Persuasion as Education for Computer Security

Alain Forget School of Computer Science S & Human-Oriented Technology Lab & Hu Carleton University Ottawa, Canada aforget@scs.carleton.ca

Sonia Chiasson School of Computer Science & Human-Oriented Technology Lab Carleton University Ottawa, Canada chiasson@scs.carleton.ca

Robert Biddle Human-Oriented Technology Lab

> Carleton University Ottawa, Canada robert_biddle@carleton.ca

Abstract: Most organisations realise the importance of computer security, yet many struggle with how to teach and influence their users to behave securely. Despite existing research on new instructions and security measures, users create memorable but insecure passwords. In an effort to teach users how to behave more securely, we present the Persuasive Authentication Framework, which applies persuasive technology to authentication mechanisms. Furthermore, we describe some examples of how the framework can be applied to existing authentication systems.

Introduction

Despite existing research on improved instructions and authentication schemes, users still do not create secure passwords (Yan et al. 2004), either because they do not understand how to do so or because strong passwords are too difficult to remember and use. Furthermore, there are numerous social factors and pressures influencing users to behave insecurely (Weirich & Sasse 2001). Finally, security is a *secondary task* for most users (Whitten & Tygar 1999), often viewed as impeding the completion their primary task. Therefore, users desire security to be quick and easy.

Rather than continuing to drill users with instructions and cumbersome security measures, we propose using Persuasive Technology (PT) to teach users how to behave more securely and why it is important to do so. By applying PT principles to existing authentication methods, we can directly influence and motivate users to create more secure passwords. In doing so, we empower users to better protect themselves and their organisations from attackers. In this paper, we present the Persuasive Authentication Framework; a framework for applying PT principles to authentication schemes requiring user-chosen passwords. Our goal is to help users create stronger passwords while educating them on how to behave more securely. This applied method of coaching enables users to continue creating strong passwords even when using authentication systems that do not make use of PT.

The paper is structured as follows: We first offer some background on past efforts to help users to choose and remember more secure passwords, as well as an outline of persuasive technology and some of its current applications. Next, we describe our framework for utilising PT to educate users on how to behave more securely. The paper concludes with a general discussion and some final comments about the applicability of our framework.

Background

Usable security is a new research area combining human-computer interaction and computer security. A few inherent properties of computer security make for a challenging user interface design process and make it difficult to teach and motivate users to behave in the desired manner. While these often lead to the dismissal that "users are unmotivated", the issues require closer investigation:

- Security is a secondary task (Whitten & Tygar 1999); if the security tasks hinder users' primary tasks, they will bypass the security to accomplish said tasks.
- Users also have a poor mental model of security due to the complexity of security systems (Chiasson et al. 2006). They typically underestimate or misunderstand the consequences of insecure actions. In fact, they may not even realise that their actions put them (or others) at risk in the first place.

- Computer security also suffers from the "barn door" property (Whitten & Tygar 1999); if private information is even briefly exposed, there is no way to guarantee that it has not been compromised.
- Users are concerned about privacy and security when they can see the direct risks and impact on their lives (Shostack & Syverson 2004). Unfortunately, users typically only reach this awareness once their privacy and security has already been severely breached.

Users tend to create easily-recalled but insecure passwords (Sasse et al. 1999). While this is partially attributable to users' poor mental models of security in general, a second reason is that human memory is limited. Users are simply unable to remember a unique sequence of random characters as a password for each of their accounts. Requirements like mandatory password changes further increase the memory load. In an effort to cope, users choose very memorable but insecure passwords. Efforts at convincing users to select more secure text passwords have found only limited success. Strategies include providing instructions on the creation of secure and memorable passwords, as well as encouraging users to base their passwords on a memorable phrase (Kuo et al. 2006).

Alternative forms of authentication have been suggested to improve the memorability and security of usercreated passwords (Renaud 2005). Various forms of graphical passwords (Suo et al. 2005) have been proposed because people are better at recalling images than text (Nelson et al. 1976). For click-based graphical passwords, a password consists of clicking on a sequence of points on an image. One example is PassPoints (Wiedenbeck et al. 2005), where users clicked on five points in a given image. Unfortunately, testing found that users selected similar click-points, forming predictable "hotspots" that attackers can exploit (Thorpe et al. 2007).

