

# Akshit Sharma

## Major

 Major\_Plug

 Major

 Jaypee University of Information Technology

---

### Document Details

Submission ID

trn:oid::1:3096909387

Submission Date

Nov 30, 2024, 12:42 PM GMT+5:30

Download Date

Nov 30, 2024, 12:46 PM GMT+5:30

File Name

G36plagcheck.pdf

File Size

2.2 MB

59 Pages

13,688 Words

77,758 Characters





# 6% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.




## Filtered from the Report

- Bibliography

## Match Groups

-  **70** Not Cited or Quoted 5%  
Matches with neither in-text citation nor quotation marks
-  **0** Missing Quotations 0%  
Matches that are still very similar to source material
-  **3** Missing Citation 0%  
Matches that have quotation marks, but no in-text citation
-  **0** Cited and Quoted 0%  
Matches with in-text citation present, but no quotation marks

## Top Sources

- 4%  Internet sources
- 4%  Publications
- 2%  Submitted works (Student Papers)

## Integrity Flags

### 0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

## Match Groups

- 70 Not Cited or Quoted 5%**  
Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%**  
Matches that are still very similar to source material
- 3 Missing Citation 0%**  
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**  
Matches with in-text citation present, but no quotation marks

## Top Sources

- 4% Internet sources
- 4% Publications
- 2% Submitted works (Student Papers)

## Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet		
	ideas.repec.org		0%
2	Internet		
	docplayer.net		0%
3	Internet		
	www2.mdpi.com		0%
4	Publication		
	Deepthi S, Mamatha Balachandra, Prema K V, Kok Lim Alvin Yau, Abhishek A K. "U...		0%
5	Publication		
	Le Sun, Yueyuan Wang, Yongjun Ren, Feng Xia. "Path signature-based XAI-enable...		0%
6	Internet		
	www.hindawi.com		0%
7	Student papers		
	University of Pretoria		0%
8	Publication		
	Emily Chia-Yu Su, Han-Ming Wu. "Dimension reduction and visualization of multip...		0%
9	Publication		
	"Touchstroke: Smartphone User Authentication Based on Touch-Typing Biometri...		0%
10	Internet		
	biomedia4n6.uniroma3.it		0%

11	Internet	vdoc.pub	0%
12	Student papers	American Public University System	0%
13	Student papers	Asia Pacific University College of Technology and Innovation (UCTI)	0%
14	Internet	zdocs.ro	0%
15	Internet	booksc.org	0%
16	Internet	pearl.plymouth.ac.uk	0%
17	Internet	123dok.org	0%
18	Internet	abuhamad.cs.luc.edu	0%
19	Internet	ijrpr.com	0%
20	Internet	ksascholar.dri.sa	0%
21	Publication	"An approach to cloud user access control using behavioral biometric-based auth...	0%
22	Publication	Haijun Bao, Minghao Yuan, Haitao Deng, Jiang Xu, Yekang Zhao. "Secure multipar...	0%
23	Student papers	Queen's University of Belfast	0%
24	Internet	research.rug.nl	0%

25	Publication	Jianfeng Guan, Xuetao Li, Ying Zhang. "Design and Implementation of Continuou...	0%
26	Student papers	The University of Buckingham	0%
27	Internet	export.arxiv.org	0%
28	Student papers	De Montfort University	0%
29	Publication	Mohammed Abuhamad, Ahmed Abusnaina, Dae Hun Nyang, David Mohaisen. "Se...	0%
30	Student papers	The British College	0%
31	Student papers	University College Birmingham	0%
32	Internet	digital-library.theiet.org	0%
33	Internet	discovery.researcher.life	0%
34	Internet	www.onlinescientificresearch.com	0%
35	Publication	Attaullah Buriro, Bruno Crispo, Yury Zhauniarovich. "Please hold on: Unobtrusive ...	0%
36	Student papers	Napier University	0%
37	Internet	managementpapers.polsl.pl	0%
38	Internet	e-space.mmu.ac.uk	0%

39	Publication	Muhammad Ehatisham-ul-Haq, Muhammad Awais Azam, Jonathan Loo, Kai Shua...	0%
40	Internet	core.ac.uk	0%
41	Internet	ensias.um5.ac.ma	0%
42	Internet	etheses.whiterose.ac.uk	0%
43	Internet	oa.upm.es	0%
44	Internet	personal.science.psu.edu	0%
45	Internet	scholar.archive.org	0%
46	Publication	"Advances in Network-Based Information Systems", Springer Science and Busine...	0%
47	Publication	"Exploiting Eye Tracking for Smartphone Authentication", Lecture Notes in Comp...	0%
48	Publication	Abdulaziz Alzubaidi, Jugal Kalita. "Authentication of Smartphone Users Using Beh...	0%
49	Publication	Priya Bansal, Abdelkader Ouda. "Continuous Authentication in the Digital Age: An...	0%
50	Publication	Attaullah Buriro, Bruno Crispo, Filippo Del Frari, Jeffrey Klardie, Konrad Wrona. "C...	0%
51	Publication	Shihong Zou, Huizhong Sun, Guosheng Xu, Chenyu Wang, Xuanwen Zhang, Ruijie ...	0%

## ABSTRACT

Today, devices like smartphones, tablets, smartwatches, laptops, and desktop computers have entirely replaced any primitive machines that people used to use in their everyday activities, and their use continues to grow day by day. With an increase in the use of these devices comes an increase in security breaches with both financial and emotional implications.

Many people know about hacking but would not definitely know what it is. However, the simple precautionary measures employed by users and service providers can avert a lot of incidents where user data gets compromised. One of the most important aspects in this regard is authentication. This refers to checking or proving that the user is the same person to whom the credentials that he/she uses belong and not someone impersonating him/her. In other words, actual authentication just means verifying someone's identity. Common examples include passwords and PINs, such as what ATMs use.

The aim of this project entitled "Authentication using Behavioral Biometrics" is to provide security and usability through an unobtrusive and effortless second-factor authentication approach. This means proving a user's identity by their unique behavioral characteristics based on how they hold a mobile device or how they interact with their device. This notion recognizes that every person has a distinct way of interacting with the devices—the grip with which they hold their phone, the pressure they apply while swiping or typing, and so on.

The project exclusively focuses on the behavior of mobile phone use and employs certain behavioral patterns, like how the user holds the device or moves it and how he or she swipes. It leverages the uniqueness of those traits to generate a very strong authentication mechanism.

The project also entails continuous authentication. That is, it not only implements the kind of static, conventional authentication such as passwords but also other forms such as biometric data, using which it continuously checks for the authenticity of a user. These authentication methods are hard to use social engineering to compromise or any other means.



# CHAPTER 1

## INTRODUCTION

### 1.1 INTRODUCTION

Authentication **is the** actual process of putting in place entry restrictions or bringing about authorization so that only the right person is allowed into a particular system or service. It can be considered as a necessity in today's world filled with security threats. To that regard, different types of authentications used by people today are passwords, fingerprints, pins, and one-time passwords. Yet, with improvement in technology, negative elements have also become intelligent and have made some sophisticated means to breach the above-mentioned methods using network attacks (like man-in-the-middle-attacks) as well as social engineering. We hear new scams every time wherein people are made to give their authentication credentials on the spur by pressure, greed, or fear.

Authentication usually falls into three broad categories:

- Something You Know

This includes using passwords or a PIN as an identity verification method. The individual validates identity by remembering a particular set of characters forming a specific passphrase or code. This is a popular method since it induces a low bar for entry; however, it is prone to theft. Leaked data are normally subject to replacement or modification, but if the data were previously abused, no remedy is available. Hence, vital yet non-exhaustive.

- Something You Are

Examples include fingers or faces. They refer to any specific biometric tool used for identification. They are part of the human body and cannot be stolen easily. But it is not immune to hackers who copy fingerprints or cheat on face



recognition systems. Biometric data are fixed, and when they are compromised, there is no alternative or simply no alternative to changing them.

- Something You Have

This criterion has to do with physical objects, including hardware security keys, smart cards, or tokens, which help prove a person's identity. This is most frequently used in high-security applications, such as using USB keys to access devices or the nuclear codes used by the world leaders. The one disadvantage, however, is that of theft, especially when there aren't any supplementary means that safeguard this physical item.

These authentication factors, notwithstanding their downfalls, are considered relevant in the protection of systems and services. Through understanding their strengths and weaknesses, one can build layered security strategies that allow risk mitigation.

Two of the basic factors in authentication are combined in most security-sensitive applications to improve protection. For instance, in ATM transactions, there is a physical card (something you have) paired with a PIN (something you know). Similarly, while signing up for important online services, one may have a password (something you know), and an OTP sent to a registered mobile device (something you have).

Of the three factors, "something you are" - like fingerprints or iris recognition—is typically deemed the strongest, as these traits are extremely difficult to duplicate.

A subclass of "something you are" includes behavioral biometrics, a little-explored but promising area. Behavioral biometrics involves analyzing and recognizing an individual's traits within behavior. It includes gait analysis, which studies how a person walks. Human behavior presents identifiable but subtle variations due to various reasons in other contexts. Some machine learning models could be trained to identify and learn these changes accurately, thus making them reliable as identification factors.

Unlike the imprinted biometrics like fingerprints which can easily be replicated using mold formation, behavior-oriented biometric systems cannot be replicated so easily. The

unique behavior pattern of an individual cannot be replicated by the exact model of behavioral patterns such as how they use a device or move through an environment.

Such biometrics are one of the only authentication mechanisms that can continuously authenticate a user. It is an authentication technique that periodically validates the user's identity after the point of entry, different from authentication techniques such as passwords, fingerprints, or PINs, which authenticate entry to the system only once. Thus, behavioral systems provide constant verification and, therefore, prevent risks such as takeover by an unauthorized user after initial access.

Another feature which makes the behavioral system interesting is the fact that it adapts with time to changes in user habits. So, behavioral systems evolve with the user, preventing the need for frequent updates and changing passwords or PINs that would normally require manual updates to improve security. All this comes at a time when simplicity and convenience are highly prized features in the present age, adding to the increasing attractiveness of behavioral biometrics.

## 1.2 PROBLEM STATEMENT

In this age of digitization, where almost every aspect of life revolves around technology, data security has become vital. Authentication is an integral process because it helps in safeguarding sensitive information; it is the process of verifying user's identity. Traditional authentication methods like passwords, PINs, even multi-factor authentication (MFA) would fail as techniques for application by the intruder and attackers are becoming advanced. Hackers sometimes depend on phishing and social engineering schemes, combined with new attack strategies, which always increase security breaches.

MFA has enhanced this by using passwords (something known) and OTPs (something owned device/ phone number/ email ID) as different factors, while at the same time it's causing a huge inconvenience for users. For instance, logging into an account requires several steps: password, a code sent to you, and verification through a device for a limited time. All of this proves futile when an ill-intentioned person knows or can tap into an already

opened account, or even forging cookies from websites. This creates loopholes in the effectiveness of these traditional ways.

To deal with such risks, service providers rely on backend monitoring systems, such as systems that try to detect and flag suspicious behavior through user activity patterns. These systems tend to consume high amounts of computing resources and require high investments in machine learning models. They also usually include real-time full rolling recordings of the users' activities, thereby bringing concerns to privacy. Because an end user would lose trust due to his fear of being spied on by a service provider, this could easily be rebutted because they would say that it is just for the enhancement of security.

Trade-off problems usually occur when security, usability, and privacy come into consideration. Most of the present authentication mechanisms impose an additional burden on the users by requiring them to perform additional steps for validating their identities and make use of an overhead resource consuming system which may impinge upon user's trust. Such circumstances have created a very urgent requirement for solutions that would be capable of meeting strong security, mild inconvenience to users, and privacy without incurring a resource burden on the provider.

Behavioral biometrics differ from conventional methods of authentication in offering a seamless, trustworthy, and privacy-preserving mode of collective authentication to resolve the challenges. Behavioral biometrics is different because while all standard authentication modalities authenticate an activity, we authenticate a user when using a device with behavioral biometrics. With the help of behavioral data from how users work, e.g., the way he or she holds his or her phone, swipe or touch gestures, it has created a completely unique digital fingerprint for every user.

This method would solve several problems by behavioral traits:

- Convenient for Users: Eliminates the hassle of repetitive two-factor authentication while effectively authenticating in the background.

- Service Provider Efficiency: Service providers would be able to validate users without necessarily having to monitor them extensively or spending resources on backend systems.
- Privacy Assurance: Traceability is based on physical behavioral data, not activities. Unlike traditional systems that track activities, behavioral biometrics take only behaviors as their data sources-and of course, these behaviors are not even usable for advertisers, thus respecting user privacy.

