Secure and Usable Out-Of-Band Channels for *Ad hoc* Mobile Device Interactions

Ronald Kainda, Ivan Flechais, and A.W. Roscoe

Oxford University Computing Laboratory
Wolfson Building, OX1 3QD Oxford
{ronald.kainda,ivan.flechais,bill.roscoe}@comlab.ox.ac.uk

Abstract. Protocols for bootstrapping security in *ad hoc* mobile device interactions rely on users' ability to perform specific tasks such as transferring or comparing fingerprints of information between devices. The size of fingerprints depends on the level of technical security[1] required by a given application but, at the same time, is limited by users' inability to deal with large amounts of data with high levels of accuracy. Large fingerprints provide high technical security but potentially reduce usability of protocols which may result in users making mistakes that compromise security. This conflict between technical security and usability requires methods for transferring fingerprints between devices that maximise both to achieve acceptable effective security. In this paper, we propose two methods for transferring fingerprints between devices. We conducted a usability and security evaluation of the methods and our results show that, in contrast to previous proposals, our methods are both usable and resistant to security failures.

Key words: Security, Usability, Device pairing, Out-Of-Band channel

1 Introduction

Protocols for bootstrapping security in *ad hoc* mobile device interactions rely on users to transfer fingerprints between devices. We call these protocols *Human-Interactive Security Protocols* (HISP). In these protocols, the level of security achieved may depend on the amount of human effort expended. There is, therefore, a need for balancing between theoretical security requirements and usability of HISP.

HISP use two channels; a high-bandwidth (normal) channel such as Bluetooth and WiFi and a low-bandwidth Out-Of-Band (OOB) channel. Limited bandwidth on OOB channels means that they are not suitable for transferring large amounts of data such as cryptographic keys. Cryptographic keys and other public information is transmitted over the normal channel while only limited information such as fingerprints of cryptographic keys is transmitted over OOB channels. Messages on the normal channel can be overheard, deleted, or modified

[1] This is the theoretical security of a protocol or system based on mathematical proofs.

by an attacker. On the other hand an attacker has no control over messages on the OOB channel.

Technical security of HISP is such that the odds of being successfully attacked are 1 in 2^b , where b is the size of the fingerprint. Choosing the value of b that is sufficiently large for a given application gives guarantees on technical security. However, these protocols require users to deal with fingerprints. Users' actions bring about threats that are not covered by mathematical assertions. For example, will users compare fingerprints accurately? Will they bother to compare and not skip this step? Can they be duped into associating devices whose fingerprints do not match? How large a value of b can they effectively deal with? In this paper we propose two OOB methods that maximise usability without compromising security. The paper is organised as follows; we discuss current proposals of OOB channels and their limitations in Section 2, and present our proposed methods in Section 3. A usability study and results are presented in Section 4, followed by a discussion of some of the application scenarios for our methods in Section 5 and conclude in Section 6.

2 Current OOB Channels and Limitations

Current OOB channels can be grouped into 5 main categories: compare and confirm, copy and enter, auxiliary device methods, short range directed channels, and timing methods. *Compare and confirm* requires users to compare fingerprints displayed on two or more devices and indicate a match or otherwise on participating devices while *copy and enter* requires users to copy a fingerprint displayed on one device to other participating devices. Auxiliary device methods depend on secondary devices, e.g. data cable, to transfer or compare fingerprints between devices. Short range directed channel methods use wireless transmission technologies to transfer information between devices. Examples include infra-red [1, 2] and light [3, 4]. Timing methods rely on transmission of information within well timed intervals such that an intruder finds it hard to synchronise and successfully attack the association. Proposed methods include shaking devices [5] and pressing buttons in response to stimulus [6].

2.1 Limitations

A Choice Between Security and Usability Currently proposed OOB channels force one to choose between security and usability. *Compare and confirm* has been found both a usable and preferred method but also susceptible to security failures [7–9]. Other methods such as timing and auxiliary device methods are also susceptible to security failures though they do not offer the same level of usability as *compare and confirm*. On the other hand, *copy and enter* is not susceptible to security failures but requires users to type at least 6 digits. This is particularly problematic in applications where a user needs to carry out device associations several times a day. We feel that one does not have to choose between usability and security but both have to be embedded in an OOB channel.