Persuasive Technology is a new psychological framework proposed by B.J. Fogg at Stanford University. Fogg defines persuasive technology (PT) as "interactive computing systems designed to change people's attitudes and behaviours" (Fogg 2003). He describes sets of persuasive tools, media, and social cues that products may leverage to encourage users to behave in the desired manner. Persuasive tools render tasks quicker and easier to accomplish, persuasive media can convey messages through numerous representations, and persuasive social cues can help products appear friendly, knowledgeable, and trustworthy.

Each of the three aforementioned persuasive roles is associated with a set of persuasion strategies. Which strategies to employ depends on many factors, such as the topic, medium, target audience, and desired persuasive strength. Elements of persuasive technology are already evident in existing educational endeavours. For example, an application encouraging literacy in Chilean children (Lucero et al. 2006) utilises PT. By captivating the children's imagination through multimedia and fun activities, they can be persuaded to learn more about the characters in the application, leading the children to read and write about them.

Many other e-learning systems also apply strategies that employ persuasion. The "Clinical Nursing Practicum" (Lai et al. 2006) was taught through a mobile learning environment on students' PDAs, granting easy access to detailed information and contact with instructors as needed. The persuasive elements in this system included reduction, timely suggestions, and customisation.

Perhaps the most significant application of PT is in building thought-provoking *persuasive games* (Bogost 2007). Persuasive games aim to educate players about political, socio-economical, nutritional, environmental, and other topical issues. Although general persuasion has been used in various applications, PT has yet to be systematically applied to educating users about security.

The Persuasive Authentication Framework

Since current methods of assisting users to create more secure passwords have proven ineffective, we propose employing persuasive technology to educate users on creating memorable passwords that are also secure. Persuasion is a means of influencing others and is used to a certain degree in most daily interactions. With respect to computer security, our goal is to assist, motivate, influence, and educate users on:

- how to choose a password likely to be secure on any system;
- the importance of security and the actual threats and consequences to the users and their organisation;

• how to behave securely and utilise coping strategies to minimise the memory load.

Whereas it is impractical to have a human coach available at all times to influence user behaviour whenever computer security tasks arise, the computer will always be present. We therefore propose that the coaching and persuasion responsibilities be shifted to the computer. The computer also has the advantages of being consistent, persistent, precise, and potentially unobtrusive for some of the more subtle types of persuasion. Furthermore, since we are dealing with sensitive information such as passwords, having the computer provide the advice avoids having to reveal private information to another human.



Figure 1: The Persuasive Authentication Framework

The Persuasive Authentication Framework (Figure 1) consists of five principles based on Persuasive Technology. These principles guide the use of persuasion in authentication mechanisms to teach users how to authenticate more securely. While focusing on authentication, these principles are generalisable to other areas of security as well. We suggest that several principles be applied in concert, since several together will have more persuasive power than just one. We now describe each principle and give examples of how it can be applied to authentication mechanisms. For each principle, we identify the characteristics of usable security that are addressed:

The Simplification Principle: Authentication tasks should be made as simple as possible. This includes reducing the process to the smallest number of actions as well as reducing the complexity of the remaining tasks. By simplifying the task of authentication, users can more easily form an accurate mental model of the authentication process. Since the burden of completing the task has been reduced to an acceptable level, users will then be less likely to try to circumvent the security mechanisms, even though security is a secondary task. The optimal outcome of simplification is that the desired actions form the "path of least resistance", meaning it is easier for users to perform the authentication properly than it is to evade it.

A strategy often employed in usable security is to make security interfaces "transparent" (Chiasson et al. 2006) by hiding as much as possible from users. In our opinion, this is a misguided goal that often leads to more confusion as it usually translates into insufficient feedback for users. Simplification offers an alternative to transparency that reduces the burden on users without removing vital interface cues. The simplification principle is based on the PT tools of reduction and tunnelling.

For example, password managers (Yee & Sitaker 2006) reduce the burden on users by having the computer generate and remember complex passwords for them. Users are only responsible for entering one master password to activate the program, yet each of their accounts is protected by a distinct complex password generated by the password manger. This exemplifies the path-of-least-resistance concept; rather than having to deal with multiple passwords, users only have to maintain one strong password for the password manager.