It contributes to the basic premise of behavioral biometric test: authenticate users, not their actions. That would require such a low effort from the users and yet ensure that only an authenticated person would be able to use his data.

It also decreases the extent of intrusive monitoring and will make the establishment of their association with a service provider more reliable. Behavioral biometrics use individualized behaviors, which are so rarely modeled by other humans, to establish identity as opposed to watching all users for security purposes.

### 1.3 OBJECTIVES

This project aims to create a second-factor authentication method for users and service providers that is unobtrusive and seamless. The aim is to cater for applications that require lots of sensitive data such as banking, online shopping, employee portals, and specialized services like parental control applications, as high-security applications, through unique human behavior.

The project aims to analyze a) user holding and movement of the phone during its usage, and b) individual behavior resulting from using this touch interface, e.g., swiping and tapping, which therefore included different subtle factors making it a unique trait of each individual. For example, height and hand size can influence how a phone is held, while habitual swipe or tap gestures are determined by personal preference, screen size, or even the past use of the same device or application.

In fact, one's lifestyle or activities help to shape these behaviors. For instance, a guitarist's muscle memory from practice might reflect in the way they use their phone; or a gamer who has been swiping continually might show different swipes. Such variations form the basis for a robust yet adaptable authentication system for the factor that makes these behaviors: something physical, habitual, and contextual.

The project, apart from being adaptive to changing behaviors over time, is naturally designed to enable the user to evolve their habits through changes in lifestyle, use of new devices, or other such reasons. Continuous updating of these natural patterns and their learning keeps the authentication type safe but user-friendly. This complex personalization along with adaptability would, in theory, make behavioral biometrics as much as, if not more secure than traditional biometrics like fingerprints or iris scans while being virtually impossible to copy or forge.

The project seeks to address the following key objectives:

**1. Develop a seamless second-factor authentication mechanism:**

Provide an additional layer of security for mobile applications without interrupting the user experience. This will be designed to run continuously in the background to ensure ongoing verification of the user's identity.

**2. Analyze behavioral patterns for unique authentication:**

It should include these two main data aspects- the way the user holds and moves their mobile phone while in operation, and how he performs touch interactions such as swipes, taps, and gestures.

**3. Include the diverse kinds of behavior in authentication:**

Difference in physiological characteristics and habitual preference as well as context during the usage of a device. For example, hand size differences, world activity classes performed (such as musicianship or gaming), and screen size would all result in behavioral differences.

#### **4. Enable defense against user behavioral changes**

Acclimatize the system gradually to the alterations in a user's habits over time so that hassle-free adaptation will be within the authentication measures and will still prove correct and reliable over time for the changing interaction patterns.

#### **5. Enable high-level security and tough forgery resistance**

The very complexity and subtlety of behavioral traits are what inform the authentication system's capability of being difficult to replicate, allowing for security levels on par with the traditional biometrics like fingerprints and iris scans.

#### **6. Customer usability and customer trust**

Strike a balance between privacy invasion and data misuse on the one hand and security and convenience on the other, thus providing a smooth user experience.

### **1.4 SIGNIFICANCE AND MOTIVATION OF THE PROJECT WORK**

Depending on digital devices and online services makes the lives of users convenient, but it also increases the risk of unauthorized access and security breaches. Traditional authentication mechanisms, while effective in the past, are now showing their limitations against the growing sophistication of malicious actors. Passwords, PINs, and even multi-factor authentication methods are often insufficient to counter evolving threats such as phishing, social engineering, and session hijacking.

In addition, these security measures create lots of inconvenience for users. Authentication processes like filling a password combined with checking an OTP within a limited time easily become bulky, and they end up failing under such extreme conditions where devices are logged into, and only unauthorized individuals access them. Perhaps a better, easy-to-use-yet-secure solution is called for here.

Behavioral biometrics offers promise as a solution. It measures more than the traditional spectrum of the user's access point: it authenticates the person through capturing unique behavioral patterns with respect to how they hold or interact with a device; continuous

authentication; and preemptive saving users the extent of multiple manual authentication repetitions. Being identified based on naturally occurring behavior makes it fall in the user's realm of habits and provides the solutions with an intuitive and friendly approach.

### **1. Evolved Threat Landscape:**

Cyber-attacks happen more often nowadays and are ever more sophisticated, and traditional authentication methods can no longer ensure security. Hence the motivation to develop a robust system for emerging challenges.

### **2. User comfort:**

The project eliminates the enhanced burdens of repetitive manual authentication and gives a ceaseless experience that is safe.

### **3. Continuous Authentication:**

Behavioral biometrics is the method of continuing verification, in contrast to conventional methods, to which a user attests once only. Thus, it offers safety even if a device is stolen or compromised after login.

### **4. Adaptation to User Behavior:**

With time, the way a user acts while interacting changes. Hence, one of the driving forces behind this project is the creation of an authentication system that adapts to behavior changes making it accurate and reliable in the long term.

### **5. Privacy Preservation:**

Our project revolves around finding solutions that enable increased security along with guaranteeing privacy in order to build a trust platform between users and service providers.

### **6. Resource Efficiency for Service Providers:**

Behavioral biometrics are both economical and scalable for service providers because they do away with the necessity of monitoring back-end activities and require very little heavy resource-demanding systems for effects of anomalies.

## 1.5 ORGANIZATION OF THE PROJECT REPORT

There are six main chapters in this report as far as the project is concerned. Below is a summary of the same in place of those chapters:

### **Chapter 1 - Introduction**

This chapter is about the project in general and even indicates the problems posed by traditional approaches and how they can be redeemed by means of behavioral biometrics. It contains the defining problem statements, their objectives with respect to the significance and motivation for the work done, and lastly, it introduces the rationale for using behavioral biometrics as a credible solution to authentication.

### **Chapter 2-Literature Survey**

This chapter contains the entire literature we researched for this project, including previous authentication works, biometric techniques, and everything done in advances along with studies into behavioral biometrics. The stress was primarily focused on research done in the last ten to fifteen years and the identification of gaps in existing methods that this project intends to fill.

### **Chapter 3: System Development**

The project has been extensively designed and developed as mentioned in this chapter. It consists of the requirements analysis followed by project design architecture in addition to the explanation on ways of data preparation. It also illustrates implementation aspects by showcasing algorithms, tools and techniques used, while addressing challenges of system implementation during development and the solutions adopted to address those issues.

### **Chapter 4: Testing**

This chapter shows project testing techniques applied during project inception. It discusses generally the testing strategies used, prepared test cases, and their results. To evaluate this system's efficiency, extensive stress testing in various situations was conducted.



## **Chapter 5: Results and Evaluation**

This chapter presents the results obtained from the behavioral biometrics system with an elaborate examination of the real results obtained vis-a-vis interpretation thereof as well as comparison against currently available solutions. The pros and cons of the proposed approach are also mentioned.

## **Chapter 6: Conclusion and Future Scope**

The project gets pulled to a finish here, most summarily, and presented are all the findings and contributions of the work to the authentication and data security domains. And, further, it highlights the limitations of the work and places forward the future scope for further development and research into behavioral biometrics.



## CHAPTER 2

### LITERATURE SURVEY

#### 2.1 OVERVIEW OF RELEVANT LITERATURE

##### 2.1.1 INTRODUCTION

The body of this literature review on behavioral biometrics includes studies focusing on its application in continuous and non-intrusive authentication of users. Behavioral biometry detects one's identity by unique human features and conditions such as patterns of interaction with the device and motion behaviors. It incorporates machine learning, human-computer interaction, and cyber security.

Many studies explore the issues and opportunities with behavioral traits authentication. For example: One area focuses on using touch interaction data from 'moving' motion sensors to identify behavioral features in humans and their invulnerability to other forms of attack, such as password phishing. Other science works include the application of detecting movement based on the accelerometer and gyroscope to create user movement and conclude the actual possibility of strategizing around continuous and adaptive authentication systems.

There is also a growing discussion on privacy and data security, with mentions of anonymized and encrypted methods of data gathering.

Finally, the literature discusses advanced and feature extraction and machine learning algorithms that have been developed for use in behavioral biometrics. Such advances include development scientific studies, such as from simple statistical models to advanced deep learning architecture, which may respond to dynamic behavior changes of users.

S. No.	Author & Paper Title [Citation]	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Key Findings/ Results	Limitations/ Gaps Identified
1.	O. L. Finnegan <sup>1</sup> , et al. The Utility of Behavioral Biometrics in User Authentication and Demographic Detection: A Scoping Review	Systematic Reviews (2024)	NA	Highlighted touch gestures and motion as the most used methods for authentication. Provided insights for child screen time studies.	Low study quality (average score 5.5/14); most studies focused on adult populations, limiting applicability to children
2.	Le SUN, Yueyuan WANG, Yongjun REN & Feng XIA Path Signature-Based XAI-Enabled Network Time Series Classification	Science China Information Sciences (2024)	5G-NIDD, INTF, ISCX_app, ISCX_tra, TSD1, TSD2	Improved explainability and accuracy in classifying network time series data for network automation. Outperformed existing methods.	Limited to network data; potential challenges in adapting to behavioral biometrics or non-network domains.
3.	Emily Chia-Yu Su, Han-Ming Wu Dimension Reduction and Visualization of Multiple Time Series Data: A Symbolic Data Analysis Approach	Computational Statistics (2024)	NA	Enhanced visualization of temporal trajectories for multiple time series, effective for short datasets in bioinformatics and finance.	Assumes predefined intervals; challenges in extending to irregularly sampled or highly dynamic time series data.
4.	Priya Bansal and Abdelkader Ouda Continuous Authentication in the Digital Age: An Analysis of Reinforcement Learning and Behavioral Biometrics	Computers (2024)	NA	Achieved high training accuracy (94.7%-100%) and test accuracy (81.06%-93.5%). Low Equal Error Rate (0-0.11) across sessions	Focused only on keystroke dynamics; applicability to multimodal behavioral biometrics is unexplored.
5.	Yris, P., et al. Deep features fusion for user authentication based on human activity	IET Biometrics (2023)	UCI-HAR (consisting 30 subjects)	A new approach of signal-to-image transformation was introduced where time sensor data were converted into 2D images, enabling use of deep learning architectures.	The dataset was yet limited. Only the way user would use smartphone device in a stationary position which is not very true to the nature.

21	6.	Praveen, R., et al. <b>Continuous user authentication on smartphone via behavioral biometrics: a survey</b>	Springer Nature <b>Multimedia Tools and Applications (2023)</b>	NA	Gives an idea over different datasets and features to be used. Reviews possible attacks on behavioral biometrics.	Theoretical approach of the concept with the deep practical knowledge yet to be tested.
34	7.	Stragapede, G., et al. <b>Mobile behavioral biometrics for passive authentication.</b>	<b>Pattern Recognition Letters 157</b> - Elsevier (2022)	HuMldb Database	Explored feasibility of unimodal and multimodal biometric traits. Implemented these using recurrent neural networks (RNNs).	When touch stroke and magnetometer data are combined, the model shows discriminative behaviors and accuracy drops
4	8.	Shihong, Z., et al. <b>A Robust Continuous Authentication System Using Smartphone Sensors and Wasserstein Generative Adversarial Networks</b>	Communication <b>Security</b> in Socialnet-Oriented Cyber Spaces (2021)	HOMG	Usage of WGANs (Wasserstein Generative Adversarial Networks). Providing comparison among the classifiers.	Lack of real-world testing makes it limited for the usage. Especially not being able to extensively consider challenges like sensor noise or unpredictable user.
42	9.	Lavanya, B., et al. Impact of Behavioral Biometrics on Mobile Banking System	Advances in Computational ECE (2021)	NA	<b>Provided a better overview of the types of</b> attack and how <b>this</b> can be used in the banking sector	Centrally worked only around biometrics such as eye movement and fingerprints hand geometry etc.
6	10.	Jianfeng, G., et al. <b>Design and Implementation of Continuous Authentication Mechanism Based on Multimodal Fusion Mechanism</b>	Communication <b>Security</b> in Socialnet-Oriented Cyber Spaces (2021)	15000 mouse operations	Proposed MFCA system for behavior collections. Also provided with the exact features to look into when applying the classification model.	Limited behavior types, major focus was only <b>on keystroke, mouse movement and application usage.</b>
23	11.	Ioannis, S., et al. <b>Behavioral biometrics &amp; continuous user authentication on mobile devices: A survey</b>	<b>Information Fusion (2021)</b>	NA	Gave background and information regarding sensors. Provided with the current challenges and the machine learning algorithms used and their limitations.	A more theoretical approach than a practical one.

