Influence of User Choice on Perception of Wireless Connection Genuineness and Security

Chris Bevan[1], James Mitchell[1], Tim Kindberg[2], Eamonn O'Neill[1], Jim Grimmett[1], Danaë Stanton Fraser[1], and Dawn Woodgate[1]

[1] University of Bath, UK
eamonn@cs.bath.ac.uk,
[2] Matter 2 Media Ltd, Pervasive Media Studio, Leadworks, Anchor Square, Harbourside, Bristol, UK, BS1 5DB
tim@matter2media.com

Abstract. People's perceptions of the security and bona fides of urban pervasive services, and their trust in them, do not necessarily match the reality of a given service. Taking WiFi hotspots as an example, this study investigated the effects on users' perceptions of a service's genuineness and security of allowing the user to choose the message used in the *Physical Interlock* device association protocol. Users were significantly more confident in the genuineness of the wireless network and the ability of *PI* to defend them against a Man-in-the-Middle attack if they contributed directly to the creation of the message. However, user creation of the entire message did not significantly affect user perceptions compared to user creation of half the message combined with system generation of the other half. Thus, messages that combine partial user generation with partial system generation may give people confidence in secure ad-hoc associations between their personal devices and urban pervasive services, while ensuring a known lower bound on message strength.

Keywords: Pervasive Computing, Security, Privacy, Trust, User Perception, Device Association, Device Authentication

1 Introduction

Urban pervasive computing brings new aspects to threats that we are already familiar with from the desktop computing environment, such as phishing attacks. Phishing occurs when a service seems to originate from a situation or environment that users are likely to trust, but in fact originates from an attacker. A feature of the pervasive phishing attack is that the service may seem to be associated with a trusted setting in the urban environment. A well known form of this threat is the man-in-the-middle (MITM) attack on WiFi hotspot users. The attacker provides a WiFi access point from her mobile device that, through a measure as simple as a situationally relevant network SSID, people will associate with, for example, the cafë they are in. The attacker may attempt to install malware, obtain personal information such as credit card details, or otherwise eavesdrop and interfere with network traffic without the victim being aware.

Since Stajano & Anderson's *Resurrecting Duckling* [13], a number of protocols have been developed to facilitate secure ad-hoc device associations that combat the *Man-In-The-Middle* (MITM) attack [1]. Sometimes, these protocols validate the device association after the exchange of a secret key, requiring the user to compare data displayed by the associated devices which are identical if and only if there is no MITM. To make the comparison tractable to humans, the authenticating data has been represented in various forms of visual ([8], [5]), auditory [12] and tactile ([11]) feedback. A review of such techniques has recently been offered [7], and a comparative study has been conducted on a number of pairing methods [6].

A recent example of an ad-hoc device association protocol is the *Physical Interlock* (PI) protocol ([4], [3]), a variation of the Interlock protocol [9]. *PI* involves the selection of a one-time secret message. In the original implementation of *PI*, the message is sent from the mobile device that is connecting to a host venue's wireless network, to a display linked to that venue's physical premises (e.g. a wall-mounted LCD screen). The user is then able to verify the authenticity of the connection by comparing the original message with the one that appears on the display. However, unlike protocols where the user can have no choice over the data being compared since it directly depends on a key chosen by the system, the *PI* protocol can be adapted so the user chooses none, some or all of the message to be compared.

In [3], we found the *PI* protocol to be capable of enabling ad-hoc device connections that were perceived as trustworthy and secure against MITM. However, users raised concerns about the complexity of the user interaction and the public visibility of the message comparison. Here, we investigate a new variant of the *PI* protocol that addresses these issues and provides the user with the ability to create some or all of the message. In this new version of *PI*, the public display is replaced by the screen of the user's own mobile phone. Further, the number of discrete steps required to complete the authentication process is reduced without compromising the underlying security of the protocol.

We report a user study of how user choice in the generation of the verification message affected user perceptions of the security value of *PI* in two respects: *genuineness*, i.e. that the connection made was with the intended host venue's network, and *security*, i.e. that the connection made was free of MITM. In this study, three variants of the revised *PI* protocol were compared, in which users were able to choose half, all or none of a required four-word message, the system choosing the remaining words.

2 Interaction Design

2.1 Protocol Procedure

The *PI* authentication procedure involves the user's mobile computer and mobile phone. Taking a cafe-based scenario, the authentication procedure is as follows:

1. On her computer, the user scans for local wireless networks using the *PI* user interface and chooses a network which seems to belong to the cafe.
2. Using her mobile phone, the user scans a 2D barcode printed on a physical authentication card supplied by the cafe staff. The URL embedded in the barcode directs the phone's browser to an SSL-secured website using a cellular data connection.
3. The user establishes a connection between her phone and her computer by selecting an avatar on her phone's display that matches one shown on her computer's display.
4. Using her computer, the user (or system; see below) generates a four-word message.
5. The user is instructed by her computer to check for a prompt on the phone-based UI before transmitting her message to the authentication system.
6. Finally, the user compares the messages displayed on both her phone and computer (figure 1).

