Towards Understanding the Link Between Age and Smartphone Authentication

Lina Qiu
University of British Columbia
Vancouver, Canada
lqiu@ece.ubc.ca

Alexander De Luca
Google
Zürich, Switzerland
adeluca@google.com

Ildar Muslukhov
University of British Columbia
Vancouver, Canada
ildarm@ece.ubc.ca

Konstantin Beznosov
University of British Columbia
Vancouver, Canada
beznosov@ece.ubc.ca

ABSTRACT

While previous work on smartphone (un)locking has revealed real world usage patterns, several aspects still need to be explored. In this paper, we fill one of these knowledge gaps: the interplay between age and smartphone authentication behavior. To do this, we performed a two-month long field study (N = 134). Our results indicate that there are indeed significant differences across age. For instance, younger participants were more likely to use biometric unlocking mechanisms and older participants relied more on auto locks.

ACM Reference Format:
Lina Qiu, Alexander De Luca, Ildar Muslukhov, and Konstantin Beznosov. 2019. Towards Understanding the Link Between Age and Smartphone Authentication. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3290605.3300393

1 INTRODUCTION

Advances in technical capabilities of modern smartphones enable storing (and access to) large amounts of data, some of which might be sensitive or private in nature [14]. Due to these devices' small sizes and high mobility, unauthorized access to sensitive data has become a realistic threat. For instance, it has been shown that 1 out of 5 users in the US has accessed someone else's mobile phone without permission [13].

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05.
https://doi.org/10.1145/3290605.3300393

To protect smartphones from unauthorized access, mobile operating systems provide secure device locking, however, most smartphone users tend to choose no or easy-to-guess unlocking secrets [3]. Inconvenience of the currently deployed unlocking methods, lack of motivation, and lack of awareness about the sensitivity of the data stored on their devices are often used to justify why a secure lock is not used [4, 7, 14].

Recent studies [5, 6, 10, 12] have started to shed light on how and when smartphone users employ unlocking mechanisms, some of which were conducted in the wild. For instance, Mahfouz et al. [12] investigated authentication process parameters, such as time to unlock, authentication error rates, and types of apps used within each session. Harbach et al. [5] focused on exact authentication speed, error counts, and types of errors among different unlocking mechanisms, with a study sample containing 134 participants. While these efforts have provided first valuable insights into studying performance of smartphone unlocking mechanisms in the wild, there are still plenty of unknowns.

One of these unknowns is the relation between age and authentication behavior. A recent online survey on secure smartphone locking across eight countries suggested that age might be linked to differences in locking behavior of users [6] (e.g., whether or not they use secure locks). It is, however, still unclear how age affects smartphone unlocking, because real world data is missing. In order to fill this knowledge gap and to provide initial insights into how age correlates with smartphone authentication, we conducted a longitudinal field study with 134 participants of diverse backgrounds. The main goal was to investigate what behavioral patterns are linked to age and which are not.

The results of our study reveal that age indeed significantly correlates with unlocking behavior. For example, older participants were less likely to use "Fingerprint" authentication. They also interacted with their devices less frequently,

which means that they are less frequently exposed to their unlocking mechanism.

2 RELATED WORK

Despite the existence of a variety of smartphone unlocking methods (e.g., PIN, alphanumeric passwords, Android unlock patterns, biometrics), a recent study [4] showed that most smartphone users likely underestimate the sensitivity of their data and how they access it with their devices. Thus, many users do not protect their devices, and with it their data, properly [6]. One potential reason for this is the inconvenience caused by frequent unlocking and the respective time this takes [7]. As a consequence, many alternative systems were proposed to make the authentication process easier [1, 8, 9, 15]. SnapApp [1], which provides a time-constrained quick-access option, is an example that reduces the authentication workload by keeping users logged-in in a more secure way than having their device unlocked all the time.

Researchers and designers need data on how people use their devices in the real world, in order to design new mechanisms that are in line with users' real needs. A few studies have started to provide insights into this [5, 6, 10, 12].

Mahfouz et al. [12] studied how different smartphone use patterns correlate with the time it takes users to unlock their device, how often users make a mistake during authentication, and which authentication methods users choose for device unlocking. The findings suggest that users who lock their devices interact with them more frequently and for longer sessions than those who do not. In addition, the cost of unlocking is low when compared to overall smartphone usage and users do not mind adopting unlocking methods with a higher error rate (e.g., Android unlock pattern), as long as they allow faster input of the unlocking secret.

Hintze et al. [10] investigated the number of interactions per day, the average interaction duration, and the total daily device usage time by using a state machine based on screen on/off events. Here the authors analyzed mobile device data logs from 1,960 Android smartphones (from the Device Analyzer project [16]). The authors report that on average, participants interacted with their devices 57 times a day, among which 43% were actual unlocks, and the daily device usage time was 117 minutes.

In a month-long field study done by Harbach et al. [5], the authors collected data from a subset of PhoneLab users, all of which were affiliated with a university. The authors instrumented LG Nexus 5 smartphones to study the performance of Android unlocking mechanisms in situ. They found that "PIN" users take longer to unlock while committing fewer errors than "pattern" users, who tend to unlock more frequently and are more prone to errors. However, on average, "PIN" and "pattern" users spend a similar amount of total unlocking time. In addition, the authors offer a benchmark against which newly designed unlocking mechanisms can be evaluated.