36	12.	Abuhamad, M., et al. Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey.	IEEE Internet of Things Journal (2021)	NA	Explored motion based, gait based, keystroke dynamic based and touch based methods. Sensor and orientation of a device held by a user also one of the methods.	A survey paper that puts into account major works done in the field and the future aspects instead of providing with any new approach.
13	13.	Alsaadi, I. Study on most popular behavioral biometrics, advantages, disadvantages and recent applications.	International Journal of Scientific & Technology Research (2021)	NA	Goes into details regarding different behavioral biometrics techniques, including voice recognition, gait recognition and keystroke dynamics.	Only a theoretical approach over the concept of behavioral biometrics.
24	14.	Tian Linger Xu, et al. Finding Structure in Time: Visualizing and Analyzing Behavioral Time Series	Frontiers in Psychology (2020)	MATLAB	Introduced methods to analyze temporal patterns in high-density multimodal behavioral data. Demonstrated utility in psychological studies.	Limited scalability to very large datasets; primarily designed for Matlab users.
20	15.	Alzubaidi, A., & Kalita, J. Authentication of smartphone users using behavioral biometrics.	IEEE Communications Surveys & Tutorials (2016)	NA	Aimed to study the existing studies and approaches in the field of behavioral biometrics and different types of behavioral authentication.	More of a review paper so instead of learning a new methodology it was a survey of other research papers.
35	16.	Buriro, A., et al. Touchstroke: Smartphone user authentication based on touch-typing biometrics.	Springer International (2015)	12 Subjects 30 Patterns 2160 Samples	Takes into account two human behaviours: how the phone is held and how a user enters a 4-digit PIN	Limited number of users makes it have a bias over
19	17.	Zheng, N., et al. You are how you touch: User verification on smartphones via tapping behaviors.	IEEE (2014)	80 Subjects 25 times per subject	Proposed a mechanism called non-intrusive user verification.	Did not take into the user's behavioral change in time.

18.	Li, L., Zhao, X., & Xue, G. Unobservable re-authentication for smartphones.	IEEE (2013)	75 Subjects	Analyzes and learns user finger movement patterns, running in the background to monitor and compare these movements to the owner's patterns for continuous verification without user intervention.	Even though there was a higher accuracy in their first testing database but then they realized that for a higher dataset the accuracy was reaching to a threshold with increasing size.
19.	Zhao, X., et al. Continuous mobile authentication using a novel graphic touch gesture feature.	IEEE (2013)	30 Subjects 300 Touch Gestures	Utilizes Android's standard API to capture and observe touch gestures. Upon screen contact by users, it records traces as raw touch samples.	Limited size of data and the users only performed 6 most commonly used touch gestures over a long period of time hence there was a lack of consideration of outlier. And the device used was also the same.
20.	De Luca, A., et al. Touch me once and I know it's you! Implicit authentication based on touch screen patterns.	ACM (2012)	26 Subjects (Lab Test)  645 valid authentications  2790 Attacks	Conducted two studies: a short-term lab study focused on basic security techniques, like using a stroke, and a long-term study that employed password patterns with passive authentication.	Lack of attackers for some of the participants is undesirable since the data collected was lab tests and it was basically entering the pattern, so it does lack the main purpose which was to provide continuous authentication.
21.	Saevarane, H., & Bhattachakosol, P. Authenticating user using keystroke dynamics and finger pressure.	IEEE (2009)	10 Subjects  3000 values of finger pressure	Implemented a method, which detects and recognizes users in terms of hold-time, inter-key and finger pressure. This used both keystroke and touchscreen patterns.	Not a very good approach since the results rejected legitimate users with high probability and led to constant re-authentication of the user.

Table 2.1 : Literature Review

## 2.2 SUMMARY AND KEY GAPS IN LITERATURE

**O. L. Finnegan<sup>1</sup>, et al. [1]** The scoping review talks about **the current state of the art** behavioral biometrics **for user authentication** and demographic feature detection, stressing its potential applications in screen time research. The behavioral biometrics use the built-in smartphone sensors so as to provide continuous, non-intrusive user authentication without any requirement of additional hardware. The paper provides with the codes and programming environment so as any novice working in this direction is able to grasp the foundational methods and working easily. The dataset though here is in sense artificial majority of which is conducted among adult population in controlled lab settings, which limits the focus on children or a diverse population. Therefore, even the accuracy achieved is very high i.e. Touch Gesture-62%, Motion Behaviour-52% and keystroke dynamics-25%, this cannot be called upon the true accuracy of the methodologies proving to be in a low average study quality. Overall, the paper puts our research on behavioral biometrics into a new direction of public health context such as tracking the screen time usage.

**Le SUN, Yueyuan WANG, Yongjun REN, Feng XIA [2]** This paper proposes recursion which is a novel **path signature-based** model for **network time series classification**, specifically modelled for cybersecurity and automated network management. It combines Explainable Artificial Intelligence (XAI) techniques to address the difficulties in feature extraction and model interpretability alongside improving the accuracy of classification. The framework incorporates two modules:

- 1. Verbe:** which is a data augmentation module that uses 1D-CNN to maintain and improve data flow.

- 2. SigRNN:** which is feature extraction module which uses path signature to achieve explainability and efficiency.

This framework is designed to manage data imbalance and simplify very complex data. Along with that it transfers the time-series data to interpretable paths and signature that enables insights to the temporal dynamics and decision-making process.

8

8

30

1

**Emily Chia-Yu Su, Han-Ming Wu [3]** This research introduces the Path Point Approach (PPA) which is novel symbolic data framework for the reduction of dimensions and visualization of multiple time series data. This approach help improve the state-of-the-art methods like Principal Component Analysis (PCA) and Sliced Inverse Regression (SIR) with the help of resenting the time series as geometric path point, that enables a better exploration and summarization of temporal dynamics. The PPA converts multiple time series data into time-dependent intervals which are marked by a starting and ending points(values), visualizing the segments in 2d and 3d space.

3

**Priya Bansal and Abdelkader Ouda [4]** This analysis proposes a Reinforcement Learning (RL) based model for continuous authentication, which combines behavioral biometrics like keystroke dynamics with anomaly detection. The method uses, Markov Decision Process (MDP) framework and Double Deep Qc-Network(DDQN) for improving the security by continuously verifying the behavior of user during sessions. The main point of this includes extraction of dynamic data, and summary features from behavioral data that is going to serve as the input to the RL model. Here the system surpasses traditional supervised learning methods, achieving test accuracy if 81.06% and low Equal Error Rate of 0.323. The study also mentions the adaptability of RL models, allowing them to flexibly change with evolving user behaviors, hence making them well placed for the real-world applications. For future work, the paper suggests integrating autoencoders for advanced feature extraction, exploring data augmentation techniques and optimizing system for real-time scalability.

39

**Yris, P., et al. [5]** The paper introduces a creative approach for user authentication using deep feature extraction based upon the human activity data. The focus here is on the continuous authentication through behavioral biometrics. The research transforms the time-series data into a 2-Dimentional images using a reversible single-to-image transformation technique. This in turn enables the application of pre-trained deep learning models like ResNet-101, Shuffle Net, Google Net and DarkNet-53. The data set used, i.e. UCI-HAR, contains accelerometer and gyroscope data for six physical activities which are: walking, sitting, standing, laying, walking upstairs and walking upstairs.



44

The results showed that Shuffle Net achieved the lowest Equal Error Rate (EER) of 3.58% when multiple activities were combined, showing exceptional accuracy compared to its peer methods. The study highlights the potential of Deep Learning in enhancing security for continuous authentication, emphasizing its ability to improve performance by merging different activities and using a diverse dataset. The future goals of the paper would be centered around refining data augmentation methods, creating innovative attack simulations and expanding the usability of biometric system for enhanced security in real-world applications.

**Praveen, R., et al. [6]** The research presents a comprehensive survey for continuous user authentication technique that can be used on smartphones, revolving around the concept of behavioral biometrics. The paper reviews the methodologies and challenges tied with the unimodal and multimodal authentication systems. These systems extract and use features such as touchscreen patterns, keystroke dynamics, motion sensor data, and behavioral profiling. The survey provides an insight into the publicly available datasets, state-of-the-art models, and performance evaluation techniques. This all emphasizes the significance of continuous monitoring to identify outliers and vulnerabilities in entry-point authentication methods. The main results stresses upon the effectiveness of multimodal approaches in enhancing accuracy and robustness against attacks. It also discusses the challenges faced in the journey like data set limitations, energy efficiency and system scalability. Future research includes exploring adaptive models, enhancing data privacy, and tackling security concerns to develop reliable, real-world implementation of continuous authentication models.

45

**Stragapede, G., et al. [7]** The study inquiries about the use of mobile behavioral biometrics for passive authentication, embarking on the restrictions of traditional methods such as fingerprints, PINs and iris sensing (also called facial recognition). It evaluates the unimodal and multimodal approaches that have been focused on the HuMIdb database, which is a comprehensive dataset that holds user interactions, such as typing, scrolling and tapping, as well as background sensors that are, accelerometer, gyroscope, magnetometer etc. This research implements LSTM RNN models with a triplet loss function:

$$\mathcal{L}_{TL} = \max \{0, d^2(\mathbb{V}_A, \mathbb{V}_P) - d^2(\mathbb{V}_A, \mathbb{V}_N) + \alpha\}$$

for individual modalities and use score-level fusion for multimodal systems. Main outcomes points that the keystroke dynamics and magnetometer data are the most distinct features, that achieved the Equal Error Rates (EER) of 4-9% when fused together. The results highlight the benefits of multimodal fusion in increasing authentication accuracy while maintaining the usability of the modals. The paper also pinpointed couple of challenges that include balancing the device and user-specific features and acknowledging the issue of data scarcity. The future goals aims to improve the methods of data collection as well as integrating synthetic data to enhance the robustness of the model. This paper spotlight upon the potential of behavioral biometrics as a secondary to complementary factor in mobile authentication systems.

**Shihong, Z., et al. [8]** The paper presents a way of continuous authentication system while utilizing the technique of Wasserstein Generative Adversarial Networks (WGAN) for the augmentation of data in order to improve the robustness of sensor-based authentication systems on smartphone devices. The method leverages motion sensors like accelerometer, gyroscope and magnetometers to capture the user behaviors in activities such as reading, writing, and navigation. Main features include:

1. The generation of realistic synthetic sensor data by WGAN, this in turn addresses the issue of data scarcity.
2. Using a Convolutional Neural Network (CNN) to study and extract the deep features from sensor data.
3. Utilizing four different classifiers: Random Forest, One-Class SVM, Decision Tree, and K-Nearest Neighbors to train on the extracted features.

The modal created has been evaluated on the HMOG dataset. The results of which achieved the Equal Error Rates (EER) between 3.68% and 6.39%. This all showed its effectiveness in various or diverse range of user activities. Future goal aims in enhancing the diversity of data along that the quality of synthetic samples. Along with that improving the performance of the model for scenarios that doesn't share the limelight. The study calls for the potential of collaborating deep learning along with adversarial networks for the effectiveness of smartphone authentication

**Lavanya, B., et al. [9]** The study presents the integration of behavioral biometrics into mobile banking systems in order to highlight the increasing threats of fraud and enhancing the user authentication. This all based upon the unique human actions like keystrokes, the rhythm of typing, scrolling patterns and other interactions, that enables the continuous and non-intrusive user verification throughout one session. This paper spotlights on the advantages of behavioral biometrics over the traditional methodologies like passwords, fingerprints and PINs, stressing over it's ability in providing a frictionless as well as a secure and a wholesome user experience. The patterns of behavior are to be monitored in real-time to identify outliers and anomalies so as to reduce fraudulent activities. Implementing these systems can reduce the user hesitancy regarding mobile banking adoption by increasing security and usability. The research brings challenges like spoofing attacks, privacy of data and the need for an accurate anti-spoofing method. Future work points on improving machine learning algorithms, that integrate behavioral biometrics with a multimodal authentication system and acknowledging scalability for broader deployment in the banking and finance. Behavioral biometrics have been presented as a transformative solution for a secure, seamless and user-friendly digital transactions.

