A Comparative Study to Evaluate the Usability of Context-based Wi-Fi Access Mechanisms

Matthias Budde, Till Riedel, Marcel Köpke, Matthias Berning and Michael Beigl

TECO, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany,
email: budde@teco.edu, www: http://www.teco.edu/people/budde

This paper presents a comparative study of six different tag and context based authentication schemes for open Wi-Fi access. All of the implemented methods require only a smartphone and an *HTML5* capable webbrowser, making them interchangeable and easy to incorporate into existing infrastructure. We recruited 22 participants for the study and used two standardized questionnaires as well as additional metrics to assess whether further investment in a systematic usability analysis seems prudent. The evaluation shows that suitable alternatives for Wi-Fi authentication exist and points out their limitations and opportunities.

Keywords: Universal Access, Practical Security, Usability, User Experience, User Study, Interfaces, Device Association, Authentication, Wi-Fi, Smart Environments, Context

1 Introduction

Future computing environments – as driven by the notions of ubiquitous computing and ambient intelligence – are expected to give rise to information technology that is embedded in everyday life and is spontaneously formed from ubiquitous devices, objects, and services that we can easily access. While humans understand how to access physical resources at their disposal, it is often harder in a digital world. Accessible digital resources are a key factor to assistance and inclusion, especially in public spaces. Open Wi-Fi access, e.g. through hotspots, creates many issues (access control, liability and legal issues) on the user and institutional side, as malicious parties cannot be kept out of the network. This is why today, usually username and password, entered into a web interface (a.k.a. *captive portal*), are required to access wireless infrastructure.

 Definitive version published and presented in *Proceedings of the 16th International Conference on Human-Computer Interaction (HCII 2014)*, Creta Maris, Heraklion, Crete, Greece, 22-27 June 2014. The final publication is available at link.springer.com.
DOI: 10.1007/978-3-319-07446-7_44 10 http://dx.doi.org/10.1007/978-3-319-07446-7_44

Providing seamless wireless network service is not only about network quality but also about user experience and ease of access has become an important factor for quality of life. Adding an extra burden on users – particularly technically non-literate users or ones with special needs – actually excludes many people from access, especially when using complicated *username:password* schemes with media breaks. As a step towards the proliferation of accessible networks, this work evaluates the usability of ways for associating handheld devices to Wi-Fi networks, while also regarding implementation and practical feasibility.

2 Related Work

A great variety of schemes have been proposed in the past to pair mobile devices for spontaneous interaction, many of which could also be applied for the use case of authenticating to a Wi-Fi hotspot. One solution proposed is the use of Out-of-Band (OOB) information to establish a shared secret. Holmquist et al. [6] as well as Mayrhofer et al. [12] proposed to couple devices using their accelerometers by shaking them simultaneously. This, however, is not easily applicable to systems involving static infrastructure. A method to generate a shared secret from ambient audio was investigated by Sigg et al. [16]. These technical publications do not take human factors into account, which we targeted in this work. Several other authors conducted comparative user studies on the usability of device pairing or Wi-Fi setup. Kostiainen et. al. [9] used formative interviews to assess user needs in home network access control and proposed a conceptual system for setup and management. Both Uzun et. al. [18] and Kainda et al. [8] analyzed device pairing by different textual interfaces. Kumar et. al. [10] included variants of eight device pairing methods in a large study implemented on a common platform. The usability assessment is mostly based on automatically logged user actions with no qualitative feedback. Ion et al. [7] used mock-ups to investigate the usability of device pairing methods with regard to perceived security needs in different real-life situation. Their findings indicate that the usability as well as user preference depend on the context of the pairing situation.

3 Methodology

In this work, we conducted a user study to explore six different methods suitable to replace classic *username:password* authentication for public Wi-Fi. In addition to token based methods we adapted we adapted two relevant OOB techniques and included a recent approach using surface-confined Wi-Fi [3] as context based methods. The methods were selected to fulfill the following constraints:

- *Non-mediated*: The user can perform the login at any time by himself.
- *Intuitive*: Wi-Fi hotspots are broadly used by ordinary non-expert users.
- *Platform-independent*: Mobile devices and OSs are diverse and volatile.

- *Explicit*: The system can record consent of the user when connecting.