Fig. 1. *Physical Interlock* message comparison interface, showing both computer and mobile phone-based UI.

The authentication procedure we describe above was designed to be as simple as possible while supporting the *PI* protocol. The shared encryption keys used in *PI* are derived from keying material generated in step 1 above, in which the user joins the wireless network using 802.1X/EAP-TTLS authentication. After the message has been generated in step 4, it is encrypted and the partial ciphertext is sent over the wireless network. The prompts in step 5 of the interaction ensure, as required by the *PI* protocol, that the authentication system has received the first part of the ciphertext before the final part is sent. A more detailed description of the underlying *PI* protocol (and its application to authenticating public 802.11 networks) is provided in [4], on which this implementation is directly based.

2.2 Message Generation

The *PI* protocol requires a one-time secret message to be created. For this study, we instantiated the message as a set of four English words. When two or four of the message words were chosen by the system, the words were chosen at random from a set of 1,055 without repetition, resulting respectively in ${}^{1,055}P_2 \approx 10^6$ or ${}^{1,055}P_4 \approx 10^{12}$ choices. While an attacker may try to guess the secret message, the protocol restricts him to a single attempt [4]. An incorrect guess will cause mismatched messages to be displayed on the users' phone and computer.

Although the size of the source word set was not disclosed to participants, we attempted to ensure that the perceived message strength was both sufficient and (as far as possible) constant between participants. To reduce the possibility of coherent sentence fragments being generated (possibly suggesting that messages were chosen from a smaller set of four-word phrases), we chose a word set containing only nouns. We further constrained the length of words in the set to be between 5 and 10 letters, to avoid a situation in which only very short words were chosen (implying a smaller word set) while also avoiding extremely long words which may have been more difficult to compare between displays.

When users choose some or all of the words in the message, the strength of the message becomes far more difficult to quantify. Participants were encouraged to *"...choose words which would be hard for someone to guess, so try not to choose a sentence or words that go together"*. In order to maintain some similarity between user–chosen and system–generated words, we placed constraints on user inputs which were similar to the parameters used by the system. Repetition was forbidden and only words between 5 and 10 letters in length were accepted. Given these constraints, we expect that each word chosen by users would contain entropy in excess of the most pessimistic average of 0.6 bits per character (i.e. 3 to 6 bits per word) observed by Shannon in [10], but below the estimate of 12 to 18 bits per word suggested by Burr et al. [2] when considering user-selected alphabetic passwords. This compares to a known quantity of entropy contributed by each word chosen by the system of approximately 10.0 bits, i.e. $\log_2(n)$, where n is initially 1,055 and falls by one for each unique choice from the set of words.

3 Study Design

We conducted a user study to examine the effect of offering users partial or full choice over the message required by the revised *PI* protocol. The experimental evaluation was conducted in a lab setting, in which a wireless network access point was implemented with the SSID (Service Set IDentifier) *AvatarellisWireless*. The limitations of lab studies are important to note, especially with regards to the study of subjective factors such as perceived security and user trust. But, while a lab setting can provide only indications of our participants' consideration of security issues, the possible effects on ecological validity of using a lab rather than a field setting were constant across all our experimental conditions.

3.1 Method

The *PI* protocol is designed to provide user-verifiable evidence of a secure end-to-end connection with a wireless digital service. We tested two relevant hypotheses:

H1: As the proportion of words that the user is able to choose for the message increases relative to the number of words generated by the system, user confidence in the *genuineness* of the service will increase.

H2: As the proportion of words that the user is able to choose for the message increases relative to the number of words generated by the system, user confidence in the *security* of the service will increase.

The term *genuineness* refers to the degree to which participants feel that the service they are using is bona fide and provided by the venue. The term *security* refers to the degree to which participants feel that the service they are using is resistant to MITM. The principle of the MITM attack was explained to participants prior to their exposure to *PI*.

The independent variable (**IV**) was *message choice*, i.e. the degree to which participants were able to generate their own message. Three conditions of *message choice* were created (and all other factors remained constant):

1. **No user choice.** The system chooses all four words in the message, randomly selected without repetition from a set of 1055 English nouns.
2. **Half user choice.** The system chooses two words and the user chooses two words. User choice was facilitated by free-text entry with syntactic constraints (see section 2.2).
3. **Full user choice.** The user chooses all four words, with the same constraints as in the previous condition.