Harbach et al. [6] also conducted a global-scale survey on Google Consumer Surveys (GCS) with 8,286 participants from 8 countries to investigate whether users' attitudes towards smartphone unlocking differed between various nationalities. The findings pointed towards the need that demographic differences, including both nationality and age, should be considered when designing new authentication systems for smartphones. The authors also conclude that despite the differences between nationalities, inconvenience of unlocking is still one of the major reasons for low adoption rate of current authentication systems, especially for older users.

While previous work provided manifold insights on smartphone authentication in situ, all the aforementioned real world studies suffer from samples skewed towards predominantly tech-savvy, young participants. That is, they do not provide data to understand whether and how age is linked to smartphone authentication. To fill this gap, we conducted a field study with a more representative participant pool with regard to age and other factors. We focused our analysis on how smartphone authentication behaviour correlates with participants' age, while considering the effect of other demographic covariates like gender.

3 METHODOLOGY

At the beginning of the study, participants installed a custom-built application on their smartphones and ran it for 60 days or more (the first 60 days were used for the analysis). The study app ran in the background and collected relevant usage statistics.

At first launch, the application presented the consent form and allowed participants to opt out of providing certain types of data (e.g., activity data). Afterwards, the app directed participants to an entry survey to collect basic demographic data. In addition, participants were asked to report which unlocking mechanism they were currently using.

During the study, the app recorded all lock and unlock events, whether they were auto locks or manual locks, and logged the start and end time stamps of each user session. In addition, the app collected user activity data such as whether unlocking happened while being still or on the move (if opted-in). We define a *user session* as the time between a device unlock and the corresponding lock event (either auto lock or manual).

In order to make sure that the data collection process was robust, we conducted a pilot study with six participants for 15 days and fixed the (few) bugs we found. We applied and obtained approval from the research ethics board in our university before conducting the pilot.

Data Transmission

For data confidentiality, we encrypted all data logged throughout the day with a symmetric encryption key, generated at run time. We encrypted this key with a hard-coded public key, and then appended it to the encrypted data logs before submitting them to the back-end server. Encrypted logs were uploaded to our back-end server once a day, around midnight.

Data Analysis

Our research objective was to investigate how (un)locking behaviours, such as choice of unlocking mechanism or error likelihood, correlate with age. For this, we also took into account other demographic covariates like gender. To answer this research question, we conducted regression analyses among different variables. We applied (multiple) linear regression models for continuous response variables, and logistic regression models for categorical response variables, with a p-value threshold of 0.05. We validated all required assumptions, e.g., normality of residuals and no/little multicollinearity, before applying the the regression models. We performed data transformation, e.g., natural log, square root, on the response variables, and data centralization on the independent variables when necessary, to assure the data meets all assumptions required for applying corresponding regression models. For the analysis, we averaged participants' data across the 60 days whenever there were several data points. For example, we averaged session length across all recorded sessions per participant. Details of all data metrics (e.g., how we computed the variables for each statistical test) can be found in Section 5.

4 PARTICIPANTS

We recruited participants from North America (US and Canada) through Amazon Mechanical Turk (MTurk), Twitter, Facebook, our university mailing lists, and The Sample Network (TSN),[1] etc. In order to be eligible for the study, participants had to be Android users and 19 years or older. To prevent potential recruitment biases, we avoided using terms like "authentication", "privacy" and "security" in our advertisement.[2]

Out of 185 people that completed the study, we had to remove some for different reasons including a technical issue due to which our back-end server was silently failing for 10 days. Another reason for removal was answer ambiguity in the start and exit surveys. These exclusions resulted in a pool of 137 participants. In addition, we removed the 3 password users from the statistical analysis due to being too small of a sample. We thus report our analysis of data from 134 participants. As shown in Table 1, the majority of our participants (57%) was recruited from TSN, while 27% were recruited through MTurk and the rest were from the other five platforms. All data used in the analysis was collected between December 8, 2016 and August 10, 2017.

Each participant who ran the study app for 60 days received a compensation of USD 40 and was entered into a raffle for an iPad Pro 2. We additionally provided a report to each participant with a summary of how they used their smartphone during the study.

We compared our study sample to the smartphone population in the US reported by the Pew Research Center [2]. Statistical results did not reveal any significant differences between our participants' demographics and those presented in the Pew Research Center report, in terms of age (χ^2 = 20, p = .22), gender (χ^2 = 2, p = .16), education levels (χ^2 = 15, p = .24), and annual salary (χ^2 = 20, p = .22). Based on this, we claim that our study sample was relatively representative. All demographics are shown in Table 1.

Parameter	Property	# of participants
Gender	Female	79
	Male	55
Age	19-24	15
	25-34	42
	35-44	26
	45-54	30
	55-63	21
Education	Less than High School	1
	High School	56
	Professional School	23
	University (Bachelor's)	33
	Master or PhD	16
	Other	5
Occupation	Managers	8
	Professionals	29
	Clerical Support Workers	16
	Service and Sales Workers	19
	Craft and Trades Workers	7
	Machine Operators	1
	Elementary Occupations	3
	Students	15
	Self-employed	4
	Unemployed/Retired/Disabled	32
Annual salary	Less than $30,000	42
	$30,000-$49,999	21
	$50,000-$74,999	35
	$75,000-$99,999	15
	$100,000+	16
	Prefer not to specify	5
Recruiting platform	The Sample Network	77
	Amazon Mechanical Turk	36
	Mailing Lists at the University of British Columbia	15
	Other (e.g., Twitter)	6

Table 1: Participant demographics, N = 134.

