Fingerprint Based Approach to avoid Data Exfiltration in Healthcare Applications

By

Abdul Mutal

A thesis submitted to the faculty of Information Security Department, Military College of Signals, National University of Sciences and Technology, Rawalpindi in partial fulfillment of the requirements for the degree of MS in Information Security

August 2022

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS Thesis written by **Abdul Mutal** Registration No. **00000275593**, of **Military College of Signals** has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulations/MS Policy, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS degree. It is further certified that necessary amendments as pointed out by GEC members and local evaluators of the scholar have also been incorporated in the said thesis.

Signature: _____

Name of Supervisor **Asst Prof Dr Shahzaib Tahir**

Date: _____

Signature (HOD): _____

Date: _____

Signature (Dean/Principal) _____

Date: _____

Declaration

I hereby declare that no segment of work presented in this thesis has been presented in support of an additional award or qualification either at this institution or elsewhere.

_____

MS Student

Dedication

“In the name of Allah, the most Beneficent, the most Merciful"

I dedicate this thesis to all who support and stood by me during the hard time faced throughout my journey. I dedicate this to my mother, sister, and teachers who supported me each step of the way.

Acknowledgments

All praises to Allah for the strengths and His blessing in completing this thesis.

I would like to convey my gratitude to my supervisor, Dr. Shahzaib Tahir for his supervision and constant support. His invaluable help of constructive comments and suggestions throughout the experimental and thesis works are major contributions to the success of this research. Also, I would thank my committee members; Asst Prof Dr. Fawad Khan, and Major Sohaib Khan Niazi for their support and knowledge regarding this topic.

Last but not the least, I am highly grateful to my parents, spouse, and sisters. They have always stood by my dreams and aspirations and have been a great source of inspiration for me. I would like to express gratitude them for all their support, love, and care through my times of stress and excitement.

Abstract

In the past few years, the PUFs are used for authentication in different IoT devices. Data exfiltration is common among different android applications and results in severe damage to individuals as well as organizations. This thesis describes the implementation of generating strong encryption keys for data at runtime. As we all know that mobile devices are computationally less strong and are not feasible to create a strong key for encryption. Our implementation uses mobile device fingerprints such as PUFs to generate strong symmetric key along with other user unique attributes to encrypt the documents. In this research we only target the healthcare applications as malicious actors are targeting this industry to get benefits.

Healthcare applications are quite vulnerable for decades as the malicious actors are targeting vulnerabilities in mobile applications to retrieve sensitive data. The documents or images contains ePHI (electronic Protected healthcare information) and PII (Personally Identifiable Information) which can be used malicious application to threat or kill a person. Our Research focused on implementation advance security features for android application in order to mitigate these types of attacks.

The thesis purely focused on the security of mobile application of healthcare organizations. As mobile devices are computationally weak and cannot generate strong encryption key with high entropy, so we used in built features of device to generate strong encryption key at runtime to encrypt user sensitive data in order to mitigate data exfiltration. The research completely focusses on generating strong encryption key by using Physically Unclonable functions like gyroscope, accelerometer or magnetometer values to generate symmetric key for encrypting user documents or labs and radiology images and reports.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

Malicious actors utilize data ex-filtration to access, collect, and transmit sensitive information. Data ex-filtration can be carried either remotely or manually, and it is typically difficult to detect since it mimics legitimate network traffic. Financial data, customer details, and innovative proprietary information are all common targets.

Unfortunately, to enter a network, hackers stealing data, and avoid detection, an attacker does not need to employ very complex tools; this is true for both malicious cyber groups and less skilled threat actors, and especially for malevolent insiders.

On the cyber-crime front, there has been a shift in motivations and complexity in assault tactics. Insiders, spammers, worm, and virus authors are today considered more of a nuisance than a threat due to their decreasing impact. As activists such as Lulzsec and Hackers engage campaigns to cause system disruption, online defacement, and information exposure, activism is perceived as having more worldwide recognition and defamation goals. Although activists are still a nuisance, organized crime, industrial espionage by rivals, and cyber warfare by nation governments appear to be the most serious concerns. Advanced recurring challenges and nation-state actors are also included in this new breed of savvy criminals, whether they are stealing intellectual property, or targeting basic infrastructure for economic and geopolitical reasons. The Botnet, Fireball, Xenon, and Duqu assault campaigns demonstrate the ferocity of emerging threats and the sophistication of attacker cheval chevaliers. Attack tools have also emerged. Phishing assaults frequently use botnets, backdoors, and malware as primary weapon. Professional programmers are increasingly writing malware and hacking tools, which are designed to circumvent discovery by security mechanisms.

Organizations all around the world are generating massive volumes of data, which they must store and safeguard against fraud, destruction, and abuse. Data is very important asset for the company and every organization working therein. Data leak dangers exist whether the data is housed on-premises or in the cloud. These risks can come from hostile insiders with authorized access to company assets, as well as malevolent outsiders with no such authorization**.**

This document is intended to assist you in developing a finger-based strategy to identifying and preventing data leakage from your company. It includes the different ways in which information exists within an organization, as well as how remote attackers use widely available forms of information interchange found in most organizations, as well as more complex modes targeted at circumventing most security mechanisms. The main point here is that an attacker's arsenal is diversified, and skilled adversaries like Advanced Persistent Threats aren't confined to using traditional methods to gather data from corporate networks. They are capable, motivated, and have the resources to employ more complex channels.

As a result, security precautions in an enterprise must account for these varied data ex-filtration mechanisms. The paper examines both direct and indirect ex-filtration routes in this regard and illustrates how they work.

Data leaking across these routes can be discovered and blocked. When it regarding data ex-filtration, the straightforward method appears to be the most effective, as many organizations have done so are not designed to effectively combat extortion attempts. Security solutions are frequently focused on perimeter defense and fail to consider how to identify and interrupt persistent attackers' actions, particularly their attempts to breach data resources, until they have acquired a foothold. The most popular method of ex-filtration is through outbound FTP connections. We found that these ex-filtration mechanisms are used in more than half of the data breach events we investigated. It mixes in with typical network traffic, making it difficult to differentiate from actual user activity. Attackers also implement a variety of information ex-filtration tactics, varying from an indiscriminate flee dump to highly thorough and thoughtful filtering of the techniques are used to extract just the most important and high-value information.