The last constraint distinguishes the evaluated methods from access control based on *Geofencing* [14] or overprovisioning [4], which both allow confining Wi-Fi networks to certain physical boundaries, enabling location based access. Although it can be argued that *implicit* access is generally preferable, additional to fulfilling legal contstraints, the proposed explicit methods do not require overprovisioning in the infrastructure and are therefore also deployable in an light-weight and interchangable fashion.

For platform-independence, we selected methods that can solely be implemented using web technology, thus easily be adapted to user needs and incorporated into *captive portals*. For this paper we implemented all on top of *HTML5* features like `GetUserMedia()` and `GetDeviceMotion()`. This has the aditional benefit that the schemes are interchangeable and a set of different methods can be offered to respect user preferences or hardware constraints. The source code of the implementations will be released as part of the *Global Public Inclusive Infrastructure* (GPII) component repository and are freely available[1].

Username:password This authentication scheme is well-known and standard today. A user enters his credentials into a form and presses a login-button to gain access. We included it as a base line and to back the obvious hypothesis that this scheme is not well suited for handheld devices.

QR Codes By encoding a URL containing the login credentials into a QR code it can be scanned with any camera phone. We accessed the camera image through *HTML5* and decoded it using a *JavaScript* library (*jsqrcode*).

NFC Login information can also be stored as URL in a *Near Field Communication* (NFC) tag and accessed remotely with many modern phones. An alternative would have been NFC-based *Wi-Fi Protected Setup*, which we refrained from using due to known vulnerabilities [19] and lacking personalizability.

2DST Sheet A detailed description of the *Two-Dimensional Signal Transmission* (2DST) waveguide sheet was published in previous work [3]. To log on, users connect to the open Wi-Fi and the *captive portal* page prompts them to place the device on the sheet (see Figure 1) and acknowledge the coupling. The device can subsequently be removed from the sheet and access is granted.

Figure 1: Phone, 2DST sheet.

Kinect We adapted the *Point&Control* system [2] that uses the *Microsoft Kinect* (first generation) for user-device association based on the user context. The *captive portal* page prompts the user to press a button and raise an arm to connect. If the gesture is matched by the Kinect, access is granted. For this usability test we used this very simple

[1] `https://github.com/teco-kit`

scheme, that does not match the accelerometer pattern of the phone with the *Kinect* model for added contextual prove, which is possible on most modern devices also using *HTML5* [5].

Audio Context The last context-based method employs ambient audio as OOB channel, as used in *Pintext* paring [16]. The entropy of generated fingerprints generally makes them suitable to be used as a shared secret [15]. To log in, a user presses a button on the *captive portal* and both phone and server record for eight seconds. A server synchronizes[2] both recordings, calculates fingerprints and compares them. If they are sufficiently similar, access is granted.

3.1 Task and Session Structure

Participants were asked to connect to an open Wi-Fi, authenticate their device using the given mechanism and open a browser to access a web page. For this, they were given a *Samsumg Galaxy SIII* smartphone, which supported all of the six methods. The test sessions were conducted in an office at our lab – in German or English, depending on the subject's preference. Participants were welcomed and guided to the test room one at a time (1-on-1 moderated sessions, see Figure 2). The test's setup and intention were introduced and subjects read and signed a privacy statement. Subsequently, the moderator collected demographic data (age, gender, etc.) and some information on the subject's habits regarding technology use (frequency of handheld Wi-Fi access, usage of public Wi-Fi hotspots, etc.) by means of a pre-test questionnaire. After that, the main phase of the test commenced: Subjects, in turn, completed the task – i.e. log on to the Wi-Fi and open a website – for each method and fill in a questionnaire. The order of the six methods was shuffled to avoid biases from practice or fatigue. Each method was explained beforehand and written descriptions of how to proceed were available throughout the test. After repeating the three steps for each method, the participant was asked to fill in the post-test questionnaire. Sessions took between 52 and 100 minutes, with an average of 69 minutes.

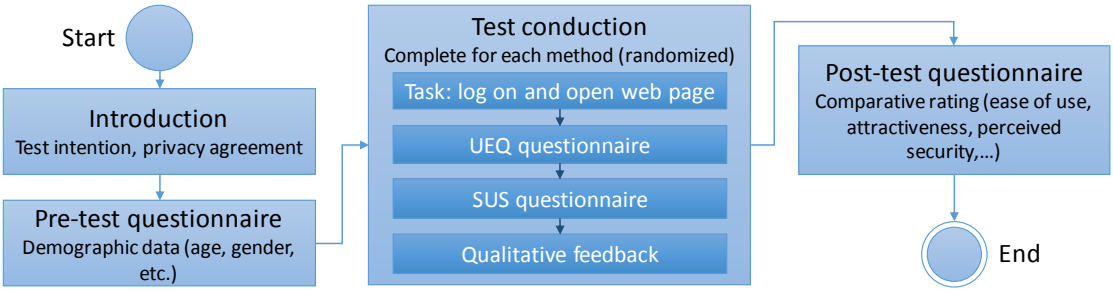

Figure 2: Session structure each participant ran through during the user study.