5 RESULTS

As mentioned before, we conducted regression analyses to identify usage patterns that were significantly correlated with age, while considering other demographic covariates

[1] The Sample Network was used for recruitment in the US only.
[2] Full content of the advertisement can be found in the supplementary material.

(gender, education, occupation, annual salary, and unlocking mechanism). For example, we checked whether there is a correlation between used unlocking mechanism and age, with gender as a covariate. For each interaction pattern, we applied different combinations of the independent variables (e.g., age and unlocking mechanism, age and gender). We report the best fitting prediction models, based on metrics including R-squared, adjusted R-squared, and predicted R-squared. Table 2 shows an overview of these results and the adjustments used (e.g., for "session length" the best fit was achieved by applying log).

Authentication Mechanism Selection

To understand how age correlates with used unlocking mechanism, we conducted a multinomial logistic regression analysis of correlation between the types of unlocking mechanisms and age, with gender as a covariate. We used "Swipe/None" as the reference category for lock type and "Female" for gender. We found that age significantly correlated with participants' choice of "Fingerprint" vs. "Swipe/None", while there was no significant difference in their choice of "Android Pattern" and "PIN" vs. "Swipe/None". With one-year increase in age, our participants were 6.3% less likely to choose "Fingerprint" as their unlocking mechanism, than "Swipe/None". In addition, we found that gender significantly correlated with participants' choice of "PIN" vs. "Swipe/None", but not their choice of "Android Pattern" and "Fingerprint" vs. "Swipe/None". Specifically, male participants were 230.8% more likely to choose "PIN" than "Swipe/None". Table 3 presents the details of the fitted model.

Error Rates

Since "Swipe/None" users cannot make unlocking errors due to the nature of the method, we removed these participants from the error analysis, which reduced the analyzed sample to 87 participants. We fitted a multiple linear regression model of the square root of error rates to investigate if error rates correlated with age. While applying the model, we considered the correlation between the error rates and age, the type of unlocking mechanism, as well as the interaction

Interaction Pattern	*Age Matters?*
Used Unlocking Mechanism	Yes
sqrt(Error Rates)	No
Auto/Manual Locks	Yes
log(Session Lengths)	Yes
log(Number of Sessions Per Day)	Yes
sqrt(Daily Usage Times)	Yes
Whether authentication happened at still / on the move	Yes

Table 2: Overview of the usage patterns that were and were not significantly correlated with age.

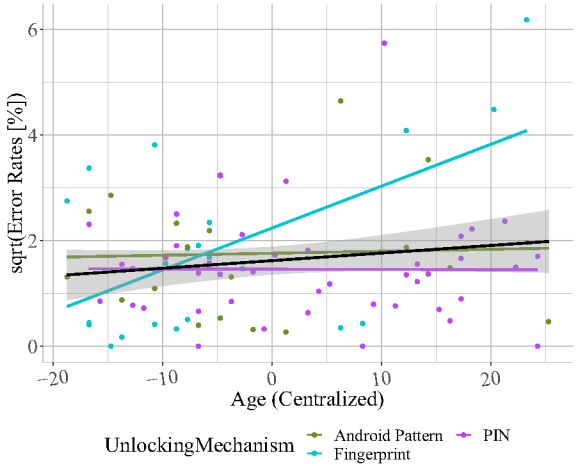

Figure 1: Multiple linear regression model of the square root of error rates on age (centralized), unlocking mechanism (base category "Android Pattern") and their interaction effect: $sqrt\,(ErrorRates) = 1.758 + 0.004 * CentralizedAge + 0.477 * Fingerprint - 0.302 * PIN + 0.076 * CentralizedAge * Fingerprint - 0.004 * CentralizedAge * PIN$, $N = 87$ (*Fingerprint* **and** *PIN* **are binary variables with values {0, 1}). The black line represents the fitted model, and the grey area represents the 95% confidence interval. The other lines represent the simple linear regression models of the square root of error rates on age for participants in each unlocking mechanism group.**

effect between those two factors. We used "Android Pattern" as the reference category. The results show that while the main effect of both age and unlocking mechanism were not significant ($F\,(1, 81) = 2.07$, $p = .15$ vs. $F\,(2, 81) = 1.63$, $p = .20$ respectively), and the interaction effect between age and unlocking mechanism was significant ($F\,(2, 81) = 5.13$, $p = .008$). Specifically, "Fingerprint" participants were more likely to have unlocking errors than "Android Pattern" participants, with the likelihood increasing for older participants. Overall, the model explained 10.98% of the variance ($R2 = .16$, $F\,(5, 81) = 3.12$, $p = .01$). Figure 1 shows the interaction effects and gives an overview of the fitted model.

Auto/Manual Locks

We removed "Swipe/None" participants from the analysis on auto/manual locks, because in Android OS, the auto lock setting is not enabled for them. Furthermore, we removed 25 participants who manually disabled auto lock, either at the beginning of the study or during the study. Overall, we analyzed locking behavior based on data collected from 62 participants.

