

Unlocking Personal Characteristics: Harnessing Keystroke Dynamics for Identification on Mobile Devices

James Campbell

A thesis presented for the degree of
Doctor of Philosophy



University of East Anglia
School of Computing Sciences

September 2024

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

Abstract

The widespread adoption of mobile devices has led to rapid technological advancements. These advancements have made the addition of motion sensors, such as accelerometers and gyroscopes, critical components of a mobile device.

These sensors offer novel methods of identification, which can be employed to harness this technology providing a transparent, yet more secure process. One such method, keystroke dynamics, has been present across more traditional physical devices, such as keyboards for a long time, and has more recently bridged across into virtual keyboards and mobile devices.

This research explores the potential of keystroke dynamics, augmented by motion sensor data, to infer personal characteristics such as name, age, gender and handedness with high accuracy. Three key research questions are addressed to form the basis of the thesis:

1. To what extent can keystroke dynamics be utilised in order to infer a person's name on a mobile device?
2. What effect does the inclusion of accelerometer and gyroscopic data alongside keystroke dynamics have on the ability to successfully in-

fer a person's name and soft biometric features on a mobile device?

3. To what extent does the volume of data per user help to improve the accuracy of the prediction of name and soft biometric features?

An innovative approach to identification is presented alongside key contributions of the work, which include:

- Novel approach to identity data using smartphone motion sensors.
- A new methodology and experimental approach (including data capture framework).
- Bespoke data sets of motion data that can be anonymised and shared with the wider community.
- A novel algorithm for predicting a letter utilising smartphone motion sensors.

A comprehensive methodology is employed, combining data collection across multiple studies with detailed machine learning and manual analysis to provide high accuracy scores. The results demonstrate that name can be inferred with significant accuracy (83.20%), whilst also achieving high accuracy rates for age (87.74%), gender (82.56%) and handedness (93.18%). Furthermore, increasing the volume of collected data per user led to further improvements in prediction accuracy.

The findings highlight the potential for sensor-enhanced identification, offering improved security while maintaining usability and transparency. This research contributes to the field of behavioural biometrics

by advancing the understanding of sensor-enhanced mobile-based identity inference, paving the way for more secure, transparent and seamless methods of identification.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Contents

1	Introduction	18
1.1	Motivation	18
1.2	Contributions	20
1.3	Thesis Organisation	21
2	Technical Background and Related Work	23
2.1	Literature Review Methodology	23
2.2	Background Information	25
2.2.1	Accelerometers	27
2.2.2	Gyroscopes	28
2.2.3	Combining Accelerometers and Gyroscopes	29
2.3	Quality of Information	30
2.4	Identification and Authentication	33
2.4.1	User Identification	33
2.4.2	Authentication	34
2.4.3	Continuous Authentication	39
2.5	Behavioural Biometrics	41
2.5.1	Keystroke Dynamics	51
2.5.2	Soft Biometrics	55

2.6	Keystroke Dynamic Inference Utilising Accelerometer and Gyroscope	57
2.7	Machine Learning	68
2.7.1	Machine Learning in Relevant Studies	70
2.7.2	Metrics	73
2.8	Identification Ethics	75
2.9	Summary	77
3	Data Collection Application	81
3.1	Experimental Design	81
3.2	Pilot Study Application	82
3.2.1	Implementation	82
3.3	Main Studies Application	96
3.3.1	Implementation	96
3.4	Google Play Store	103
3.4.1	Pilot Study	104
3.4.2	Main Study 1	106
3.4.3	Main Study 2	108
3.5	Datasets	110
3.6	Summary	112
4	Pilot Study	113
4.1	Introduction	114
4.2	Participants	114
4.3	Data Preparation and Analysis	117
4.4	Findings	119
4.4.1	Soft Biometric Findings	119

4.4.2	Keystroke Dynamic Findings	123
4.5	Discussion	127
5	Evaluation of the addition of further sensor input to en- hance accuracy	130
5.1	Introduction	131
5.2	Participants	131
5.3	Data Analysis	133
5.3.1	Data Preparation and Analysis	133
5.3.2	Predictive Model	135
5.3.3	Machine Learning	139
5.4	Findings	141
5.4.1	Soft Biometric Findings	142
5.4.2	Keystroke Dynamic Findings	144
5.5	Discussion	155
6	Evaluation of an increase in data per user to enhance accuracy	158
6.1	Introduction	159
6.2	Participants	159
6.3	Data Preparation and Analysis	162
6.4	Findings	164
6.4.1	Soft Biometric Findings	164
6.4.2	Keystroke Dynamic Findings	166
6.5	Discussion	170
7	Evaluation of a combination of datasets to further en-	

hance accuracy	173
7.1 Combination Findings	173
7.2 Data Preparation and Analysis	175
7.2.1 Soft Biometric Findings	177
7.2.2 Keystroke Dynamic Findings	180
7.3 Discussion	183
8 Conclusions and Future Work	185
8.1 Research Questions	185
8.1.1 Research Question 1	186
8.1.2 Research Question 2	187
8.1.3 Research Question 3	188
8.1.4 Discussion	189
8.2 Comparison of Experiments	190
8.2.1 Soft Biometrics	190
8.2.2 Keystroke Dynamics	191
8.3 Discussion of Practical Applications	192
8.4 Discussion of Contributions	193
8.5 Limitations, Future Work and Research Opportunities	194
8.5.1 Limitations	194
8.5.2 Future Work and Research Opportunities	195
References	197
Appendices	208

List of Tables

2.1	Google Scholar search terms, completed on 30.03.2024 and papers accessed on 01.04.2024 for review of abstract and downloaded if considered for further analysis. Further reviews conducted after this initial access.	24
2.2	Chosen sentences for the data collection as per Figure 2.4, shown in alphabetical order with letter frequency. Characters were not case sensitive in the application.	32
2.3	Suitability comparison of behavioural biometrics. Adapted and reduced from <i>R. V. Yampolskiy and V. Govindaraju. Taxonomy of behavioural biometrics. In Behavioural Biometrics for Human Identification: Intelligent Applications, pages 1–43. IGI Global, 2010. [65]</i>	44
2.4	Explanation of suitability for behavioural biometrics. Based on Table 2.3 taken and adapted from <i>R. V. Yampolskiy and V. Govindaraju. Taxonomy of behavioural biometrics. In Behavioural Biometrics for Human Identification: Intelligent Applications, pages 1–43. IGI Global, 2010. [65]</i>	46

2.5	Generated text vs. actually typed text. <i>Taken from Beta-logger: Smartphone Sensor-based Side-channel Attack Detection and Text Inference Using Language Modelling and Dense MultiLayer Neural Network by Javed et al.[35]. . .</i>	65
2.6	Classification accuracies for two datasets (41 features vs 71 features) <i>Taken and re-created from Keystroke dynamics on Android platform by Antal et al. [4]</i>	72
3.1	Datasets and the corresponding experiments/chapters. . .	111
4.1	Pilot Study - Classifier accuracy, precision, recall and F_1 scores based on an average of 50 runs.	121
4.2	Pilot Study - Name inference classifier accuracy scores - Average of 50 runs (10% Test Size).	123
4.3	Pilot Study - Name inference classifier accuracy scores - Average of 50 runs (15% Test Size).	124
4.4	Pilot Study - Average accuracy and frequency per letter - Manual analysis.	127
5.1	Experiment 2 - (Main Study 1) - Soft biometric classifier accuracy, precision, recall and F_1 scores - Average of 50 runs.	143
5.2	Experiment 2 - (Main Study 1) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (10% Test Size).	146

5.3	Experiment 2 - (Main Study 1) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (15% Test Size).	146
5.4	Experiment 2 - (Main Study 1) - Name inference threshold accuracy scores	150
5.5	Experiment 2 - (Main Study 1) - Name inference improved threshold accuracy scores	154
6.1	Experiment 3 - (Main Study 2) - Soft biometric classifier accuracy, precision, recall and F_1 scores - Average of 50 runs.	165
6.2	Experiment 3 - (Main Study 2) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (10% Test Size).	167
6.3	Experiment 3 - (Main Study 2) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (15% Test Size).	168
6.4	Experiment 3 - (Main Study 2) - Name inference threshold accuracy scores	169
7.1	Experiment 2 and 3 - (Main Studies 1 and 2) - Soft biometric classifier accuracy, precision, recall and F_1 Scores - Average of 50 runs.	178
7.2	Experiment 2 and 3 - (Main Studies 1 and 2) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (10% Test Size).	180

7.3	Experiment 2 and 3 - (Main Studies 1 and 2) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (15% Test Size).	181
7.4	Experiment 2 and 3 - (Main Studies 1 and 2) - Name inference threshold accuracy scores.	182
8.1	Highest accuracy scores per experiment across soft biometrics.	191
8.2	Highest accuracy scores per experiment across name. . .	192

List of Figures

2.1	Smartphone Sales Globally 2007 - 2022/23. Taken from https://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007/	26
2.2	Visual representation of a 3-Axis accelerometer. [40]	27
2.3	Visual representation of a 3-Axis gyroscope. [41]	28
2.4	Chosen Uni-grams to be used in the data collection of the application for all of the experiments proposed.	30
2.5	Most Common Letters in the English Language in Descending Order. Taken from http://letterfrequency.org/	31
2.6	Multi-Factor Authentication (MFA) diagram showing knowledge, token and biometric principals alongside common applications.	36
2.7	A User and machine authentication process. Re-created from <i>Comparing Passwords, Tokens, and Biometrics for User Authentication</i> by L. O’Gorman [43].	38
2.8	The SilentSense Framework. Taken from <i>Continuous User Identification via Touch and Movement Behavioral Biometrics</i> by Bo et al [7].	40

2.9	Keystroke dynamic illustrations showing common measurements. Taken from <i>Identity Theft, Computers and Behavioural Biometrics</i> by Moskovitch et al [42].	54
2.10	Three Steps of KDA framework: enrolment, classifier building, and user authentication. <i>Recreated from Keystroke dynamics-based authentication for mobile devices</i> by Hwang et al [32].	60
3.1	Participant information screen from the data collection application.	84
3.2	Application instructions screen from the data collection application.	84
3.3	Demographic data capture screen from the data collection application.	85
3.4	Experiment data capture screen from the data collection application.	87
3.5	Unique identifier screen from the data collection application.	88
3.6	Pilot study flowchart showing data collection and storage.	95
3.7	Main studies flowchart showing data collection and storage.	103
3.8	Google Play Store analysis graph - Experiment 1.	105
3.9	Google Play Store analysis graph- Experiment 2.	107
3.10	Google Play Store analysis graph - Experiment 3.	109
4.1	Pilot Study - Number of participants for gender, bucketed by age.	115
4.2	Pilot Study - Number of participants split by handedness.	116

4.3	Pilot Study - Percentage coverage (average) sorted by length of name.	126
5.1	Spread of participants data with age and gender - Main Study 1	132
5.2	Spread of participants handedness data - Main Study 1 .	132
5.3	Keyboard quadrants as determined by the predictive model analysis.	136
5.4	Keyboard with quadrants assigned for manual analysis. .	136
5.5	Keyboard with quadrants assigned and X/Y values for manual analysis.	137
5.6	Keyboard with quadrants assigned and median values for manual analysis.	138
5.7	Keyboard with mapped values for manual analysis. . . .	139
6.1	Spread of participants data with age and gender - Main Study 2	161
6.2	Spread of participants handedness data - Main Study 2 .	161
7.1	Spread of participants data with age - Combination of Experiments 2 and 3 - (Main Studies 1 and 2).	174
7.2	Spread of participants handedness data - Combination of Experiments 2 and 3 - (Main Studies 1 and 2).	174

List of Algorithms

1	Novel rotational threshold algorithm	149
2	Improved novel rotational threshold algorithm	152

Acknowledgements

Firstly, I would like to thank my entire supervisory team that has been with me on this journey: Prof. Oli Buckley, Dr Rameez Asif and Prof. Richard Harvey for all their help and patience.

I would also like to thank my amazing friends, who have been beyond supportive over these last few years and have kept me going, so special thanks to: Debbie, James, Gav, Lucy, Dan and Mollie. A huge thank you goes to Sally for her support and guidance, and of course a massive thanks to Danny, who has given me enough laughter and support to last a lifetime.

A very special thank you to my incredible Mum and Dad, who have been a constant source of inspiration and admiration, (and of course Barley as well for her endless need for belly rubs). Also, not forgetting my wonderful dog Paddy who has kept me company 'til the small hours on numerous occasions.

Finally, and most importantly, the biggest thanks I could ever physically give go to my incredible wife Becca. Without her love and support through all the late nights and stressful times, I have no idea how I would have managed. Thank you.

List of Publications

- J.Campbell and O.Buckley. A Glass Case of Emotion: Identity, Motion Sensors, and Your Smartphone. In *Breakthroughs in Digital Biometrics and Forensics*, pages 89-106. Springer, Charm, 2022.
- S. Earl, J.Campbell, O.Buckley. Identifying soft biometric features from a combination of keystroke and mouse dynamics. In *International Conference on Applied Human Factors and Ergonomics*, pages 184-190. Springer, 2021.
- S. Earl, J.Campbell, O.Buckley. Investigating what you share: Privacy perceptions of behavioural biometrics. In *International Conference on Human- Computer Interaction*, pages 408-415. Springer, 2021.

Chapter 1

Introduction

Keystroke dynamics are a type of behavioural biometric which describe the unique typing pattern of a user. Typically, these are measured on a physical keyboard using timings, however, with the increase in mobile device usage [37] virtual keyboards can be used to obtain these measurements. Subsequently, accelerometers and gyroscopes are now regular inclusions in mobile devices, and with these more advanced technologies, come more sophisticated methods of obtaining the required timings.

1.1 Motivation

Identification, and the subsequent authentication of a user is one of the key challenges any system faces, as we discuss in Chapter 2. Many methods of identification exist, with examples such as PINs and passwords widely adopted across all industries. Unfortunately, these methods only prove that the person being identified has the correct password or PIN

and not that they are indeed that person.

Keystroke dynamics have already been identified as a potential solution to this problem, as they make for a more secure method of identification and one that is employed in various scenarios. Recognising unique typing patterns and behaviours, and being able to identify users via this method, can help to detect impostors, even if they have stolen a password. In addition, large amounts of data can be inferred from these typing patterns and, as we present in this thesis, can be used to further strengthen identification and enhance security.

This thesis, aims to approach mobile keystroke dynamics in a novel way, focusing on identification via name, age, gender and handedness inference with the inclusion of accelerometer and gyroscopic measurements to increase accuracy and reduce errors. The aim is that this could eventually be used in practical applications to improve security for users, whilst remaining as transparent as possible.

To do this, key questions were posed in an attempt to guide the research.

1. To what extent can keystroke dynamics be utilised in order to infer a person's name on a mobile device?
2. What effect does the inclusion of accelerometer and gyroscopic data alongside keystroke dynamics have on the ability to successfully infer a person's name and soft biometric features on a mobile device?
3. To what extent does the volume of data per user help to improve the accuracy of the prediction of name and soft biometric features?

A total of three experiments were conducted to answer the research questions above. Initially, a pilot study was completed which aimed to infer a user's name from the way they type on a smartphone by utilising keystroke dynamics.

Following from the pilot study, a main study was completed in order to incorporate accelerometer and gyroscope readings as well as a number of other data points. This experiment was designed to allow identification of the user based on the movement of the device as they type.

Finally, there was the third study, which was similar in structure to the second study and had the same planned outcomes. However, the inclusion of more data sets per user was adopted to provide a better chance of identification.

1.2 Contributions

A number of contributions are presented within this thesis, however some key contributions were identified which can be seen below:

- **Novel approach to identity data using smartphone motion sensors.** From this research we have produced a novel approach on extracting/understanding identity data on smartphones using motion sensors. From this we are able to infer large amounts of data about a user without explicit permissions utilising buttons instead of a keyboard.
- **A new methodology and experimental approach (including data capture framework).** This thesis presents details of this new methodology and experimental approach.

- **Bespoke data sets of motion data that can be anonymised and shared with the wider community.** We produced a large data set with varied users that can be anonymised and used as a set of data for future experiments.
- **A novel algorithm for predicting a letter utilising smartphone motion sensors.** This novel algorithm utilises a threshold model for rotational data to allow for prediction of letter. This is further detailed in subsequent chapters.

1.3 Thesis Organisation

The remainder of this thesis is structured as follows. Chapter 2 investigates the technical background and relevant literature around the research to provide a solid grounding in the topic, as well as identified gaps for contributions. Chapter 3 consists of the detailed breakdown of the application that was created for both the pilot and main studies. Chapter 4 details the investigation of replicating name inference from keystroke dynamics on a mobile device with results and findings around this and a discussion. Chapter 5, consists of a detailed breakdown of the evaluation of the addition of further sensor input to enhance accuracy, including results and findings. Chapter 6 mirrors that of Chapter 5 with the longitudinal study to evaluate the increase of data per user and the effect on accuracy, again with results and findings. Chapter 7 presents a combined analysis of the main datasets from chapters 5 and 6, to provide an insight into results within a larger dataset. Finally, Chapter 8 looks at what conclusions have been drawn from the research and potential points

that warrant further study. References and Appendices then follow.

Chapter 2

Technical Background and Related Work

In this chapter, the relevant background information is discussed in connection to this thesis. The concepts of behavioural biometrics and keystroke dynamics are introduced, along with many of the current key contributions to the field. Crucial elements such as machine learning, ethics and keystroke dynamic inference are also discussed in detail. Providing a solid basis to further explore the research presented in this thesis. Furthermore, problems and limitations are discussed, which form the basis of the research.

2.1 Literature Review Methodology

To ascertain the gap that this thesis is addressing, we first need to review the relevant literature on the topic. To ensure that every effort has been

made to include all relevant literature, a methodology was used to review and search for specific research which is detailed below, alongside the more generalised information that was relevant to the research.

Firstly, multiple searches were conducted and combined to narrow down the results to a more manageable level. Google Scholar was used to conduct these searches, that can be seen in Table 2.1 below.

Table 2.1: Google Scholar search terms, completed on 30.03.2024 and papers accessed on 01.04.2024 for review of abstract and downloaded if considered for further analysis. Further reviews conducted after this initial access.

Search Term(s)	Number of Results
Keystroke Dynamics	39,200
Keystroke Dynamics + Mobile	21,800
Keystroke Dynamics + Mobile + Infer	10,400
Keystroke Dynamics + Mobile + Infer + Accelerometer	6,180
Keystroke Dynamics + Mobile + Infer + Accelerometer + Android	2,730

Looking at Table 2.1 above, the number of relevant pieces of literature was reduced to allow for a more thorough and focused review.

Of the 2,730 results remaining, the first 100 were reviewed for their suitability. This decision was made due to the large amount of literature available, and Google Scholar's relevance ranking when displaying results. These rankings were not necessarily on number of citations, but more on the relevance to the search criteria entered.

The 100 papers were then filtered down further, focusing on those

with the most relevance to the proposed research questions. Following a review of each paper, resulted in a final selection of 14 papers. The full list of papers that were reviewed can be seen in Appendix 1. Papers which were deemed as not suitable, contained research that was not entirely relevant to the proposed research questions. Items such as acoustic signal based inference and papers that focused on swipe patterns or PIN (Personal Identification Number) based applications were not considered relevant and therefore, were not reviewed further.

2.2 Background Information

The use of mobile devices is prevalent in modern society [57], as shown by the dramatic increase in mobile phone sales over the last 13 years. Figure 2.1, shows this increase has occurred in a relatively short time period, from 2007 when there were roughly 122 million units sold, compared to 2022 where there were nearly 1.4 billion. As the use of mobile devices increases, so do the advances in technology, as shown by the increase in mobile performance. For example, from the initial iPhone, released in 2007 which had a 412MHz processor to the latest iPhone 15 Pro which has a 3.78GHz processor [56], we have seen an over 800% increase in frequency.

Such technological advances have made the inclusion of sophisticated pieces of hardware such as accelerometers and gyroscopes, viable and common, inclusions to mobile devices. The addition of these devices opens up new possibilities for the identification¹ of users.

¹Identification, with regards to technology, can be defined as “...the ability to identify uniquely a user of a system or an application that is running in the system.” [33].

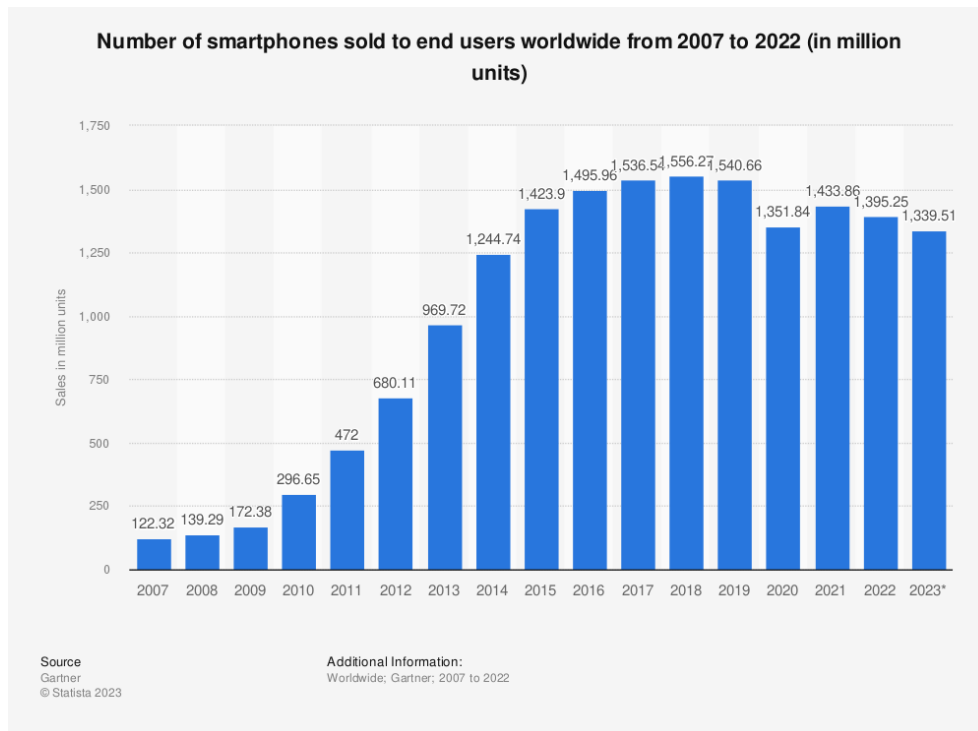


Figure 2.1: Smartphone Sales Globally 2007 - 2022/23. Taken from <https://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007/>

This research explores a combination of keystroke dynamics, focusing on dwell time², as well as data from the accelerometer and gyroscope, location, and size of press on the screen in order to infer the typing data of a user. With this data, many different soft biometric identifiers can also be implied, such as handedness, age and gender [12] as well as inferring name. All of which can be used for identification purposes.

To fully understand the research discussed throughout this thesis, the background information and review of literature serve as an overview of key topics which are essential for the understanding needed.

²Dwell time can be summarised as the length of time a key on a keyboard is pressed for.

2.2.1 Accelerometers

An accelerometer is a device that “measures proper acceleration” [60]. Proper acceleration is defined as the physical acceleration of an object in a linear motion. This is measured by sensors in the accelerometer. The standard configuration for an accelerometer, within a mobile device, as can be seen in Figure 2.2 is a 3-axis set up [40]. This measures the proper acceleration of the device along three axis, (x) , (y) and (z) .

The (x) , (y) and (z) values are referred to as the Earth coordinate frame and is aligned based on gravity and standard magnetic orientation. Within this coordinate system, (x) represents the east-west direction (where east is positive). In addition to this, (y) represents the north-south direction (where north is positive), and (z) represents the up-down direction, perpendicular to the ground (where up is positive) [24].

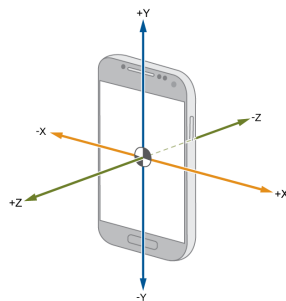


Figure 2.2: Visual representation of a 3-Axis accelerometer. [40]

Accelerometers are used to measure the directional movement of the device. The gyroscope however, adds an additional dimension to the information supplied by the accelerometer, by tracking rotation or twist, along the (x) , (y) and (z) axis.

2.2.2 Gyroscopes

In addition to the inclusion of accelerometers within mobile devices, these are paired with gyroscopic devices in order to get a more accurate and responsive output.

With regards to a definition, the Oxford English Dictionary states that a gyroscope is “A device consisting of a wheel or disc mounted so that it can spin rapidly about an axis which is itself free to alter in direction. The orientation of the axis is not effected by tilting of the mounting, so gyroscopes can be used to provide stability or maintain a reference direction in navigation systems, automatic pilots, and stabilisers” [55]. Whilst this a good definition for a traditional gyroscope, the device we see included in modern mobile devices is very different to its early counterparts.

A modern gyroscope follows the same premise as traditional mechanical ones, although the ones in a smartphone are MEMS (Micro-Electro-Mechanical Systems) gyroscopes. These are smaller versions of the traditional concept which is embedded on electronic boards [68].

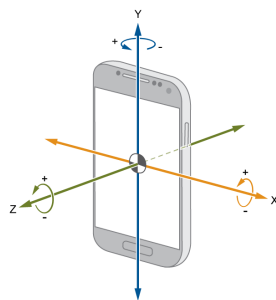


Figure 2.3: Visual representation of a 3-Axis gyroscope. [41]

As can be seen in Figure 2.3 [41], a gyroscopic device measures the movement along the same three axis as the accelerometer, (x) , (y) and

(z). However, the gyroscope measures the rotation or twist of these axis. The rotation around the (z) axis is referred to as the Alpha value and is set to 0° when the device is pointing magnetic north. When the device is then rotated counter-clockwise, the alpha value increases. The rotation around the (x) axis is referred to as the Beta value and this is set to 0° when the device is perpendicular to the earth. As the top of the device tilts towards the earth, the beta value increases. Finally, the gamma value refers to the rotation around the (y) axis. The gamma value is set to 0° when the left and right edges are perpendicular to the earth. This then increases as the right hand side of the device is tilted towards the earth [24].

2.2.3 Combining Accelerometers and Gyroscopes

One of the more prominent issues of only having either the accelerometer or gyroscope is the balance between responsiveness and quality of the output. Whilst modern mobile devices are powerful, providing a solution which is both responsive and with a quality output significantly drains resources. For example, a modern mobile device is expected to be able to auto-rotate the screen in line with user movements or adjust images in real time, such as through the camera application. Both of which require a very responsive piece of equipment with a large amount of processing power. Using only one of the devices therefore, could result in the device auto-rotating with a significant delay, or promptly but rotating too far. One solution therefore, due to the limited amount of processing power available, is to use a combination of the devices which work in sync. The devices share resources and each support the outputs of the other to

provide a clean and responsive reading as required. Using both devices will be essential to this research to ensure as much data as possible is obtained.

2.3 Quality of Information

To provide as natural a typing experience as possible, whilst still focusing on ensuring the quality of data collected is high, text will be displayed on screen, which is then copied by the user into a text field. In order to ensure that the data obtained is fair and unbiased, the text is inclusive of all of the letters within the English language, with an uneven distribution of the letters with the highest frequencies to simulate real world text. The chosen text, which is a collection of uni-grams³, can be seen below and ensures there was no additional bias in the chosen text. Three sentences were also chosen to account for potential spelling mistakes, as the delete key was removed to make the input cleaner. Uni-grams were chosen over other *n-grams*⁴ for the initial data collection to evaluate their effectiveness with the identification of a user.

Jackdaws love my big sphinx of quartz.

The five boxing wizards jump quickly.

Pack my box with five dozen liquor jugs.

Figure 2.4: Chosen Uni-grams to be used in the data collection of the application for all of the experiments proposed.

The uni-grams above were chosen due to the length of text. Asking

³A uni-gram is a sequence of 1 letter in isolation.

⁴In the fields of computational linguistics and probability, an n-gram is a contiguous sequence of n items from a given sample of text or speech

a user to enter a detailed passage of text into a mobile device could be time consuming and discourage user participation or completion of the experiment. Secondly they cover a good representation of the spread of letters in the English language. Whilst they cover every letter, they also cover a higher frequency of the more common letters. A figure showing the highest frequency letters in the English language can be seen below in Figure 2.5.

e t a o i n s r h l d c u m f p g w y b v k x j q z

Figure 2.5: Most Common Letters in the English Language in Descending Order. Taken from <http://letterfrequency.org/>

Compared to the letter frequencies of the English language above, the following Table 2.2 shows the frequencies of those within the chosen uni-grams for the research, with the highest ranking letters highlighted.

Table 2.2: Chosen sentences for the data collection as per Figure 2.4, shown in alphabetical order with letter frequency. Characters were not case sensitive in the application.

Letter	Frequency In Text	Letter	Frequency in Text
A	5	N	3
B	3	O	6
C	3	P	3
D	3	Q	3
E	5	R	3
F	3	S	4
G	3	T	3
H	3	U	5
I	9	V	3
J	3	W	3
K	3	X	3
L	3	Y	3
M	3	Z	3

As can be seen from the table above, the most common letters, in descending order of frequency for the passage of text chosen were: *I*, *O*, *A*, *E*, *U*, *S*. 83% of the top letters from the uni-grams which were utilised, correspond to 71% of the top letters in the English language, with the exception of *U*. We can therefore state with confidence that this is an accurate depiction of real world text.

2.4 Identification and Authentication

Identification of a user is a key challenge, and one which is the primary focus of this research. By using sensors such as the accelerometer and gyroscope, combined with behavioural biometrics such as keystroke dynamics, we can hope to identify users with a good level of accuracy on a mobile device. Authentication goes hand in hand with identification, and therefore must be understood, despite not being the key focus of this research.

2.4.1 User Identification

User Identification is regarded as the ability to identify a user that is using or accessing a system [33]. This is a key requirement for all systems, as it ensures that the actions which are carried out are those of a specific user. Unfortunately, due to the ever increasing number of systems in use, all with varying levels of security and with a consistent rise yearly in cybercrime [27], it is becoming increasingly likely that a user could not be who the system believes they are. Ensuring robust identification is paramount, as the potential harm that can be done by an incorrect identification is severe and wide reaching.

User Identification can be broadly categorised under two distinct approaches, reactive and proactive, according to the time identification occurs [50]. Reactive identification occurs after the interaction with the system, and is predominantly used in cybercrime forensics. Proactive identification, which is the more common of the two, occurs at the time of the user interacting with the system. For the purposes of this re-

search, reactive identification will be utilised, due to the nature of the experimentation and analysis.

Many methods exist for proactive identification, such as utilisation of cookies, IP address, User ID and so on [50]. Unfortunately, whilst these are convenient and often provide a method of transparent identification⁵ [50] they are easily spoofed or bypassed resulting in a potentially malicious user accessing the system.

Clearly the identification of a user is therefore not enough, and must be complemented with additional information in order to successfully allow access. Authentication, which goes hand in hand with identification is therefore, also key in protecting systems, as it allows for a greater chance of a true positive⁶ than just identification alone.

2.4.2 Authentication

Authentication, which is “The process or action of **verifying** the identity of a user or process.” [55], goes hand in hand with identification and forms the basis for the security of a system. Where identification is concerned around the identity of the user or process, authentication focuses on ensuring that the user or process which has been identified is actually that user or process.

Authentication can be broken down into three broad subcategories, as specified by Teh et al. [59], which are; Knowledge, Token and Biometrics.

Knowledge refers to the most common of all authentication techniques

⁵Transparent Identification occurs when a user is identified without being prompted for a source of logon information.

⁶A True Positive is an outcome where the model correctly predicts the positive class.[25]

which is the password or PIN. Whilst this is the easiest of authentication techniques, it has the disadvantages of being forgotten, as well as spoofed, relatively easily [1]. Yan et al. [66] theorise that due to human limitations with regards to the complexity of a password, and the recollection of this, that these limitations likely will compromise a password's security as the user will probably keep a paper version of it, which can be lost or stolen. This is due to Human memory being limited for sequences of items to around 7 (+/-2).

As knowledge is a fairly weak technique for authentication, multiple factors can sometimes be employed to increase the likelihood of a true positive authentication. The amount of factors utilised can vary depending on the system in question. Single-factor authentication often consists of a knowledge item such as a password or PIN. Unfortunately, a single factor in isolation is no longer widely accepted, with many systems turning to MFA (Multi-Factor Authentication)⁷. MFA uses a combination of methods as can be seen in Figure 2.6 below.

⁷Multi-factor authentication uses a combination of Knowledge, Token and Biometrics

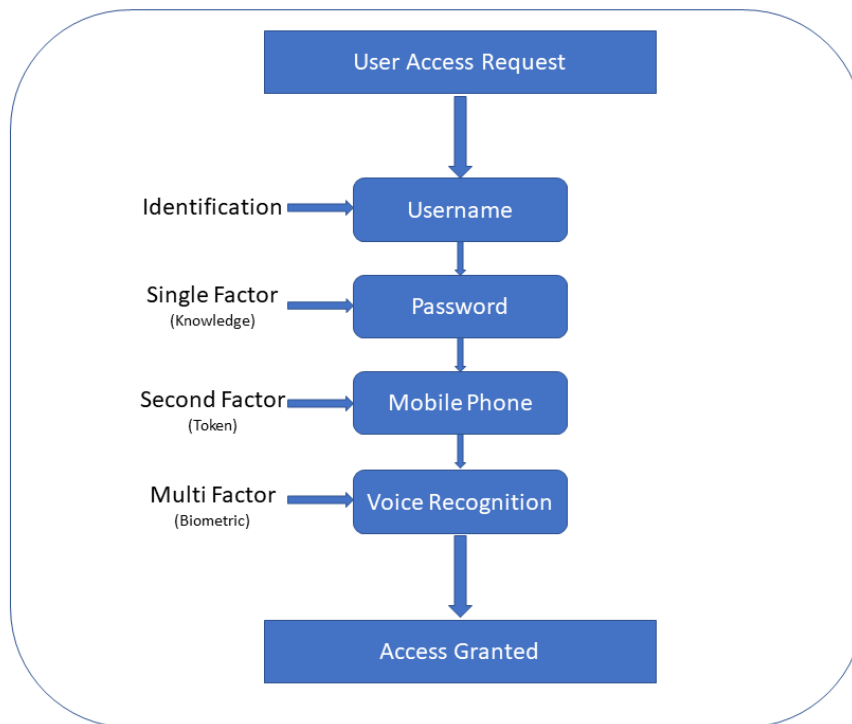


Figure 2.6: Multi-Factor Authentication (MFA) diagram showing knowledge, token and biometric principals alongside common applications.

The above figure shows how a combination of authentication factors can be employed to authenticate a request. This is just one example of many various combinations.

Tokens, such as Smart cards or Minidevices, provide a different approach to authentication, and whilst they are cheap and have a simple deployment, they are easily lost, stolen or spoofed. This is highlighted by an analysis of a paper originally presented by Hwang et al. [31], in which, they suggested a user authentication system using smart cards. Analysis of this system, presented by Chan et al. [13] found that legitimate users can easily create valid credentials without knowing the secret key of a system. As a result, a legitimate user can impersonate other users. This

shows that there are instances where smart cards can be spoofed and even bypassed.

Finally, Biometrics such as a Fingerprint, Voice or Keystroke can be used. These automatically deter sharing and are unique and also unforgeable. Unfortunately, these are historically costly to implement as an additional security layer, and therefore not appropriate for all contexts. Some advantages of biometric keys instead of passwords are listed below which are presented by Li et al [38]:

- Biometric keys can not be lost or forgotten;
- Biometric keys are difficult to copy or share;
- Biometric keys are hard to forge or distribute;
- Biometric keys can not be guessed easily.

L. O’Gorman [43] presents key discussions around authentication, including limitations and problems. An example of a User and Machine authentication process can be seen below in Figure 2.7.

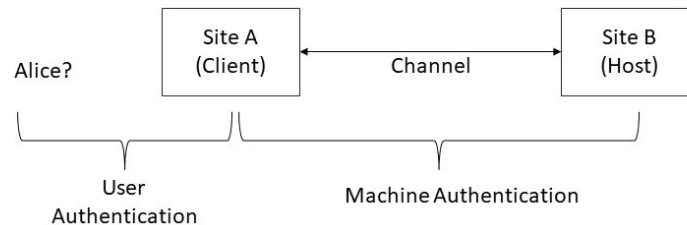


Figure 2.7: A User and machine authentication process. Re-created from *Comparing Passwords, Tokens, and Biometrics for User Authentication* by L. O’Gorman [43].

As can be seen in Figure 2.7 above, there are two key types of authentication, that of a user and that of a machine. User authentication, is concerned with the user logging into the site being who they say they are. The machine focuses on authenticating the request from the client site to the host, which is arguably an easier task. User authentication is particularly difficult due to the balance between providing correct authentication, as well as the ease of use for the User.

To correctly authenticate a user, the process needs to (ideally) combine multiple factors as mentioned previously, which encompass Knowledge, Token and Biometrics as defined by Teh et al. [59]. Whilst no method is completely infallible, a combination of methods increases the probability that the user is correctly identified and subsequently authenticated.

2.4.3 Continuous Authentication

Continuous authentication is the process of repeatedly authenticating a user. As mentioned previously, traditional systems authenticate the user only once, usually at log-in. Whilst this log-in can have multi-factor authentication and protection attached to it, once access is unlocked, there is no way of discerning the user's identity. This is useful because it means that a system can constantly check and safeguard against an imposter using the system and gaining access utilising another user's credentials.

Many methods exist for continuously identifying and authenticating a user; however, most centre around behavioural biometrics. Behavioural biometrics, as discussed in Section 2.5 are innate behaviours which are unique to a user. Bo et al. [7] discuss the use of behavioural biometrics for continuous identification/authentication on mobile devices, and achieve this by the use of their *SilentSense* framework which looks to exploit the user touch behaviour biometrics and the micro-movements of the device caused by user's screen-touch actions [7]. The framework for this method of identification can be seen below in Figure 2.8.

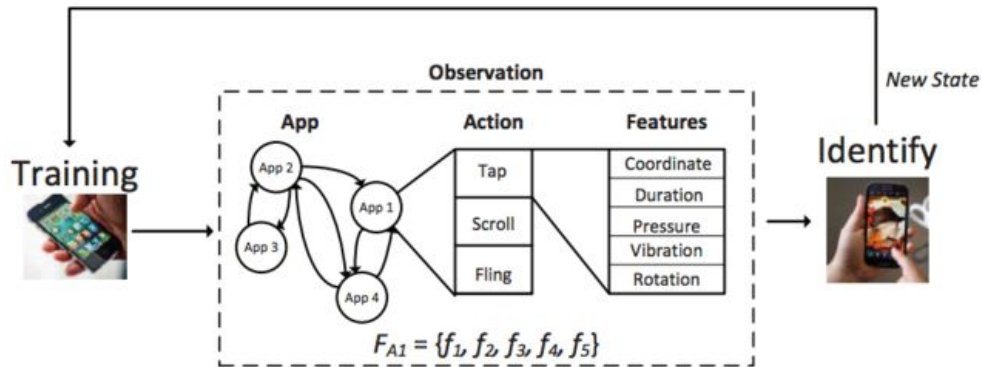


Figure 2.8: The SilentSense Framework. Taken from *Continuous User Identification via Touch and Movement Behavioral Biometrics* by Bo et al [7].

The framework uses the following method to identify and authenticate a user. “While the owner is using the phone, it is feasible to establish a behaviour model through automatically learning. When interacting happens, the system evaluates the probability of being the owner, and updates the evaluation with increasing observations to determine the identity silently and automatically” [7].

Whilst the method mentioned above is theoretically a good approach, this does raise some questions around real world applications. Depending on the tolerances, or just general user habits, this may result in false positive results making the device less secure than it previously would have been. Or render the device inoperable even if the owner is using it.

Looking at identification and authentication as a whole, this research does somewhat combine these, however it more closely aligns with identification and therefore this is the key focus.

2.5 Behavioural Biometrics

Behavioural biometrics are a set of innate behaviours, which are unique to a user. Predominant amongst these are keystroke dynamics, which will be analysed in detail below.

As discussed by Yampolskiy et al. [65], there are many behavioural biometrics which can be used in order to identify users. These can range from physical actions such as the gait/stride of a user or lip movement, to actions which are more technologically based such as GUI⁸ interaction and game strategy. Yampolskiy et al. go on to discuss the most popular of these biometrics, as well as providing a classification and detailed discussion around these.

Whilst there are a wide range of techniques available, these are all subject to a certain group of characteristics in order to be classified as ‘good’ biometrics. These characteristics, which were defined by Jain et al. [34] are listed as the following:

- Universality;
- Uniqueness;
- Permanence;
- Collectability;
- Performance;
- Acceptability;

⁸GUI stands for Graphical User Interface and is a key component of many pieces of software.

- Circumvention.