1.2 Motivation and Problem Statement

The critical data present in healthcare applications especially android phones are vulnerable to data exfiltration which leads up to think about the new security loopholes in these applications. Help industry to identify and mitigate latest threats to medical data which is vulnerable to different data breach attacks. The fundamental issue in spectrum sensing is to precisely differentiate between a primary user and secondary user. There is a possibility for malicious applications to intercept the data and dodge the system into believing that it's the legitimate request.

Hence, there is a need for an effective security control over healthcare mobile applications in an easy manner so that applications residual data procures more security. After detection of attacker, the utmost requirement is of a mechanism which enables a node to do its effective defense.

1.3 Research Objectives

The main objectives of thesis are: -

- Study and analyze the various data siphoning or exfiltration techniques along with their existing prevention techniques.
- Study and analyze the various features, uses and types of PUFs.
- Use Fingerprint along with PUFs to authenticate users which will add security in mobile applications
- Propose a solution that combines PUFs and defensive techniques for data siphoning to mitigate multiple attacks via a single solution or improve an existing solution.

1.4 Thesis Contribution

To the best of our knowledge the technique used in this thesis research has not been used for handling data exfiltration attacks in mobile applications for healthcare.

The main contributions of this work are as follows.

- We propose and build a secure login mechanism for android application based on physically unclonable functionality and user passwords.
- Next, we generate keys based on PUFs as mobile devices have less computation powers.
- Then based on encryption keys to encrypt documents uploaded by the user. So that data exfiltration can be avoided.
- Unlike existing work, we have considered multiple PUFs sensors to generate unique keys for each document to evict key generation by adversaries.

1.5 Scope and Limitations

This research is mainly focused on healthcare mobile applications security. The research provides better security measures against data exfiltration, data siphoning and unauthorized data access. Its only applicable to healthcare mobile applications so that other malicious applications can't access critical data present in mobile devices.

Different studies have been carried out for analyzing the data exfiltration techniques for different states of data. In [1] the authors have reviewed different exfiltration attack vectors like passive monitoring of network channels, exploiting system vulnerabilities, spyware and malware, phishing, cross site scripting as well as physical theft. Then they classified the countermeasures according to the types of data (in use, in transit or at rest) and then mapped the countermeasures to the attack vectors. According to the carried-out research, different types of techniques exist to counter the existing data siphoning or exfiltration attack vectors. There can be preventive, detective, and investigative techniques. Our scope of study is only about the preventive countermeasure techniques. Further the scope can be expanded to other areas of industries where it can be applicable to other industry areas to achieve high security controls on residual data.

1.6 Thesis Organization

The thesis is structured as follows:

- Chapter 2 contains the literature reviewed in the thesis. The current implementation of PUFs in different kind of key generation mechanisms in mobile devices or IoTs. Strong key generation mechanisms to attain high entropy level. Different data exfiltration practices by adversaries.

- Chapter 3 contains the proposed scheme or methodology to generate keys using PUFs for avoiding data exfiltration. The Encryption techniques used for healthcare documents security.

- Chapter 4 covers the implementation of key generation in android application with enhanced security features and APIs that communicate with backend server and store encrypted data in databases.

- Chapter 5 contains test carried out on the PUFs and android security feature to enhance data security and avoid data siphoning. The simulation results involving defense actions and test results are defined in this chapter.

- Chapter 6 marks the end of the document. The conclusion and future work areas are revealed in this chapter.

Chapter 2

Literature Review

Data ex-filtration is a hotly debated research topic, with several security techniques described in the literature for preventing and combating data theft. Current Data Exfiltration Preventative Measures Systems (DEPSs) were studied, with an emphasis on their difficulties and solutions.

Potential areas for further research, according to the authors, include content analysis, internal abuse, and handset security [1]. In the same vein, the authors included a rundown of the most well-known relevant data strategies that malware may use. A classification of data exfiltration avoidance solutions was recently published in both business and academic research [2].

The causes and motives for data leaking, as well as the types of released information and knowledge leakage pathways, were discussed in this study. The authors [2] conducted a rigorous analysis of known data exfiltration attack routes and countermeasures with the goal of reporting the present state of the art in this field and identifying research needs for future study. Various techniques scan known channels that might be abused by attackers for stealing personal information, such as email messages, HTTP/HTTPS downloading and uploading, DNS protocol, and some elements of the TCP/IP protocol suite, in order to identify and halt data exfiltration attacks. For example, the authors presented a machine-learning-based DNS exfiltration detection technique that covers both DNS tunneling and reduced malware exfiltration. The technique investigation is based on principal domains, which allows for the screening of lawful DNS-based services. Two DNS tunneling utilities and two DNS unauthorized access malware implanted in massive DNS traffic are used to test this strategy. DPI (deep packet inspection) has been utilized by several other technologies to monitor network activity and identify data exfiltration. This approach scans all outbound traffic for important data that provides a greater degree of detection by ensuring that no data packet is left unchecked [3].

2.1 Avoiding Ex-filtration data techniques

Access Control: The goal of these strategies is to limit access to just those which are permitted [3]. Its techniques can take a variety of forms, depending on the methodology taken when giving access: allocating privileges based on roles (Position Multifactor Authentication) and attributes (Essential element Access Control).

Cryptography: A variety of cryptographic approaches are used to protect privacy, and they may be divided into three categories: Secret Key Cryptography (DES), which utilizes the same key to encrypt and decrypt, Public Key Cryptography (PAC), which uses two separate keys, ROSA, and Hash Function, which is an irrevocable function that creates a fixed-size output data from an unfixed-size input data [4].

Anonymization: This approach is frequently used until distribution and analysis procedures with the goal of data sanitization, also known as de-identification [4]; it reduces the precision of the data and conceals the identity of patients. Due to its multiple characteristics [5], including as decentralization, transparency, open-source, autonomy, immutability, and anonymity, blockchain has recently expanded outside the financial industry and has become a popular solution for decentralization and privacy concerns in the smarter healthcare area.