By default, auto lock is set to 5 seconds. Our analysis revealed that out of the 62 participants, 26 kept the default auto lock timeout, while 18 reduced the auto lock time to 0,

Variable	*Estimate*	*Std. Err.*	*z value*	*Odds Ratio*	*Estimate*	*Std. Err.*	*z value*	*Odds Ratio*	*Estimate*	*Std. Err.*	*z value*	*Odds Ratio*
	Swipe/None vs. Android Pattern				**Swipe/None vs. Fingerprint**				**Swipe/None vs. PIN**			
Intercept	0.390	0.959	0.407	1.477	1.214	0.977	1.242	3.366	-0.415	0.808	-0.514	0.66
Age	-0.034	0.023	-1.49	0.966	-0.065	0.025	-2.572*	0.937	-0.002	0.018	-0.126	0.998
Gender=Male	0.341	0.585	0.584	1.407	1.056	0.561	1.884	2.876	1.196	0.452	2.644*	3.308

Table 3: Multinomial logistic regression model: how age and gender correlated with participants' choice of unlocking mechanism. A "*" denotes significance ($p < .05$).

which locks the device immediately after it enters sleep mode. Since we were unable to differentiate auto locks from manual locks when the timeout was set to 0, we further excluded those 18 participants from the analysis. This reduced our participants pool for this analysis down to 44.

To evaluate whether participants relied more on auto locks or manual locks, we calculated the percentages of auto locks over the total number of locks per day per user. We fitted a multiple linear regression model to investigate if the average percentage of auto locks correlated with age, while considering gender (with "Female" serving as the base category) and its interaction with age as covariates. The results in Figure 2, show that only the main effect of age ($F(1, 40) = 7.46$, $p = .009$) was significantly correlated with people's locking behaviors, but not the main effect of gender ($F(1, 40) = 1.86$, $p = .18$) and the interaction effect between age and gender ($F(1, 40) = 2.26$, $p = .14$). Specifically, older people relied more on auto locks, with an increase of 5.09% for females and 16.06% for males for every 10-year increment of age. Overall, age explained 16.63% of the variance ($R2 = .22$, $F(3, 40) = 3.86$, $p = .02$).

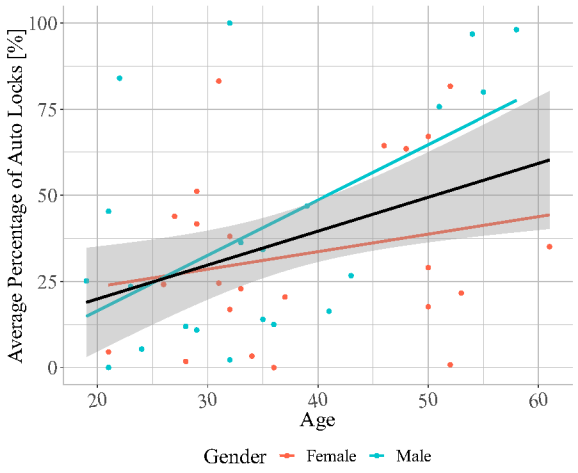

Figure 2: Multiple linear regression model of the percentages of auto locks on age, gender (base category "Female") and their interaction effect: $Percentage of AutoLocks = 13.303 + 0.509 * Age - 28.906 * Male + 1.097 * Age * Male$, $N = 44$ **(*Male* is a binary variable with values {0, 1}). The black line represents the fitted model, and the grey area represents the 95% confidence interval. The other lines represent the simple linear regression models of the percentage of auto locks on age for female and male participants.**

Session Lengths

The analyzed dataset contained 257,437 user sessions in total. On average, a session lasted 10.89 minutes ($SD = 12.80$ minutes). We applied multiple linear regression analysis to test if age, gender and the interaction effect between those two factors correlated with the log of averaged session length. Figure 3 gives an overview of the fitted model, with "Female" serving as the reference category. The results indicate that the main effect of both age and gender, as well as their interaction effect were significantly correlated with the log value of session length ($F(1, 129) = 6.72$, $p = .01$ vs. $F(1, 129) = 6.91$, $p = .01$ vs. $F(1, 129) = 6.31$, $p = .01$ respectively), and those factors accounted for 11.37% of the variance ($R2 = .13$, $F(3, 129) = 6.65$, $p < .001$). Furthermore, a 10-year increase in age corresponds to a growth of 0.9% for female participants and 41.76% for male participants in session length.

Number of Sessions Per Day

On average, participants had 32 smartphone sessions a day ($SD = 26$). We applied multiple linear regression analysis to test if age and unlocking mechanism correlated with the average number of sessions each day. Figure 4 gives an overview of the fitted model, with "Swipe/None" serving as the reference category. The results reveal that both age ($F(1, 129) = 37.99$, $p < .001$) and unlocking mechanism ($F(3, 129) = 4.14$, $p = .008$) were significant factors. Younger participants interacted with their devices more frequently than older participants (i.e., a 10-year increase in age corresponds to a deduction of 25.17% in total number of sessions per day), and "Swipe/None" users had significantly more sessions (approximately 39.51%) than those using "PIN". Overall, the two independent variables explained 25.87% of the variance ($R2 = .28$, $F(4, 129) = 12.60$, $p < .001$).