[2]In our tests synchronization significantly slowed down this scheme, as acurate time sync could not be realized in *HTML*.

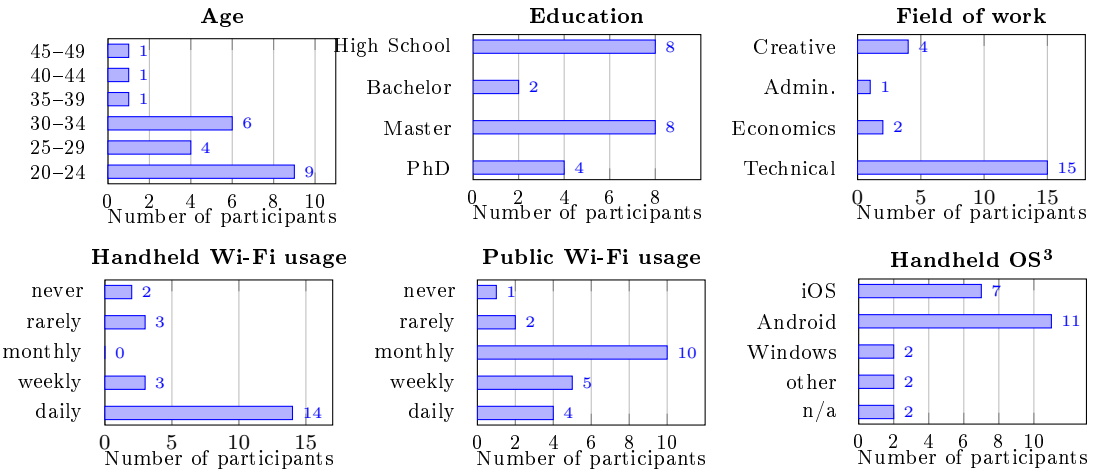

Figure 3: (Top) Participants by age, completed level of education and field of work or studies. (Bottom) Frequency of Wi-Fi network access with a handheld device respectively using open Wi-Fi hotspots, as well OS's installed on subjects' personal devices[2].

3.2 Participants

We recruited 22 participants aged between 20 and 48, eight of them female. All of them attended voluntarily without being offered a reward. Figure 3 shows the data on demographics and participants' habits collected using the pre-test questionnaire. The subjects composed a well-educated group, accustomed to the use of mobile devices and working in different fields, the majority pursuing technical professions. None were security experts. All participants reported accessing the Internet once or more per day. Most (14) also daily used handheld mobile devices to connect to Wi-Fi networks, some weekly (3). Three participants said that they rarely accessed Wi-Fi with a handheld device and two never at all. Most subjects often used public Wi-Fi hotspots, only one never did, and two rarely. All others accessed public Wi-Fis at least monthly, four even daily. Overall, the subject group is suitable for an initial assessment of the selected methods, as they are digitally literate and familiar with the presented task.

3.3 Questionnaire Design

The study we conducted is mostly summative, with additional formative aspects. We were looking to find out whether the proposed solutions could in principle satisfy user needs. To quickly assess both usability and user experience of the tested schemes, we considered different standardized questionnaires. The *System Usability Scale (SUS)* [1] has been applied to a wide range of systems in the past 20 years, from printers over phones to desktop and web applications. Subjects express their level of agreement to ten simple statements using a five-point Likert. It is slim, short term viable and yields

[2]Three participants chose multiple options.

a single score as result, which is already generalizable at relatively small sample sizes [17]. As alternative, we looked at the *User Experience Questionnaire (UEQ)* [11]. It consists of 26 pairs of opposing attributes (e.g. *annoying* and *enjoyable*). Users express their agreement with them on a seven-point Likert. The UEQ yields six different scores for the categories *attractiveness*, *perspicuity*, *dependability*, *efficiency*, *stimulation*, and *novelty*. We decided to use both UEQ and SUS, as they can be filled in quickly and we were interested to see if they yielded consistent results, since the SUS focuses on usability while the UEQ aims at assessing the whole user experience.

