Sonifying Internet Security Threats

Akbar Siami Namin
Computer Science Department
Texas Tech University
Lubbock, TX 79409, USA
akbar.namin@ttu.edu

Rattikorn Hewett
Computer Science Department
Texas Tech University
Lubbock, TX 79409, USA
rattikorn.hewett@ttu.edu

Keith S. Jones
Psychology Department
Texas Tech University
Lubbock, TX 79409, USA
keith.s.jones@ttu.edu

Rona Pogrund
College of Education
Texas Tech University
Texas School for the Blind and Visually Impaired (TSBVI)
Austin, TX, 78756, USA
rona.pogrund@ttu.edu

CHI'16 Extended Abstracts, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-4082-3/16/05.
http://dx.doi.org/10.1145/2851581.2892363

Abstract
The Internet enables users to access vast resources, but it can also expose users to harmful cyber-attacks. It is imperative that users be informed about a security incident in a timely manner in order to make proper decisions. Visualization of security threats and warnings is one of the effective ways to inform users. However, visual cues are not always accessible to all users, and in particular, those with visual impairments.
This late-breaking-work paper hypothesizes that the use of proper sounds in conjunction with visual cues can better represent security alerts to all users. Toward our research goal to validate this hypothesis, we first describe a methodology, referred to as sonification, to effectively design and develop auditory cyber-security threat indicators to warn users about cyber-attacks. Next, we present a case study, along with the results, of various types of usability testing conducted on a number of Internet users who are visually impaired. The presented concept can be viewed as a general framework for the creation and evaluation of human factor interactions with sounds in a cyber-space domain. The paper concludes with a discussion of future steps to enhance this work.

Author Keywords
Sonification; Internet security threats; usability testing; visually impaired

ACM Classification Keywords
H.5.2. **[User Interfaces]**: User-centered design, user interface management systems

Introduction

The Internet plays a key role in today's life and day-to-day activities. While involvement of the Internet in daily activities improves the quality of life, it also makes users vulnerable to various cyber-attacks. For instance, during phishing attacks, Internet users often receive emails that appear to be sent from legitimate and reputable entities, which claim that the users' accounts will be suspended unless they click on the provided email link. Clicking on the link will take users to inappropriate or malicious Websites. Given the prevalence of and risks pertinent to cyber-attacks, it is imperative that users be informed when they are being attacked. To draw users' attention to potential security attacks immediately, technologies and visual cues can offer help in detecting and also determining the risks associated with cyber threats. It is preferable to obtain more assistive guidance through these visual cues in learning about applying response strategies to lessen the damage caused by these attacks. Although scientific evaluations of these security warnings exist (e.g., usability of security warnings [8], visual cues to prevent SSLtripping [9], bank customer perception [10]), human factors studies in this area are still rare. A potential downside of using only visual cues to inform users about cyber attacks is that they are not available to all users, in particular, those with visual impairments. As a complement to visual cues, the use of sounds to represent certain events and effectively alarm users may help better inform users about suspicious events occurring during navigation.

The goal of our research is to investigate (1) how practical it is to convey security alerts to the targeted victims through sonification, (2) what type of sonification, auditory icons or earcons, works better for addressing the needs of Internet users with special needs (e.g., assistive technologies for persons who are visually impaired), and (3) whether Internet users find these sounds to be accessible and easy to use. To address these challenges, this late-breaking-work paper presents a methodology for sonifying cyber-security threats to warn users about cyber-attacks. More specifically, one based on non-speech sounds, i.e., earcons, was chosen because screen readers for visually impaired Internet users predominantly output speech sounds. It was believed that speech-based threat indicators could further complicate the already complicated task of using screen readers [12, 13, 14]. For the purpose of our research, this late-breaking-work paper focuses on sonifying three security threats and cues: i) *phishing*, ii) *malvertising*, and iii) *form-filling or typing sensitive information* into a form.

