Accessible *and* Secure?
Design Constraints on Image and Sound Based Passwords

Marcia Gibson
University of Bedfordshire
Park Square, Luton, UK
marcia.gibson@beds.ac.uk

Marc Conrad
University of Bedfordshire
Park Square, Luton, UK
marc.conrad@beds.ac.uk

Carsten Maple
University of Bedfordshire
Park Square, Luton, UK
carsten.maple@beds.ac.uk

Karen Renaud
University of Glasgow
17 Lilybank Gardens, Glasgow, UK
karen@dcs.gla.ac.uk

Abstract

When members of society cannot access the World Wide Web, or the information and services it contains in a meaningful or useful way, they can become digitally excluded. Many factors have been highlighted as having an effect on the likelihood of exclusion, including psychological, material and skills related barriers. In this paper, we consider the role played by authentication systems in the divide. In light of the widely researched tension between aspects of usability and security in authentication, we identify a number of conflicting accessibility and security goals as manifested in image and sound based schemes.

1. Introduction

The term 'digital divide' is often used to describe technological inequalities that exist within and between nations [18]. Although causes for digital divides are an ongoing area of research, many factors have been highlighted as having an effect. Van Dijk and Hacker [7] conceptualise these as belonging to four categories of access obstacle that must be overcome:

Lack of any digital experience ('psychological access'.) Causes include lack of interest, computer fear, and unattractiveness of new technologies. The effects are especially prevalent in older, unemployed, and those people with low levels of literacy or education [7]. Lack of experience can self-perpetuate; as less experienced members report increased fears of online privacy violation and fraud [8].

No possession of computers and network connections ('material access'.) Members must reside in a locality where an Internet infrastructure is available. They must also have the financial ability to purchase an access device and Internet connection. It is often noted that many areas of the developing world, as well as poorer, rural areas within developed nations have been largely neglected in terms of investment in the former [12, 18]. Mobile telephones have been identified as a possible solution to both financial and geographical barriers, since they are typically cheaper to purchase and can carry out many of the functions performed by Web applications [12]. Here it is not argued that these should be used to access the Web and related content; but rather that equivalent services designed for, and delivered by, the medium should be made available. Particularly as members of society who are economically disadvantaged tend to own less sophisticated handsets [2].

Lack of digital skills ('skills access'). This is caused by "insufficient user-friendliness and inadequate education or social support; digital skills not only as the skill to operate computers and network connections, but also as the skill to search, select and process information from a superabundance of sources" [7].

Lack of significant usage opportunities ('usage access'). Finally, Van Dijk and Hacker anticipate that a new barrier will appear between those who systematically use and benefit from advanced digital technologies and the more difficult applications and services, and members only using basic digital technologies for simple applications, with a large proportion of use for the purpose of entertainment. A conceptually similar definition is provided by Zittrain [23]. Where a differentiation is made between these contrasting modes of functionality and use. Here, it is suggested that there is a continuum of flexibility between 'generative' technologies and 'appliances'.

Generative technologies are those which provide the end-user a high degree of freedom and control over components, how they are used and their purpose. Appliances, on the other hand, offer minimal user choice, and some functionality is not provided in favour of ease of use. The trend towards 'appliancisation' [3] hence carries the risk of reduced availability of development opportunities.

2. Authentication Systems and Their Role in the Divide

The widespread proliferation of Web-based applications, makes it increasingly necessary that before specific online content may be accessed, the requesting entity must first *authenticate* themselves in order to prove that they hold sufficient privileges to view or interact with the data or service being sought. For this reason, authentication systems can be considered the initial gateway to interaction with many Web sites, services and applications. There is clearly some futility in the development of universally usable content, only to have its access blocked in practice by an overly obstructive authentication procedure. Therefore we can also consider that a successful authentication experience is the gateway to unrestricted participation in the digital society. For this reason, it is important also, that we can develop methods to authenticate users without further exacerbating inequalities. Typically, users are granted access via one or more of the following mechanisms [16]:

Firstly, systems that authenticate users based on what they are, or what they do, encompassing any system that utilises behavioural and physiological biometrics to identify a user; secondly, those that authenticate given the user's knowledge of a particular secret, characteristically systems that utilise passwords and pass-phrases of various types; and finally, systems that authenticate on the basis of what the user has, where a physical token such as swipe card or USB dongle is required to authenticate.