Aside from summative data from the two questionnaires, we also collected three to five qualitative statements about what the subjects liked and disliked about each system. All of this was done separately for each of the methods. In addition, we constructed a short post-test questionnaire which prompted the subjects to directly compare the systems with each other, by ranking them regarding their ease of use, perceived security and attractiveness. Participants were also asked which of the systems (if any) they would recommend to friends or acquaintances. Finally, they were given the possibility to specify additional free text comments. The moderator also recorded any unprompted statements made throughout the test. Aside from the questionnaires, we recorded the number of attempts needed to complete the task and the time to do so.

3.4 Data Cleansing

Overall, our implementations proved to run stably and the conduction of the study went smoothly. In two cases however, we experienced software problems that led to difficulties in completing the task: A software crash interrupted the task completion in six cases of the *Audio Context* login and an error dialog was shown. In these cases, the task was repeated and participants were instructed to disregard the first failed attempt. Failed tries were not included in the results for task times or number of attempts. In two instances involving the 2DST sheet, a software bug prevented the correct recognition of the device by the sheet, leading to an unusually high number of tries (12 respectively 8 attempts). To avoid skewed results, the data from these two runs was removed from the set.

Regarding the final three ranking questions, subjects explicitly were given the option to rank two systems equally by assigning the same ordinal number. However this lead to some participants using *competition ranking* (i.e. leaving a gap in the ranking when several systems tied) and others using *dense ranking* (no gaps). To reach a realistic ranking when averaging over all participants, we transformed the data to *fractional ranking* scores, as those have the property that the ranking numbers' sum is the same as under strict ordinal ranking.

4 Results

This section presents the results of our analyses. First, we show the quantitative metrics (attempts, task time), followed by the SUS and UEQ scores. Finally, comparative statements and qualitative feedback ratings are presented.

Table 1: Automatically collected metrics (sorted by median time per attempt).

	# Attempts					Task time (overall)					Task time (per attempt)				
Meth.	min	max	med.	mean	conf.[3]	min	max	med.	mean	conf.	min	max	med.	mean	conf.
NFC	1	1	1	1.00	*n/a*	2*s*	15*s*	6*s*	6.5*s*	1.36	2*s*	15*s*	6*s*	6.5*s*	1.36
Kinect	1	4	1	1.64	0.38	7*s*	57*s*	11*s*	19.3*s*	5.66	7*s*	17*s*	10*s*	10.5*s*	1.16
2DST	1	3	1	1.45	0.33	7*s*	56*s*	14*s*	19.3*s*	5.81	7*s*	34*s*	11*s*	13.2*s*	2.65
QR	1	3	1	1.55	0.33	10*s*	91*s*	33*s*	44.9*s*	11.90	6*s*	81*s*	26*s*	30.4*s*	7.51
Pwd	1	2	1	1.09	0.42	34*s*	153*s*	51*s*	60.3*s*	1.67	34*s*	111*s*	51*s*	54.4*s*	1.25
Audio	1	2	1	1.18	0.17	77*s*	381*s*	126*s*	157.3*s*	33.43	77*s*	215*s*	119*s*	130.3*s*	16.25

4.1 Quantitative Performance Metrics

Table 1 shows the number of attempts and the time needed to perform the login task (both overall and averaged per attempt). NFC was the only scheme which took all subjects only one attempt, all others had to be repeated at least once by at least one participant. While this was seldomly necessary for *username:password* (2, mistyping) and *Audio Context* (4, fingerprints too different), it happened more often with the 2DST sheet (6, device removed from sheet to early). More than a third of the subjects had to repeat their attempt using QR codes (8, recognition failed) or the *Kinect* (9, tracking failed). Regarding task times, all methods except *Audio Context* performed significantly faster than entering passwords. Again, NFC stood out, closely followed by *Kinect* and 2DST sheet. The use of QR codes still took only half the time of entering text credentials, while the audio login took much longer due to the long processing times.

