Towards Device-to-User Authentication: Protecting Against Phishing Hardware by Ensuring Mobile Device Authenticity using Vibration Patterns

Rainhard Dieter Findling
JRZ u'smile
University of Applied Sciences Upper Austria
Softwarepark 11
A-4232 Hagenberg, Austria
rainhard.findling@fh-hagenberg.at

Rene Mayrhofer
JRZ u'smile
Johannes Kepler University Linz
Altenbergerstrasse 69
A-4040 Linz, Austria
rene.mayrhofer@jku.at

ABSTRACT

Users usually authenticate to mobile devices before using them (e.g. PIN, password), but devices do not do the same to users. Revealing the authentication secret to a non-authenticated device potentially enables attackers to obtain the secret, by replacing the device with an identical-looking malicious device. The revealed authentication secret could be transmitted to the attackers immediately, who then conveniently authenticate to the real device. Addressing this attack scenario, we analyze different approaches towards mobile device-to-user (D2U) authentication, for which we provide an overview of advantages/drawbacks, potential risks and device authentication data bandwidth estimations. We further analyze vibration as one D2U feedback channel that is unobtrusive and hard to eavesdrop, including a user study to estimate vibration pattern recognition using a setup of $\sim$7 bits per second (b/s). Study findings indicate that users are able to distinguish vibration patterns with median correctness of 97.5% (without taking training effects into account) – which indicates that vibration could act as authentication feedback channel and should be investigated further in future research.

ACM Classification Keywords

D.4.6. Security and Protection: Authentication; H.1.2 User/Machine Systems: Human factors

Author Keywords

Phishing hardware; mobile authentication; vibration; feedback.

INTRODUCTION

Nowadays, personal mobile devices have access to a vast amount of private data and services worth protecting. Therefore, mobile devices are usually locked when idle and have to be unlocked before usage (e.g. by entering a PIN, password or unlock pattern). Although users frequently authenticate to devices before usage, devices usually don't do the same to users (e.g. revealing a shared secret to users, so that they are assured that the device is in fact the correct one). This potentially enables attackers to obtain the unlocking secret by using identical-looking, but malicious devices. In case users try to authenticate with such a malicious device (which is under the remote control of attackers), their authentication secret would be revealed and could be transmitted to attackers immediately. If the attackers have brought the original device under their control previously, the obtained secret could be used to further unlock the real device.

As such an attack using malicious devices works by deceiving users (similar to web-based phishing attacks), we refer to these attacks as *phishing hardware attacks*, and to the devices simply as *phishing hardware*. As with all phishing attacks – including mobile phishing hardware attacks – the malicious instance just needs to mock the real instance until authentication credentials have been revealed. It is already too late if users recognize moments later that they are interacting with a wrong device – especially, if the real device is already out of their reach.

Additional issues are that a) virtually all mobile device models are strongly standardized (including possible customizations in software and look-and-feel), and identical copies of all these models can be easily obtained by attackers; b) phishing hardware attacks cause devices to be swapped – hence for attackers there is no loss in terms of hardware; and c) individual/personal customization (e.g. screen wall paper, sounds, even hardware customizations as stickers on the device) could as well be duplicated easily by attackers for the mock device. Obtaining information about the target device and its features for creating phishing hardware can be done by attackers without physical access to the device, e.g. by inconspicuously taking pictures of the phone (e.g. while it is lying on a table).

Mobile devices authenticating to their users (as users do to mobile devices) would be an effective measure against such attacks. We therefore focus on device-to-user (D2U) authentication for frequent and "everyday" usage, specially for mobile devices, which attackers could replace more easily with phishing hardware than traditional computers. In this paper, we provide an overview of possible D2U authentication concepts, including vibration, which – in comparison to other potential channels – is harder to eavesdrop. Summarizing, the contributions of this paper are:

MUM'15, November 30–December 02, 2015, Linz, Austria.
Copyright © 2015 ACM ISBN/14/04...$15.00.
DOI string from ACM form confirmation

- We provide an overview of possible D2U authentication approaches and compare their advantages and drawbacks, including estimated bandwidth and possible risks.