The Personalisation Principle: Customised information for individual users typically offers a more personal and engaging experience, which could be more persuasive than generic information. Users are concerned with security and privacy if they understand the implications and consequences of their actions. By offering well-timed personalised advice relating to the individual's needs, preferences, or context-of-use, the system can provide details about why the users' current behaviour is insecure and how it can be modified to be more secure. Because the information is personalised, it is likely to help improve users' mental model of security and help them understand the relevance of behaving securely. The PT tools of tailoring and suggestion form the basis of the personalisation principle.

For example, users could provide some general interests (sports, music, cartoons, etc.) to a system that would customise a mnemonic phrase (Kuo et al. 2006) to help users remember a system-assigned random password. The given mnemonic phrase could further include something relating to the website or system itself, helping users to mentally link their mnemonic phrase and password to the system. This teaches users a coping strategy for remembering passwords that they can then apply to other passwords as well.

The Monitoring Principle: When aware that they are being observed, users are more likely to perform the desired behaviour. A system tracking user performance or status can report it directly to the users, who may then adjust their behaviour in accordance with security policies. The system should provide the opportunity for users to learn what should be done to start behaving more securely. This monitoring can be automated and done entirely by the system or can report to administrators who then take action. Furthermore, events that threaten security often happen in the background, over a long period of time, or as a result of a series of user actions. These events may not be obvious to users. In these cases, monitoring can help the system recognise these circumstances and bring them to the users' attention.

There is also the additional concern of the "barn door" property where even the slightest slip can compromise computer security (Whitten & Tygar 1999). These events may not be perceived by users as a cause for concern, which makes it important that the system raises an alert. It also provides the opportunity for intervention to teach how and why this behaviour is unacceptable. Users who modify their behaviour can then see the effect as their reported performance improves. The monitoring principle stems from the self-monitoring and surveillance tools described in PT.

For example, a system monitoring user activity could detect when users begin to enter a password (or other sensitive information) into a non-password field (or other risky circumstance) and warn them about the dangers of entering this sensitive information in the wrong place. Through immediate feedback, users could become more attentive when entering sensitive data, thus keeping private information secure.

The Conditioning Principle: Computer security is concerned about potential threats and risks to the system. However, most users have little direct experience with the consequences of an attack. When users perform a mental risk analysis, they do not believe that the probability of being attacked outweighs the additional burden of correctly performing the security tasks. In these cases, we need to artificially induce the correct behaviour because the users' natural environment does not support it. With user authentication, we want to convince people to use secure passwords even though it is a secondary task. Applying various forms of reinforcement can help shape the desired behaviour or convert existing behaviours into habits. For users to learn from any conditioning strategy, there should be other techniques at work to help users understand how to create effective passwords in order to receive the rewards for behaving securely. The conditioning PT tool is the foundation of the conditioning principle. Examples of conditioning inducements in authentication systems include:

- Longer sessions without requiring users to re-enter their password.
- A customised icon and access to extra features and benefits.
- A smiley face with encouraging messages like, "Your password is awesome! Good job!"
- Having faster system response.

The Social Interaction Principle: Authentication is an activity that typically occurs in isolation; users enter a secret password while sitting at their computer. In other areas of physical security, social norms influence behaviour and encourage users to behave securely. For example, someone may think twice about trying to enter a building without the proper credentials when there is a security guard or others nearby. The social interaction principle advocates repositioning user authentication as a social activity in order to leverage these social norms.

Users are more likely to be persuaded by a system that appears to share similar attitudes, traits, personality, and social membership. Such traits can be conveyed through language that best matches the users' own style, conveying a sense of "team". Positive and supportive language, such as personally greeting, befriending, and praising users, may further compel them to begin or continue behaving securely. Additionally, the system can be positioned to represent authority, potentially adding more persuasive power. The social interaction principle is based on several PT cues for computers as social actors.

For example, users can be taught that their own insecure behaviour puts others at risk. Through wording and presentation of the security system, users may develop a sense of belonging and duty towards their organisation. For example, organisation members can be told:

- Insecure accounts compromise not only their own account but the entire system.
- Everyone is counting on them to do their part.
- Their efforts at keeping the organisation secure are crucial and appreciated.
- "Other employees have passwords *this* strong. You don't want to be the weakest link."