2.2 Healthcare data ex-filtration security systems

Because medical data is more sensitive than other categories, the usage of smart health has become a major source of data breaches. According to Risk-Based Security's 2021 Mid-Year Security Breach Quick View Report, 238 healthcare data infringements were disclosed in the first six months of 2021, putting the healthcare sector in first place as the most clearly violated economic sector [6]; additionally, "hacking" or "Unauthorized Access" is the most common breach type, highlighting the significance of data network access. Many strategies, including the proposed techniques, Security Systems, and Cryptography, are utilized for this purpose; the authentication procedure is also employed as a solution to enable secure access to medical data.

With the advent of block-chain, other technologies such as public block chain technology and services in place have been introduced to the mix. Block chain network includes an

authorization layer, which adds a layer of protection to the standard block chain, while agreements are used to manage the privileges to a patient's [7].

Zhong et al. [8], on the other hand, designed an efficient essential element encryption (ABE) approach that outsources portion of the encoding and decoding to edge nodes and enables attribute modifications, allowing for flexible right control. This method has been tested and reviewed at various levels of security and has proven to be more efficient for devices with limited resources than the typical ABE schema. Additionally, Onesimu et al. designed a privacy-preserving data collecting technique for IoT-based healthcare services founded on the clustering-based anonymous mode and formulates the detection mechanism as client-server-to-user to secure privacy on both ends.

Regarding dictionary attacks, the most reliable solutions utilized by modern PWM systems are salted scrambling or encoding the customer information [9]. By gaining illegal access to data stored in the database, hackers will be unable to rapidly discover the users' credentials. Despite their benefits, these PWM systems have a number of security flaws. For example, if an attacker gains access to a database containing user information, they can decrypt or decode it using a variety of computational approaches. The fundamental cause of this problem is that contemporary PWM systems employ algorithms that are well-known and widely available.

Authors examine a variety of PUFs developed in prior studies [10]. Memory PUFs are one sort of PUF that may be created using readily available memory techniques such as Flash, MRAM, and SRAM. In [10] and [11], SRAM PUFs were found independently and at the same time. Two cross-coupled inverters with two stable states make up a typical SRAM cell. The physical mismatch between the two symmetrical portions of the SRAM circuit is always caused by manufacturing variations. The power-up behavior is in charge of this random physical disparity. The reactions of the SRAM PUFs are consistent over power-up cycles for the majority of SRAM cells. Few SRAM cells, however, have a weak or no preference [12], which are referred to as fuzzy cells in this article. The authors of this study execute the approach suggested in utilizing commercial SRAMs for the first-time in. An attacker cannot just read the complete PUF arrays and use the material to expose the stored PWs in the DB since the PUF is guarded in the server. Nonetheless, the approach is subject to MIM attacks, in which the PWs can be revealed by repeatedly observing the same network traffic.

For the first time, Lamport [13] used One-way Hash Chains as a cryptographic primitive. Hash Chains have low transmission overheads and certain aspects of public-key encryption. As a result, protocol designers have utilized them in a variety of security applications, including one-time password creation and IoT device authentication [14].

The finiteness of the hash length is the fundamental problem while designing an HC-based security system. At the one side, if the HC is too long, the memory-computation cost for producing each output grows linearly with n, where n is the HC's length. A low number of n, but at the other hand, causes the HC to quickly drain, necessitating reinitialization. An attacker cannot just read the complete PUF array and then use the knowledge to expose the stored PWs in the DB since the PUF is guarded in the server. However, the approach in [13] is open to MIM attacks, in which the PWs can be revealed by repeatedly observing the same network traffic. Moreover, in the protocol described in [13], the PUF is only utilized to create PUF answers, which are then saved in the database as content.

The protocol proposed offers various advantages, including the ability for the server to authenticate users without knowing their passwords. The server, on the other hand, should hash and read the PUF many times. Long Hash Chains, on the one hand, add to the APG's memory-computation cost while linked to the server. As a consequence, the verification time is extended, and the number of customers that may be managed is lowered. A low number of N, on the other hand, causes the hash chain to quickly deplete, necessitating reinitialization. As a result, this technique is incompatible with authentication circumstances in which at minimum one of the enabling communications is a device with limited power or memory. For each authenticating request, the protocol in [15] needs reading the PUF numerous times.

Data exfiltration is a severe issue that many businesses face across the world. Data exfiltration has affected a number of important firms in recent years, including Google, Hotmail, the Pentagon, the Iran nuclear plant, and US military contractors and banks. Firewall, intrusion prevention systems, infiltration prevention strategies, firewalls, anti-virus, and anti-malware are some of the current approaches for preventing these dangers. Despite widespread use of smart devices, criminals continue to cause havoc on businesses and individuals by stealing important data.

These solutions either utilize white lists, spam filters, signature-based scanning, or behavioral analysis of programs, which are insufficient to resist assaults based on zero-day vulnerabilities, according to the findings of this investigation. Due to their multimodal approach, which includes using people or hardware within organizations, malware that conceals and devastates itself after a set period of time, the malware's ability to disguise as legitimate programs to avoid detection, and the malware's desire to speak with aggressors to gain additional payloads or directions, data exfiltration strikes in most instances very harder to identify.

2.3 Ex-filtration in healthcare applications

Another recent security development is the usage of security issue and event monitoring systems, which collect network activity from sources including IDS, NIDS, antivirus, and firewall event logs. The gathered events are then subjected to statistical correlation in order to detect potential dangers. Nevertheless, this approach is ineffective in detecting sophisticated assaults and has a limited time frame within which event linkages and hence occurrences occurring over a longer period of time would never be associated. As a result, a well-planned attack disguised as a sequence of seemingly unconnected events can never be discovered [16].

Honey pots and honey nets are used to detect malicious nodes and other wireless assaults in internet accessible and intranet environments. Honey files, on the other hand, were used to identify unauthorized access to resources, while factors which have an impact and honey users were used to trace down compromised credentials. All of these deception strategies, on the other hand, fail to account for the complex movements used by experienced and professional attackers, as [17] illustrates. Honey files, for example, are made up of scripts that run when the document is viewed and submit the information to the data acquisition system. A skilled attacker will determine not to access data and instead ex-filtrate and examine it from a machine that is not connected to the monitoring device.