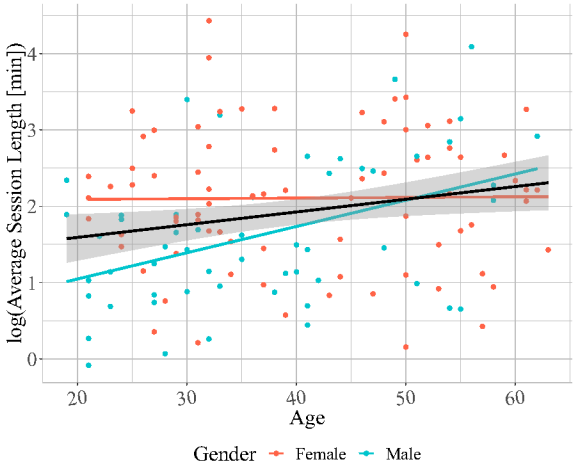

Figure 3: Multiple linear regression model of log of average session length on age, gender (base category "Female") and their interaction effect: $log\,(AverageSessionLength) = 2.073 + 0.0009 * Age - 1.715 * Male + 0.034 * Age * Male$, $N = 133$ **(*Male* is a binary variable with values {0, 1}). The black line represents the fitted model, and the grey area represents the 95% confidence interval. The other lines represent the simple linear regression models of the percentage of auto locks on age for female and male participants.**

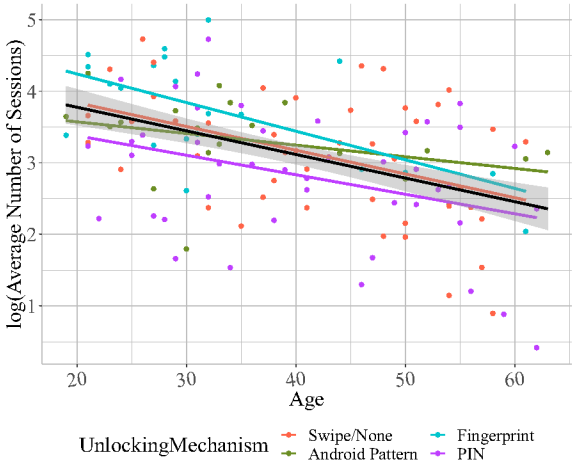

Figure 4: Multiple linear regression model of log of average number of sessions on age and unlocking mechanism (base category "Swipe/None"): $log\,(AverageNumberofSessions) = 4.338 - 0.029 * Age + 0.037 * (AndroidPattern) + 0.353 * Fingerprint - 0.333 * PIN$, $N = 134$ **(*AndroidPattern*, *Fingerprint* and *PIN* are binary variables with values {0, 1}). The black line represents the fitted model, and the grey area represents the 95% confidence interval. The other lines represent the simple linear regression models of the log of average number of sessions on age for participants in each unlocking mechanism group.**

Daily Usage Times

We fitted a multiple linear regression model among the square root of the average amount of daily device usage time for participants and the three factors age, unlocking mechanism, and gender, while using "Swipe/None" and "Female" as the reference categories. The results show that both age and gender significantly correlated with daily usage time ($F\,(1, 127) = 6.55$, $p = .01$ vs. $F\,(1, 127) = 7.74$, $p = .006$ respectively), whereas unlocking mechanism did not ($F\,(3, 127) = 1.36$, $p = .26$). In particular, older participants tended to interact with their smartphones for a significant shorter amount of time per day, while female participants tended to have longer daily usage times than male participants, as shown in Figure 5. Overall, the model explained 9.20% of the variance ($R2 = 0.13$, $F\,(5, 127) = 3.68$, $p = .004$).

Activity While Unlocking

Of the 134 study participants, 116 opted in to provide activity data. To identify user activities during unlocks, we mapped their timestamps. We used one minute as a threshold, meaning that if there were no activity records that occurred one minute before or after an unlock event, then we considered that we failed to detect the type of activity for this unlock. When there were multiple activity records within the one-minute timeframe, we selected the closest (time-wise) activity as the one that user undertook during the unlock.

Among all unlocking events collected for the 116 participants, we removed the unlocks that were tagged with "unknown" activity and those in which we failed to detect activities for, which were around 50.1%. Figure 6 shows the distribution of activity types over the remaining 49.9% of unlocks (106,030 in total). We found that more than half of these unlocks happened when participants' devices were still, whereas the other 43.8% of unlocks happened while the device was moving. Only 0.26% of unlocks happened while participants were on a bicycle. The distributions of activities for each unlocking mechanism did not differ notably from each other.

We removed the activity type "tilting",[3] which we considered as not significant enough of a move to influence unlocking performance. Then we categorized all activities except *still* as *move*. As shown in Table 2, we found that age was a good predictor for whether authentication was used while at *still* or on the *move* (details below).

[3]Tilting the device around the horizontal or vertical axes.

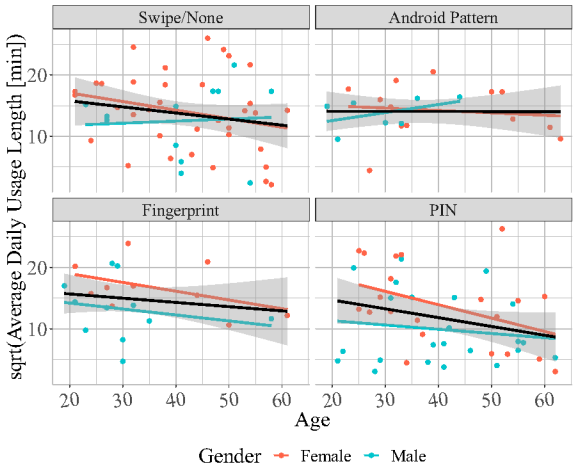

Figure 5: Multiple linear regression model of the square root of average daily usage length on age, unlocking mechanism (base category "Swipe/None") and gender (base category "Female"): $sqrt\,(AverageDailyUsageLength)$=18.869-0.108∗ Age+0.161 ∗ ($AndroidPattern$)+1.019 ∗ $Fingerprint$-1.256 ∗ PIN-2.867∗$Male$, N = 133 **(**$AndroidPattern$, $Fingerprint$, PIN **and** $Male$ **are binary variables with values {0, 1}). The black line represents the fitted model for participants in each unlocking mechanism group, and the grey area represents the 95% confidence interval. The other colored lines represent the simple linear regression models of the square root of average daily usage length on age for female participants and male participants in each unlocking mechanism group.**