4.2 SUS and UEQ scores

Figure 4 and Table 2 show the UEQ category results, the overall UEQ score as well as the SUS score over all participants. When looking at SUS scores, a percentile rank of below 60.0 is considered poor and an indicator for severe usability problems, while values over 80.0 are generally good [17], 100.0 being the maximum possible. Regarding the UEQ score, values below -0.8 indicate a negative rating, over 0.8 a positive one, and in between neutral. However, Schrepp et al. [13] point out that the actual interpretation of the ratings depends on the weight of the categories in the concrete application and intended user group. For normal end users, they regard *attractiveness* as most important, followed by *perspicuity* and *dependability*, and thirdly *efficiency*. The authors also provide a benchmark which is based on 163 studies with a total of 4818 participants and sets the score boundary between above-average and below-average different for the individual categories (*attractiveness*: 1.09, *perspicuity*: 0.9, *dependability*: 1.06, *efficiency*: 0.84 , *stimulation*: 1.0, and *novelty*: 0.63).

Basing our interpretations on these preliminary considerations, the ratings show a slightly different picture than the performance metrics before: Judging from the SUS scores, for *Kinect* and QR, no compelling conclusion can be drawn from the SUS score. Both NFC and the 2DST sheet login can be considered good systems with no apparent usability problems, while *Audio Context* seems to have severe issues. *Username:password*

[3] All confidence intervals ($\pm$) in this work are constructed at a 95% confidence level.

Table 2: Mean UEQ category and SUS scores, including confidence[4], and median SUS.

Meth.	UEQ category scores												SUS score		
	attrac.		perspic.		depend.		effic.		stimul.		novel.				
	mean	conf.	mean	conf.	mean	conf.	mean	conf.	mean	conf.	mean	conf.	median	mean	conf.
Pwd	**-0.85**	0.45	1.56	0.50	1.22	0.31	0.00	0.51	-1.30	0.41	-2.50	0.32	63.8	66.3	6.02
QR	0.94	0.48	1.76	0.33	1.11	0.41	0.88	0.46	0.59	0.41	0.76	0.44	73.8	77.7	5.56
NFC	**2.11**	0.36	**2.72**	0.21	**2.02**	0.40	2.24	0.39	1.60	0.37	1.81	0.40	**95.0**	**91.6**	4.45
2DST	**1.67**	0.38	**2.31**	0.34	1.30	0.45	1.78	0.35	1.39	0.45	1.90	0.45	**86.3**	**83.5**	4.72
Kinect	0.75	0.60	1.83	0.43	**0.22**	0.50	1.13	0.38	1.48	0.40	2.16	0.29	72.5	69.9	7.67
Audio	-0.12	0.60	**1.27**	0.48	-0.15	0.32	-0.42	0.39	0.24	0.42	2.10	0.41	55.0	58.5	8.70

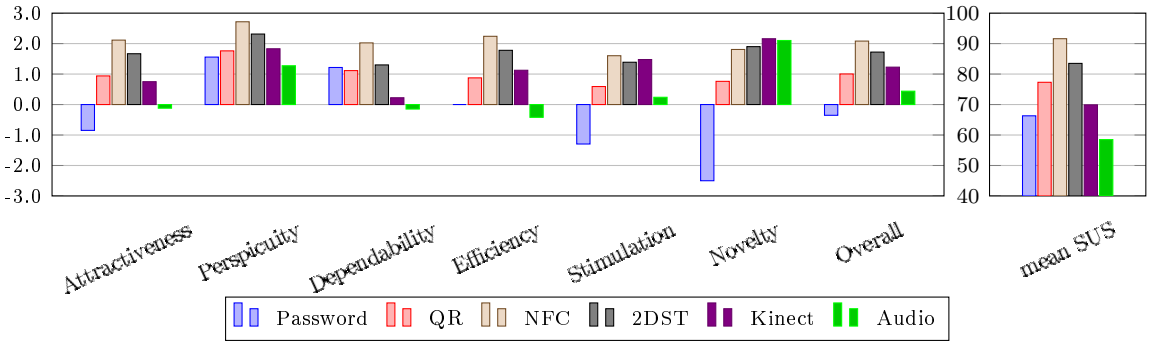

Figure 4: Mean scores for each access method: UEQ categories (left) and SUS (right).

barely scores better than ambient audio, only just exceeding 60.0 points.

When looking at the UEQ, we see a difference between the overall UEQ and the SUS score: The overall UEQ rating for the *Kinect* based system is better than that of QR codes, and *Audio Context* scores higher than *username:password*.