The behavior of the traffic that will be travelling through the network will be observed via heuristics scanning. There will be two forms of conduct: regular behavior and abnormal behavior. If any unusual activity is detected, the algorithm will first provide an alarm before taking precautionary measures, such as cutting off the link between the

hacker-controlled command and control centers and the victim's computer systems. On either hand, traffic that behaves as predicted will be permitted to pass via the network.

The HIPAA Security Rule mandates that healthcare organizations use organizational, architectural, and technical protections to preserve the integrity, authenticity, and accessibility of confidentiality. Data encryption is essential, for instance, if data must be conveyed to that other person or organization and there is a considerable danger of unlawful disclosure. In the end, in respect of HIPAA, there is an acknowledgement that health care personnel must be ever alert in terms of privacy and security.

It promised billions of dollars to establish a nationwide accessible medical records system, imposed a data infringement reporting requirement, and expected evidence of tiered "functional use" of the system by specified dates. The US Health & Human services Services' Office for Civil Rights keeps track of unauthorized protected health care information breaches impacting 500 or more people in a publicly available database. Healthcare institutions, understandably, want the capacity to recognize when they have been hacked and are compelled to reveal this information publicly. Given the over 2,000 data breaches revealed by the US Health and Human Services data breach web site in June 2017, the prospect of public humiliation and reputational harm appears to have been insufficient.

A new interoperability electronic health record facilitating the safe sharing of healthcare information across all engaged providers across the country, as well as offering consumers online access to their own EHR, was a key project in healthcare. To be eligible for subsidy payments, healthcare organizations have to demonstrate meaningful usage. For all healthcare providers, having an interoperable EHR raises important data privacy and security concerns, such as safeguarding your own patient data, securing patient data transferred from other providers, and assuring HIPAA privacy and security conformity by cloud EHR providers.

The National Health Service also has a set of security rules and recommendations for public sector enterprises in England, Scotland, and Wales, with regional variations. Mandatory yearly information governance training and the implementation of protocols for withdrawing access privileges to information and services when an employee is dismissed are among the provisions. The Care Record Guarantee ensures that patient

information is kept private and secure, while the institute Confidentiality Code of Practice establishes guidelines for sharing information with other organizations.

2.4 Cloud based solutions in healthcare

Health organizations are transferring a variety of assignments to cloud customer services, much like the rest of the world. As a result, healthcare IT providers and other industry players are increasingly depending on cloud-based storing documents and integration services, cloud-based electronic healthcare services, and other cloud-based services to share healthcare data. In the most recent report, HIMSS Analytics saw widespread use of numerous types of cloud solutions by healthcare, with exchange of health information technology systems and back-office technologies dominating the present workload. The use of cloud services creates a security paradigm of shared responsibility, a demand for innovative ways to safeguard confidential material, and a plethora of new security problems.

Cybercrime, phishing, and security breaches all affect fundamental healthcare systems, putting a healthcare organization's capacity to provide routine and emergency treatment to patients at risk. When doctors cannot access a patient’s medical record, a health center has to turn people away patients on an emergency, nursing staff have to regress to manual methods they haven't been received training to use, the dispensary can't get timely prescription alerts, and medical equipment are locked and inoperable, it can be a matter of life and death. Patients, their relatives, and officials all anticipate near-perfect uptime and accessibility in these life-critical situations.

Medical treatment is no longer the realm of the generalist, but rather a complicated partnership involving several medical professionals from various organizations interacting through separate IT systems. Once different hospitals and outpatient clinics are taken into consideration, healthcare companies have several geographical locations. Thousands of workstations, specialty medical equipment with embedded operating systems, specialist medical software, mobile devices, and both on-premises and cloud-based services may be found in a modern hospital. An ever-changing roster of medical practitioners uses shared workstations, and the urgency of the task necessitates the usage of generic user credentials rather than individualized user accounts. This implies that systems have been left wide open. Hypersensitive patient data is constantly travelling in

and out of health care systems, thanks to the demand for interoperable electronic health records.

2.5 Threats to healthcare

- **Ransom ware**

Ransom ware is a huge threat to all businesses' data and systems, but it is particularly dangerous to healthcare organizations because to the life-or-death repercussions of being unable to operate a hospital or other institution. Modern ransom ware is capable of not only infecting the initial system, but also automatically sniffing out more susceptible targets throughout the network, so healthcare workers can't afford to make a single mistake. While medical records is one of the most useful resources on the black market, ransom ware allows thieves to get paid right away without having to sell anything.

- **Theft of one's identity**

Healthcare records provide all of the information about a person that is required for identity theft. In terms of security, the healthcare business has a dismal track record.

- **Theft of data and patient information**

Not everyone who works for your organization is an ethical and fair healthcare professional who is committed to improving patient health and expanding the organization's impact: others are concealing malicious intentions. These nefarious insiders have access to the juicy data and may have enhanced access credentials to the same systems, resulting in data breaches and patient data loss. But it's not only nefarious insiders who are harmful; it's also well-intentioned employees who leave computers logged in but physically unsecured, neglect to lock a lockbox when no one is around, or transmit a spreadsheet with PHI to the incorrect person. The industry's endemic characteristics, such as the failure to adopt encryption and the usage of unique users, are partly to blame.

- **Spear phishing**

Spear phishing is an email-based assault in which specially designed emails with malicious files or links are delivered to important employees. These emails are particularly convincing since they look to have been issued from a trustworthy source. Spear phishing is a common method used by criminals looking for specific information on medical developments or treatments. When attempting to target management, attackers will send an email that is connected to a current event or business policy. Other corporate members will be targeted with attacks that take a different structure and focus on areas such as human resources and IT updates. Spear phishing emails are exceedingly difficult to detect at the email gateway due to their highly personalized character.

2.6 Mobile devices and Healthcare

IT administrators have a difficulty with the recent surge in mobile devices and their accessibility to users for personal and business usage. Mobile devices have changed the way healthcare workers deliver treatment by allowing them to be more flexible and providing instant access to patient details, allowing them to spend more time with patients. Employees may utilize their personal devices to fulfill their jobs if IT Administrators do not deploy the relevant mobile device for the job or are hesitant to incorporate an MD into the workplace. If a healthcare practitioner uses a personal device to access patient information, such as a Smartphone, tablet, or USB drive, the device is vulnerable to theft or is not password protected.

