Human-in-the-loop: rethinking security in mobile and pervasive systems

Vassilis Kostakos
University of Madeira / Carnegie Mellon University
Funchal 9000-319, Portugal
vassilis@cmu.edu

Eamonn O'Neill
University of Bath
Department of Computer Science
Bath BA2 7AY, UK
eamonn@cs.bath.ac.uk

CHI 2008, April 5 – April 10, 2008, Florence, Italy
ACM 978-1-60558-012-8/08/04.

Abstract
In this paper we argue that pervasive systems introduce human-driven security vulnerabilities that traditional usability design cannot address. We claim that there is a need to understand better the appropriate role of humans in the context of pervasive systems security, and to develop quantifiable and measurable concepts that describe humans and their relationship with our systems. Here, we highlight mobility and sociability as two new sources of security vulnerabilities for pervasive systems, and describe our method for developing quantifiable metrics for these concepts.

Keywords
Security, usability, pervasive and ubiquitous systems, mobility, sociability.

ACM Classification Keywords
H1.2 Models and principles: User/Machine Systems. K6.5 Management of computing and information systems: Security and protection.

Introduction
Since as early as 1975 researchers have pointed out the need for a better understanding of the role of humans in the security of computer and information systems, calling for "psychological acceptability" [17]. With the subsequent establishment of HCI as a

discipline came the realisation that humans are an integral, yet weak link in the security loop of computer systems. The unique role of humans is due to the fact that they are unfit for storing strong cryptographic keys, they are slow and inaccurate when performing cryptographic operations, and choose to circumvent security mechanisms in order to carry out their work [2,9].

Traditional HCI literature has considered the human-in-the-loop security issues as a design problem in need of appropriate interfaces, interactions, and policies [3]. Similarly, security experts have demonstrated the social engineering aspects of humans-in-the-loop and have called for better employee training and increased awareness [16]. More recently, researchers have begun to explore the trade-offs between usability and security in more detail, in the context of both desktop and pervasive systems [e.g. 5].

In this paper we argue that pervasive systems introduce human-driven security vulnerabilities that traditional usability design cannot address. We claim that there is a need to understand better the appropriate role of humans in the context of pervasive systems security, and to develop quantifiable and measurable concepts that describe humans and their relationship with our mobile and pervasive systems.

Aiming for security

Secure computer systems provide confidentiality, integrity, and availability [e.g. 19]. In other words, a secure system does not allow unauthorised reads, unauthorised writes, and always provides access to authorised users. These concrete objectives have enabled security researchers to make progress in the context of constant changes in computer capabilities and computer use. In fact, while today we are researching systems that provide multiple devices per user in an everyday environment, security is still focused on achieving the same unchanged goals within these new conditions.

A further aspect that has remained unchanged throughout the history of security research has been humans, both in their capacity and behaviour. Judged from a security perspective, humans have bad memory (cannot remember many passwords, choose easy-to-crack passwords), bad habits (write their passwords under the mouse mat), are easily persuaded to break protocol, and can easily be confused by tricks in the user interface [6,14].

To address the persistent friction between security objectives and human properties, a number of strategies has been followed. Initially, systems were made extremely secure at the cost of usability [e.g. 18]. Recently, anecdotal evidence about the Windows Vista operating system has raised similar concerns [2]. A second, and arguably more fruitful approach, has been to understand users' "weaknesses" and design appropriate systems. This has yielded solutions that address, for example, users' weak memory by utilising keychains or photographic passwords [4,20], weak attention to detail by focusing on detecting patterns [6] and inability to perceive wireless signals by utilising short-range technologies [e.g. 11].

More recently, the focus on mobile and pervasive computing has produced quite interesting security solutions utilising pervasive technology. Examples include solutions that involve shaking devices before using them [13], using tags and physical objects as keys [15], relying on extremely short-range wireless

technologies for secure authentication [12], or even using light and sound patterns to verify wireless connections [10].

Security in mobile and pervasive systems has by necessity moved away from the desktop and has spilled into the everyday world and everyday objects. At the moment, however, it is difficult to abstract away from the specific security solutions presented in the literature and describe the role of humans. This is due to the fact that solutions are aimed at specific contexts and systems, and do not transfer well. For instance, while shaking mobile devices can be a fun yet secure way of establishing a secure channel, the same is not true when attempting to communicate with a large plasma display. Similarly, while tagging and short-range communications are secure mechanisms, users' don't always want to communicate with people in physical vicinity.

As designers, we seek guidelines that can be applied to different contexts. To derive such guidelines, we need to consider security solutions from a high-level perspective, and identify those design features that transfer across contexts. For instance, are humans best utilised as "shakers" of things, tag-finders, or perpetual pattern detectors? To derive appropriate guidelines, we consider how we can draw on the lessons learned in developing desktop systems.