- We analyze vibration as one such D2U feedback channel in detail, including a user study on how well vibration patterns can be distinguished.

DEVICE-TO-USER AUTHENTICATION

Mutual authentication principles (both parties authenticating with each other) are well established in machine-to-machine (M2M) communication, such as web technologies (e.g. TLS [10]). In contrast to M2M authentication, user-to-device (U2D) or device-to-user (D2U) authentication is limited by certain human factors, such as cognitive load (the effort and difficulty of remembering long and complex secrets), limited channel bandwidth (exchange of larger portions of information takes longer), and computational limitations (e.g. cryptographic mathematics, which humans can hardly do without aid of computers).

As with most security mechanisms involving human factors, on the one hand, these limitations result in a trade-off between security and usability: increasing security decreases usability and vice versa [2]. On the other hand, with D2U authentication currently de facto not being employed on mobile devices, even authentication approaches focusing on usability at the cost of security will cause a security gain.

Related Work

One approach to D2U authentication is by devices visually revealing secret information to users to ensure they can be trusted. One example are web-based banking systems, where after logging in, users are presented a previously defined secret to ensure authenticity of the service they are interacting with. Another example is displaying variations of secret images to the user to assure authenticity of user interfaces and computer systems [12, 13]. The main drawback of such approaches is being prone to shoulder surfing attacks (an attacker visually observing secret information revealed to the user by the device – without requiring physical access to the device).

Other related approaches deal with human verifiable authentication on pairing devices (e.g. Bluetooth pairing in general [4]) or pairing of devices with restricted in- and output capabilities (e.g. pressing a button on device A in the same pattern a LED blinks on device B [7] or shaking devices together [9]). In contrast to these mechanisms, which are intended to be employed once during device pairing (hence, reduced usability is experienced only once and the risk of e.g. being shoulder surfed can be avoided by additional effort), our approach is intended to be used frequently. Consequently, usability drawbacks through additional effort would impact users more frequently.

Device-to-User Authentication Approaches

Combining capabilities of current mobile devices and human sensing, different D2U authentication approaches seem possible (see table 1). All of them could be employed standalone or merged into a single hybrid approach. Further, all of these could be used for the device revealing authentication information to the user before, during, or after the user authenticates to the device.

	See	Hear	Feel	Smell	Taste
Visual	+	-	-	n.a.	n.a.
Sound	-	+	-	n.a.	n.a.
Vibration	-	o	+	n.a.	n.a.

Table 1: Possible D2U authentication approaches with strong (+), weak (o) and few/no correlation (-) with human sensing capabilities.

Visual

One obvious D2U authentication is to show authentication information visually, e.g. on the mobile device display. Notification elements could be used as well (e.g. the LED usually indicating the reception of messages or calls). While displays feature higher channel bandwidth, notification elements could show information even when the screen is off (which does not seem to be an advantage in terms of security). Similar to the concept of showing a secure authentication image to the user [12, 13], this approach is prone to shoulder surfing.

Sound

Analogous to using visual information, authentication information can be revealed via sound. For example, HAPADEP [15] uses a human recognizable MIDI codec transporting 240 bits of information in 3.4 s (∼70 b/s), which seems sufficient for D2U authentication tasks. Similarly to visual approaches, sound is prone to attackers observing authentication information without physical access to the device.

Vibration

Information emitted by device vibrators can conceptually be observed by a) feeling the vibration and b) hearing noise caused by vibrators – given a quiet environment. In contrast to previous concepts, vibration cannot be visually observed by attackers, which is a valuable advantage in terms of security. It further is unobtrusive as users don't need to look at the screen or have to listen to sounds in a possibly noisy environment [1]. A drawback is attackers potentially being able to observe vibration pattern sounds in quiet environments. While this could be exploited to obtain secret information, it is likely still more complicated than e.g. overhearing authentication via dedicated sound or observing secret information displayed on mobile device screens via shoulder surfing. We are currently not aware of any research stating channel bandwidth of users distinguishing vibration patterns. This is, together with its favorable security properties, why we conduct a user study on evaluating how well preliminary vibration patterns can be recognized by users.