If a hospital staff loses a patient's information, the hospital is held accountable. If a USB drive containing about 25,000 patient information is stolen or lost, the penalty to the hospital in fines might be as much as $6 million or more. Legal expenses, notification to impacted patients, and the cost of ID monitoring services are also possible penalties. 17

Sectors of health care and education health IT administrators must bridge the security and smart phone use divide. Uncontrolled mobile device access, verification of users requesting access to a hospital's web server, how to secure mobile devices that contain patient data, unsecured network connectivity or cellular networks, and protection against unauthorized breaches of lost or stolen devices are all areas of concern.

Denial of Service assaults may be vulnerable to some devices that are extremely sensitive to battery capacity, such as those that are implanted. As a result of the DOS assaults, the

device may be constantly awakened from its battery-saving "sleep" mode, shortening its lifespan and prompting the patient to have an earlier operation for its replacement than otherwise required. This threat vector can only be mitigated by internal design features that are specially designed to DOS assaults. Engineers must take this issue in mind while designing technologies that may survive DOS assaults.

Another risk in medical devices, patient management software, and general hospital IT is the theft of personally identifiable medical information, which could be given to unauthorized agencies such as the media, health insurers, private investigators, lawyers, and others, causing embarrassment to the patient or other interested parties. This private information might be exploited for nefarious purposes like profit on online forums or identity theft.

A review of data leakage identification and avoidance is presented by Shabtai et al [18]. They present a taxonomy for data leakage prevention solutions, examine numerous data leakage prevention solutions in business and academia, emphasize the causes, regions, durations and motivational factor for data loss, and at the conclusion predict future prospects for study in DLP. The literature study by Raman et al. [19] outlines the DLP problem and emphasizes several fundamental techniques from the literature. Encryption, access control, the semantic gap in DLP, and cooperation are identified as important issues, and grouping and social media network analysis are tested as prospective study fields in DLPs. Raman et al. focus on data exfiltration issues, and as a result, only three solutions from the academic papers are described.

Yang and colleagues provide a method for gathering news items and other reports about data thieves from web sources [20]. The acquired data is analyzed to discover data thief behavior patterns so that companies may be alerted to the observed trends. There are five steps in the procedure: (1) a compilation of news pieces from the internet (2) condition characterized (3) text retrieval of proper names (Name, residence, mail, and age) (4) creation of a theft record by classifying named entity recognition into categories like time, place, and loss (5) Examining the theft record to learn more about numerous aspects of data theft. This method provides individuals and organizations with some insight into hacker behavior. However, the method has drawbacks, such as the fact that it does not cover all cases of data exfiltration.

Trostle [21] is a network security system that focuses on detecting and preventing exfiltration. The author uses the example of a malevolent process having access to data that it wants to transfer out of a network, which it may do via a timing broadcaster, full effects on a vector that its outside collaborator can watch. Rather than datagram or network monitoring, they propose an encrypted mix-network, in which packets supposed to head to the same intended destination have seemingly distinct destination addresses generated for them, and source addresses are procedurally chosen quantities, making network analysis by some outside observer considerably more difficult. This 'fizzing' strategy to preventing exfiltration is similar to the data-masking remedies developed for network content management, but it applies them across the whole network.

By outlining a communication network based on the use of mobile software agents and supplying middleware to facilitate it, Van't Noordende et al.[22] advocate for a more associated with sudden of communication and information controls. Sensitive information is stored within 'restricted rooms' in its formation. Parties seeking to obtain access to a piece of the sensitive information send an autonomously agent to run remotely, which has access to personal records and may search for the specific information required by its owner. When a virtual machine has recognized the information, it wants to return towards its owner, it sends it to a gate keeping agent, who can filter the communication based on local policy concerns, collect fees, or arrange non-disclosure agreements.

2.7 Framework of traditional fingerprint methods

In present practice, the data used for authentication is stored in only one centralized database. When a user provides the needed evidence of identification, the authentication server verifies it and authorizes the user access. A user is requested to provide a password while attempting to access his account on a standard web application. The online application traditionally stores information about the user's account and password. The programmed checks the saved password to the provided password when the user provides his password during the log-in procedure. If they verify, the user is allowed application access. In other words, a single system stores all of the information required to authenticate the user.

Enrollment and recognition are the two basic steps of biometric-based authentication systems. The enrolment stage provides a digital representation of an individual's biometric feature, which is subsequently saved in a centralized system database as a biometric template. The system requires that the procured investigate biometric template be paired against an exactly as planned (in the biometric identification) or all templates (in the identification mode) stored in the centralized database during the matching process, which can actually function in two modes: validation and identification.

As a result, such systems become the single point of failure for protecting digital identities. To put it another way, if an attacker obtains access to the online application or the biometric centralized database, they can extract enough information to jeopardize the user's digital identity. Furthermore, because many users use the same pass code or biometric feature across several apps, disclosing their identity on one hacked database might lead to unauthorized access to other accounts and transactions.

2.7.1 Data Ex-filtration techniques

Data ex-filtration may be done in a variety of ways, and these tactics are growing more complex over time to remain one point ahead of data protection solutions. Some of the procedures are simple, and others are more complicated and need the use of tools to apply to the files in question. The authors will use a variety of scenarios in this study, including encrypted data, combine broadcasting, and formatting file executable. The section provides an overview of the data ex-filtration methods used in this study in order to highlight the flaws.

Encryption is now commonly used to safeguard communication, personal information, and other sensitive data. Furthermore, as seen in the very first second cases, encryption may be employed to prevent unwanted data transit. In the first situation, the test files were encrypted with the software Kleopatra using a Windows version of the GNU Privacy Guard utility; in the second scenario, the data were encoded with WinRAR 256 AES encrypted communications.

Merge streamed methods were utilized in the third, fourth, and fifth situations to simply decompress, modify, and modify the test files' extensions.