Universality: *Universality is concerned that behavioural biometrics are dependent on specific abilities that are possessed by different people to a different degree, or not at all. However it does depend that the whole of the group, so in this case for the security of a system, possess the biometrics in some state or another.*

Uniqueness: *Uniqueness explains that, since only a small set of different approaches exist to performing any task, verification is the key application of this instead of identification. For example, there are only a certain amount of ways that you can open a bottle, so whilst you would not be able to identify someone uniquely from the way they opened a bottle, you could verify that the way which was used, is consistent with the usual pattern of the user.*

Permanence: *Behavioural biometrics as a whole exhibit a low degree of permanence as they measure behaviour which changes with time, such as faster ways of accomplishing the same task. Systems of identification and authentication should therefore consider ways to adjust to the changing behaviour as required.*

Collectability: *Collecting behavioural biometrics is usually unobtrusive to the user. The method of data collection should usually be fully automated where possible and in most cases, the user will not even be aware that the data is being collected. This, whilst providing excellent usability for the user, does raise ethical concerns, which are addressed later.*

Performance: *Overall, behavioural biometric verification accuracy is very good for a number of techniques, however the identification accu-*

racy is low, particularly as the number of users increases.

Acceptability: *Furthering the issues raised in the discussion around collectability, whilst behavioural biometrics are easy to collect, resulting in a high acceptability, users may object due to the lack of privacy surrounding these.*

Circumvention: *It is relatively difficult to bypass behavioural biometric systems as it requires the intimate knowledge of users innate behaviours. However, once this has been obtained, the bypassing of the system is exceptionally easy. This is why it is best to combine this with traditional methods of securing a system such as a password, as well as encrypting users biometric data [34].*

These key characteristics raised by Jain et al. [34] will be considered when choosing which behavioural biometrics to adopt for further research. We will now apply these characteristics to many behavioural biometrics techniques which are discussed by Yampolskiy et al. [65] to determine the suitability of each technique for the research. Whilst this table is based on older research, a good range of techniques are presented. Further analysis building upon this original study, such as Alsaadi [3] help to reassure that this table is still very much relevant. They go on to examine a subset of these techniques and provide various advantages and disadvantages.

Table 2.3: Suitability comparison of behavioural biometrics. Adapted and reduced from *R. V. Yampolskiy and V. Govindaraju. Taxonomy of behavioural biometrics. In Behavioural Biometrics for Human Identification: Intelligent Applications, pages 1–43. IGI Global, 2010.* [65]

Classification	Suitability	Required Hardware
Audit Logs	Unsuitable	Computer
Biometric Sketch	Unsuitable	Mouse
Blinking	Unsuitable	Camera
Calling Behaviour	Unsuitable	Phone
Dynamic Facial Features	Unsuitable	Camera
Email Behaviour	Unsuitable	Computer
Gait/Stride	Unsuitable	Camera
Game Strategy	Unsuitable	Computer
GUI Interaction	Unsuitable	Computer
Haptic	Unsuitable	Haptic
Keystroke Dynamics	Suitable	Keyboard
Lip Movement	Unsuitable	Camera
Mouse Dynamics	Unsuitable	Mouse
Network Traffic	Unsuitable	Computer
Programming Style	Unsuitable	Computer
Registry Access	Unsuitable	Computer
Signature/Handwriting	Unsuitable	Stylus
Storage Activity	Unsuitable	Computer
Tapping	Suitable	Sensor
Voice/Speech/Singing	Unsuitable	Microphone

Based on the analysis of different behavioural biometrics above in Table 2.3, the key biometrics applicable to this research are that of keystroke dynamics, and tapping, which are explored in more detail below. Suitability was assessed by considering the research question, but primarily the method of delivery for the experiment as the focus of the research is around mobile devices. Taking this into account, we can see the following Table 2.5 which outlines why the particular biometrics were deemed as suitable or not.

Table 2.4: Explanation of suitability for behavioural biometrics. Based on Table 2.3 taken and adapted from *R. V. Yampolskiy and V. Govindaraju. Taxonomy of behavioural biometrics. In Behavioural Biometrics for Human Identification: Intelligent Applications, pages 1–43. IGI Global, 2010.* [65]

Biometric	Suitability	Reason for Classification
Audit Logs	Unsuitable	Refers to logs in a device which would not effect keystroke dynamics or be a clear supplemental metric to add.
Biometric Sketch	Unsuitable	This looks at how a user draws a picture or sketch and is not directly related to keystroke dynamics and is therefore unsuitable.
Blinking	Unsuitable	Refers to blinking which is not directly correlated to keystroke dynamics and therefore deemed unsuitable.
Calling Behaviour	Unsuitable	Refers to how a user behaves when making phone calls, which is not directly linked to keystroke dynamics and therefore deemed unsuitable.

Dynamic Facial Features	Unsuitable	Similar to blinking, this looks at how a user's facial features change by using a camera to record, and is not directly linked to keystroke dynamics and therefore deemed unsuitable.
Email Behaviour	Unsuitable	Concerned with how a user behaves when emailing, and is not directly linked to keystroke dynamics and therefore deemed unsuitable.
Gait/Stride	Unsuitable	Concerned with the specific walking characteristics of a user, which as this is not directly linked to keystroke dynamics, this is deemed unsuitable.
Game Strategy	Unsuitable	Concerned with how a user behaves when playing a game, and is not directly linked to keystroke dynamics and therefore deemed unsuitable.

GUI Interaction	Unsuitable	Concerned with how a user behaves when interacting with a Graphical User Interface (GUI), and is not directly linked to keystroke dynamics and therefore deemed unsuitable.
Haptic	Unsuitable	Concerned with how a user behaves when touching a device, and is not directly linked to keystroke dynamics and therefore deemed unsuitable.
Keystroke Dynamics	Suitable	This is the main focus of this experimentation and focuses on how a user interacts with a keyboard, which is suitable for this research.
Lip Movement	Unsuitable	Concerned with how a user's lips move, and is not directly linked to keystroke dynamics and therefore deemed unsuitable.
Mouse Dynamics	Unsuitable	Concerned with how a user moves the mouse and interacts with this, and is not directly linked to keystroke dynamics and therefore deemed unsuitable.

Network Traffic	Unsuitable	Concerned with network traffic on a device, and is not directly linked to keystroke dynamics and therefore deemed unsuitable.
Programming Style	Unsuitable	Concerned with how a user behaves when programming, and is not directly linked to keystroke dynamics and therefore deemed unsuitable.
Registry Access	Unsuitable	This looks at the behaviours of how the registry is accessed, which is not directly linked to keystroke dynamics and therefore deemed unsuitable.
Signature/Handwriting	Unsuitable	Concerned with the specific handwriting characteristics of a user, which as this is not directly linked to keystroke dynamics, this is deemed unsuitable.
Storage Activity	Unsuitable	Looks at the activity within storage of the device and how this is utilised, which is not directly linked to keystroke dynamics and is deemed unsuitable.

Tapping	Suitable	Concerned with how a user taps on a sensor, and is technically linked to keystroke dynamics and therefore deemed suitable.
Voice/Speech/Singing	Unsuitable	Concerned with a user's voice, speech or singing which is not linked directly to keystroke dynamics and therefore deemed unsuitable.

2.5.1 Keystroke Dynamics

Keystroke dynamics are a type of behavioural biometric which describe the unique typing pattern of a user. Early forms can be traced back to the 1800s, where morse code users could be distinguished from one another based on their use of the device [54]. Keystroke dynamics are one of the more reliable methods of identification and authentication as they cannot be forgotten. Behaviours are learned slowly over time by repetition and consistency and these make for desirable characteristics. There are three key metrics to record when capturing keystroke data; the keycode, timestamp of keypress and timestamp of key release [22]. From these three metrics we can produce a vast number of different measurements such as dwell time, which is a key metric for this research.

1. The keycode is the code or label of the actual key which has been pressed in order to identify the specific key and its location on the keyboard.
2. The timestamp of the key press is the system time taken when the key is pressed, usually in milliseconds.
3. The timestamp of the key release corresponds to the release of the key.

By measuring these key metrics, dwell and flight time can also be calculated. Dwell time, which is defined as the length of time a key is pressed for, as well as flight time, which is the time between releasing a key and pressing the next one. Dwell and flight time are two key metrics which help to work out a users typing rhythm as these form a regular cadence

and are the most common metrics in studies which have been reviewed as part of this research. Whilst it is important to understand both of these key measurements, in this research, the choice has been made based on previous relevant literature to only utilise dwell time.

In order to visualise this concept, below is an equation to demonstrate the recording of the keycode, timestamp of keypress and key release, as described by Giot et al. [22] in Equation 2.1. From this, we can extract our dwell time, flight time and keycode.

$$\left\{ \begin{array}{l} (keycode_i, event_i, time_i), \forall i, 0 \leq i < n \\ keycode_i \in Z \\ event_i \in \{PRESS, RELEASE\} \\ time_i \in N \end{array} \right. \quad (2.1)$$

Giot et al. [22] go on to explain the four basic values which are captured for recording keystroke dynamics. Below is the equation used to note duration, which can be used to calculate the dwell and flight time.

$$duration = time\{event = RELEASE\} - time\{event = PRESS\} \quad (2.2)$$

Once this initial equation has been understood, we can then begin to understand the four key measurements that are captured as part of any keystroke dynamics capture. These are shown below in the form of four equations, as well as the associated notation.

PR. (Timing vector containing the duration of each press.)

$$\forall_i, 1 \leq i \leq n, \quad PR_i = duration_i \quad (2.3)$$

PP. (*Latency* - Difference of time between the press of each key.)

$$\forall_i, 1 \leq i < n, \quad PP_i = time_{i+1}\{event_{i+1}=PRESS\} - time_i\{event_i=PRESS\} \quad (2.4)$$

RR. (*Latency* - Difference of time between the release of each key.)

$$\forall_i, 1 \leq i < n, \quad RR_i = time_{i+1}\{event_{i+1}=RELEASE\} - time_i\{event_i=RELEASE\} \quad (2.5)$$

RP. (*Latency* - Difference of time between the release of one key, and the press of the next.)

$$\forall_i, 1 \leq i < n, \quad RP_i = time_{i+1}\{event_{i+1}=PRESS\} - time_i\{event_i=RELEASE\} \quad (2.6)$$

Looking at the above equations for the latencies, we can ascertain that the combination of the results from these equations, can be used to calculate the dwell and flight time of a user, thus identifying their typing rhythm. Upon successful identification of the typing rhythm and keystroke dynamic data we can then infer different behavioural biometrics such as handedness, age and gender.

To illustrate key measurements with regards to keystroke dynamics, the below figure, which was taken from a paper presented by Moskovitch et al. [42], clearly illustrates the different measurements which are represented as equations above.

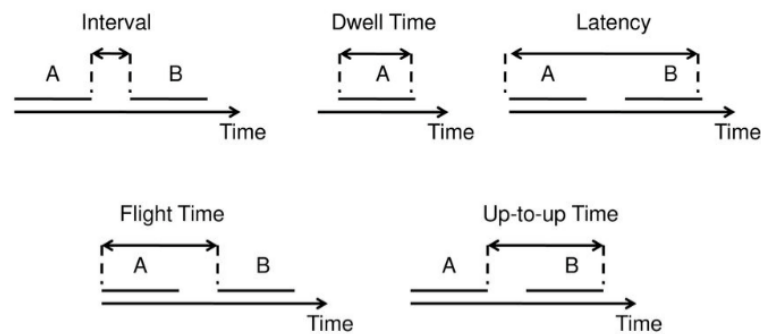


Figure 2.9: Keystroke dynamic illustrations showing common measurements. Taken from *Identity Theft, Computers and Behavioural Biometrics* by Moskovich et al [42].

Looking at keystroke dynamics in practice, Maiorana et al. [39] propose an authentication method that utilises keystroke dynamics at a certain point. By having a pre-recorded ‘rhythm’ of the user typing their password, when a different user attempts to type their password, these measurements are then recorded and compared. If the difference is found to be too great, the user is blocked from accessing the account.

Ponnusamy et al. [47] use a common method that has been seen throughout the relevant literature of setting up a ‘fingerprint’ of a user by getting them to type a specific phrase, in this case, 10 times. During this they collected key press and release data. From this, they extracted the hold-down time and pressure during typing. This was then stored and compared to a user trying to login to the system. A classifier then determines if the user matches or not with an up to 80% success rate. Interestingly, despite the lack of accelerometer and gyroscope data they managed to achieve a good accuracy.

Another similar experiment utilising keystroke dynamics on a mobile device was conducted by Trojahn et al. [61]. They theorise around

utilising keystroke dynamics as a second factor without the need for additional hardware as is seen typically with multi factor authentication. In addition, multiple other factors can be measured when utilising a touch screen device such as pressure during typing, finger tip size and how the device was held/which hand was preferred. Furthermore, features such as holding times and straightness can be used in relation to swipe keyboards. Swipe keyboards require one single press of the screen with the user “swiping” to the different keys required. Within this context, holding time is the time a user spends in one specific letter whilst searching for the next letter to select. Straightness relates to the angle of the line between letters and how direct the path is. All of these various dynamics can be measured and utilised in order to further identify and authenticate users.

2.5.2 Soft Biometrics

Soft biometrics, refers to physical or behavioural traits that can be described by humans and are focused around identity. They are not unique to one specific person and usually, in relevant literature to this thesis, focus around age, handedness and gender [12].

As a large part of this research focuses around soft biometrics, and their subsequent inference. Below is an investigation into key literature surrounding this topic, alongside the prevalence to this research.

Shen et al. [53] look to infer handedness on a physical keyboard from utilising keystroke dynamics. They centre around three key measurements in their research from the key events. Firstly, the hold time, which is the duration between the key down and key up events of a specific key.

The hold time is also referred to as dwell time and is the key metric that will be used in this research for analysis. Secondly, they look at latency times which is the duration between the key down events of two letters. Finally, they look at the interval times which is the duration between key up and key down events. Interestingly, they were able to successfully model user's handedness with an accuracy of 87.75%. This work demonstrates the potential accuracy with a reasonably sparse dataset. The work in this thesis will look to build on this idea when considering virtual keyboards.

Giot and Rosenberger [21] focus on gender inference utilising keystroke dynamics, again on a physical keyboard. They present a minimum accuracy rate of 87.31%. Interestingly, the accuracy scores for the dwell time are slightly higher than this at 87.80%. Whilst not the highest metric in the results, which was that of the fusion of the four keystroke dynamics measurements explained in Section 2.5.1 below, it does provide a very high accuracy score. Which, alongside that of the study from Shen et al. [53] shows promise.

Cascone et al. [12] utilise three different datasets to calculate various soft biometric traits via machine learning classifiers. They calculated gender, age and user-touch experience. There were different accuracies per dataset with 92% being the highest across the datasets for gender, 84% for age and 81% for user-touch experience. The classifiers that were used over these datasets were Adaboost, Decision tree, Random forest, SVM (Support Vector Machine), k-NN (k-Nearest Neighbours) and GaussianNB (Gaussian Naive Bayes).

Pentel [46] presents a lowest accuracy score of 82%, and highest of

92% when predicting age by utilising keystroke dynamics, again on a physical keyboard. Hold or dwell time was again used amongst other measurements to help provide these accuracy scores. Once again, as can be seen by the accuracies of all of the three previous papers, dwell time has been used to great effect when assisting the accuracy scores and therefore will be used in this research to attempt to infer these specific soft biometrics.

From the relevant literature reviewed above, we will focus on the inference of age, gender and handedness for soft biometrics. There has been a high level of accuracy with the use of dwell time to predict these features and as such, this is the measurement that we will utilise in this research. The general motivation behind the inference of these metrics is to narrow down who a user of a device might be. For example, if we can know that a user of a device is of a certain age, gender and handedness; combined with name, then this is a far more powerful combination.

2.6 Keystroke Dynamic Inference Utilising Accelerometer and Gyroscope

Looking to the specific focus of this thesis, the inference of keystroke dynamics will rely heavily on receiving accurate and responsive data from the accelerometer and gyroscope. In addition, the size of the touch on the screen, as well as the location, will also be recorded in order to help distinguish between users as well as to identify different soft biometrics.

Buckley et al. [9] present a highly accurate method of inferring a

user's name utilising keystroke dynamics on a physical keyboard. They utilise bi-grams to produce accuracy results of above 51.34% for balanced accuracy, up to a peak of 70.83%. Whilst this study does not use a mobile device, it does provide a good use of various classifiers and scientific method that can be utilised during this research. This study also displays the possibilities for name inference using these methods. Whilst 51.34% does not seem much higher than random chance, it is in fact significantly better when we consider we have x number of bi-grams in the English language, and they are pulling out a high percentage of those from a name out of the larger pool of x .

Cai et al. [10] present a number of concerns around the security implications of certain aspects of keylogging, particularly those associated with the use of device motion data. The paper highlights the lack of knowledge of most users around the implications that can occur from the unauthorised use of device motion data.

Unfortunately, accelerometer and gyroscopic data is available to any website, under the W3C *DeviceOrientationEvent* [64] which allows web applications to access these through javascript on a website, without the requirement of installing any software.

They then go on to explain around the *TouchLogger* software they have created which “TouchLogger infers the landing locations of the typing finger based on the device orientation and then looks up the corresponding keys based on the current soft keyboard configuration.” [10]. This is achieved by utilising the Alpha, Beta and Gamma angles as described earlier [24], as well as utilising the vibrations recorded by the motion sensors. This research was impressively successful with a rate

of 70% inference, however this was based on a numeric keyboard rather than a full QWERTY keyboard.

This identifies a significant gap in the research field and one which this thesis aims to bridge. The research presented will focus on utilising a full QWERTY keyboard instead of a purely numerical one, due to an improved availability of data points. This provides a greater opportunity for the accurate identification of a user.

Owusu et al. [44] discuss and focus mainly around the lack of security concerning the accelerometer and gyroscopic functions within a smart device. The research concentrates around two key aspects of inference, Area Mode and Character. Area Mode Inference is around the program attempting to infer the section of the screen which was pressed from the movements of the motion sensors, where as Character Inference is concerned with inferring the exact character which was pressed. This research is an interesting premise and the research presented will cover a portion of both of these as we will look to infer the characters and by that the name.

Hwang et al. [32] present the use of Keystroke Dynamics on mobile devices as a method of authentication as an alternative to conventional passcode systems. The shift towards different methods of authentication such as Keystroke Dynamic Authentication (KDA) is due to the inherent risks associated with commonly adopted 4-digit PIN codes. With a maximum number of 10,000 combinations, these are susceptible to brute force attacks and, if breached, could potentially result in data or financial loss. Hwang et al. [32] go on to identify a pattern based approach to KDA which involves three key steps as can be seen below in Figure 2.10.

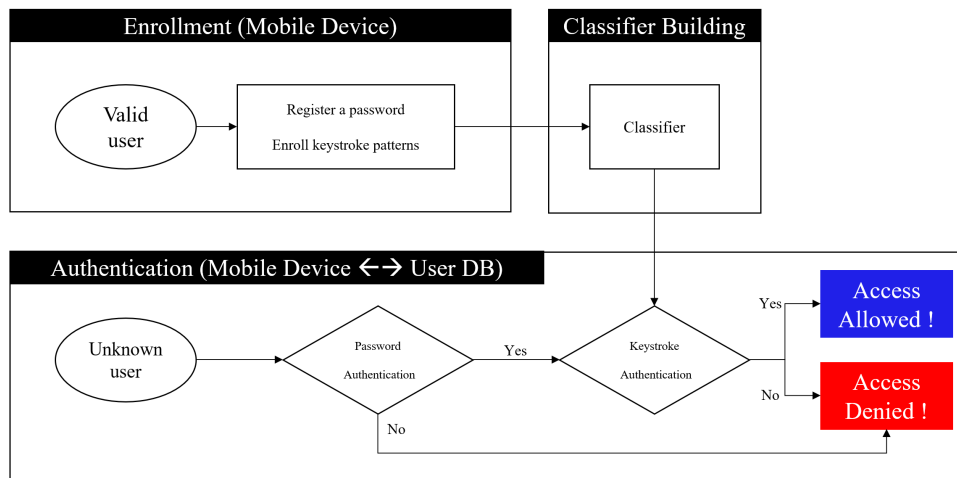


Figure 2.10: Three Steps of KDA framework: enrolment, classifier building, and user authentication. *Recreated from Keystroke dynamics-based authentication for mobile devices by Hwang et al [32].*

As can be seen above, the first step is to enrol a specific keystroke pattern which is recorded based on a user-chosen string. Following this, the timings are evaluated and a classifier is built based on these timings. Finally, the authentication classifier will measure the timing of the password and ensure that it is within the allowed threshold of the previously recorded pattern. If these arguments result in being true, authentication completes and is successful. When compared with other work, such as that of Clarke and Furnell [15] we find a greatly reduced error rate, with a smaller amount of patterns required for training. This method will be explored in the research by the means of a threshold model, which will take in the various data collected and experiment with a plus/minus variable to create a threshold.

Karnan and Krishnaraj [36] present a similar idea to that of Hwang et al. [32] by utilising a pattern-based approach for authenticating a PIN. This research also explores the possibility of incorporating a finger

or palm print in order to provide a further level of security. Accuracy results were good with a success rate of 92.8% in detecting imposters whilst utilising the hybrid approach of fingerprint, palmprint and keystroke dynamics.

Antal et al. [4] explore enhancing existing password authentication mechanisms. In addition to implementing keystroke dynamics on a mobile device, they go further by ascertaining the effect various other inputs will have on the authentication. This includes additional factors such as pressure of touch on the screen and size of touch on the screen. This method of including additional data has been adopted within this study and, as such, the size of the touch on the screen has been recorded. Unfortunately, a large number of Android⁹ devices do not include a scale for the pressure of the touch on the screen: this is usually represented as a binary value of either off or on. Therefore, the pressure of the touch has not been recorded due to the open nature of the data collection and the differing devices it has been collected on.

Hodges and Buckley [29] explore inferring key presses on an android device by utilising the motion sensors. Using a very small training data set of less than 140 characters they were able to identify over 81% of the bi-grams which were in the test set and nearly 30% of bi-grams that were not. Furthermore, factoring a similar element in to future experiments will provide a further test of whether or not we can identify any semblance of personal information based on motion sensor data. This presents another potential gap in the field which this research will address. We will be utilising a smaller dataset than that presented by Hodges and Buckley

⁹Android is a mobile device operating system that was created by Google.

[29] in that we only utilise a maximum of 94 characters. The hope is that this reduced length will make the method more feasible in a practical application in the future by keeping the load small and therefore creating as little disruption to the usual processing of the device as possible. We will also be utilising singular letters instead of bi-grams to attempt to remove the dependency on character pairs. Additionally, we will not only augment the keystroke dynamic data with the gyroscope and accelerometer, but also additional measurements such as size of press on the screen.

Crawford and Ahmadzadeh [16] present research that takes into account a user's position when completing the authentication, as this can have a significant impact on the accuracy scores. They go on to present evidence of the change in user typing patterns depending on whether or not they are standing, sitting or walking whilst typing. In this experiment we do not look to identify the position of a user. Again, this may be a key consideration in future research as this could have a large impact on the accuracy, which may need to be accounted for.

Giuffrida et al [23] present a method of enhancing keystroke dynamics utilising the accelerometer and gyroscope (Sensor-enhanced Keystroke Dynamics (SKD)) on two pre-determined passwords. They then present a tool called UNAGI, which is a fixed text authentication system based on SKD. This takes the sensor and keystroke dynamics readings and applies them to a feature extraction module, which in turn then passes this to training and detection modules. From their findings they discovered they could drastically enhance the accuracy of keystroke dynamics by including the sensor data. They report that there is not a large difference in accuracy from using sensor data in isolation, to using a com-

bination of sensor data and keystroke timings. The main experiment in this thesis will follow a similar approach by enhancing the keystroke measurements with sensory inputs to provide a greater accuracy for authenticating users.

Cai and Chen [11] show good accuracy on a numeric keyboard focusing on 4-digit PINs. They reported a circa 90% probability that an inferred key belongs to a set of three keys (including the correct key). Additionally, they reported that the inference scores for a QWERTY keyboard were always lower than a numerical keyboard. This can be explained by the distinct size difference between the keys on a numerical only keyboard versus a QWERTY keyboard. There is therefore a much lesser chance of noticing the subtle differences in movement between letters, compared to the number keys. That being said, they did manage to successfully infer between 30%-33% of keystrokes within 26 letters. With numbers only, the average inference accuracy was as high as 55%.

Yuksel et al. [67] performed their research on iOS (Apple's mobile operating system). This was a prototype application only which meant they had participants use one device in a controlled setting. This is less of a real world application of the software but did record very good results.

The participant typed randomly generated words shown on the screen for 1 minute, whilst collecting accelerometer and gyroscope data during the experiment. Before the data collection, age and gender of the participants were gathered. Looking to results, k-NN gave them the best results with 100% accuracy. They also utilised ANN (Artificial Neural Network), SVM and RFC (Random Forest Classifier), all of which produced results over 98% accuracy. Keystroke data was not collected at

all, only the sensor data. The model was trained with their data and then the user attempted to ‘log in’ using the same password and the movements were compared to the trained model. As has been seen with other relevant literature, this is a very popular method of proving the authentication of a system by registering the user before hand.

Javed et al. [35] used 10 participants for data collection on five phones. The data collection was tightly controlled and involved the participant sitting, holding the phone and typing explicitly with thumbs. Whilst this resulted in all participants typing in the same way, this could be forced behaviour and not how the user normally types, which could give false results. As with Yuksel et al. [67], only sensor data was utilised with no keystroke data captured. They utilised a DMNN (Dense Multi-layer Neural Network) to analyse and produce data. They then output this generated text compared with the actual typed text as can be seen below in Table 2.5.

Table 2.5: Generated text vs. actually typed text. *Taken from Betalogger: Smartphone Sensor-based Side-channel Attack Detection and Text Inference Using Language Modelling and Dense MultiLayer Neural Network by Javed et al.[35].*

No	Model	Text
1	Typed text	“well that was probably one of the hardest thing i have ever had to do in my life”
	Generated text	“well thst was probablz one of the hardest thing i haxe ever had to do in mz life”
2	Original text	“going through our pictures made me realize hoe stuoid i was to ruin everything we had”
	Generated text	“going through our pictures made me realixe hoe stuoid i was to ruin everzthing we had”
3	Original text	“happy birthday to my favorite person in the world when i would be there celebrating when you”
	Generated text	“happz birthdaz to mz faxorite person in the world when i would be there celebrating when zou”
4	Original text	“A Quick Brown Fox jump over the lazy Dog”
	Generated text	“a quick brown foy jump over the laxz dog”

The Table shows that even though some letters were typed incorrectly the overall meaning of the text is displayed. Whilst this may be inaccurate for passwords, actual sentences of text can be easily deciphered. Overall, an accuracy of 91.14% was achieved at the highest, with a lowest of 79.63% utilising different classifiers [35]. Whilst this does show a novel way of inferring text, this method will not be utilised as it does not fit within the research remit of inference of name and soft biometric features. However, this may be utilised in future work to expand upon

the research in this thesis.

Vaishnav et al. [62] propose a system called KDSmart (Keystroke Dynamics Smart) which utilises three phases; a registration phase, login phase and testing phase. The registration phase captures the keystrokes and touch data. They utilised two hundred participants for the experiment and captured typing speed, flight time, dwell time, error rate and finger size during the registration phase. The login phase, then compares the login keystroke data with the registered data. Finally, the final phase compares the touch data with the registered data during use. If at any point the usage data is not within the specified threshold parameters then the application will close. This approach is one we have seen quite commonly in the relevant literature around creating a registered set of data to compare with for entry into a system.

Huh et al. [30] propose two techniques to reduce user frustration with long term usage of keystroke authentication. Firstly, a pair-wise correlation between the accelerometer and gyroscope sensor values. Secondly an on device feature extraction method to compute DTW (Dynamic Time Warping) measurements. Public dataset evaluation showed increasing FRRs¹⁰ over time, which signified increasing user frustration, and with periodic model retraining they were able to maintain an FRR of around 2.5% showing that the reduction in frustration was successful.

Bedogni et al. [6] obtained the area pressed on the screen with an accuracy of >80% in some scenarios by utilising accelerometer and magnetometer and touch events on the screen. They performed word recog-

¹⁰FRR stands for False Rejection Rate and is the probability that a correct user, is incorrectly rejected. This is sometimes referred to as a false negative prediction.

dition by correlating subsequent touch events and building a tree, which provided the probability of a given word to be classified. They then evaluated the Levenshtein distance from the recognised word to the real one to calculate the success. This distance is the number of simple changes that need to be applied to string a to change it to string b .

Gurary et al. [28] show an accuracy of up to 97% to identify a user out of a set of 15. To do this, duration of touch, time since last touch, relative x and y location of the press, size of press, magnitude of acceleration on press, relative x and y location of release, size of touch on release, magnitude of acceleration on release and the maximum, minimum and average acceleration during the touch were all recorded. This was done explicitly with the accelerometer and no gyroscopic data was recorded. There was a total of 15 participants for mobile, with 15 for tablet, with both repeating the same phrase of ‘mary had a little lamb’ with the difference of spaces being included on mobile, but not on tablet. This was typed a minimum of 20 times by each participant, and then machine learning was used, specifically k-NN, decision tree and naive Bayes classifiers.

Gabralla [20] presents an accuracy of between 94% and 98.7% by utilising a proposed dense DNN (Deep Neural Network) classifier. They had participants use one device and type a password of ‘tie5Roanl’ 51 times per subject. There were a total of 56 subjects, so a total of 2856 records in total. Out of these, 71 features were captured per dataset with the main features being: the hold time, up time, down time, pressure, finger area, average duration of hold, average pressure and average finger area. The proposed DNN model had better accuracy than SVM, k-NN

and ANN which were all compared.

Finally, Buckley et al. [8] present research containing a total of 25 participants, each typing 132 characters each using a standard Android keyboard. This was completed twice as training data. Next, a third piece of text that contained dynamically generated bigrams was typed which would be used as a validation piece. Whilst the typing experiment was conducted, motion sensor data of accelerometer, gyroscope and rotation was recorded as well as times of key presses. Soft biometrics were also captured at the beginning of the experiment which included age, number of hands used to type, fingers or thumbs and comfortability rating with a smartphone keyboard. Overall the results were very promising, with a naive accuracy of 46.5% and 64.7% for the bigram accuracy for the training text. And a 9% naive accuracy and 16.7% bigram accuracy for the dynamically generated text. Whilst my research will not capture the comfortability rating, or utilise bigrams it does capture extra data such as the size of press on the screen and also the x/y location of the press on the screen. The actual keystrokes are also intercepted thanks to the button usage for a keyboard, which they were not in this particular research. Finally, this experiment was also constrained to a single device, which this research looks to build upon by utilising the participants own device, and is a gap that has been identified.

2.7 Machine Learning

Machine Learning is the study of algorithms used to perform tasks and processes without instruction. The knowledge to complete the processes

and tasks comes from patterns and knowledge that the algorithm has ‘learned’ [2].

Machine learning, or rather the process used by machine learning algorithms, can be split into a vast array of categories, however, two of the main learning classifications are supervised and unsupervised learning. Supervised learning is to take a set of labelled ‘training data’, which is essentially a set of sample data. This is used to encourage predictions or decisions from the program which can then be used against real data of a similar subset. This process is then repeated periodically in order to improve the effectiveness of the algorithm and to further ‘train’ it. Unsupervised learning focuses on data which only has inputs, and attempts to find commonalities and patterns in the data. There is no response to feedback, instead, the commonalities and patterns are looked to be identified in the new set of data, thus confirming the learning’s.

One of the many advantages of machine learning is the ability to analyse vast quantities of data in order to identify complex patterns, which would be otherwise unrecognisable. Due to this ability, and the need for this research to analyse vast quantities of User typing data, various machine learning techniques will be employed.

The techniques used will be based in the Python programming language, and will utilise the SciKit-Learn python library [45]. There are a number of applicable key concepts within machine learning which will be utilised within this research.

Supervised learning, which is the preferred choice due to the nature of the research, can be classified into the four following steps:

1. Gather the data. Supervised learning needs the data to be in a

collection of pairs e.g. (input/output). Inputs are the data where as the outputs are the labels or a real number, or even a vector.

2. Convert the input into a feature vector. This assigns a single value to each feature position.
3. Convert the output labels from human-readable text. Some algorithms need output labels to be numbers.
4. Apply the dataset to the learning algorithm to get the model

2.7.1 Machine Learning in Relevant Studies

Due to the large amounts of data, predominantly numerical, that are processed when analysing keystroke dynamics on mobile devices, machine learning is of the utmost importance. In order to compare the effectiveness of the proposed research to that of other studies, the accuracy will be recorded so that it can be compared at a later date. In addition we will collect the precision, recall and F_1 scores to provide further insight into the results.

Revett et al. [49] explore the concept of utilising a variety of machine learning algorithms to effectively measure and compare keystroke dynamics experiments. In this particular study they utilise a PNN (Probabilistic Neural Network) as the method used to complete the authentication from the data to either allow or deny access. This supervised learning technique minimises training time as long as the resource is available. This method presents problems that mobile devices do not necessarily have the memory required in order to effectively carry out this algorithm on-

board the device. Looking forward, the inclusion of such authentication algorithms needs to balance both effectiveness and processing power.

Finally, Antal et al. [4] present keystroke dynamics on the android platform and utilise a varied range of machine learning techniques to effectively analyse the data collected. A table taken from the research can be seen below in Table 2.6 showing the different techniques employed via the WEKA toolkit and their levels of accuracy.

Table 2.6: Classification accuracies for two datasets (41 features vs 71 features) *Taken and re-created from Keystroke dynamics on Android platform by Antal et al. [4]*

Classifier	Accuracy using time based features	Accuracy using time and touchscreen based features
	H+DD+UD+AH (41 features)	H+DD+UD+P+FA+AH+AP+AFA (71 features)
Naive Bayes	50.15% (2.86)	78.93% (2.63)
Bayesian Networks	75.95% (2.65)	91.94% (1.73)
C4.5(J48)	54.79% (3.84)	69.02% (3.32)
k-NN (IBk)	41.07% (2.83)	72.98% (2.25)
SVM(LibSVM)	61.71% (3.22)	88.33% (1.87)
Random Forest	82.53% (2.53)	93.04% (1.65)
MLP	53.01% (3.39)	86.26% (2.19)

Analysing the research above and comparing this to previous research undertaken with colleagues [17][18], a combination of machine learning algorithms are to be used in this thesis. This will consist of decision trees, random forest with n values of 10 and 100, SVC (Support Vector Classifier) with the default RBF¹¹ kernel, k-NN with the default value of five for k and GaussianNB in order to achieve the most effective analysis of results.

It is also important to note, that to avoid cherry picking or bias when selecting results, certain measures have been put in place. All results have been included for the range of classifiers and all were run on the same machine to ensure fairness across the board. Additionally, the train/test split function within SciKitLearn [45] was utilised to avoid making conscious choices when splitting the data.

2.7.2 Metrics

Based on the above analysis of relevant literature, certain metrics will be used to determine the success of the research. Overall, the main metric will be accuracy, with precision score, recall score and F_1 score also used to provide extra insight into the results achieved.

Accuracy, which will show the ability of the machine learning to correctly predict the right labels, is the most important metric that we will focus on. Additionally, this is commonly used in most relevant literature around keystroke dynamics, as a measure of success of the prospective methods.

¹¹RBF stands for Radial Basis Function Kernel and is utilised commonly, and is the default kernel for SVC.

Precision score, which can be seen in algebraic form below in Equation 2.7, is the number of true positives, over the number of true and false positives totalled.

$$P = \frac{T_p}{T_p + F_p} \quad (2.7)$$

Recall score, which again can be seen in algebraic form below in Equation 2.8, is the number of true positives, over the number of true positives and false negatives totalled.

$$R = \frac{T_p}{T_p + F_n} \quad (2.8)$$

Finally, F_1 score, is the harmonic mean of both precision and recall, and the algebraic notation for this can be seen below in Equation 2.9.

$$F_1 = \frac{2T_p}{2T_p + F_p + F_n} \quad (2.9)$$

With the three metrics above, we expect to see high results that are in line with a high accuracy score. High precision in the results obtained demonstrates that we are returning accurate results. Subsequently, high recall demonstrates that the results returned are correct (positive). Whilst F_1 score is a blend of these, and it is therefore important to ensure we have high scores across the board to demonstrate the success of the research.

It is also important to note, that for each of these metrics, the weighted average will be used instead of the macro average. This decision was taken due to the nature of the imbalanced dataset, where

some names appear more commonly than others.

2.8 Identification Ethics

As part of our understanding, it is important to consider the ethical applications of the use of keystroke dynamics for identification. Looking at the GDPR (General Data Protection Regulations) [63], we can see that biometric data as a whole, is categorised as ‘special category’ data, where used for identification purposes. As we have previously established, keystroke dynamics is a subset of behavioural biometrics, and therefore falls under this categorisation. Specifically these regulations call for, amongst other items, explicit consent when processing special category data, which goes some way to protect users and their data in this regard.

Where this research is concerned, and is the case with similar applications, this presents two main points to be addressed surrounding the data capture and processing aspect.

Firstly, if we were to capture the data secretly, without explicitly notifying the user. This would obviously be highly unethical and could present a security risk for the user. Having PII (Personally Identifiable Information) processed without their knowledge could even be used to spoof keystrokes. Rebera et al. [48] discuss that, although these innate behaviours are unknown to us and can be used for a secure and un-invasive method of identification. These can be somewhat easily spoofed, questioning the security and robustness of these techniques, and how a user can stop their behavioural biometric data from being captured and used maliciously.

Multi-Modal biometric systems for identification are a possibility to prevent this. These are concerned around combining multiple types of biometrics such as, fingerprint, iris scan etc. with innate behavioural biometrics such as keystroke dynamics in order to reduce the possibility of spoofing and/or misuse [51]. This, although a more secure method, compounds the previous concern around the data stored about an individual and increases this by including other biometrics. Whilst this will not be explored in the planned research of this thesis, it is important to consider for potential future work.

Secondly, as discussed, explicit consent must be gained to process this data, which does raise some problems with the effectiveness of the identification. If a user is aware that they are being monitored and their data recorded, they are more likely to become aware of their typing patterns. This behaviour is seen in the medical industry with something referred to as the ‘Hawthorne Effect’ [52] or ‘observer bias’. To address this concern, we will explicitly call out include the collection of keystrokes and the inference of these within the terms and conditions, which have to be agreed before proceeding.

It is important to consider that the experiment will be available to users globally on the Google Play Store and the ethical considerations surrounding this. As such, any user from any country can access the application and complete the experiment. As per the requirements from Google, a privacy policy detailing data retention and processing has been completed alongside extensive testing and review from Google before allowing the application to be published. As this adheres to Google’s ethical requirements, these potential ethical implications are accounted

for. Furthermore, as the application conforms to the GDPR as discussed above, users that are based within Europe are covered as well.

Unfortunately, looking at research previously conducted with Earl et al. [18] we found that although users are conscious about their privacy and have concerns around behavioural biometrics for identification, they do not necessarily translate this into real world practice. Many users, as found in the research, will still value the convenience of having their personal data stored over their belief in privacy. Furthermore, a large number of participants were unable to accurately recall how long their data would be stored, despite having explicitly stated they had read and agreed to the privacy policy at the start of the experiment where this data was displayed.

We therefore must present the explicit consent request and ensure this is agreed before proceeding with the experiment, however, from the previous research conducted this should not produce any observer bias due to users not necessarily reading and understanding before agreeing to the terms set out.

2.9 Summary

This review has covered many of the topics critical to the research such as machine learning, keystroke dynamics and the inference of these. Furthermore, ethics have been discussed as a key consideration. We have also deduced that dwell time will be a key metric for inference, and that soft biometric features of age, handedness and gender will be used.

For machine learning throughout the thesis, we have chosen the key

classifiers that will be utilised based on numerous studies and relevant literature. This will result in a reliable and solid collection of classifiers with which to ascertain the effectiveness of the proposed methods.

Keystroke dynamics have also been critically reviewed alongside behavioural biometrics to determine those most suitable for the research. Furthermore the individual measurements and metrics that we will utilise have been selected to complement the behavioural and soft biometrics.

Various gaps have also been identified from the review of the literature, which will form the contributions of this thesis. To summarise, these are:

- Research presented by Cai et al. [10] provides an interesting presentation into device based inference on mobile devices. This paper does however focus on using a numerical keypad which has only 10 digits (0-9) and the accelerometer and gyroscope inference of where on the screen was touched. This research will focus on a full QWERTY keyboard and augment the typical keystroke dynamics measurements with motion data from the accelerometer and gyroscope.
- Hodges and Buckley [29] present a promising method of keystroke dynamic inference on a mobile device, which again presents another potential gap in the field which this research will address. We will attempt to use fewer data than that presented by Hodges and Buckley in that we only utilise a maximum of 94 characters, which is 32.86% lower. The hope is that this reduced length will make the method more feasible in a practical application in the future. Ad-

ditionally, we will not only augment the keystroke dynamic data with the gyroscope and accelerometer, but also additional measurements such as size of press on the screen. Finally we will focus on single characters, rather than bigrams to hopefully remove the dependency on two character strings.

- Multiple papers referenced above show a particular method of only utilising one device in a strict environment. Whilst this may produce more consistent results, this is not feasible to real world applications. As such, this research will utilise the participants' own device and in a setting of their choice. By allowing for the experimental application to be downloaded and completed on any device on the Google Play Store this will fulfil this requirement and the experiment can be completed at any time during the data collection window.

These gaps will form the focus of the pilot and main studies as we look to successfully implement the keystroke dynamic inference from our mobile device, with the goal of inferring name, age, handedness and gender with a reasonable degree of accuracy. We also must ensure that the data collection is as invisible as possible so that a participant is unaware they are being recorded, within ethical constraints, as this will enable us to get the highest possible accuracy of reading. This again will be a consideration in future work if the research is applied to practical methods, as this will effect a users typing.