Limited Application Even though mobile interactions are ephemeral and both scenario and context may change from one interaction to the next, current proposals seem to focus on one scenario and context. For example, timing methods and short range directed channels may work well in pairwise associations but difficult to extend to group scenarios or where devices are a certain distance apart. In addition, the range of affordances on mobile devices is diverse ranging from netbooks and smart phones to Bluetooth hands-free sets. Current proposals of OOB channels focus on devices of similar affordances such as cameras, accelerometers, laser beamers and readers. There is anecdotal evidence that shows that majority of *ad hoc* mobile interactions occur either between devices of sufficient input/output capabilities or between an input/output rich device and an input/output constrained device. With this in mind, it is possible to exploit devices with rich input/output interfaces to leverage limitations of input/output constrained devices. Current proposals, however, fail to take this into account which has led to an assumption that most mobile interactions occur between devices with poor input/output interfaces. We feel that, because of changing context and scenarios, OOB channels should be general purpose to cover as many scenarios and contexts as possible and, by taking advantage of the fact that one device in the association is likely to have rich input/output interfaces, better methods that also apply to interface constrained devices can be developed.

Unmotivated User Property Users tend to avoid security tasks because they are not their primary goals in most scenarios. While OOB channels rely on users to accurately perform critical security tasks, many do not compel users to perform required tasks. For example, in *compare and confirm*, users can ignore comparing but accept that the fingerprint is matching. Even vigilant users may be distracted, accidentally press a wrong button, or merely miss the comparison. In fact, in their study of pairing methods, Kainda *et al.* [7] reported that some participants were concerned that with *compare and confirm* it was easy for a security failure to occur because the method did not force them to compare accurately.

3 Proposed Methods

The methods of *compare and confirm* and *copy and enter* seem to fall on two opposite ends of a spectrum. *Compare and confirm* is usable but subject to security failures where as *copy and enter* is secure but demands more effort from users. Our methods are aimed at balancing security and usability of HISP to achieve acceptable effective security. We discuss our proposed methods below.

3.1 Word-Matching and Number-Typing

Word-matching and number-typing is based on the fact that *copy and enter* is not subject to security failures but is regarded as difficult to use. While earlier

work has argued that typing short strings on devices with limited input interfaces is hard for most users, the popularity of Short Message Service (SMS) is an indication that users are comfortable with such a task. We are, however, cognisant of the lack of motivation from users to type strings for the sake of security. *Word-matching and number-typing* is, therefore, aimed at offering the same level of security as *copy and enter* but only requiring users to type a smaller number of digits.

Word-matching and number-typing uses two locally stored dictionaries of words that are phonetically distant [10]. Associating devices generate words representing a fingerprint using a method similar to S/Key one time password [11]. One device displays a set of words with numbers assigned to them in increasing order starting from 1. Other devices randomly display one of the locally generated words and prompt a user to enter the number assigned to that particular word on the other device. The user types the digit and the device displays the next word until all the words have been displayed. The device then checks if the assigned numbers are correct and informs user of the result.

In comparison with *copy and enter*, our method requires a user to type 2 digits for a fingerprint of 20 bits (for 1024 word-dictionaries) while the former requires about 7 digits to achieve the same level of security. Our method makes it easier to transfer reasonable amount of information with limited human effort to achieve high levels of security. To transfer 40 bits, for example, our method requires typing 4 digits while *copy and enter* requires typing about 12 digits.

Potential problems to *word-matching and number-typing* include display of duplicate words and user prediction of the digit to be typed for the last word on the list. We counter the first problem by using two dictionaries, similar to PGPfone proposal [10]. This ensures that when two consecutive bit sequences are similar, two different words will be produced as the first bit sequence will result in a word from the first dictionary and the second bit sequence in a word from the second dictionary. In current protocols, two words are sufficient. The second problem is due to the fact that once the user has entered the first word, she knows that only one other word remains and hence will have no motivation to check if that word actually exits. To counter this threat, an extra word is displayed together with a fingerprint but devices ask users to enter positions of words that correspond to the fingerprint only.

3.2 Repeated Numeric Comparison

Compare and confirm lacks the ability to compel users to compare fingerprints accurately. Similar to pop up boxes prompting users to either accept an 'invalid' certificate or not, users will figure out that by pressing a button that indicates acceptance that fingerprints match they will be able to accomplish their primary task.

To compel users to carry out comparison without undue effort, *repeated numeric comparison* is a two step process. In addition to a fingerprint, an authenticating device generates a random string of similar format as the fingerprint. The authenticating device then randomly chooses to display either its fingerprint or

the random value. The user compares and indicates whether the string displayed on the authenticating device matches that on the other device. An authenticating device then displays the remaining string and the user does the comparison again.