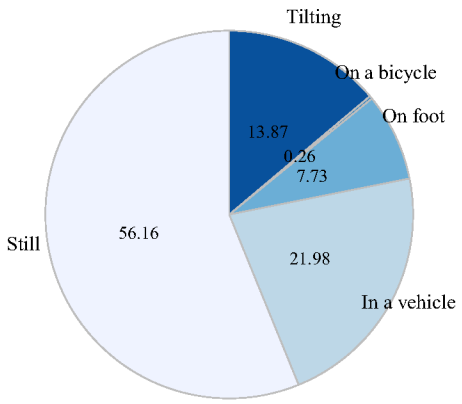

Figure 6: Distribution of activities during unlock, N = 116.

Error Rates While At Still vs. On the Move. To assess how error rates correlated with activity types, we again excluded participants who used "Swipe/None", as errors are not possible with this unlocking method.

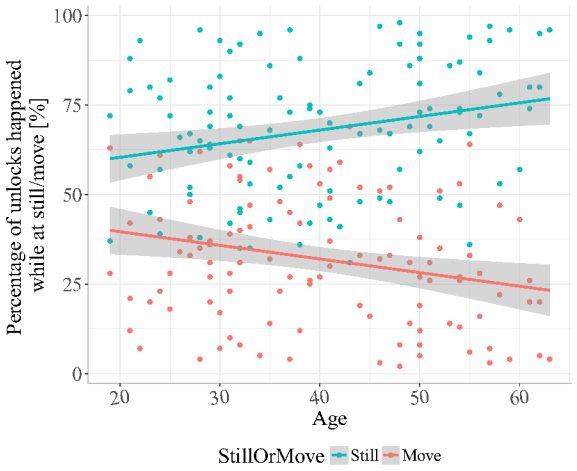

Figure 7: Distribution of unlocks that happened while at still/move, N = 116**. The lines represent the fitted simple linear regression model of the percentage of unlocks on age for each activity category, and the grey area represents the 95% confidence interval.**

We then counted the number of failed unlocking attempts and total unlocking attempts that each participant performed, while they were in "still" and "move" states. Afterwards, we calculated the likelihood of making an error, for each participant for each of the two activity categories. However, a paired-samples t-test did not reveal any significant differences in the likelihoods that participants would make an error, while their devices were *still* (M = 4.37, SD = 7.16) and *moving* (M = 4.78, SD = 7.27); t (70) = 1.08, p = .28.

How Age Predicts Whether Authentication is Used While At Still vs On the Move. To evaluate how often participants unlocked while they were in still and moving states, we calculated the percentage of unlocks for each activity category per user over the whole study period. Figure 7 shows the distribution of unlocks at still and on the move, with respect to participants' age. To further understand how age interacted with activity types (*still/move*), we fitted a binary logistic regression model to predict whether authentication would be used when users were at still and on the move based on age, with *move* as the reference activity category. The analysis revealed that age had a significant effect on predicting whether authentication happened while at still or on the move. Specifically, with one-year increase in age, participants were 1.5% more likely to use authentication when they were *still* than moving. Table 4 gives an overview of the fitted model.

Variable	*Estimate*	*Std. Err.*	*z value*	*Odds Ratio*
Intercept	.078	.024	3.268*	1.081
Age	.015	.001	23.772*	1.015

Table 4: Binary logistic regression model: whether authentication is used while at still vs on the move by age. An "*" denotes significance ($p < .05$).

6 DISCUSSION

The results presented in this paper provide the first real world data and detailed insights on the link between individual smartphone authentication patterns and age, thus, corroborating and extending on the previous study [6] that hypothesized that age might be an important factor. In addition, we investigated a broader scope of unlocking mechanisms, than previous work [5, 6, 10, 12], in terms of how participants' in-situ (un)locking behaviour differs among unlocking mechanisms. We are the first to extend these findings to "Fingerprint" users, despite that we had relatively smaller "Fingerprint" sample size (about 56% of participants' devices supporting such mechanism).

Age Makes a Difference

First of all, and most importantly, our data suggests that smartphone usage patterns, indeed, in many cases significantly correlate with age.

Specifically, we found a significant correlation of age with the following parameters: the choice of unlocking mechanism, session length, number of sessions per day, daily usage length, auto/manual lock, state of movement (at still vs. on the move), the choice of unlocking mechanism at still vs. on the move. In addition, we found that participant's gender also significantly correlated with their session length, daily usage length, and the choice of the unlocking mechanism.

That is, our results provide the first quantitative support (based on field measurements of real-world behaviour) for previous recommendations that, while designing new authentication systems for smartphones, demographic differences, such as age, should be taken into account [6]. Furthermore, we emphasize that our work not only confirms the link between age and in-situ authentication behaviour, but also extends the findings to gender. We thus invite more studies to investigate (other) demographic factors, and suggest future work could focus on recruiting even more representative (particularly age-wise) samples to provide further evidence.