Interlock authentication

For all mentioned possible D2U authentication channels, there exist multiple variants of how to integrate D2U authentication with U2D authentication. The first is to have the device authenticate to the user before the user authenticates to the device. On the one hand, this ensures users that it is the correct device they are revealing their authentication secret to. On the other hand, in case attackers get physical access to the device (without being aware of the user authentication secret, so they cannot unlock the device), they would be able to observe the D2U authentication secret – and could later mock it too, using a phishing hardware device. If instead the user authenticates to the device first, and afterwards the device to the user, phishing hardware attacks are possible, as the device only authenticates after the user authentication secret has been fully revealed.

A more promising variant would be using the interleaving “interlock” information exchange [6, 11] to integrate user-to-device and D2U authentication. Interleaving authentication information is well known and in active use in a variety of areas (e.g. to prevent different types of attacks on network communication and key exchange protocols [8]). Interleaving could start with the device revealing the first authentication part to the user, right before the user starts authentication to the device (e.g. when the screen is turned on). Successive parts would be revealed only if the user enters correct authentication information. Here, the difficulty could again lie with the human factor: users experience a potentially increased authentication effort and are required to stop entering further authentication information to the device, if the device does not reveal itself as their trusted device.

VIBRATION PATTERN RECOGNITION

There exist several studies of M2M communication using mobile device vibration as communication channel, which state the channel bandwidth in the range of 10s b/s [17] to 100s b/s [1, 3, 14]. In contrast to M2M communication, at the time of writing, we are not aware of any vibration channel bandwidth analysis that involves humans and devices (e.g. how much information a human can possibly extract from machine vibration patterns). We therefore conduct a user study to evaluate how well a preliminary vibration code for D2U authentication can be correctly recognized by participants.

Preliminary Vibration Code

The main limitation of vibration for user friendly D2U authentication is duration: if authentication takes noticeably longer when incorporating device authentication, the vibration feedback will possibly not be employed by users. As mobile U2D authentication usually takes in the range of 1.5–3.5 s (depending on the employed unlocking approach) [5, 16], we restrict ourselves to a window of this size. For example, using a 4 digit PIN for user authentication with an estimated duration of 2 s would result in revealing the next digit to the device about every 0.5 s. This 0.5 s window could be used to reveal a part of the D2U authentication information via vibration. Based on these limitations and a preceding, preliminary study on which vibration types and timings are easy to be distinguished correctly, a prototypical vibration test code was derived. Consequently, with more in-depth insights to human vibration pattern recognition capabilities this code (and its bandwidth) could likely be improved.

The preliminary vibration code contains 1–2 groups of vibrations, with each group consisting of up to 3 single vibrations (see figure 1). The second group in allowed to be empty (containing no vibrations), while the first group must contain at least one vibration. This results in our test code being able to transport one of a total of $3 \cdot 4 = 12$ different patterns per transmission. Vibration and pauses between vibrations of the same group are of 60 ms duration. Pauses between vibrations of different groups are of 200 ms duration. This setup results in an average pattern duration of 465 ms, which would be within the hypothetic 0.5 s time frame for feedback with a 4 digit PIN entered in 2 s – and which results in a bandwidth of ~7.7 b/s. Subsequently, we depict patterns as the amount of vibrations contained in each group, e.g. “3 2” for the first group containing 3, the second 2 vibrations, or “2” the first group containing two vibrations and the second being empty.