Desktop security mechanisms have been designed with two sets of constraints in mind: security objectives and human characteristics. While both remain unchanged when considering pervasive systems, the shift from desktop to pervasive computing introduced further considerations. A first step towards determining appropriate ways of integrating humans in the design of secure pervasive systems is to identify those unique characteristics of pervasive systems that set them apart from desktop systems.

Of course, understanding and designing for human behaviour has been a tenet of HCI. Yet while HCI security research has mostly focused on human memory, attention, cognition, and error, pervasive systems are embracing a palette of human behaviour that cannot be described, at least for now, only in those terms. As we move away from computer systems that deal with files and databases, towards real-world everyday systems that support shopping, cooking and exercise, our systems are exposed to an increasing range of human characteristics. As a knock-on effect, these characteristics are introducing potential security vulnerabilities in pervasive systems.

New sources of security vulnerabilities

The success of secure yet usable solutions relies on our ability to quantify and measure those human characteristics that result in security compromises. Human memory and attention, for example, are well understood in the context of security. Thus, we know that security is enhanced if users are asked to remember a picture rather than a cryptic word [4], or if humans have to detect changes in patterns rather than changes in detail [6]. Such knowledge can guide us in developing secure systems. In our work, we have identified mobility and sociability as two of the many aspects of human behaviour that introduce security vulnerabilities in pervasive systems. We focus on these two characteristics because we are increasingly capable of capturing them on a large scale, they are well-understood, and easy to quantify.

Security implications of mobility
While security mechanisms are designed to control the flow of information within the digital infrastructure, we currently have no robust way of detecting the movement of information within, say, urban space. Consider the example of multi-gigabyte mobile storage devices. When being carried, they may be considered not just as storage but also as bandwidth. This bandwidth is effectively uncontrolled by computer systems as it consists of data that have left a host computer and will end up wherever and whenever the user decides. Additionally, this bandwidth is vulnerable to hijacking: stealing one's storage device and reading its contents can become a security threat.

Furthermore, although security protocols are being developed to enable computers to deal with such situations [10], users on the move typically have to make security-related decisions based on incomplete or inaccurate knowledge: file exchange over proximity-based channels (Bluetooth and Infrared), logging on to wireless access points, using public terminals, or using one's NFC phone to pay for tickets.

The counterpart of mobility, embeddedness, also has potential security implications. Embedded and fixed systems such as kiosks, art installations, and monitoring equipment are all open to physical attack. In addition to vandalism, we must also consider physical hijacking where the attackers take control of the device and manipulate it. Similarly, we can consider car sensors installed by insurance companies, taxi meters inspected by the government, or cargo ships' sensors, all of which are pervasive systems that assist humans who may be motivated to contravene. In fact, the presence of unknown or untrustworthy humans becomes a real security problem because our users will interact and socialise with them.

Security implications of sociability
Humans are social animals, and as part of everyday life humans socialise. Designers of desktop systems have attempted to address social exploitation, also known as social engineering. The effectiveness of social engineering attacks [14], and reports of people's failure to comply with organisational security policies, demonstrate that social behaviour can lead to security compromises. Attackers exploit internalised norms [8], either by pretending to be a co-worker in need of help, or someone in authority. Similarly, many email scams and phishing attacks use the same approach [7]. The potential for social engineering greatly increases when we consider pervasive systems. Users can potentially be tricked or "helped" into associating with the wrong wireless device, touching their phone on an insecure reader or scanning a malicious tag.

It is often assumed that one way of reducing exposure to social engineering is to limit the role that people have in the secure system. Such a solution, however, is not desirable in the case of pervasive systems, as it would defeat the very purpose of designing systems for everyday life. Further solutions presented in the literature typically involve institutionalised procedures, such as developing rules for mutual authentication, providing support for reporting incidents, developing appropriate ethics, and holding security awareness campaigns [2,8], which can be inappropriate for everyday life settings.

Quantifying mobility and sociability

In our research we have begun to measure both mobility and sociability on a city scale. We have

deployed a number of computers that carry out constant Bluetooth scanning across the city of Bath in the UK. Utilising Bluetooth technology, which closely maps to users' mobility and presence, we are able to uniquely identify movement patterns across a city. Additionally, we are able to measure sociability by analysing the patterns of encounter using social network tools. We are still quite a long way from being able to address design issues in relation to security, mobility and sociability. However, at the moment we are identifying the ground truth, and at the same time establishing a scientific basis by developing an array of metrics which can be captured automatically.

Mobility metrics
Users and cities have unique patterns of mobility. Considered from an ego-centric perspective, useful metrics are distance travelled (km per day) and speed (km per hour). When considering mobility from an exo-centric perspective, flow becomes a useful metric (people/hour), as well as visit duration and encounter duration (in the form of a time-based distribution).

Sociability metrics
In terms of sociability, we are able to automatically capture the city's and people's social groups and patterns of encounter. Social structures can be examined from an ego-centric or exo-centric perspective and involve measures such as group or community size, and number of singles vs. couples. Numerous concrete metrics can be adopted from traditional social network analysis such as average degree (number of people someone interacts with), betweenness (0 to 1 indicating the importance of a person as a link in the chain of information spreading) and closeness (0 to 1 indicating the reachability of a person within the social network).