Regarding the most relevant UEQ categories, the use of passwords scores high in the categories *perspicuity* and *dependability*, while receiving low ratings for *attractiveness* and *efficiency*. This illustrates the additional information that can be drawn from the UEQ scores. Similar observations can be made for the other systems: While the QR method scores more or less average in all remaining categories, it performs well regarding *perspicuity*. Both NFC and the 2DST sheet login clearly score high in all categories. According to the UEQ scores, the main shortcoming of the *Kinect* scheme is *dependability*, which is in line with the observations made from the performance metric *number of attempts* while the *efficiency* score adequately reflects the short completion times of the scheme. *Audio context* scores badly across the board, with the exception of the categories *perspicuity* and *novelty*. An aspect clearly differing from the performance metrics are the non-task related hedonic quality aspects, that express how *novel* and *stimulating* users percieve a system to be [11]. Not surprising, the password scheme has a very low rating here, while the context-based methods score high.

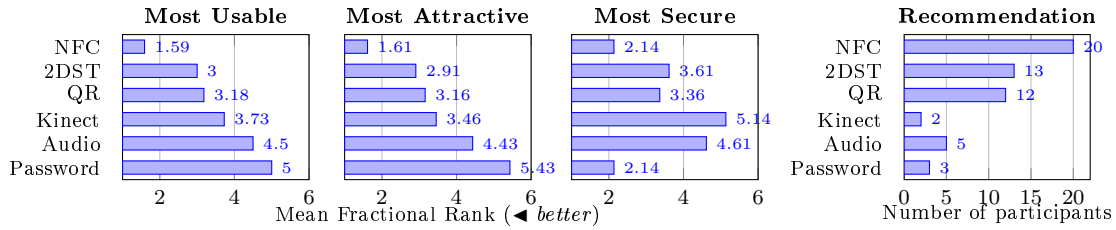

Figure 5: Subjective ranking of the methods regarding usability, perceived security and attractiveness (left) and which schemes participants would recommend (right).

4.3 Comparison & Preferences

The last section of the post-test questionnaire specifically prompted the users to rank the six methods compared to each other. The resulting ranking (see Figure 5) is consistent for the two aspects *usability* and *attractiveness*: NFC clearly spearheads, followed by the 2DST sheet, QR codes, *Kinect* and then *Audio Context*. The classic *username:password* scheme came out last. The rankings so far are in agreement with the UEQ *attractiveness* rating.

Interestingly, the *perceived security* ranking paints a different picture: Generally, participants felt that the token based methods were more secure than the context-based schemes (see Figure 5).

Altogether, almost all (20) participants stated that they would recommend NFC as Wi-Fi access method to friends or acquaintances. More than half would recommend using QR codes (12) or the 2DST sheet for contextual access (13). Notably, despite the otherwise rather poor ratings of the audio based scheme, still five participants would recommend the system, more than the classic password (3) or the *Kinect* based scheme (2).

4.4 Qualitative Statements & User Feedback

As expected, the study confirmed that *username:password* credentials are not suitable for Wi-Fi authentication on mobile devices. While users rate the method as *intuitive* (8) and secure (10), almost all users said that *smartphones are inadequate for entering random strings* (18). As unique property, subjects saw that *text is memorizable* ($P06$), providing an abstract token. Regarding QR codes, we observed that opinions were divided: Participants both describe the method to be *intuitive* (15) and *complicated* (9) as well as *fast* and *slow* (9 each). Some (6) also stated that they either experienced or anticipated problems in bad lighting. ($P07$: *the shadows of hand and phone interfered*). Regarding the necessary permission for the browser to access the camera, $P22$ reported: *giving deep system access without exactly knowing what is going on makes me uneasy.*

Concerning NFC, participants predominantly gave positive comments (see Table 3): Most subjects characterized it as *intuitve* (18) and *fast* (17). However, a few users (5) also voiced technical concerns, ranging from fear of high energy consumption ($P14$: *NFC always on?*) or losing the tag to security concerns, as the phone did not ask permission

before opening the webpage. Participant $P01$ also specifically disapproved that *one has to use both hands*. Some users (4) also pointed out that not every device features NFC (e.g. $P15$: *my iPhone doesn't have an RFID reader*).
Regarding the 2DST sheet, users reported that it was *innovative/cool* (6) as well as *fast* (13) and *easy* (21) – emphasizing that they needed to do very little ($P06$: *that's it?*). Some users perceived the system as limited in terms of being fixed to a location (3) and one subject did not grasp the concept of context-based access ($P01$: *big and bulky and would not fit in my bag*). Another user was concerned regarding possible radiation from the sheet.