**Jianfeng, G., et al. [10]** The paper proposes a Multimodal Fusion Continuous Authentication (MFCA) system, this combines multidimensional behavioral biometrics, this includes keystroke dynamics, mouse movements, and patterns of application usage, in regard to the continuous identity verification. Unlike one-time authentication, this approach ensures continuous user validation, looking out for unauthorized users when detecting anomalous behaviors. The MFCA system utilizes a trusted model to test the user behavior over time, which dynamically updates trust scores so as to determine user legitimacy. The main features of this model are:

1. A combination of keystrokes, mouse and application usage data using a parallel multi-classifier mechanism.
2. Tracking and adjusting the extracted trust scores based on behavior alignments, triggering alerts when scores doesn't meet the threshold.
3. Employing tailored recognition models for each modal, like, Decision Trees for keystrokes, SVM for mouse movements and Naive Bayes for application usage.

The key results showcase high accuracy and timely detection of unauthorized users. Future works focus on expanding the datasets, improving the capabilities of real time, and

strengthening the scalability of model for diverse user group. This research tells about the potential in the multimodal fusion for enhancing security along with maintaining user convenience at the same time.

**Ioannis, S., et al. [11]** The research is all about the combined study of behavioral biometrics and continuous authentication technologies for mobile devices. There are seven different modalities in which the behavioral biometrics is based upon: touch gestures, walking gait, keystroke dynamics, behavioral profiling, hand waving, power consumption and multimodal fusion. The study highlights their application for seamless, real-time user verification throughout the usage of mobile device sessions. This review inspects methodologies for the collection of data, extracting features and the performance matrices of the machine learning models across these modalities. It also directs toward the adversarial attack vectors that target the Behavioral Biometrics and Continuous Authentication along with proposing some countermeasures to enhance the robustness and the security. Some of the key findings highlighted here are:

1. Combining multiple biometrics consistently helps in improving the authentication accuracy.
2. Machine Learning models like SVM, LSTM and Random Forests have been found to work efficiently across all modalities used.
3. Continuous Authentication systems are useful after the initial login process since they monitor the ongoing behavior of user which offers a much more enhanced security in comparison with the traditional authentication methods.

There are some challenges which have been consistent among the papers reviewed like:

- a) The limitation in the amount of publicly available dataset for the purpose of research.
- b) The limited addressing in the variability of user behavior among contexts.
- c) Privacy concerns computational compromises in order to implement the Behavioral Biometrics and Continuous Authenticated systems.

The future goals include, creating standardized metrics, expansion of datasets, and integrating stronger privacy-preserving mechanisms. This paper provides valuable knowledge and works in the direction of behavioral biometrics and continuous authentication technologies, proving the insights over the foundational structures to work on along with remapped challenges one might face in the journey.

**Abuhamad, M., et al. [12]** This survey comprehends sensor-based continuous authentication methods using behavioral biometric on smartphone devices. In these papers around 140 studies have been summarized. This study primarily focusses on six biometric modalities: motion, gait, keystroke dynamics, touch gestures, voice and multimodal systems. Leveraging the sensors which have been embedded in smartphones i.e. accelerometer, gyroscope etc., this research provides study over how to enable transparent, non-intrusive user authentication keeping in mind the convenience of user. Some of the notable conclusions drawn in this review are:

1. The approach where we combine multiple modalities provides the best accuracy results and are found to be resilient towards adversarial attacks since it combines various biometric features.
2. Continuous Authentication provides a more secure environment for smartphones since it monitors behaviors throughout a session, unlike the traditional methods relying on one-time login.
3. Most common metrics used as benchmark for accuracy are False Acceptance Rate (FAR) and Equal Error Rate (EER).

A couple of challenges have also taken a spotlight here, like:

- a) Acknowledging privacy concerns sensor-based attacks.
- b) Handling of high computation and memory overhead on a real-time basis.
- c) Limited availability of the datasets for large scale and diverse user population.

Future plans are focused on improving multimodal systems, developing a mechanism to preserve privacy, and optimizing resource efficiency for a large scale deployment. This paper puts behavior biometrics as a foundation of secure and seamless authentication for the future of mobile ecosystems.

**Alsaadi, I. [13]** This review gives a thorough analysis of behavioral biometric technologies, featuring their advantages, limitations and the ways they can be used. Taking advantage of distinct human behaviors like voice patterns, gait, keystrokes and signature recognition, behavioral biometrics authenticates a user. Which is in pretty contrast with the conventional methods we use right now like passwords and PINs. Behavioral Biometrics provides subtle,

seamless and more secure techniques since they are difficult to copy and are being monitored for a session. The main outcomes that we draw from this paper are:

1. Keystroke dynamics and gait shows a promising security method for application since it has low hardware requirements and has better ability to operate in real-time.
2. voice is effective but these can be influenced easily by factors like sickness and aging.
3. Signature recognition, though widely used, lacks consistency when one's signature changes over time.

The challenges faced here are:

1. the reliability of the technique is based on the consistency in the user behavior, if it finds any outlier the algorithm might collapse.
2. The privacy concern attached to tracking and the database highlights the privacy concerns of the technique.

Some applications that we found in the paper were:

1. These can be used in departments like online banking, surveillance and forensic investigations.
2. some of the latest implementations will be voice-controlled smart devices, real-time gait recognition, and mobile based continuous authentication systems.

Future prospects that we saw were further research into multimodal systems, to improve the diversity in dataset, and developing a privacy-preserving mechanisms in order to increase the efficiency and usability in real-world applications.

**Tian Linger Xu, et al [14]** This paper presents some practical tools and techniques for analyzing high density behavioral time series data, focusing on challenges posed by the multimodal and dynamic nature of human behavior. The study provides four different modules for visualizing and analyzing temporal patterns, focusing on methods like burstiness calculation-which is used to quantify temporal regularity, classifying behaviors into periodic, random or bursty- , cross-recurrence quantification analysis-a method used for the identification of recurrent patters and asymmetric in dyadic interaction across the time scales-, and granger causality- which examines directional influence among interdependent variables in multimodal datasets. The study provides useful insights over as to some

methodologies for the analysis of a time series data and hence finds importance in frameworks that are used for uncovering meaningful patterns and relations in high density temporal data.

**Alzubaidi, A., & Kalita, J. [15]** This extensive paper reviews behavioral biometrics as a pillar for continuous authentication on mobile devices. The primary focus in this is on seven biometric modalities: handwaving, touchscreen, gait, voice, signature and profiling. These methods gives an ongoing user authentication and validation with the help of analyzing unique behavioral patterns which have been extracted from device interactions. Key finding provided are:

1. inclusive methods to evaluate as the existing studies are based on the size of data, the type of classifiers use and the matrices of preformation.
2. Major techniques that provide the best results are support vector machines (SVM), neural networks and statistical models for the identification of user.

Some challenges faced were:

1. The variation in the behavior of the user affects the performance.
2. The requirements for the computation are too high for a real-time system.
3. The concerns related to privacy and the spoofing threats remain unsolved.

Furthermore, the study recommends improving the diversity of the dataset, creating a technique so as to optimize energy and efficiency for smartphone usage.

**Buriro, A., et al. [16]** In this paper we present Touch stroke, which is a novel bi-modal biometric authentication system used for smartphones. It uses touch-typing patterns and the behavior associated with the movement of phones. Leveraging the built-in smartphone sensors such as accelerometer, gyroscope, and magnetometer, in combination with keystroke dynamics, the system is able to enhance the accuracy of user authentication keeping in mind the ease of use of the security technique. The main points here would be as:

1. The data has been collected for six different scenarios from 12 users, which captures the timing of touch and typing alongside the patterns drawn from the movements.

2. Then the features have been statistically extracted i.e. in form of mean, standard deviation, skewness and kurtosis from the sensor data fused with keystroke timing features.
3. The fusion of motion sensor with touch proved an improved classification accuracy and a pretty strong foundation against mimicry.

The evaluation matrices are based upon the True Acceptance Rate (TAR), False Rejection Rate (FRR) and False Acceptance Rate (FAR). Variability in user behavior from holding the phones and typing styles along with the limited availability of dataset proves to be the biggest challenges here. The paper also proposes Behavior Biometrics to be applicable as primary or supplementary authentication mechanism for improved smartphone security and offer a seamless integration with existing devices without the need of any additional hardware. For future goals, the paper suggests towards the expansion of the dataset, so as to include more users and activities, and explore more sensor fusion techniques for improved performance and scalability.

**Zheng, N., et al. [17]** The study recommends non-intrusive user verification mechanisms for smartphone devices using the behaviors of tapping. Main aim her is to enhance the security of passcode-based authentication. By studying the tapping patterns, including acceleration, pressure, size, and time, the system can identify whether the authenticating user is legitimate owner of device or an imposter. This study is done by over 80 users, collecting data on 4-digit and 8-digit PINs, which is then applied to a one-class learning classifier for the verification and validation of users. The lowest Equal Error Rate (EER) achieved was 3.65% which is one of the most accurate matrices for the system, contributing to the fact that features like acceleration and pressures are highly effective in distinguishing the users. With these key findings the approach is able to integrate seamlessly with the existing passcode techniques without the requirement of any further hardware. This all ensures transparency and the ease of use for the technique. Though dure to the limited datasets, various distance and uncommon features for authentication have not been able to be added in the list of features and are to be termed as outliers, this does provide a big issue in improving the technique and providing the method as a secondary way of authentication, as the application of the proposed model is recommended. Therefore, the paper suggests addressing issues like user behavior variability, and optimization of classifier for real-time application.



**Li, L., Zhao, X., & Xue, G. [18]** The research instigates a continuous and unobtrusive re-authentication mechanism for smartphones, so as to enhance the security but not disrupt user experience. The system uses behavioral biometrics, like grip patterns, touchscreen interactions and device movement, to verify a user's identity during normal usage. By taking in the built-in smartphone sensors, the method ensures passive monitoring and removing the need for explicit authentication actions like entering passwords or PINs. The framework is designed to detect unauthorized access seamlessly in real time by monitoring natural user behavior. This behavior data is collected through various sensors and the evaluation matrices is used for accuracy among multiple scenarios so that we get to know what works best to pass the threshold of security. The potential scenarios where this technology can be used is mobile banking, sensitive data access and enterprise applications. Offering a scalable and user-friendly alternative to the old and traditional static authentication methods.

**Zhao, X., et al. [19]** This study presents a state of the art approach for **mobile user authentication using Statistical Touch Dynamics Images(STDI)**, using behaviors like touch gestures and tactile pressure dynamics. STDI extract feature instances by learning variations of these behaviors person to person, and acknowledging the challenges of computational efficiency and variability in touch behaviors. The STDI framework is a new method that combines movement dynamics and tactile pressure data in order to represent touch gestures visually in 2D, it reduces the computational overhead compared to previously used methods like Graphic Touch Gesture Features(GTGF). This all is evaluated on the UH-TOUCH dataset, which is a collection of 78 subjects across six touch gestures like slide, pinch, zoom etc. The results showed the low Equal Error Rates(EER) of 9.7% for verification and high Recognition Rates (RR) of 82.3% demonstrating improved performance over GTGF **and other state of the art methods**. With **the** results it showcased we can see that this method is suitable for continuous authentication, promising a secure and seamless user verification and validation on smartphones and can be used in both single-user and multi-user environments. Even though with these accuracy the method still limits itself in portions like the variation in touch behaviour due to different device handling style and computational complexity for large scale datasets. This provied us with the future goal to improve the real-time processing and enhancing usability in practical applications along with that studying further potential in

STDI and integration of additional modalities increasing the accuracy and complication for mimic attacks.

**De Luca, A., et al. [20]** This paper presents an implicit authentication mechanism for smartphones enhancing the security of password patterns by analyzing how each user input their patterns on a touchscreen. The system here collects data such as pressure, speed, size, and coordinates during the pattern input, using Dynamic Time Warping (DTW) to compare user behavior against the stored reference state. In this paper another layer of security is added which is how the pattern is drawn instead of just the correctness of the pattern. Under this there were two different studies held, one where user need to unlock study with four gesture types and the other a password pattern study in the realistic scenario, these were conducted so as to access the effectiveness of the system. One of the challenges the paper faced was the issue of high false acceptance rates which was up to 50% in some configurations and how the user variability over time affects the recognition performance of the system. This all proposes increasing the accuracy with dynamic reference sets and exploring some Machine Learning methods along with conducting long-term studies to validate the effectiveness and the strength of the system. Even though with the challenges and limitations the system still demonstrated promising environment for enhancing mobile security with minimal user inconvenience.