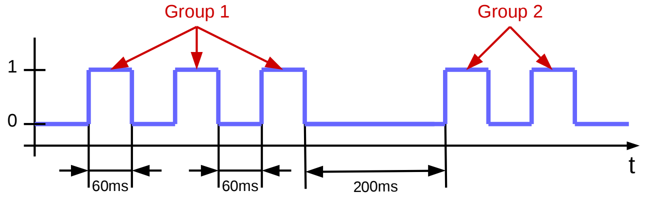


Figure 1: Example pattern “3 2” from our preliminary vibration code, with 0 and 1 indication no vibration and vibration, respectively.

Vibration Pattern Recognition Study Setup

The preliminary vibration code has been implemented in an Android application[1] for the successive user study. The application features two modes: in trial mode, users can trigger all different vibrations as they wish and learn how they feel. In test mode, users are assigned a randomly chosen vibration pattern and have to decide for further, also randomly chosen vibration patterns, if this was their assigned pattern.

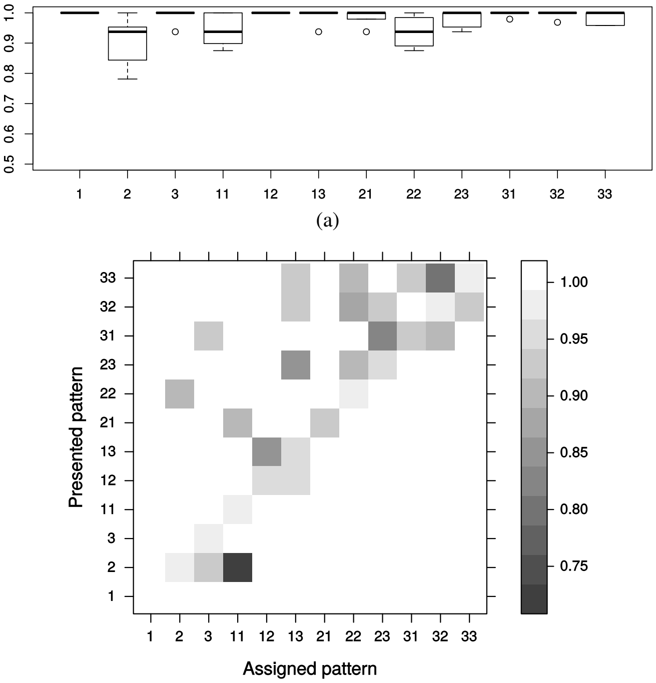
12 people participated in the study and were allowed to try out the application in trial mode as long as they wished. Each participant did at least 12 vibration patterns recognition sets in test mode, where for each test set they were assigned a random pattern and had to decide for 16 further random patterns (which they could trigger only once), if it was their assigned pattern. The probability of the test pattern being the assigned pattern was set to $\frac{5}{16}$. This setup resulted in 898 and 1614 recognitions of assigned and non-assigned patterns, respectively[2].

Vibration Pattern Recognition Results

Vibration pattern recognition rates over all users (see figure 2a) indicate that our vibration patterns can successfully be distinguished. There seems to be no trend of shorter or longer patterns being recognized correctly with higher probability. Instead, recognition correctness involving vibration patters “2”, “1 1” and “2 2” seem to be lower, compared to recognition not involving these patterns. The distribution of recognition correctness over assigned and presented patterns (see figure 2b) indicates their assigned patterns are likely recognized correctly

[1] The application code is open source and publicly available at https://github.com/mobilesec/device-to-user-authentication-vibration-bandwidth.

[2] Detailed study results are publicly available at https://www.usmile.at/downloads/.



(a)

(b)

Figure 2: Participants' recognition rates of vibrations patterns as a) recognition correctness per codes involved and b) distribution of recognition correctness over assigned and presented codes.

if presented to users (median correctness of 97.5%). It further indicates that correctly recognizing non-similar patterns as being different is even likelier (e.g. patterns such as "1" and "3 3" have been distinguished without a single error). But it also indicates, that there is a tendency of users to incorrectly recognize non-assigned patterns as their assigned ones, if patterns are similar. For example, pattern "2" and "2 1" have frequently been mis-recognized as "1 1" (error rates of 27% and 9%), pattern "1 3" as "1 2" (15%), pattern "2 3" as "1 3" (14%), or pattern "3 2" as "3 3" (20%). The resulting median recognition rate over all assigned and non-assigned patterns is 97.5%.