Biometrics and tokens can provide sufficient levels of usability [14] and convenience when implemented with care, and when information about the needs and abilities of the user population is available. However, these can be unsuitable for implementation over the Internet, where user *dis*-abilities and requirements, as well as access devices and goals can vary considerably. Here, hardware-centric approaches may increase the divide, as it is possible they would generate new dependencies and possible exclusions of minority groups. Contributing factors include, cost of purchase [19], lack of installation and operating knowledge, or specifically in the case of biometrics, incompatibility between the characteristics to be measured and those of the person requesting access [5]. Mechanisms in place to cope with malfunction or failure may be insufficient for remote use and instantaneous access usually cannot be achieved until the hardware is forwarded (possibly collected by hand or delivered by mail) to the person requesting it.

If we seek to promote an inclusive Web for those who are economically, physically or cognitively impaired, it seems that software-centric approaches may well hold the key to success.

2.1. The Relationship between Security and Access

In order to maximise inclusivity on the Web, we must support users in overcoming four barriers. As discussed in section 1, these are: Psychological access, material access, skills access and usage access. In authentication systems we can disregard the usage access barrier, since it is concerned with how software is used and repurposed for use, which in general, will happen only after the authentication event has taken place. Facilitating a reduction of the divide via authentication design would therefore require the incorporation of design decisions that will aid users in overcoming the first three barriers.

In terms of psychological access, this might be achieved by appealing to all age groups and to the educated *and* to the less educated, these points essentially mean that the system must offer a high degree of *accessibility* [20]. In addition, we need to promote feelings of self-efficacy and self-confidence, especially in the novice user, i.e. offer a high degree of *usability*. Finally, we must engender the feeling that the protected asset will be secure, in order to counteract heightened fear and risk perceptions experienced by this user group. This not only relies on the actual protection offered by the system against attacks; but also on its *usability*. If the interface does not work well, this is likely to negatively impact the perceived impression of overall functionality [8].

To promote material access, authentication systems must be inexpensive or, better still, *free* to use. This means that the outlay for purchasing the technology itself should be low, and also that the required bandwidth should not be too great. It should also be possible to adapt systems for use via a mobile telephone.

To better enable skills access, systems once again, must provide a high degree of usability and accessibility. Learnability, protection from errors and the ability to recover from errors is especially important on the Web, where many users will not have access to assistance should they encounter difficulties. Authentication systems must also be efficient (i.e. quick to use once learned), especially important on the Web, where users will point their browser elsewhere, should they feel it takes too long to reach their goals [9]. Finally, the operation of the user interface itself should be easy to pick up at a proficient standard after a period away from the system (i.e. it must be memorable), as many users will access Web sites infrequently.

To ensure that the system is usable for as many people as possible, it should support multiple presentation formats (for example visual, or sound), so that loss or impairment of one physical sense will not affect the ability to operate it. Finally, the system should include support for multiple languages and character sets and instructive content should be clear, and free of jargon.

In summary, the properties of an authentication system to aid users in overcoming access barriers are: (i) secure, (ii) low in cost, (iii) usable and, (iv) accessible. Fulfilment of these goals should be achievable in theory. In practice, however, the design of authentication systems often cannot meet them all simultaneously and in full. This is because the role of authentication is two-fold. In addition to allowing the legitimate user access, these systems also must prohibit the same for the illegitimate user (who we refer to as an *attacker*). The result is a natural tension between some properties of access and those of security.