**Saevanee, H., & Bhattarakosol, P. [21]** This study proposes a state-of-the-art mechanism for authentication by combining keystroke dynamics and finger pressure to enhance the security of mobile devices. This technology stresses reviewing on how users type and the pressure that is exerted on the keys, creating a behavioral biometric profile for each user. For the keystroke dynamics it measures the timing information **such as dwell time (time a key is held) and flight time (time between the keystrokes)**. For finger pressure it captures the pressure levels that are applied during typing to provide an additional layer of authentication. It also fuses both the techniques for far better accuracy and the strength of the method against impersonation. With keystrokes and pressure sensors this is ideal for desktop systems and to enhance security for sensitive applications such as online banking and secure communication channels. Since the method uses finger pressure, there is a dependency on a consistent

hardware to ensure reliable collection of data, with this challenge the limited data availability also poses an issue, hence the author proposes further research over the machine learning algorithms to refine the system and to study more on methods to use so as to integrate a diverse range of data in the evaluation method so as to enhance the accuracy of the mechanism.

# CHAPTER 3

## SYSTEM DEVELOPMENT

### 3.1 REQUIREMENTS AND ANALYSIS

The project started with an exhaustive identification and analysis of its functional and non-functional requirements to successfully carry out the implementation of the behavioral biometrics-based authentication system. At the moment, the project intends to get varieties and high-quality behavioral data collected from users. That data would form the bedrock for tuning the machine learning models that would be used in the next stages of the project.

In essence, this system includes interactions collected from mobile devices-for example, touch gestures and movement pattern diversity. The current dataset has behavioral data of 25 participants and contains more than 2.5 million rows of diverse usage pattern and demographic representation.

#### 3.1.1 REQUIREMENTS

##### 1. Functional Requirements

- Data Collection Application:

A build that is tailored for mobile applications to collect sensor and interaction data by installing hooks to touch gestures (swipe or tap) and to patterns of device movement (e.g., accelerometer and gyroscope).

- Secure Data Storage:

Local on-device data storage or transfer mechanisms for secure cloud-based storage.

- Data Validation and Preprocessing:

Simple validation checks ensure correctness and usability of the data collected, including anomaly removal and filtering for incomplete entries.

## 2. Non-functional Requirements

- Reliability:  
Data collection applications must be functioning consistently with no loss of data over a wide spectrum of devices or operating systems.
- Performance:  
The application should function in a way that is not battery-hungry or performance compromising on the user's device.
- User Privacy:  
Anonymization of all data collected, with the requirement that users explicitly agree to participate in the information collection process.

## 3. Hardware Requirements

- Mobile devices:  
Android smartphones or tablets with built-in sensors such as accelerometers, gyroscopes, magnetometers, rotation sensors and other sensors, and a touch screen.
- Environment of Development:  
A powerful computer incorporating a new CPU and adequate storage to process and store huge datasets.

## 4. Software Requirements

- Programming Language:  
The use of the cross-platform language Flutter for developing the mobile application.
- Database Management:  
Use of Firebase to ensure that data collected remains safe and retrievable.
- Development Tools:  
Use Android Studio or Visual studio Code for developing and debugging applications.

## 5. Skills Required

- Mobile Application Development:  
Experience with cross-platform application development with Flutter.
- Sensor Integration:  
Skills that can be adapted to understanding how data can be collected from mobile device sensors, such as accelerometers and gyroscopes.
- Data Handling:  
Familiarity with Database Management software such as Firebase to store and retrieve data securely.
- User Experience Design:  
Ability to design a data collection application that is easy to use and intuitive.

It, therefore, guarantees the collection of quality behavioral data as a prerequisite for the phases that follow model training and development of the authentication system.

### 3.1.2 ANALYSIS

#### 1. Feasibility

At this stage in the project, developing an authentication system that uses behavioral biometrics would be technically possible. This is proven by the successful creation of the data collection application as well as the collection of around 25 lakh rows of data from 25 different participants. This would take the project into its subsequent stages, concentrating on model training and building a continuous authentication system. Simple techniques being mobile sensors and Firebase make deployment feasible.

#### 2. Challenges

- Data Correlation and Variation:  
The existing collection so far has 25 subjects. To make this collection as meaningful as one can find, it is highly necessary to increase the sample size and ensure diversity in demographics and types of devices used. Collection of data while respecting privacy and ethical guidelines is also challenging.

- Data Cleaning and Preprocessing:  
Incomplete, noisy, or irregular data may be present in the dataset, which may impair the model's quality of output. Preprocessing has to be efficient and provide the analysis and preparation of data to machine learning model.
- Device Variability:  
Devices have different screen sizes, resolutions, and hardware differences that make user behaviors vary. Such a difference would be required to be put in place when carrying out the analysis and development of the model.
- Resource and Processing Limitations as a Barrier:  
Heavy computation and training large datasets may require much-needed machine resources at this stage. Optimizing algorithms of resource management would be a requisite.
- Adaptation over Time:  
This is a long-term goal for the system and will need more elaborate retraining schemes for future challenges to ensure the system will respond to changing user behavior over time.

### 3. Potential Benefits

- Enhanced safety:  
Behavioral biometrics are not only difficult to duplicate, but also come in a tough form in an aspect of authentication. Thus, it offers good access protection against unauthorized entry.
- Improved ease of use:  
It provides a non-intrusive way of behavioral biometric measurement and thus avoids the duplicate manual authentication steps, which are very useful in terms of usability efficiency for the access to protected content.

- Continual authentication:  
The traditional methods provide authentication only at the time of login, but this new approach continues to authenticate access continuously, providing a more secure system, especially in case of a stolen device.
- Scalable:  
One can easily scale it from using banking and online shopping to corporate use and niche applications such as parental controls.
- Preservation of Privacy:  
Behavioral biometrics has not tracked any delicate personal information, unlike most traditional security techniques. Thus, it allows much better privacy preservation for users.
- Adapt to a User Behavior:  
The system adapts to changing usage patterns of the user due to time, which would provide long-term reliability and effectiveness.

By addressing the challenge and providing the advantages, this project could really improve authentication systems security, user-friendliness, and respect for privacy.

## 3.2 PROJECT DESIGN AND ARCHITECTURE

This project currently focuses primarily on design and architecture at the stage of data collection, which is the basis of the behavioral biometrics system. The primary goal is to develop an application that can acquire data efficiently with high quality, usability, privacy, and scalability at the future stages.

### 3.2.1 METHODOLOGY

#### 1. Data Collection:

The app at its current stage collects user data. For us to be able to do that, we had to develop an application that collects the user data, stores it and then shares it.



- Sensor Integration:

The mobile application collects sensor data via accelerometer, gyroscope and other sensor readings to determine how the different chevron patterns are displayed. It also collates the touch interaction data via modality of several swipe gestures, taps, and hold durations.

- User Behavior Tracking:

The data will be gathered naturally as the users engage in behaviors found in their devices. Rather than employing predefined tasks, it will allow capturing true behavior by running in the background.

- Diversity of Users:

The app collected data among a test population of 25. The cumulative number of people reached within the current data set is about 25 lakhs, and this is adequate for the data in the future.

## 2. Data Anonymization and Privacy:

All user data are anonymized from the very beginning of collection to avoid having personal information stored such as by giving each participant a unique identifier and recording behavioral metrics only. Giving participants information regarding the data collected and should get their consent to the usage of data before adding them to the dataset.

## 3. Application Design:

The application design is very straightforward, intuitive, and simple, making it easy for users to use. It has simple options to start and end data collection and a status indicator for active sensors. The data will be logged in CSV format. Each row refers to a unique interaction instance. Data points like timestamp, sensor readings, and touch characteristics will be noted.

#### 4. Secure Data Storage:

Data is temporarily stored on the local device of the user and then periodically uploaded to a secure Firebase database. This ensures minimum data loss in application breaks or during interruptions in operation. Data encryption techniques are employed during transfer to prevent unauthorized access.

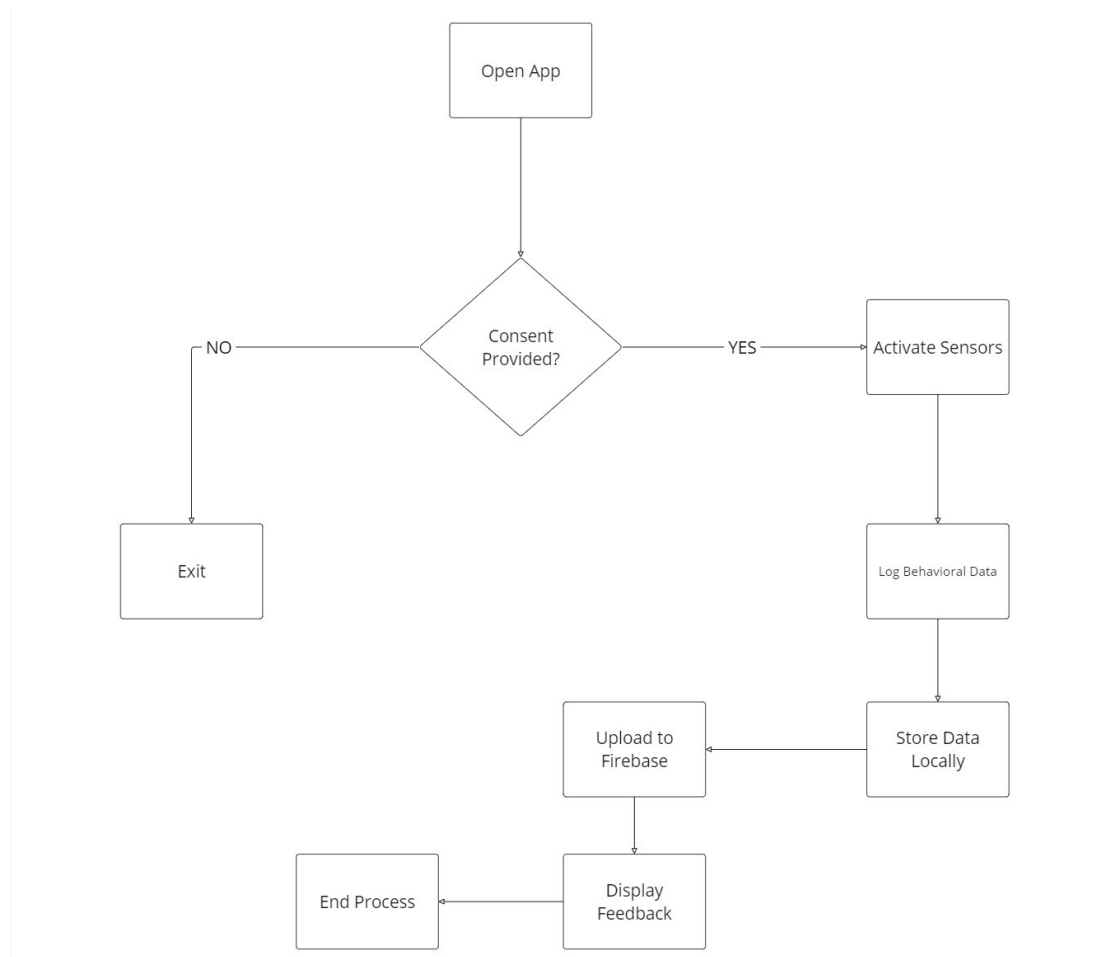


Figure 3.2.1 : Flow graph of the project.

#### 3.2.2 SYSTEM ARCHITECTURE

The system architecture for data collection can be divided into the following significant components:

### 1. Mobile Application:

- Frontend: Built on Flutter to support cross-platform implementation so it can be accessed seamlessly on an Android device.
- Backend: Real-time connections and secure storage via Firebase.

### 2. Sensor Data Pipeline:

- Accelerometers, Gyroscopes and Magnetometer Supported Data: This collects motion and orientation from a device.
- Touch Interface Sensors: This logs every touch interaction, including swipe acceleration, angle, and pressure.

### 3. Data Processing and Storage:

- Local Processing: Raw data from sensors is preprocessed on the device to remove noise and achieve consistency before being uploaded.
- Cloud Storage: This processed data is uploaded onto Firebase for centralized access and further analysis of this data.

### 4. User Feedback Mechanism:

The application relays the progress of data collection through notification feedback to both parties.