There are 4 possible outcomes from the user's input. Indicating MM (Matching-Matching) for both strings, DD (Different-Different) for both strings, DM, and MD. The authenticating device will only accept one input; M for a fingerprint value and D for random string. A user, however, does not know which one is the fingerprint and which one is the random string hence is forced pay attention to comparing otherwise device association fails.

4 Usability Evaluation of the Methods

To evaluate usability and security of the two methods, we conducted a laboratory experiment. The aim of the study was not only to evaluate how usable the methods are but also compare our results with those previous studies. To this regard, we designed our laboratory experiment similarly to [7–9].

4.1 Experimental Design

We used a repeated measure design, using counterbalancing, with 28 participants (12 male and 16 female) of varying age, education and professional background. Participants were recruited via web advertisement and mailing lists. We asked participants to participate in a quiz, using two mobile phones, where they were first required to establish a secure connection. We developed the quiz application and device association interfaces using Java 2 Micro Edition (J2ME). The application also logged participant's actions. Other data were collected through questionnaires and discussions. After Scenario Questionnaires ASQ [12] (a 3 question, seven point Likert-type questionnaire) were completed by each participant after each method. Also End of Experiment Questionnaires, where participants indicated which of the two methods they thought was easy and which one they felt was difficult, were used. In line with previous studies, we also asked participants which method they preferred between the two.

4.2 Results

Word-matching and number-typing had mean completion time of 12.7 seconds and 3.6% non-security failures while *repeated numeric comparison* had mean completion times of 13.4 seconds and 7% non-security failures. Similar to *copy and enter*, security failures are difficult to simulate for both methods.

A statistical test using paired t-test showed no statistical significance between the methods in completion times at 95% confidence interval with $t(55) = 0.53$, $p = 0.598$. In terms of user ratings, 93% of participants indicated that *word-matching and number-typing* is ease to use compared to 89% for *repeated numeric comparison*. In addition 57% of participants preferred *word-matching*

and number-typing compared to 25% who preferred *repeated numeric comparison.*

Participants also rated each method on the three main elements of usability—efficiency, effectiveness, and satisfaction—on a 7 point scale. For each participant, we summed and averaged scores for each method. As this data is ordinal, we only report median, mode, and a Wilcoxon test statistic. *Word-matching and number-typing* had a mode of 7 and median of 6.5 (min = 2, max = 7) while *repeated numeric comparison* had a score of 7 (min = 4, max = 7) for both the mode and median. A Wilcoxon test showed that there is no statistical significance in ratings between the methods (Z = -0.275 and p(2-tailed) = 0.78).

In the post study discussion, participants indicated that the methods were similar in terms of ease-of-use and preferences were just for one's 'taste' and not to do with how the method is used. There was, however, concern over *repeated numeric comparison* from two participants who are dyslexic—a condition that affects one's ability to deal with digits.

4.3 Comparison with Earlier Methods

We compare our results to that of Uzun *et al.* [9] and Kainda *et al.* [7]. A common result in these studies is that *compare and confirm* had the lowest completion time and ranked as most usable. *Copy and enter* was found as most preferred (personal choice) in Uzun's work while *compare and confirm* was found to be the most preferred in Kainda's work. We compare our methods to results from these two studies on three crucial issues: efficiency (completion times), satisfaction, and security.

Efficiency: With an additional step in *repeated numeric comparison*, we did not expect participants to complete an association process using the method in a time comparable to previous studies. To our surprise, participants completed the process with an average time of 13.4 seconds compared to 16.4 seconds reported in Uzun's work. In additional, while users have to look up a word on the list before typing its position (in *word-matching and number-typing*), completion times were not affected with a reported average of 12.7 seconds compared to 13 seconds in Uzun's work.

User rating: We used various methods through which participants rated our methods. Our methods were rated higher in terms of ease of use, 93% for *word-matching and number-typing* and 83% for *repeated numeric comparison*, compared to *compare and confirm's* 40-45% and *copy and enter's* 20-37% in previous studies. On a rating scale of 1 (worst) to 7 (best), both our methods were rated above 6.5 while in Kainda's study both *compare and confirm* and *copy and enter* were rated below 6.3.

Security: Unlike *compare and confirm*, both our methods are resistant to security failures. While *copy and enter* is resistant to security failures as well, it does not offer room for larger fingerprints—typing 6 digits is already deemed difficult to use—hence limited in its application. *Word-matching and number-typing* offers room for larger fingerprints than currently possible by both *compare and confirm* and *copy and enter.* For example, for the same amount of human

effort (typing 6 digits), *word-matching and number-typing* can offer security of 2^{60} entropy while *copy and enter* can only offer 2^{20}.