On the test files, the split package capability was employed in the sixth manner. Dividing a file system is a way of assisting with huge downloads by dividing the file to smaller portions of a size specified by the user. The next option was to append the sensitive test file into a non-confidential file using the TYPE command to show the received data into yet another file, which is accessible in Windows operating systems. Using Win hex software, eight methods of hex editor software were utilized to change the file's binary using the forms of (PDF, DOCX). The procedure entails opening the test data in a hex editor and erasing the file's starting in a technique known as eliminating magic number.

2.7.2 Existing Prevention tecniques

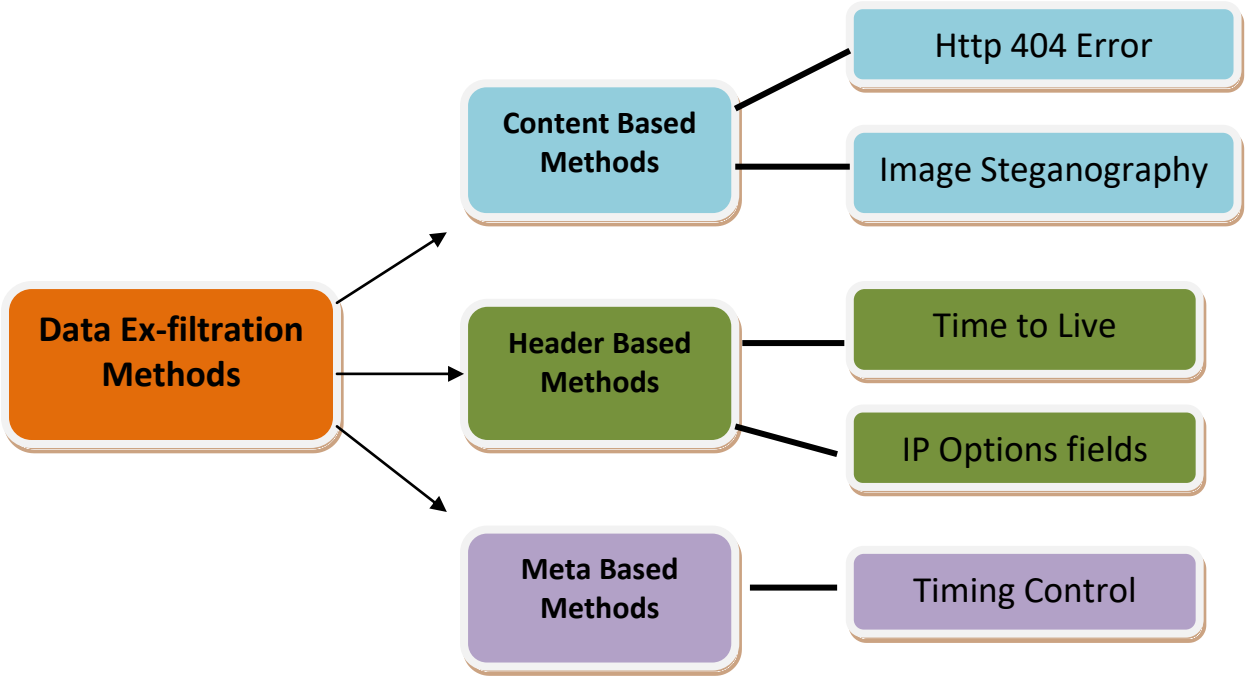

Figure 1 Existing Preventive Tecniques

2.8 Physical Unclonable Function

The concept of identifying items, systems, and people based on fundamental random physical properties is not new. Human fingerprint identification stretches back to at least the eighteenth century, spawning the discipline of biometrics. Random sequences in paper and optically tokens were utilized for unique identifier of money notes and strategic weaponry in the 1980s and 1990s. Physical one-way variables, physical randomized functions, and eventually physical non-clone functions were developed at the dawn of the twenty-first century as a formalization of this notion. In the years after its launch, a growing number of different types of PUF's have been suggested, with a trend toward more integrated designs.

PUFs are unique dissimilarities that are introduced in a hardware device or chip during its manufacturing. They are intrinsic to a device. Just like a fingerprint is unique for every person, so is the PUF for a hardware chip. The main feature of PUFs is that they are unclonable. This gives PUFs an edge over other hardware-based security systems as the hacker cannot clone an intrinsic feature of a device. PUFs work on the principle of Challenge-Response Pairs (CRP). A challenge (an external stimulus) is applied, and a response (an output) is received. For every different challenge there will be a unique response that will be different for different devices even if the same challenge is applied.

Moreover, PUFs can provide a low-cost solution for generation of cryptographic keys from the device rather than the conventional methods, where the secret keys are produced and distributed by the server and stored in the device memories. PUFs are mostly used in IoT based security systems as it is a low cost, light weight, and robust solution.

Our aim is to combine features of PUFs with fingerprint-based authentication for healthcare application to prevent data exfiltration mechanisms and present a novel framework or a solution that can help in mitigating multiple data exfiltration attacks using a single approach.

PUFs are used for two basic purposes:

- Authentication at a cheap cost
- Key creation that is secure

2.8.1 Characteristics of Physical Unclonable Function

- Robustness in the face of potential threats.
- In cryptography applications, good statistical qualities are critical (CRPs uniqueness and uniformity).
- Number of CRPs compared. Area Occupied (strong PUF: exponential increment in the number of CRPs by increasing utilized computational resources vs. weak PUF: linear increment in the number of CRPs by increasing utilized computational resources)
- FPGA implementation is simple (attempt to make it possible to adapt the PUF after implementing the IoT devices and updating it to newly encountered attacks).

Traditional PUF designs, such as RO PUFs, claimed unclonable fingerprints but were subject to deep learning attacks and altered physical attributes. Newer designs have made significant strides in resolving these challenges. They are still not well understood and must be put to the test from many angles in order to assess their benefits and drawbacks. The goal of PUF protocols was to allow lightweight safe authorization for Internet of things without requiring encryption or secure memory. However, at minimum a few of these two approaches is used in the majority of the suggested protocols.

