Investigating Handedness in Air Signatures for Magnetic 3D Gestural User Authentication

Abdallah El Ali
Media Informatics and Multimedia Systems Group
University of Oldenburg
Germany
abdallah.el.ali@uni-oldenburg.de

Hamed Ketabdar
Quality and Usability Lab
Deutsche Telekom Innovation Laboratories
Berlin, Germany
ketabdar.hamed@gmail.com

MobileHCI '15 Adjunct, August 24-27, 2015, Copenhagen, Denmark
ACM 978-1-4503-3653-6/15/08.
http://dx.doi.org/10.1145/2786567.2793691

Abstract
Balancing usability and security in user authentication is essential to the adoption of any authentication method. Magnet-based Around Device Interaction (ADI), in allowing 3D gestural signatures around the device, has been shown to be a secure method for user authentication. In this paper, we further verify the usability and security of this method against video-based shoulder surfing attacks, and test the hypothesis that 2-handed gestural signatures could provide an additional layer of security. Our results showed that while 1-handed signatures are to an extent secure and usable, 2-handed input provides a poor tradeoff between usability and security. We discuss the role of 2-handed gestural input for user authentication on mobile devices.

Author Keywords
Magnet, usability, security, user authentication, 1-handed, 2-handed, 3D gestures

ACM Classification Keywords
D.4.6 [Security and Protection]: Authentication; H.5.2 [User Interfaces]: Input Devices and Strategies

Introduction
Magnet-based Around Device Interaction (ADI) [6] is a novel interaction method which also allows gestural interaction in the whole 3D space around the device by deform-

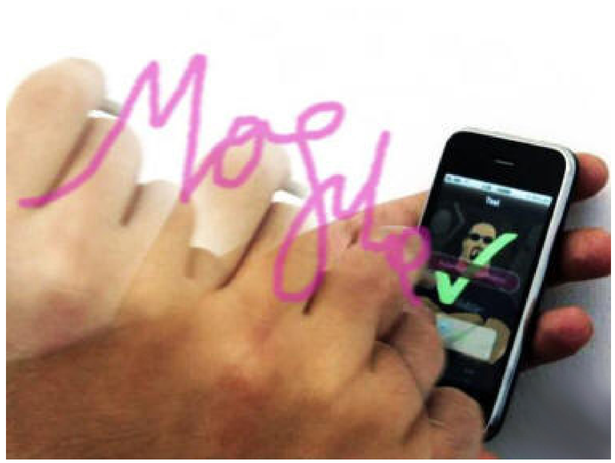

Figure 1: An illustration of a user defining a magnet-based gestural air signature.

ing an embedded compass sensor's magnetic field . Under magnet-based ADI, a regular, properly-shaped magnetic material held in hand (e.g. bar shaped, ring) is used to influence the compass (magnetometer) sensor in (mobile) devices by different 3D gestures. This allows for touchless interaction around the device.

Previous work [8] has shown a need for additional layers of security for different settings, including entering a PIN code to access ATMs or unlock smartphones. Indeed, in a survey with 465 participants asking about security methods on mobile phones, Ben-Asher et al. [2] found that only 26.7% of respondents perceived PIN-based entry methods to be a secure method of user authentication. ADI promises a fast, secure and natural method for user authentication. Despite that 3D gesture authentication is not generally perceived (as assessed by a web survey) as providing a high level of security by users [8], previous work has shown that 3D gestures are in fact quite secure against video-based shoulder surfing attacks [1, 7].

Research Question

As an extension to ongoing work on investigating the usability and security tradeoff of this gestural authentication method, in this paper we present results on the difference between using 1-handed versus 2-handed signatures for user authentication. The motivation behind testing this difference is that we hypothesized that while 2-handed signatures may be initially more difficult to perform, they would provide an additional layer of security that leverages human motor performance given user training. Moreover, we expect users to perceive 2-handed gestural authentication as more secure than performing air signatures with only one hand, given the hypothesized difficulty in an attacker accurately reproducing a 2-handed gesture. In a 2-handed scenario, users do not require any additional setup aside from training to perform a signature using both hands, which is feasible in a mobile setting.