Table 3: Amount of positive and negative comments (both prompted and unprompted).

Method	Pos.	Neg.	Ratio
NFC	82	25	3.28
2DST	73	37	1.97
QR	57	48	1.19
Kinect	70	62	1.13
Passwd	46	47	0.98
Audio	48	78	0.62

As for the *Kinect*, positive and negative comments nearly balanced each other: On one side, users saw the system as *fast* (13), *intuitive* (14) and *cool* (13). On the other, some people were embarrassed (8, e.g. $P08$: *don't want to jump about and attract attention in public*) and some voiced their concerns on being recorded and possible privacy implications (2, e.g. $P05$: *sense of being under surveillance*). Others again liked the aspect of performing an activity in order to log on (8). Regarding the *Ambient Audio* login, the users' main issue was the *long wait* (21), followed by technical concerns, such as interference from *handling noises* or unreliability in *silent ambiences* (2, e.g. $P11$: *especially problematic for mutes*). Four users expressed disbelief that the method would work at all ($P13$: *I have the feeling that this will often fail*). Regarding the security, the method felt both *insecure* (5, e.g. $P18$: *I guess that many false positives occur*) and *very secure* (4) to the users. As with the *Kinect*, privacy concerns were also expressed (2). On the other hand, participants characterized the system as *innovative* (8) and *intuitive/magic* (12, e.g. $P07$: *great, I don't have to do anything*).

General comments mostly concerned lack of understanding regarding context-based access applicability (e.g. $P18$: *I don't see use cases, except maybe in trains*).

5 Discussion

We implemented six *HTML5*-based techniques for associating handheld devices to Wi-Fi networks and evaluated them regarding their usability. As expected, it backed that *username:password* credentials are unsuitable for handheld Wi-Fi access, and that the other five schemes may – to a varying degree – present viable alternatives.

Our study yielded some interesting results: A general observation is that purely summative studies may not reflect the full range of relevant aspects. Although both SUS and UEQ seem suitable to determine if severe issues exist, caution should be exercised when using them to rank systems. While we can see that they generally show similar tendencies, the UEQ addresses the whole user experience and its categories can provide helpful additional insights regarding the area possible problems may reside in. This is especially true for systems whose SUS score lies in the "gray area" between 60.0 and 80.0 percent. Placing too much emphasis on mere speed or completion rates as a factor

may be misleading as well, especially in the context of usable security. In this area, it is important to augment standard usability testing with some metric that specifically addresses aspects like perceived security, trust, etc.

The multitude of different user statements revealing interesting issues – both actual and perceived – underlines the importance of also collecting qualitative feedback. While for the two best performing schemes (NFC and 2DST) most metrics are in agreement with each other, singular issues are revealed by other metrics, such as qualitative feedback for NFC or the number of attempts with the sheet. For some methods, only the full range of metrics paints an adequate picture for the assessment of the system, especially those on which opinions are divided: The *Kinect* achieved average to high summative ratings and the second to best completion times, but subjects perceived it as insecure and least recommended it. As for the *Audio Context* scheme, despite mostly bad ratings and performance, more than half of the subjects regarded it as intuitive and almost a quarter would recommend it. This illustrates that multiple metrics also allow discerning between fundamental issues and specific problems that can e.g. be attributed to the implementation and may be remedied in the future.

An important realization regarding methods involving cameras or microphones was that many participants voiced their concerns on being recorded and possible privacy implications, as well as a sense of being under surveillance. We conclude that systems involving video or audio recording should probably be avoided, and if such methods are considered, it is important to convey to the users that their privacy is protected. Furthermore, deep system access by web apps (e.g. camera or sensor access or automatically opening scanned URLs) should be transparent and only occur with user consent. When considering methods involving visible activity, embarrassment is an important factor, even if the *Kinect* based scheme caused both positive and negative feedback regarding the activity. We conclude that the applicability of such a system depends strongly on the situation, user group and maybe also cultural aspects. As a design guideline for systems that involve little interaction (*Ambient Audio*) and/or losing focus of the screen (as sometimes seen with *Kinect* or *2DST*), non-visual feedback such as an auditory signal is advisable to indicate success of the association process.