2.8.2 Types of PUFs

The fact that PUFs created during the last decade have generally split into two broad categories has resulted in these two applications. "Strong PUFs" and "weak PUFs" are the two types of PUFs. Authentication is usually done with strong PUFs, whereas key storage is done with weak PUFs.

2.8.2.1 Weak PUF

Weak PUFs, also known as physically obscured keys, are the first class of PUFs that take use of manufacturing variance. These PUFs might be viewed as PUFs that effectively digitize a circuit's "fingerprint." A digital signature is created as a result of this precise measurement, which may be utilized for cryptographic reasons. The PUF can only be

questioned by one or a tiny proportion of tasks since the fingerprint signature stay mostly unchanging. This relates to having a scope of one or a limited number of inputs in the black-box definition above. As a result, it will have a relatively narrow range, as each challenge should always provide the same answer. Weak PUFs have the following characteristics:

1. a limited percentage of CRPs (linearly connected to the number of features whose behavior is affected by manufacturing variation); a limited percentage of CRPs (sequentially related to the number of modules whose behavior is affected by manufacturing variation); and a limited number of CRPs (linearly pertaining to the components whose behavior is affected by
2. The reaction is consistent and resilient in the face of changing environmental circumstances and many readings, ensuring that a problem always produces the same result.
3. Responses are unexpected and heavily influenced by the device's inherent manufacturing variability.
4. Manufacturing two gadgets with much the same actual fingerprint is impractical.

2.8.2.2 Strong PUF

Strong PUFs are distinguished from weak PUFs by their ability to sustain a high number of CRPs. As a result, a powerful PUF may be directly validated without the use of cryptographic hardware. The following are the recommendations for a strong PUF: a large sufficient based authentication space so that an opponent cannot encompass all CRPs in a fixed amount of time (preferably, exponential in the number of challenge bits); a large sufficient based authentication space and that an opponent cannot encompass all CRPs in a fixed amount of time (presumably, exponential in the number of challenge bits);

1. Multiple readings, steady responses to the environment
2. With a quadratic sample of dynamically determined CRPs, an adversary cannot foresee the response to a fresh, randomly chosen challenge.
3. It is not possible to produce two PUFs with identical reactions.

4. The readout just shows the answer and no further information about the PUF's internal operation.

2.8.3 Fingerprint based PUFs to authenticate users

The majority of computer hackers in the industry sector have centered on data loss. However, because sensors enable IoT devices to communicate with the rest world, cyber-attacks can now invade the physical world, posing a risk of bodily injury. A chemical facility, a nuclear power plant, or a multibillion-dollar industrial process can all be destroyed, putting the lives of its employees and neighbors in jeopardy.

In this case, ciphering data (which decreases the risk of transmitting erroneous data) could not be enough. Counterfeiting must be avoided at all costs, as false sensors seldom exceed the high standards set by genuine sensors. Fake sensors can also be used as a backdoor for industrial espionage.

If the sensor networks presented here are duplicated, the resulting phony sensors will not be able to imitate them since the manufacturing attributes of the device cannot be cloned to make the same PUF. They are still unable to extract the security code from the stolen Helper Data in this situation. Because many sensors used in important operations must have high accuracy and specificity, they should indeed be validated and checked on-site by qualified specialists on a regular basis. It is critical to avoid impersonation attempts by adversaries posing as qualified specialists during calibration and testing. Otherwise, attackers or rivals might be able to manipulate sensor behavior for nefarious purposes. It's also crucial to avoid repudiation assaults since an authorized professional might turn malignant.

2.8.4 Applications of Finger based PUF devices

- **Compliance of Regulation**

Many rules aimed at preventing occupational and environmental hazards stipulate those specific variables must be measured to reasonable standards (for example, levels of noise, radioactivity, carbon-dioxide emissions, etc.). Inspectors of the environment, occupational health, and safety conduct on-site analyses of such factors in order to protect

the environment, people, and employees. They wear network devices that have been calibrated and validated so that they can capture the required measurements.

Because the sensors presented here relay ciphered observations to an access point, data manipulation to circumvent rules is more difficult. In this case, ciphering might not be sufficient. Fake sensors that collect misleading data, as well as imposter inspectors that pose as actual inspectors should be avoided since otherwise, unsafe situations can be certified as safe.

- **Soldiers uniform adherence**

Sensors integrated into army uniforms lessen the weight load put on soldiers and provide them additional opportunity to complete tasks effectively. Sensors may track troops' whereabouts as well as their health and physiological characteristics, such as pulse rate, respiration, and blood pressure. Rapid access to this data might enable mission commanders and soldiers survive otherwise devastating enemy strikes. This information should be delivered in ciphered form since it is sensitive. Cryptanalysis, on the other hand, might not be enough. In this circumstance, fake devices should not be employed. An attacker cannot imitate a soldier using the suggested sensors since their biometric traits do not match the authentic soldier's template.

- **In controlling of Arms**

The issue of weapons control is extremely important for peace, security, and stability. To prevent their careless or criminal usage, production, and trafficking, arms should be recognized. Ballistic testing (length, course, and conduct of their bullets) or detecting if an arm wanders outside a specified region can both benefit from a sensor node inserted in an arm. Arm licenses are used to identify those who are allowed to use arms.

If the suggested sensor nodes are placed in arms, they will be able to recognize both the arm and its user. The arm may be proven lawful since the sensors show their validity. False data supplied by unauthorized guns can be discovered more quickly if they send encrypting measurements.

- **Healthcare system and PUF**

Everything (even live objects) will be accessible, felt, and interlinked inside the worldwide, adaptive, living architecture of the Internet, according to its vision. The following three areas of protection in a technological based platform should be taken into account: infrastructure, connectivity, and system model. Hardware security refers to the physical cyber security of PUF equipment, whereas communication security refers to the security of Unclonable applications. Depending on the system paradigm, the cyber security of each IoT technology may differ. For example, the privacy of the concerned entities should be addressed in certain applications (e.g. VANET, cellular technologies, finger-based healthcare system), whereas secure data ability to connect and key management are necessary in many applications.

Recently, systems such as e-healthcare monitoring have gotten a lot of interest from the scientific community, where the device's security and efficacy are critical.

