Don't Listen! I Am Dictating My Password!

Shaojian Zhu, *Yao Ma, *Jinjuan Feng, and Andrew Sears

Department of Information Systems
UMBC
Baltimore, MD 21250
+1-410-455-3883
{szhu1, asears}@umbc.edu

*Computer and Information Sciences Department
Towson University
Towson, MD 21252
+1-410-704-3463
{yma1, jfeng}@towson.edu

ABSTRACT
Speech recognition is a promising alternative input technology for individuals with upper-body motor impairments that hinder the use of the standard keyboard and mouse. A recent long-term field study found that the users employed speech techniques for a variety of tasks beyond generating text documents [1]. One challenge with hands-free speech-based interactions is user authentication, which requires the users to speak their user IDs and passwords character by character. Unfortunately, speaking a password presents both security and privacy threats as well as usability problems. To address this challenge, we propose a new speech-based authentication model. An initial proof-of-concept prototype has been implemented and a pilot study was conducted. Preliminary results suggest several problems for further examination.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces, K.4.2 [**Computers and Society**]: Social Issues – *Assistive technologies for persons with disabilities.*

General Terms: Design, Human Factors

Keywords: Speech technology, physical impairment, authentication

1. INTRODUCTION
A significant number of people have physical impairments which cause the loss of specific motor functions and limit their use of hands, arms or other body parts [2]. Those impairments cause significant difficulties for these individuals when using computers via a traditional keyboard and mouse. Speech-recognition, which employs human speech as input, is a promising alternative which may help these individuals interact with computers [3].

In one of our early speech-recognition field studies, which investigated how users with no physical impairments and users with upper body physical impairments use speech technologies in home environments, we observed that individuals with physical impairments benefit from the speech-recognition technology and are more satisfied with speech-based systems as compared to the users without physical impairments [1]. However, we also found that security and privacy mechanisms are becoming a major obstacle for the users of hands-free speech applications. Nowadays authentication is an integral part of many applications and web systems. When people use speech as input for the authentication process, they often speak their user names and passwords character by character. The system captures the audio input and transforms it into textual information before the required information can be entered in the cursor-focused textbox (usually the password box which uses asterisks to hide the original plain-text password).

Several problems arise from this approach. First, speech recognition errors cause significant difficulties when entering passwords (e.g. Character 't' may be captured as word 'tea' or char 'p' because of pronunciation similarity). To make this problem even worse, most password input boxes use asterisks to hide characters for security reasons, so the user has no way to detect recognition errors as they occur. If the first character is incorrectly recognized, the user has to enter the entire password before getting any feedback. If a password is rejected by the system, the user would have no idea whether the rejection is due to entering the wrong password or one or more recognition errors. Another fundamental problem with this approach is that the authentication information can be overheard by any people nearby, making personal information of the user vulnerable.

In order to address these problems, especially the problem of someone overhearing the user as they dictate their passwords, while maintaining a speech-based authentication approach, we proposed an indirect speech-based authentication model. We have developed an early 'proof-of-concept' prototype and completed an initial user evaluation.

2. AN INDIRECT SPEECH-BASED AUTHENTICATION MODEL
The essential idea of the new approach is the adoption of a query-response, bidirectional interactive model instead of the traditional unidirectional approach. We take advantage of the computer-to-user channel to pose questions to the user. Then users need to respond to these questions by combining a system-provided detail with information contained in their secret password (a manual encryption process). There are multiple ways to combine the system facts with password information, with the initial 'proof-of-concept' prototype adopting arithmetic operations (e.g., addition) with the assumption that the user's password only contains numbers. In the next stage, we plan to extend the design to

ASSETS'09, October 25-28, 2009, Pittsburgh, Pennsylvania, USA.
ACM 978-1-60558-558-1/09/10.

accommodate passwords containing both numbers and letters. The basic process is illustrated in Fig. 1.

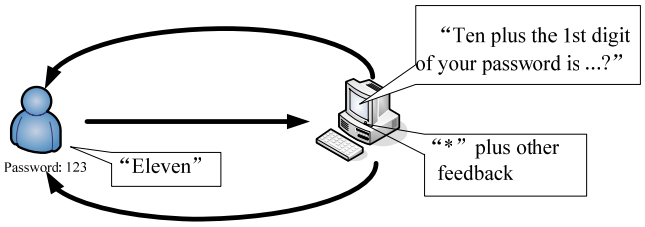

Figure 1. Illustration of the authentication process.