Many of these principles are based in psychology because they are aimed at influencing people's behaviour. Rather than offering strictly technical solutions to authentication problems, the Persuasive Authentication Framework recognises that users play an active role in computer security and offers a means of influencing these users to behave appropriately. Since traditional methods of influencing users through education or imposing unreasonable restrictions have not been very successful, we suggest that a system that subtly persuades users and offers concrete advice may be more successful.

Applying the Persuasive Authentication Framework

We have begun applying our Persuasive Authentication Framework to security mechanisms in order to test its effectiveness. Here we give two examples of how two existing authentication mechanisms, text-based passwords and click-based graphical passwords, can be enhanced using persuasiveness.

A text-password system could adopt a "Wheel of Fortune" or "hangman" scheme during password creation where the system randomly places a small number of uncommon characters into the password (e.g. " $_ x _^{-} V$ _"), allowing users to choose the remaining characters. Should the user prefer different characters in different positions, pressing a "shuffle" button would randomly choose a new set of uncommon characters and positions. The purpose of inserting random characters at random positions is to make passwords harder for attackers to guess.

A similar system could be adopted for click-based graphical passwords, such as PassPoints (Wiedenbeck et al. 2005) or Cued Click-Points (Chiasson et al. 2007b). When creating a graphical password, users could be guided in selecting their click-points by lightly shading the entire image, except for a small area known as the *viewport* (see Figure 2). Users can only choose a click-point within this randomly-positioned viewport. If they are unable to find a suitable click-point, they can press the "shuffle" button to randomly reposition the viewport. The most straightforward and quickest action is to select a click-point from the first viewport. However, someone determined to reach a specific click-point can repeatedly shuffle until the viewport reaches that area.

These schemes employ the following three principles from our Persuasive Authentication Framework:

- *The Simplification Principle:* By anchoring a few random uncommon characters in the user's text password, the system removes the immediate need for users to devise their own secure password strategy. Furthermore, users can learn by example that the insertion of random uncommon characters in passwords leads to greater security, which they can then apply to other passwords. The viewport scheme simplifies the password creation task by providing a smaller area in which users may choose a point. The viewport discourages users from choosing hotspots as their click-points since the shuffle button will likely need to be pressed numerous times before the viewport falls on the one click-point users find the most memorable. This tedious selection process for insecure click-points persuades users to choose more random, and hence more secure, click-points.
- The Personalisation Principle: Knowing the characters and their positions are random suggests to users that these selections are unique and were chosen especially for them. This leads users to feel their password is more secure, motivating them to comply with using a password with the given inserted characters, as well as possibly applying the learnt scheme themselves to their other passwords. Similarly, in the viewport scheme, users believe that the viewport's initial random position is unique and placed especially for them, leading them to feel the area is more secure. Likewise, they are motivated to choose a click-point in the

initial viewport, rather than shuffle. Since users are unlikely to have ever used a click-based graphical password, they are particularly open to advice on choosing click-points. The viewport makes its suggestion at the most opportune moment; when users are first faced with creating a graphical password.

• *The Conditioning Principle*: Continually pressing the shuffle button, in the hopes of finding some desired set of characters and positions, or to select one particular click-point, can be very tedious. Rendering common letter combinations and hotspots unattractive choices trains users that such decisions result in poor security. Furthermore, since it is the fastest way to create a password, complying with the system's first suggestion appeals to users since it allows them to complete this secondary task as quickly as possible.

The hangman scheme guides users in selecting their password while largely preventing the use of common words, and limits password reuse since any newly created password will have different starting characters. Users learn that uncommon character combinations and randomness improves the security of their passwords; a concept that can also be applied to other passwords. Both the hangman and viewport schemes could be respectively implemented for any text-based password system or click-based graphical password system.



Figure 2: Screenshot of the viewport-CCP Create Password interface persuading users to choose a click-point within the highlighted area (*viewport*). (Pool image from PD Photo 2007)

We have recently implemented the viewport scheme for Cued Click-Points (CCP) (Chiasson et al. 2007b), shown in Figure 2. Using data from an in-lab user study with 20 participants, where a total of 184 trials were completed, we compared the viewport-CCP click-points to those collected from our earlier PassPoints (Chiasson et al. 2007a) and CCP studies (Chiasson et al. 2007b). We found that the viewport-CCP click-points were much more uniformly distributed, reducing hotspots. Since most participants used the shuffle button sparingly, the viewport mostly remained in its initial random position, lowering the chance of participants selecting hotspots as their click-points. This shows that the security viewport was successful in persuading most users to choose more random click-points, and thus taught users to create more secure passwords.