In the next section, we will discuss the application that was created for both the proof of concept study and the main studies. This includes the

design and implementation, as well as the procedure for the participants.

Chapter 3

Data Collection Application

In this chapter, we cover the application that was created for data collection across the three experiments that were conducted as part of this research. We go into detail around how the application functions and also look at the various store performance across the three experiments. The chapter is split into two; firstly, the pilot study application which looked to recreate keystroke dynamics on a mobile device. Secondly, we look at the main studies application which included the accelerometer and gyroscope measurements to enhance the accuracy and collect more data from participants.

3.1 Experimental Design

The data collection was divided into three key experiments, corresponding to the research questions posed at the start of this research. This allowed for an initial pilot study to test the method and also user up-

take. Following this, the second experiment aimed for more comprehensive data collection after the method was defined. Additionally, it enabled the usage of a user's own device in order to provide a more realistic usage scenario. Finally, experiment three allowed for even more data collection, focusing on quantity of data per user, with a smaller subset of users.

3.2 Pilot Study Application

Below, is a detailed explanation as to how the data collection method for each of the experiments was created. Additionally, we look at the procedure the participant has to go through, alongside the Google Play store analysis for the application.

3.2.1 Implementation

Design

The application was created using Android (Java implementation). Android was chosen due to the ease in accessing sensitive user data, compared with that of iOS devices. Furthermore, all of the relevant literature that shaped the experiments also used Android for these reasons.

For the pilot study, the aim was to capture keystroke dynamics measurements (focusing on dwell time as this was most commonly utilised in the relevant literature) to provide a baseline for the other experiments to measure against. Additionally, since creating the keyboard to capture the data would be a significant challenge due to interception of keyboard

events as explained below, this would need to be overcome early in the research to provide a solid data collection platform from which to base the rest of the experiments.

To create the application, the aim was to have as few screens and as quick an experience as possible to allow for the maximum chance of a user completing the experiment. From research, we know that a person's attention span in terms of technology is ever decreasing [58], therefore having as short an experiment as possible, whilst still capturing all the data required, was key. To do this, the experiment was limited to five screens in total. These were:

- Participant Information Statement;
- Application Instructions;
- Demographic Data Capture;
- Experiment Capture;
- Participant Identifier.

Initially, the participant information screen and application instructions appear as screens one and two to give the user an overview of the experiment and to explain how to complete the process. These were designed to be as clear and concise as possible whilst ensuring that everything was explained as per the University's ethical requirements (please see Appendices 2, 3 and 4).

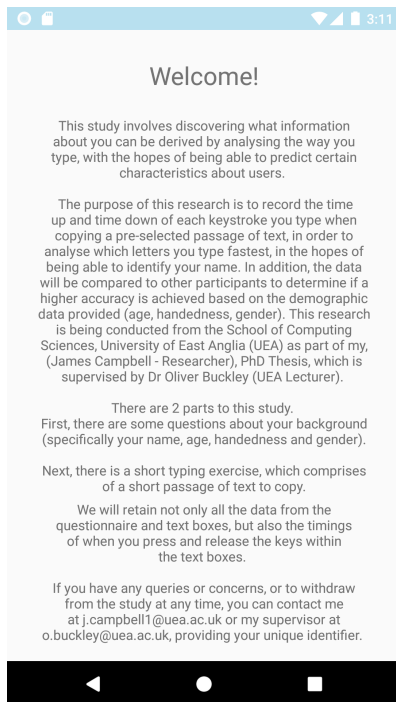


Figure 3.1: Participant information screen from the data collection application.

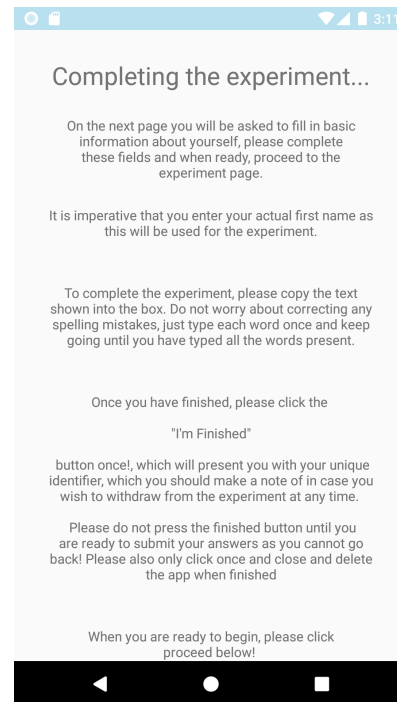


Figure 3.2: Application instructions screen from the data collection application.

Once the participant continued past these screens and agreed to the information provided, demographic data was captured. The main reason for capturing this was to provide us with data to analyse the keystroke dynamics and to provide for the machine learning. As can be seen below, the name was required alongside age, handedness and gender. These were chosen as drop down fields to allow for a uniform response for data analysis but also to keep the data as anonymous as possible.

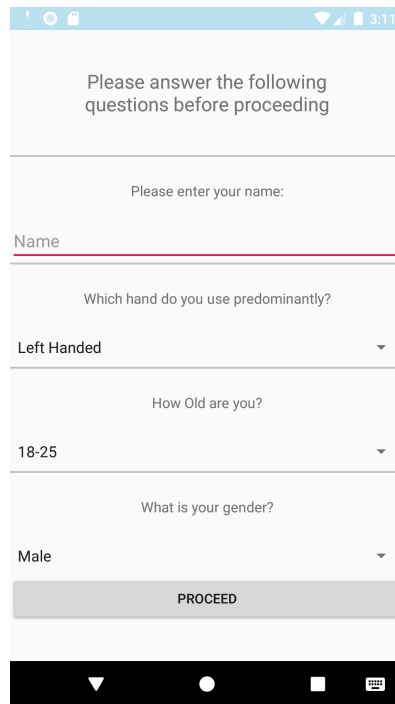


Figure 3.3: Demographic data capture screen from the data collection application.

Next, we moved on to the data capture screen for the experiment. There were a number of considerations when implementing the data capture. Initially, the main consideration was around the keyboard. It was evident that a custom keyboard was needed to prevent things like swipe typing which would invalidate the results of the experiment, therefore a custom solution was developed.

To capture the key press, key release and keycode, buttons were chosen. Intercepting stock Android keyboard events was the initial goal, however these are not able to be accessed due to the differing nature of implementations on keyboards across devices. Interestingly physical keyboard events can be intercepted through an Android device, but not the

on-screen keyboard. Another alternative option which was attempted was to create a custom keyboard in order to handle the key events manually instead of relying on default solutions, however this resulted in the same behaviour of being unable to intercept the events.

The final decision was to build a QWERTY soft keyboard using buttons, which gave the user a familiar interface. A custom build allowed us to capture the system time in milliseconds of the key press and release, as well as being able to capture the key code. The interface is shown in Figure 3.4, and as can be seen, some keys and functionality were omitted to simplify the data collection. All punctuation, numbers, deletion and even capitalisation keys were removed, with only the space bar being allowed and the letters all recorded as capitals. Despite the removal of these various keys, the keyboard was still shaped in the way that a regular mobile QWERTY keyboard would be so as not to interfere with the participant's normal typing layout. The keyboard also scaled to fill the participant's screen as a normal mobile QWERTY keyboard would.

The sentences which would be used were also a key consideration when creating the experiment. Through the relevant literature studied, it was ascertained that the three sentences chosen would be those as listed in Figure 2.4 (see previous chapter).

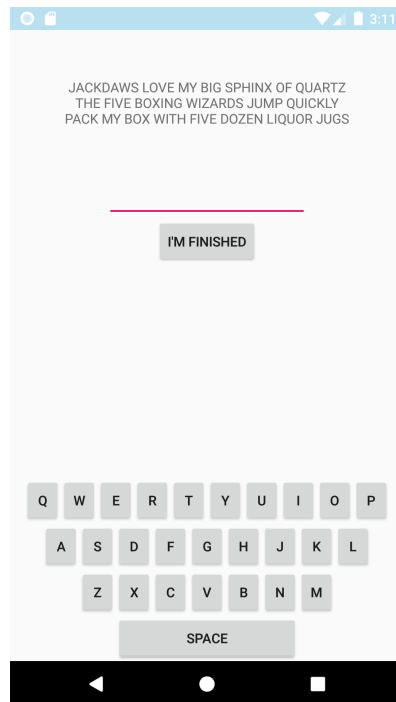


Figure 3.4: Experiment data capture screen from the data collection application.

Finally, the participant identifier screen was presented as can be seen in Figure 3.5 below. This was presented to the participant upon completion and corresponds to the randomly generated text file name which was sent to the data collection server. This was given so that a participant, as explained in the instructions, could withdraw from the experiment at any time as required.

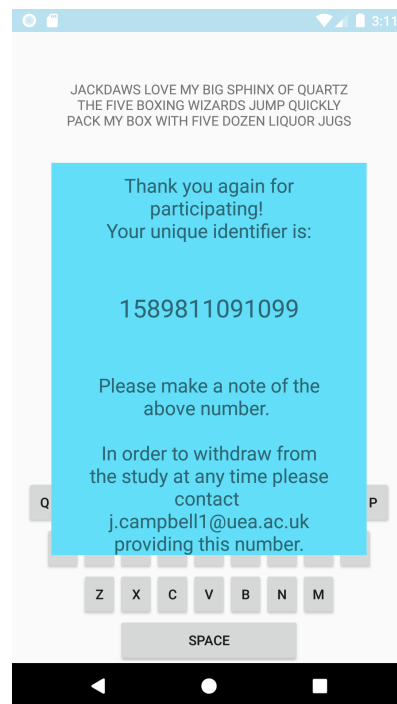


Figure 3.5: Unique identifier screen from the data collection application.

Visual Design and Accessibility

A number of design decisions were made in order to make the application easier to use for those who may have accessibility requirements which are detailed below.

To record the demographic data, as little typing was required as possible, which meant that drop down menus with limited options were chosen to collect the majority of the data.

For text, 'sp' or 'scalable pixels' were used as a font size which allows the font to scale in size, dependent on the participants device and their font settings. This allows for the user's preferences to take precedence and makes the experiment more accessible for those with visual impair-

ment. For example, if the participants system font size is set larger or smaller than the default, the text will scale in the application to allow the participant to complete the research with their preferences enabled.

The font used was ‘Roboto’ which is the default Android font choice. It is a sans-serif font which is designed to specifically be accessible to allow people with dyslexia and other accessibility requirements to read the font more easily.

The actual design of the application with a plain background and dark font was chosen to provide a high contrast to allow for ease of reading. Subsequently, the light blue colour was chosen as the pop up window to provide an even more contrasting colour to again allow for ease of reading.

Programmatic

Following the design of the experiment above, below are key programmatic call outs from creating the application which includes the keyboard and transmitting the text files to the storage solution.

As discussed, a custom keyboard made of buttons had to be produced to successfully capture the required actions. To do this, each button had to be individually created and laid out using an invisible table to ensure the alignment was correct. In addition, the width of each key was set to be equal to that of an average keyboard. A shortened example of this can be seen in the code below; Listing 3.1.

```
1 <TableRow
2     android:layout_width="match_parent"
3     android:layout_height="match_parent"
```

```
4         android:gravity="center">
5
6         <Button
7             android:id="@+id/button_Q"
8             android:layout_width="38sp"
9             android:layout_height="wrap_content"
10            android:text="Q" />
11 </TableRow>
```

Listing 3.1: Keyboard layout code for the data collection application.

Once this layout was successfully created, the buttons were initialised and an `OnClickListener` was assigned to each button. The function of this listener was to link to a switch statement, which can be seen in the code below in Listing 3.2. The switch statement, depending on which button was pressed, would append the corresponding letter to the text field, so that the user could see they were typing in real time as you would normally expect. This was done to not only mimic regular keyboard functionality but also to allow a user to track their progress along the sentences that they were required to type.

```
1 //Initialising each button in the corresponding layout.
2 Button button_A = findViewById(R.id.button_A);
3
4
5 //Switch statement to append to the EditText (Text Field)
   on the layout (shortened).
6 public void onClick (View v){
7     switch(v.getId()){
8         case R.id.button_A:
9             EditText et_content = findViewById(R.id.
```

```

    et_content);
10         et_content.append("A");
11         break;
12     case R.id.button_B:
13         EditText et_contentB = findViewById(R.id.
    et_content);
14         et_contentB.append("B");
15         break;
16     }
17 }

```

Listing 3.2: Button initialisation and listener for the data collection application.

Looking onto capturing the keystroke dynamic data, the key press and release data was stored in an `ArrayList` which was then output to a `.txt` file alongside the demographic information which was recorded. A switch statement was again created to capture the key presses which involved assigning another listener to each individual button. If there was then a motion event of either up or down the code would capture the key which was pressed, direction of motion (up or down) and the current system time in milliseconds. A code snippet for the space bar can be seen below in Listing 3.3, this was then repeated for each key.

```

1 button_SPACE.setOnTouchListener( new View.OnTouchListener
    (){
2     public boolean onTouch(View v, MotionEvent event){
3         switch (event.getAction()){
4             //For a Key Press do the following
5             case MotionEvent.ACTION_DOWN:
6                 {

```

```
7         long millisDown = (System.  
currentMillis());  
8         keypress.add("SPACE - Down");  
9         keypress.add(String.valueOf(millisDown));  
10        }  
11        break;  
12        //For a Key Release do the following  
13        case MotionEvent.ACTION_UP:  
14        {  
15            long millisUp = (System.currentMillis  
());  
16            keypress.add("SPACE - Up");  
17            keypress.add(String.valueOf(millisUp));  
18        }  
19    }  
20    return false;  
21 }  
22 });
```

Listing 3.3: Keystroke capture code for the data collection application.

In order to store the captured data, a text file was created on the participants device. Initially, the application was designed to run on a single device and to output each text file to the device storage. Unfortunately, due to the COVID-19 pandemic, this had to be swiftly edited to utilise remote data capture. A number of different options were investigated, however the method chosen was to link an Amazon S3 (Simple Storage Service) bucket. This bucket allowed for the .txt file from the user to be stored in the cloud securely, so that all of the results could be collated in one place and then downloaded at a later date. This was especially

important as the application was deployed globally on the Google Play Store.

To perform this function, default code was utilised from the Amplify `TransferUtility` library. This resulted in getting the text file which had been written to internal storage and then utilising the `TransferObserver` utility to perform the transfer to S3 [5]. Please see Listing 3.4 below.

```
1
2 //Build the transfer by assigning the file to go and the
   instance location of S3.
3 public void uploadWithTransferUtility(String filename,
   String content) {
4     String fileName = filename + ".txt";
5     TransferUtility transferUtility =
6         TransferUtility.builder()
7             .context(getApplicationContext())
8             .awsConfiguration(AWSMobileClient.
   getInstance().getConfiguration())
9             .s3Client(new AmazonS3Client(
   AWSMobileClient.getInstance()))
10            .build();
11
12     //Try, Catch to write the text file to the storage
13     File file = new File (Environment.
   getExternalStorageDirectory().getAbsolutePath(),
   fileName);
14     try {
15         BufferedWriter writer = new BufferedWriter(new
   FileWriter(file));
16         writer.append(content);
17         writer.close();
```

```
18     }
19     catch(Exception e) {
20         Log.e(TAG, e.getMessage());
21     }
22
23     //Function to actually upload the file to the AWS
24     S3 bucket.
25
26     TransferObserver uploadObserver =
27         transferUtility.upload(
28             "public/"+ fileName,
29             new File(file.getAbsolutePath()));
```

Listing 3.4: Write to storage and upload to S3 bucket code for the data collection application.

Finally, below is a code snippet (please see Listing 3.5 below) for when the participant selects the ‘I’m Finished’ button at the end of the typing experiment. We take the values of the demographic data and a string of the keypress data. We also set the filename to the exact time on their system in milliseconds to allow for a unique identifier. We store the unique identifier and finally we also upload the file using the method shown above. This then also activates the popup window which displays the unique identifier to the participant.

```
1     public void onClick(View v) {
2         String content = (value + value2 + value3
3         + value4 + keypress.toString());
4         String filename = String.valueOf((System.
5         currentTimeMillis()));
6         uploadWithTransferUtility(filename,
7         content);
```



```

5         identifierString = filename;
6         onButtonShowPopupWindowClick(v,
7         identifierString);
        }

```

Listing 3.5: Finish button click application code.

An example data file that the author has edited for display purposes can be seen in Appendix 5.

Figure 3.6 shows the collection of data and how this ties into the application and storage.

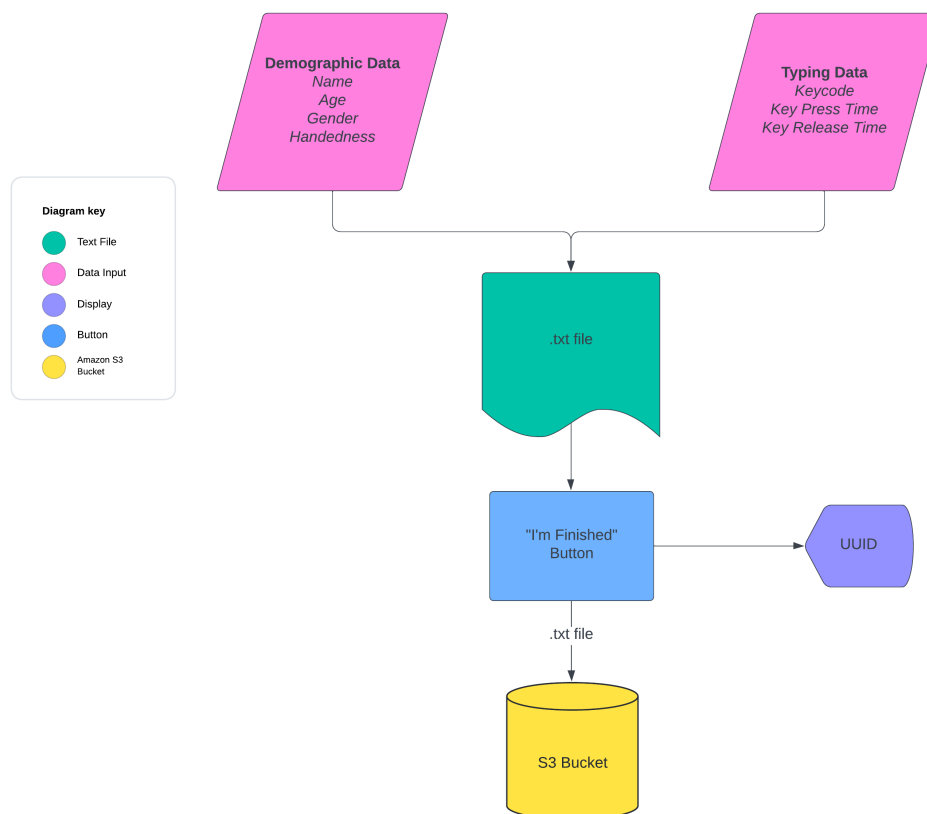


Figure 3.6: Pilot study flowchart showing data collection and storage.

In the pink boxes above we can see the data which is input throughout

the experiment from both the demographic and typing data sections. Next, these are stored in a text file and once the ‘I’m Finished’ button is pressed, a UUID (Universal Unique Identifier) is output to the display and the data is stored in the correct S3 bucket.

3.3 Main Studies Application

Similar to the pilot study discussed above, below is a detailed explanation as to how the experiment was created and designed for the remaining two experiments. These subsequent experiments, two and three respectively, utilised an identical application, therefore this below explanation serves for both applications concurrently.

3.3.1 Implementation

Design

These experiments utilised the same user interface as the pilot study, as described previously. Therefore, all screens and information were kept identical, with the noticeable changes happening on the back end.

When looking at the information recorded, there were a number of extra items of data capture, as well as the measurements from above (Name, Handedness, Gender, Age, Key code and Key press/release). To summarise, these additional measurements were:

- Screen Height;
- Screen Width;
- Size of Screen Press;

- Location of the Press on Screen (x and y);
- Tilt and Rotation of the Device (expressed as x , y and z coordinates).

In order to include the above data, modifications were made to the existing code. These have been detailed in the programmatic section below.

Programmatic

Looking at the list above, to obtain the screen height and width, the following code snippet (please see Listing 3.6 below) was utilised. This code utilised the system metrics functionality to get the pixels for the height and width of the screen being used and stored these as an integer. This was recorded to aid with the analysis around the screen press location.

```
1 final int Scrwidth = Resources.getSystem().
    getDisplayMetrics().widthPixels;
2 final int Scrheight = Resources.getSystem().
    getDisplayMetrics().heightPixels;
```

Listing 3.6: Device screen width and height code for the data collection application.

The size of screen press was calculated by utilising `event.getSize()` which returns the size of a touch, nested within the `OnTouchListener` and stored as a float. This was included to allow the registering of different size of presses to potentially tell if multiple fingers were being used, as well as providing more information to the machine learning for inferring the gender of a user based on hand size (finger press size).

To gather the x and y location of the screen press, again the event class was utilised with the `getRawX` and `getRawY` methods which returned two integers signifying the coordinates.

Finally, accelerometer and gyroscope data, as can be seen below in Listing 3.7, was more complex to obtain. The `sensorManager` was initialised to obtain the accelerometer and gyroscope readings with a normal delay. Two listeners were assigned to the sensors to capture the data. The data was captured on key press so in order to gather the data, the accelerometer and gyroscope data was stored in an array with a value corresponding to X , Y and Z . Then, whenever the user touched the screen, these variables were output to the text file with their current values.

```
1 //Initialise the sensorManager and the Accelerometer and
   Gyroscope and assign listeners.
2 sensorManager = (SensorManager) getSystemService(Context.
   SENSOR_SERVICE);
3
4 accelerometer = sensorManager.getDefaultSensor(Sensor.
   TYPE_ACCELEROMETER);
5 sensorManager.registerListener(MainActivity.this,
   accelerometer, SensorManager.SENSOR_DELAY_NORMAL);
6
7 gyroscope = sensorManager.getDefaultSensor(Sensor.
   TYPE_GYROSCOPE);
8 sensorManager.registerListener(MainActivity.this,
   gyroscope, SensorManager.SENSOR_DELAY_NORMAL);
9
10 //Store the X,Y and Z data in array.
11
12 public void onSensorChanged(SensorEvent sensorEvent) {
```

```
13     if (sensorEvent.sensor.getType() == Sensor.  
TYPE_ACCELEROMETER){  
14         mSensorX = sensorEvent.values[0];  
15         mSensorY = sensorEvent.values[1];  
16         mSensorZ = sensorEvent.values[2];  
17     }  
18  
19     if (sensorEvent.sensor.getType() == Sensor.  
TYPE_GYROSCOPE){  
20         mSensorX2 = sensorEvent.values[0];  
21         mSensorY2 = sensorEvent.values[1];  
22         mSensorZ2 = sensorEvent.values[2];  
23     }  
24 }  
25  
26 //Output the data to text file.  
27  
28 switch (event.getAction()) {  
29     case MotionEvent.ACTION_DOWN: {  
30         touched = true;  
31         screenpress.add(" X: " + mSensorX + " Y: " +  
mSensorY + " Z: " + mSensorZ);  
32         screenpress.add(" X Rotation: " + mSensorX2 + " Y  
Rotation: " + mSensorY2 + " Z Rotation: " + mSensorZ2);  
33     }  
34 }
```

Listing 3.7: Accelerometer and gyroscope code for obtaining sensor readings.

Looking at the code, the switch statement for the `onTouchListener` included a much larger array of data for each press. We can see that

below in the code snippet (Listing 3.8). Firstly, the on touch listener was set for the whole view. This means that whenever there was a touch event on the screen, this was activated. We store the *X* and *Y* coordinates of the screen press as discussed earlier as well as the size of the touch on the screen. Following this, we then add a large amount of data to the text file including the location of the press on the screen, the time for calculating dwell time, size of the press on the screen and finally the rotation and accelerometer data from the sensors.

```
1      v.setOnTouchListener(new View.OnTouchListener() {
2          public boolean onTouch(View v, MotionEvent
3              event) {
4              {
5                  int x = (int) event.getRawX();
6                  int y = (int) event.getRawY();
7                  float size = event.getSize();
8
9                  switch (event.getAction()) {
10                     case MotionEvent.ACTION_DOWN: {
11                         touched = true;
12                         long millisDown = (System.
13                             currentTimeMillis());
14                         screenpress.add("Down X: " + x
15                             );
16                         screenpress.add("Down Y: " + y
17                             );
18                         screenpress.add(String.valueOf
19                             (millisDown));
20                         screenpress.add("Size: " +
21                             size);
```

```
16         screenpress.add(" X: " +
mSensorX + " Y: " + mSensorY + " Z: " + mSensorZ);
17         screenpress.add(" X Rotation:
" + mSensorX2 + " Y Rotation: " + mSensorY2 + " Z
Rotation: " + mSensorZ2);
18
19     }
20     break;
21     case MotionEvent.ACTION_UP: {
22         touched = true;
23         long millisUp = (System.
currentTimeMillis());
24         screenpress.add("Up X: " + x);
25         screenpress.add("Up Y: " + y);
26         screenpress.add(String.valueOf
(millisUp));
27         screenpress.add("Size: " +
size);
28         screenpress.add(" X: " +
mSensorX + " Y: " + mSensorY + " Z: " + mSensorZ);
29         screenpress.add(" X Rotation:
" + mSensorX2 + " Y Rotation: " + mSensorY2 + " Z
Rotation: " + mSensorZ2);
30     }
31
32     }
33 }
34 return false;
```

35

}

Listing 3.8: Motion switch statement to append the data collected to the array for output.

The data was then transferred via text file to an Amazon S3 bucket which could then be accessed at a later date for data analysis. This process is identical and involved writing the text file to local storage and then transferring this once the participant was finished with the experiment.

For the final experiment, Firebase [26] was utilised to send daily push notifications to all participants with the application installed to remind them to complete the experiment. This was beneficial and resulted in a large number of participants producing multiple data sets across different days, as was the purpose of this iteration of the experiment.

An example data file that the author has edited for display purposes can be seen in Appendix 6.

As with the pilot study application, in order to understand the flow of the application more clearly, a diagram has been included below in Figure 3.7. This shows the collection of data and how this ties into the application and storage.

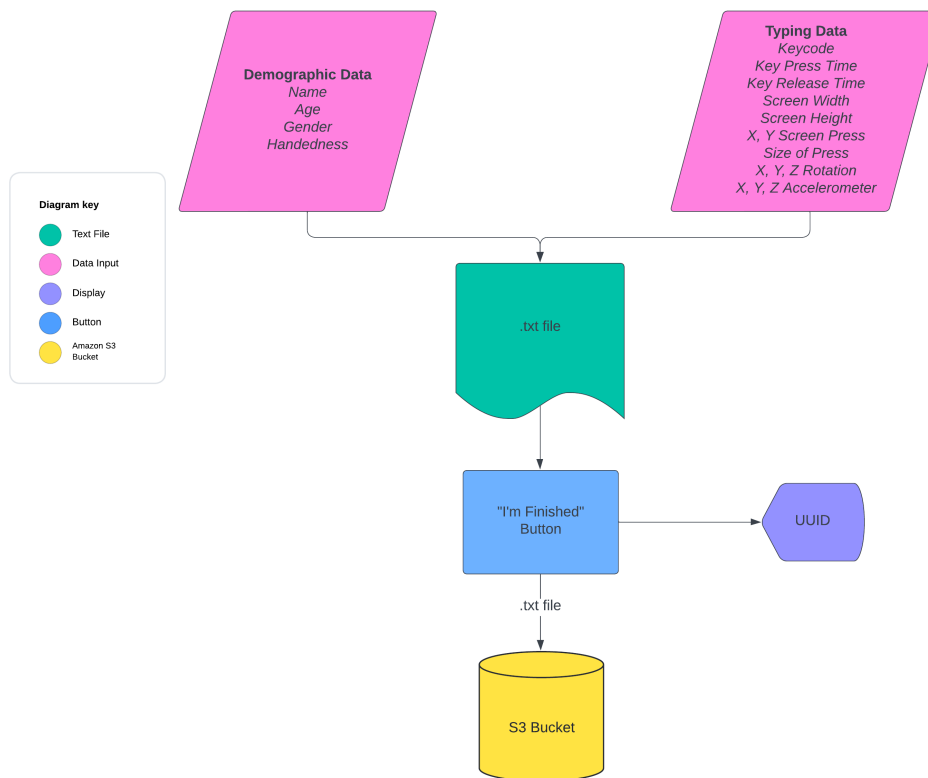


Figure 3.7: Main studies flowchart showing data collection and storage.

Looking to the flow chart above, the data which was input into the application can be seen in the pink boxes at the top. This was then stored into the text file and once the ‘I’m Finished’ button is pressed, a UUID was presented to the participant and the data was stored in the S3 bucket.

3.4 Google Play Store

As the application was deployed to the Google Play Store due to the COVID-19 pandemic, it meant a large range of participants were able to complete the experiment from different countries. To provide more

insight into this, the store page analysis is shown below for the three experiments that were conducted.

3.4.1 Pilot Study

Below in Figure 3.8 we can see the statistics for the application that was used to collect the data. The application was available on the Google Play Store globally for any users to download. For setup, social media was used to attract participants to the study.

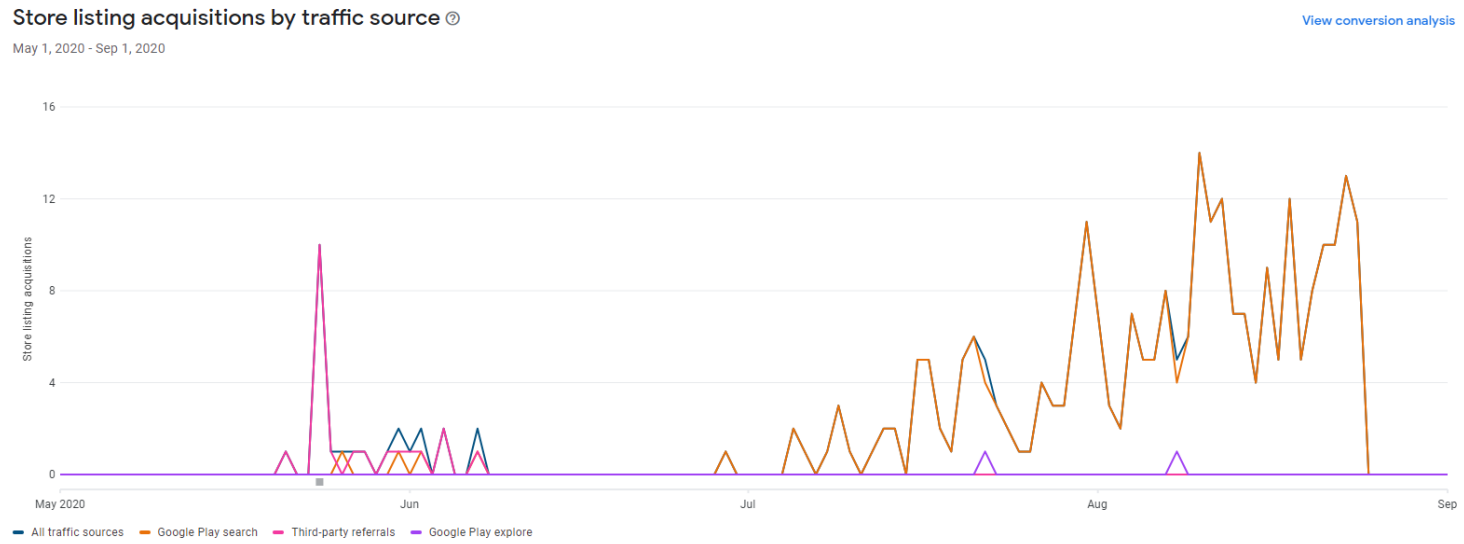


Figure 3.8: Google Play Store analysis graph - Experiment 1.

There were 695 unique visitors to the store’s page for the application, with a total of 289 unique downloads, which is interesting considering the 39 usable sets of data we were able to collect (roughly 390 keystrokes). This points to a large quantity of users maybe downloading the application out of curiosity or not understanding the instructions fully, or potentially just not being bothered enough to participate or finish the experiment.

Unfortunately, we were unable to identify why this occurred within the experiment, as we did not see those results come through to the received participants data.

3.4.2 Main Study 1

Below in Figure 3.9 we can see the statistics for the application that was used to collect the data. As with the pilot study, the application was available on the Google Play Store globally for any users to download. For setup, and to gain participants, the experiment was advertised on social media.

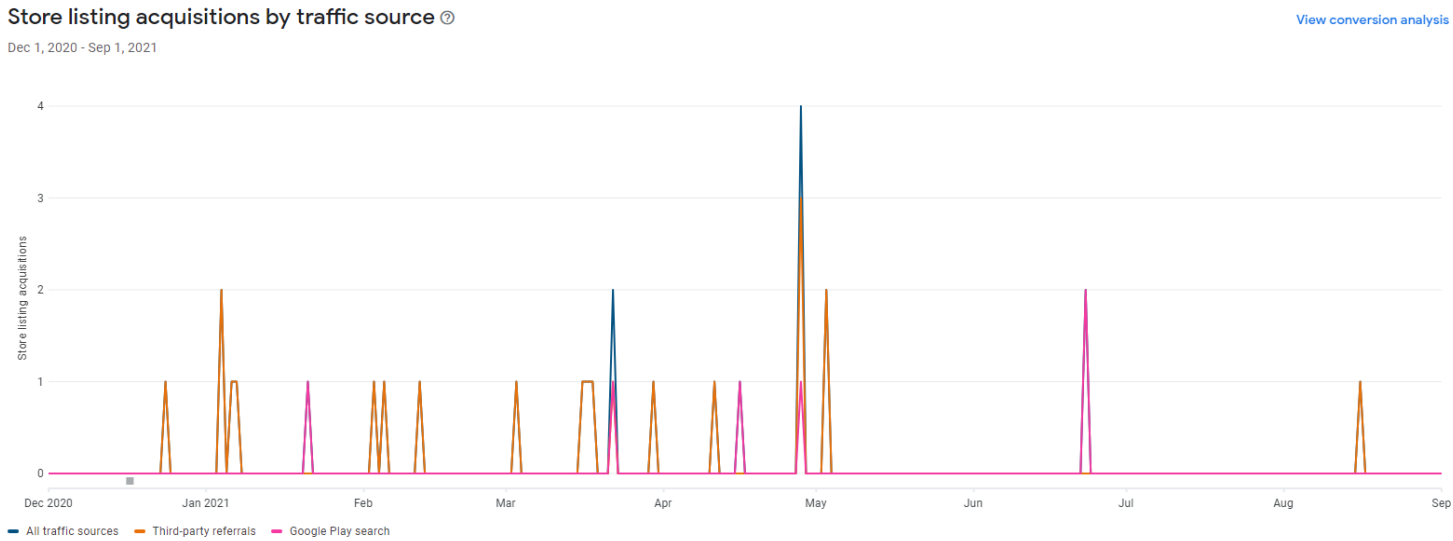


Figure 3.9: Google Play Store analysis graph- Experiment 2.

As can be seen, we had a total of 66 unique visitors to the store’s page for the application, with a total of 27 unique downloads, which is a significantly smaller number than the pilot study. This being said, we did end up with more usable data than previously obtained. This was likely due to the fact that some participants completed the study multiple times and also completed it on the primary research device, due to the lifting and reduction in COVID-19 social distancing restrictions.

3.4.3 Main Study 2

Below in Figure 3.10 we can see the statistics for the application that was used to collect the data. As before, the application was available on the Google Play Store globally for any users to download. For setup, as per the other two applications, the experiment was advertised on social media to gain participants.

Store listing performance

Store listing visitors ⓘ
1,847

Store listing acquisitions ⓘ
672

Store listing conversion rate ⓘ
36.38%
 +36.38% ▲ vs. previous period

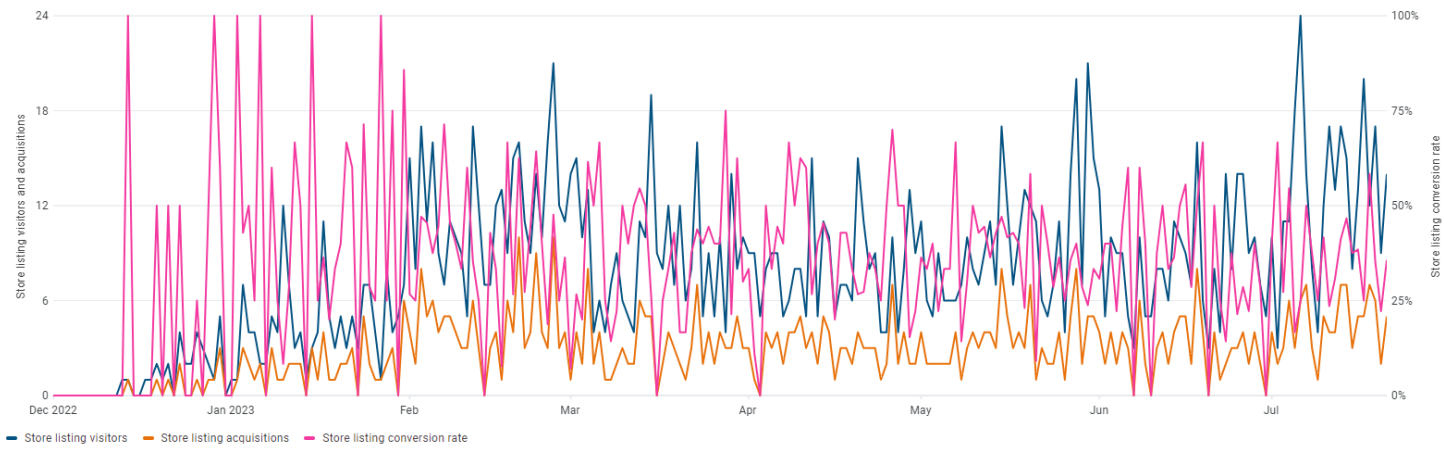


Figure 3.10: Google Play Store analysis graph - Experiment 3.

As can be seen, we had a total of 1,847 unique visitors to the store’s page for the application, with a total of 672 unique downloads, which is far greater than the previous experiments combined. This therefore looked positive to help us achieve the greater spread of data we required to increase the accuracy of our predictive machine learning models.

Unfortunately, due to a large amount of incomplete data, a large number of responses were discarded and could not be utilised in the analysis. The data was discarded due to the length of the results in the file. Restrictions were predetermined as to if a data set was valid or invalid. These were, length of inputs, number of characters and size of file. If they did not meet these requirements for data gathering, then they were removed. We did however end up with a good amount of data at 46 pieces (roughly 460 keystrokes), this was helped by multiple users completing the experiment more than once which is what was the hope of the experiment and which was specified in the recruiting of participants.

3.5 Datasets

To avoid confusion, a table (see Table 3.1) has been created which can be seen below. This table aligns each of the datasets captured with the experiments and analysis to provide a concise summary and a reference point moving forward.

Table 3.1: Datasets and the corresponding experiments/chapters.

Chapter	Experiment	Data Set(s) Used	Participants	Data Captured
4	Pilot Study	Pilot Study Dataset	39	Name, Age Range, Gender, Handedness, Keycode, Time Down Press, Time Up Press
5	Main Experiment 1	Main Experiment 1 Dataset	68	As Above + Accelerometer x , y and z , Gyroscope x , y
6	Main Experiment 2	Main Experiment 2 Dataset	46	and z , Size of Press on Screen, Location of Press
7	Combination Experiments	Main Experiment 1 and 2 Datasets Combined	114	on Screen (x and y),

3.6 Summary

To summarise, we have looked at the data collection platform in depth for each of the experiments, alongside the design choices and the programmatic implementation. As can be seen throughout the chapter, this is a novel data capture framework which can be built upon and utilised for future experimentation. In addition, the ability to store the results of the data collection into the cloud (Amazon S3) mean that this can be utilised by researchers across the globe and reach a wide range of participants due to the compliance with requirements for Google Play Store publication. To see the ethical approval for each application, please see Appendices 2, 3 and 4.

Chapter 4

Pilot Study

In this chapter, the pilot study is discussed, which was formed as a basis to answer the first research question, based on the gaps in the literature that was reviewed. To recall, the first research question of this thesis was ‘to what extent can keystroke dynamics be utilised in order to infer a person’s name on a mobile device?’. We focus on the findings from the pilot study, which whilst we achieved good results, show the need for the inclusion of further data to augment the results in the subsequent experiments. This chapter also goes into detail around the results achieved and their impact on future studies within this thesis. The dataset utilised in this study and chapter can be seen in Table 3.1.

4.1 Introduction

As mentioned above, the analysis and discussion of relevant literature, indicated the need for further experimentation in the field of keystroke dynamics on mobile devices. From the literature, a large number of experiments focused on the ability to infer a PIN or guess what a user had typed, but very few focused on actually inferring information about a user, such as name. To fully explore both the usefulness of keystroke dynamics as a form of identification and the potential implications, a pilot study was employed to bring to understand the various limitations and complications.

This pilot study focused on recreating keystroke dynamics on a mobile device (with the inclusion of name and soft biometric inference), with accelerometer and gyroscopic data to be added in subsequent studies. This would therefore provide a basis off which to build the future studies, but also a point of comparison to improve accuracy.

Firstly, we focus on the participants for the study, looking at the split from the various demographic data available. Following this, we then analyse the results collected and then discuss the outcomes of this experiment as we move forward onto the inclusion of motion data to enhance the accuracy.