Despite these errors, using our preliminary vibration code with an average bandwidth of ∼7.7 b/s, and our results showing an median successful vibration pattern distinguishing rate of 97.5%, we infer that vibration patterns could serve as valuable D2U authentication channel.

After finishing the study, about 50% of participants stated that they believed they used hearing vibration patterns in combination with feeling them to decide if it was their assigned pattern. This indicates that hearing and feeling are used together for recognizing vibration patterns. Consequently, future research should investigate human vibration pattern recognition capabilities from only feeling patterns (e.g. with suppressing vibration sounds for participants or having them listening to music), as well as from only hearing patterns. Although the latter represents the scenario of attackers possibly being able to overhear secret vibration authentication information, we argue that this is likely still more complicated than e.g. overhearing dedicated sound or observing secret information displayed on mobile devices via shoulder surfing.

CONCLUSION

In this paper we tackled the problem of mobile devices not authenticating to their users, which could be exploited to steal the user authentication secret using phishing hardware attacks. We discussed different possible D2U authentication concepts – and found vibration to be specially interesting for mobile devices, as it cannot be observed visually by attackers. Consequently, we evaluated how well users are able to distinguish vibration patterns on mobile devices, using a preliminary vibration code of ∼7.7 b/s in a user study. Results show that users were able to distinguish vibration patterns with average correctness of 97.5%. Further, they confirm intuition that patterns observed as being more similar to each other also seem harder to be distinguished correctly. From these findings we conclude that vibration could act as valuable and potentially hard-to-eavesdrop D2U authentication feedback channel. Participants further stated they used hearing vibration patters too to decide if it was their assigned pattern. Consequently, future research should investigate i.a. human vibration pattern recognition capabilities by only hearing or feeling them – with the latter representing a possible attack scenario of attackers in quiet environments observing vibration authentication information by hearing it. Therefore, future research should investigate the design of robust and distinguishable vibration patterns as well as secure exchange of information between users and devices, with the whole process being user verifiable and user friendly.

ACKNOWLEDGMENTS

This work has been carried out within the scope of *u'smile*, the Josef Ressel Center for User-Friendly Secure Mobile Environments, funded by the Christian Doppler Gesellschaft, A1 Telekom Austria AG, Drei-Banken-EDV GmbH, LG Nexera Business Solutions AG, NXP Semiconductors Austria GmbH, and Österreichische Staatsdruckerei GmbH.

REFERENCES

1. Joshua Adkins, Genevieve Flaspohler, and Prabal Dutta. 2015. Ving: Bootstrapping the Desktop Area Network with a Vibratory Ping. In *The 2nd ACM Workshop on Hot Topics in Wireless (HotWireless'15)*. Paris, France.

2. Lorrie Faith Cranor and Simson Garfinkel. 2008. *Security and Usability*. O'Reilly Media.

3. M. Hansen, R. Hill, and S. Wimberly. 2012. Detecting covert communication on Android. In *Local Computer Networks (LCN), 2012 IEEE 37th Conference on*. 300–303. DOI: http://dx.doi.org/10.1109/LCN.2012.6423634

4. Robin Heydon. 2012. *Bluetooth Low Energy: The Developer's Handbook*. Prentice Hall.

5. Daniel Hintze, Rainhard Dieter Findling, Sebastian Scholz, and René Mayrhofer. 2014. Mobile Device Usage

Characteristics: The Effect of Context and Form Factor on Locked and Unlocked Usage. In *Proc. MoMM 2014: 12th International Conference on Advances in Mobile Computing and Multimedia*. ACM Press, New York, NY, USA, 105–114. DOI: http://dx.doi.org/10.1145/2684103.2684156