We have conducted a series of in-depth usability tests on five Internet users who were visually impaired. The sample size of five subjects may seem small, but it has been accepted to be sufficiently large for human factor evaluation [5]. Furthermore, it serves our purpose to obtain initial findings without wasting efforts in a fruitless path. Our testing examined whether users with visual impairments a) could clearly hear the indicator when triggered, b) could identify what kind of cyber threat a given indicator was meant to convey, and c) reacted properly to the indicator. The results show that it is possible to develop sonified cyber-security threat indicators that users intuitively understand, even with minimal experience. The results suggest that sonified cyber-security threat indicators could be part of a solution to the problem of how to warn Internet users about cyber-security threats.

Designing Sound-Featured User Interfaces

Research indicates that individuals with disabilities use the Internet for communication, online banking and shopping, and entertainment [1]. Those who design effective computer-human interfaces to warn users, and in particular those with visual impairments, in the context of cyber-attacks, face several challenges. First, while the technique for sonifying warnings is well studied, most work deals with other application domains (e.g., medicine, transportation, disaster warnings). No study has specifically investigated sonifying warnings about cyber-attacks, which can be more abstract and take many forms, some of which are difficult to understand or differentiate the semantics

and the situations, and some may be unknown to novice users. The design must be able to convey the warning at different levels of danger to users with diverse backgrounds. Second, users with visual impairments often use screen readers, which auditorily present information that normally would be graphically displayed on a computer screen (e.g., conveying text via speech). Most often, users with visual impairments utilize standard applications, such as Microsoft Internet Explorer, in conjunction with a system-wide screen reader such as Job Access with Speech (JAWS), which speaks the elements of the computer's interface aloud. The screen reader user hears synthesized speech as she navigates a document or the Internet that reads what is on the computer screen. Human linear-like perception of sounds is limited compared to the nature of visual perceptions. Third, the visual cues that individuals without visual impairments rely on to detect cyber-attacks can be lost when screen readers translate the system's interface into a verbal description. Specifically, screen readers mainly verbalize content and ignore the interface's appearance. For instance, it would be difficult for users with visual impairments to know that those around them can see what they are typing into the system because the information is being typed into an insecure field [2].

Earcon-Based Sonification Approaches

Communication with sounds has been of great interest to the CHI community. Examples include gesture sonification to support reflective craft practice [16], the use of voice in Web browsers [19], and the use of sounds in mobile devices [21]. Furthermore, the CHI community has a long-time interest in designing accessible interfaces for people with disabilities [17, 18, 20], which makes this venue a fit for this work.

There are three basic approaches to designing non-speech sounds or earcons [6]: the *representational*, *abstract,* and *semi-abstract* approaches [3]. The representational approach uses natural sounds from events in one's environment to convey information. For example, to convey the threat of a security breach, one could play the sound of an old creaky door. The abstract approach is to use sounds that are synthesized from basic sound components and do not stem from events in one's environment to convey information. Combining pitches with simple rhythm and pitch design can be used to represent complex information. For example, to distinguish information related to security, privacy, and user-interface control, one could associate each type of information with a different instrument [11]. The semi-abstract approach is a mixture of the representational and abstract approaches. For example, a natural sound such as an old creaky door could be combined with a simple pitch, which could vary in frequency to denote threat severity.

User-Centric Information to Convey

It is necessary to determine what exactly the sonification should convey. We identify three approaches to convey information through sonification, namely, i) the *"concept"* or *"meaning"* of associated cyber-security threats. For example, a sonification should make users aware that they are experiencing a phishing attack, as opposed to a different type of attack; ii) the *"consequences"* of cyber-security threats for users. For example, a sonification should make users aware that they are experiencing an attack that could compromise the integrity of their sensitive data; in contrast to the first possibility, users should not be made aware of the specific type of attack that they are experiencing; and iii) the *"actions"* that users should take in response to the current cyber-security threat. For example, a sonification should be designed so as to discourage users from continuing their actions (e.g., by clicking on a link to a malicious file).

Study Method: The Creation of Sonification

The concepts described above constitute the steps for sonifying cyber threats. It consists of identifying: a) a set of security threats to sonify, b) a sonification approach, and c) sounds to present the intention and meaning of the cyber threat from users' perspective.