4.2 Participants

In total, 39 usable data sets were analysed, although a larger quantity of data was collected. Unfortunately a large proportion of the data collected was unusable due to incomplete data sets which did not provide more

than 10% of the letters required. This data was unfortunately discarded, as in order for this method to work, a full range of data from the user was needed where they had typed all letters in the alphabet at least once. In these cases where the data was discarded, multiple letters were missing, meaning that we couldn't accurately attempt to infer a name or fastest letters, if we did not have all the data. From the data that was able to be utilised, there was a split of 53.85% female, with 46.15% of participants being male. Below are two graphs seen in Figures 4.1 and 4.2 which summarise the key demographic data of the usable data sets.

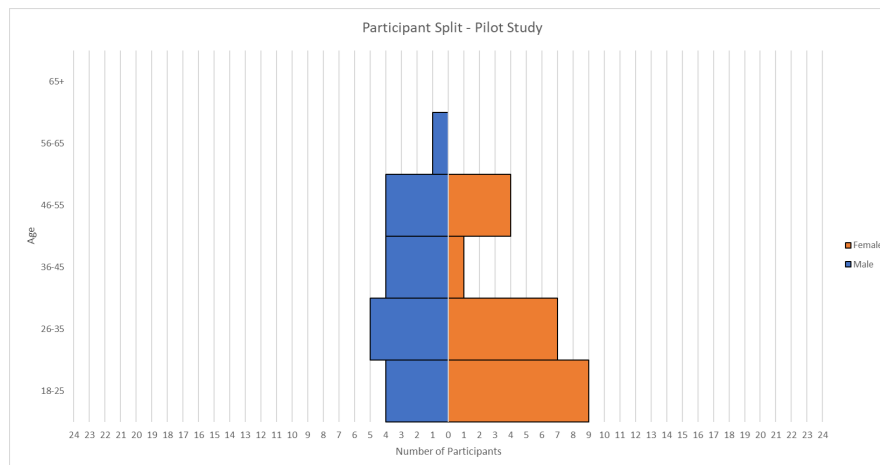


Figure 4.1: Pilot Study - Number of participants for gender, bucketed by age.

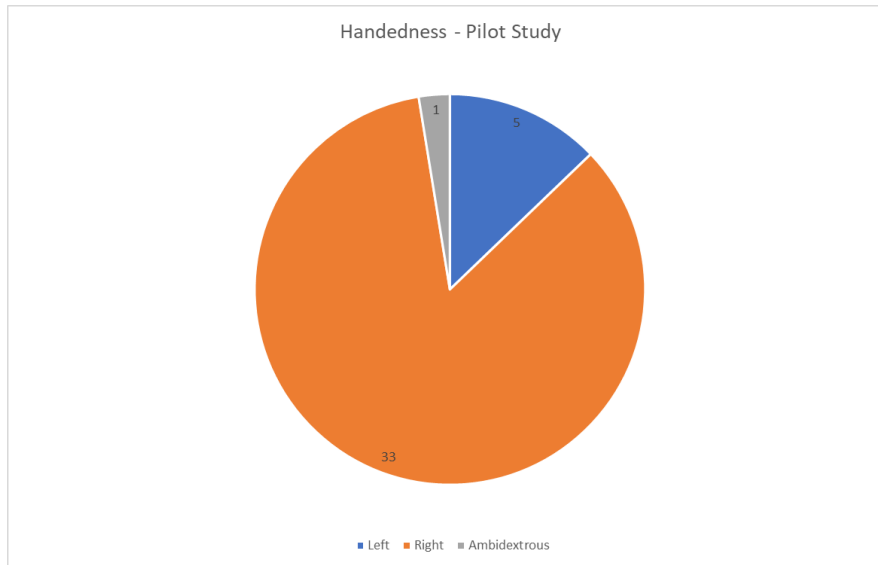


Figure 4.2: Pilot Study - Number of participants split by handedness.

As can be seen from the figures above, there was a fairly even split in terms of male versus female participants in the study, with a slight favour towards female participants. Looking at national statistics, we see that the population of the UK has a female proportion of approximately 51% and a male proportion of 49% [14]. This does not take into account other genders and is purely based on sex assigned at birth. We did have an option for other genders, which were combined into a singular option, however no participants were present that identified with this gender. The age categories were chosen so that we had ten years for each grouping, with the exception of the lower and upper brackets (18-25, 26-35, 36-45, 46-55, 56-65, 65+). This split is commonly seen in related literature and this is why this approach was adopted for the bracketing of the ages in this thesis. Whilst the age split was mostly towards the younger brackets, we do see a split which does cover some

older participants.

When looking at handedness, we also had a fairly accurate split in terms of what we would expect to see, despite the severe class imbalance against left handed and ambidextrous participants.

It is important to note, for this study and the subsequent studies, we did not measure disability or comfort with the device. There was an assumption made that the user would be comfortable with their own device, and as such would be set up to cater to any potential disabilities that may effect their use. This is another area for further study.

4.3 Data Preparation and Analysis

Before analysing the data, cleaning and preparation was needed to ensure a uniform input for the machine learning which was later utilised. To do this, capitalisation and spaces were standardised across the name field for all of the data to ensure that the machine learning was able to interpret the data correctly. Additionally, splitting and removal of miscellaneous characters took place, which were used in the code for splitting the data, such as array parentheses, to parse a clean data-set. Once this formatting had been completed, dwell time was calculated from the data to provide a common data point for analysis. This resulted in the letter and dwell time being output to then be able to be parsed to the machine learning, or to be analysed manually.

To analyse the data for the soft biometric features and the name, a number of classifiers from the SciKitLearn library were utilised [45]. From previous research conducted with colleagues [17], as well as reviewing the

relevant literature mentioned in Section 2, the following classifiers were used:

- Gaussian NB;
- Decision Trees;
- Random Forest ($n = 10$, $n = 100$);
- SVC (Support Vector Classifier);
- k-NN.

To prepare the data for the machine learning, the shortest data-set was chosen as the length with which to trim all the other data-sets, which was 92 characters in length. Once this was completed, we encoded the data utilising OneHotEncoder from the SciKitLearn library [45] and used a test split of 15% as this was the most commonly used in relevant studies. An average was then taken from 50 loops of the machine learning to ascertain which would provide the best accuracy scores for both name and soft biometric features. To do this, a for loop was included that re-ran the machine learning 50 times. Regardless of the scores, these were output into a .csv file and an average was computed using functions, these are then the scores presented in the research. For each of the classifiers, the dwell time and letter were analysed, alongside the particular soft biometric. In this particular experiment, inferring a user's name was the key objective and therefore dwell time made the most logical sense due to it representing the speed of which certain letters were pressed. For the soft biometrics, in addition to accuracy we recorded precision, recall and F_1 score.

The analysis for the pilot study utilises the first data set that was collected, which we can refer to as data set one. This data set, consisted of 39 individual sets of data and did not include any sensor data, just key pressed, dwell time and basic personal information as discussed in the previous chapter.

4.4 Findings

The findings can be broadly split into two main categories, those of soft biometrics, which concern identifying the age, gender and handedness of a user and those of keystroke dynamics, which look to guess letters of a name based on the speed of the keystrokes present as well as machine learning techniques. These will both be discussed in turn below focusing on the detailed results achieved, and the methods undertaken to identify them.

4.4.1 Soft Biometric Findings

As can be seen in Table 4.1 below, we achieved a positive range of results with varying accuracy scores across the range of features.

As a whole, age accuracy scores were not as high due to the categorisation of data. On reviewing previous studies, a common practice was to use older or younger than a particular age in the median of the data in the hopes of improving accuracy scores by completing the analysis as a binary value. For this particular study however, there were a total of six age categories that a participant could belong to meaning a reduction in accuracy occurred. Despite this expected reduction in accuracy, we were

still able to predict age with a highly positive accuracy, far better than random chance.

Looking at gender results as a whole, we also could see some positive results with one in particular being better than random chance. There was also a suitable variance of different genders, with a split as mentioned above of 53.85% female, and 46.15% male.

Finally, as mentioned above, there was a severe class imbalance for handedness due to only 1 participant being ambidextrous and few left handed participants. Whilst this will have effected accuracy, this class imbalance is similar to the handedness split found normally, so this was to be expected. As such, no adjustments were made. Despite this we can see we had a very high accuracy level when predicting handedness.

Table 4.1: Pilot Study - Classifier accuracy, precision, recall and F_1 scores based on an average of 50 runs.

Classifier	Scores											
	Handedness				Age				Gender			
	Accuracy	Precision	Recall	F_1	Accuracy	Precision	Recall	F_1	Accuracy	Precision	Recall	F_1
Gaussian NB (Gaussian Naive Bayes)	84.22%	72.54%	84.22%	77.80%	35.02%	45.64%	35.02%	34.08%	47.34%	58.68%	47.34%	47.48%
Decision Trees	69.68%	73.38%	69.68%	70.16%	23.02%	24.78%	23.02%	20.82%	62.34%	66.98%	62.34%	60.10%
Random Forest (n = 10)	87.64%	78.62%	87.64%	82.54%	29.00%	22.84%	29.00%	22.98%	46.34%	42.26%	46.34%	39.58%
Random Forest (n = 100)	81.22%	67.56%	81.22%	73.62%	28.36%	29.10%	28.36%	24.22%	41.34%	40.20%	41.34%	34.34%
SVC (Support Vector Classifier)	84.62%	73.62%	84.62%	78.42%	21.70%	7.94%	21.70%	11.14%	44.68%	23.28%	44.68%	30.04%
KNeighbours (K Nearest Neighbours)	81.20%	69.92%	81.20%	74.86%	28.02%	27.68%	28.02%	25.14%	42.30%	43.44%	42.30%	39.68%

Looking at the results in more detail, with regards to accuracy scores for handedness, all of the results from the various classifiers used were above 69%, which was significantly better than a random chance. As can be seen in the table, Random Forest ($n = 10$) achieved the best score at 87.64%. Looking to the precision, recall and F_1 scores, these are all high and in line with the accuracy scores achieved, meaning that we produced accurate and positive results throughout the classifiers.

In looking at the accuracy scores for age, we can again see that all of the scores were better than a random chance. As mentioned previously, from reviewing similar research, frequently studies analyse age as a binary form (under or over the age of 30) which would theoretically result in a higher accuracy. For this study, age was split into six categories, with a hope of being more specific in the results, and as we saw in the relevant literature for bracketing ages in similar studies. The data provided a highest accuracy of 35.02% for GaussianNB, which was better than a random chance. Unfortunately, the precision, recall and F_1 scores are not as positive with precision dropping to as low as 7.94% which shows highly inaccurate results for one of the classifiers. This explains the low score achieved for this particular classifier.

Finally, with regards to gender, the highest score was better than a random chance at 62.34% with Decision Trees, however this was the only classifier to do so. Unfortunately, the rest of the results did drop below a random chance with the lowest being 41.34%. Interestingly, there was a 5% difference in accuracy between the two variations of Random Forest, with a value of 10 or 100 for n . Precision, recall and F_1 scores are more consistent and performed better than those for age, however these are still

low. It is hoped that with the inclusion of accelerometer and gyroscope data in the main experiment that these will improve.

4.4.2 Keystroke Dynamic Findings

This section is split into two key parts, firstly, that of machine learning classifiers and finally looking at more manual analysis methods.

Machine Learning

In order to establish a comparison across all of the experiments, all of the classifiers were run against name as they were for all of the experiments to show the progression from the lack of inclusion of the accelerometer, gyroscope and other sensor data, to the results with these extra inputs included. The average of 50 runs was utilised and both a 10% and 15% test split was used to compare results. The tables showing these results can be seen below.

Table 4.2: Pilot Study - Name inference classifier accuracy scores - Average of 50 runs (10% Test Size).

Classifier	Accuracy
GaussianNB	0.00%
Decision Trees	0.00%
Random Forest (n = 10)	0.00%
Random Forest (n = 100)	0.00%
SVC	0.00%
KNeighbours	0.00%

Table 4.3: Pilot Study - Name inference classifier accuracy scores - Average of 50 runs (15% Test Size).

Classifier	Accuracy
GaussianNB	0.00%
Decision Trees	0.00%
Random Forest (n = 10)	0.00%
Random Forest (n = 100)	0.00%
SVC	0.00%
KNeighbours	0.00%

As can be seen in the results above in Tables 4.3 and 4.2, there was a 0% accuracy score for predicting name across both the test splits utilised on all classifiers. The suspected reason for this is due to a lack of data, with only 39 usable data sets in total and only 3.9 to 5.9 sets of data for testing depending on the percentage utilised this is obviously a fairly small amount. To analyse the data further, manual analysis will be employed to provide hopefully better results than the machine learning. Precision, recall and F_1 scores were not calculated due to the accuracy scores that were achieved.

Manual Analysis

Further, manual analysis was completed on the data for name inference from the keystroke dynamics from the participants. To prepare the data, the same pre-processing was used as with the machine learning. Initially, the length of the data sets were limited to 92 letters as some data was incomplete, so this provided the largest usable set of data, in line with

the shortest response received. This would also ensure a fair analysis over all the participants. Following this, the dwell times were sorted into descending order with the corresponding key code. The results for the 'space' button were removed and then finally the top 20% fastest were analysed for each participant.

It is important to note that each participant was analysed in isolation as to normalise differences in typing speeds, and not introducing a bias to participants who are more familiar with typing than others. This resulted in high levels of accuracy for each user, as can be seen in Table 4.4 below.

The main focus of the analysis was to determine how many letters were present that were common across the participants name and the fastest letters they typed as can be seen in the graph below in Figure 4.3. This was to enable us to potentially infer a participants name utilising these keystrokes, or to at least prove a positive correlation between letters in name and speed of typing. We found that 97.44% of participants had one or more letters present in their top 20% fastest keystrokes that were also in their name. This shows that there is a direct correlation between the speed of a user typing certain letters and their name, meaning that technically, we are able to infer a user's name from the letters they type fastest. The probable reason for this correlation would be that people type their own name frequently, as well as other key words that they type. This method could be used to infer more than just a participants name, such as passwords, partner or children's names and so on, depending which letters are typed most frequently.

This can be explained and rationalised by Fitts and Posner's [19]

three stage model. In this, a person goes through three stages to learning a motor skill; cognitive, associative and autonomous. Subsequently, as the stages progress, the attention required diminishes, resulting in autonomous behaviour without thinking. Whilst this initial application was related to sports, the same theory can be applied to typing.

There were a number of different participants with differing length of name, from three at the shortest, all the way to nine for the longest. Below, in Figure 4.3, we can see the average percentage of letters which were present in a name, based on the number of letters in the name.

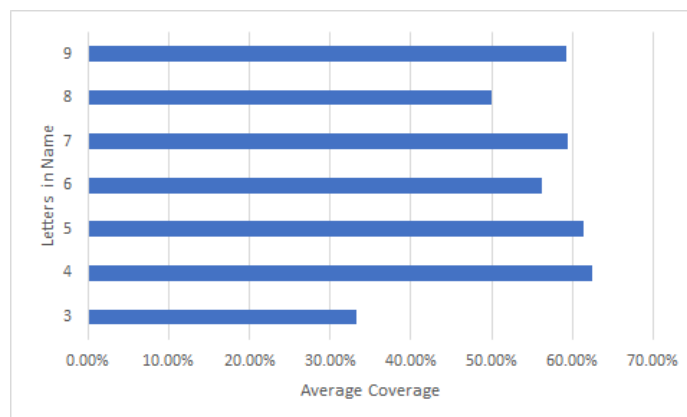


Figure 4.3: Pilot Study - Percentage coverage (average) sorted by length of name.

Interestingly, there was a small variance across many of the coverage scores of different name lengths showing that the method was consistent. Coverage, in this instance, refers to the percentage of letters in a name that were in the fastest keystrokes analysed. Names of four and five letters had the best coverage according to the data collected, although this may be due to this being the most frequent name length (see Table 4.4).

Table 4.4: Pilot Study - Average accuracy and frequency per letter - Manual analysis.

Letters in Name	Average	Frequency
3	33.33%	3
4	62.50%	4
5	61.43%	13
6	56.25%	8
7	59.52%	6
8	50%	1
9	59.26%	3

When looking at the data above, there are some inconsistencies. With regards to names of length three, these results were 33.33%, 0% and 66.66% respectively. Additionally, names of length eight had only one participant at a score of 50%.

Eliminating these outlying results, we can see that there was a coverage variance of only 11.11% showing the method was fairly robust and can be used with a certain degree of confidence.

4.5 Discussion

The results above of the pilot study show that it is possible to predict a User's gender, age and handedness from the way they type on a mobile device, with a high degree of accuracy. Looking at soft biometrics results which can be seen from Table 4.1, each class had at least one classifier which produced a score better than random chance. This was a

very positive outcome considering that there was a relatively small data set of 39 participants, combined with limited data compared to planned subsequent studies.

As mentioned previously, when looking at inferring name, a significantly large proportion of participants (97.44%) had letters from their name present in the top 20% of fastest keystrokes. Table 4.4 shows that there was also a high level of coverage with regards to inferring name, meaning that with these letters, we could potentially piece together a User's name from the data collected. Furthermore, we could identify other potentially sensitive information such as partner or child names, passwords and so forth from analysing the fastest letters that were typed.

Unfortunately, the machine learning results provided a 0% accuracy, precision, recall and F_1 score on average for inferring name. The suspected reasoning for this, as discussed briefly above, was due to the lack of data sets in general, as well as from users with the same names. In order to improve on this, a larger quantity of data will be sampled in the main study, alongside the accelerometer and gyroscope data to hopefully provide better results. This larger data set should provide us with a larger test set and subsequently a larger training set. Additionally, if a participant completes the experiment multiple times, we should obtain more data per name allowing us to hopefully predict name with a higher accuracy.

Overall, the results were positive from the pilot study, despite the lacking machine learning results for inferring name. The results that were produced due to manual analysis have provided fantastic results and allow us to happily continue with the inclusion of further data over

a wider range of participants in the next experiment.

As mentioned above, in order to further develop this study, experiment two which is discussed in the subsequent chapter, focuses on the addition of accelerometer and gyroscope data to further refine the method used, and to add additional context. We also hope to include a larger range of participants to provide a wider range of data for analysis. The same machine learning classifiers will be used both on the soft biometric features and also to hopefully infer name with a reasonable degree of accuracy to progress the research.

Chapter 5

Evaluation of the addition of further sensor input to enhance accuracy

This chapter discusses the analysis and results from the first main study, which corresponds to the second research question that we posed at the start of this thesis. To recall, the question was ‘what effect does the inclusion of accelerometer and gyroscopic data alongside keystroke dynamics have on the ability to successfully infer a person’s name and soft biometric features on a mobile device?’ We also explore the key differences and advances from the pilot study which was discussed in the previous chapter. The dataset utilised in this study and chapter can be seen in Table 3.1.

5.1 Introduction

The discussion in Chapter 2, alongside the results from the pilot study in Chapter 4, indicated the need for further experimentation of keystroke dynamics on mobile devices. In particular, the inclusion of gyroscope and accelerometer readings to provide enhancement to the data collected, and with the intention of increasing the accuracy with which we can infer a user's name or demographic features.

As can be seen below, this study captured not just data from the accelerometer and gyroscope, but also additional inputs, such as the size of the press on the screen, as well as x and y co-ordinates of the press, to see if these provide further useful context.

5.2 Participants

A total of 68 usable data sets were analysed, although as with the pilot study, over 100 data sets were discarded due to being incomplete and providing less than 20% of the data required. Below, in Figure 5.1 and Figure 5.2 is the split of data based on handedness, age and gender which shows a total split of 69.18% Male and 30.82% Female. As per UK population statistics [14], this is not indicative of what we would normally see in the population.

From the data below, we can see that the majority of all participants were right handed, with one ambidextrous user. We do have some left handed participants, and this class imbalance is what we would expect to see. Additionally, a large proportion of participants were in the 18-25 and 26-35 age brackets, with significantly fewer in the older brackets.

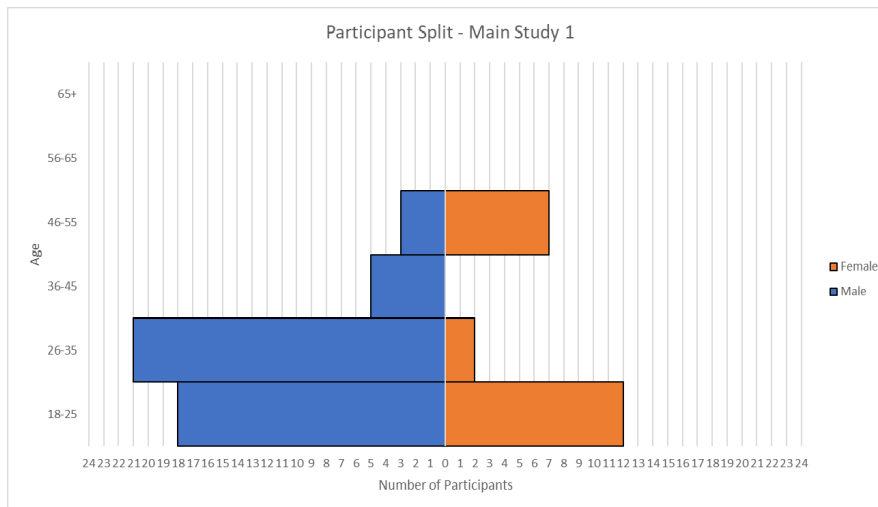


Figure 5.1: Spread of participants data with age and gender - Main Study 1

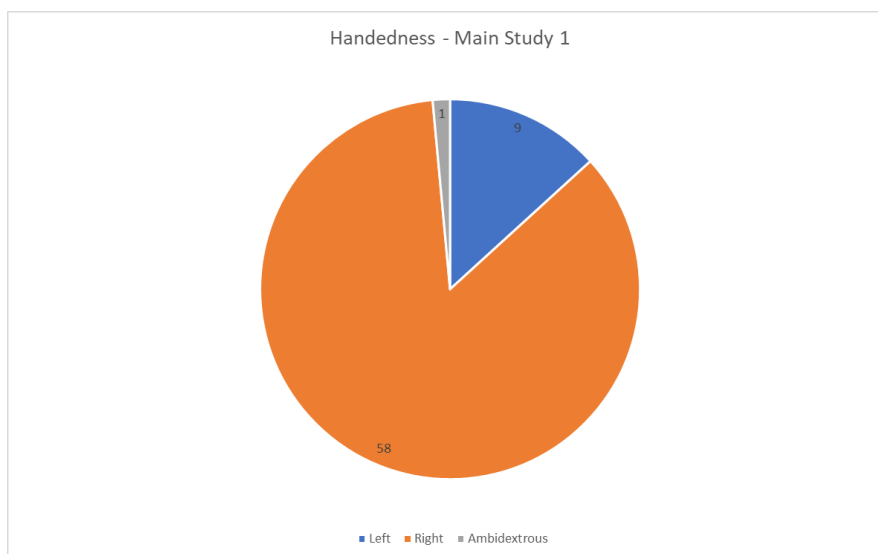


Figure 5.2: Spread of participants handedness data - Main Study 1

For both genders¹ present, we had a fairly broad range of age ranges, but without those above the 55 years bracket. This could be due to a

¹As with the other studies, a third option of ‘other’ was provided however no participants chose to proceed with this classification.

number of reasons, however, we did not actively recruit participants of a specific background or characteristic so we could therefore not control this. This is also more feasible for a realistic scenario in which we would not be able to control exact participants. Unsurprisingly, most of our data for both genders is within the lower age ranges for the survey.

5.3 Data Analysis

Below details the different analysis methods that were employed, including both the predictive model and the regular machine learning as with the previous experiment.

5.3.1 Data Preparation and Analysis

As with the pilot study, before analysing the data, cleaning and preparation was needed to ensure a uniform input for the machine learning which was later utilised. To do this, capitalisation and spaces were standardised across the name field for all of the data to ensure that the machine learning was able to interpret the data correctly. Additionally, splitting and removal of miscellaneous characters which were used in the code for organising the data, such as array parentheses, occurred to parse a clean data set. Once this formatting had been completed, dwell time was calculated from the data to provide a common data point for analysis. This resulted in the letter and dwell time being output to then be able to be parsed to the machine learning, or to be analysed manually.

To analyse the data for the soft biometric features and the name, a number of classifiers from the SciKitLearn library were utilised [45]. From

previous research conducted with colleagues [17], as well as reviewing the relevant literature mentioned in Section 2, the following classifiers were used:

- Gaussian NB;
- Decision Trees;
- Random Forest ($n = 10$, $n = 100$);
- SVC;
- k-NN.

Again, to prepare the data for the machine learning, the shortest data-set was chosen as the length with which to trim all the other data-sets, which was 92 characters in length. Once this was completed, we encoded the data utilising OneHotEncoder from the SciKitLearn library [45] and used a test split of 15% as this was the most commonly used in relevant studies, and an average was taken from 50 loops of the machine learning to ascertain which would provide the best accuracy scores for both name and soft biometric features. To do this, a for loop was included that re-ran the machine learning 50 times. Regardless of the scores, these were output into a .csv file and an average was computed using functions, these are then the scores presented in the research. For each of the classifiers, the dwell time and letter were analysed, alongside the particular soft biometric and name. As with all the experiments, inferring a user's name was the key objective and therefore dwell time made the most logical sense due to it representing the speed of which certain letters

were pressed. For the soft biometrics and name, in addition to accuracy we recorded precision, recall and F_1 score.

The analysis for the first main study utilised a different data set, which we can refer to as data set two. This data set consisted of 68 individual sets of data and included a variety of sensor data, as well as key pressed, dwell time and basic personal information as discussed in the application chapter.

5.3.2 Predictive Model

The main basis behind the predictive model was the notion of using a binary tree style method of analysis. This involved segmenting the keyboard into sections, with a baseline set for the centre of the keyboard. Once the keyboard had been divided into sections, based on the movement of the device, this then determined the quadrant the keystroke was located in and then also the letter.

The space bar on the keyboard was ignored, as these keystrokes are not taken into account during the analysis. The remaining keyboard was then divided up into four unique sections, these were:

Q W E R T A S D F
 Y U I O P H J K L
 A S D F Z X C V
 H J K L V B N M

Figure 5.3: Keyboard quadrants as determined by the predictive model analysis.

As can be seen above, there was some overlap of the keyboard, due to how the quadrants were assigned. Additionally, the letter ‘G’ which was in the center of the keyboard essentially presented as a fifth section as this was not counted in any of the other quadrants. A diagram can be seen in Figure 5.4 below.

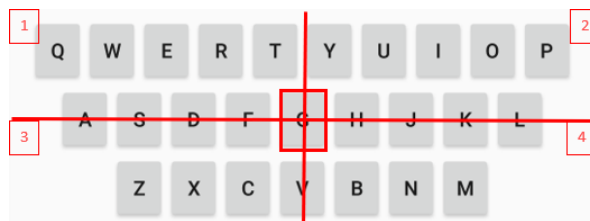


Figure 5.4: Keyboard with quadrants assigned for manual analysis.

Once the keyboard had been divided into sections, this allowed us to discard 75% of the keyboard from the X , Y and Z rotation data provided by the gyroscopic sensor. This provided a positive or negative value and based on these we know which way the phone rotated when the key was pressed, this can be seen in Figure 5.5 below.

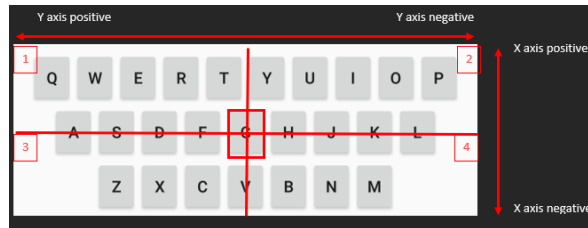


Figure 5.5: Keyboard with quadrants assigned and X/Y values for manual analysis.

To further refine the quadrants, maximum, minimum and average values were taken from the keystroke data and assigned to the corners and middle of the keyboard. For example, the largest positive and negative X axis coordinates would be our maximum/minimum for the movement of the X , Y and the same for our Z . This also applied to the averages which would provide our midpoint of the sections. The values were calculated by utilising the first sentence of the three sentences that were typed, or more specifically, the first occurrence of each letter. From these values we used these as the ‘training data’ to then conduct our blind analysis.

Once we were able to divide the keyboard up into smaller sections, this theoretically improved the accuracy. We then analysed the data from a user to attempt to infer and predict the keystrokes. In order to make analysis easier, we then mapped the median values for the X , Y and Z values onto the diagram as can be seen below in Figure 5.6.

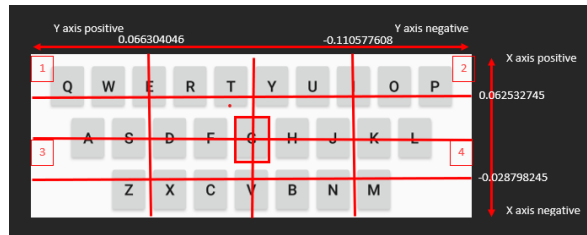


Figure 5.6: Keyboard with quadrants assigned and median values for manual analysis.

To provide more context, detailed manual analysis continued as follows:

1. Firstly, the maximum, minimum and median values were calculated for the particular user in question by utilising the first occurrence of each letter typed.
2. Next, the data was hidden to only show the next 30 keystrokes and only the X , Y and Z rotation values.
3. Based on these values, we then mapped each one onto the keyboard manually to predict a letter with a number.
4. Any ‘*space*’ letters at the end of the analysis were disregarded.
5. A letter was then noted and this continues until the data has been analysed.
6. Finally, we then compare the predicted letters against those that were actually typed.

This process is displayed below with one of the sample sets of data in Figure 5.7.

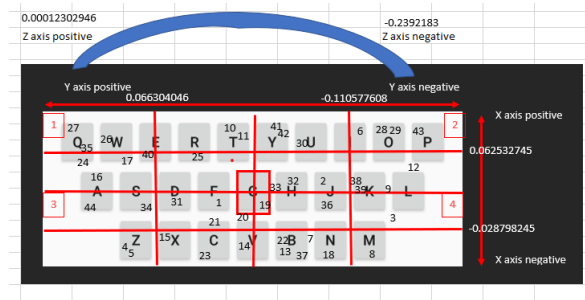


Figure 5.7: Keyboard with mapped values for manual analysis.

5.3.3 Machine Learning

As with the pilot study, machine learning was again employed to provide a more programmatic approach to the analysis in comparison with the statistical methods also utilised. To do this, much of the same cleaning was completed, alongside additional formatting to account for the extra data such as accelerometer and gyroscope. Once again, dwell time was calculated from the data to provide a common data point for analysis, with the addition of the accelerometer and gyroscope data.

As with the pilot study, the same classifiers were utilised in order to attempt to provide significant results, again utilising dwell time but with the inclusion of the gyroscopic data in the form of a dictionary. As before, a selection of training data split percentages were used and an average of 50 runs was taken for each classifier.

Finally, percentages were employed to provide a threshold for the data. This was achieved by taking the first occurrence of each letter, (utilising the dwell time and gyroscopic data from the previous analysis), and using this as effectively training data as a base point. From this,

a percentage threshold was set either side of the data, both positive and negative. This percentage varied across the X , Y and Z values to provide the best results possible from the data. This approach would theoretically account for a slight deviation in the position of the phone whilst still maintaining a small enough percentage threshold to provide an accurate prediction. Please see the code snippet below in Listing 5.1.

```
1     #Check to see if the letter is within n% + or - from
    the first occurrence of the letter.
2     if abs(float(letter[1])) >= abs(float(letterXMinus5))
    and abs(float(letter[1])) <= abs(float(letterXPlus5))
    and abs(float(letter[2])) >= abs(float(letterYMinus5))
    and abs(float(letter[2])) <= abs(float(letterYPlus5))
    and abs(float(letter[3])) >= abs(float(letterZMinus5))
    and abs(float(letter[3])) <= abs(float(letterZPlus5)):
3     print("Predict: " + trainingData + " Actual: " +
    letter[0])
```

Listing 5.1: Machine learning - Initial threshold code

Following analysis of this, a second threshold model was developed in order to encompass not only the X , Y and Z rotational values, but also the size, X , Y and Z accelerometer readings and X and Y coordinates of the touch on the screen. The code listing below (please see Listing 5.2) for this can be seen which includes these in the if statement as above. This code takes the same approach as the previous threshold model and uses the output from the previous machine learning which encompasses dwell time and all measurements for each key pressed. The same percentage threshold of n plus or minus was set which then applied to all of the thresholds analysed against the letter.

```

1  #Check to see if the letter is within n% + or - from
   the first occurrence of the letter.
2  if abs(float(letter[2])) >= abs(float(downXMinus5))
   and abs(float(letter[2])) <= abs(float(downXPlus5)) and
   abs(float(letter[3])) >= abs(float(downYMinus5)) and
   abs(float(letter[3])) <= abs(float(downYPlus5)) and abs
   (float(letter[4])) >= abs(float(sizeMinus5)) and abs(
   float(letter[4])) <= abs(float(sizePlus5)) and abs(
   float(letter[5])) >= abs(float(letterXMinus5)) and abs(
   float(letter[5])) <= abs(float(letterXPlus5)) and abs(
   float(letter[6])) >= abs(float(letterYMinus5)) and abs(
   float(letter[6])) <= abs(float(letterYPlus5)) and abs(
   float(letter[7])) >= abs(float(letterZMinus5)) and abs(
   float(letter[7])) <= abs(float(letterZPlus5)) and abs(
   float(letter[8])) >= abs(float(letterXMinus5R)) and abs
   (float(letter[8])) <= abs(float(letterXPlus5R)) and abs
   (float(letter[9])) >= abs(float(letterYMinus5R)) and
   abs(float(letter[9])) <= abs(float(letterYPlus5R)) and
   abs(float(letter[10])) >= abs(float(letterZMinus5R))
   and abs(float(letter[10])) <= abs(float(letterZPlus5R))
   :
3  print("Predict: " + trainingData + " Actual: " +
   letter[0])

```

Listing 5.2: Machine learning - Improved threshold code

5.4 Findings

Below the findings from the main experiment are discussed, categorised between soft biometric findings and keystroke dynamic findings. Ad-

ditionally, the keystroke dynamics section is further sub-categorised between the different machine learning classifiers, as well as further machine learning and predictive models that were employed.

5.4.1 Soft Biometric Findings

As with the previous study, a number of machine learning classifiers were used and can be seen below. A train/test split of 15% was used as with the pilot study, and an average was taken from 50 runs. Again, for each of the classifiers, the dwell time and letter were analysed alongside the particular soft biometric.

Table 5.1: Experiment 2 - (Main Study 1) - Soft biometric classifier accuracy, precision, recall and F_1 scores - Average of 50 runs.

Classifier	Scores											
	Handedness				Age				Gender			
	Accuracy	Precision	Recall	F_1	Accuracy	Precision	Recall	F_1	Accuracy	Precision	Recall	F_1
Gaussian NB (Gaussian Naive Bayes)	87.94%	86.14%	87.94%	86.20%	69.86%	70.30%	69.86%	64.74%	77.42%	76.56%	77.42%	72.98%
Decision Trees	89.02%	87.88%	89.02%	87.54%	57.24%	59.26%	57.24%	55.30%	73.28%	73.92%	73.28%	71.40%
Random Forest (n = 10)	91.90%	87.38%	91.90%	89.08%	61.24%	64.02%	61.24%	55.12%	78.58%	78.54%	78.58%	78.32%
Random Forest (n = 100)	91.00%	86.58%	91.00%	87.90%	62.96%	65.82%	62.96%	56.56%	80.74%	80.48%	80.74%	76.24%
SVC (Support Vector Classifier)	86.50%	75.88%	86.50%	80.62%	65.14%	68.32%	65.14%	58.68%	76.22%	76.66%	76.22%	70.38%
KNeighbours (K Nearest Neighbours)	82.36%	74.48%	82.36%	77.94%	45.56%	42.00%	45.56%	41.30%	66.58%	68.48%	66.58%	64.08%

Looking at handedness, Random Forest ($n = 10$) came out as the top classifier with a 91.90% accuracy score, however all of the classifiers produced accuracy scores that were better than chance, and all above 82%. When looking at the precision scores, we can see that we achieved a high accuracy across all classifiers which helps to validate the accuracy scores achieved. For recall, again we achieved high scores across the board showing we are returning positive results. Finally, the F_1 scores being high solidifies this analysis and the scores obtained.

For age, which was split into 6 distinct categories, GaussianNB produced the highest accuracy score with 69.86%. Again, this is much better than random chance. When looking at precision scores, we scored fairly highly overall, with a few scores dropping below 60% but the scores remained in line with the accuracy. Recall again produced good scores in line with accuracy and the F_1 scores again confirm this.

Finally, gender, which was split into two categories, has again provided an accuracy score better than random chance from random forest ($n=100$) at 80.74%. All of the classifiers for each accuracy score performed much better than random chance. Looking at the precision, recall and F_1 scores we can see that we scored highly here meaning that we are returning accurate and positive results in line with the high accuracy scores.

5.4.2 Keystroke Dynamic Findings

This section is split into two main sections, firstly, that of the developed predictive model, which uses a modified binary tree approach to analyse the data. Finally, we look at the different classification results from the

machine learning and the results that were achieved.

Predictive Model

To enable us to determine the viability of a manual predictive model, a total of five users were analysed using this method and an average taken. This was to identify if this method produced good results before continuing with the rest of the participants. Averaged across the group, we had a total correct percentage prediction of 5.24%. Splitting the data separately, two of the five participants had no correct predictions. The remaining three had separate accuracy scores of 6.66%, 11.53% and 8.00%.

To ensure there was no cherry picking of results, the participants were chosen completely at random utilising a random number generator. Using the generated numbers, data was selected and anonymised before the analysis began.

Unfortunately, the average accuracy prediction was not sufficient to pursue this analysis method further. Therefore, in order to produce potentially better results, as well as a less manual method of analysis, machine learning was employed.

Machine Learning

As mentioned above, we proceeded with a variety of machine learning classifiers that were used for the soft biometrics analysis in order to provide a more automated approach to the analysis for inferring name.

As can be seen from the below results, a comparative 10% split provided far better results than 15%.

Table 5.2: Experiment 2 - (Main Study 1) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (10% Test Size).

Classifier	Accuracy	Precision	Recall	F_1
GaussianNB	32.10%	28.68%	32.10%	29.28%
Decision Trees	28.08%	28.24%	28.08%	27.70%
Random Forest (n = 10)	25.76%	26.18%	25.76%	25.38%
Random Forest (n = 100)	35.50%	32.76%	35.50%	33.28%
SVC	28.60%	27.82%	28.60%	26.64%
KNeighbours	25.16%	23.18%	25.16%	23.70%

Table 5.3: Experiment 2 - (Main Study 1) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (15% Test Size).

Classifier	Accuracy	Precision	Recall	F_1
GaussianNB	32.48%	30.78%	32.48%	30.02%
Decision Trees	26.86%	28.96%	26.86%	27.04%
Random Forest (n = 10)	26.48%	25.60%	26.48%	24.42%
Random Forest (n = 100)	31.08%	29.84%	31.08%	29.10%
SVC	27.02%	28.38%	27.02%	26.02%
KNeighbours	18.72%	18.02%	18.72%	17.40%

As can be seen in Table 5.2 above, there was a high percentage accuracy score to be able to theoretically infer a user's name from the dwell time and gyroscopic data. This was achieved using 96 characters or less of data which, with a 35.50% highest accuracy score was very positive. Whilst this was not as high a score as we would like, it shows

promise based on the small amount of data that was available for this study. Therefore, hopefully with more data, the accuracy would improve. Looking to precision, recall and F_1 scores, these are all in line with the accuracy scores produced however as with accuracy, these were not as high as we would have liked. These scores also illustrated the need for further experimentation with a larger, qualitative data set to hopefully provide better results.

As with all of the previous machine learning, an average of 50 runs was taken to provide the data. The data set had a total of 41 unique names, with 68 overall participants. Applying the accuracy score of 35.50% to a set of 41 unique names provides us with a statistic that is better than random chance.

We can also see above that all of the classifiers were better than random chance at predicting name. This suggests that it is feasible to predict a users name from the way they type on a smartphone.

Threshold Model

In order to further improve the results, we employed a threshold to allow for programmatically assessing the gyroscopic data. This involved taking the data in the form of the letter typed, followed by the X, Y, Z rotational co-ordinates and attempted to predict a letter based on the threshold provided. This then repeated for each participant in the experiment.

To express an algorithm of this novel process, the below has been created for ease of understanding, along with accompanying explanations.

As can be seen in the first part of the algorithm below, for the first occurrence of each letter in the dataset (l in n), we store the letter, and

the X , Y and Z rotation data as α . As an output, we expect to return the predicted letter which is returned as β and the actual letter as γ .

Following this, we create percentage plus and minus variables as a threshold for each of the X , Y and Z rotation data, which was plus and minus $n\%$. This was then used to test the following values in the dataset to see if they matched the threshold. If the numbers being tested fell within that threshold, the letter was output to the console along with the actual letter in question.