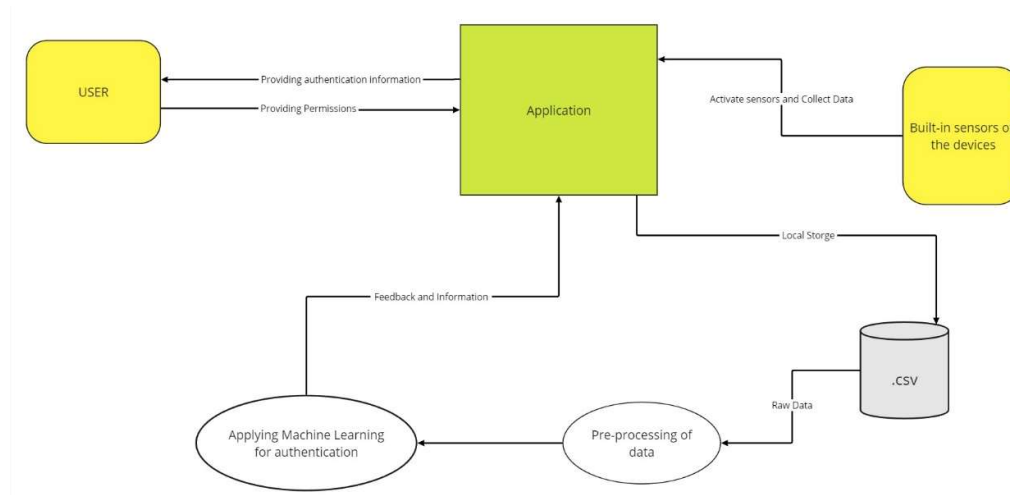


Figure 3.2.2 : Data Flow Diagram (DFD) of the project.

### 3.3 DATA PREPARATION

Preparation of all relevant data before it can be utilized for analysis, or the training of machine learning models is necessary for the behavioral biometrics' dataset dealing with 25 participants. It is a self-made dataset comprising 5,563,802 rows collected by a mobile application, which include sensor data along with interaction data. It contains numerous columns such as accelerometer, gyroscope, magnetometer, rotation vectors, tilt detector, motion, auto-rotation, and touch data. It is observed that the data sizes vary among different participants, which indicates the need for thorough preprocessing.

#### 3.3.1 DATA COLLECTION OVERVIEW

The raw dataset includes the following attributes:

- Timestamp: Records the time of each interaction or sensor reading.
- Accelerometer (X, Y, Z): Measures the acceleration of the device in three dimensions.
- Gyroscope (X, Y, Z): Captures the angular velocity in three axes.
- Magnetometer (X, Y, Z): Detects the magnetic field in three dimensions.
- Rotation Vector (X, Y, Z): Represents the device's orientation.

- Tilt Detector (X, Y, Z): Measures the tilt of the device.
- Autorotation (X, Y, Z): Tracks changes in screen orientation.
- Motion (X, Y, Z): Logs specific motion patterns.
- Last Touch (X, Y): Captures the last touch point on the device's screen.

### 3.3.2 CHALLENGES IN DATA COLLECTION

1. **Varying Data Sizes:** The size of collected data varies from one participant to the other owing to a vast difference in usage time, compatibility with the devices, or differing interaction styles of the users.
2. **Noise Data:** Sensor readings are often mixed with noise from environmental causes, device hardware variability, or inadvertent user actions, which must be filtered out.
3. **Missing or incomplete data:** Some rows could have missing sensors values or timestamps which could affect downstream analysis unless handled properly.
4. **Timestamps synchronization:** It is critical that all sensor data are synchronized and tied to the same interaction instance.
5. **Users Imbalance Contribution:** Some participants have contributed datasets many times larger than those of others, which could bias results when training the model.

### 3.3.3 THE NEXT STEPS IN DATA COLLECTION

#### 1. Data Cleansing:

There is a need to remove duplicate entries in the dataset so that there is no redundancy in the recorded data. Similarly, missing sensor values need to be

filled in using interpolation methods or removing incomplete rows if they are a small fraction of the entire data.

## **2. Data Filtering:**

The noise that is coming to the sensor readings must be reduced, while preserving the useful signals using filters such as moving average and low pass. Then, we can find the outliers by statistical approaches like using the z-score to detect and eliminate those that are so far out of bounds from where they are expected.

## **3. Data Synchronization:**

We will synchronize compatible timestamps across all sensor columns such that every row will signify one occurrence of user interaction.

## **4. Normalization and Scaling:**

The next step is normalizing sensor reading to the same range for example, [0, 1] or -1 to 1 will make the features homogeneous and invariant to different units and magnitudes.

## **5. Balancing the Dataset:**

Finally, we can balance the dataset for the ones having imbalance amount of data according to the needs: down-sampling users with too much data or synthesizing data for users with too few samples.

# **3.4 IMPLEMENTATION**

The current stage of implementation includes developing and deploying a data collection app supporting easy collection of behavioral data from mobile devices. It was designed to collect sensor data, touch interaction data, and movement patterns, all in a user-friendly context with functional privacy protections. Implementation included selection of suitable tools, architecture design for the app, and integration of diverse functionalities to aid in data collection.

### 3.4.1 APPLICATION DEVELOPMENT

This application for gathering data was created using Flutter, a cross-platform framework. It is intended for running a mobile application on devices using Android. The clear objective of this application was to develop a very powerful yet simple means of collecting real-time sensor and touch interaction data. The most development was oriented towards the widest possible scenario compatibility with minimum resource consumption and the highest data integrity during collection. The application has Firebase integrated into it for this purpose: it gives a very secure way for data cloud storage, thus ensuring scalability and centralized access to the data. The entire development was carried out considering user privacy and consent, with a clear mechanism in place to anonymize and encrypt the data collected. This application is efficient, scalable, and ready for the next phases of the project.

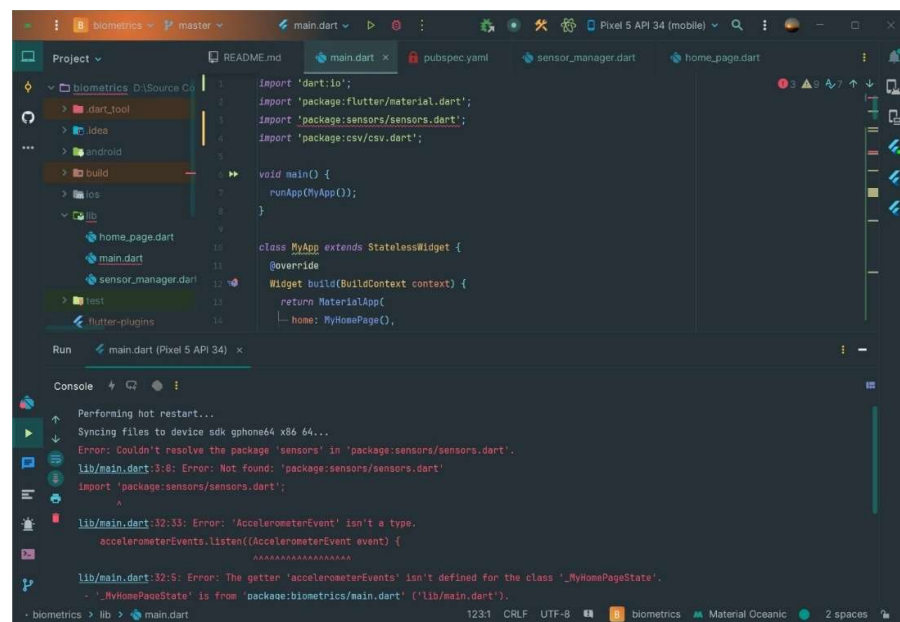


Figure 3.4.1 : The application in the initial stages of development.





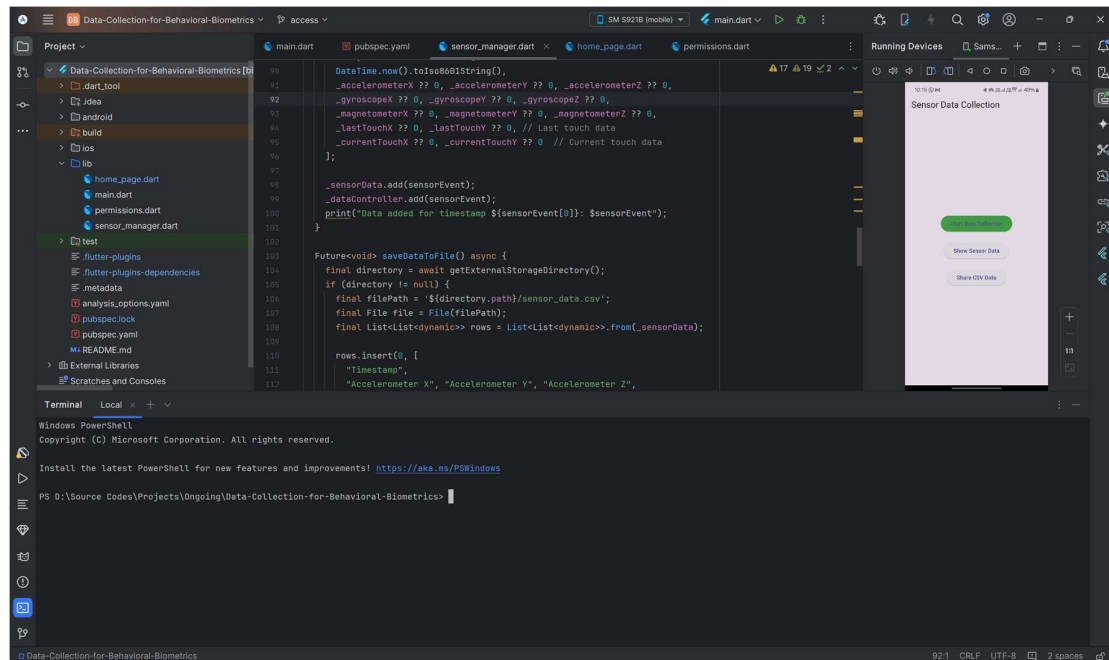


Figure 3.4.4 : Integrating the other sensors into the app.

## Features of the application

- Collection of Sensor Data:

The application uses the built-in sensors of devices such as accelerometer, gyroscope, magnetometer, and rotation vector to acquire real-time motion and orientation information. It also collects touch interaction data including the coordinates of the last touch on the screen.

- Real Time Data Logging:

Logging of data occurs as CSV format and each row corresponds to one sensor reading i.e. one interaction instance. Time stamp information is added to properly order data in time sequence.

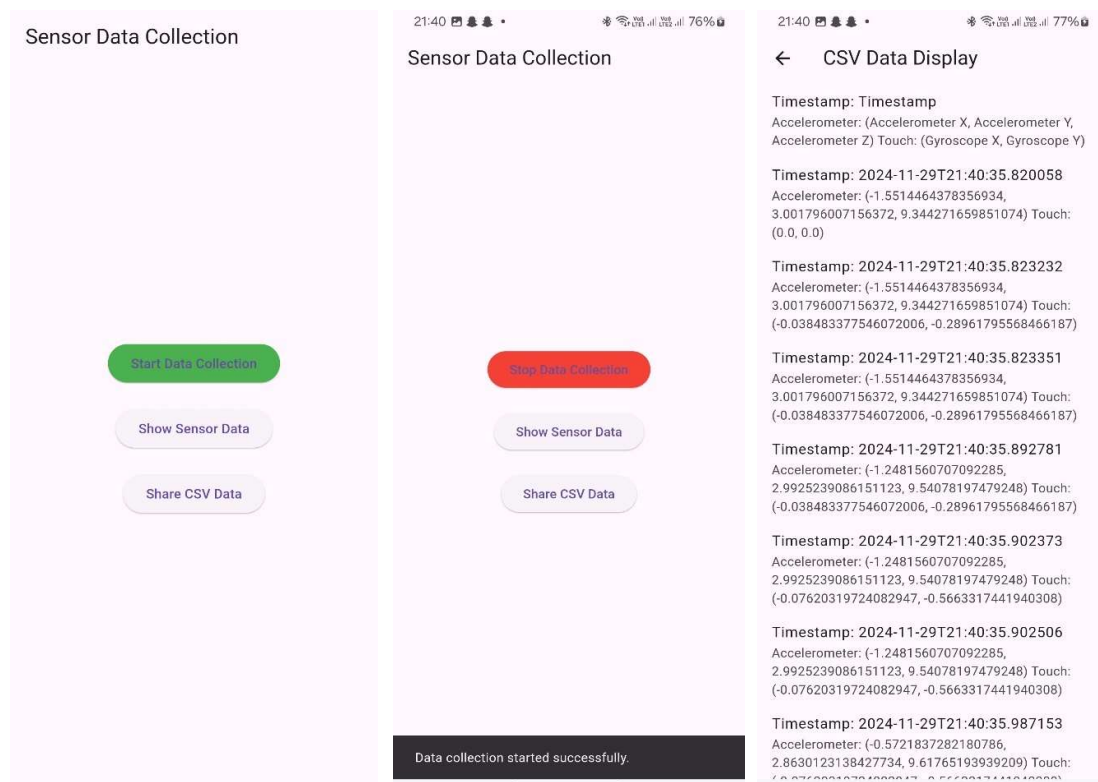
- User Interface (UI):

Developed minimalistic and user-friendly interface using flutter acting clearly to start and stop data collection. Real-time status indicators inform users about active sensors and collecting data.