Since the primary role of authentication systems is to secure what can often be sensitive or business critical information, designers all too often will find themselves in a situation where they are forced to choose security goals over their conflicting access-oriented counterparts.

One well researched example manifestation of the trade-off is in the security and memorability of the text based password. Here, the more invulnerable a password is to attacks that utilise shortcomings in length and predictability; the less usable it becomes, because memorability decreases [17, 15].

Much focus has been placed in recent years on the development of user-centered [24] or usable security where a number of alternative systems have been trialled. Some rely on recognition of images from a challenge set [6, 1, 11]. Since humans have superior picture memory, these mechanisms have the potential to perform better than traditional passwords and are an area of promising research.

The number of images available to create image based passwords is large when compared to the number of printable characters that can be generated using common key combinations on a keyboard, these schemes can be shown to provide greater resistance to statistical attacks than an equivalent length alphanumeric password. In addition, they can also show a high rate of successful registrations and logins (E.g.[6]). However, they do encompass inherent drawbacks of their own, in that they cannot be used by those who are blind and are likely to be difficult for those who are partially sighted [10]. Image based passwords are also restricted in that they cannot be used in situations where it is not possible to use a screen, such as when authentication is required over the telephone [4].

Sound based passwords have been suggested as accessible alternatives to compliment image based schemes. Early user trials suggest that they can provide a comparable level of security and memorability [13][22].

Much research has been carried out into the trade-off between usability and security, however there is less research available addressing the existence of similar trade-offs between security and other components of inclusivity in the context of alternative authentication schemes. In the following section, we carry out a brief analysis with regard to the possibility of a trade-off with accessibility.

3 Conflict Analysis

We compare accessibility checkpoints (as defined by the current W3C Web content accessibility guidelines [20]) with a threat model consisting of known attacks that may be carried out against image and sound based passwords. Here, we consider only those attacks that would result from the design of the interface, or that would occur at the interface, and not the wider security context (for example, we do not consider here the type or level of encryption used during communication between client and server, as this bears no direct relationship with user interaction). Where it is found that adherence to a specific accessibility checkpoint results in an increased risk of successful attack, we record this as the manifestation of a conflict between the properties of accessibility and security.

Before assessing conflicts, we must decide on which of the many types of image and sound based authentication systems should undergo the analysis. We opt to consider a system which requires the user to specify a preference for using either a sound or image based password depending on their ability and skill set.

On selection of the system type, in order to create a password, the user would be presented with a challenge set of sounds or images from which they select a smaller subset to form a password sequence (this is similar to systems as described in [13] and [6]). At authentication the user is presented with a challenge set again, and is asked to indicate which of these are also members of their selected password set.

If the user interacting with the sound based option is not alone, he or she inserts earphones into the access device, and any would-be intruder observing the interaction will get no clues as to the identity of the selected audio clips.

Shielding image based passwords from onlookers is more involved, here we opt to consider systems wherein the image to be selected, and the control that the user selects, are delineated (similar to the design described in [21]). Additionally, in order to ensure an attacker does not simply recreate the interaction, ignoring the password elements altogether, we consider that the placement of elements is "shuffled" within the interface between log in sessions.

Details of the W3C checkpoints that were found to increase the risk of attacks are listed below:

- **1.2.2:** "Captions are provided for all prerecorded audio content in synchronized media, except when the media is a media alternative for text and is clearly labeled as such"
- **1.2.3:** "An alternative for time-based media or audio description of the prerecorded video content is provided for synchronized media, except when the media is a media alternative for text and is clearly labeled as such"
- **1.2.6:** "Sign language interpretation is provided for all prerecorded audio content in synchronized media"
- **1.3.3:** "Instructions provided for understanding and operating content do not rely solely on sensory characteristics of components such as shape, size, visual location, orientation, or sound"
- **1.4.1:** "Color is not used as the only visual means of conveying information"
- **1.4.3:** "The visual presentation of text and images of text has a contrast ratio of at least 4.5:1, except for the following:

  1. Large Text: Large-scale text and images of large-scale text have a contrast ratio of at least 3:1;
  2. Incidental: Text or images of text that are part of an inactive user interface component, that are pure decoration, that are not visible to anyone, or that are part of a picture that contains significant other visual content, have no contrast requirement.
  3. Logotypes: Text that is part of a logo or brand name has no minimum contrast requirement."