Algorithm 1: Novel rotational threshold algorithm

```

Data:  $l = (\alpha, \alpha_x, \alpha_y, \alpha_z)$     /* Letter and Rotational Data */
Data:  $n$           /* Percentage Threshold Defined by User */
Result:  $\beta$                                      /* Predicted Letter */
Result:  $\gamma$                                      /* Actual Letter */

1  $xThreshold+$   $\leftarrow \alpha_x + n\%$ 
2  $yThreshold+$   $\leftarrow \alpha_y + n\%$ 
3  $zThreshold+$   $\leftarrow \alpha_z + n\%$ 
4  $xThreshold-$   $\leftarrow \alpha_x - n\%$ 
5  $yThreshold-$   $\leftarrow \alpha_y - n\%$ 
6  $zThreshold-$   $\leftarrow \alpha_z - n\%$ 
7 foreach  $l$  in  $n$  do
8   if  $\alpha_x \geq xThreshold- \wedge \alpha_x \leq xThreshold+ \wedge \alpha_y \geq$ 
    $yThreshold- \wedge \alpha_y \leq yThreshold+ \wedge \alpha_z \geq$ 
    $zThreshold- \wedge \alpha_z \leq zThreshold+$  then
9     Predicted Letter:  $\beta$ 
10    Actual Letter:  $\gamma$ 
11  end
12  continue
13 end

```

From this, we then iterated through the rest of the list and produced a percentage fit for the remaining data. As can be seen below, the percentages varied in success. Therefore, if the test data was within a particular threshold of the training data, we would successfully predict the letter. For example, if the threshold was 5% we would take a value of plus and minus 5% of the original value and then test the letter if it was within

this range.

As can be seen in Table 5.4 below, we ran the whole dataset through for each percentage threshold from 10% in 5% increments up to 50% to see if we could successfully predict the letter based on gyroscopic data alone.

Table 5.4: Experiment 2 - (Main Study 1) - Name inference threshold accuracy scores

Threshold	Correct	Incorrect	Percentage
10%	0	23	0%
15%	0	84	0%
20%	2	177	1.13%
25%	9	394	2.28%
30%	14	657	2.13%
35%	26	1069	2.43%
40%	38	1632	2.33%
45%	66	2389	2.76%
50%	88	3384	2.60%

As can be seen above, we had fairly poor results with the best accuracy being 2.76%. This could be due to a number of reasons, such as the user holding the phone on a flat surface and not rotating. Additionally, we were only utilising the rotational data to create the thresholds, which omits potentially useful additional data to supplement the analysis. In order to improve this model, more data points were included into a revised model to improve the accuracy.

Improved Machine Learning Threshold Model

As above with the initial threshold model developed, the same steps were taken with the analysis of the improved model, with the inclusion of the additional data captured. As discussed, this model includes the addition of size of press data, X , Y and Z accelerometer data as well as the X and Y press on the screen location.

Below is an algorithm of this improved threshold model to help with ease of understanding, along with accompanying explanations.

As can be seen in the first part of the algorithm below, for the first occurrence of each letter in the dataset (l in n), we store the letter, and the X , Y and Z rotation data, X , Y and Z movement data, X and Y press on the screen location, and size as α . As an output, we expect to return the predicted letter which is returned as β and the actual letter as γ .

Following this, we create percentage plus and minus variables as a threshold for each of the data, which was plus and minus $n\%$. This was then used to test the following values in the dataset to see if they matched the threshold. If the numbers being tested fell within that threshold, the letter was output to the console along with the actual letter in question.

Algorithm 2: Improved novel rotational threshold algorithm

Data: $l = (\alpha, \alpha_x, \alpha_y, \alpha_z, \alpha_{x^r}, \alpha_{y^r}, \alpha_{z^r}, \alpha_{x^d}, \alpha_{y^d}, \alpha_s)$
 /* Letter/Data */
Data: n /* Percentage Threshold Defined by User */
Result: β /* Predicted Letter */
Result: γ /* Actual Letter */
 /* Rotation Thresholds */
 1 $xRotationThreshold+ \leftarrow \alpha_{x^r} + n\%$
 2 $yRotationThreshold+ \leftarrow \alpha_{y^r} + n\%$
 3 $zRotationThreshold+ \leftarrow \alpha_{z^r} + n\%$
 4 $xRotationThreshold- \leftarrow \alpha_{x^r} - n\%$
 5 $yRotationThreshold- \leftarrow \alpha_{y^r} - n\%$
 6 $zRotationThreshold- \leftarrow \alpha_{z^r} - n\%$
 /* Movement Thresholds */
 7 $xThreshold+ \leftarrow \alpha_x + n\%$
 8 $yThreshold+ \leftarrow \alpha_y + n\%$
 9 $zThreshold+ \leftarrow \alpha_z + n\%$
 10 $xThreshold- \leftarrow \alpha_x - n\%$
 11 $yThreshold- \leftarrow \alpha_y - n\%$
 12 $zThreshold- \leftarrow \alpha_z - n\%$
 /* Location on Screen Thresholds */
 13 $xPressThreshold+ \leftarrow \alpha_{x^d} + n\%$
 14 $xPressThreshold- \leftarrow \alpha_{x^d} - n\%$
 15 $yPressThreshold+ \leftarrow \alpha_{x^y} + n\%$
 16 $yPressThreshold- \leftarrow \alpha_{x^y} - n\%$

```

/* Size Thresholds */

sizeThreshold+ ←  $\alpha_s + n\%$ 
sizeThreshold- ←  $\alpha_s - n\%$ 

foreach  $l$  in  $n$  do
  if  $\alpha_{x^d} \geq xPressThreshold - \wedge \alpha_{x^d} \leq$ 
     $xPressThreshold + \wedge \alpha_{y^d} \geq yPressThreshold - \wedge \alpha_{y^d} \leq$ 
     $yPressThreshold + \wedge \alpha_s \geq sizeThreshold - \wedge \alpha_s \leq$ 
     $sizeThreshold + \wedge \alpha_x \geq xThreshold - \wedge \alpha_x \leq xThreshold +$ 
     $\wedge \alpha_y \geq yThreshold - \wedge \alpha_y \leq yThreshold + \wedge \alpha_z \geq$ 
     $zThreshold - \wedge \alpha_z \leq zThreshold + \wedge \alpha_{x^r} \geq$ 
     $xRotationThreshold - \wedge \alpha_{x^r} \leq xRotationThreshold + \wedge \alpha_{y^r} \geq$ 
     $yRotationThreshold - \wedge \alpha_{y^r} \leq yRotationThreshold + \wedge \alpha_{z^r} \geq$ 
     $zRotationThreshold - \wedge \alpha_{z^r} \leq zRotationThreshold +$  then
    | Predicted Letter:  $\beta$ 
    | Actual Letter:  $\gamma$ 
  end
  continue
end

```

From this, as before with the previous model, we iterated through the rest of the list and produced a percentage fit for the remaining data. As can be seen below, the percentages varied in success. Therefore, if the test data was within a particular threshold of the training data, we would predict successfully the letter. For example, if the threshold was 5% we would take a value of plus and minus 5% of the original value and then test the letter if it was within this range.

Looking at Table 5.5 below, we ran the whole dataset through for each percentage threshold from 10% in 5% increments up to 50% to see if we could successfully predict the letter based on gyroscopic data as well as the additional data points.

Table 5.5: Experiment 2 - (Main Study 1) - Name inference improved threshold accuracy scores

Threshold	Correct	Incorrect	Percentage
10%	2	27	6.90%
15%	3	97	3%
20%	3	158	1.86%
25%	4	281	1.40%
30%	7	456	1.51%
35%	9	663	1.34%
40%	22	906	2.37%
45%	32	1232	2.53%
50%	41	1658	2.41%

As can be seen from the results, these were much improved over the initial algorithm that was created, where we had a slightly higher accuracy score for 10% over the previous analysis of 6.90% which was a 150% increase in accuracy. Overall we also had a significantly reduced amount of incorrect predictions, with a good level of predictions over the previous model. Whilst this was still producing less accurate results than we would like, it was a marked improvement, and with a larger set of data in the subsequent experiment, should hopefully improve the accuracy.

5.5 Discussion

Overall, we had positive results from the main study, which scored higher than those in the pilot study. The inclusion of accelerometer and gyroscope data, alongside a wider group of participants helped to significantly improve the results obtained.

With regards to soft biometrics, the same classifiers were used and each one of the features had results better than random chance that can be seen in Table 5.1. We can see that all of the results for accuracy handedness were well over random chance, with 91.90% being the highest result. For age, again all accuracy scores were better than random chance due to the split of six distinct categories for age rather than the binary approach of under or over 30 employed by a vast number of comparable experiments. Our best result for accuracy for age came at 69.86% which as stated, was far better than random chance. Finally, for gender accuracy, we again had far better results than random chance, with the highest being 80.74%. Additionally, our precision, recall and F_1 scores were all high meaning that our results were accurate and positive.

Looking to our keystroke dynamic findings, there were three key areas of analysis: the predictive model, machine learning, and the threshold machine learning. For the predictive model, we produced some results but these were far below the accuracy that we hoped to achieve with 11.53% as the best result. Therefore, this analysis method was discontinued and we moved over to machine learning. For the bulk classifiers as we had used for the soft biometric analysis, these produced some good results with a test/train split of 10% providing the best results. Our

best classifier was GaussianNB with a result of 35.50% accuracy, which in a data set of 68 participants with 41 unique names was a much better result than random chance. Unfortunately with the precision, recall and F_1 scores we did not achieve the results we would have liked, which correlates to the accuracy scores achieved. Finally, the threshold scores across a range from 10% to 50% threshold in 5% increments produced results that were again far below that which we would expect. Whilst the improved model did provide more accurate results, this was still a lot lower than would be expected. It was therefore hoped that this would improve with the inclusion of a larger data set in the following experiment.

From the results above, we can infer that the inclusion of the accelerometer and gyroscope data has significantly improved the accuracy with which we are able to identify a user's name from the way they type on a smartphone. From analysing the results from the threshold model, unfortunately the percentages were either too low to be able to predict a letter, or too high so that it included a larger range of letters meaning we ended up with a considerable amount of false predictions.

Despite the improvement in prediction going from near zero to just under 36%, we still believe that this number can be greatly improved with a larger range of data, and this is to be the focus of the next and final experiment. The experiment was set up to require participants to produce more data over a wider range of days with the hope of creating more training and test data for the classifiers to accurately predict a user's name. Within this data we had 68 total pieces of data, which whilst a larger quantity than the first experiment, this is still below the data level we would like to achieve. This was the main reason that we

feel was why the results were not as good as expected, despite achieving the objectives. We can therefore confidently answer the research question associated with the study in a positive manner, which is discussed at the end of this thesis.

Chapter 6

Evaluation of an increase in data per user to enhance accuracy

This chapter discusses the findings from the final main study of this research which was linked to our final research question. As a reminder, the question was ‘To what extent does the volume of data per user help to improve the accuracy of the prediction of name and soft biometric features?’ We discuss results in detail and the application of these, as well as participant information, key findings and results from the experiment. The dataset utilised in this study and chapter can be seen in Table 3.1.

6.1 Introduction

The results from the previous study, again indicated the need for further experimentation. With the inclusion of the gyroscope and accelerometer readings, we achieved improved results on those present in the pilot study. From the previous analysis completed, it is predicted that these results could be improved further to provide a more robust approach to identification. We aimed to increase the accuracy by focusing on a larger set of data per user to parse to the machine learning.

As can be seen in the following paragraphs, we collected a larger quantity of data per user. Whilst the overall dataset numbers were lower, far more data per user was collected. Participants were asked to complete the experiment multiple times and over a number of days, in order to allow us to gain a significant amount of data per participant to assist with the identification. This then enabled us to train the model more sufficiently to be able to predict users with greater accuracy.

6.2 Participants

A total of 46 data sets were analysed, and as with the other studies, a large proportion of data was discarded due to being incomplete. We had a total of 1,964 submissions with only 2.34% of the data being usable. Most of this was ‘spam’ data, as the application was available globally on the Google Play Store, anyone could download the application and participate. Most of the ‘spam’ data was submitted without any information at all and just a blank text file.

Looking at the data that was able to be analysed, in Figure 6.1 and

Figure 6.2 we can see the split of users who were in the different age categories, gender and handedness. A large proportion of both the male and female participants were all within the younger age brackets, with few in the older sections, and none in the oldest sections of 55-65 and 65+. Interestingly, in the male participants, there were only the three youngest age brackets, with no participants taking part over the age of 45. For female participants, there were distinctively more in the 18-25 age bracket than the other two, with no participants in the 36-45, or older two age brackets either.

Whilst the application was present on the Google Play Store globally so that anyone could participate, the primary recruitment came from the author's social media (LinkedIn and Facebook) and as such this may have had an effect on the ages of the participants. Theoretically, this could be due to the network of contacts being younger across the author's social media, as well as some elderly participants being more wary or uninterested in taking part.

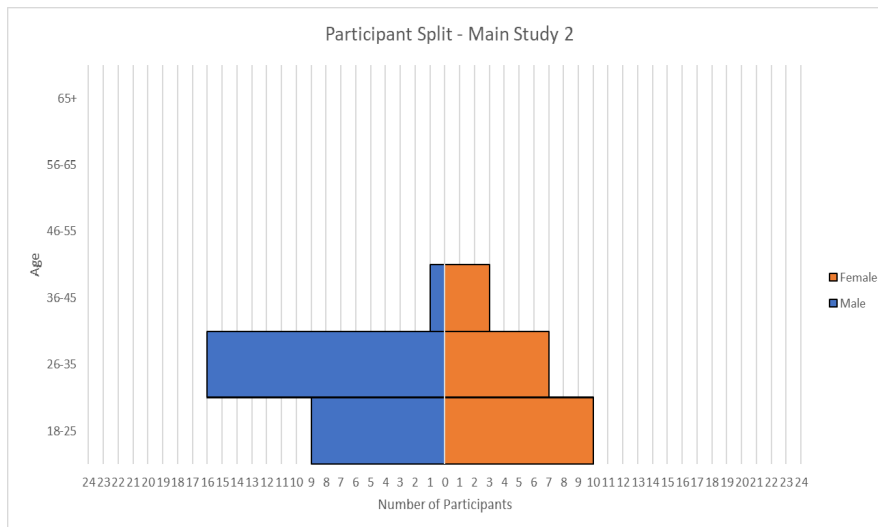


Figure 6.1: Spread of participants data with age and gender - Main Study 2

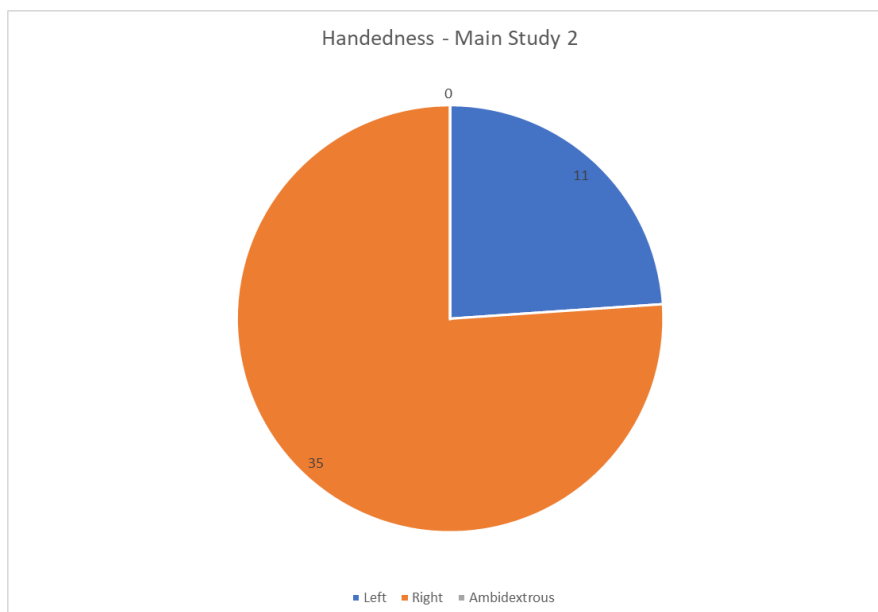


Figure 6.2: Spread of participants handedness data - Main Study 2

Looking at the data as a whole, we had no ambidextrous participants. Unfortunately, as we did not specifically select candidates for the exper-

iment, it was therefore not possible to target ambidextrous users to take part.

With regards to gender, we had a total split of 56.52% male participants and a 43.48% response from female participants. There is no particular class imbalance here and we have a good mix of both left and right handed participants in both gender categories. It is important to note that this population skew, whilst close in terms of percentages, it is not indicative of the UK population [14].

6.3 Data Preparation and Analysis

As with the first main study, before analysing the data, cleaning and preparation was needed to ensure a uniform input for the machine learning which was later utilised. To do this, capitalisation and spaces were standardised across the name field for all of the data to ensure that the machine learning was able to interpret the data correctly. Additionally, splitting and removal of miscellaneous characters which were used in the code for splitting the data, such as array parentheses, occurred to parse a clean data set. Once this formatting had been completed, dwell time was calculated from the data to provide a common data point for analysis. This resulted in the letter and dwell time being output to then be able to be parsed to the machine learning, or to be analysed manually.

To analyse the data for the soft biometric features and the name, a number of classifiers from the SciKitLearn library were utilised [45]. From previous research conducted with colleagues [17], as well as reviewing the relevant literature mentioned in Section 2, the following classifiers were

used:

- Gaussian NB;
- Decision Trees;
- Random Forest ($n = 10$, $n = 100$);
- SVC;
- k-NN.

Again, to prepare the data for the machine learning, the shortest data-set was chosen as the length with which to trim all the other data-sets, which was 92 characters in length. Once this was completed, we encoded the data utilising OneHotEncoder from the SciKitLearn library [45] and used a test split of 15% as this was the most commonly used in relevant studies, and an average was taken from 50 loops of the machine learning to ascertain which would provide the best accuracy scores for both name and soft biometric features. To do this, a for loop was included that re-ran the machine learning 50 times. Regardless of the scores, these were output into a .csv file and an average was computed using functions, these are then the scores presented in the research. For each of the classifiers, the dwell time and letter were analysed, alongside the particular soft biometric and name. As with all the experiments, inferring a user's name was the key objective and therefore dwell time made the most logical sense due to it representing the speed of which certain letters were pressed. For the soft biometrics and name, in addition to accuracy we recorded precision, recall and F_1 score.

The analysis for the second main study again utilised a different data set, which we referred to as data set three. This data set, consisted of 46 individual sets of data and included a variety of sensor data, as well as key pressed, dwell time and basic personal information as discussed in the application chapter. This data set differed to data set two by having more repeat participants, so we collected more data per user.

6.4 Findings

Next, the findings are discussed from the second main experiment, which have once again been categorised between soft biometrics and keystroke dynamics. Keystroke dynamics findings have been further split between the machine learning classifiers, as well as the threshold model.

6.4.1 Soft Biometric Findings

As with the pilot and main studies, the same group of machine learning classifiers were utilised for the soft biometric features analysis. Once again, a test split of 15% was used and the average taken from 50 runs for each classifier. Dwell time was once again used to calculate the prediction alongside the letter and other measurements for each soft biometric.

Table 6.1: Experiment 3 - (Main Study 2) - Soft biometric classifier accuracy, precision, recall and F_1 scores - Average of 50 runs.

Classifier	Scores											
	Handedness				Age				Gender			
	Accuracy	Precision	Recall	F_1	Accuracy	Precision	Recall	F_1	Accuracy	Precision	Recall	F_1
Gaussian NB (Gaussian Naive Bayes)	93.18%	92.92%	93.18%	91.68%	87.74%	89.72%	87.74%	86.50%	82.56%	86.90%	82.56%	80.66%
Decision Trees	88.68%	90.56%	88.68%	88.14%	63.62%	72.82%	63.62%	63.06%	74.80%	78.44%	74.80%	73.96%
Random Forest (n = 10)	87.46%	87.62%	87.46%	84.80%	61.94%	69.70%	61.94%	58.56%	73.42%	80.42%	73.42%	71.26%
Random Forest (n = 100)	90.58%	90.72%	90.58%	88.80%	69.90%	73.86%	69.90%	66.26%	75.36%	80.70%	75.36%	72.90%
SVC (Support Vector Classifier)	91.52%	90.96%	91.52%	89.80%	61.60%	68.70%	61.60%	57.08%	71.92%	74.92%	71.92%	67.10%
KNeighbours (K Nearest Neighbours)	87.20%	89.76%	87.20%	86.92%	54.84%	61.66%	54.84%	52.88%	69.62%	75.38%	69.62%	68.08%

Looking at Table 6.1 above, we can see that GaussianNB was the most successful classifier for age, gender and handedness.

The handedness scores for all classifiers were consistently high across the group, with the highest at 93.18%. The consistently high accuracy of these scores is promising and allows us to confidently predict handedness. This is solidified by the consistently high precision, recall and F_1 scores meaning that we have highly accurate and positive results.

Looking at age, 87.74% was a significantly high score, especially taking into consideration the split of different age categories that were present in the dataset analysed. Additionally, whilst all the other results are between 54% to 69%, these were still very good scores and show a robustness to the prediction across the data. Again, as with handedness, we can see we had high precision, recall and F_1 scores across the board for age meaning we also had highly accurate and positive results for this feature.

Finally, looking at gender with a 82.56% accuracy score was again far better than random chance. We also had consistently high scores across all of the classifiers, again resulting in a robustness to the predictive method. As with the handedness and age scores, we had consistently high precision, recall and F_1 scores for all of the gender results too meaning again we had highly accurate and positive results.

6.4.2 Keystroke Dynamic Findings

This section is split into two main sections, firstly, that of the machine learning classifiers that were utilised for soft biometrics. Secondly, we look at the machine learning threshold model to help to analyse the

data.

Machine Learning

Initially, we chose to again run the same group of classifiers as with the previous study and soft biometrics. To test which would derive the best results, a test split of both 10% and 15% were utilised to determine the better test size for this analysis.

Table 6.2: Experiment 3 - (Main Study 2) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (10% Test Size).

Classifier	Accuracy	Precision	Recall	F_1
GaussianNB	80.40%	81.74%	80.40%	79.32%
Decision Trees	70.00%	71.24%	70.00%	68.96%
Random Forest (n = 10)	79.20%	77.04%	79.20%	76.62%
Random Forest (n = 100)	83.20%	81.48%	83.20%	81.06%
SVC	42.80%	34.08%	42.80%	35.60%
KNeighbours	46.40%	46.80%	46.40%	44.04%

Table 6.3: Experiment 3 - (Main Study 2) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (15% Test Size).

Classifier	Accuracy	Precision	Recall	F_1
GaussianNB	78.28%	76.58%	78.28%	75.42%
Decision Trees	67.56%	68.58%	67.56%	65.38%
Random Forest (n = 10)	68.02%	69.80%	68.02%	66.16%
Random Forest (n = 100)	67.70%	69.00%	67.70%	65.00%
SVC	41.48%	33.00%	41.48%	34.14%
KNeighbours	49.36%	47.98%	49.36%	46.06%

As can be seen in Table 6.2 above, a 10% split produced mostly better results and with GaussianNB as the most accurate classifier by a distinct margin across both test sizes. The analysis was completed with 107 characters or less due to the length of the shortest piece of data, which is improved over the first set of data analysed in experiment two.

We had a total of 46 pieces of data with 12 unique names across the group which resulted in more data per user which in turn has produced more accurate analysis and prediction. Once again, an average was taken from 50 runs for each classifier to produce a fair and average result.

Looking at the number of participants and the accuracy scores from the classifiers, we can see that all produced results were better than random chance. With a highest accuracy score of 83.20% we had a very robust method of predicting name based on the keystrokes, dwell time and extra inputs. Looking at precision scores, we can see that overall we had highly accurate results on most of the classifiers, with the scores dropping slightly on those that were not as accurate. For recall, we again

had mostly high recall scores meaning we had positive results. The F_1 score for all of these also matches the same pattern as with the accuracy, precision and recall scores, producing a positive set of results overall.

Threshold Model

As with the main study, we employed the improved threshold model (see Algorithm 2) which utilised a programmatic threshold which ingested the data in the form of letter, followed by the X , Y , Z rotation, as well as size, accelerometer data and screen press data. A full explanation can be read in the previous chapter.

We then analysed from 10% to 50%, increasing in 5% increments analysing the amount of correct and incorrect predictions.

Table 6.4: Experiment 3 - (Main Study 2) - Name inference threshold accuracy scores

Threshold	Correct	Incorrect	Percentage
10%	2	152	1.30%
15%	11	284	3.73%
20%	11	452	2.38%
25%	18	671	2.61%
30%	28	980	2.78%
35%	30	1373	2.14%
40%	38	1748	2.13%
45%	53	2254	2.30%
50%	75	2898	2.52%

As can be seen from Table 6.4 above we had a highest accuracy of

3.73% when utilising the threshold algorithm. This is interestingly for the 15% threshold, which was a narrow margin around each keystroke. From looking at the number of incorrect predictions, this was the probable cause of the higher level of accuracy due to the sheer number of incorrect predictions for the higher thresholds. With a 50% threshold both positive and negative of the actual value, this ended up encompassing a very large range of keystrokes, resulting in a worse accuracy.

Overall the results from the threshold model, despite having more data per user, were not consistent with the results we would wish to see.

6.5 Discussion

After reviewing the results in detail, the accuracy scores significantly passed those of both the pilot and first main studies. From this we can deduce that the higher density of data per participant or name, significantly increased the accuracy, especially within the smaller dataset for experiment three.

Looking at soft biometrics, as before, the same classifiers were used in order to ensure a level comparison across experiments. Looking at Table 6.1 we can see that all of the results for handedness were well into the >87% to <93% accuracy range which was similar to those of the main experiment completed previously. Where we see a real change was in the accuracy scores of age and gender. Age, which was split into six different categories as with previous experiments scored far above random chance with the highest score being 87.74% which was a significant uplift on the best score from the main study. Finally, gender scored far above

that of the previous experiment with an accuracy score of 82.56%. This, whilst also being better than random chance, was a slight improvement on the previous results obtained. As above, the precision, recall and F_1 scores were all consistently high meaning we achieved highly accurate and positive results.

Moving on to keystroke dynamic findings in Tables 6.3 and 6.2, with the exclusion of the threshold model which was fairly inaccurate, we can again see a marked improvement upon the previous results obtained. Interestingly, the difference in accuracy scores between a 10% test size compared with a 15% test size only produced marginally better results for the smaller test size. GaussianNB yielded the best accuracy score for predicting name with 83.20% which was positive considering the 12 unique names over 46 pieces of data. We also have highly accurate and positive results as described by the precision, recall and F_1 scores achieved.

Following the classifier scores, the improved threshold model (see Table 6.4) was again utilised. Unfortunately, despite the increase in data per person, the model was slightly more effective, producing a top accuracy score of just 3.73% at inferring the correct letter based on the threshold of the X , Y and Z rotation data as well as size, X and Y press coordinates and the X , Y , Z accelerometer data.

From the results we have produced, we can confidently say that the inclusion of a higher frequency of data per name, significantly increased the accuracy of the prediction for each user. The comparison of the data will be fully analysed in subsequent chapters, however, a marked increase from 35.50% as a highest accuracy name prediction from the first main study to 83.20% was a 134.37% improvement. Unfortunately, whilst

the accuracy scores of the machine learning classifiers was improved, the threshold analysis remained poor with a high number of false predictions.

Despite this, we can positively say that we have again answered the research question corresponding to this study, which will be explained in detail at the end of the thesis.

Chapter 7

Evaluation of a combination of datasets to further enhance accuracy

In this chapter we look at a combination of the data collected in both of the main studies to provide us with a larger dataset to analyse as can be seen in Table 3.1. The results are then presented with subsequent discussion and analysis.

7.1 Combination Findings

Due to the amount of data collected in experiment three in isolation, and in order to analyse a larger dataset, the data from experiments two and three were combined to give us a dataset with a total of 114 pieces of data. The hope was that with this larger dataset we were able to prove the robustness of the model, with a bigger split of ages, handedness and

gender.

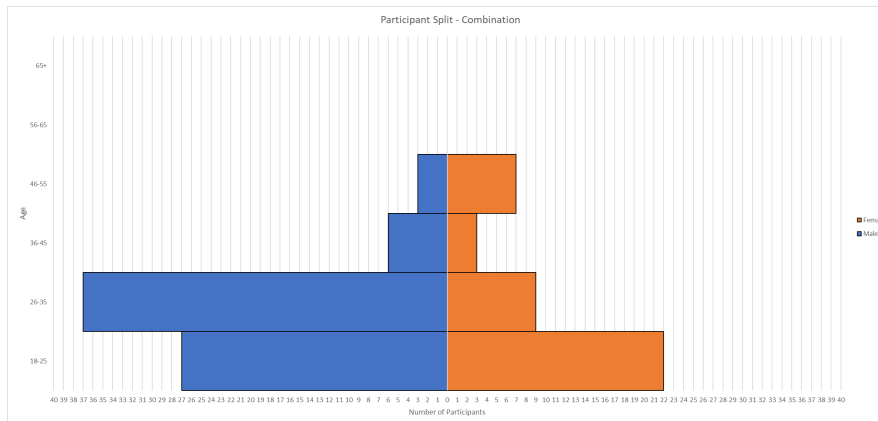


Figure 7.1: Spread of participants data with age - Combination of Experiments 2 and 3 - (Main Studies 1 and 2).

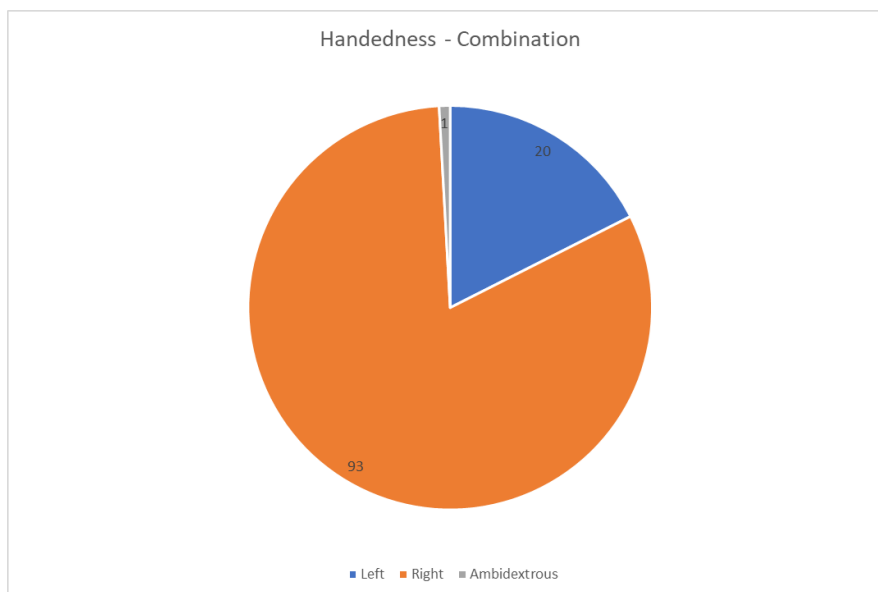


Figure 7.2: Spread of participants handedness data - Combination of Experiments 2 and 3 - (Main Studies 1 and 2).

Looking at the data that was combined, in Figure 7.1 and Figure 7.2 we can see the split of users who were in the different age categories,

gender and handedness.

Overall for handedness in the combined data set we had a percentage split of 81.58% right handed, with 17.54% left handed and 0.88% ambidextrous, which whilst not exactly a sample of the UK population, this is closer to what we would expect to see.

Looking at gender, we can see we had a split of 64.04% male, and 35.96% female. This is not indicative of the UK population as per the statistics seen [14].

7.2 Data Preparation and Analysis

As with the previous studies, before analysing the data, cleaning and preparation was needed to ensure a uniform input for the machine learning which was later utilised. To do this, capitalisation and spaces were standardised across the name field for all of the data to ensure that the machine learning was able to interpret the data correctly. Additionally, splitting and removal of miscellaneous characters which were used in the code for splitting the data, such as array parentheses, occurred to parse a clean data set. Once this formatting had been completed, dwell time was calculated from the data to provide a common data point for analysis. This resulted in the letter and dwell time being output to then be able to be parsed to the machine learning, or to be analysed manually.

To analyse the data for the soft biometric features and the name, a number of classifiers from the SciKitLearn library were utilised [45]. From previous research conducted with colleagues [17], as well as reviewing the relevant literature mentioned in Section 2, the following classifiers were

used:

- Gaussian NB;
- Decision Trees;
- Random Forest ($n = 10$, $n = 100$);
- SVC;
- k-NN.

Again, to prepare the data for the machine learning, the shortest dataset was chosen as the length with which to trim all the other data-sets, which was 92 characters in length. Once this was completed, we encoded the data utilising OneHotEncoder from the SciKitLearn library [45] and used a test split of 15% as this was the most commonly used in relevant studies, and an average was taken from 50 loops of the machine learning to ascertain which would provide the best accuracy scores for both name and soft biometric features. To do this, a for loop was included that reran the machine learning 50 times. Regardless of the scores, these were output into a .csv file and an average was computed using functions, these are then the scores presented in the research. For each of the classifiers, the dwell¹ time and letter were analysed, alongside the particular soft biometric and name. As with all the experiments, inferring a user's name was the key objective and therefore dwell time made the most logical sense due to it representing the speed of which certain letters were pressed. For the soft biometrics and name, in addition to accuracy we recorded precision, recall and F_1 score.

¹As a reminder, dwell time is defined as the length of time a key is pressed for.

The analysis for the combination of the main study data sets utilises a combination of data sets two and three, which we can refer to as data set four. This data set, consisted of 114 individual sets of data and included a variety of sensor data, as well as key pressed, dwell time and basic personal information as discussed in the application chapter.

7.2.1 Soft Biometric Findings

For the combined findings, we utilised the same 15% test split, alongside the same group of classifiers utilised previously. The average was once again taken from a total of 50 runs per classifier, as with the other experiments to allow for a detailed comparison.

Table 7.1: Experiment 2 and 3 - (Main Studies 1 and 2) - Soft biometric classifier accuracy, precision, recall and F_1 Scores - Average of 50 runs.

Classifier	Scores											
	Handedness				Age				Gender			
	Accuracy	Precision	Recall	F_1	Accuracy	Precision	Recall	F_1	Accuracy	Precision	Recall	F_1
Gaussian NB (Gaussian Naive Bayes)	88.66%	86.34%	88.66%	85.76%	73.44%	74.92%	73.44%	71.00%	77.44%	84.00%	77.44%	74.40%
Decision Trees	86.36%	86.14%	86.36%	85.54%	62.68%	65.54%	62.68%	61.48%	74.66%	76.76%	74.66%	73.88%
Random Forest (n = 10)	89.38%	88.56%	89.38%	87.16%	62.92%	67.38%	62.92%	60.06%	73.84%	78.72%	73.84%	69.48%
Random Forest (n = 100)	90.76%	90.26%	90.76%	88.84%	63.72%	70.22%	63.72%	60.62%	76.30%	80.76%	76.30%	71.94%
SVC (Support Vector Classifier)	90.16%	89.74%	90.16%	88.16%	60.68%	73.20%	60.68%	56.10%	75.96%	83.60%	75.96%	71.80%
KNeighbours (K Nearest Neighbours)	78.08%	70.64%	78.08%	72.94%	44.38%	43.16%	44.38%	42.16%	69.46%	71.08%	69.46%	68.34%

Looking at Table 7.1 above, we can see that we produced a high accuracy score across all features. Compared to the standalone results for experiments two and three, we had fairly similar accuracy scores across the board which are mostly better than experiment two, however, we were slightly less accurate than those in experiment three in isolation. This is what we would expect given the accuracy scores of experiment two in isolation and how these scores would translate into the combination results by lowering the accuracy slightly.

For handedness, Random Forest ($n = 100$) was the top classifier with a 90.76% accuracy score, with all other classifiers producing scores above 78%. This was a reassuringly high score across a large dataset, proving the robustness of the method for analysing handedness. For precision, recall and F_1 scores we scored highly across the board meaning that we once again had highly accurate and positive results.

Age, again scored highly with a 73.44% accuracy with GaussianNB being the most accurate classifier. Due to the combination of data, there were in fact a high number of different age brackets and all other classifiers scored highly, apart from k-NN which came in just under 45%. Despite this lower score for one classifier, all results were still much higher than a random chance. As with the handedness scores, we again achieved highly accurate and positive results as confirmed by the precision, recall and F_1 scores.

Finally, with regards to gender, we again scored highly across all classifiers, with GaussianNB coming in highest with 77.44%. As with the other features, gender scored consistently highly across all classifiers once again proving the robustness of the method. To solidify these scores,

the precision, recall and F_1 scores were also high.

7.2.2 Keystroke Dynamic Findings

Once again, this section is split into two subsections. Firstly we have the standard machine learning model utilising the previous classifiers decided upon. Secondly, we have the machine learning threshold model.

Machine Learning

With the classifiers below, we utilised both a 10% and 15% split again to keep the experiments analysis the same to allow us to accurately compare results.

Once again, the split of 10% yielded on average better results than that of 15% which was due to the increase in training data.

Table 7.2: Experiment 2 and 3 - (Main Studies 1 and 2) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (10% Test Size).

Classifier	Accuracy	Precision	Recall	F_1
GaussianNB	49.72%	43.42%	49.72%	43.68%
Decision Trees	42.20%	43.96%	42.20%	42.14%
Random Forest (n = 10)	41.12%	39.24%	41.12%	38.40%
Random Forest (n = 100)	47.40%	41.74%	47.40%	42.94%
SVC	28.84%	25.90%	28.84%	24.72%
KNeighbours	27.40%	26.56%	27.40%	25.12%

Table 7.3: Experiment 2 and 3 - (Main Studies 1 and 2) - Name inference classifier accuracy, precision, recall and F_1 scores - Average of 50 runs (15% Test Size).

Classifier	Accuracy	Precision	Recall	F_1
GaussianNB	48.20%	42.64%	48.20%	42.46%
Decision Trees	42.72%	46.66%	42.72%	43.38%
Random Forest (n = 10)	42.56%	41.66%	42.56%	39.74%
Random Forest (n = 100)	47.80%	42.54%	47.80%	42.46%
SVC	27.40%	25.10%	27.40%	23.44%
KNeighbours	21.46%	20.34%	21.46%	19.18%

Table 7.2 shows that we had over a 49% accuracy for predicting name, which was significantly better than random chance considering the vast amount of unique names across the entire group of data. Even the lowest of the scores from KNeighbours was still better than a random chance within the data. This compounded the previous assumptions from experiment two and experiment three in isolation around being able to predict a users name from the way they type on a smartphone, and has solidified this with an even greater accuracy. There were a total of 51 unique names in the data combined for experiments two and three, from a total of 114 pieces of data. Looking at precision, recall and F_1 scores these are all in line with the accuracy scores for each one meaning that the results were accurate and positive.

Whilst this was nowhere near as high of an accuracy as experiment three in isolation, we still produced a highly accurate result. We expected to reduce accuracy on this, given the sparsity of specific participant's data being diluted across the dataset. This however was still a positive result

and confirms the feasibility of this method of prediction.

Machine Learning Threshold Model

Finally, the improved threshold model was once again utilised (see Algorithm 2) as with the previous studies data separately. Combining the data resulted in the largest number of correct predictions per percentage, however we also had a very high number of incorrect predictions which resulted in our threshold accuracy percentages remaining in a similar ballpark to those previously analysed, as with error rate.

Table 7.4: Experiment 2 and 3 - (Main Studies 1 and 2) - Name inference threshold accuracy scores.

Threshold	Correct	Incorrect	Percentage
10%	4	182	2.15%
15%	14	367	3.67%
20%	14	633	2.16%
25%	22	1008	2.14%
30%	36	1540	2.28%
35%	45	2235	1.97%
40%	68	2973	2.24%
45%	103	3965	2.53%
50%	140	5284	2.58%

Unfortunately the 3.67% accuracy score that we achieved was lower by 0.06% from that of experiment three in isolation. However, the significant increase in data that was analysed with only a 0.06% variance shows that the model is able to cope with larger amounts of data and still

provide a similar result, proving the robustness, albeit with low accuracy scores. Theoretically, the reduction in accuracy could be accounted for from the increase of data per name, not necessarily being from the same participant. With experiment two, we had large quantities of data being submitted per participant over many days, with the first experiment there were similar names but likely different participants; therefore, the skewing of these has possibly degraded the performance slightly.

7.3 Discussion

Due to the similar nature of the experiment datasets, the combined analysis was done over the full range of data, which produced a total of 114 unique pieces of data. The higher density of data per person helped to increase the accuracy scores over those of experiment two, bringing them just below those from experiment three in isolation.

The soft biometrics findings were largely similar with a slight drop in accuracy over the full dataset but still producing a higher range of results than experiment two in isolation. Looking at Table 7.1 we can see that GaussianNB had the highest classifier accuracy scores for age and gender whereas Random Forest (n=100) scored highest in handedness. We achieved a 90.76% accuracy for handedness, which was similar to that achieved in isolation for both experiments. Looking at age, we achieved 73.44% which was a few percent better than that of the result from experiment two in isolation. Whilst it was less than that achieved from experiment three, it was still far better than random chance. Finally, with regards to gender we achieved a very similar result to that of experiment

two in isolation, with 77.44%. Whilst this was worse than experiment three in isolation, it was still by far better than random chance, and it also means that we can consistently predict accurately over a larger data set. As mentioned above, the precision, recall and F_1 scores were all in line with the accuracy scores for each soft biometric and achieved high results across the board.

Moving next to keystroke dynamic findings for name in Table 7.3, we again focused on the main group of classifiers and the threshold model. Looking at the results, the highest accuracy was for GaussianNB producing a score of 49.72%. Whilst this was lower than experiment three in isolation, it did provide an improvement on experiment two and also was far better than random chance with a total of 51 unique names across 114 pieces of data. The precision, recall and F_1 scores for inferring name were also positive, however they were not as high as experiment three in isolation as has been discussed.