5 Applications of Proposed Methods

We highlight some of the scenarios and contexts in which our methods may be applied.

Close/distant devices: With devices close to each other users can directly read a fingerprint off the screen of a device displaying it. In a scenario where devices involved are far apart, we assume that there is a secondary channel (e.g. audio/video channel, SMS etc) through which the fingerprint will be communicated. For example, a user at one end may read out a fingerprint for another user via telephone.

Input/output constrained devices: Common device association scenarios involve a device with rich input/output interfaces and an input/output constrained device. For both our method, a display and a single button is sufficient for interface constrained devices. In *word-matching and number-typing*, a device with rich interfaces will display 3 words and user will be required to press the button 3 times consecutively to enter the numeric value 3 , for example. Since 2 words provide sufficient entropy for security, users will not be required to enter numbers higher than 3. Similarly, in *repeated numeric comparison* users will be required to press the button in a particular way, 3 seconds press for example, to accept that fingerprints match and a short press to reject.

Group associations: In a group association scenario, e.g. meeting, a user with a device displaying a fingerprint can read it out to other users in the group. In *word-matching and number-typing*, words in the dictionary should be phonetically distant to avoid confusion between similar sounding words.

6 Conclusion

In secure systems, it is crucial to balance usability and security. The problem we have encountered in device association methods is that on one hand we have methods that are usable but are subject to security failures while on the other we have methods that are secure but are perceived as difficult to use. It is therefore crucial to do one of the two: improve security of usable but insecure methods or improve usability of secure but difficult to use methods. In this paper, we have attempted to do both and compared our results with previous findings. Our findings show that our proposed methods provide the needed security in device association tasks without putting undue workload on the users.

References

1. Balfanz, D., Smetters, D.K., Stewart, P., Wong, H.C.: Talking to strangers: Authentication in ad-hoc wireless networks. In: In Symposium on Network and Distributed Systems Security (NDSS '02), San Diego, California. (2002)

2. Feeney, L.M., Ahlgren, B., Westerlund, A.: Demonstration abstract: Spontaneous networking for secure collaborative applications in an infrastructureless environment. In: International conference on pervasive computing (Pervasive 2002). (2002)
3. Mayrhofer, R., Welch, M.: A human-verifiable authentication protocol using visible laser light. In: ARES '07: Proceedings of the The Second International Conference on Availability, Reliability and Security, Washington, DC, USA, IEEE Computer Society (2007) 1143–1148
4. Saxena, N., Ekberg, J.E., Kostiainen, K., Asokan, N.: Secure Device Pairing based on a Visual Channel (Short Paper). In: SP '06: Proceedings of the 2006 IEEE Symposium on Security and Privacy, Washington, DC, USA, IEEE Computer Society (2006) 306–313
5. Mayrhofer, R., Gellersen, H.: Shake well before use: Authentication based on Accelerometer Data. In: Proc. Pervasive 2007: 5th International Conference on Pervasive Computing. Volume 4480 of LNCS., Springer-Verlag (May 2007) 144–161
6. Saxena, N., Uddin, B., Jonathan, V.: Universal Device Pairing Using an Auxiliary Device. In: Symposium on Usable Privacy and Security (SOUPS). (July 2008)
7. Kainda, R., Flechais, I., Roscoe, A.: Usability and Security of Out-Of-Band Channels in Secure Device Pairing Protocols. In: SOUPS '09: Proceedings of the 5th symposium on Usable privacy and security. (2009)
8. Kobsa, A., Sonawalla, R., Tsudik, G., Uzun, E., Wang, Y.: Serial Hook-ups: A Comparative Usability Study of Secure Device Pairing Methods. In: SOUPS '09: Proceedings of the 5th symposium on Usable privacy and security. (2009)
9. Soriente, C., Tsudik, G., Uzun, E.: BEDA: Button-Enabled Device Association. In: In International Workshop on Security for Spontaneous Interaction (IWSSI). (2007)
10. Juola, P.: Whole-word Phonetic Distance and the PGPfone Alphabet. In: Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on. Volume 1. (Oct 1996) 98–101 vol.1
11. Haller, N.M.: The S/KEY one-time password system. In: Proceedings of the Symposium on Network and Distributed System Security. (1994) 151–157
12. Lewis, J.R.: IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. Int. J. Hum.-Comput. Interact. **7**(1) (1995) 57–78