- **3.3.1:** "Error Identification: If an input error is automatically detected, the item that is in error is identified and the error is described to the user in text"
- **3.3.4 and 3.3.6:** Checkpoint 3.3.4 specifies that, "For Web pages that cause legal commitments or financial transactions for the user to occur, that modify or delete user-controllable data in data storage systems, or that submit user test responses, at least one of the following is true:

  1. Reversible: Submissions are reversible.
  2. Checked: Data entered by the user is checked for input errors and the user is provided an opportunity to correct them.
  3. Confirmed: A mechanism is available for reviewing, confirming, and correcting information before finalizing the submission."

  3.3.6 is identical, but specifies this for *all* Web pages, not only those that cause legal commitments or financial transactions to occur.

3.1 Threat – User writes password down, attacker finds it

- No conflicts found.

3.2 Threat – Attacker watches on as the user enters password credentials

- **Conflicts with checkpoints 1.2.2, 1.2.3 and 1.2.6:**

  Points 1.2.2 and 1.2.3 are concerned with providing captions for synchronized media, which the W3C defines as, "Audio or video synchronized with another format for presenting information and/or with time-based interactive components, unless the media is a media alternative for text that is clearly labeled as such". Additionally, 1.2.6 is concerned with the providing sign language descriptions. In sound based passwords, if we create a control that plays a tune (time-based media) that we interact with to signal our selection (i.e. is an interactive component), then we must also supply captions for that media. If the tune in question contains identifying information – for example lyrics, the caption containing these or sign language equivalent would also be visible to any onlooker, diffusing the security benefits offered by the use of earphones to shield the interaction.

  In a sound based system it might be possible to separate the media itself from playback and password selection controls. In this case, the audio component is no longer classified as "synchronized media" given the W3C definition, and we can disregard these checkpoints. This is unlikely to cause difficulties for a hearing impaired user (who would be unlikely to have opted for the sound based password, and would use the image based alternative instead).

- **Conflict with checkpoint 1.3.3:** If neither sound or image can be used to enable users to understand and operate content, we are left with the option of text. However by providing text equivalents, we again introduce new a visual output channel. Making this attack more plausible in sound based passwords, which would otherwise be protected by the use of earphones.

- **Conflict with checkpoint 3.3.4:**

  Here, we regard an incorrect entry as the selection of a non-password element during authentication. Allowing the submission to be reversed or checked may make the password more vulnerable to online brute

force attacks (discussed in Section 3.3 below). However, if confirmed (for example by displaying the selected elements to the user to confirm before submission) this would require selections to be echoed to the screen, thus increasing the risk of this attack being carried out with success.

3.3 Threat – Brute Force Attack (Online)

In this attack, the inauthentic user selects sounds or images from the challenge set at random until the correct password sequence is obtained.

- **Conflict with checkpoints 1.4.1 and 1.4.3:** Fulfilment of these checkpoints results in a reduction in the number and variety of images that can be used to form a password in an image based scheme. A smaller image set leads to increased risk of success for this attack. However, the sheer number of images available that would still meet these criterion should negate the effect of this conflict in practice.
- **Conflict with checkpoint 3.3.1:** This checkpoint is concerned with protecting the user from making errors, and allowing them to recover should they occur. If we notify the user that they have made a selection error (i.e. incorrect password element selection), in text, this does not pose a problem. However, if we identify the item that was in error (as specified in the checkpoint description), we are forced to identify to the user which incorrect element has been selected. If we do this, we aid any attacker in making a better educated guess on any subsequent reattempt.
- **Conflict with checkpoints 3.3.4 and 3.3.6:** Again, we regard an "incorrect entry" as the selection of a non-password element. Allowing user submissions to be reversed or checked would increase vulnerability to this attack. The input must therefore be confirmed. However, this would create a vulnerability for observation attacks to occur in image based schemes.