Magnetic Gestural Authentication Framework

Our magnetic gestural authentication framework allows using the embedded magnetic sensor (or magnetometer) of a smartphone as a means of authenticating users. A piece of magnet when moved close enough to a smartphone can influence the compass sensor. Whenever a user performs a new signature around the device (illustrated in Fig. 1 on a iPhone 3GS), the compass sensor registers the temporal patterns of magnetic field along its three axes. A time derivative function is applied to sensor readings in order to highlight changes in the pattern of magnetic field, and remove effects of earth's magnetic field (which is almost constant). The sequence of vectors is divided into overlapping windows for gesture recognition. In order to match templates, we adapt a template matching algorithm called multi-dimensional Dynamic Time Warping (DTW) [9] to analyze different 3D magnetic signatures. DTW is suitable for measuring similarities between two signal sequences that may vary in time or speed, and can operate with a limited number of templates and still achieve very accurate results.

We used DTW to compare the multi-dimensional time series signals with pre-recorded templates of the user's signature for authentication. If the distance of a new input gesture with respect to the prerecorded signature is less than some threshold, the person is considered as a legitimate user and granted access to a smartphone or protected device. In our prototype, in order to define an authentication gesture or magnetic signature, the user arbitrarily moves an appropriate permanent magnet (e.g., a magnetic token/stylus or ring magnet) around the device along 3D trajectories.

Methods

Usability Study

Aside from collecting signal data, to measure the usability and user experience (our dependent variables) of interaction using the magnetic gestural authentication system, we collected: a) System Usability Scale (SUS) [3] responses b) NASA-TLX questionnaire [4] responses c) Likert-scale questions about the perceived usability differences between 1-handed and 2-handed signatures. These different questionnaires were administered to more fully assess the usability of this authentication method under the different handedness conditions. We recruited 20 participants (14 m, 6 f) aged between 20-38 (M_{age}= 29.7; SD_{age}= 5). Most were right-handed (18/20). Study was carried out at the usability lab at Telekom Innovation Laboratories. Each session took between 45-60 min. For recording signatures, we used our magnetic gestural authentication prototype. To obtain precise magnetometer signal information, the SHAKE SK6 [5] sensor was used instead of an iPhone 3GS/4®. The SHAKE sensor is able to sense magnetic fields from proximate magnetic material and transmit the data to a PC over a Bluetooth connection. Each participant was provided with a foam model with an embedded SHAKE sensor and a pole labeled magnet. To record a signature, they had to press and hold the button on the SHAKE sensor (Fig. 2(a)), and then perform their gestural signature using the magnet (Fig. 2(b)). They could take as long as they wanted to practice and define their chosen signature. For 2-handed signatures, participants were asked to make a simultaneous dual-handed gesture for their air signature.

(a) (b)

Figure 2: (a) Foam-based mobile device mockup with embedded SHAKE sensor box (b) pole labelled magnet used for making a gestural signature.

Security Study

We built on previous work [7] and designed a follow-up controlled experiment to assess the vulnerability of our method against video-based shoulder surfing attacks. Under this scenario, we assume the worst case scenario where the

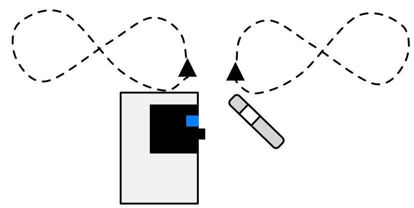

Figure 3: Example of an easy 2-handed air signature.