As illustrated in Fig. 1, the system asks the user to solve a simple addition problem based on the first digit of his password. Instead of saying the original character (one), the user speaks the encrypted password 'eleven'. Since the encryption process applied each time the password is entered, people overhearing the user cannot guess the original password. As long as people cannot hear the system prompt, the user's password is secure. In this scenario, the user would use earphones to hear the system's prompt so that the other people do not hear it.

Another design consideration is to provide sufficient feedback when spoken words are captured. By adding additional feedback (e.g., visual, audio or other modality) regarding the recognized words, instead of the default of displaying asterisks, it is easier to detect and recovery from both user errors (e.g., addition mistakes) and system errors (e.g., recognition errors). We designed the system so that every time the user answers a question based on one digit of the password, the system shows the recognized spoken number on the screen.

Most authentication systems only store encrypted passwords using one-way encryption methods (e.g. MD5), leaving no way to reveal the original plain-text passwords easily. Therefore, one limitation of our model is the assumption that our system knows the plain-text passwords. This limitation would be addressed, perhaps by having our system serve as an intermediary, much like password assistants build into many web browsers.

3. METHODOLOGY

We conducted a small pilot study to evaluate the proposed speech-based authentication method. The objective was to collect some initial data on how users interact with the authentication solution. In particular, we would like to examine whether the user can easily and accurately respond to the manual encryption questions. Ten participants with no cognitive or perceptual impairments took part in the study. Two participants are female, and all are native English speakers. The average age of the participants is 21.8 (stdev: 1.93). We use a Dell-GX270 with Windows Vista^{R} as the testing computer, and a Sony DR-115 high-quality headset for speech input and audio output.

During the study, participants first entered a 3-digit password that they defined. A longer password could be used for increased security, but three digits are sufficient for an initial proof of concept evaluation. The investigator demonstrated the use of the system and explained the task to the participants. All participants completed a training task with the assistance of the investigator. In order to enter the three-digit password, the participant needs to answer three questions posed by the system, one for each digit of the password. After the training session, participants completed ten authentication tasks. At the end of the study, participants responded to a questionnaire, providing demographic information and satisfaction ratings regarding the application.

4. EARLY RESULTS

Through our study observations and participant comments, we found that this authentication model can effectively address the problem of openly dictating passwords. Participants used earphones to receive the questions from the computer and spoke encrypted passwords such that observers received minimal information regarding their passwords.

However, the ease of use of this system needs further improvement. We found that speech recognition errors remain a challenge for users. Four participants complained that the recognition error rate was too high, and 50% of the participants were observed to experience frequent recognition errors when attempting to enter the number 'thirty'.

Although participants did not report too much difficulty answering the addition and subtraction questions, we observed that the arithmetic calculations placed additional cognitive demands on the user. During the authentication process, users must listen carefully to the system for the questions, retrieve the required password digit, and complete the associated arithmetic calculations, and speak their answer. The task may be demanding for older users or users with cognitive impairments.

While our pilot study only required a 3-digit numeric password, increased security is possible with longer passwords and a larger character set. Many systems require more complicated passwords, often involving letters, numbers, and perhaps even symbols. The proposed solution needs to be enhanced and expanded to address more realistic passwords.

5. ACKNOWLEDGMENTS

We would like to thank National Institute on Disabilities and Rehabilitation Research for funding this project through grant number H133G050354 and National Science Foundation through grant number NSF DUE 0830865.

6. REFERENCES

[1] Hu, R., Zhu, S., Feng, J., and Sears, A. 2009. Evolving requirements for speech applications: Lessons learned from a speech study. Technical paper.

[2] National Spinal Cord Injury Statistical Center. 2009. The 2009 Annual Statistical Report for the Model Spinal Cord Injury Care Systems, retrieved on June 1, 2009 at http://www.spinalcord.uab.edu/show.asp?durki=19775

[3] Sears, A., Young, M., and Feng, J. 2007. Physical Disabilities and Computing Technologies: An Analysis of Impairments. In J. Jacko, and A. Sears, (eds) The Human-Computer Interaction Handbook. CRC Press.