As with the individual experimentation, the threshold model was once again attempted in Table 7.4, however yielded similar results with 3.67% accuracy.

From the results collated, we can again confidently say that we can accurately predict name, age, handedness and gender for each participant with a high degree of accuracy. Unfortunately, the results did reduce in accuracy due to the larger spread of data over the participants, however they were still very accurate, especially for 114 pieces of data.

Chapter 8

Conclusions and Future Work

In this chapter, we conclude the thesis and discuss the implications and importance of the conducted experiments and analysis. Each of the studies is compared and discussed to provide the outcomes across the studies, as opposed to isolated analysis. Finally, potential options for future research is discussed to allow us to examine the continued application of this novel field of inference utilising keystroke dynamics on mobile devices.

8.1 Research Questions

At the beginning of this thesis, we posed three questions that would form the basis for the experimentation and contributions. These questions were chosen to both enhance and contribute to the knowledge that currently exists, but also to follow on from one another, resulting in a coherent series of research and experimentation. As a reminder, the three research questions posed were as follows:

1. To what extent can keystroke dynamics be utilised in order to infer a person's name on a mobile device?
2. What effect does the inclusion of accelerometer and gyroscopic data alongside keystroke dynamics have on the ability to successfully infer a person's name and soft biometric features on a mobile device?
3. To what extent does the volume of data per user help to improve the accuracy of the prediction of name and soft biometric features?

8.1.1 Research Question 1

To what extent can keystroke dynamics be utilised in order to infer a person's name on a mobile device?

The first question, which corresponded with the pilot study, looked at inferring name by recreating keystroke dynamics on a mobile device. Keystroke dynamics, which are widely utilised on physical keyboards, were created via the use of buttons within an Android application which can be fully understood in previous chapters.

Upon analysis of the results from the pilot study, we can see that 97.44% of the 39 unique data sets had letters which were both present in their name, and the top 20% of their fastest keystrokes. We can therefore see there is a direct correlation between the speed of someone typing and their name, which we can surmise is due to the frequency of typing this. Furthermore, looking at the soft biometric analysis from the pilot study, we can see that we were able to predict handedness with a high degree of accuracy (87.64%), as well as age from a selection of six distinct categories to 35.02% and finally gender with a 62.34%

accuracy. These are all better than random chance, meaning that we have a positive prediction accuracy which is higher than just guessing. The accuracy scores were much better than random chance, specifically across handedness and in the case of age, due to the categorisation being split into six categories. A high accuracy score means that we have a better chance at predicting the biometrics and therefore achieving the aims of the research. These were all calculated using dwell time for each letter to establish the speed at which the user typed the letter to produce the results.

8.1.2 Research Question 2

What effect does the inclusion of accelerometer and gyroscopic data alongside keystroke dynamics have on the ability to successfully infer a person's name and soft biometric features on a mobile device?

Looking at the second question posed, which corresponded to the first main study that was completed, we looked at the inclusion of accelerometer and gyroscopic data, as well as location and size of press on the screen to enhance the accuracy and analysis. From this, we then swapped to a more automated approach utilising machine learning instead of the fastest keystrokes analysis, as we had a much larger set of data to analyse. Looking at a comparison of soft biometric features however, we can see a marked improvement over that of the pilot study with just using dwell time. For handedness, we produced a highest accuracy score of 91.90%, for age, 69.86% and finally for gender, 80.74%. These obviously

all being better, quite significantly, than random chance. When it comes to the name inference via the machine learning, we were able to produce a 35.50% accuracy of predicting a user's name from the dwell time, combined with accelerometer and gyroscopic data. This score, whilst good from a total of 41 unique names over 68 pieces of data, we felt could still be improved on.

8.1.3 Research Question 3

To what extent does the volume of data per user help to improve the accuracy of the prediction of name and soft biometric features?

Finally, looking at question three, we increased the data per user in order to provide the machine learning algorithms with a greater set of data to improve accuracy. From the analysis, looking at the total data obtained, we had 46 sets of data with a total of only 12 unique names. This increase in data per user, resulted in a much higher accuracy than just the diluted data from experiment two. For soft biometrics, we achieved a 93.18% for handedness, 87.74% for age over six categories, and finally 82.56% for gender. On the name inference side, again utilising the same dwell time and accelerometer and gyroscopic data as the first main study to keep the analysis similar, we achieved an 83.20% accuracy at predicting name. Looking at the improvement in accuracy of 134.37% from the first to the second main study we can confidently say that having higher volumes of data per user significantly impacts the scores in a positive way.

8.1.4 Discussion

As can be seen from the discussion above, we have managed to answer and prove, each question individually as we set out to do. For question one, looking at ‘to what extent can keystroke dynamics be utilised in order to infer a person’s name on a mobile device?’ We can confidently say that to a very high extent, 97.44% of users have keystrokes in their top 20% of fastest keystrokes that are also in their name with a high coverage threshold as can be seen in the results. It is therefore feasible to say that with a list of the most common names, such as from census data, and the analysis of the fastest keystrokes from people that we could recreate the experiment and infer names successfully. Looking onto question two, looking at ‘what effect does the utilisation of accelerometer and gyroscopic data in place of keystroke dynamics have on the ability to successfully infer a person’s name on a mobile device?’. We can see that we are able to significantly increase the accuracy of our predictions, from looking at which users have letters in their fastest keystrokes in their name, to actually being able to predict the user’s name with a 35.50% accuracy is a notable advancement. When looking at the soft biometrics for experiment two, we can see that we have managed to significantly improve our accuracy results over handedness, age and gender. Finally, when analysing question three, ‘to what extent does the inclusion of larger amounts of data per user help to improve the accuracy of the prediction of name and soft biometric features?’. We can actually understand that it has improved the accuracy a vast amount across all predictions, both name and soft biometric. For name, we have an ac-

curacy of 83.20% which is over double that of the previous experiment, due to a reduction in participants and an increase in the quantity of data per user. When looking at soft biometrics, we can also see that we have similar scores on handedness, but significantly across age and gender. As noted above, the 134.37% improvement in accuracy score for inferring name between the first and second main studies confirms the research hypothesis.

8.2 Comparison of Experiments

Due to the disparity between the pilot study and main experiments, we will compare these in two different ways. Firstly, the soft biometric features accuracy scores will be compared across all experiments including the combination of experiment two and three data. Following this, we will compare the name prediction for experiments two, three and the combined data as well to analyse this.

8.2.1 Soft Biometrics

As discussed, below in Table 8.1 we can see the top accuracy scores for each experiment, across handedness, age and gender.

Table 8.1: Highest accuracy scores per experiment across soft biometrics.

Experiment	Accuracy Scores		
	Handedness	Age	Gender
Pilot Study	87.64%	35.02%	62.34%
Main Experiment 1	91.90%	69.86%	80.74%
Main Experiment 2	93.18%	87.74%	82.56%
Combined Main Experiments	90.76%	73.44%	77.44%

Looking at the results above, we can see that inline with our research question, that providing a greater amount of data with fewer participants did in fact provide us with a much higher accuracy overall. With regards to the combination data, similar accuracy's were kept across handedness and gender. However, age managed to produce much better results than that of experiment one data in isolation. We can comfortably say that after analysis, main experiment two in isolation produced the best results across the board due to the increased amount of data per participant, with a lesser ratio of smaller data sets. This also confirms that the inclusion of the movement data from the mobile device, consistently adds a higher level of accuracy to the prediction of all features.

8.2.2 Keystroke Dynamics

In order to accurately compare the keystroke dynamics results, below in Table 8.2 we can see the best accuracy scores for each experiment, main study 1 and 2, as well as the combination findings.

Table 8.2: Highest accuracy scores per experiment across name.

Experiment	Accuracy Scores
	Name
Main Experiment 1	35.50%
Main Experiment 2	83.20%
Combined Main Experiments	49.72%

Looking at the table above, we can see that again, main experiment two was by far the best, due to the smaller data set and increased data per user. We can however, see that overall we increased the results of the first experiments accuracy by just over 14% with the combined study. This goes a long way to answer question three from our research objectives, which focused on the addition of more data sets per user to improve accuracy and this is quite clearly shown.

8.3 Discussion of Practical Applications

Whilst the research presented in this thesis is theoretical in the current form, there are some practical applications that can be applied based on the research as it stands. These have been detailed below to explain in more detail, however most of the desired practical applications are limited due to the computational power required to run machine learning classifiers, especially on a mobile device.

- *Auditing Shared Devices.* Based on the current state of running the machine learning based on data collected on device, we could use this for auditing on a shared device. For example, if a shop utilised

a mobile device for daily usage and certain actions were needed to be audited, the methods presented could be used to determine who was using the device, and narrow this down with various biometric identifiers also.

- *Lost Device Finding.* Technologies exist that allow for a device to be located with a fairly high degree of accuracy if lost via GPS (Global Positioning System) location services. If this device was still active, we could again use the methods presented as they stand to narrow down who the person who could have the mobile device could be. This would hopefully improve accuracy of location; if for example we knew a mobile device was in a specific location, we could also narrow down that there was a male of age 18-25 who is left handed has the device.

8.4 Discussion of Contributions

The first contribution we discussed was the novel approach to identity data using smartphone motion sensors. As can be seen in the main experiments, we were able to consistently predict a user's name, age, gender and handedness with significantly high levels of accuracy. We were able to do this with just getting a user to type three sentences, and from this, predicted name with an accuracy of 83.20%, age with an accuracy of 87.74%, gender with an accuracy of 82.56% and finally handedness with an accuracy of 93.18%. This novel approach has provided accuracy scores for inference that are higher than any found in existing studies from relevant literature.

As part of this experimentation we produced a new methodology and experimental approach including a new data capture framework as part of the research. Subsequently, we also created a bespoke data set of motion data that can be further anonymised and shared with the wider community in order to advance understanding around keystroke dynamics and inference on mobile devices.

Finally, we presented a novel algorithm for predicting a letter using various sensor inputs. This has been specifically detailed in Algorithm 2 which can be seen in a previous chapter. Whilst this did not produce the level of accuracy in results that we would like, this is a novel algorithm which could potentially show promise in future iterations.

8.5 Limitations, Future Work and Research Opportunities

Below we discuss limitations around the research to highlight challenges and areas for improvement. Following this we then look to future work and research opportunities that the author would like to pursue.

8.5.1 Limitations

Unfortunately there were some limitations in the research which restricted the overall efficiency and these are something that would ideally be worked on in future, but which would require significant changes to the research.

Firstly, the main limitation was that only Android devices were utilised

and not iOS which cut down the number of participants significantly. Ideally, a cross-device platform would be developed, supporting not only Android but also iOS devices. Unfortunately this is unlikely to be possible for a native application due to the restrictions Apple put in place for their operating system, which is why relevant studies all utilise Android. This may be possible utilising a web based application, but further testing would be needed.

Secondly, another key limitation is that the keyboard utilised was actually a keyboard made of buttons instead of a custom keyboard. Custom keyboards and other various methods were attempted to capture the keystrokes and other data however only buttons allowed the capture of the timings needed for the research. In order for the keyboard to be utilised across other applications on the device, a significant change would have to be made and alternative options would need to be explored.

Finally, the last key limitation of the research was around the machine learning. Due to the vast computational power required to run the machine learning, it was not possible to perform on device. This was certainly a limitation as it meant that we could only provide reactive identification that was off-device. To improve this in the future, cloud based machine learning could be utilised to provide a solution which is closer to real time.

8.5.2 Future Work and Research Opportunities

Looking forward to potential future work following this research, there are a number of promising avenues that this author would like to explore, and have been detailed below.

Keystroke Dynamics and Name Inference

As can be seen from the analysis of the pilot study, there is a high probability that the top 20% of fastest keystrokes produced by a user, also contain letters from their name, as surmised, this is due to the frequency they type their own name. In order to expand this avenue of research, looking at a user's partner's name would be an interesting expansion of this, as theoretically they type this even more than they do their own name. This could also be expanded to potentially children or even pets, as a number of pieces of private information, could be discerned from this, which could be utilised in password cracking attempts. This would allow the research to expand significantly and would be extremely interesting to understand what intelligence can be obtained around a person just from the way they type. It could also theoretically be applied to numbers as well to look at identifying PIN or even credit/debit card numbers.

Bi-grams and Tri-grams

Whilst uni-grams were chosen for this particular research as it presented a novel method of inference utilising keystroke dynamics and sensor data, bi-grams and tri-grams would be looked into in future iterations of the research to hopefully present even higher accuracy scores. Whilst bi-grams have been successful in a large proportion of research papers, a small number of papers look into tri-gram usage, which could potentially be a significant boost in accuracy. However, this hypothesis would need to be confirmed through additional research.

Inclusion of Further Keystroke Dynamic Metrics

Whilst Dwell time was chosen for this specific research due to the performance across related literature, there are a number of additional measurements which could be utilised to potentially deliver even better results than those achieved. In Figure 2.9 these various measurements can be seen which could be employed such as interval, latency, flight time and up-to-up time. The author will look to expand on this thesis research in future papers to include these additional measurements.

References

- [1] H. Alhakami and S. Alhrbi. Knowledge based authentication techniques and challenges. *International Journal of Advanced Computer Science and Applications*, 11(2):727–732, 2020.
- [2] E. Alpaydin. *Machine learning*. MIT press, 2021.
- [3] I. M. Alsaadi. Study on most popular behavioral biometrics, advantages, disadvantages and recent applications: A review. *Int. J. Sci. Technol. Res*, 10(1):15–21, 2021.
- [4] M. Antal, L. Z. Szabó, and I. László. Keystroke dynamics on android platform. *Procedia Technology*, 19:820–826, 2015.
- [5] AWS. Class `transferutility.builder`. <https://aws-amplify.github.io/aws-sdk-android/docs/reference/com/amazonaws/mobileconnectors/s3/transferutility/TransferUtility.Builder.html>, 2018.
- [6] L. Bedogni, A. Alcaras, and L. Bononi. Permission-free keylogging through touch events eavesdropping on mobile devices. In *2019 IEEE International Conference on Pervasive Computing and Com-*

-
- munications Workshops (PerCom Workshops)*, pages 28–33. IEEE, 2019.
- [7] C. Bo, L. Zhang, T. Jung, J. Han, X.-Y. Li, and Y. Wang. Continuous user identification via touch and movement behavioral biometrics. In *2014 IEEE 33rd International Performance Computing and Communications Conference (IPCCC)*, pages 1–8. IEEE, 2014.
- [8] O. Buckley, D. Hodges, M. Hadgkiss, and S. Morris. Keystroke inference using smartphone kinematics. In *Human Aspects of Information Security, Privacy and Trust: 5th International Conference, HAS 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings 5*, pages 226–238. Springer, 2017.
- [9] O. Buckley, D. Hodges, J. Windle, and S. Earl. Clicka: Collecting and leveraging identity cues with keystroke dynamics. *Computers & Security*, 120:102780, 2022.
- [10] L. Cai and H. Chen. Touchlogger: Inferring keystrokes on touch screen from smartphone motion. *HotSec*, 11(2011):9, 2011.
- [11] L. Cai and H. Chen. On the practicality of motion based keystroke inference attack. In *International Conference on Trust and Trustworthy Computing*, pages 273–290. Springer, 2012.
- [12] L. Cascone, M. Nappi, F. Narducci, and C. Pero. Touch keystroke dynamics for demographic classification. *Pattern Recognition Letters*, 158:63–70, 2022.

-
- [13] C.-K. Chan and L.-M. Cheng. Cryptanalysis of a remote user authentication scheme using smart cards. *IEEE Transactions on Consumer Electronics*, 46(4):992–993, 2000.
- [14] D. Clark. Population of the united kingdom from 1953 to 2022, by gender. <https://www.statista.com/statistics/281240/population-of-the-united-kingdom-uk-by-gender/#:~:text=In%202022%20the%20population%20of,increased%20by%20approximately%208.2%20million.,2024>.
- [15] N. L. Clarke and S. M. Furnell. Authentication of users on mobile telephones—a survey of attitudes and practices. *Computers & Security*, 24(7):519–527, 2005.
- [16] H. Crawford and E. Ahmadzadeh. Authentication on the go: Assessing the effect of movement on mobile device keystroke dynamics. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 163–173, 2017.
- [17] S. Earl, J. Campbell, and O. Buckley. Identifying soft biometric features from a combination of keystroke and mouse dynamics. In *International Conference on Applied Human Factors and Ergonomics*, pages 184–190. Springer, 2021.
- [18] S. Earl, J. Campbell, and O. Buckley. Investigating what you share: Privacy perceptions of behavioural biometrics. In *International Conference on Human-Computer Interaction*, pages 408–415. Springer, 2021.

-
- [19] P. M. Fitts and M. I. Posner. *Human performance*. Brooks/Cole, 1967.
- [20] L. A. Gabralla. Dense deep neural network architecture for keystroke dynamics authentication in mobile phone. *Adv. Sci. Technol. Eng. Syst. J*, 5(6):307–314, 2020.
- [21] R. Giot and C. Rosenberger. A new soft biometric approach for keystroke dynamics based on gender recognition. *International Journal of Information Technology and Management*, 11(1-2):35–49, 2012.
- [22] R. Giot, M. El-Abed, and C. Rosenberger. Keystroke dynamics overview. In *Biometrics*. IntechOpen, 2011.
- [23] C. Giuffrida, K. Majdanik, M. Conti, and H. Bos. I sensed it was you: authenticating mobile users with sensor-enhanced keystroke dynamics. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pages 92–111. Springer, 2014.
- [24] Google. Device orientation and motion. <https://developers.google.com/web/fundamentals/native-hardware/device-orientation/>, 2019.
- [25] Google. Classification: True vs false and positive vs negative. <https://developers.google.com/machine-learning/crash-course/classification/true-false-positive-negative>, 2020.

-
- [26] Google. Firebase in-app messaging. <https://firebase.google.com/docs/in-app-messaging>, 2023.
- [27] C. Griffiths. The latest 2024 cyber crime statistics. <https://aag-it.com/the-latest-cyber-crime-statistics/#:~:text=Since%202001%2C%20the%20victim%20count,standing%20in%202021%20at%20%24787%2C671.>, 2024.
- [28] J. Gurary, Y. Zhu, N. Alnahash, and H. Fu. Implicit authentication for mobile devices using typing behavior. In *Human Aspects of Information Security, Privacy, and Trust: 4th International Conference, HAS 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings 4*, pages 25–36. Springer, 2016.
- [29] D. Hodges and O. Buckley. Reconstructing what you said: Text inference using smartphone motion. *IEEE Transactions on Mobile Computing*, 18(4):947–959, 2018.
- [30] J. H. Huh, S. Kwag, I. Kim, A. Popov, Y. Park, G. Cho, J. Lee, H. Kim, and C.-H. Lee. On the long-term effects of continuous keystroke authentication: Keeping user frustration low through behavior adaptation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 7(2):1–32, 2023.
- [31] M.-S. Hwang and L.-H. Li. A new remote user authentication scheme using smart cards. *IEEE Transactions on consumer Electronics*, 46(1):28–30, 2000.

-
- [32] S.-s. Hwang, S. Cho, and S. Park. Keystroke dynamics-based authentication for mobile devices. *Computers & Security*, 28(1-2):85–93, 2009.
- [33] IBM. Identification and authentication. https://www.ibm.com/support/knowledgecenter/en/SSFKSJ_7.5.0/com.ibm.mq.sec.doc/q009740_.htm, 2019.
- [34] A. K. Jain, R. Bolle, and S. Pankanti. *Biometrics: personal identification in networked society*, volume 479. Springer Science & Business Media, 2006.
- [35] A. R. Javed, S. U. Rehman, M. U. Khan, M. Alazab, and H. U. Khan. Betalogger: Smartphone sensor-based side-channel attack detection and text inference using language modeling and dense multilayer neural network. *Transactions on Asian and Low-Resource Language Information Processing*, 20(5):1–17, 2021.
- [36] M. Karnan and N. Krishnaraj. A model to secure mobile devices using keystroke dynamics through soft computing techniques. *International Journal of Soft Computing and Engineering (IJSCCE) ISSN*, pages 2231–2307, 2012.
- [37] F. Laricchia. Number of smartphones sold to end users worldwide from 2007 to 2022. <https://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007/>, 2023.
- [38] C.-T. Li and M.-S. Hwang. An efficient biometrics-based remote

-
- user authentication scheme using smart cards. *Journal of Network and computer applications*, 33(1):1–5, 2010.
- [39] E. Maiorana, P. Campisi, N. González-Carballo, and A. Neri. Keystroke dynamics authentication for mobile phones. In *Proceedings of the 2011 ACM Symposium on Applied Computing*, pages 21–26, 2011.
- [40] Mathworks. Accelerometer. <https://uk.mathworks.com/help/supportpkg/android/ref/accelerometer.html>, 2018.
- [41] Mathworks. Gyroscope. [https://uk.mathworks.com/help/supportpkg/android/ref/gyroscope.](https://uk.mathworks.com/help/supportpkg/android/ref/gyroscope..), 2018.
- [42] R. Moskovitch, C. Feher, A. Messerman, N. Kirschnick, T. Mustafic, A. Camtepe, B. Lohlein, U. Heister, S. Moller, L. Rokach, et al. Identity theft, computers and behavioral biometrics. In *2009 IEEE International Conference on Intelligence and Security Informatics*, pages 155–160. IEEE, 2009.
- [43] L. O’Gorman. Comparing passwords, tokens, and biometrics for user authentication. *Proceedings of the IEEE*, 91(12):2021–2040, 2003.
- [44] E. Owusu, J. Han, S. Das, A. Perrig, and J. Zhang. Accessory: password inference using accelerometers on smartphones. In *Proceedings of the Twelfth Workshop on Mobile Computing Systems & Applications*, page 9. ACM, 2012.
- [45] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al.

-
- Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.
- [46] A. Pentel. Predicting user age by keystroke dynamics. In *Artificial Intelligence and Algorithms in Intelligent Systems: Proceedings of 7th Computer Science On-line Conference 2018, Volume 2* 7, pages 336–343. Springer, 2019.
- [47] V. Ponnusamy, W. C. Hong, A. Yichiet, R. Annur, and G. M. Lee. Keystroke dynamics in mobile platform. In *Proceedings of the 2019 2nd International Conference on E-Business, Information Management and Computer Science*, pages 1–8, 2019.
- [48] A. P. Rebera, M. E. Bonfanti, and S. Venier. Societal and ethical implications of anti-spoofing technologies in biometrics. *Science and engineering ethics*, 20(1):155–169, 2014.
- [49] K. Revett, F. Gorunescu, M. Gorunescu, M. Ene, P. S. T. Magalhães, and H. D. d. Santos. A machine learning approach to keystroke dynamics based user authentication. *International Journal of Electronic Security and Digital Forensics*, 2007.
- [50] D. Sahu and D. S. Tomar. End user identification through proactive techniques. In *International Conference on Information Science (ICIS)*, pages 234–238, 2014.
- [51] S. A. Schuckers. Spoofing and anti-spoofing measures. *Information Security technical report*, 7(4):56–62, 2002.

-
- [52] P. Sedgwick and N. Greenwood. Understanding the hawthorne effect. *Bmj*, 351, 2015.
- [53] C. Shen, H. Xu, H. Wang, and X. Guan. Handedness recognition through keystroke-typing behavior in computer forensics analysis. In *2016 IEEE Trustcom/BigDataSE/ISPA*, pages 1054–1060. IEEE, 2016.
- [54] S. Singh. *The code book*, volume 7. Doubleday New York, 1999.
- [55] C. Soanes and A. Stevenson. *Concise oxford English dictionary*, volume 11. Oxford University Press Oxford, 2004.
- [56] D. Specifications. Device specifications. <https://www.devicespecifications.com/en>, 2024.
- [57] Statista. Smartphone sales globally 2007-2022. <https://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007/>, 2023.
- [58] K. R. Subramanian. Myth and mystery of shrinking attention span. *International Journal of Trend in Research and Development*, 5(1): 1–6, 2018.
- [59] P. S. Teh, A. B. J. Teoh, and S. Yue. A survey of keystroke dynamics biometrics. *The Scientific World Journal*, 2013.
- [60] R. F. Tinder. Relativistic flight mechanics and space travel. *Synthesis lectures on engineering*, 1(1):1–140, 2006.
- [61] M. Trojahn and F. Ortmeier. Toward mobile authentication with keystroke dynamics on mobile phones and tablets. In *2013 27th*
-

International Conference on Advanced Information Networking and Applications Workshops, pages 697–702. IEEE, 2013.

- [62] P. Vaishnav, M. Kaushik, and L. Raja. Design an algorithm for continuous authentication on smartphone through keystroke dynamics and touch dynamics. *Indian Journal of Computer Science and Engineering*, 13(2):444–455, 2022.
- [63] P. Voigt and A. Von dem Bussche. The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 2017.
- [64] W3C. Device orientation and motion. <https://www.w3.org/TR/orientation-event/#def-deviceorientation>, 2024.
- [65] R. V. Yampolskiy and V. Govindaraju. Taxonomy of behavioural biometrics. In *Behavioral Biometrics for Human Identification: Intelligent Applications*, pages 1–43. IGI Global, 2010.
- [66] J. Yan, A. Blackwell, R. Anderson, and A. Grant. Password memorability and security: Empirical results. *IEEE Security & privacy*, 2(5):25–31, 2004.
- [67] A. S. Yuksel, F. A. Senel, and I. A. Cankaya. Classification of soft keyboard typing behaviors using mobile device sensors with machine learning. *Arabian Journal for Science and Engineering*, 44:3929–3942, 2019.
- [68] Q. Zheng, L. Dong, D. H. Lee, and Z. Gao. Active disturbance

rejection control for mems gyroscopes. In *2008 American control conference*, pages 4425–4430. IEEE, 2008.

Appendices

Appendix 1: Table showing the reviewed papers for the relevant literature as per the searches completed in Table 2.1.

Paper Title	Author	Suitable?	Narrative
I sensed it was you: authenticating mobile users with sensor-enhanced keystroke dynamics	C Giuffrida, K Majdanik, M Conti, H Bos	Yes	Abstract suggests that this is suitable for further review.
Reconstructing what you said: Text inference using smartphone motion	D Hodges, O Buckley	Yes	Abstract suggests that this is suitable for further review.
On the practicality of motion based keystroke inference attack	L Cai, H Chen	Yes	Abstract suggests that this is suitable for further review.
Keystroke inference using smartphone kinematics	O Buckley, D Hodges, M Hadgkiss, S Morris	Yes	Abstract suggests that this is suitable for further review.

Keystroke Dynamics in Mobile Platform	V Ponnusamy, WC Hong, A Yichiet, R Annur	Yes	Abstract suggests that this is suitable for further review.
Classification of soft keyboard typing behaviors using Mobile device sensors with machine learning	AS Yuksel, FA Senel, IA Cankaya	Yes	Abstract suggests that this is suitable for further review.
Betalogger: Smartphone sensor-based side-channel attack detection and text inference using language modeling and dense multi-layer neural network	AR Javed, SU Rehman, MU Khan, M Alazab	Yes	Abstract suggests that this is suitable for further review.
Design An Algorithm For Continuous Authentication On Smartphone Through Keystroke Dynamics And Touch Dynamics	P Vaishnav, M Kaushik, L Raja	Yes	Abstract suggests that this is suitable for further review.

On the Long-Term Effects of Continuous Keystroke Authentication: Keeping User Frustration Low through Behavior Adaptation	JH Huh, S Kwag, I Kim, A Popov, Y Park	Yes	Abstract suggests that this is suitable for further review.
Permission-free keylogging through touch events eavesdropping on mobile devices	L Bedogni, A Alcaras, L Bononi	Yes	Abstract suggests that this is suitable for further review.
Continuous user identification via touch and movement behavioral biometrics	C Bo, L Zhang, T Jung, J Han, XY Li	Yes	Abstract suggests that this is suitable for further review.
Implicit authentication for mobile devices using typing behavior	J Gurary, Y Zhu, N Alnash, H Fu	Yes	Abstract suggests that this is suitable for further review.

Touch keystroke dynamics for demographic classification	L Cascone, M Nappi, F Narducci, C Pero	Yes	Abstract suggests that this is suitable for further review.
Dense Deep Neural Network Architecture for Keystroke Dynamics Authentication in Mobile Phone	LA Gabralla	Yes	Abstract suggests that this is suitable for further review.
Keystroke biometric system for touch screen text input on android devices optimization of equal error rate based on medians vector proximity	P Gautam, PR Dawadi	No	Abstract suggests that this is not suitable for further review.
Inferring user profile attributes from multi-dimensional mobile phone sensory data	Z Yu, E Xu, H Du, B Guo, L Yao	No	Abstract suggests that this is not suitable for further review.

I know what you type on your phone: Keystroke inference on android device using deep learning	L Bo, L Fengjun, W Guanghui	No	Abstract suggests that this is not suitable for further review.
Authentication on the go: Assessing the effect of movement on mobile device keystroke dynamics	H Crawford, E Ah-madzadeh	No	Abstract suggests that this is not suitable for further review.
Mobile keystroke dynamics for biometric recognition: An overview	E Maiorana, H Kalita, P Campisi	No	Abstract suggests that this is not suitable for further review.
Smartwatch-based keystroke inference attacks and context-aware protection mechanisms	A Maiti, O Armbruster, M Jadliwala, J He	No	Abstract suggests that this is not suitable for further review.

Alphallogger: Detecting motion-based side-channel attack using smartphone keystrokes	AR Javed, MO Beg, M Asim, T Baker	No	Abstract suggests that this is not suitable for further review.
Detecting Mobility Context over Smartphones using Typing and Smartphone Engagement Patterns	S Chatterjee, A Bhowmik, A Singh	No	Abstract suggests that this is not suitable for further review.
The rise of keyloggers on smartphones: A survey and insight into motion-based tap inference attacks	M Hussain, A Al-Haiqi, AA Zaidan, BB Zaidan	No	Abstract suggests that this is not suitable for further review.
Understanding keystroke dynamics for smartphone users authentication and keystroke dynamics on smartphones built-in motion sensors	H Lee, JY Hwang, DI Kim, S Lee, SH Lee	No	Abstract suggests that this is not suitable for further review.

Enhanced free-text keystroke continuous authentication based on dynamics of wrist motion	B Li, H Sun, Y Gao, VV Phoha	No	Abstract suggests that this is not suitable for further review.
A parameterized model to select discriminating features on keystroke dynamics authentication on smartphones	H Lee, JY Hwang, S Lee, DI Kim, SH Lee, J Lee	No	Abstract suggests that this is not suitable for further review.
When good becomes evil: Keystroke inference with smartwatch	X Liu, Z Zhou, W Diao, Z Li, K Zhang	No	Abstract suggests that this is not suitable for further review.
Activity Context Detection during Smartphone Keyboard Interactions: A Machine Learning Approach	K Sakellariou, G Licitra, E Ferrante, J Tomczak	No	Abstract suggests that this is not suitable for further review.

Practicality of accelerometer side channels on smartphones	AJ Aviv, B Sapp, M Blaze, JM Smith	No	Abstract suggests that this is not suitable for further review.
Keysens: Passive user authentication through micro-behavior modeling of soft keyboard interaction	B Draffin, J Zhu, J Zhang	No	Abstract suggests that this is not suitable for further review.
Privacy implications of accelerometer data: a review of possible inferences	JL Kröger, P Raschke, TR Bhuiyan	No	Abstract suggests that this is not suitable for further review.
Your eyes reveal your secrets: An eye movement based password inference on smartphone	Y Wang, W Cai, T Gu, W Shao	No	Abstract suggests that this is not suitable for further review.

Using unrestricted mobile sensors to infer tapped and traced user inputs	T Nguyen	No	Abstract suggests that this is not suitable for further review.
Keypad entry inference with sensor fusion from mobile and smart wearables	Y Liu, UM Qureshi, GP Hancke	No	Abstract suggests that this is not suitable for further review.
Niffler: A context-aware and user-independent side-channel attack system for password inference	B Tang, Z Wang, R Wang, L Zhao, L Wang	No	Abstract suggests that this is not suitable for further review.
Actions speak louder than (pass) words: Passive authentication of smartphone users via deep temporal features	D Deb, A Ross, AK Jain	No	Abstract suggests that this is not suitable for further review.

KEYSTROKE CLASSIFICATION OF MOTION SENSOR DATA	M PATRICKS	No	Abstract suggests that this is not suitable for further review.
EyeteLL: Video-assisted touchscreen keystroke inference from eye movements	Y Chen, T Li, R Zhang, Y Zhang	No	Abstract suggests that this is not suitable for further review.
Visible: Video-assisted keystroke inference from tablet backside motion.	J Sun, X Jin, Y Chen, J Zhang, Y Zhang, R Zhang	No	Abstract suggests that this is not suitable for further review.
Toward mobile authentication with keystroke dynamics on mobile phones and tablets	M Trojahn, F Ortmeier	No	Abstract suggests that this is not suitable for further review.

Secure Keyboards Against Motion Based Keystroke Inference Attack	S Du, Y Gao, J Hua, S Zhong	No	Abstract suggests that this is not suitable for further review.
Keylistener: Inferring keystrokes on qwerty keyboard of touch screen through acoustic signals	L Lu, J Yu, Y Chen, Y Zhu, X Xu	No	Abstract suggests that this is not suitable for further review.
Analyzing the effectiveness of touch keystroke dynamic authentication for the Arabic language	SA Alsuhibany, AS Almuqbil	No	Abstract suggests that this is not suitable for further review.
A Review on Smartphone Keystroke Dynamics as a Digital Biomarker for Understanding Neurocognitive Functioning	TM Nguyen, AD Leow, O Ajilore	No	Abstract suggests that this is not suitable for further review.

Continuous user authentication on smartphone via behavioral biometrics: a survey	PK Rayani, S Changder	No	Abstract suggests that this is not suitable for further review.
Touch-dynamics based behavioural biometrics on mobile devices—a review from a usability and performance perspective	E Ellavarason, R Guest, F Deravi	No	Abstract suggests that this is not suitable for further review.
Real-time smartphone activity classification using inertial sensors—recognition of scrolling, typing, and watching videos while sitting or walking	S Zhuo, L Sherlock, G Dobbie, YS Koh, G Russello	No	Abstract suggests that this is not suitable for further review.
User authentication on mobile devices: Approaches, threats and trends	C Wang, Y Wang, Y Chen, H Liu, J Liu	No	Abstract suggests that this is not suitable for further review.

Introducing touchstroke: keystroke-based authentication system for smartphones	G Kambourakis, D Damopoulos	No	Abstract suggests that this is not suitable for further review.
Android based Access Control Systems using Sensory-Data	R Saranya, C Sownthararajan, R Suriya	No	Abstract suggests that this is not suitable for further review.
Mimicry attacks on smartphone keystroke authentication	H Khan, U Hengartner, D Vogel	No	Abstract suggests that this is not suitable for further review.
Meta-heuristic optimization and keystroke dynamics for authentication of smartphone users	ESM El-Kenawy, S Mirjalili, AA Abdelhamid, A Ibrahim	No	Abstract suggests that this is not suitable for further review.

Learning-based Practical Smartphone Eavesdropping with Built-in Accelerometer.	Z Ba, T Zheng, X Zhang, Z Qin, B Li, X Liu, K Ren	No	Abstract suggests that this is not suitable for further review.
A new feature scoring method in keystroke dynamics-based user authentications	DI Kim, S Lee, JS Shin	No	Abstract suggests that this is not suitable for further review.
Smartphone user authentication using touch dynamics in the big data era: Challenges and opportunities	L Jiang, W Meng	No	Abstract suggests that this is not suitable for further review.
Spearphone: A speech privacy exploit via accelerometer-sensed reverberations from smartphone loudspeakers	SA Anand, C Wang, J Liu, N Saxena	No	Abstract suggests that this is not suitable for further review.

Adaptive human-machine interactive behavior analysis with wrist-worn devices for password inference	C Shen, Y Chen, Y Liu, X Guan	No	Abstract suggests that this is not suitable for further review.
Keystroke dynamics-based authentication using unique keypad	M Choi, S Lee, M Jo, JS Shin	No	Abstract suggests that this is not suitable for further review.
Passive sensing of affective and cognitive functioning in mood disorders by analyzing keystroke kinematics and speech dynamics	F Hussain, JP Stange, SA Langenecker	No	Abstract suggests that this is not suitable for further review.
Motion Sensor-based Privacy Attack on Smartphones	SA Anand, C Wang, J Liu, N Saxena	No	Abstract suggests that this is not suitable for further review.

Touchstroke: Smartphone user authentication based on touch-typing biometrics	A Buriro, B Crispo, F Del Frari, K Wrona	No	Abstract suggests that this is not suitable for further review.
Passive sensing of affective and cognitive functioning in mood disorders by Analyzing keystroke kinematics and speech dynamics	F Hussain, JP Stange, SA Langenecker	No	Abstract suggests that this is not suitable for further review.
Active authentication on mobile devices via stylometry, application usage, web browsing, and GPS location	L Fridman, S Weber, R Greenstadt	No	Abstract suggests that this is not suitable for further review.
Inferring Smartphone Users' Handwritten Patterns by using Motion Sensors.	WH Lee, J Ortiz, B Ko, RB Lee	No	Abstract suggests that this is not suitable for further review.

Pin skimmer: Inferring pins through the camera and microphone	L Simon, R Anderson	No	Abstract suggests that this is not suitable for further review.
Affective state prediction from smartphone touch and sensor data in the wild	R Wampfler, S Klingler, B Solenthaler	No	Abstract suggests that this is not suitable for further review.
Associations between smartphone keystroke dynamics and cognition in MS	MH Chen, A Leow, MK Ross, J DeLuca	No	Abstract suggests that this is not suitable for further review.
Reran: Timing-and touch-sensitive record and replay for android	L Gomez, I Neamtiu, T Azim	No	Abstract suggests that this is not suitable for further review.

Contextual Authentication: Using Mobile Phone Movements to Authenticate Owners Implicitly	Y Badin	No	Abstract suggests that this is not suitable for further review.
GazeRevealer: Inferring password using smartphone front camera	Y Wang, W Cai, T Gu, W Shao, I Khalil	No	Abstract suggests that this is not suitable for further review.
Towards device independent eavesdropping on telephone conversations with built-in accelerometer	W Su, D Liu, T Zhang, H Jiang	No	Abstract suggests that this is not suitable for further review.
TapSnoop: Leveraging tap sounds to infer tapstrokes on touchscreen devices	H Kim, B Joe, Y Liu	No	Abstract suggests that this is not suitable for further review.

Accelword: Energy efficient hotword detection through accelerometer	L Zhang, PH Pathak, M Wu, Y Zhao	No	Abstract suggests that this is not suitable for further review.
Improving reliability: User authentication on smartphones using keystroke biometrics	Y Wang, C Wu, K Zheng, X Wang	No	Abstract suggests that this is not suitable for further review.
Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey	M Abuhamad, A Abusnaina, DH Nyang	No	Abstract suggests that this is not suitable for further review.
A survey of keystroke dynamics biometrics	PS Teh, ABJ Teoh, S Yue	No	Abstract suggests that this is not suitable for further review.

CASTRA: Seamless and unobtrusive authentication of users to diverse mobile services	DM Shila, K Srivastava	No	Abstract suggests that this is not suitable for further review.
Keystroke dynamics based recognition systems using deep learning: A survey	A Tewari	No	Abstract suggests that this is not suitable for further review.
Dialerauth: A motion-assisted touch-based smartphone user authentication scheme	A Buriro, B Crispo, S Gupta, F Del Frari	No	Abstract suggests that this is not suitable for further review.
Kollector: Detecting Fraudulent Activities on Mobile Devices Using Deep Learning	L Sun, B Cao, J Wang, W Srisa-an	No	Abstract suggests that this is not suitable for further review.

A report on personally identifiable sensor data from smartphone devices	M Fanourakis	No	Abstract suggests that this is not suitable for further review.
A novel word-independent gesture-typing continuous authentication scheme for mobile devices	M Smith-Creasey, M Rajarajan	No	Abstract suggests that this is not suitable for further review.
Machine learning techniques for implicit interaction using mobile sensors	MF Md Noor	No	Abstract suggests that this is not suitable for further review.
A Short Survey: Behavioral Authentication Using Mobile Sensors	ABA Ali, V Ponnusamy, A Sangodiah	No	Abstract suggests that this is not suitable for further review.

ResearchIME: A mobile keyboard application for studying free typing behaviour in the wild	D Buschek, B Bisinger, F Alt	No	Abstract suggests that this is not suitable for further review.
An implicit identity authentication system considering changes of gesture based on keystroke behaviors	J Wu, Z Chen	No	Abstract suggests that this is not suitable for further review.
Inferring Touch From Motion in Real World Data	P Bissig, P Brandes, J Passerini	No	Abstract suggests that this is not suitable for further review.
Tracking Mobile Web Users Through Motion Sensors: Attacks and Defenses.	A Das, N Borisov, M Caesar	No	Abstract suggests that this is not suitable for further review.