adversary has full access to HD video evidence of the different angles of performed signatures. Since 2D cameras are widely available to attackers, these were used instead of depth cameras (e.g., Kinect©). To ensure that the adversary has sufficient information on the performed gestural signatures, the video recordings of signatures from the usability study were provided for attackers to target. Signature videos were shown from 4 angles: front, left, right, rear. To make this scenario realistic, we put a restriction on the number of security attacks, where an adversary was allowed only 3 total attacks. To determine which signatures from the usability study were easy or difficult, two independent coders were asked to make a checklist amongst the resulting (post data cleaning) 15 signatures for easy and difficult signatures. Among those, they selected 1-handed and 2-handed signatures for each. The experimenter then selected two easy signatures (1- and 2-handed) and two difficult signatures (1-handed and 2-handed) for attackers to target in this study. The selected difficult 2-handed gestures were not symmetrically mirrored gestures (cf., Fig 3 that shows an easy 2-handed signature), to ensure that some of the tested 2-handed gestures are not easy to forge.

Foregoing design decisions led to a within-subject factorial (2 x 2) design, where all participants had to forge gestural signatures, and signature difficulty (2 levels: easy vs. difficult) and handedness (1-handed vs. 2-handed) were within-subjects factors. Participant assignment was randomized, and order of presented videos was counterbalanced. As in the usability study, participants were given a short training and all relevant hints for forging, such as grasping the foam SHAKE device with the correct position and orientation and how the magnet was held (marker on magnet always up). Four videos of each signature (each with 4 different view angles) were shown to each participant, and then asked to forge the targeted signatures. There was no restriction to the study duration, where participants could watch the videos as many times as desired. They could speed/slow down the videos, as well as step through each frame individually. They were also given a notepad and pen to draw the gestured signature motion if they wanted. 20 participants (11 m, 9 f) aged between 20-34 (M_{age}= 27.1; SD_{age}= 3.6) were recruited. Participants were all were right-handed. Study procedure was the same as the usability study. Aside from collecting magnetometer signal data, we also gave participants a few Likert-scale questions that asked participants about the perceived security differences in forging 1- and 2-handed signatures.

Results

Recall

As in 2D ink-based signatures, a person's signature varies each time it is performed to some degree. To define a 3D magnetic signature and check the repeatability, the user is required to enter a signature template five times. Average DTW distance of all templates is then calculated and used as the main signature. To find the acceptable threshold for authentication, we found the min and max thresholds for accepting a signature as a participant's own, and the acceptance ratio across all participants. The lower the threshold (θ) value, the higher the acceptance rate. For original signatures made, thresholds across 15 participants recall attempts across 1- and 2-handed signatures are shown in Fig. 4. At $\bar{\theta}$ = 2.4, the percentage of successful login attempts for 1-handed signatures is 95.3%, and 80% for 2-handed signatures. This illustrates the difficulty users might have faced in recalling their own signature in the 2-handed condition, and importantly highlight that this authentication method may not be suitable for unlocking smartphones which occurs several times per day [10].

Authentication Speed

Authentication speed was measured from the start of performing a signature to the release of the button on the SHAKE sensor. For this analysis, only the successful recall attempts were considered. The average authentication speed in milliseconds for 1-handed and 2-handed signatures are shown in Table 1. For both types of authentication, the average speed was greater than 3s. In comparison with in-the-wild PIN-based methods [10], this shows that this authentication method may be too slow for frequent, daily use, regardless of the handedness condition.

Signature	Mean	SD
1-handed	3238	1311
2-handed	3234	1336

Table 1: Authentication speed in milliseconds (m/s) for 1-handed and 2-handed signatures.