Discussion

Through the Persuasive Authentication Framework we have described, we hope to teach users how secure passwords can be created, empowering them to continue behaving securely even when not using a framework-enhanced authentication mechanism. There are several advantages to using PT to educate users rather than providing traditional security instructions and imposing rules:

- A security mechanism with built-in PT teaches users how to create secure passwords in the context where the password will be used. Teaching secure behaviour in a classroom or meeting room is out of context and is unlikely to be effective.
- In regular systems, users often do not understand the security rules imposed upon them and as a result of frustration, they will either try to circumvent the security mechanism or fulfill the new requirements as minimally as possible. Using persuasion, we hope users will be more willing to comply with security rules.
- PT is interactive, giving users the opportunity to learn by doing. Traditional lecturing methods do not include hands-on ways to apply knowledge in the very environment it is meant to be used.
- PT leverages our innate cognitive abilities such that users may not even be aware they are being persuaded or taught. This minimises resistance to new security measures, particularly since properly implemented PT principles should result in a system that is easier to use.

Persuasive technology must be applied with great care, because there is always a risk of annoying users to the point that they rebel against the system. One example of this is a system proposed by Brustoloni & Villamarín-Salomón (2007) intended to help protect users against phishing emails (which attempt to trick users into divulging private information, such as online bank account credentials and credit card numbers). Their warning dialog boxes changed each time they were displayed, forcing users to carefully read them before deciding on the appropriate response. Furthermore, users were required to provide justification for their actions and were audited by administrators who quarantined users who behaved insecurely. Although the authors did not explicitly use PT, many of the principles presented in this paper can be seen in their system. However, these principles are being severely misused such that users will quickly become angered at the very system that is supposed to protect them. When using the Persuasive Authentication Framework and persuasive technology in general, it is crucial to employ tactics aimed to help, assist, and teach users not only how to perform the desired action, but also why such action is beneficial for them and their organisation. Employing PT to force, scare, and coerce users to do one's bidding is counter to the intended purpose of persuasive technology and our framework.

Finally, with security applications there is always the risk of leaking information to attackers. If not implemented carefully, features intended to help users and increase the usability of the system can often be leveraged by attackers to help them break into the system. Any additional cues provided to users must be fully evaluated from a security perspective to ensure that security is not compromised. This is why the viewport is placed randomly rather than purposefully placed away from hotspots. If the viewport avoided hotspots when shuffling, attackers could learn which points are most popular by simply watching the areas the viewport avoids when shuffling.

Conclusion

To date, attempts at educating users to behave securely have had only limited success. Although persuasive technology has been successful in training users in numerous other domains, we have not yet seen it applied to security. Through the Persuasive Authentication Framework, we are proposing a new methodology of utilising persuasive technology to educate users to better protect themselves from attackers. This offers a new versatile strategy for teaching and influencing users to behave more securely.

The hangman and viewport persuasion schemes are examples of how the Persuasive Authentication Framework can be applied. The viewport-CCP's preliminary user study results demonstrate that our framework shows promise as an effective tool in educating users to create more secure passwords. We look forward to further refining our implementation ideas as well as applying the framework to other forms of authentication. Persuasion need not stop at authentication however; our framework can also be utilised to educate users about security certificates, phishing, encryption, malware, and many other contemporary security issues. Although users currently have little to no understanding of how best to use the security measures employed to protect them, we hope to change that fact using our Persuasive Authentication Framework.

References

Bogost, I. (2007). Persuasive Games. Cambridge, MA, USA: MIT Press.

Brustoloni, J.C., & Villamarín-Salomón, R. (2007). Improving Security Decisions with Polymorphic and Audited Dialogs. 3rd Symposium on Usable Privacy and Security (SOUPS), 2007, ACM Press, New York, NY, USA. 76-87.