The metrics we indicate here can be used to guide the design of pervasive systems, but more importantly can be sensed and utilised by pervasive systems in real time. For instance, a system may utilise a map of a city's social relationships to calculate its social "proximity" to another entity, and thus adapt its security behaviour. Similarly, systems may look for deviations from patterns of mobility and sociability.

Ongoing work
Our scenarios of utilising mobility and sociability metrics make a crucial assumption: that we are able to link those metrics to security considerations. While we know that cognitive overload, or password complexity may cause users to write passwords on sticky notes, we currently do not know how to relate mobility and sociability metrics to security considerations. Having developed the mechanisms to sense these metrics, our ongoing efforts are focused on systematically relating those metrics to security considerations. In doing so we aim to identify design and operational guidelines that relate to objective and measurable metrics. This will be a crucial step in developing secure pervasive systems that people can use in their everyday lives.

Acknowledgements
We wish to thank Tim Kindberg for his guidance and insightful comments. This research is funded by the UK Engineering and Physical Sciences Research Council grant EP/C547683/1 (Cityware: urban design and pervasive systems).

References
[1] Adams, A. and Sasse, M.A (1999). Users are not the enemy. Communications of the ACM, 42(12): 41-46.

[2] Atwood, J (2006). Windows Vista: Security through endless warning dialogs. http://www.codinghorror.com/blog/archives/000571.html

[3] Brostoff, S., and Sasse, M.A. (2001). Safe and sound: a safety-critical approach to security. NSPW, New Mexico, USA, pp. 41-50.

[4] Chiasson, S., Biddle, R., van Oorschot, PC. (2007). A Second Look at the Usability of Click-Based Graphical Passwords. Symposium on usable privacy and security, Pittsburgh, USA.

[5] Cranor, L. & Garfinkel, S. (2005). Security and Usability: Designing Secure Systems that People Can Use. O'Reillly Media, Inc.

[6] Dhamija, R. & Tygar, J.D. (2005). The Battle Against Phishing: Dynamic Security Skins. Symposium on usable privacy and security, Pittsburgh, USA.

[7] Dhamija, R., Tygar, J. D., and Hearst, M. (2006). Why phishing works. Proc. CHI 2006, 581-590.

[8] Flechais, I., Riegelsberger, J., Sasse, M.A. (2005). Divide and Conquer: the role of trust and assurance in the design of secure socio-technical systems. Proc. New security paradigms workshop 2005, Lake Arrowhead, California, USA.

[9] Kaufman, C., Perlman, R., and Speciner M. (2002). Network Security: Private Communication in a Public World, 2nd ed., Prentice-Hall.

[10] Kindberg, T. & Zhang, K. (2002). Validating and Securing Spontaneous Associations between Wireless Devices. Proc. 6th Information Security Conference (ISC'03), LNCS 2581, Springer, 44-53.

[11] Kostakos, V., O'Neill, E., Shahi, A. (2006). Building Common Ground for Face to Face Interactions by Sharing Mobile Device Context. Proc. LOCA 2006, Dublin, Ireland. LNCS 3987, Springer, 222-238.

[12] Kostakos, V. and O'Neill, E. (2007). NFC on mobile phones: issues, lessons and future research. Proc. Percom 2007, Pervasive RFID/NFC Technology Workshop (Pertec 2007), NY, USA, 367-370.

[13] Mayrhofer R. & Gellersen, H. (2007). Shake well before use: Authentication based on accelerometer data. In Proc. Pervasive 2007, LNCS 4480, Springer, 44-161,

[14] Mitnick, K. D. & Simon, W. L. (2003). The Art of Deception: Controlling the Human Element of Security. John Wiley & Sons Inc.

[15] O'Neill, E., Thompson, P., Garzonis, S. and Warr, A. (2007). Reach out and touch: using NFC and 2D barcodes for service discovery and interaction with mobile devices. In Proc. Pervasive 2007, LNCS 4480, Springer, 19-36.

[16] Poulsen K: 'Mitnick to lawmakers: People, phones and weakest links', (March 2000) — http://www.politechbot.com/p-00969.html

[17] Saltzer, J.H. & Schroeder, M.D. (1975). The Protection of Information in Computer Systems. Proc. IEEE, 63(9):1278–1308.

[18] Smith, S.W. (2003). Humans in the Loop. IEEE Security & Privacy, 1(3):75-79.

[19] Stajano, F. (2003). Security for Whom? The shifting security assumptions of pervasive computing. Proc. Software Security - Theories and Systems, LNCS 2609, Springer, 16-27.

[20] Yee, K.P. & Sitake, K. (2006). Passpet: convenient password management and phishing protection. Proc. Symposium on usable privacy and security, Pittsburgh, USA.