ROC Analysis of Magnetic Signature Data

To validate the security of the gestural authentication method, we used Equal Error Rate (EER) to measure accuracy. This is the rate at which False Acceptance Rate (FAR) and False Rejection Rate (FRR) are equal. We made use of all recalled signatures defined in the usability study, as well as forgeries made in the security study. We have 15 (signers) x 3 (recall samples) x 2 (handedness) = 90 cases for genuine recall signatures, and 19 (forgers) x 3 (attack samples) x 2 (signatures) x 2 (handedness) = 228 total attacks. To calculate EER, for each forgery threshold value the corresponding FAR and FRR were derived. To plot the Receiver Operating Characteristic (ROC) curve, the True Acceptance Rate (TAR) was calculated (100 - FRR). All (FAR,TAR) pairs were used to plot ROC curves (shown in Fig. 5). For 1-handed signatures, the EER is 8.8% at threshold value of 2.5, which shows that the magnetic gestural authentication system does not provide sufficient security for use in a commercial setting, despite moderate usable access for 1-handed authentication. However, for 2-handed signatures, the EER is 16.7% at threshold value of 2.5. This shows that 2-handed gestural input has a poor security and usability tradeoff.

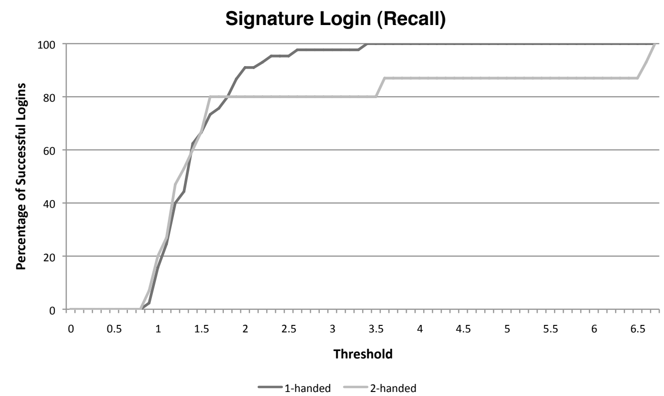

Figure 4: For login attempts across users (N = 15), with θ = 6.5, the percentage of successful logins is 100% successful.

System Usability Scale Responses

Measured SUS responses across participants (N=20) were calculated according to [3], and analyzed in terms of average score frequency distributions. Results are shown in Fig. 6. For 1-handed signatures, only few participants

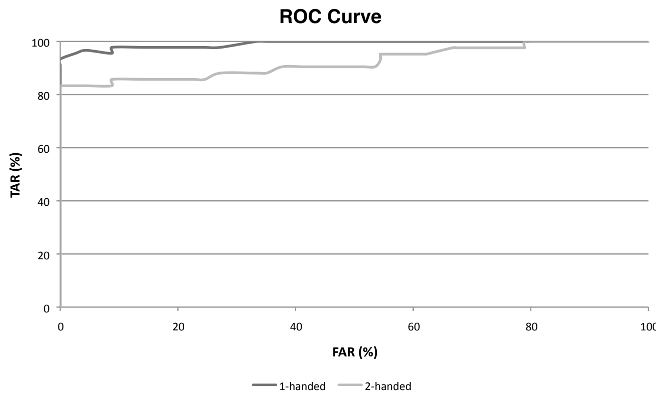

Figure 5: ROC curve for magnetic gestural authentication system for 1-handed and 2-handed signatures.

(6/20) gave a score greater than 70, which indicates that the tested magnetic gestural authentication prototype is not yet ready for use in the consumer market. This is not surprising, given that the system was still a prototype (involving bulky and light foam models with embedded SHAKE sensors and a complicated toolkit interface for recording gestures on a PC). For 2-handed signatures, the usability scores are even lower, with only one participant giving a score greater than 70. While the same argument applies for 2-handed signatures, namely that the tested system is still in a prototype stage, it nevertheless shows that 2-handed signatures for magnetic gestural authentication is not a usable method of authentication.