- Data Storage and Security: The data is stored locally and periodically uploaded to Firebase for secure centralized storage. It's anonymized for user privacy, and all transfers are encrypted.

The application is also performance tuned to have low battery consumption and low resource usage which is important for long-term data collection sessions. It has implemented strong error handling procedures to prevent any data loss in case of crashing or interruptions of the application.

In essence, this application is the backbone of the data collection phase in this whole project: it promises highly qualitative, heterogeneous, and secure behavioral data always in an experience that is easy to use yet respects privacy.



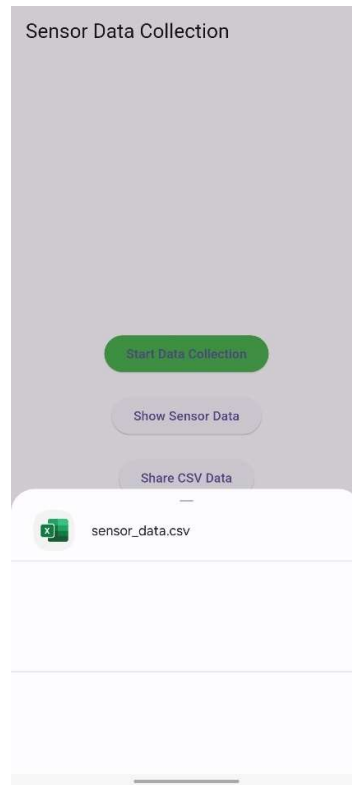


Figure 3.4.1.5 : The final, functional application.

## 3.4.2 TOOLS AND TECHNOLOGIES

### 1. Flutter framework:

Selection of this framework is due to its cross-platform capability with the ability of the Android devices to run the app. Widgets were used in designing the UI to maintain compatibility and responsiveness of the apps across devices.

### 2. Firebase:

The backend for real-time data storage and synchronization. It also offers scalability and secure data handling.

### 3. Sensor plugins:

sensor\_plus: A Flutter plugin for access to data from the device sensors, such as accelerometer, gyroscope, and magnetometer. shared\_preferences: For local cache of the application settings.

#### 4. Programming Language:

Dart: The primary programming language used for Flutter application development.

### 3.5 KEY CHALLENGES

Data collection application development and deployment for behavioral biometrics had many challenges both technically and with usage by the users. The key challenges are outlined below with the solutions which were deployed to address them:

#### 1. Data variability across devices

Sensor data collected from different devices presented very high variability when it comes to the hardware specifications. These specifications include sensor sensitivity, screen size, and resolution for the devices themselves. Hence, it is difficult to keep the whole dataset standard.

**Solution:** Standardization of data intervals and formats will be performed during data preprocessing. Procedures for preprocessing were laid out to be uniform for all devices prior to any analysis by normalizing sensor readings.

#### 2. Battery Consumption

By using sensors such as accelerometers, gyroscopes, and magnetometers in any device, it continuously led it to draw a lot of battery power which does not allow long time usage by participants.

**Solution:** Optimization of polling rates of the sensors achieved a trade-off in terms of detail in the data versus energy consumption. The system was programmed to run in the background.

### 3. Storage Permissions

Some storage permission restrictions in recent Android updates were causing problems for locally saving or even creating collected data files.

**Solution:** Scoped storage methods were implemented to be compliant with all the latest Android updates. End-users were guided towards the granting of necessary permissions with in-app prompting and further explanations.

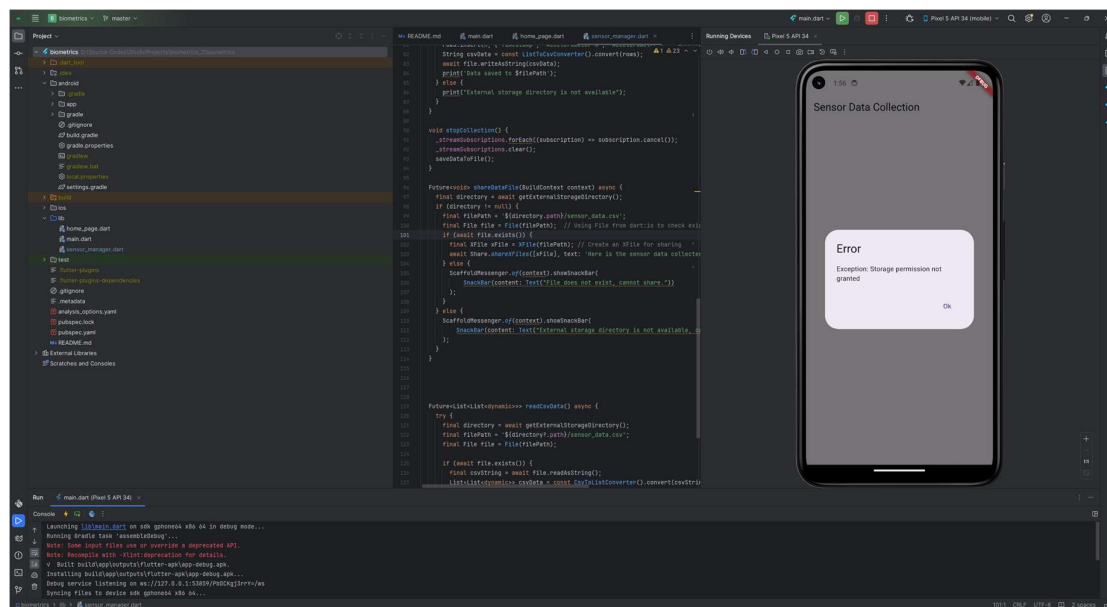


Figure 3.5.1 : Unable to access storage due to an android permission issue.

### 4. Large Data Volume

A big volume of data would fail to manage, store, or even transfer, since there were still over 5.5 million rows of sensor data having been collected during these twenty-five participants.

**Solution:** The app was designed so that it held data locally dense in compressed CSV files so that it could reduce the storage. Data upload to Firebase in batches was then set up to leverage upload speed and bandwidth savings associated with the scale capabilities of the platform.

### 5. File Not Found After Recording

Sometimes the recorded sensor data files were not available due to either an improper saving of files or a misalignment in file paths.

**Solution:** Robust file management system has been integrated into the application. It saves files in combinations of unique timestamps and paths, and verification added at file creation confirms that the file has been successfully created. Cases where the files were not found would have exception handler mechanisms to provide fallback options for re-saving or notifying the user.

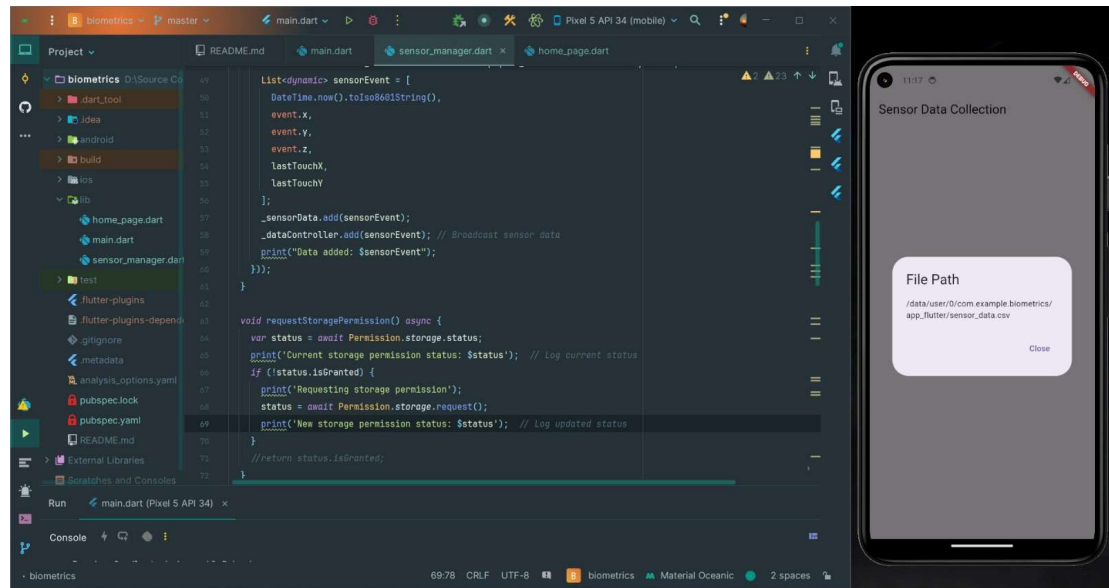


Figure 3.5.2 : The app shows that the file has been saved, but it was nowhere to be found.

# CHAPTER 4

## TESTING

### 4.1 TESTING STRATEGY

The testing based on the data collection application was primarily concentrated on checking whether the application effectively delivers exact and complete data as well as ensuring that privacy and security standards are met. The testing strategy was focused on aspects such as functionality, data integrity, performance, compatibility, and user experience, to ensure that the application would very tightly be able to collect high-quality behavioral data in the future.

#### 4.1.1 FUNCTIONAL TESTING

Functional testing was done to check the working of core application features like sensor integration, data logging, and upload of collected data with security assurance; besides, each individual component, such as accelerometer, gyroscope, magnetometer, and touch data modules, was tested to verify accurate data capture. The sensor readings were checked against expected outputs to ensure precision and completeness. Data logging tests that the collected data were accurately recorded in CSV format aligned by timestamp for usability during future analysis.

#### 4.1.2 PERFORMANCE TESTING

Performance testing targets efficiency and resource consumption of the app while using it for long periods. The app was tested for long hours to see the difference in battery life and CPU utilization memory consumption. Optimizing sensor polling rates, the application is meant to reduce battery drain without sacrificing data quality. Stress tests consisted of continuous data collection for hours to determine app stability and capacity for managing large datasets while taking data. So that participants can use the app for long data collection sessions without any interruption.

### 4.1.3 COMPATIBILITY TESTING

To make sure that the application would work generally on a broad spectrum of Android devices with different hardware and software configurations, it was tested thoroughly across them. Test devices had various combinations of screen size and resolution, sensor specification, and operating system (ranging from Android 9 to Android 12). Identify, leaving any device-specific issues such as sensor sensitivity or access permission for storage unresolved. Test compatibility ensured performance consistency of the application across all the devices under tests, making it ready for launch to users from various backgrounds.

### 4.1.4 USABILITY TESTING

The application interfaces were tested with actual users for the assessment of simplicity, clarity and ease of use. The feedback was collected on how intuitive the controls for were initiating and terminating data collection and the overall navigation experience. Participants considered the app to be easy to use without much guidance and real-time feedback on the data collection status. Each of the feedback was considered towards user-friendliness improvement so that even a non-technical person can participate without any hassle.

## 4.2 TEST CASES AND OUTCOMES

### 4.2.1 TEST CASE 1: SENSOR DATA LOGGING

**Objective:**

Verify that the app accurately captures sensor data (e.g., accelerometer, gyroscope) and logs it in the correct format with timestamps.

**Test Steps:**

1. Started the app and enabled data collection.
2. Performed actions such as shaking the phone, rotating it, and swiping on the screen.
3. Stopped data collection and inspected the generated CSV file for accuracy.



### **Expected Outcome:**

The app should log data from all sensors in the correct format with properly aligned timestamps.

### **Actual Outcome:**

Sensor data was logged accurately, and timestamps were consistent across all sensors. CSV files were generated as expected, with no missing values.

**Status: Pass**

## **4.2.2 TEST CASE 2: TOUCH INTERACTION DATA**

### **Objective:**

Verify that the app accurately captures touch data (e.g., current/last touch, swipes) and logs it in the correct format with timestamps.

### **Test Steps:**

1. Started the app and enabled data collection.
2. Swiped across the screen.
3. Stopped data collection and inspected the generated CSV file for accuracy.

### **Expected Outcome:**

The app should log touch in the correct format with properly aligned timestamps.

### **Actual Outcome:**

Touch data was logged accurately, and timestamps were consistent across all sensors. CSV files were generated as expected, with no missing values.

**Status: Pass**

### 4.2.3 TEST CASE 3: VERIFY COMPATIBILITY ACROSS DEVICES

**Objective:**

Verify that the app works on multiple android versions, since we need to account for population having a variety of different android devices.

**Test Steps:**

1. Used different devices and install the app.
2. Checked if the app is working.

**Expected Outcome:**

The app should work across different devices.