The two dependent measures were the degree to which participants considered the network connection in each condition to be *genuine* (**DV1**) and *secure* (**DV2**).

24 participants (m=15, f=9, modal age range=18-29), were recruited by email and general opportunity sampling. The participants were sourced predominantly from staff and students of the University of Bath.

Participants were run individually. Each participant was provided with a laptop computer to use at a table and a mobile phone (HTC *myTouch 3G*) with 3G Internet connectivity and 2D barcode reader software. An *Avatarelli's* branded WiFi service "connection card" was also provided, with a 2D barcode printed on it. It was explained that the laptop and mobile phone would in practice be the participant's own and that the "connection card" would have been offered by a member of staff in a venue such as a cafe. Each experimental trial began by compensating for different levels of knowledge of WiFi networks. The participant was given an instruction sheet which included a short explanation of the potential threats involved in unsecured WiFi use, specifically the MITM attack.

Participants were told: *There are two ways in which wireless connections are susceptible to attack: 1. Somebody might 'listen-in' to the wireless communications made between your laptop and a genuine access point to the Internet. 2.*

Somebody might have created an entirely fake network to which you can still connect, potentially giving away passwords or other information and inadvertently giving the faker access to your computer. The experiment then proceeded in two phases.

Phase 1: The participant was asked to evaluate the evidential value of the *PI* protocol in terms of genuineness and security. For the three conditions of *message choice* (presented in counterbalanced order), she connected to the *AvatarellisWireless* network and followed the *PI* protocol for whichever of the three conditions applied.

Phase 2: The participants rank-ordered the three conditions of *message choice* in terms of the degree to which they considered them as being *genuine* (**DV1**) and *secure* (**DV2**). Measures for each DV were collected separately using a graphical web-based interface that allowed participants to create a relative ranking score for each condition by clicking and dragging icons that represented each of the three conditions along a horizontal axis. A semi-structured interview was also conducted with the participant based around two questions:

- When I chose some or all of the words, my confidence that I was connected to the right network was: [less confident - more confident] (**DV1**).
- When I chose some or all of the words, the security of my connection to the network was: [less secure - more secure] (**DV2**).

4 Results

Rank scores for **DV1** and **DV2** fell in a range from 0 to 850. Prior to analysis, any two icons that overlapped were adjusted to be equal at their mid-point.

4.1 Effect of Message Choice upon Confidence in Network Genuineness

Genuineness scores were analysed across the three conditions using a one-way repeated measures analysis of variance (ANOVA). A significant main effect of *message choice* was observed [Wilks' Lambda = .507, $F(2,22) = 10.677$, $p<0.01$], and a multivariate partial eta squared value of .49 suggested a large effect size. Subsequent pairwise comparisons (repeated measures *t*-tests) were then performed on scores between successive condition pairs. There was a significant difference in *genuineness* scores between the *no choice* (M=450.04, SD=183.18) and *half choice* (M=670, SD=151.78) conditions; $t(23) = -4.31$, $p<0.01$. However, no significant difference was found between scores for the *half choice* and *full choice* conditions; $t(23) = -0.92$, $p=0.37$ n.s. Figure 2 shows mean scores for each level of *message choice*.

4.2 Effect of Message Choice upon Perception of Protocol Security

Security scores were analysed across the three conditions using a one-way repeated measures analysis of variance (ANOVA). A significant main effect of *message choice* was observed [Wilks' Lambda = .501, $F(2,22) = 10.960$, $p<0.01$], and

a multivariate partial eta squared value of .50 suggested a large effect size. Pairwise comparisons (repeated measures t-tests) performed on scores between successive condition pairs indicated a significant difference in security scores between the *no choice* (M=490, SD=142.44) and *half choice* (M=641.79, SD=134.32) conditions (t(23) = -4.60, $p<0.01$), but no significant difference between the *half choice* and *full choice* conditions (t(23) = -0.52, p=0.61 n.s.). Figure 2 shows mean scores for each level of *message choice*.