3.4 Threat – Low and Slow

In this attack, the inauthentic user selects sounds or images from the challenge set at random, but distributes their guesses over a number of accounts to avoid detection.

- **Conflict with checkpoint 1.4.3** The result of fulfilling this checkpoint is that in limiting the type of images that can be used to form a password in an image based scheme, the number available would be reduced. A smaller image set leads to increased risk of success for this attack. However, the large number of images available that would still meet these criterion should negate the effect in practice.

3.5 Threat – Prediction Attack

In this attack, the attacker knows the authentic user's taste in music and images and attempts to predict which formative password elements they would have selected.

- No conflicts found.

4 General Design Constraints

As a result of the above analysis, we also identify a number of accessibility criteria that do not conflict with security, but that impact the way image and sound based schemes should be implemented, these are discussed below.

4.1 Design Constraints Resulting from Checkpoints 2.1.1 and 2.1.3

Checkpoint 2.1.1 specifies that, "All functionality of the content is operable through a keyboard interface without requiring specific timings for individual keystrokes, except where the underlying function requires input that depends on the path of the user's movement and not just the endpoints". Checkpoint 2.1.3 is identical, except without the exception for dependance on the path of the user's movement.

As a result, is it is not advisable to design a sound based password selection procedure whereby the user is requested to specify a point in time during the playback of a clip.

4.2 Design Constraints Resulting from Checkpoints 2.2.1 and 2.2.3

These checkpoints are concerned with allowing users enough time to complete their tasks. In order to minimise the risk of online brute force attacks, developers should set an upper bound for maximum number of authentication attempts, and not a time limit. In sound based passwords, users should be allowed to select a password element even after its playback has been completed, so as to allow enough time for the selection to be made.

5 Conclusion

Usable, secure, accessible and low cost authentication systems are key in promoting digital inclusion on the Web. Considering existing tensions between some security and usability properties in this context, we are led to question whether similar trade-offs may also exist between security and accessibility. After analysis, we find that these properties **do** conflict in image and sound based schemes, and we discuss the design constraints that emerge.

Our findings suggest that if we wish to fulfil the requirements of the current Web Content Accessibility Guidelines, we cannot simply devise a system wherein the enrolling user is given an initial choice of whether to select from an audible or image based alphabet. This is because the wording of a number of checkpoints specifies that text be used as an alternative to visual or audio content. Use of text alternatives would certainly increase accessibility for deaf-blind users, however it seems this would simultaneously result in increased vulnerability to a number of attacks at the interface.

We find that in order to implement alternative schemes securely, we must seek to exploit loopholes in the current accessibility specification – which we believe is never a good strategy. A future solution might be to develop a system that offers a third option for the user to select a traditional text based password. However, if alternative password systems are to become widely deployed, and in light of the conflicts identified; a more sustainable solution might be to appeal to the W3C's Web Accessibility Initiative working group to revise their recommendations to better satisfice conflicting security and accessibility requirements.

References

[1] S. Brostoff and A. Sasse. Are passfaces more usable than passwords? a field trial investigation. In S. McDonald, editor, *People and Computers XIV - Usability or Else! Proceedings of HCI 2000*, pages 405–424. Springer, 2000.

[2] C. Bure. Digital inclusion without social inclusion: The consumption of information and communication technologies (ICTs) within homeless subculture in scotland. *Journal of Community Informatics*, 2(2), 2006.