Touchsignatures: identification of user touch actions and PINs based on mobile sensor data via javascript	M Mehrnezhad, E Toreini, SF Shahandashti	No	Abstract suggests that this is not suitable for further review.
Exploring orientation and accelerometer sensor data for personal authentication in smartphones using touchscreen gestures	A Jain, V Kanhangad	No	Abstract suggests that this is not suitable for further review.
Going through the motions:{AR/VR} key-logging from user head motions	C Slocum, Y Zhang, N Abu-Ghazaleh	No	Abstract suggests that this is not suitable for further review.
An Accelerometer-based Privacy Attack on Smartphones.	R De Prisco, A De Santis, R Zaccagnino	No	Abstract suggests that this is not suitable for further review.

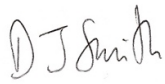
A survey on touch dynamics authentication in mobile devices	PS Teh, N Zhang, ABJ Teoh, K Chen	No	Abstract suggests that this is not suitable for further review.
A multi-faceted approach to user authentication for mobile devices—using human movement, usage, and location patterns	DM Shila, K Srivastava, P O’Neill	No	Abstract suggests that this is not suitable for further review.
I know what you type: Leaking user privacy via novel frequency-based side-channel attacks	R Song, Y Song, S Gao, B Xiao	No	Abstract suggests that this is not suitable for further review.
Modeling interactive sensor-behavior with smartphones for implicit and active user authentication	Y Chen, C Shen, Z Wang, T Yu	No	Abstract suggests that this is not suitable for further review.

Stationary mobile behavioral biometrics: A survey	A Ray-Dowling, D Hou, S Schuckers	No	Abstract suggests that this is not suitable for further review.
Fine-grained and context-aware behavioral biometrics for pattern lock on smartphones	D Shi, D Tao, J Wang, M Yao, Z Wang, H Chen	No	Abstract suggests that this is not suitable for further review.
Speechless: Analyzing the threat to speech privacy from smartphone motion sensors	SA Anand, N Saxena	No	Abstract suggests that this is not suitable for further review.
Stealing your android patterns via acoustic signals	M Zhou, Q Wang, J Yang, Q Li, P Jiang	No	Abstract suggests that this is not suitable for further review.

Appendix 2: Pilot Study ethical approval 09.03.2020. Ref: SCI-CMP/1920/R/17.

Research Ethics Application

1. Application summary

Applicant to complete			
Applicant Name	James Campbell		
Email	J.Campbell1@uea.ac.uk		
School	CMP	PGR	
Name of Supervisor / PI	Dr Oliver Buckley		
Short title of project	Identification and Authentication Utilising Keystroke Dynamic Inference		
Start date	October 2018	End date	October 2023
Outcome: for SCI-REC use only			
Lead Reviewer		Review date	9 Mar 2020
Decision	Approve	Approve after revision	Reject
Conditions (if any) or reasons for rejection	Approved subject to the participant information sheet email addresses being altered to the normal form and the institutional affiliation including "School of Computing Sciences"		
Approved by (Chair SCI-REC)	D J Smith	Signature	

2. Checklist

A. Does the research use an interview or questionnaire survey?	Yes
If so, does it ask for any personal information?	Yes
B. Does the research offer advice or guidance to people?	No
C. Does the research involve children, vulnerable adults or their carers?	No
D. Does the research record or observe people's behaviour?	No
E. Has this research been previously considered by another REC?	No
F. Does the research involve the analysis of personal data collected by others?	No
G. Will the researcher carry out fieldwork alone while away from UEA?	Yes
H. Will participants be paid or offered a reward for participating?	No
I. Does the project rely on data supplied from external sources?	No
I.1 Does the research use sensitive data? (commercially confidential, military, ...)	No
I.2 Does the research use confidential data? (e.g. medical records)	No
II.3 Is the research covered by the consent given when the data were collected?	No
J. Are special arrangements needed for the storage and retention of the data?	No
K. Will the research have an impact on the environment?	No
K.1. Will the research be carried out in an environmentally sensitive area, or SSSI?	No
K.2. Have appropriate steps been taken to gain permission to access the field site(s)?	No
K.3. Will samples be collected and removed in sufficient quantities to have a negative physical/environmental impact on the site and/or its ecosystem?	No
K.4. Does the fieldwork involve sampling rare/endangered or harmful taxa/species?	No
K.5. Will the fieldwork significantly disrupt the site and/or its environment?	No
K.6. Will the research involve transporting samples/specimens between countries or across other significant boundaries?	No

All applications must be approved by the Research Ethics Committee before beginning data generation or approaching potential research participants.

3. Application form

This form is for all staff and students who are planning research in areas of interest to the Schools of the UEA Faculty of Science, excluding animal, tissue and medical-related research which are covered by other committees. Any research involving NHS patients must be approved by the relevant NHS ethics committee.

SCI-REC has particular expertise in topics related to computing (data science, machine learning, cybersecurity, online privacy, ...), environmental impacts, and science education.

Please refer to the guidance notes at the end of this form when preparing your application as this can clarify what the committee needs to see about your project and can avoid any unnecessary requests for further information.

- **Briefly describe the scope and aims of your project in language understandable by a wide audience. (250 words maximum)**

The aim of this piece of research is to determine if a participant's name can be inferred from the way they type. This research will analyse the results achieved by asking the user to copy a pre-selected passage of text to ascertain which letters they type fastest, in the hopes of being able to identify their name.

In addition to the name of a participant, comparisons will be made around other demographic data such as (age, handedness and gender) to determine the impact this has on the correct prediction of a name.

Participants will be using a custom made Android Application on a Samsung Galaxy S8+ Smartphone to complete the experiment in it's entirety. The Android Application records the keystrokes (time up/time down) of a participant into a text file, along with some basic demographic questions which can be seen in the supporting materials.

This research is being conducted from the University of East Anglia (UEA) as part of my, (James Campbell - Researcher), PhD Thesis, which is supervised by Dr Oliver Buckley (UEA Lecturer). Please see the attached supporting materials for further clarification. As mentioned this is a pilot study for my PhD Thesis and, as such, the sample size will be a maximum of 100 participants.

- **Briefly outline the proposed research methods. (250 words maximum)**

The research method will be based around the interaction with the Android Mobile Phone Application, both with the questionnaire and experiment itself.

Upon loading the application, the participant will be presented with the Participant Information Statement, which they are required to have read and understood before proceeding. Following the PIS, is a short questionnaire in which the participant is required to enter their first name, and select from drop down fields for their age, gender and handedness. Once this questionnaire has been completed the experiment begins and the participant is presented with a passage of text and the on screen keyboard to use to copy the text into a field. Upon completion the participant presses a submit button (which saves the copied text to a .txt file on the phone's memory card containing the questionnaire answers and key timings). Finally a unique identifier (in the form of a random 12 digit number) is presented to the participant, with which they can withdraw from the research at any time.

Once the participant has completed the experiment, they will be reminded that they can withdraw at any time, and thanked for their participation. If a participant decides to end the experiment early then they will be free to leave at any time.

The participant data will then be moved onto a secured, encrypted hard drive so that it can be analysed at a later date.

- **Briefly explain how participants will be recruited. (250 words maximum)**

Any participants, as long as they are over the age of 18, will be recruited, and no demographic details will be considered in choosing participants who are suitable for the study. There will be a conscious effort however to ensure that there is a wide spread of participants from differing demographics.

Participants will be recruited via two main methods.

- Social Media
 - The researcher (James Campbell) will utilise the contacts on his social media accounts (Facebook + LinkedIn) to recruit friends, family and colleagues to participate in the study, after providing them with full Participant Information Statement up front. Before the experiment begins, the Participant Information Statement will be provided again to ensure the full understanding and agreement from the participant before undertaking any aspect of the experiment.
- Place of Work
 - The researcher (James Campbell) will also utilise emails within his place of work to recruit colleagues to participate in the study, after providing them with full Participant Information Statement. These colleagues will be contacted via internal email at the researcher's place of work, with full disclosure of the experiment before requesting participation from work colleagues. Before the experiment begins, the Participant Information Statement will be provided again to ensure the full understanding and agreement from the participant before undertaking any aspect of the experiment.

Please note – A maximum of 100 participants will be recruited for the study using the combination of recruitment methods above. Whilst the maximum is 100 participants, a more realistic figure of around 70-75 participants is expected, depending on the appetite for the research.

- **State who will have access to the data and what measures will be adopted to maintain participants' confidentiality or anonymity. (250 words maximum)**

The data will be accessible exclusively by the researcher on this project (James Campbell) as well as the researcher's Supervisor (Dr. Oliver Buckley), and any associated researchers within the Cyber Research Group within UEA. In order to protect anonymity, only first names are being recorded, in addition, participants will be identified by using a unique identifier (number) to reduce the likelihood of a correlation between participant and data. Furthermore, the analysis will be completed after the data collection period to reduce the likelihood that anonymity will be compromised.

If publication is to occur in the future, then participants will be protected via the use of amendments to ensure anonymity.

All of the results from the experiments and questionnaires will be secured on an encrypted hard drive, and the data on it will be kept in line with UEA data management policies and the DPA (Data Protection Act).

- **Give details of how the consent of participants will be (250 words maximum), e.g. through information sheets and consent forms, oral or other approach. Copies of all forms should be submitted with the application (do not include the text of these documents in this space).**

The consent of the participants will be gathered with a consent form on the landing page of the study, after reading the Participant Information Statement before the questionnaire or demographic information is gathered. This information and consent form can be seen in the attached supporting materials. In order to ensure the consent of participants is legitimate, the user of the application is not allowed to progress with the study until they have agreed to proceed. The form will also include details of the purpose of the study and the methods of withdrawal, as well as any further clarification or background information which will be provided orally before requesting participation.

Oral confirmation that the user is happy to proceed will also take place before handing the participant the device and again before the experiment begins.

- **If any payment or incentive will be made to any participant, please explain what it is, how anonymity will be maintained (given the need to record recipients' details for HMRC compliance) and provide the justification (no more than 250 words).**

N/A

- Please add here any other ethical considerations the ethics committee may need to be made aware of (no more than 250 words).
 - Are there any issues here for who can or cannot participate in the project?
 - If you are conducting research in a space where individuals may also choose not to participate, how will you ensure they will not be included in any data collection or adversely affected by non-participation?

There are no issues for those who can or cannot participate in the project as all participation is on a one to one basis and as such can be stopped at any time. If a participant wishes not to participate then the experiment (if started) will be stopped immediately and the data will be removed with the participant watching. Oral confirmation will be requested that the participant is happy that the data has been removed, and contact details will be provided in case of any follow up questions.

- What risks or costs to the participants are entailed in involvement in the research and how will these risks be managed?
 - Are there any potential physical, psychological or disclosure dangers that can be anticipated?
 - What is the possible harm to the participant or society from their participation?
 - What arrangements for the care and protection of participants have been made?

There are no risks or costs to the participants.

- Does this research have environmental implications? Please refer to the University's Research Ethics Guidance Note: [Research with a Potential Impact on the Environment](#) for further details. Identify any significant environmental impacts arising from your research/project and the measures you will take to minimise risk of adverse impact.

There are no environmental implications for this research, other than the electricity being used to charge the mobile device before beginning the experiment.

- Will your research involve investigation of or engagement with terrorist or violent extremist groups, people engaged in or supporting activities that compromise computer security or other activity that may normally be considered harmful or unlawful? Please provide a full explanation if the answer is 'yes'.

N/A

- **If any of the work is being carried out away from UEA, describe the arrangements been made to ensure the safety of the researchers.**

Outside of UEA, I will only be running the experiment in my usual working environment or home environment. Outside of this, a UEA member of staff will be provided with my location, how long I intend to spend gathering data, and emergency contact details. I will also check in periodically whilst working away from the above specific environments.

- **State any precautions being taken to protect the health and safety of other researchers and others associated with the project (as distinct from the participants or the applicant).**

N/A

- **If the research is being undertaken outside the UK:**
 - **Have any necessary research permissions, in-country ethical clearance, visas, ... been obtained?**
 - **Have appropriate steps been taken to identify and minimize any risks?**
 - **Are the researchers aware of health and safety advice, and other travel advisory notices?**
 - **If the researchers are travelling to conduct the research, have they taken out travel and health insurance for the full period of the research? If not, why not.**
 - **Has formal permission or a research permit been sought to conduct this research? If not, please explain why this is not necessary and/or appropriate**

N/A

- **Describe the steps that will be taken and the techniques that will be used to ensure that the anonymity and confidentiality of participants will be preserved, including the measures will be implemented to prevent de-anonymisation.**

In order to protect anonymity, the only piece of potentially identifiable demographic information which is being recorded is the first name. In addition, participants will be identified by using a unique identifier (number) to reduce the likelihood of a correlation between participant and data. Furthermore, the analysis will be completed after the data collection period to reduce the likelihood that anonymity will be compromised. For other pieces of data which are being recorded (age, handedness and gender) these will be recorded in the format of a drop down box, so that the answers are generic and cannot be used to identify a specific participant.

Applications with missing or incomplete sections will be returned to the applicant for resubmission.

Supporting documents	Tick to confirm
Participant Information and consent	Yes
Other supporting documents (e.g. questionnaires, interview/focus group questions, stimulus materials, observation checklists, letters of invitation, recruitment posters, etc.)	Yes

4. DECLARATION:

Please complete the following boxes with YES, NO, or NOT APPLICABLE:

<i>For student applicants:</i> I have discussed the planned work with my supervisor, who is in agreement with this application.	YES
I am aware of the relevant sections of the GDPR (2018): https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/ and Freedom of Information Act (2005).	YES
The purpose and procedures of the research, and the potential benefits and costs of participating (e.g. the amount of their time involved), will be explained to participants.	YES
It is clear to participants who is conducting the research (name, School and UEA email), why the research is being carried out (e.g. funded by UKRC, PhD, dissertation project, ...) and who to contact about concerns (e.g. PI, supervisor, Head of School – name and UEA email).	YES
Participants will be informed of how the data they provide will be used and shared, their ability to withdraw from the study and what happens to their data if they withdraw.	YES
Special efforts will be made to be sensitive to differences relating to age, culture, disability, race, sex, religion and sexual orientation amongst research participants when planning, conducting and reporting on the research.	YES

Data generated by the research (e.g. transcripts of research interviews) will be kept in a safe and secure location and will be used purely for the purposes of the research project (including dissemination of findings). No-one other than research colleagues, professional transcribers and supervisors will have access to any identifiable raw data collected, unless written permission has been explicitly given by the participant.	YES
All appropriate steps will be taken to protect the privacy and ensure the anonymity and non-traceability of participants.	YES

I am satisfied that all ethical issues have been identified and that satisfactory procedures are in place to deal with those issues in this research project. I will abide by the procedures described in this form.

5. Guidance notes:

General points

Any research, dissertation or project carried out at UEA that involves working with people or animals - either directly or indirectly - must obtain ethics approval before work starts. Failure to do so is a Research Misconduct matter.

Many applications can be processed quickly, but work that falls outside the scope of SCI-REC (a sub-committee of the UEA REC) will be referred elsewhere. Work that involves medical patients, or NHS staff issues that may affect health and well-being, must be approved by a NHS REC and Research Governance Committee. Work with NHS staff on non-sensitive matters (e.g. use of IT) needs SCI-REC ethics approval and NHS Research Governance approval. Plenty of time must be allowed for these processes.

The most important issues in considering the ethical dimensions of a project are:

- **Appropriateness of methods.** Are the methods proposed appropriate (e.g. not unduly intrusive, or time-consuming) for the gains in knowledge and understanding expected,
- **Experimental subjects and consent.** These are *indicative* topics:
 - How will you recruit subjects?
 - How many will be recruited? (justified in relation to the aims of the survey and the analysis methods)
 - How will you obtain the informed consent of your subjects?
 - How will they be informed of their options to withdraw and of any risks or benefits from participating?

Incentives. If you intend providing any reward (e.g. money voucher) for participating in the work, you must collect the names, addresses and signatures of every participant who is rewarded. This information must not be linked to the survey responses in any way.

Complete the sections of the form that are relevant to your project and leave the others blank (or N/A).

Notes on the form

3.1 Provide the reviewers with sufficient information to understand the main motivations and goals of the project, so that they can understand how the work needing ethics approval contributes to the research or other outcomes.

3.2 This is a clear description of what will be done. It should be possible (in principle) for an experienced researcher, knowing nothing about the project, to read this and complete the work.

3.3 Include the desired and minimum numbers, methods of recruitment (word of mouth, social media announcements, email, posters, ...) and any restrictions on participants (e.g. age, previous experience).

3.4 Most projects restrict access to the data to the researcher, supervisor (if applicable) and research colleagues for research purposes only. Any wider access or planned use must be very clearly described. It is important to consider how the published results or any data shared outside the research team might be

combined with other data (e.g. from social media) to de-anonymise it and compromise the privacy or security of participants.

3.5 Participants can only give meaningful consent if they are provided with the relevant information in a form that they can understand.

3.6 UEA is obliged under UK tax law to hold the names and addresses of everybody who receives a payment for taking part in research of this type. Typically, arrangements may be made to keep a list of recipients in a secure location (e.g. School Local Support Office) for the time required by HMRC. This data must never be stored with or be linkable from any anonymized data.

3.7 Please describe anything not covered elsewhere that may have ethical implications for the project.

3.8 If the research involves risks beyond those of everyday life they must be described here. This is particularly important for work that may lead to participants disclosing activities that could be in breach of the law (in other countries as well as the UK), recalling traumatic events, cause unwarranted anxiety, ... It is expected that the researchers consider the likelihood and severity of these risks and mitigate them through a combination of participant information, screening and possible interventions at critical moments.

3.9 Will the research be carried out in an environmentally sensitive area or area of Special Scientific Interest, or involve crossing one? Give details of how oversampling, harm to endangered flora and fauna, and other disruption to the field site and ecosystems will be avoided if applicable and details of how the relevant regulations on the transportation of samples and specimens have been respected, if applicable

3.10 Research with many of these groups is restricted or proscribed by UK legislation ("Prevent") and so must be thought out very carefully.

3.11 As a minimum, it is expected that there are arrangements to check that researchers let an independent person know when and where they are working, and that they make contact at the end of each session to report their safety. The procedures for dealing with an unexpected absence should be clearly described.

3.12 Similar considerations as for 3.11

3.13 This is to ensure that appropriate measures have been taken to ensure the health and safety of researchers and that the University is not put in breach of its legal and other obligations.

3.14 Maintaining appropriate privacy and anonymity is increasingly difficult as large volumes of data are available online, particularly through social media (where many people share data much more widely than they would normally consider desirable). It is therefore important that these issues are fully taken into account in the research design.

Attachments

Recruitment letters and emails. Copies of the text of recruitment emails, letters, introductory remarks, etc. must be attached

Questionnaire. Copies of all questionnaires, interview forms etc. must be attached. Any questionnaire or information sheet should provide participants with:

- Information on the aims of the project and questionnaire,
- How long it should take to complete,
- What will happen to the information they provide,
- What will happen if they withdraw part way through,
- What will happen to the information they provide,
- Contact details of the investigator and supervisor or Head of School (name and UEA email)

This should provide sufficient detail to allow them to decide whether or not to participate.

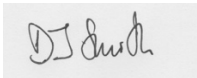
Participation information and consent. Any participant information sheets, consent forms, etc. that will be used in the research. For online surveys it should be clearly stated at the beginning that by continuing the survey participants agree to participate and for the data they provide to be used.

Other documents. Any task specifications, guidance, observation recording forms, ... that will be used in the research.

Appendix 3: Main Study 1 ethical approval 17.12.2020. Ref: SCI-CMP/1920/R/17. Same ethics reference was utilised due to the same application previously having approval, but amended with more data collection.

Research Ethics Application

1. Application summary

Applicant to complete			
Applicant Name	James Campbell		
Email	J.Campbell1@uea.ac.uk		
School	CMP	PGR	
Name of Supervisor / PI	Dr Oliver Buckley		
Short title of project	Identification and Authentication Utilising Keystroke Dynamic Inference		
Start date	October 2018	End date	October 2023
Outcome: for SCI-REC use only			
Lead Reviewer	D J Smith	Review date	
Decision	Approve	Approve after revision	Reject
Conditions (if any) or reasons for rejection	revisions are OK - DJS		
Approved by (Chair SCI-REC)	D J Smith	Signature	

2. Checklist

A. Does the research use an interview or questionnaire survey?	Yes
If so, does it ask for any personal information?	Yes
B. Does the research offer advice or guidance to people?	No
C. Does the research involve children, vulnerable adults or their carers?	No
D. Does the research record or observe people's behaviour?	No
E. Has this research been previously considered by another REC?	Yes
F. Does the research involve the analysis of personal data collected by others?	No
G. Will the researcher carry out fieldwork alone while away from UEA?	No
H. Will participants be paid or offered a reward for participating?	No
I. Does the project rely on data supplied from external sources?	No
I.1 Does the research use sensitive data? (commercially confidential, military, ...)	No
I.2 Does the research use confidential data? (e.g. medical records)	No
II.3 Is the research covered by the consent given when the data were collected?	No
J. Are special arrangements needed for the storage and retention of the data?	No
K. Will the research have an impact on the environment?	No
K.1. Will the research be carried out in an environmentally sensitive area, or SSSI?	No
K.2. Have appropriate steps been taken to gain permission to access the field site(s)?	No
K.3. Will samples be collected and removed in sufficient quantities to have a negative physical/environmental impact on the site and/or its ecosystem?	No
K.4. Does the fieldwork involve sampling rare/endangered or harmful taxa/species?	No
K.5. Will the fieldwork significantly disrupt the site and/or its environment?	No
K.6. Will the research involve transporting samples/specimens between countries or across other significant boundaries?	No

All applications must be approved by the Research Ethics Committee before beginning data generation or approaching potential research participants.

3. Application form

This form is for all staff and students who are planning research in areas of interest to the Schools of the UEA Faculty of Science, excluding animal, tissue and medical-related research which are covered by other committees. Any research involving NHS patients must be approved by the relevant NHS ethics committee.

SCI-REC has particular expertise in topics related to computing (data science, machine learning, cybersecurity, online privacy, ...), environmental impacts, and science education.

Please refer to the guidance notes at the end of this form when preparing your application as this can clarify what the committee needs to see about your project and can avoid any unnecessary requests for further information.

- **Briefly describe the scope and aims of your project in language understandable by a wide audience. (250 words maximum)**

The aim of this piece of research is to determine if a participant's name can be inferred from the way they type (which is in turn inferred from sensors). This research will analyse the results achieved by asking the user to copy a pre-selected passage of text to ascertain which letters they type fastest, in the hopes of being able to identify their name.

In addition to the name of a participant, comparisons will be made around other demographic data such as (age, handedness and gender) to determine the impact this has on the correct prediction of a name.

Participants will be using a custom made Android Application on their personal mobile device to complete the experiment in it's entirety. The Android Application records the keystrokes (time up/time down), the size of the screen in pixels, the x and y location of presses on the screen, the size of the press on the screen and finally the tilt and rotation of the device, of a participant into a text file, along with some basic demographic questions which can be seen in the supporting materials.

This research is being conducted from the School of Computing Sciences, University of East Anglia (UEA) as part of my, (James Campbell - Researcher), PhD Thesis, which is supervised by Dr Oliver Buckley (UEA Lecturer). Please see the attached supporting materials for further clarification.

- **Briefly outline the proposed research methods. (250 words maximum)**

The research method will be based around the interaction with the Android Mobile Phone Application, both with the questionnaire and experiment itself.

Upon loading the application, the participant will be presented with the Participant Information Statement, which they are required to have read and understood before proceeding. Following the PIS, is a short questionnaire in which the participant is required to enter their first name, and select from drop down fields for their age, gender and handedness. Once this questionnaire has been completed the experiment begins and the participant is presented with a passage of text and the on screen keyboard to use to copy the text into a field. Upon completion the participant presses a submit button (which saves the copied text to a .txt file on the phone's memory card containing the questionnaire answers and key timings, this is then uploaded by the application to a secure AWS S3 data storage bucket on the cloud to negate the need for the user to send an email). Finally a unique identifier (in the form of a random 12 digit number) is presented to the participant, with which they can withdraw from the research at any time.

Once the participant has completed the experiment, they are reminded that they can withdraw at any time, and thanked for their participation. If a participant decides to end the experiment early then they can close down the app at anytime and not continue the experiment.

The participant data will then be downloaded onto a secured, encrypted hard drive so that it can be analysed at a later date.

- **Briefly explain how participants will be recruited. (250 words maximum)**

Any participants, as long as they are over the age of 18, will be recruited, and no demographic details will be considered in choosing participants who are suitable for the study. There will be a conscious effort however to ensure that there is a wide spread of participants from differing demographics.

Participants will be recruited via two main methods.

- Social Media
 - The researcher (James Campbell) will utilise the contacts on his social media accounts (Facebook + LinkedIn) to recruit friends, family and colleagues to participate in the study, after providing them with full Participant Information Statement up front. Before the experiment begins, the Participant Information Statement will be provided again to ensure the full understanding and agreement from the participant before undertaking any aspect of the experiment.
- Google Play Store
 - The researcher (James Campbell) will also the Google Play Store to recruit participants to the study. The Android Application will be uploaded to the store for members of the public to download as they choose. In doing this, it will enable the research to reach a diverse range of participants and provide an easy method of delivery.

Please note – A maximum of 300 participants will be recruited for the study using the combination of recruitment methods above. Whilst the maximum is 300 participants, a more realistic figure of around 150 participants is expected, depending on the appetite for the research.

- **State who will have access to the data and what measures will be adopted to maintain participants' confidentiality or anonymity. (250 words maximum)**

The data will be accessible exclusively by the researcher on this project (James Campbell) as well as the researcher's Supervisor (Dr. Oliver Buckley), and any associated researchers within the Cyber Research Group within UEA. In order to protect anonymity, only first names are being recorded, in addition, participants will be identified by using a unique identifier (number) to reduce the likelihood of a correlation between participant and data. Furthermore, the analysis will be completed after the data collection period to reduce the likelihood that anonymity will be compromised.

If publication is to occur in the future, then participants will be protected via the use of amendments to ensure anonymity.

Upon initial collection, the data will be stored in an AWS (Amazon Web Services) S3 (Simple Storage Service) Bucket. This will ensure that the data is secure and will also negate the necessity for a user to email the text file over to me, which reduces the likelihood of an error occurring in transit of the data.

All of the results from the experiments and questionnaires will then be secured on an encrypted hard drive once the data collection has ended, and the data on it will be kept in line with UEA data management policies and the DPA (Data Protection Act).

- **Give details of how the consent of participants will be (250 words maximum), e.g. through information sheets and consent forms, oral or other approach. Copies of all forms should be submitted with the application (do not include the text of these documents in this space).**

The consent of the participants will be gathered with a consent form on the landing page of the study, after reading the Participant Information Statement before the questionnaire or demographic information is gathered. This information and consent form can be seen in the attached supporting materials. In order to ensure the consent of participants is legitimate, the user of the application is not allowed to progress with the study until they have agreed to proceed. The form will also include details of the purpose of the study and the methods of withdrawal.

- **If any payment or incentive will be made to any participant, please explain what it is, how anonymity will be maintained (given the need to record recipients' details for HMRC compliance) and provide the justification (no more than 250 words).**

N/A

- **Please add here any other ethical considerations the ethics committee may need to be made aware of (no more than 250 words).**
 - **Are there any issues here for who can or cannot participate in the project?**
 - **If you are conducting research in a space where individuals may also choose not to participate, how will you ensure they will not be included in any data collection or adversely affected by non-participation?**

There are no issues for those who can or cannot participate in the project as all participation is completed on the user's terms and as such can be stopped at any time. If a participant wishes not to participate, then the experiment (if started) can be stopped immediately by closing the app and the data will be removed upon removal of the application.

If the user decides that after they have participated they wish to withdraw from the study, then they can email either myself or my supervisor with their unique identifier. Once this has been received the data will be removed immediately and confirmation will be sent to the user.

- **What risks or costs to the participants are entailed in involvement in the research and how will these risks be managed?**
 - **Are there any potential physical, psychological or disclosure dangers that can be anticipated?**
 - **What is the possible harm to the participant or society from their participation?**
 - **What arrangements for the care and protection of participants have been made?**

There are no risks or costs to the participants.

- **Does this research have environmental implications? Please refer to the University's Research Ethics Guidance Note: [Research with a Potential Impact on the Environment](#) for further details. Identify any significant environmental impacts arising from your research/project and the measures you will take to minimise risk of adverse impact.**

There are no environmental implications for this research, other than the electricity being used to charge the mobile device before beginning the experiment.

- **Will your research involve investigation of or engagement with terrorist or violent extremist groups, people engaged in or supporting activities that compromise computer security or other activity that may normally be considered harmful or unlawful? Please provide a full explanation if the answer is 'yes'.**

N/A

- **If any of the work is being carried out away from UEA, describe the arrangements been made to ensure the safety of the researchers.**

N/A

- **State any precautions being taken to protect the health and safety of other researchers and others associated with the project (as distinct from the participants or the applicant).**

N/A

- **If the research is being undertaken outside the UK:**
 - **Have any necessary research permissions, in-country ethical clearance, visas, ... been obtained?**
 - **Have appropriate steps been taken to identify and minimize any risks?**
 - **Are the researchers aware of health and safety advice, and other travel advisory notices?**
 - **If the researchers are travelling to conduct the research, have they taken out travel and health insurance for the full period of the research? If not, why not.**
 - **Has formal permission or a research permit been sought to conduct this research? If not, please explain why this is not necessary and/or appropriate**

N/A

- **Describe the steps that will be taken and the techniques that will be used to ensure that the anonymity and confidentiality of participants will be preserved, including the measures will be implemented to prevent de-anonymisation.**

In order to protect anonymity, the only piece of potentially identifiable demographic information which is being recorded is the first name. In addition, participants will be identified by using a unique identifier (number) to reduce the likelihood of a correlation between participant and data. Furthermore, the analysis will be completed after the data collection period to reduce the likelihood that anonymity will be compromised. For other pieces of data which are being recorded (age, handedness and gender) these will be recorded in the format of a drop down box, so that the answers are generic and cannot be used to identify a specific participant.

Applications with missing or incomplete sections will be returned to the applicant for resubmission.

Supporting documents	Tick to confirm
Participant Information and consent	Yes
Other supporting documents (e.g. questionnaires, interview/focus group questions, stimulus materials, observation checklists, letters of invitation, recruitment posters, etc.)	Yes

4. DECLARATION:

Please complete the following boxes with YES, NO, or NOT APPLICABLE:

<i>For student applicants:</i> I have discussed the planned work with my supervisor, who is in agreement with this application.	YES
I am aware of the relevant sections of the GDPR (2018): https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/ and Freedom of Information Act (2005).	YES
The purpose and procedures of the research, and the potential benefits and costs of participating (e.g. the amount of their time involved), will be explained to participants.	YES
It is clear to participants who is conducting the research (name, School and UEA email), why the research is being carried out (e.g. funded by UKRC, PhD, dissertation project, ...) and who to contact about concerns (e.g. PI, supervisor, Head of School – name and UEA email).	YES
Participants will be informed of how the data they provide will be used and shared, their ability to withdraw from the study and what happens to their data if they withdraw.	YES
Special efforts will be made to be sensitive to differences relating to age, culture, disability, race, sex, religion and sexual orientation amongst research participants when planning, conducting and reporting on the research.	YES

Data generated by the research (e.g. transcripts of research interviews) will be kept in a safe and secure location and will be used purely for the purposes of the research project (including dissemination of findings). No-one other than research colleagues, professional transcribers and supervisors will have access to any identifiable raw data collected, unless written permission has been explicitly given by the participant.	YES
All appropriate steps will be taken to protect the privacy and ensure the anonymity and non-traceability of participants.	YES

I am satisfied that all ethical issues have been identified and that satisfactory procedures are in place to deal with those issues in this research project. I will abide by the procedures described in this form.

5. Guidance notes:

General points

Any research, dissertation or project carried out at UEA that involves working with people or animals - either directly or indirectly - must obtain ethics approval before work starts. Failure to do so is a Research Misconduct matter.

Many applications can be processed quickly, but work that falls outside the scope of SCI-REC (a sub-committee of the UEA REC) will be referred elsewhere. Work that involves medical patients, or NHS staff issues that may affect health and well-being, must be approved by a NHS REC and Research Governance Committee. Work with NHS staff on non-sensitive matters (e.g. use of IT) needs SCI-REC ethics approval and NHS Research Governance approval. Plenty of time must be allowed for these processes.

The most important issues in considering the ethical dimensions of a project are:

- **Appropriateness of methods.** Are the methods proposed appropriate (e.g. not unduly intrusive, or time-consuming) for the gains in knowledge and understanding expected,
- **Experimental subjects and consent.** These are *indicative* topics:
 - How will you recruit subjects?
 - How many will be recruited? (justified in relation to the aims of the survey and the analysis methods)
 - How will you obtain the informed consent of your subjects?
 - How will they be informed of their options to withdraw and of any risks or benefits from participating?

Incentives. If you intend providing any reward (e.g. money voucher) for participating in the work, you must collect the names, addresses and signatures of every participant who is rewarded. This information must not be linked to the survey responses in any way.

Complete the sections of the form that are relevant to your project and leave the others blank (or N/A).

Notes on the form

3.1 Provide the reviewers with sufficient information to understand the main motivations and goals of the project, so that they can understand how the work needing ethics approval contributes to the research or other outcomes.

3.2 This is a clear description of what will be done. It should be possible (in principle) for an experienced researcher, knowing nothing about the project, to read this and complete the work.

3.3 Include the desired and minimum numbers, methods of recruitment (word of mouth, social media announcements, email, posters, ...) and any restrictions on participants (e.g. age, previous experience).

3.4 Most projects restrict access to the data to the researcher, supervisor (if applicable) and research colleagues for research purposes only. Any wider access or planned use must be very clearly described. It is important to consider how the published results or any data shared outside the research team might be

combined with other data (e.g. from social media) to de-anonymise it and compromise the privacy or security of participants.

3.5 Participants can only give meaningful consent if they are provided with the relevant information in a form that they can understand.

3.6 UEA is obliged under UK tax law to hold the names and addresses of everybody who receives a payment for taking part in research of this type. Typically, arrangements may be made to keep a list of recipients in a secure location (e.g. School Local Support Office) for the time required by HMRC. This data must never be stored with or be linkable from any anonymized data.

3.7 Please describe anything not covered elsewhere that may have ethical implications for the project.

3.8 If the research involves risks beyond those of everyday life they must be described here. This is particularly important for work that may lead to participants disclosing activities that could be in breach of the law (in other countries as well as the UK), recalling traumatic events, cause unwarranted anxiety, ... It is expected that the researchers consider the likelihood and severity of these risks and mitigate them through a combination of participant information, screening and possible interventions at critical moments.

3.9 Will the research be carried out in an environmentally sensitive area or area of Special Scientific Interest, or involve crossing one? Give details of how oversampling, harm to endangered flora and fauna, and other disruption to the field site and ecosystems will be avoided if applicable and details of how the relevant regulations on the transportation of samples and specimens have been respected, if applicable

3.10 Research with many of these groups is restricted or proscribed by UK legislation ("Prevent") and so must be thought out very carefully.

3.11 As a minimum, it is expected that there are arrangements to check that researchers let an independent person know when and where they are working, and that they make contact at the end of each session to report their safety. The procedures for dealing with an unexpected absence should be clearly described.

3.12 Similar considerations as for 3.11

3.13 This is to ensure that appropriate measures have been taken to ensure the health and safety of researchers and that the University is not put in breach of its legal and other obligations.

3.14 Maintaining appropriate privacy and anonymity is increasingly difficult as large volumes of data are available online, particularly through social media (where many people share data much more widely than they would normally consider desirable). It is therefore important that these issues are fully taken into account in the research design.

Attachments

Recruitment letters and emails. Copies of the text of recruitment emails, letters, introductory remarks, etc. must be attached

Questionnaire. Copies of all questionnaires, interview forms etc. must be attached. Any questionnaire or information sheet should provide participants with:

- Information on the aims of the project and questionnaire,
- How long it should take to complete,
- What will happen to the information they provide,
- What will happen if they withdraw part way through,
- What will happen to the information they provide,
- Contact details of the investigator and supervisor or Head of School (name and UEA email)

This should provide sufficient detail to allow them to decide whether or not to participate.

Participation information and consent. Any participant information sheets, consent forms, etc. that will be used in the research. For online surveys it should be clearly stated at the beginning that by continuing the survey participants agree to participate and for the data they provide to be used.

Other documents. Any task specifications, guidance, observation recording forms, ... that will be used in the research.

Appendix 4: Main Study 2 ethical approval 23.11.2022. Ref: ETH2223-0921.

Ethics ETH2223-0921 : Mr James Campbell

Date Created	23 Nov 2022
Date Submitted	23 Nov 2022
Date forwarded to committee	23 Nov 2022
Researcher	Mr James Campbell
Category	PGR
Supervisor	Prof Oli Buckley
Faculty	Faculty of Science
Current status	Approved

Ethics application

Applicant and research team

Principal Applicant

Name of Principal Applicant
Mr James Campbell

UEA account
njc12hmu@uea.ac.uk

School/Department
School of Computing Sciences

Category
PGR

Primary Supervisor

Name of Primary Supervisor

[Prof Oli Buckley](#)

Primary Supervisor's school/department
School of Computing Sciences

Project details

Project title

Identification and Authentication Utilising Keystroke Dynamic Inference

Project start date

01 Jan 2023

Project end date

01 Oct 2023

Describe the scope and aims of the project in language understandable by a non-technical audience. Include any other relevant background which will allow the reviewers to contextualise the research.

The aim of this piece of research is to determine if a participant's name can be inferred from the way they type (which is in turn inferred from sensors). This research will analyse the results achieved by asking the user to copy a pre-selected passage of text to ascertain which letters they type fastest, in the hopes of being able to identify their name.

In addition to the name of a participant, comparisons will be made around other demographic data such as (age, handedness and gender) to determine the impact this has on the correct prediction of a name.

Participants will be using a custom made Android Application on their personal mobile device to complete the experiment in it's entirety. The Android Application records the keystrokes (time up/time down), the size of the screen in pixels, the x and y location of presses on the screen, the size of the press on the screen and finally the tilt and rotation of the device, of a participant into a text file, along with some basic demographic questions which can be seen in the supporting materials.

This research is being conducted from the School of Computing Sciences, University of East Anglia (UEA) as part of my, (James Campbell - Researcher), PhD Thesis, which is supervised by Dr Oliver Buckley (UEA Lecturer). Please see the attached supporting materials for further clarification.

Provide a brief explanation of the research design (e.g. interview, experimental, observational, survey), questions, methodology, and data gathered/analysis. If relevant, include what the participants will be expected to do/experience.

The research method will be based around the interaction with the Android Mobile Phone Application, both with the questionnaire and experiment itself.

Upon loading the application, the participant will be presented with the Participant Information Statement, which they are required to have read and understood before proceeding. Following the PIS, is a short questionnaire in which the participant is required to enter their first name, and select from drop down fields for their age, gender and handedness. Once this questionnaire has been completed the experiment begins and the participant is presented with a passage of text and the on screen keyboard to use to copy the text into a field. Upon completion the participant presses a submit button (which saves the copied text to a .txt file on the phone's memory card containing the questionnaire answers and key timings, this is then uploaded by the application to a secure AWS S3 data storage bucket on the cloud to negate the need for the user to send an email). Finally a unique identifier (in the form of a random 12 digit number) is presented to the participant, with which they can withdraw from the research at any time.

Once the participant has completed the experiment, they are reminded that they can withdraw at any time, and thanked for their participation. If a participant decides to end the experiment early then they can close down the app at anytime and not continue the experiment.

The participant data will then be downloaded onto a secured, encrypted hard drive so that it can be analysed at a later date.

Detail how any adverse events arising in the course of the project will be reported in a timely manner.

Any adverse events will be reported immediately to the committee that has given approval via email within 24 hours of occurrence. As results will be uploaded in real time, these will be monitored and reported on as necessary.

Will you also be applying for Health Research Authority approval (HRA)?

No

Indicate if you are applying for approval for an experiment to be conducted in the School of Economics' Laboratory for Economic and Decision Research (LEDR).

No

Is the project?:

none of the options listed

Does the project have external funding administered through the University's Research and Innovation Services (RIN)?

No

Will the research take place outside of the UK?

No

Will any part of the project be carried out under the auspices of an external organisation, or involve collaboration between institutions?

No

Do you require or have you already gained approval from an ethics review body external to UEA?

No

Does this new project relate to a project which already has ethics approval from UEA?

Yes

If yes, provide the name of the UEA ethics approval body.

[SCI S-REC \(Faculty of Science Research Ethics Subcommittee\)](#)

If yes, provide the date of the ethics approval.

17 Dec 2022

If yes, provide the UEA ethics application reference number, if allocated.

SCI-CMP/1920/R/17

Research categories

Will the project include primary data collection involving human participants?

Yes

Will the project use secondary data involving human participants?

Will the project involve the use of live animals?

No

Will the project have the potential to affect the environment?

No

Will the project have the potential to affect culturally valuable, significant or sensitive objects or practices?

No

Will the project involve security sensitive research?

No

Will the project involve a generative Artificial Intelligence (AI) tool?

Human participants - selection and recruitment

How many Participant Groups are there who will receive tailored participant information?:

One

Name of Participant Group 1.

Participants

How will the participants be selected/recruited?

A small set of participants will be actively recruited, as long as they are over the age of 18, and no demographic details will be considered in choosing participants who are suitable for the study. There will be a conscious effort however to ensure that there is a wide spread of participants from differing demographics.