Chiasson, S., Biddle, R., & van Oorschot, P.C. (2007). A Second Look at the Usability of Click-based Graphical Passwords. 3rd Symposium on Usable Privacy and Security (SOUPS), 2007, ACM Press, New York, NY, USA. 1-12.

Chiasson, S., van Oorschot, P.C., & Biddle, R. (2007). Graphical Password Authentication Using Cued Click-points. 12th European Symposium On Research In Computer Security (ESORICS), 2007, Springer-Verlag, Berlin Heidelberg, Germany.

Chiasson, S., van Oorschot, P.C., & Biddle, R. (2006). A Usability Study and Critique of Two Password Managers. 15th USENIX Security Symposium, 2006, USENIX, Berkeley, CA, USA. 1-16.

Fogg, B.J. (2003) Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco, CA, USA: Morgan Kaufmann.

Kuo, C., Romanosky, S., & Cranor. L.F. (2006). Human Selection of Mnemonic Phrase-based Passwords. 2nd Symposium on Usable Privacy and Security (SOUPS), 2006, ACM Press, New York, NY, USA. 67-78.

Lai, C., Wu, C., & Chen, S. (2006). A Mobile Learning Environment to Support the Clinical Nursing Practicum. *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (E-Learn) 2006.* Association for the Advancement of Computing in Education (AACE), Chesapeake, VA, USA. 695-700.

Lucero, A., Zuloaga, R., Mota, S., & Muñoz, F. (2006) Persuasive Technologies in Education: Improving Motivation to Read and Write for Children. *1st International Conference on Persuasive Technology, 2006*, Springer-Verlag, Berlin Heidelberg, Germany. 142-153.

Nelson, D.L., Reed, V.S., & Walling, J.R. (1976). Pictorial Superiority Effect. Journal of Experimental Psychology: Human Learning and Memory, 2 (5), 523-528, 1976.

PD Photo. (2007). http://pdphoto.org/ Accessed August 2007.

Renaud, K. (2005). Evaluating Authentication Mechanisms. In Cranor, L.F., & Garfinkel, S. (Eds.), *Security and Usability* (pp. 103-128). O'Reilly Media, Sebastopol, CA, USA. 2004.

Sasse, M.A., Brostoff, S., & Weirich, D. (1999). Transforming the 'Weakest Link': A Human-Computer Interaction Approach to Usable and Effective Security. *BT Technical Journal*, 19 (3), 122-131.

Shostack, A., & P. Syverson. (2004). What Price Privacy? (and why identity theft is about neither identity nor theft). In Camp, L.J., & Lewis, S. (Eds.), *Economics of Information Security* (pp. 129-142). Kluwer Academic Publishers, Norwell, MA, USA. 2004.

Suo, X., Zhu, Y., & Owen, G.S. (2005). Graphical Passwords: A Survey. 21st Annual Computer Security Applications Conference (ACSAC), 2005, IEEE Computer Society, Los Alamitos, CA, USA. 463-472.

Thorpe, J., & van Oorschot, P.C. (2007). Human-Seeded Attacks and Exploiting Hot-Spots in Graphical Passwords. *16th USENIX Security Symposium, 2007, USENIX, Berkeley, CA, USA.* 103-118.

Weirich, D., & Sasse, M.A. (2001). Pretty Good Persuasion: A first step towards effective password security in the real world. *New Security Paradigms Workshop (NSPW). 2001, ACM Press, New York, NY, USA. 137-143.*

Whitten, A., & Tygar, J.D. (1999). Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. 8th USENIX Security Symposium, 1999, USENIX, Berkeley, CA, USA. 169-183.

Wiedenbeck, S., Waters, J., Birget, J.C., Brodskiy, A., & Memon, N. (2005). PassPoints: Design and longitudinal evaluation of a graphical password system. *International Journal of Human-Computer Studies*, 63 (1-2), 102-127.

Yan, J., Blackwell, A., Anderson, R., & Grant, A. (2004). Password Memorability and Security: Empirical Results. *IEEE Security* & *Privacy Magazine, Sept-Oct 2004*, 2 (5), 25-31.

Yee, K., & Sitaker, K. (2006). Passpet: convenient password management and phishing protection. 2nd Symposium on Usable Privacy and Security (SOUPS), 2006, ACM Press, New York, NY, USA. 32-43.