6. Tim Kindberg, Chris Bevan, Eamonn O'Neill, James Mitchell, Jim Grimmett, and Dawn Woodgate. 2009. Authenticating Ubiquitous Services: A Study of Wireless Hotspot Access. In *Proceedings of the 11th International Conference on Ubiquitous Computing (UbiComp '09)*. ACM, New York, NY, USA, 115–124. DOI: http://dx.doi.org/10.1145/1620545.1620565
7. M. Long and D. Durham. 2007. Human Perceivable Authentication: An Economical Solution for Security Associations in Short-Distance Wireless Networking. In *Computer Communications and Networks, 2007. ICCCN 2007. Proceedings of 16th International Conference on*. 257–264. DOI: http://dx.doi.org/10.1109/ICCCN.2007.4317829
8. Rene Mayrhofer and Hans Gellersen. 2007. On the Security of Ultrasound as Out-of-band Channel. In *Parallel and Distributed Processing Symposium, 2007. IPDPS 2007. IEEE International*. IEEE, 1–6.
9. R. Mayrhofer and H. Gellersen. 2009. Shake Well Before Use: Intuitive and Secure Pairing of Mobile Devices. *Mobile Computing, IEEE Transactions on* 8, 6 (2009), 792–806. DOI:http://dx.doi.org/10.1109/TMC.2009.51
10. Eric Rescorla. 2000. *SSL and TLS: Designing and Building Secure Systems*. Addison-Wesley Professional.
11. Ronald L. Rivest and Adi Shamir. 1984. How to Expose an Eavesdropper. *Commun. ACM* 27, 4 (April 1984), 393–394. DOI:http://dx.doi.org/10.1145/358027.358053
12. P.C. Roberts, L.P. Benofsky, W.G. Holt, L.H. Johnson, M.J. Bryant, and N.I. Nussbaum. 2009. Systems and methods for demonstrating authenticity of a virtual machine using a security image. (July 21 2009). https://www.google.com/patents/US7565535 US Patent 7,565,535.
13. P.C. Roberts, L.P. Benofsky, W.G. Holt, L.H. Johnson, B.M. Willman, and M.J. Bryant. 2010. Systems and methods for determining if applications executing on a computer system are trusted. (May 18 2010). https://www.google.com/patents/US7721094 US Patent 7,721,094.
14. Nirupam Roy, Mahanth Gowda, and Romit Roy Choudhury. 2015. Ripple: Communicating through Physical Vibration. In *12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15)*. USENIX Association, Oakland, CA, 265–278. https://www.usenix.org/conference/nsdi15/technical-sessions/presentation/roy
15. Claudio Soriente, Gene Tsudik, and Ersin Uzun. 2008. HAPADEP: Human-Assisted Pure Audio Device Pairing. In *Information Security*, Tzong-Chen Wu, Chin-Laung Lei, Vincent Rijmen, and Der-Tsai Lee (Eds.). Lecture Notes in Computer Science, Vol. 5222. Springer Berlin Heidelberg, 385–400. DOI: http://dx.doi.org/10.1007/978-3-540-85886-7_27
16. Emanuel von Zezschwitz, Paul Dunphy, and Alexander De Luca. 2013. Patterns in the wild: a field study of the usability of pattern and pin-based authentication on mobile devices. In *Proc. of the 15th international conference on Human-computer interaction with mobile devices and services (MobileHCI '13)*. ACM, New York, NY, USA, 261–270. DOI: http://dx.doi.org/10.1145/2493190.2493231
17. T. Yonezawa, J. Nakazawa, and H. Tokuda. 2015. Vinteraction: Vibration-based information transfer for smart devices. In *Mobile Computing and Ubiquitous Networking (ICMU), 2015 Eighth International Conference on*. 155–160. DOI: http://dx.doi.org/10.1109/ICMU.2015.7061059