Age and Unlocking Security

One of our results related to age being a good predictor for some choices related to smartphone unlocking mechanism selection, is that older participants were less likely to enable "Fingerprint" authentication on their devices. In addition, older participants significantly more likely relied on the auto lock feature, despite the used unlocking mechanisms, which is in line with [5].

The auto lock feature very specifically represents a trade-off between security and usability. The related question is whether users make educated choices, when making these decisions, or whether other factors are more prominent in the decision-making process.

As a consequence, we argue that future research should look into the popularity of auto lock among older users and their awareness of respective security implications.

Design and Evaluate for Movement

Another interesting result to highlight is that age was a good predictor for whether device unlocking was used on the move compared to at still.

Our results show that, overall, about 44% of unlocks happened with the devices moving. Furthermore, despite a trend in our data towards more authentication taking place in still states with increasing age, authentication on moving devices was a common task across all age groups.

While one might hypothesize that participants were more likely to make unlocking errors while moving, our findings do not support such a claim. The numbers of failed unlocking attempts when participants' devices were still and when they were moving were very similar, also across age groups. This might indicate that current unlocking mechanisms are robust against errors on the move.

A consequence of this and the fact that people do regularly authenticate themselves while moving, means that newly-designed unlocking mechanisms should be as robust in the presence of movement as existing mechanisms, in order to be acceptable for smartphone unlocking. This requirement not only influences how unlocking systems should be designed (e.g., constant eye contact might not be possible) but also how they are evaluated in studies (in-lab vs. in the wild).

No One-Fits-All Solution

As mentioned before, we showed that age significantly correlated not only with the choice of unlocking mechanism but also whether and how features such as auto lock were used and whether authentication took place on still or moving devices.

Previous research [7] has concluded that due to user preferences, it is hard (or impossible) to find or build one authentication system that caters to all user needs. Our results add to these previous findings by providing evidence that these usage differences are also linked to age and gender.

While we cannot make claims about the cause of these differences, our findings suggest that these differences are important factors to consider. The findings further our understanding of why offering different unlocking mechanisms

for smartphones seems to be the right thing to do. Users have different preferences and usage behaviour, and this approach allows users to pick the mechanism that best fits their needs and the way they are using their devices.

7 LIMITATIONS

While we tried our best to mitigate any potential problems with the study setup, the study has a few limitations that need to be kept in mind when interpreting the results.

First, our server was silently down for 10 days at the end of May, when most of our participants were supposed to complete the study.This unexpected issue forced us to exclude those participants (N = 43) whose data we failed to recover. This could have potentially biased our study sample. However, we do not consider this accident as a major threat to the validity of our results, since we did not find a significant difference between our sample demographics distributions and the US smartphone ownership distributions reported by Pew Research Center.

In addition, the use of MTurk for recruiting could potentially bias our study sample, since MTurk workers tend to be more privacy self-aware than average people [11]. However, we argue that avoiding "privacy" and "security" terms in our advertisement and only a small portion (about one fourth) of participants of the sample coming from MTurk helped limit this bias. Again, as our sample demographic distributions were not significantly different from the US smartphone population, we do not consider this affecting our results' validity either.

To determine session length, we measured the time between the screen on and off (with *keyguard* removed) events. Since this measurement did not remove the screen off timeout from the session length calculations, the reported length of sessions could potentially be longer than of the real sessions. Our comparisons on session lengths among groups might also be biased, as different participants would have set the screen off timeout to different values, varying from 0 seconds to 24 hours.

Due to technical limitations, we were not able to detect the used unlocking mechanisms and the states of auto lock setting programmatically for participants using Android 6 (N = 64) and 7 (N = 31) devices. Therefore, we asked those participants about their unlocking mechanisms (Android 6 and above users) and the auto lock settings (Android 7 and above) in the entry and exit surveys. As a consequence, the correctness of our collected data partially depends on the recollection of the participants. However, the data gives us no reason to believe that they would have incorrectly reported these attributes.

On average, the confidence level of detected activities (still, in a vehicle, on foot, etc.) is 73.31 (SD = 25.16, Min = 21, $Median$ = 75, Max = 100). Therefore, the results of activity analysis are highly dependent on the accuracy of the Android activity recognition API.

Finally, the small "Fingerprint" sample size (75 participants' devices supported Fingerprint-based authentication) could limit the generalizability of our Fingerprint-related findings. Another caveat is that the awareness of biometrics authentication scheme among individuals, especially the seniors, could bias our results as well.

8 CONCLUSION

In this paper, we add to the research field of real world unlocking behavior by providing the first detailed look at how age correlates with smartphone users' in situ unlocking behavior. We conducted a longitudinal field study with 134 participants from North America, who installed our study app on their Android phones and ran it for 60+ days. The results of the study suggest that age does indeed significantly correlate with different factors related to smartphone unlocking. In particular, we observed that older users interacted with their devices less frequently. In addition, older participants relied more on the auto lock features and were less likely to enable "Fingerprint" authentication on their devices. We conclude that when designing new authentication systems, varying age-related usage patterns should be taken into consideration.

Based on our results, an interesting area for future work is to explore security perceptions of unlocking mechanisms across age groups. This would allow to identify whether the decisions that people make are conscious or whether they are influenced by other parameters such as usability related factors.

9 ACKNOWLEDGMENTS

This research was supported by funding from Google and the NSERC Engage Grant (EG). We would like to thank the anonymous reviewers for their helpful insights and suggestions.