NASA-TLX Responses
To investigate differences in workload incurred on participants performing 1- and 2-handed signatures, participants filled in the NASA-TLX questionnaire [4] after recording their original signature. Mean responses and confidence

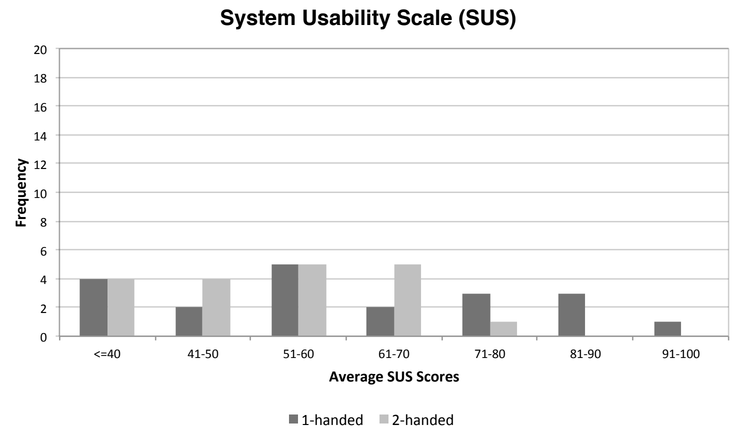

Figure 6: Frequency distribution of mean SUS responses across participants (N= 20) to the magnetic gestural authentication system in 1-handed and 2-handed conditions.

intervals are shown in Fig. 7. We found a significant difference between mean scores ($t(19) = 4.5$, $p < .001$), where mean Subjective Workload for 1-handed signatures is 6.73 compared with 10.2 for 2-handed signatures. These results show that 2-handed signatures incur high workload on participants, which harms the usability of this dual-handed interaction method.

Questionnaire Responses
Perceived Security: To further assess differences between 1- and 2-handed air signatures, users' feedback was gathered via 7-point Likert-scale questionnaire responses for both studies. We report Median (Md) and Interquartile Ranges (IQR). For the usability study (N=20), participants were asked about their perceived difference between 1- and 2-handed air signatures. Participants found both 1-

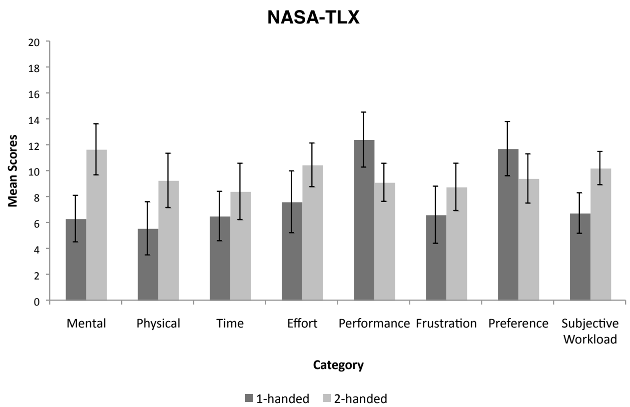

Figure 7: Mean NASA-TLX workload responses [range 0-20] across participants (N= 20) to the magnetic gestural authentication system using 1-handed and 2-handed signatures. Capped error bars represent 95% confidence intervals.

handed (Md=5, IQR=3.75-6) and 2-handed signatures (Md=4, IQR=4-5.25) secure enough for their mobile devices. They also stated that 2-handed signatures cannot be easily forged (Md=4, IQR=2.75-5). In the security study (N=20), after participants attempted to forge signatures, they were also asked about the differences between 1- and 2-handed signatures. Participants found both 1-handed signatures difficult to forge (Md=5, IQR=4.75-6) as well as 2-handed signatures (Md=6, IQR=4-7). When asked about whether these signatures provide sufficient security for their mobile devices, participants found both 1-handed signatures (Md=5, IQR=4-6) as well as 2-handed signatures (Md=5.5, IQR=4-7) secure enough.

Form Factor: To gain insight into how the participants in the usability study found the form factor of the SHAKE foam model, we explicitly asked participants about this in the exit interview. More than half (13/20) of participants found the designed foam model acceptable as a probe into the actual use of this authentication method (*P5: "The foam model was realistic. At one point I imagined it would be a smartphone."*), while the rest (7/20) found it too bulky and/or too light. Nevertheless, even those participants who found it too thick or too light conceded that they served as believable surrogates for smartphones (*P9: "Shape is too big, it is long and thick. But it was fine."*), which meant that we did have external validity, given the prototype stage this authentication system is currently at.