6 Conclusion & Future Work

We believe that good integration of classic usability studies and metric based analyses, as well as an analysis of other requirements on the user side (hardware features, OS, etc.) and the operators side (infrastructure, maintenance cost, etc.) needs to be conducted in order to come to a meaningful assessment of the viability of a method. As we collected very different statements regarding concerns and the acceptance of systems from technically versed and unseasoned users, we conclude that usable security systems should ideally be evaluated in a real-world context to assess which methods users would actually choose and keep using.

In future work we plan to study the most suitable methods addressing especially technically non-literate users and users with special needs. We also plan to further evaluate the

barriers that hinder including novel, multimodal *HTML*-based context and user sensing methods – such as the alternative login methods presented in this work – in real applications. This includes further assessment of the potential of basing such systems on web technology, as this makes systems easy to interchange or include into existing applications in a modular fashion.

Acknowledgements

This work was partially funded by the European Union under project *Prosperity4All*, grant *610510*. We thank Daniel Karl for his implementation support as well as all study participants, especially Klaus Rümmele and his staff.

References

[1] Brooke, J.: SUS - A quick and dirty usability scale. Usability evaluation in industry 189 (1996)

[2] Budde, M., Berning, M., Baumgärtner, C., Kinn, F., Kopf, T., Ochs, S., Reiche, F., Riedel, T., Beigl, M.: Point & Control – Interaction in Smart Environments: You Only Click Twice. In: UbiComp '13 Adjunct. pp. 303–306. ACM (2013)

[3] Budde, M., Köpke, M., Berning, M., Riedel, T., Beigl, M.: Using a 2DST waveguide for usable, physically constrained out-of-band Wi-Fi authentication. In: 2013 ACM Conference on Pervasive and Ubiquitous Computing. pp. 221–224. ACM (2013)

[4] Faria, D.B., Cheriton, D.R.: No long-term secrets: Location-based security in overprovisioned wireless lans. In: Hot Topics in Networks (HotNets-III) (2004)

[5] Hauber, M., Bachmann, A., Budde, M., Beigl, M.: jActivity: Supporting Mobile Web Developers with HTML5/JavaScript Based Human Activity Recognition. In: 12th International Conference on Mobile and Ubiquitous Multimedia. ACM (2013)

[6] Holmquist, L., Mattern, F., Schiele, B., Alahuhta, P., Beigl, M., Gellersen, H.: Smart-Its Friends: A Technique for Users to Easily Establish Connection Between Smart Artefacts. In: Ubicomp'01 (2001)

[7] Ion, I., Langheinrich, M., Kumaraguru, P., Čapkun, S.: Influence of User Perception, Security Needs, and Social Factors on Device Pairing Method Choices. In: SOUPS'10. ACM (2010)

[8] Kainda, R., Flechais, I., Roscoe, A.W.: Usability and security of out-of-band channels in secure device pairing protocols. In: SOUPS'09 (2009)

[9] Kostiainen, K., Rantapuska, O., Moloney, S., Roto, V., Holmstrom, U., Karvonen, K.: Usable access control inside home networks. In: WOWMOM. pp. 1–6 (2007)

[10] Kumar, A., Saxena, N., Tsudik, G., Uzun, E.: A comparative study of secure device pairing methods. Pervasive and Mobile Computing 5(6), 734–749 (2009)

[11] Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. HCI and Usability for Education and Work (2008)

[12] Mayrhofer, R., Gellersen, H.: Shake well before use: Authentication based on accelerometer data. In: Pervasive computing (2007)

[13] Schrepp, M., Olschner, S., Schubert, U.: User Experience Questionnaire Benchmark Praxiserfahrungen zum Einsatz im Business-Umfeld. In: Usability Professionals '13

[14] Sheth, A., Seshan, S., Wetherall, D.: Geo-fencing: Confining Wi-Fi coverage to physical boundaries. In: Pervasive'09 (2009)

[15] Sigg, S., Budde, M., Ji, Y., Beigl, M.: Entropy of Audio Fingerprints for Unobtrusive Device Authentication. In: Context 2011. LNCS (2011)

[16] Sigg, S., Schuermann, D., Ji, Y.: PINtext: A Framework for Secure Communication Based on Context. In: MobiQuitous 2011 (2011)

[17] Tullis, T., Albert, W.: Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics. Elsevier Science (2010)

[18] Uzun, E., Karvonen, K., Asokan, N.: Usability analysis of secure pairing methods. In: Financial Cryptography and Data Security, pp. 307–324. Springer (2007)

[19] Viehböck, S.: Brute forcing Wi-Fi Protected Setup. Wi-Fi Protected Setup. (2011)