REFERENCES

[1] Daniel Buschek, Fabian Hartmann, Emanuel Von Zezschwitz, Alexander De Luca, and Florian Alt. 2016. SnapApp: Reducing Authentication Overhead with a Time-Constrained Fast Unlock Option. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3736–3747. http://doi.acm.org/10.1145/2858036.2858164

[2] Pew Research Center. 2013. Smartphone Ownership 2013. http://www.pewinternet.org/2013/06/05/smartphone-ownership-2013/

[3] Ivan Cherapau, Ildar Muslukhov, Nalin Asanka, and Konstantin Beznosov. 2015. On the Impact of Touch ID on iPhone Passcodes. In *Proceedings of the 2015 Symposium on Usable Privacy and Security (SOUPS '15)*. USENIX Association, Ottawa, ON, Canada, 257–276. https://www.usenix.org/conference/soups2015/proceedings/presentation/cherapau

[4] Serge Egelman, Sakshi Jain, Rebecca S. Portnoff, Kerwell Liao, Sunny Consolvo, and David Wagner. 2014. Are You Ready to Lock? Understanding User Motivations for Smartphone Locking Behaviors. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*. ACM, New York, NY, USA, 750–761. http://doi.acm.org/10.1145/2660267.2660273

[5] Marian Harbach, Alexander De Luca, and Serge Egelman. 2016. The Anatomy of Smartphone Unlocking: A Field Study of Android Lock Screens. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4806–4817. http://doi.acm.org/10.1145/2858036.2858267

[6] Marian Harbach, Alexander De Luca, Nathan Malkin, and Serge Egelman. 2016. Keep on Lockin' in the Free World: A Multi-National Comparison of Smartphone Locking. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4823–4827. http://doi.acm.org/10.1145/2858036.2858273

[7] Marian Harbach, Emanuel Von Zezschwitz, Andreas Fichtner, Alexander De Luca, and Matthew Smith. 2014. It's a Hard Lock Life: A Field Study of Smartphone (Un)Locking Behavior and Risk Perception. In *Proceedings of the 2014 Symposium on Usable Privacy and Security (SOUPS '14)*. USENIX Association, Menlo Park, CA, USA, 213–230. https://www.usenix.org/conference/soups2014/proceedings/presentation/harbach

[8] Eiji Hayashi, Sauvik Das, Shahriyar Amini, Jason Hong, and Ian Oakley. 2013. CASA: Context-Aware Scalable Authentication. In *Proceedings of the 2013 Symposium on Usable Privacy and Security (SOUPS '13)*. ACM, New York, NY, USA, 3:1–3:10. http://doi.acm.org/10.1145/2501604.2501607

[9] Eiji Hayashi, Oriana Riva, Karin Strauss, A. J. Bernheim Brush, and Stuart Schechter. 2012. Goldilocks and the Two Mobile Devices: Going Beyond All-Or-Nothing Access to a Device's Applications. In *Proceedings of the 2012 Symposium on Usable Privacy and Security (SOUPS '12)*. ACM, New York, NY, USA, 2:1–2:11. http://doi.acm.org/10.1145/2335356.2335359

[10] Daniel Hintze, Rainhard D. Findling, Muhammad Muaaz, Sebastian Scholz, and René Mayrhofer. 2014. Diversity in Locked and Unlocked Mobile Device Usage. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct)*. ACM, New York, NY, USA, 379–384. http://doi.acm.org/10.1145/2638728.2641697

[11] Ruogu Kang, Stephanie Brown, Laura Dabbish, and Sara Kiesler. 2014. Privacy Attitudes of Mechanical Turk Workers and the U.S. Public. In *Proceedings of the 2014 Symposium on Usable Privacy and Security (SOUPS '14)*. USENIX Association, Menlo Park, CA, USA, 37–49. https://www.usenix.org/conference/soups2014/proceedings/presentation/kang

[12] Ahmed Mahfouz, Ildar Muslukhov, and Konstantin Beznosov. 2016. Android Users in the Wild: Their Authentication and Usage Behavior. *Pervasive and Mobile Computing* 32 (2016), 50–61.

[13] Diogo Marques, Ildar Muslukhov, Tiago Guerreiro, Luís Carriço, and Konstantin Beznosov. 2016. Snooping on Mobile Phones: Prevalence and Trends. In *Proceedings of the 2016 Symposium on Usable Privacy and Security (SOUPS '16)*. USENIX Association, Denver, CO, USA, 159–174. https://www.usenix.org/conference/soups2016/technical-sessions/presentation/marques

[14] Ildar Muslukhov, Yazan Boshmaf, Cynthia Kuo, Jonathan Lester, and Konstantin Beznosov. 2013. Know Your Enemy: The Risk of Unauthorized Access in Smartphones by Insiders. In *Proceedings of the 2013 International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '13)*. ACM, New York, NY, USA, 271–280. http://doi.acm.org/10.1145/2493190.2493223

[15] Oriana Riva, Chuan Qin, Karin Strauss, and Dimitrios Lymberopoulos. 2012. Progressive Authentication: Deciding When to Authenticate on Mobile Phones. In *Proceedings of the 2012 USENIX Security Symposium (USENIX Security '12)*. USENIX Association, Bellevue, WA, USA, 301–316. https://www.usenix.org/conference/usenixsecurity12/technical-sessions/presentation/riva

[16] Daniel T. Wagner, Andrew Rice, and Alastair R. Beresford. 2014. Device Analyzer: Large-Scale Mobile Data Collection. *ACM SIGMETRICS Performance Evaluation Review* 41 (2014), 53–56.