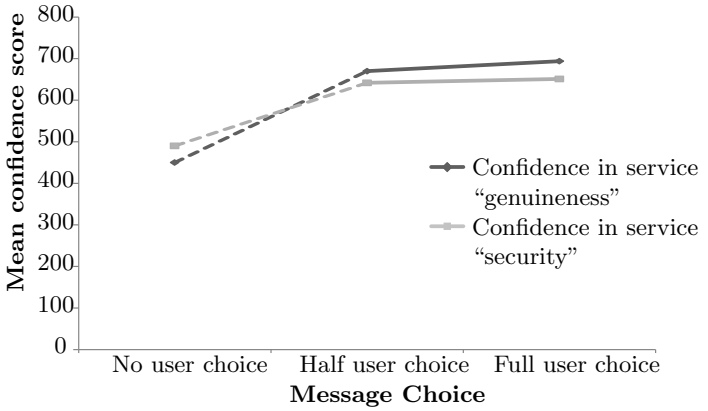

Fig. 2. Confidence in service *genuineness* and *security* across conditions of *message choice*. Significant pairwise comparisons (t-test) are highlighted with a dashed line.

5 Conclusion

Trust is an important factor in people's relationships with urban pervasive computing. On the one hand, users may mistakenly trust a malevolent service and therefore be vulnerable to attacks; on the other hand, people may distrust a bona fide service and not use it. Users' *perceptions* of the security and bona fides of urban pervasive services, and their trust in them, do not necessarily match the reality of a given service. This study investigated the effects on users' perceptions of a service's genuineness and security of allowing the user to choose the message used in the *Physical Interlock* device association protocol. We found that users were significantly more confident in the genuineness of the wireless network and the ability of *PI* to defend them against MITM if they contributed directly to the creation of the message. However, user creation of the *entire* message did not significantly affect user perceptions compared to user creation of *half* the message combined with system generation of the other half. Thus, the results of the study tend to support our hypotheses but that support is limited to a distinction between choice and no choice.

When users were allowed to choose the message, partially or completely, their confidence in the connection's genuineness and security rose. A user may be right

or wrong in thinking that she has done better than the system would have done. But while *perceived* security is rising, *actual* security could be rising or falling. The user's input might generate almost no entropy - or it might generate much more than the system choices (if the user has a vocabulary much larger than 1,055 words). As with so many aspects of computing, introducing humans into the loop causes uncertainty and complexity. When considering security strength, we typically want to know the weakest message the user could choose. We simply cannot know what actually happens to "real security" when we introduce user choice, except within very wide bounds. Thus, in applications where user memorization is not required, designers could consider using messages or passphrases which combine partial user generation with partial system generation since that may give people confidence in secure ad-hoc association protocols while ensuring a known lower bound on message strength.

References

1. Balfanz, D., Smetters, D., Stewart, P., Wong, H.: Talking to strangers: Authentication in ad-hoc wireless networks. In: Proc. NDSS 2002. pp. 7–19 (2002)
2. Burr, W.E., Dodson, D.F., Polk, W.T.: Electronic authentication guideline. NIST Special Publication 800, 63 (2004)
3. Kindberg, T., Bevan, C., O'Neill, E., Mitchell, J., Grimmett, J., Woodgate, D.: Authenticating ubiquitous services: a study of wireless hotspot access. In: Proc. Ubicomp 2009. pp. 115–124. ACM, New York, NY, USA (2009)
4. Kindberg, T., Mitchell, J., Grimmett, J., Bevan, C., O'Neill, E.: Authenticating public wireless networks with physical evidence. In: Proc. WIMOB 2009. pp. 394–399. IEEE Computer Society, Washington, DC, USA (2009)
5. Kindberg, T., Zhang, K.: Secure spontaneous device association. UbiComp 2003: Ubiquitous Computing pp. 124–131 (2003)
6. Kobsa, A., Sonawalla, R., Tsudik, G., Uzun, E., Wang, Y.: Serial hook-ups: a comparative usability study of secure device pairing methods. In: Proc. SOUPS 2009. pp. 1–12. ACM, New York, NY, USA (2009)
7. Kumar, A., Saxena, N., Tsudik, G., Uzun, E.: A comparative study of secure device pairing methods. Pervasive Mob. Comput. 5, 734–749 (December 2009)
8. Perrig, A., Song, D.: Hash visualization: A new technique to improve real world security. In: Proc. Cryptographic Techniques and E-commerce 1999 (1999)
9. Rivest, R., Shamir, A.: How to expose an eavesdropper. Communications of the ACM 27(4), 393–394 (1984)
10. Shannon, C.E.: Prediction and entropy of printed english. Bell System Technical Journal 30(1), 5064 (1951)
11. Soriente, C., Tsudik, G., Uzun, E.: Beda: Button-enabled device association (2007)
12. Soriente, C., Tsudik, G., Uzun, E.: Hapadep: Human-assisted pure audio device pairing. In: Information Security, pp. 385–400. Springer Berlin / Heidelberg (2008)
13. Stajano, F., Anderson, R.: The resurrecting duckling: Security issues for ad-hoc wireless networks. pp. 172–194. Springer-Verlag (1999)