, D.J. "Social Identity and Self-Categorization Processes in Organizational Contexts," *Academy of Management Review*, Vol. 25,1, (2000). 121-140.
- Huitt, W., "Maslow's hierarchy of needs", Educational Psychology Interactive, (2004) <http://chiron.valdosta.edu/whuitt/col/regsys/maslow.html> (Accessed 9 Dec 2005).
- Jianxin, Y., Blackwell, A., & Anderson, R., Password memorability and security empirical results Cambridge, *IEEE Security & Privacy*, September (2004), 26-31.
- Klemmer, R.S., Thomsen, M., Phelps-Goodman, E., Lee, R. and Landay, J.A. Where do web sites come from? Capturing and interacting with design history. In Proc. CHI 2002, ACM Press (2002), 1-8.
- Krebs, V. organizational consultant and the developer of InFLOW software. (2002) Available at <http://www.orgnet.com/IHRIM.html>
- Mather, B.D. Making up titles for conference papers. Ext. Abstracts CHI 2000, ACM Press (2000), 1-2.
- Nosseir A., Connor R., & Dunlop M. Internet Authentication Based on Personal History – a Feasibility Test. In *Proceedings of ACM world wide conference a workshop (WWW 2005)* (Tokyo WWW 2005 May).
- Qureshi, S., Supporting Electronic Group Processes: a social perspective, in (Ed) L. Olfman, Supporting Teams, Groups, and Learning Inside and Outside the IS Function. *SIGCPR/ACM*, Nashville. 1995.
- Stajano, F. *Security for Ubiquitous Computing*. WILEY: University of Cambridge, IN, UK, 2002.
- Streeter, C. L. and Gillespie, D. F. Social Network Analysis. *Journal of Social Service Research* 16 (1/2), (1992):201-222.
- Tajfel, H. "Experiments in Inter-group Discrimination," *Scientific American*, 223, 5, (1970). 96-102.
- Tajfel, H. "Social Categorization, Social Identity and Social Comparison," In Tajfel, H. (Ed.), *Differentiation between Social Groups*, UK: *Academic Press*, (1978).61-76.
- Thorne, S. L. Artifacts and Cultures-of-Use in Intercultural Communication. *Language Learning & Technology* 7/2 (2003): 38-67.
- Turner, J.C. "Social Categorization and the Self-concept: A Social Cognitive Theory of Group Behavior," In Lawler, E. L. (Ed.), *Advances in Group Processes*, UK: *JAI Press Inc.*, (1985). 77-122.
- Walther, J. B, Gay, G., and Hancock, J. T. How Do Communication and Technology Researchers Study the Internet? *Journal of Communication* 55(3) (2005):632-657
- Zellweger, P.T., Bouvin, N.O., Jehøj, H., and Mackinlay, J.D. Fluid Annotations in an Open World. Proc. Hypertext 2001, ACM Press (2001), 9-18.
- Zorn, I. Do culture and technology interact? Overcoming technological barriers to intercultural communication in virtual communities. In: *SIGGROUP Bull.* 25(2), (2005): 8-13.
- Zviran, M., and Haga, W. Cognitive passwords: the key to easy access control, *Computer Security*, 9, (1990), pp.723-736.