The selection of security threats is "user-centric" and mostly depends on the effects of the chosen threat on users. There are many different types of cyber-security threats [15]. The CIA (Confidentiality, Integrity, and Availability) classification scheme is a simple but informative way to organize them based on whether they threaten the **C**onfidentiality of information and users, tamper with the **I**ntegrity of information, or hinder the **A**vailability of services. The selection of sonification approaches largely depends on the underlying context. Several factors play important roles in selecting proper sonification approaches including: the easiness to remember, the types of audiences, the familiarity of sonifications, and the relevance to the events sonified. Similarly, the selection of sounds mainly depends on the type and severity of events being sonified [6]. For instance, an event with catastrophic consequences should be represented with louder and high frequency sounds, whereas, a minor event (cue) should be presented with a shorter sound with lower frequencies.

Case Study

This section presents a case study in which the presented sonification methodology was tested.

Participants

Five individuals who were visually impaired participated in the study. Such a sample size is sufficient to identify the majority of a system's usability problems [5]. Comparable sample sizes have also been employed in similar usability studies that involved users who were visually impaired [4]. The participants were three males and two females in the age range of 20–49. Four of the participants reported being currently employed. Of the five participants, one had a Master's degree, three had Bachelor's degrees, and one had a high school diploma. They rated their uses of screen readers from Good (2) to Very Good (3). The participants were randomly recruited from a pool of 20 users who previously completed a survey regarding Internet use, assistive device use, and cyber-security concerns [7]. Survey participants were recruited from workers and students at a special purpose school for students who are blind or visually impaired and from a state rehabilitation agency for individuals who are blind. All participants reported being blind, as opposed to having low vision. All reported using screen readers such as: JAWS, Window-Eyes, VoiceOver, NVDA, System Access, and Talkback for Android. For demographic and some other information, please refer to the survey report [7].

Choosing What to Convey

We chose to create sonifications that a) convey the concept or meaning of associated cyber-security threats, and b) convey the consequences of cyber-threats for users. It was believed that using multiple strategies would increase the likelihood of developing effective sonification.

Chosen Cyber Security Threats and Cues

We initially focused on commonly occurring threats with which most users would be familiar. Such familiarity was important to ensure the exact intention of usability testing and its effectiveness. In the end, two cyber-security threats were chosen for this study: phishing and malvertising. In the present research, phishing attacks have been defined as an attacker's attempt to trick users into giving the attacker private information (e.g., users' passwords). Similarly, malvertising (malware + advertising) refers to malware downloading attacks in which an attacker attempts to entice users to download files that contain programs that cause havoc on the user's computer. In addition, we wanted to include a cyber-security threat that involved an activity with which usability testing participants would be familiar but operated differently than phishing and malvertising. We selected entering sensitive information into an online form, hereafter referred to as form-filling, which has the potential to expose private information.

Application	Threat	Sonification	%Heard Alert in the Context	%Correctly Identified the Type of Attack	%Properly Responded to the Alert (Task Relevant)	%Correctly Ignored the Alert (Non-Task Relevant)
Tech News	Phishing	-Casting a fishing reel	100%	20%	80%	80%
Pet Shop	Malvertising	-Dropping a bomb	100%	60%	60%	80%
Online Banking	Form-Filling	-Typing on a keyboard	100%	60%	20%	80%

Table 1. The final set of sonifications[1].

Usability Testing Artifacts
Three Web applications were developed with an injected security threat along with the sounds representing the injected threat. A simple technology news Website injected with a phishing attack, an online pet shop store injected with a malvertising, and a simplified banking system annotated with a form-filling cue, were the three Web applications developed for this study. Because the JAWS screen reader was used to conduct the tests, these Web applications had to be accessible for users who were blind (e.g., each field had to be annotated with a description).