[3] L. Church and A. Whitten. Generative Usability: Security and User Centered Design beyond the Appliance. In *Procs. New Security Paradigms Workshop (NSPW'09)*, Queens College, Oxford, United Kingdom, Sept 8-11 2009. ACM.

[4] M. Conrad, T. French, and M. Gibson. A pragmatic and musically pleasing production system for sonic events. In *10th IEEE International Conference on Information Visualisation IV06 (5-7 July 2006, London)*, pages 630–635. IEEE Publications, 0-7695-2602-0, 2006.

[5] L. Coventry, A. D. Angeli, and G. Johnson. Usability and biometric verification at the ATM interface . In *CHI 2003*, New York, 2003. ACM Press.

[6] R. Dhamija and A. Perrig. Déjà vu: A user study using images for authentication. In *Proceedings of USENIX Security Symposium*, pages 45–58, Denver, Colorado, August 2000.

[7] J. V. Dijk and K. Hacker. The digital divide as a complex and dynamic phenomenon. In *50th Annual Conference of the International Communication Association*, Acapulco, 1-5 June 2000.

[8] M. Featherman and P. Pavlou. Predicting e-services adoption: a perceived risk facets perspective. *Int. J. Human-Computer Studies*, 59:451–474, 2003.

[9] K. Forcht and R. Wex. Doing business on the internet: marketing and security aspects. *Information Management and Computer Security*, 4(4):3–9, 1996.

[10] K. Franklin and J. Roberts. A path based model for sonification. *Information Visualization*, 1(1):865–870, July 2004.

[11] S. Furnell, I. Papadopoulos, and P. Dowland. A long-term trial of alternative user authentication technologies. *Information Management & Computer Security*, 12(2):178–190, 2004.

[12] N. Geach. The digital divide, financial exclusion and mobile phone technology: Two problems, one solution? *Journal of International Trade Law and Policy*, 6(1):21–29, 2007.

[13] M. Gibson, K. Renaud, M. Conrad, and C. Maple. Musipass: Authenticating me softly with my song. In *Procs. New Security Paradigms Workshop (NSPW'09)*, Queens College, Oxford, United Kingdom, Sept 8-11 2009. ACM.

[14] International Organization for Standardization. ISO9241-11: Ergonomic requirements for office work with visual display terminals (VDTs). Available from: http://www.iso.org/iso/en/iso9000-14000/index.html.

[15] D. Klein. "foiling the cracker" – A survey of, and improvements to, password security. In *Proceedings of the second USENIX Workshop on Security*, pages 5–14, Summer 1990.

[16] B. Menkus. Understanding the use of passwords. *Computers and Security*, 7(2):132–136, 1988.

[17] R. Morris and K. Thompson. Password security: A case history. *Communications of the ACM*, 22(11):594–597, November 1979.

[18] N. Selwyn. Reconsidering political and popular understandings of the digital divide. *New Media Society*, 6(3):341–362, 2004.

[19] R. Smith. *Authentication From Passwords to Public Keys*. Addison-Wesley, NJ, USA, 2002.

[20] Web Accessibility Initiative. Web content accessibility guidelines. Web document, 11 December 2008. http://www.w3.org/TR/WCAG20/.

[21] S. Wiedenbeck, J. Waters, L. Sobrado, and J. Birget. Design and evaluation of a shoulder-surfing resistant graphical password scheme. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces*, pages 177–184, New York, NY, USA, 2006. ACM.

[22] J. Wobbrock. Tapsongs: tapping rhythm-based passwords on a single binary sensor. In *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 93–96, New York, NY, USA, 2009. ACM.

[23] J. Zittrain. *The Future of the Internet – And how to stop it*. Yale University Press, 2008.

[24] M. Zurko and R. Simon. User-centered security. In *NSPW '96: Proceedings of the 1996 workshop on New Security Paradigms*, pages 27–33, New York, NY, USA, 1996. ACM.