**Actual Outcome:**

The app worked and gathered data in all the devices it was installed on.

**Status: Pass**

# CHAPTER 5

## RESULTS AND EVALUATION

The project results and evaluation part at this stage of the project comprises of the performance of the data collection app as well as the extent to which it achieved its objectives. This chapter summarizes the outcomes of the activities performed during the testing phase, as well as addresses quality in the collection of data, plus the general performance, scalability, and readiness for the following project stages.

Name	Date modified	Type	Size
Combined	29/11/2024 16:25	File folder	
01 sensor_data - Akshit Calling Akshit	14/10/2024 11:58	Microsoft Excel Comma Separat...	17,601 KB
02 sensor_data - Manisha Sharma	02/11/2024 21:53	Microsoft Excel Comma Separat...	222,331 KB
03 sensor_data - Prashast Sharma (Touch)	02/11/2024 23:47	Microsoft Excel Comma Separat...	9,041 KB
03 sensor_data - Prashast Sharma	02/11/2024 23:47	Microsoft Excel Comma Separat...	18,625 KB
04 sensor_data - Akshit Sharma	05/11/2024 19:53	Microsoft Excel Comma Separat...	24,807 KB
05 sensor_data - Subodh C. Sharma	05/11/2024 19:53	Microsoft Excel Comma Separat...	60,703 KB
06 sensor_data - Pragati Thakur	05/11/2024 19:54	Microsoft Excel Comma Separat...	92,549 KB
07 sensor_data - Yukti Sharma (Touch)	05/11/2024 19:54	Microsoft Excel Comma Separat...	2,984 KB
07 sensor_data - Yukti Sharma	05/11/2024 19:55	Microsoft Excel Comma Separat...	1,185 KB
08 sensor_data - Utkarsh Sharma	05/11/2024 19:56	Microsoft Excel Comma Separat...	205,859 KB
09 sensor_data - Narendra Sharma [LESS]	05/11/2024 19:56	Microsoft Excel Comma Separat...	167 KB
10 sensor_data - Ishant Sharma [LESS]	05/11/2024 19:57	Microsoft Excel Comma Separat...	680 KB
11 sensor_data - Kartik Sharma	05/11/2024 19:58	Microsoft Excel Comma Separat...	23,713 KB
12 sensor_data - Krishna Devi	05/11/2024 19:59	Microsoft Excel Comma Separat...	198,317 KB
13 sensor_data - Reyaansh Sharma	05/11/2024 19:59	Microsoft Excel Comma Separat...	101,915 KB
14 sensor_data - Bimla Devi Sharma	05/11/2024 20:00	Microsoft Excel Comma Separat...	135,040 KB
15 sensor_data - Aman Shrivastava	05/11/2024 20:02	Microsoft Excel Comma Separat...	228,015 KB
16 sensor_data - Ekshta Mishra	05/11/2024 20:02	Microsoft Excel Comma Separat...	105,038 KB
17 sensor_data - Somya Soni	05/11/2024 20:05	Microsoft Excel Comma Separat...	37,816 KB
18 sensor_data - Harshit Srivastava	05/11/2024 22:05	Microsoft Excel Comma Separat...	95,292 KB
19 sensor_data - Rushil Agnihotri	05/11/2024 22:07	Microsoft Excel Comma Separat...	17,891 KB
20 sensor_data - Adarsh Mahajan	06/11/2024 15:10	Microsoft Excel Comma Separat...	220,620 KB
21 sensor_data - Simran	15/11/2024 20:23	Microsoft Excel Comma Separat...	26,030 KB
22 sensor_data - Swapnil Tyagi	15/11/2024 18:45	Microsoft Excel Comma Separat...	150,342 KB
23 sensor_data - Kartikey Attri	15/11/2024 18:46	Microsoft Excel Comma Separat...	130,808 KB
24 sensor_data - Shivansh	15/11/2024 18:47	Microsoft Excel Comma Separat...	252,072 KB
25 sensor_data - Utkarsh	15/11/2024 20:23	Microsoft Excel Comma Separat...	94,905 KB

Figure 5.1 : The different sensor data collected by the data collection app.

## 5.1 RESULTS

### 1. Functionality

This app is capable of effectively capturing data from various sensors on the device, including accelerometers, gyroscopes, magnetometers, rotation vector, and touch interaction data. For every single data point, structured CSV format logging was

carried out in a database with more than 5,563,802 rows obtained from 25 participants.

## **2. Result Testing**

The application did indeed accurately record sensor readings and synchronize them by timestamp, assuring that all data was aligned across all metrics.

## **3. Compatibility**

The applications performed in a thoroughly effective manner on a variety of Android gadgets with various hardware specs and versions of the OS for the device (Android 9 through 12). No problems of substantial compatibility would seem to have been found thus proving the viability and the scalability of the app in terms of deployment across different customer bases.

## **4. Performance**

The use of the application even in extended periods of time did not seem to tax the battery at all since the polling rate for sensors was optimized for better efficiency. The application hung and crashed very seldom even for hours of continuous use.

## **5. User Feedback**

The users described the app as having favorable experience because of its ease of use and interface. Feedback from data collection in real time and at a click of a button asking permissions helped to make the application easy to use.

# **5.2 EVALUATION**

## **1. Quality of Collected Data**

Data from 25 individuals form an excellent basis for the next phases of the project. A wealth of differences in behavior patterns across the device type, usage habits, and the background ensures that diversity is present for the next training and act on machine learning model construction in subsequent phases.

## 2. Adaptability

Its feature to be able to being deployed on varied devices and OS platforms gives it robustness with respect to going into the audience. Hardware compatibility varies, enabling it to fit and materialize for several arrays of users.

## 3. Limitations

The app has the function of acquiring behavioral data, but several issues have not been resolved. For example, some participants have different data sizes because they exhibit varying levels of usage and engagement with their devices. Future versions of this project would integrate normalization and augmentation methods into the contribution of each user to improve or use data integration methods for further balancing of the data set.

## 4. Comparison to the Objectives

The results of the research can now indicate that the app has indeed met all its initial objectives. It gives all the benefits of behaviour data collected using a seamless interface yet ensures security and privacy and achieves a balance between features and user experience. Above all, it proves itself successful in collecting a quality set of data from a range of participants.

The screenshot shows a Microsoft Excel spreadsheet with the following structure:

- Columns:** Labeled A through Z. The first few columns (A-E) contain timestamps. Columns F through Z contain various sensor readings.
- Rows:** Numbered 1 through 33. Row 1 is the header row.
- Data:** The data is organized into groups of 5 rows each, corresponding to different participants or sessions. Each group starts with a timestamp in column A and continues with sensor readings in columns B through Z.
- Sensors:** The sensor readings include Accelerometer, Gyroscope, Magnetometer, and Rotation, among others.

Figure 5.2: A sample of sensor\_data.csv file collected from a participant.

### 5.3 KEY METRICS

Total Participants	25
Rows Collected	55,63,802
Columns Collected	24
Sensors Used	<ul style="list-style-type: none"> <li>• Accelerometer</li> <li>• Gyroscope</li> <li>• Magnetometer</li> <li>• Rotation Vector</li> <li>• Tilt Detector</li> <li>• Autorotation</li> <li>• Motion</li> <li>• Last Touch</li> </ul>
Compatibility	Tested on Android Versions 9 to 14

*Table 5.1: Key Metrics of the data collection app.*

## CHAPTER 6

### CONCLUSIONS AND FUTURE SCOPE

#### 6.1 CONCLUSION

The initial application for the data collection is a real milestone in our project. The application is developed for high-quality data from device sensors and touches, reaching its goal of collecting about 5,563,802 data rows from 25 different subjects. Through thorough testing and validation, the app proved its performance most widely capable and compatible with devices and operating systems in terms of reliability, accuracy, and compatibility across such a large range.

Indeed, the results demonstrated and provided signs that the app is making a solid base within which to draw behavioral data while protecting users with their privacy rights on their data. Its intuitive design and minimal consumption of resources make it ideal to collect preliminary data and for further expansion later. It collects data on which many machine learning models will train the nonintrusive, continuous preferred authentication scheme, directly relating to the overall aim of the project.

This phase completes a project on the potential for taking advantage of behavioral biometrics to provide a user-friendly and secure way of authenticating. Coupled with scalability and accessibility, a seamless replacement of traditional techniques with this will become an option. The challenges weren't many, ranging from a size variation in data to a couple of minor device-specific issues. However, these are solved through thoughtful design and iterative improvements.

## 6.2 FUTURE SCOPE

### 1. Dataset Expansion

We would continue collecting data from more diverse participants to improve robustness and generalization. As of now, the target by the end of our major project is to have data from 100 different people.

### 2. Data Enrichment

As we continue for now working on a machine learning model, simultaneously, we plan to keep adding to the data collection application – like adding more sensors (proximity, heart rate monitor and more). Advanced features such as gesture recognition could offer the possibility of increasing the richness of the data.

### 3. Model Training and Validation

The next step toward completing the project is to perform machine learning model training using the collected data for continuous authentication. These models can be further explored with the use of deep learning and ensemble methods in improving model accuracy and enhancing adaptability.

### 4. Real-Time Authentication

The next step after training the model would be to make it such that the app collects data in real-time, processes it, and predicts with a very good accuracy, the authenticity of the user.

### 5. Cross-Platform Development

Since flutter is a cross-platform framework, it would not be much effort to also make the app available to iOS users, just to get a bit more diverse data.

### 6. Optimizations

Right now, with the app running in the background constantly, it does take a significant toll on the user's smartphone. In the case of an old smartphone, it can even cause lag. This needs to be optimized.



## 7. Integration with real world applications

The final plan is to integrate the app with real-world applications for seamless authentication. For this project, we are planning to develop a prototype app.

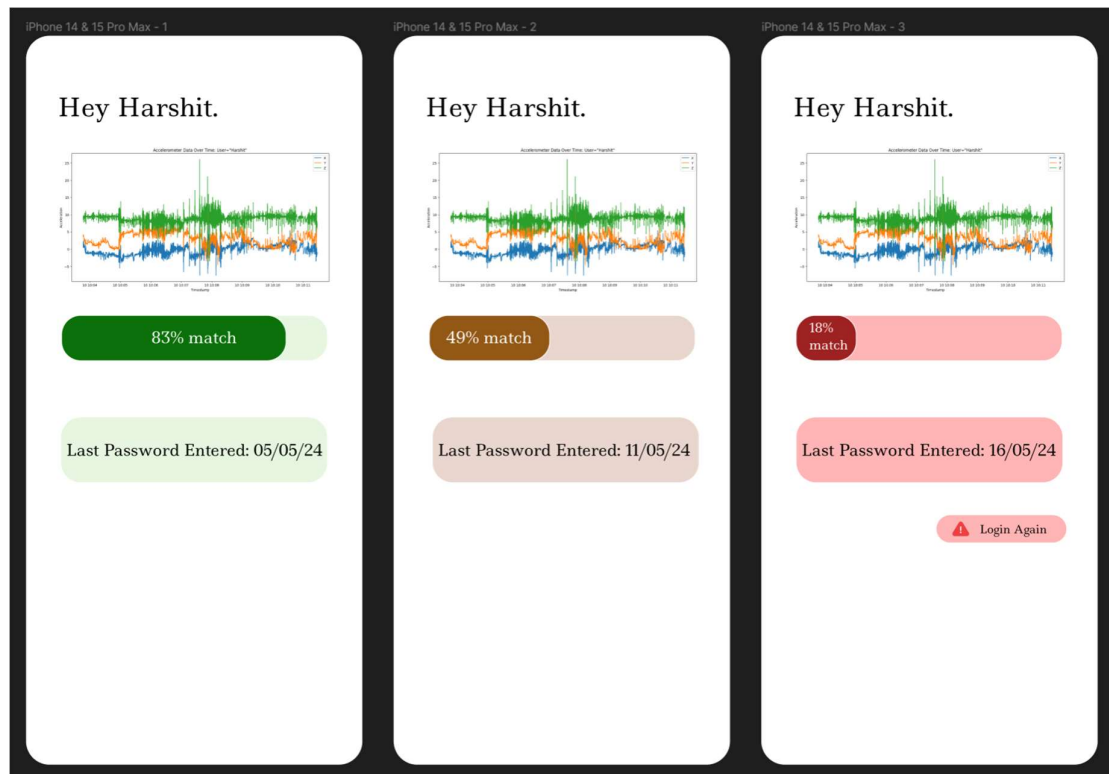


Figure 6.2.1: Figma design for the target prototype.