Tasks Given to Participants
When interacting with the technology news Web site and the online pet shop store, participants completed three tasks per site in a random order. One of the tasks required participants to tab to a link that was not associated with an alert. For example, on the technology news Web site, participants had to "find and open an article about the financial stability of Amazon's Cloud Computing Division". The target link was not associated with an alert and successful task completion did not require participants to tab past any links that triggered alerts. Such tasks are hereafter referred to as "no-alert tasks." Another type of the tasks required participants to tab to a link that was associated with an alert. For example, on the online pet shop site, participants were told "This site has a video on dog training. What kind of dog treats does the trainer in the video suggest using during training?" When participants tabbed to the video's link, an alert was triggered. Such tasks are hereafter referred to as "task-relevant alert tasks." The remaining task required participants to tab past a link that was associated with an alert in order to locate a desired link. For example, on the technology news Web site, participants had to "find and open an article about Smart Garbage." To do so, participants had to tab past a link that triggered an alert. Such tasks are hereafter referred to as "non-task-relevant alert tasks." When interacting with the online banking site, participants completed only a no-alert task and a task-relevant alert task. The structure of the banking site did not allow for the creation of a non-task-relevant alert task. No-alert tasks were included in order to test whether participants would notice the alerts, which would not have been feasible had all tasks triggered alerts. Non-task-relevant alert tasks were included in order to investigate how users would react to alerts that could be ignored, given that the users' task did not require them to select the malicious link.

Chosen Earcon-Based Sonification Approach
The *representational* sonification was chosen mainly for two reasons: First, we wanted to create sounds that could convey their *intended meanings* with *little-to-no* user training. Second, we hypothesized that, to reduce their cognitive loads, individuals with visual impairments heavily utilize natural sounds, which are known to them, to determine whether an external entity threatens them. Natural sounds and their intentions seem to be easier to remember by people whose ears are their major communication channels. Each project team member independently searched online sound repositories (e.g., www.sounddogs.com) for natural sounds that he or she thought could represent phishing, malvertising, and form-filling. We

[1]For sonification and sounds, please visit the following webpage: http://www.myweb.ttu.edu/asiamina/SonificationSounds.html

used two techniques. First, we selected sounds that played with words "*conceptually*" related to the name of the cyber-security threat (e.g., the sound of a fishing rod and reel to represent phishing). Second, we selected sounds that were associated with the consequences of a given cyber-security threat. In the end, approximately sixty percent more candidate sounds were chosen than needed. Candidate sounds were then compiled into a master list, and each team member selected and ranked three sounds that he or she considered to be the best potential sonification for each cyber-security threat.

The highly ranked sonifications for each threat was usability tested in order to investigate whether users would a) hear the indicator when triggered, b) identify what cyber-security threat a given indicator was meant to convey, and c) react appropriately to the indicator.

Results and Discussions

Table 1 describes the final set of sonifications along with some of the results obtained through the testing. The highlights of the findings are as follows:

- Participants consistently heard the alerts (Column 4). There were no problems with background noise masking the alerts.
- For "form filling" and "malware downloading," the sonification's intended meaning was sometimes, but not always, correctly identified (Column 5). This finding suggests that the current sonifications worked about as well as they could, but we could use better sonifications for those threats. For "phishing," the sonification's intended meaning was less correctly identified (Column 5). It seems that context made the intended meaning of that sonification less apparent.
- For "phishing" and "malvertising," users responded appropriately to the alerts when they were relevant to the users' current task (Column 6). The general sense, though, was that users just thought sonifications meant something bad was happening, so they should not click the link. They did not seem to have a clear sense of what was bad about what was happening. For "form filling," users frequently did not or would not respond appropriately to the alerts when they were relevant to the users' current task (Column 6). Two users did or would have just entered their Social Security Numbers without reacting to the alert at all. Two other users would have stopped what they were doing or left the page. These responses are interesting because "form filling" is a bit different than the other sonifications. Specifically, it is not meant to stop users from doing something; rather, it is meant to encourage them to be careful about what they are entering.
- Most often, users correctly ignored the alerts when they were not relevant to their current task (Column 7). A minority of the time, though, users let the non-task relevant alerts disrupt their activities. This seemed to stem from a general sense that a site was bad or risky if an alert was triggered.

Conclusions and Next Steps

This paper presented late-breaking-work research that explored whether sonified cyber-security threat indicators could be used to warn users with visual impairments about cyber-security attacks. The results of initial usability testing were promising. Specifically, the results suggested that it is possible to develop sonified cyber-security threat indicators that users intuitively understand. Future research should explore ways to optimize various facets of the sonification development process. For example, the process of finding and selecting candidate sonification was cumbersome; it would be advantageous to develop ways to automate, at least parts, of that process.

Acknowledgment

This work is supported by National Science Foundation under award number CNS-1347521. Thanks to John Rose, Miriam Armstrong, and Tim Salau for conducting the usability testing sessions.