Participants will be recruited via Social Media:

- Social Media

- o The researcher (James Campbell) will utilise the contacts on his social media accounts (Facebook + LinkedIn) to recruit friends, family and colleagues to participate in the study, after providing them with full Participant Information Statement up front. Before the experiment begins, the Participant Information Statement will be provided again to ensure the full understanding and agreement from the participant before undertaking any aspect of the experiment.

Please note – As this study is a longitudinal study, a smaller maximum amount of 30 participants will be recruited for the study using the recruitment method above. Whilst the maximum is 30 participants, a more realistic figure of around 20 participants is expected, depending on the appetite for the research.

It is also worth noting that the application will be posted on the Google Play Store to allow for the participants to download the application easily and also with confidence that the application is safe. Due to this – there is a minute chance that a member of the public may download the application.

This does not matter, as we will only be focusing on the longitudinal data and can therefore omit this data.

If appropriate, upload a copy of the proposed advertisement, including proposed recruitment emails, flyers, posters or invitation letter.

How and when will participants receive this recruitment material?

The consent of the participants will be gathered with a consent form on the landing page of the study, after reading the Participant Information Statement before the questionnaire or demographic information is gathered. This information and consent form can be seen in the attached supporting materials. In order to ensure the consent of participants is legitimate, the user of the application is not allowed to progress with the study until they have agreed to proceed. The form will also include details of the purpose of the study and the methods of withdrawal.

In terms of UEA participants only, will you be advertising the opportunity to take part in this project to?:

None of the above (i.e. UEA's Student Insight Review Group (SIRG) does not need to be informed)

What are the characteristics of the participants?

Age of 18+ is the only exclusion, other than this there are no exclusions

Will the project require the cooperation of a gatekeeper for initial access to the individuals/groups to be recruited?

No

Is there any sense in which participants might be 'obliged' to participate?

No

Will the project involve vulnerable groups?

No

Will payment or any other incentive be made to any participant?

No

Include any other ethical considerations regarding participation.

There are no issues for those who can or cannot participate in the project as all participation is completed on the user's terms and as such can be stopped at any time. If a participant wishes not to participate, then the experiment (if started) can be stopped immediately by closing the app and the data will be removed upon removal of the application.

If the user decides that after they have participated they wish to withdraw from the study, then they can email either myself or my supervisor with their unique identifier. Once this has been received the data will be removed immediately and confirmation will be sent to the user.

Human participants - consent options

By which method(s) will consent to participate in the research be obtained?:

Online Participant Information and Consent

Human participants - information and consent

Participant Information and Consent

Will opt out consent for participation in the research be used?

No

You can generate a Participant Information Text and Consent Form for this application by completing information in the Participant Information Text and Consent Form Generator tab. Alternatively you can upload your Participant Information Text and Participant Consent Form which you have already prepared. Confirm below:

Upload prepared Participant Information Text and Consent Form.

Upload the Participant Information Text and Consent Form.

Enter participant group number and name.

Participants

When will participants receive the participant information and consent request?

Once they open the application, and also presented before hand when being recruited for the study.

How will you record a participant's decision to take part in the research?

By them continuing and agreeing with the participation. If they wish to stop at any time they can withdraw from the study via their unique identifier or they can simply close the application and no data will be sent.

Human participants - method

Which data collection methods will be used in the research?:

Non-anonymous questionnaire

Other methods which involve recording or observing people's behaviour, e.g. experiments

If your research involves any of the methods (including Other) listed above, upload supporting materials.

How have your characteristics, or those of the participants influenced the design of the study or how the research is experienced by participants?

Not applicable due to the nature of the survey/experiment. No interviews etc will be conducted.

Will the project involve transcripts?

No

Will you be capturing photographs or video footage (digital assets) of individuals taken for University business?

No

Is this research using visual/vocal methods where respondents may be identified?

No

Will it be necessary for participants to take part in the study without their knowledge and consent at the time?

No

Will deception or incomplete disclosure be used?

No

Will the participants be debriefed?

No

Will substances be administered to the participants?

No

Will involvement in the project result in, or the risk of, discomfort, physical harm, psychological harm or intrusive procedures?

No

Will the project involve prolonged or repetitive testing?

No

Will the project involve potentially sensitive topics?

No

Will the project involve elite interviews?

No

Will the project involve any incitement to, encouragement of, or participation, in an illegal act (by participant or researcher)?

No

Will the research involve an investigation of people engaged in or supporting activities that compromise computer security or other activities that may normally be considered harmful or unlawful?

No

Does the research involve members of the public in participatory research where they are actively involved in undertaking research tasks?

No

Does the research offer advice or guidance to people?

No

Is the research intended to benefit the participants, third parties or the local community?

No

Provide an explanation.

The research will provide a novel method of identification which could benefit participants in the future if put into an industry application. The purpose of this research is purely hypothetical however and will not provide any immediate benefit.

What procedures are in place for monitoring the research with respect to ethical compliance?

Myself and my supervisor will monitor the research for ethical compliance. Due to the nature of the research (questionnaire and recording of keystrokes) there should not be any ethical implications, but we will watch for these during the data collection.

Does the study involve the use of a clinical or non-clinical scale, questionnaire or inventory which has specific copyright permissions, reproduction or distribution restrictions or training requirements?

No

Include any other ethical considerations regarding data collection methods.

There are no videos/photographs or any other kind of personally identifiable information other than first name. We also made sure to anonymise as best we could the options for responses in the questionnaire by utilising drop down boxes and age ranges and generic statements so that there was no way of being able to identify an individual.

Health and safety - participants

Is there a possibility that the health and safety of any of the participants in this project including a support person (e.g. a care giver, school teaching assistant) may be in question?

No

Health and safety - researcher(s)

Is there a possibility that the health and safety of any of the researcher(s) and that of any other people (as distinct from any participants) impacted by this project including research assistants/translators may be in question?

No

Risk assessment

Are there hazards associated with undertaking this project where a formal risk assessment will be required?

No

Data management

Will the project involve any personal data (including pseudonymised data) not in the public domain?

No

Will any personal data be processed by another organisation(s)?

No

Will the project involve access to records of sensitive/confidential information?

No

Will the project involve access to confidential business data?

No

Will the project involve secure data that requires permission from the appropriate authorities before use?

No

Will you be using publicly available data from the internet for your study?

No

Will the research data in this study be deposited in a repository to allow it to be made available for scholarly and educational purposes?

No

Provide details.

The data will be accessible exclusively by the researcher on this project (James Campbell) as well as the researcher's Supervisor (Dr. Oliver Buckley), and any associated researchers within the Cyber Research Group within UEA. In order to protect anonymity, only first names are being recorded, in addition, participants will be identified by using a unique identifier (number) to reduce the likelihood of a correlation between participant and data. Furthermore, the analysis will be completed after the data collection period to reduce the likelihood that anonymity will be compromised.

If publication is to occur in the future, then participants will be protected via the use of amendments to ensure anonymity.

Upon initial collection, the data will be stored in an AWS (Amazon Web Services) S3 (Simple Storage Service) Bucket. This will ensure that the data is secure and will also negate the necessity for a user to email the text file over to me, which reduces the likelihood of an error occurring in transit of the data.

All of the results from the experiments and questionnaires will then be secured on an encrypted hard drive once the data collection has ended, and the data on it will be kept in line with UEA data management policies and the DPA (Data Protection Act).

Who will have access to the data during and after the project?

The data will be accessible exclusively by the researcher on this project (James Campbell) as well as the researcher's Supervisor (Dr. Oliver Buckley), and any associated researchers within the Cyber Research Group within UEA. In order to protect anonymity, only first names are being recorded, in addition, participants will be identified by using a unique identifier (number) to reduce the likelihood of a correlation between participant and data. Furthermore, the analysis will be completed after the data collection period to reduce the likelihood that anonymity will be compromised.

If publication is to occur in the future, then participants will be protected via the use of amendments to ensure anonymity.

Upon initial collection, the data will be stored in an AWS (Amazon Web Services) S3 (Simple Storage Service) Bucket. This will ensure that the data is secure and will also negate the necessity for a user to email the text file over to me, which reduces the likelihood of an error occurring in transit of the data.

All of the results from the experiments and questionnaires will then be secured on an encrypted hard drive once the data collection has ended, and the data on it will be kept in line with UEA data management policies and the DPA (Data Protection Act).

Where/how do you intend to store the data during and after the project?

The data will be accessible exclusively by the researcher on this project (James Campbell) as well as the researcher's Supervisor (Dr. Oliver Buckley), and any associated researchers within the Cyber Research Group within UEA. In order to protect anonymity, only first names are being recorded, in addition, participants will be identified by using a unique identifier (number) to reduce the likelihood of a correlation between participant and data. Furthermore, the analysis will be completed after the data collection period to reduce the likelihood that anonymity will be compromised.

If publication is to occur in the future, then participants will be protected via the use of amendments to ensure anonymity.

Upon initial collection, the data will be stored in an AWS (Amazon Web Services) S3 (Simple Storage Service) Bucket. This will ensure that the data is secure and will also negate the necessity for a user to email the text file over to me, which reduces the likelihood of an error occurring in transit of the data.

All of the results from the experiments and questionnaires will then be secured on an encrypted hard drive once the data collection has ended, and the data on it will be kept in line with UEA data management policies and the DPA (Data Protection Act).

How will you ensure the secure storage of the data during and after the project?

The data will be accessible exclusively by the researcher on this project (James Campbell) as well as the researcher's Supervisor (Dr. Oliver Buckley), and any associated researchers within the Cyber Research Group within UEA. In order to protect anonymity, only first names are being recorded, in addition, participants will be identified by using a unique identifier (number) to reduce the likelihood of a correlation between participant and data. Furthermore, the analysis will be completed after the data collection period to reduce the likelihood that anonymity will be compromised.

If publication is to occur in the future, then participants will be protected via the use of amendments to ensure anonymity.

Upon initial collection, the data will be stored in an AWS (Amazon Web Services) S3 (Simple Storage Service) Bucket. This will ensure that the data is secure and will also negate the necessity for a user to email the text file over to me, which reduces the likelihood of an error occurring in transit of the data.

All of the results from the experiments and questionnaires will then be secured on an encrypted hard drive once the data collection has ended, and the data on it will be kept in line with UEA data management policies and the DPA (Data Protection Act).

How long will research data be stored after the study has ended?

Until the thesis has been submitted and viva is completed and the researcher has gained the qualification - at which point the data will be deleted in accordance with DPA restrictions.

How long will research data be accessible after the study has ended?

Until the thesis has been submitted and viva is completed and the researcher has gained the qualification - at which point the data will be deleted in accordance with DPA restrictions.

How are you intending to destroy the project data when it is no longer required?

Securely deleted and then scrubbed using industry standard software to ensure permanent erasure.

Appendix 5: Pilot Study - Example Data Collected.

1 Name : (James) Handedness : (Right Handed) Age : (18-25)
Gender : (Male) [J - Down, 1591220055244, J - Up,
1591220055311, A - Down, 1591220055420, A - Up,
1591220055523, C - Down, 1591220055745, C - Up,
1591220055857, K - Down, 1591220155939, K - Up,
1591220056027, D - Down, 1591220057812, D - Up,
1591220057929, A - Down, 1591220058411, A - Up,
1591220058523, W - Down, 1591220058699, W - Up,
1591220058792, S - Down, 1591220059451, S - Up,
1591220059552, SPACE - Down, 1591220061303, SPACE - Up,
1591220061429, K - Down, 1591220062265, K - Up,
1591220062349, O - Down, 1591220065541, O - Up,
1591220065625, V - Down, 1591220065931, V - Up,
1591220066024, E - Down, 1591220066140, E - Up,
1591220066244, SPACE - Down, 1591220069162, SPACE - Up,
1591220069256, M - Down, 1591220069979, M - Up,
1591220070046, Y - Down, 1591220070714, Y - Up,
1591220070790, SPACE - Down, 1591220071162, SPACE - Up,
1591220071248, B - Down, 1591220072173, B - Up,
1591220072257, I - Down, 1591220072400, I - Up,
1591220072493, G - Down, 1591220072620, G - Up,
1591220072698, SPACE - Down, 1591220073365, SPACE - Up,
1591220073433, S - Down, 1591220073976, S - Up,
1591220074069, P - Down, 1591220074121, P - Up,
1591220074174, H - Down, 1591220074374, H - Up,
1591220074443, I - Down, 1591220074601, I - Up,
1591220074680, N - Down, 1591220074795, N - Up,
1591220074883, X - Down, 1591220075035, X - Up,
1591220075153, SPACE - Down, 1591220075828, SPACE - Up,
1591220075913, O - Down, 1591220076333, O - Up,

1591220076394, F - Down, 1591220076668, F - Up,
1591220076769, SPACE - Down, 1591220076945, SPACE - Up,
1591220077022, Q - Down, 1591220078155, Q - Up,
1591220078157, U - Down, 1591220078352, U - Up,
1591220078439, A - Down, 1591220078755, A - Up,
1591220078855, R - Down, 1591220079235, R - Up,
1591220079329, T - Down, 1591220079444, T - Up,
1591220079548, Z - Down, 1591220080183, Z - Up,
1591220080285, SPACE - Down, 1591220081373, SPACE - Up,
1591220081490, T - Down, 1591220081688, T - Up,
1591220081799, H - Down, 1591220081834, H - Up,
1591220081922, E - Down, 1591220081955, E - Up,
1591220062044, SPACE - Down, 1591220082086, SPACE - Up,
1591220082191, G - Down, 1591220082759, G - Up,
1591220082836, I - Down, 1591220082991, I - Up,
1591220083087, V - Down, 1591220084306, V - Up,
1591220084399, E - Down, 1591220084489, E - Up,
1591220084561, SPACE - Down, 1591220084861, SPACE - Up,
1591220084968, B - Down, 1591220085347, B - Up,
1591220085432, O - Down, 1591220085909, O - Up,
1591220085994, X - Down, 1591220086077, X - Up,
1591220086140, I - Down, 1591220086386, I - Up,
1591220086474, N - Down, 1591220086630, N - Up,
1591220086695, G - Down, 1591220087114, G - Up,
1591220487199, SPACE - Down, 1591220087464, SPACE - Up,
1591220087548, W - Down, 1591220078133, W - Up,
15912200883216, I - Down, 1591220088399, I - Up,
1591220088436, Z - Down, 1591220088488, Z - Up,
1591220088574, A - Down, 1591220088885, A - Up,
1591220089014, R - Down, 1591220089232, R - Up,
1591220089317, D - Down, 1591220089539, D - Up,

1591220089625, S - Down, 1591220089806, S - Up,
1591220089885, SPACE - Down, 1591220091038, SPACE - Up,
1591220091147, J - Down, 1591220091444, J - Up,
1591220091530, U - Down, 1591220091646, U - Up,
1591220091708, M - Down, 1591220091849, M - Up,
1591220091919, P - Down, 1591220092177, P - Up,
1591220092229, SPACE - Down, 1591220092706, SPACE - Up,
1591220092791, Q - Down, 1591220093989, Q - Up,
1591220093990, U - Down, 1591220094054, U - Up,
1591220094150, I - Down, 1591220094274, I - Up,
1591220094370, C - Down, 1591220094580, C - Up,
1591220094655, K - Down, 1591220094682, K - Up,
1591220094752, L - Down, 1591220094934, L - Up,
1591220095021, Y - Down, 1591220095292, Y - Up,
1591220095388, SPACE - Down, 1591220096015, SPACE - Up,
1591220096108, P - Down, 1591220097024, P - Up,
1591220097094, A - Down, 1591220097177, A - Up,
1591220097288, C - Down, 1591220097535, C - Up,
1591220097622, K - Down, 1591220097656, K - Up,
1591220097744, SPACE - Down, 1591250098214, SPACE - Up,
1591220098308, M - Down, 1591220098841, M - Up,
1591220098926, Y - Down, 1591220099174, Y - Up,
1591220099259, SPACE - Down, 1591220099498, B - Down,
1591220099593, B - Up, 1591220099593, O - Down,
1591220093822, O - Up, 1591220099911, X - Down,
1591220100102, X - Up, 1591220100172, SPACE - Down,
1591222100322, SPACE - Up, 1591220100401, W - Down,
1591220100673, W - Up, 1591220100783, I - Down,
1591220100826, I - Up, 1591220100922, T - Down,
1591220100998, T - Up, 1591220101092, H - Down,
1591220101127, H - Up, 1591220101215, SPACE - Down,

1591220101363, SPACE - Up, 1591220101434, F - Down,
1591220101599, F - Up, 1591220101696, S - Down,
1591220101860, S - Up, 1591220101931, I - Down,
1591220102185, I - Up, 1591220102258, V - Down,
1591220102825, V - Up, 1591220102910, E - Down,
1591220103065, E - Up, 1591220103161, SPACE - Down,
1591220104095, SPACE - Up, 1591220104163, D - Down,
1591220104790, D - Up, 1591220104874, O - Down,
1591220105014, O - Up, 1591220105087, S - Down,
1591220106392, S - Up, 1591220106460, E - Down,
1591220106623, E - Up, 1591220106680, B - Down,
1591220106860, B - Up, 1591220106948, SPACE - Down,
1591220107246, SPACE - Up, 1591220107347, L - Down,
1591220107758, L - Up, 1591220107812, I - Down,
1591220107943, I - Up, 1591220108031, Q - Down,
1591220108261, Q - Up, 1591220108357, U - Down,
1591220109275, U - Up, 1591220109359, O - Down,
1591220179531, O - Up, 1591220109603, R - Down,
1591220109702, R - Up, 1591225509783, SPACE - Down,
1591220110765, SPACE - Up, 1591220110849, J - Down,
1591220111090, J - Up, 1591220111174, U - Down,
1591220111274, U - Up, 1591220111371, G - Down,
1591220111575, G - Up, 1591220111663, D - Down,
1591220111729, D - Up, 1591220111801]

Appendix 6: Main Study 1 and 2 - Example Data Collected.

1 Name: (James) Handedness: (Right Handed) Age: (26-35)
Gender: (Male) Height: 2042 Width: 1080 [J - Down,
1608830169943, J - Up, 1608830160025, A - Down,
1608830160033, A - Up, 1608230160107, C - Down,
1608830160326, C - Up, 1608830160393, K - Down,
1600830160426, K - Up, 1608830160494, SPACE - Down,
1608830160662, SPACE - Up, 1608830160761, D - Down,
1608830160795, D - Up, 1608830160870, A - Down,
1608830162936, A - Up, 1608830162997, W - Down,
1608830163166, W - Up, 1608830163255, D - Down,
1608830163424, D - Up, 1608830163498, SPACE - Down,
1608830164563, SPACE - Up, 1608833164628, L - Down,
1608830164805, L - Up, 1608830164888, O - Down,
1608830165047, O - Up, 1608830165088, V - Down,
1608830165298, V - Up, 1608830165382, E - Down,
1608830165500, E - Up, 1708830165583, SPACE - Down,
1608830165643, SPACE - Up, 1608830165733, M - Down,
1608830166018, M - Up, 1608830166060, Y - Down,
1608830166253, Y - Up, 1608830166329, SPACE - Down,
1608830166471, SPACE - Up, 1608830166562, H - Down,
1608830166688, H - Up, 1608830166746, I - Down,
1608830166906, I - Up, 1608830166973, G - Down,
1608830167166, G - Up, 1608830167249, SPACE - Down,
1208830167400, SPACE - Up, 1608830167491, S - Down,
1608830167927, S - Up, 1608830168037, P - Down,
1608830168146, P - Up, 1608830168703, H - Down,
1608830168370, H - Up, 1608830168454, I - Down,
1608830168605, I - Up, 1608830198663, M - Down,
1608830168806, M - Up, 1608830168897, X - Down,
1608830168939, X - Up, 1608830169031, SPACE - Down,

1608830170271, SPACE - Up, 1608840170364, O - Down,
1608830171344, O - Up, 1608830171417, F - Down,
1608830171577, F - Up, 1608830171660, SPACE - Down,
1608830171778, SPACE - Up, 1608830171887, Q - Down,
1608830171953, Q - Up, 1678830172045, U - Down,
1608830172163, U - Up, 1608830172271, A - Down,
1608830172297, A - Up, 1608830172380, R - Down,
1608830172874, R - Up, 1608830172951, T - Down,
1608850173079, T - Up, 1608830173151, Z - Down,
1608830173259, Z - Up, 1608830173360, SPACE - Down,
1608830173854, SPACE - Up, 1608830173963, T - Down,
1608830174331, T - Up, 1608830174406, H - Down,
1608830174582, H - Up, 1608830174658, E - Down,
1608830174658, E - Up, 1608830174766, SPACE - Down,
1608830174859, SPACE - Up, 1608830174959, F - Down,
1608830175069, F - Up, 1608830575135, I - Down,
1608830175378, I - Up, 1608830175411, V - Down,
1608830175596, V - Up, 1608830175691, E - Down,
1608830175747, E - Up, 1608830175830, SPACE - Down,
1608830175914, SPACE - Up, 1608830175997, SPACE - Down,
1608830176232, SPACE - Up, 1608830176307, O - Down,
1608830176517, O - Up, 1608830176617, X - Down,
1608830176640, X - Up, 1608830176725, X - Down,
1605830178317, X - Up, 1608830178392, I - Down,
1608830178392, I - Up, 1638830178467, M - Down,
1608830178577, G - Down, 1608839178652, M - Up,
1608830178660, G - Up, 1608830178710, SPACE - Down,
1608830179824, SPACE - Up, 1608930179915, W - Down,
1608830180124, W - Up, 1608830180225, I - Down,
1608830180459, I - Up, 1608830180560, Z - Down,
1608830180953, Z - Up, 1608830181054, A - Down,

1608830182965, A - Up, 1608830183038, R - Down,
1608830183340, R - Up, 1608830183433, D - Down,
1608830183608, D - Up, 1608830183691, S - Down,
1608830183817, S - Up, 1608830183891, SPACE - Down,
1608830183984, SPACE - Up, 1608830184093, F - Down,
1608830184587, F - Up, 1608830184670, U - Down,
1608830184821, U - Up, 1608830184871, M - Down,
1608830185190, M - Up, 1608830185266, P - Down,
1208830185585, P - Up, 1608830185659, SPACE - Down,
1608830186145, SPACE - Up, 1608830186244, Q - Down,
1608830186412, Q - Up, 1608830186529, U - Down,
1608830186579, U - Up, 1608830186646, I - Down,
1608830186998, I - Up, 1608730187065, C - Down,
1608830187299, C - Up, 1608830187383, K - Down,
1608830187475, K - Up, 1608830187534, L - Down,
1608830187667, L - Up, 1608830137743, Y - Down,
1608830188009, Y - Up, 1608830188103, SPACE - Down,
1608830188388, SPACE - Up, 1608830188463, P - Down,
1608830189116, P - Up, 1608830189216, A - Down,
1608830189042, A - Up, 1608830189425, C - Down,
1608830189643, C - Up, 1608830189735, K - Down,
1608830189744, K - Up, 1608830189819, SPACE - Down,
1608830190021, SPACE - Up, 1608830190112, M - Down,
1608830190406, M - Up, 1608830190489, Y - Down,
1608830190673, Y - Up, 1608830190748, SPACE - Down,
1608830190907, SPACE - Up, 1608830190999, B - Down,
1608830191117, B - Up, 1608830191184, O - Down,
1609830191335, O - Up, 1608830191410, X - Down,
1608830191493, X - Up, 1608830191602, SPACE - Down,
1608830191745, SPACE - Up, 1608830191837, W - Down,
1608830191996, W - Up, 1608830192046, I - Down,

1608830192524, I - Up, 1608830192615, T - Down,
1608830192667, T - Up, 1608830192741, H - Down,
1608830192852, H - Up, 1608830192934, SPACE - Down,
1608830194252, SPACE - Up, 1608830194332, F - Down,
1608830194576, F - Up, 1608830194643, I - Down,
1608830194718, I - Up, 1608830194784, B - Down,
1608830194919, E - Down, 1608830195003, B - Up,
1608830195010, E - Up, 1608830195085, SPACE - Down,
1608830195170, SPACE - Up, 1608830195228, D - Down,
1608830195328, D - Up, 1608830195394, O - Down,
1608830195579, O - Up, 1608830195681, Z - Down,
1608830195773, Z - Up, 1608830195847, E - Down,
1608830196007, E - Up, 1608830196090, N - Down,
1608830196141, N - Up, 1608830196249, SPACE - Down,
1608830196516, SPACE - Up, 1608830196634, L - Down,
1608830196869, L - Up, 1608830196952, I - Down,
1608830197096, I - Up, 1608830197161, Q - Down,
1608830197279, Q - Up, 1608830197362, U - Down,
1608830197496, U - Up, 1608830197580, O - Down,
1608830197739, O - Up, 1608830197839, R - Down,
1608830197957, R - Up, 1608830198024, SPACE - Down,
1608830198208, SPACE - Up, 1608830198308, J - Down,
1608830199045, J - Up, 1608830199128, U - Down,
1608830199254, U - Up, 1608830199312, G - Down,
1608830199447, G - Up, 1608830199522, D - Down,
1608830199614, D - Up, 1608830199681][735, 815,
0.043137256, -0.9385557, 9.5483675, 0.967287,
0.06319156, 0.1709989, -0.0245183, 775, 1842,
0.039215688, -0.5913859, 9.143736, 3.397476,
-0.04920764, -0.030586634, 0.012133616, 67, 1795,
0.019607844, -0.5913859, 9.143736, 3.397476,

0.04364388, 0.017060855, 0.004803234, 414, 1968,
0.058823533, -0.65842557, 9.201198, 3.3232534,
0.0784632, -0.05563211, -0.005581476, 852, 1795,
0.039215688, -0.65842557, 9.201198, 3.3232534,
0.0784632, -0.05563211, -0.005581476, 658, 2086,
0.043137256, -0.36632404, 9.1628895, 3.7422516,
0.05647205, 0.012784799, 0.010911887, 315, 1828,
0.043137256, -0.36632404, 9.1628895, 3.7422516,
-0.055927157, -0.041582208, -0.0049706106, 994, 1957,
0.027450982, -0.5889916, 9.150918, 3.4381785,
-0.01316659, -0.0061520236, -0.006192341, 986, 1957,
0.03529412, -0.53392327, 9.115005, 4.0104103,
-0.004614476, 0.07875825, -0.006192341, 50, 1835,
0.019607844, -0.6368771, 9.296969, 3.2969165,
-0.043098986, 0.017060855, -0.005581476, 160, 1651,
0.043137256, -0.94573855, 9.06233, 3.46691,
-0.025383893, -0.029364903, -0.004359745, 239, 1824,
0.039215688, -0.63208854, 9.237112, 3.1460772,
-0.041266393, 0.021336911, -0.0019162843, 636, 2087,
0.043137256, -0.5554718, 9.325701, 2.9928436,
0.016154941, -0.018980194, -8.3688545E-5, 974, 1866,
0.03137255, -0.5818088, 9.141341, 3.0886145,
-0.027827354, 0.023169508, 0.011522751, 884, 1677,
0.050980397, -0.6344828, 9.256267, 2.944958,
0.03142657, -0.005541159, -0.0049706106, 552, 1943,
0.054901965, -0.7709565, 9.067119, 3.4836698,
-0.019886106, 0.03844114, 0.001138042, 310, 1661,
0.043137256, -0.679974, 9.330489, 3.6560576,
0.014933212, 0.20337476, 0.0616137, 642, 2124,
0.03529412, 0.3902668, 9.993703, 2.8180614, 0.10839559,
-0.12343815, 0.0017489073, 840, 1945, 0.043137256,

-0.4932206, 9.474146, 2.6025767, 0.065635026,
-0.02325625, -0.033681277, 577, 1712, 0.043137256,
-0.82602483, 9.478934, 2.6097596, -0.047375042,
0.049436715, -0.0013054191, 698, 2132, 0.039215688,
-0.52913475, 9.217958, 3.502824, -0.15122214,
-0.08128845, 5.2717666E-4, 652, 1898, 0.043137256,
-0.52674043, 9.407105, 3.0215747, -0.0638684,
-6.5423665E-4, -0.009246667, 837, 1677, 0.04705883,
-0.38108397, 9.347249, 3.1556542, -0.030270817,
-0.026310578, -0.0074140714, 524, 1848, 0.050980397,
-0.7541966, 9.282603, 2.535537, 0.2122427, -0.10450133,
-0.021463972, 645, 2085, 0.039215688, -0.61293435,
9.337671, 2.7677817, -0.019275242, 0.003010955,
0.0060249637, 224, 1849, 0.039215688, -0.50758624,
9.335278, 3.189174, -0.04920764, 0.029278161,
-0.0013054191, 1030, 1688, 0.023529414, -0.5411061,
9.227535, 2.9593236, -0.02477303, -0.0049302937,
-0.009246667, 681, 1855, 0.03529412, -0.6464542,
9.337671, 2.8180614, 0.052806858, 0.055545364,
0.007857559, 848, 1701, 0.04705883, -0.67279124,
9.337671, 2.4589202, 0.072965406, -0.034251824,
0.0023597723, 778, 2005, 0.039215688, -0.37829542,
9.555551, 2.5714512, -0.02477303, -0.03852788,
0.005414099, 388, 1948, 0.039215688, -0.56265455,
9.229929, 3.0886145, -0.02477303, -0.03852788,
0.005414099, 639, 2099, 0.043137256, -0.5913859,
9.335278, 3.2107224, -0.02599476, 0.03172162,
-0.00374888, 909, 1688, 0.04705883, -0.6249057,
9.380769, 2.8994668, 2.7244588E-4, -0.015925867,
0.001138042, 397, 1837, 0.043137256, -0.6440599,
9.325701, 3.0718546, 0.023485325, -0.031197498,

-6.945538E-4, 613, 2072, 0.03529412, -0.4932206,
9.428655, 3.3112822, 0.08945877, -0.014093272,
0.004192368, 66, 1731, 0.023529414, -0.74222517,
9.325701, 3.397476, -0.006447072, 0.08547776,
0.0017489073, 687, 1715, 0.043137256, -0.60335726,
9.4095, 2.9736893, 0.044254743, -0.04891259,
-0.004359745, 75, 1826, 0.027450982, -0.83081335,
9.490906, 2.2075214, 0.005159368, 0.09464074,
0.01824227, 410, 1747, 0.04705883, -0.38308397,
9.361614, 2.5499027, -0.09196821, 0.02439124,
-0.00374888, 488, 1707, 0.04705883, -0.5458947,
9.431048, 1.8435916, -0.043098986, 0.1270166, 0.019464,
254, 1989, 0.03529412, -0.46209505, 9.234718,
2.8515813, -0.024162164, -0.13993151, -0.082550496,
656, 2116, 0.039215688, -0.6057515, 9.282603, 3.047912,
-0.018664377, -0.029364903, -0.018409645, 521, 1710,
0.04705883, -0.46209505, 9.318518, 3.213117,
-0.043098986, -0.029975768, -0.011079263, 696, 1826,
0.043137256, -0.6177229, 9.107821, 3.4022646,
0.037535224, -0.36045384, -0.10515251, 271, 1676,
0.039215688, -0.6177229, 9.107821, 3.4022646,
0.037535224, -0.36045384, -0.10515251, 628, 2098,
0.039215688, -0.57462597, 9.210775, 3.790137,
-0.012555724, -0.018980194, 0.006635829, 407, 1848,
0.043137256, -0.5243462, 9.313729, 2.8803127,
0.109006464, -0.012871542, -0.012911859, 793, 1694,
0.043137256, -0.5506832, 9.318518, 2.753416,
0.028372247, 0.009730472, -0.0019162843, 555, 1995,
0.043137256, -0.59378016, 9.284998, 3.007209,
-0.13228531, 0.07020613, -0.00374888, 276, 1750,
0.039215688, -0.59378016, 9.284998, 3.007209,

-0.02599476, 0.03416508, 0.003581503, 648, 2124,
0.04705883, -0.41420954, 8.777411, 4.0295644,
-0.32104266, -0.03486269, -0.017187916, 656, 2037,
0.039215688, 0.11013664, 7.721536, 4.810098,
-1.8451515, 0.19115746, 0.26442096, 908, 1751,
0.039215688, -0.69912827, 8.078283, 4.889109,
0.5286709, 0.07142787, 0.019464, 331, 2004, 0.03529412,
-0.69912827, 8.078283, 4.889109, 0.5286709,
0.07142787, 0.019464, 932, 1962, 0.043137256,
0.26097596, 9.775824, 1.8292259, 0.9220681,
-0.17902689, 0.1654608, 989, 1942, 0.03137255,
-0.12210801, 9.122187, 2.152453, 0.058304645,
0.12848783, -0.10148732, 323, 1949, 0.04705883,
-0.13886794, 9.074302, 3.5435266, -0.004614476,
-0.037917014, -0.0013054191, 761, 1724, 0.03137255,
-0.13886794, 9.074302, 3.5435266, -0.004614476,
-0.037917014, -0.0013054191, 774, 1980, 0.039215688,
-0.21787901, 9.057542, 3.222694, 0.10289781,
-0.07273634, -0.0025271494, 531, 1858, 0.054901965,
-0.21787901, 9.057542, 3.222694, 0.10289781,
-0.07273634, -0.0025271494, 584, 2076, 0.039215688,
-0.27294734, 9.399923, 3.1221344, -0.0015601498,
0.05859969, -6.945538E-4, 129, 1715, 0.03529412,
-0.25379312, 9.325701, 3.0742488, -0.010112263,
-0.016536733, -0.0074140714, 796, 1714, 0.023529414,
-0.117319465, 9.364009, 2.9329867, -0.030270817,
-0.004319428, -0.011079263, 230, 1951, 0.039215688,
-0.2466103, 9.395134, 2.978478, -0.07486398,
-0.0030976976, -0.008635802, 66, 1838, 0.023529414,
-0.22506183, 9.337671, 3.0766432, -0.0015601498,
0.01422826, -0.006192341, 354, 1732, 0.03529412,

-0.11492519, 9.421472, 2.588211, 0.08212839,
-0.016536733, 0.0072466945, 302, 1884, 0.039215688,
-0.21069619, 9.440625, 2.691165, -0.031492546,
0.01522826, -0.013522724, 219, 1855, 0.039215688,
-0.27055305, 9.265843, 3.4477558, -0.080972634,
-0.0030976976, -0.010468398, 652, 2080, 0.039215688,
-0.25139886, 9.284998, 3.3232534, 0.026539652,
0.006676146, -0.0019162843, 389, 1875, 0.043137256,
-0.22745611, 9.356826, 3.047912, -0.0070579373,
-0.0030976976, -0.006192341, 706, 1720, 0.04705883,
-0.31604427, 9.229929, 3.2633965, -0.007668802,
-0.0024868324, -6.945538E-4, 842, 1960, 0.039215688,
-0.05506832, 9.318518, 3.1436827, 0.027150517,
-0.014093272, 0.0029706378, 1016, 1672, 0.023529414,
-0.16281068, 9.325701, 3.0910087, 0.028372247,
-0.005541159, 0.0072466945, 639, 2094, 0.043137256,
-0.19154198, 9.3280945, 2.8803127, -0.007668802,
5.674938E-4, -0.0025271494, 63, 1718, 0.019607844,
-0.07422252, 9.390346, 2.7390504, 0.013711481,
0.015839124, -0.005581476, 714, 1736, 0.043137256,
-0.07422252, 9.452597, 2.1716073, 0.04486561,
-0.033030093, -0.005581476, 819, 1726, 0.054901965,
-0.23942748, 9.308941, 3.3615618, 0.015544077,
0.007897877, 5.2717666E-4, 428, 1978, 0.039215688,
-0.19633053, 9.421472, 2.7845416, 0.0033267722,
-0.0018759671, -0.0019162843, 867, 1839, 0.039215688,
-0.14844504, 9.318518, 2.248224, -0.011944858,
-0.011038946, -0.006192341, 971, 1821, 0.03137255,
-0.18435916, 9.454991, 2.3200524, -0.0070579373,
-0.02081279, -6.945538E-4, 608, 1722, 0.050980397,
-0.071828246, 9.361614, 2.8994668, -0.07486398,

0.0024000895, 0.009690155, 662, 2099, 0.043137256,
-0.17238778, 9.330489, 2.7175019, -0.006447072,
-0.036695287, 5.2717666E-4, 1027, 1706, 0.019607844,
-0.18196489, 9.4142885, 3.0239692, -3.3841934E-4,
-0.011649811, -0.0068032066, 95, 1854, 0.027450982,
-0.08140534, 5.44302, 2.719896, 0.02287446,
-0.032419227, 5.2717666E-4, 443, 1979, 0.043137256,
-0.083799616, 9.26345, 2.9377751, 0.028372247,
-0.012871542, -0.009246667, 889, 1812, 0.03529412,
-0.083799616, 9.26345, 2.9377751, -0.007668802,
-0.01959106, -0.0068032066, 640, 2057, 0.043137256,
-0.12929083, 9.483723, 2.5044115, 0.016154941,
0.01644999, 0.0060249637, 837, 1963, 0.03137255,
-0.16759923, 9.397529, 2.7151077, -0.031492546,
-0.001265102, -0.0074140714, 582, 1739, 0.043137256,
-0.27294734, 9.378374, 2.954535, 0.013100617,
0.020115182, -0.0019162843, 673, 2108, 0.039215688,
-0.102953814, 9.615408, 1.5658557, 0.058915507,
0.26812646, 0.072609276, 647, 1948, 0.043137256,
-0.04070267, 9.411894, 2.8084843, -0.051040232,
0.018893452, -0.008024937, 889, 1702, 0.043137256,
-0.18196489, 9.270632, 2.8491871, -0.118846275,
-0.028143173, -0.017187916, 368, 1961, 0.039215688,
-0.2083019, 9.378374, 2.7342618, -0.0070579373,
-0.021423655, 0.004803234, 605, 2109, 0.039215688,
-0.083799616, 9.529214, 2.8515813, -0.03210341,
0.0268347, -0.005581476, 148, 1672, 0.043137256,
-0.11492519, 9.550762, 2.6217308, 0.076019734,
0.003010955, -0.004359745, 821, 1696, 0.043137256,
-0.052674048, 9.447808, 2.796513, -0.010723128,
-0.0049302937, -0.0074140714, 444, 1738, 0.04705883,

-0.052674048, 9.447808, 2.796513, -0.01316659,
-0.055021245, -0.018409645, 637, 1898, 0.043137256,
-0.21548474, 9.569917, 2.317658, 0.01920927,
-0.023567117, -0.0245183, 662, 2084, 0.03529412,
-0.14126222, 9.311335, 3.0694604, -0.09013561,
-0.022645386, -0.009857533, 399, 1853, 0.04705883,
-0.17478207, 9.37598, 3.0455174, 0.08335012,
-0.080677584, 0.009079291, 822, 1712, 0.04705883,
-0.14605077, 9.340066, 2.8994668, -0.051040232,
0.03172162, 0.004803234, 584, 1964, 0.039215688,
0.023942748, 9.380769, 3.0527003, -0.020496974,
0.02805643, -0.0074140714, 238, 1688, 0.039215688,
0.023942748, 9.380769, 3.0527003, -0.020496974,
0.02805643, -0.0074140714, 641, 2118, 0.027450982,
-0.34477556, 9.332883, 3.4166303, -0.051040232,
0.007897877, -0.0013054191, 319, 1863, 0.039215688,
-0.03351985, 9.447808, 2.5115943, 0.1804777,
-0.03608442, 0.006635829, 904, 1689, 0.039215688,
-0.031125572, 9.684841, 1.8172545, -0.007668802,
-0.01836933, -0.014133588, 241, 1969, 0.039215688,
-0.25858167, 9.433443, 2.2266755, 0.074187145,
0.034775946, 0.004192368, 230, 1749, 0.039215688,
-0.32801566, 9.438231, 2.6241252, 0.009435425,
-6.5423665E-4, 0.009079291, 747, 1998, 0.03137255,
-0.11253092, 9.402317, 3.083826, -0.07730744,
0.006065281, -0.012911859, 669, 2110, 0.03529412,
-0.26337022, 9.481328, 2.7031362, 0.0033267722,
-0.018980194, 0.0023597723, 958, 1830, 0.03529412,
-0.10774237, 9.356826, 2.9593236, -0.030270817,
-0.01042808, -0.011079263, 837, 1718, 0.039215688,
-0.021548472, 9.349643, 2.4804688, -0.024162164,

0.036608543, 0.0023597723, 59, 1706, 0.019607844,
-0.339987, 9.308941, 3.0383348, -0.035768606,
0.01522826, -0.004359745, 722, 1757, 0.043137256,
-0.25858167, 9.423865, 3.284945, -0.017442646,
-0.028143173, -0.004359745, 929, 1663, 0.039215688,
-0.34477556, 9.457385, 3.301705, 0.04364388, 0.0268347,
0.004803234, 391, 1766, 0.039215688, -0.09577099,
9.531608, 2.2865324, -0.016220914, -0.015925867,
-0.009246667, 693, 2062, 0.043137256, -0.12689656,
9.423865, 2.5235655, 0.010046289, 0.0024000895,
-0.006192341, 766, 1859, 0.03529412, -0.13886794,
9.541185, 2.5331428, -0.044320717, 0.011563068,
-0.00374888, 688, 1699, 0.043137256, -0.0670397,
9.273026, 2.6744049, -0.018053511, 0.031110756,
-0.010468398, 548, 1881, 0.043137256, -0.28970724,
9.438231, 2.7366562, 0.013711481, 0.09586247,
0.022518326, 237, 1833, 0.039215688, -0.16520496,
3.452597, 6.007209, -0.033325143, 0.031110756,
0.005414099]