Discussion & Study Limitations

From our recall and ROC results, we had to reject our hypothesis that 2-handed signatures provide an additional layer of actual and perceived security. Despite allowing participants to practice as long as they desired to define their signature, we did expect the usability of 2-handed gestures to be low given the complexity in reproducing simultaneous gestures. This was verified from the SUS and NASA-TLX scores. However, it was surprising that our participants did not perceive 2-handed signatures as more secure against attacks, precisely because they are hard to reproduce.

Concerning study limitations, it is likely that the form factor of the foam device may have negatively influenced participants' perceptions of this authentication method, despite their later responses. We chose to use the SHAKE sensor due to its more precise capture of magnetometer signals, however on current smartphones this difference may be negligible and future work should ensure a more commercial form factor to avoid any user bias. Furthermore, while we assumed a strong attacker in our study, it is not uncommon that attackers would have access to depth cameras. This was however not tested, and is a limitation in our design.

Finally, since our tested authentication method by necessity requires carrying an extra token, this method may not be fitting for a truly mobile setting and for performing a common action (such as unlocking) – it is better suited for infrequent authentication actions that would benefit from an extra token-based layer of authentication in addition to the performed gestural signatures.

Future Work
Key areas for future work are firstly to investigate the repeatability of the signatures over time (e.g., 1 week or 1 month) and test whether different training methods can help users in strengthening their air gestural signatures. Furthermore, we will investigate further the role of the second hand and applied human motor performance in strengthening the security and improving the usability of gestural authentication (e.g., through unique device grasp sensing with the free hand).

Acknowledgements
This work was carried out during an internship (April to June, 2012) at Telekom Innovation Laboratories in Berlin, Germany.

REFERENCES
1. Md Tanvir Islam Aumi and Sven Kratz. 2014. AirAuth: Evaluating In-air Hand Gestures for Authentication. In *Proc. MobileHCI '14*. ACM, New York, NY, USA, 309–318.
2. Noam Ben-Asher, Niklas Kirschnick, Hanul Sieger, Joachim Meyer, Asaf Ben-Oved, and Sebastian Möller. 2011. On the need for different security methods on mobile phones. In *Proc. MobileHCI '11*. ACM, 465–473.
3. J. Brooke. 1996. SUS: A Quick and Dirty Usability Scale. In *Usability Evaluation in Industry*. Taylor & Francis., London.
4. S. G. Hart and L. E. Stavenland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*. Elsevier, Chapter 7, 139–183.
5. S. Hughes. 2006. SHAKE - Sensing Hardware Accessory for Kinaesthetic Expression Model SK6. In *SAMH Engineering Services*. Blackrock, Ireland.
6. Hamed Ketabdar, Mehran Roshandel, and Kamer Ali Yüksel. 2010. Towards using embedded magnetic field sensor for around mobile device 3D interaction. In *Proc. MobileHCI '10*. ACM, NY, USA, 153–156.
7. Alireza Sahami Shirazi, Peyman Moghadam, Hamed Ketabdar, and Albrecht Schmidt. 2012. Assessing the vulnerability of magnetic gestural authentication to video-based shoulder surfing attacks. In *Proc. CHI '12*. ACM, 2045–2048.
8. Hanul Sieger, Niklas Kirschnick, and Sebastian Möller. 2010. User Preferences for Biometric Authentication Methods and Graded Security on Smartphones. In *Proc. SOUPS '10*. ACM.
9. Gineke A. ten Holt, Marcel J.T. Reinders, and Emile A. Hendriks. 2007. Multi-Dimensional Dynamic Time Warping for Gesture Recognition. In *Proc. ASCI '07*.
10. 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 *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '13)*. ACM, New York, NY, USA, 261–270.