References

1. Stacy M. Kelly, Karen E. Wolffe. 2012. Internet use by transition-aged youths with visual impairments in the United States: Assessing the impact of postsecondary predictors. Journal of Visual Impairment & Blindness, 106(10), 597-608.
2. Holman, Lazar, Feng. 2008. Investigating the Security-related Challenges of Blind Users on the Web. Designing Inclusive Futures, Springer.
3. Meera M. Blattner, Denise A. Sumikawa, Robert M. Greenberg. 1989. Earcons and icons: Their structure and common design principles. *Human–Computer Interaction*, 4(1): 11–44.
4. Elaine Gerber. 2002. Surfing by ear: Usability concerns of computer users who are blind or visually impaired. *Access World*, pages 38–43.
5. Robert A. Virzi. 1992. Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors*, 34:457–468.
6. Thomas Hermann, Andy Hunt, John G. Neuhoff. 2011. *The Sonification Handbook*. Logos Verlag.
7. Fethi Inan, Akbar Siami Namin, Keith S. Jones, Rona Pogrund, 2016, Perception of Cybersecurity Risks for Internet Users Who Are Visually Impaired, Journal of "Educational Technology and Society" (ISSN 1436-4522).
8. Serge Egelman, Lorrie F. Cranor, Jason Hong. 2008. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. *The SIGCHI Conference on Human Factors in Computing Systems*, pages 1065–1074.
9. Dongwan Shin, Rodrigo Lopes. 2011. An empirical study of visual security cues to prevent the SSLstripping attack. T*he 27th Annual Computer Security Applications Conference*, pages 287–296.
10. Egwali A. Oghenerukeybe. 2009. Customers' perception of security indicators in online banking sites in Nigeria. *Journal of Internet Banking and Commerce*, 14(1), 1–15.
11. Stephen Brewster, Vali-Pekka Raty, Atte Kortekangas. 1996. Earcons as a method of providing navigational cues in a menu hierarchy. In *Proceedings of HCI'96*.
12. Petrie, Fisher, ONeill, Fisher, Di Segni Y. 2001. Deliverable 2.1: Report on user requirements of mainstream readers and print disabled readers.
13. Romisa R. Ghahari, Mexhid Ferati, Tao Yang, Davide Bolchini. 2012. Back navigation shortcuts for screen reader users. In Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility, pages 1–8.
14. Jonathan Lazar, Allen Aaron, Jason Kleinman, Chris Malarkey. 2007. What frustrates screen reader users on the web: A study of 100 blind users. International Journal of human-computer interaction, 22(3): 247–269.
15. Internet Security Threat Report (ISTR 20). 2015 Volume 20, Symantec Annual Report.
16. Thomas Smith, Simon J. Bowen, Bettina Nissen, Jonathan Hook, Arno Verhoeven, John Bowers, Peter Wright, Patrick Olivier. 2015. Exploring Gesture Sonification to Support Reflective Craft Practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '15) 67-76.
17. Dhruv Jain, Leah Findlater, Jamie Gilkeson, Benjamin Holland, Ramani Duraiswami, Dmitry Zotkin, Christian Vogler, Jon E. Froehlich. 2015 Head-Mounted Display Visualizations to Support Sound Awareness for the Deaf and Hard of Hearing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '15). 241-250.
18. Jonathan Lazar, Jinjuan Feng, Tim Brooks, Gennra Melamed, Brian Wentz, Jon Holman, Abiodun Olalere, Nnanna Ekedebe. 2012. The SoundsRight

CAPTCHA: an improved approach to audio human interaction proofs for blind users. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '12) 2267-2276.

19. Shaojian Zhu, Daisuke Sato, Hironobu Takagi, Chieko Asakawa. 2010. Sasayaki: Augmented voice web browsing experience. 2769-2778, *ACM ASSETS'10*.
20. Wai-ling Ho-Ching, Jennifer Mankoff, James A. Landay. 2013. Can you see what I hear? the design and evaluation of a peripheral sound display for the deaf. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '03) 161-168.
21. Ellen Isaacs, Alan Walendowski, Dipti Ranganthan. 2012. Hubbub: a sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (CHI '02) 179-186.