2.8.5 Techniques to protect data in healthcare applications

These healthcare cyber security best practices are designed to stay up with the changing threat landscape by tackling risks to privacy and confidentiality on terminals and in the cloud, as well as protecting data in transit, at rest, and in use. This necessitates a multifaceted and smart security strategy.

- **Educate Healthcare Employees**

The human factor remains one of the most serious security hazards in all businesses, but especially in healthcare. For healthcare companies, simple human mistake or neglect may have severe and costly effects. Security awareness training provides healthcare workers with the information they need to make informed decisions and exercise proper caution while managing patient data.

- **Access restriction and authentication**

User authentication is required to ensure that only authorized have access to secured data. A preferred option is multi-factor authentication, which requires users to verify that they are the person allowed to access particular data and apps using several or more verification methods, such as:

- User-specific information, such as a pass code or Security code
- Anything that only the authorized user has, such a card or a key Biometrics or other features specific to the authorized user (facial recognition, fingerprints, eye scanning)

- **Implement Data Usage Controls**

Protective data restrictions go beyond access restrictions and surveillance to guarantee that potentially harmful or malicious data activity is reported and stopped in real time. Data controls can be used by healthcare companies to prevent sensitive data from being uploaded to the internet, sent via illegal email, copied to external storage, or printed. Data discovery and categorization are critical components of this process because they allow sensitive data to be recognized and marked for the appropriate level of security.

- **Record and Monitor Your Use**

All accessibility and user information must be logged in order for providers and business partners to see who is accessing what content, apps, and other facilities, whenever, and from which gadgets and regions.

- **Data Ex-filtration research**

Both direct and indirect players in a company can engage in data ex-filtration. Direct and indirect actors were accountable for 69 percent and 34 percent of data ex-filtration events, respectively, according to Verizon research published in 2019. In recent years, data ex-filtration actors have evolved from individuals to organized groups, sometimes even backed by a nation-state with significant money and resources. Data ex-filtration instances linked to nation-state actors, for example, has climbed from 12 percent in 2018 to 23 percent in 2019.

Individuals, commercial companies, and government institutions are all affected by data theft, which costs an average of $3.86 million each incidence. It may also have a significant influence on a company's reputation. Furthermore, surrendering vital data

(e.g., sales possibilities and new product data) to a competitor might put a company's survival at risk. Since events involving nation-state actors with large budgets and resources have become more common, it may leak national security secrets.

2.9 Preventing techniques from data ex-filtration

- **Inappropriate transmission routes should be blocked**

The fewer channels you make accessible for data access, the less likely those paths will be exploited as a channel for data ex-filtration, either mistakenly or deliberately. Consider shutting all unapproved communication channels, ports, and protocols by default as a best practice, then enabling them as needed. This strategy provides a more secure standards and guidelines than one in which all entryways are allowed by default, which could also lead to difficulties like users forgetting to switch down a server's HTTP service, which can result in data being accidentally exposed over the web.

- **Avoid being a victim of phishing scams**

Phishing is a frequent method of obtaining harmful data. Preventing phishing attempts is thus a critical step in reducing the danger of data ex-filtration. To that end, it's critical to teach staff how phishing attacks operate, how to detect one, and what to do if they suspect they're being targeted. Standard practices for protection analytics techniques that can create warnings when they recognize emails, text messages, and other content that might be used in phishing attacks also aids in the prevention of this form of attack.

- **Ex-employees' data access should be revoked in a systematic manner**

When an employee's connection with a firm expires, the employee's access to IT systems should be disabled promptly. The same should be true of business partners or suppliers who may have access to internal systems while working with a firm but should no more have that access after the relationship is over. Don't put off "cleaning up" old accounts for another week or month. Make the procedure a standard component of the departing procedure.

- **Employees should be educated**

Because technological tools and automation may only go so far in stopping employees from sharing data with unauthorized parties, employees must be educated on corporate rules on data sharing as well as data security best practices.

Software solutions, on the other hand, can be used to assist detect unlawful data sharing, such as odd network behavior that signals an employee has linked a personal computer to the internet and is transmitting data to it.

Chapter 3

Proposed Methodology

3.1 Methodology

Organizations all around the world are generating massive volumes of data, which they must store and safeguard against fraud, destruction, and abuse. Data is very important asset for the company and every organization working therein. Data leak dangers exist whether the data is housed on-premises or in the cloud. These risks can come from hostile insiders with authorized access to company assets, as well as malevolent outsiders with no such authorization.

3.2 Fingerprint Authentication

What is Fingerprint well it used for both devices and human beings. Devices have unique attributes that can used with security advancements to secure user data. Physically Unclonable functions are one of them that works as device fingerprint. Now moving towards human fingerprints which we can also be called as biometric. In a progressively digitized world, password-based validation is no longer satisfactory to secure applications and software programs. Can biometrics substitute conventional passwords in user verification? Biometrics acts as safe and easy mode of authenticating users without compromising their virtual experience.

3.3 Key generation for encryption through mobile sensors

A secure service must have device authentication and key management for encrypting communication data. Based on device identity collected by a PUF, we may implement safe authentication. To avoid keeping the encryption key in a device, for instance, PUF is employed as a key generator. To extract device identification, nevertheless, conventional PUFs need specialized hardware or software. Therefore, in a situation where there are several devices and device makers, it might not be practical to adapt existing PUFs to internet of things devices. As an alternative to PUF, we may utilize the typical values of the current sensors in an internet of things device.

PUF responses are generally noisy and of low-entropy, a PUF-based key generator faces two main challenges: increasing the reliability to a practically acceptable level and compressing sufficient entropy in a fixed length key. Fuzzy extractors [7] perform exactly these two functions and can be immediately applied for this purpose, as suggested in several earlier PUF key generator proposals.

3.4 System Workflow

The below diagram illustrates the complete system workflow. The major key points are mentioned below.

1. User Login is Secured with two factor authentication. First factor is biometric authentication and other one is phone number and password. Further security of authentication is enhanced by using user behavior by implanting PUFs sensor coordinated while login.
2. After successfully logged into system the user can access his data files by pressing the download button. Here the user must generate the symmetric encryption key in order to see his uploaded document. The symmetric encryption key is generated based on the PUF values and user biometric value.

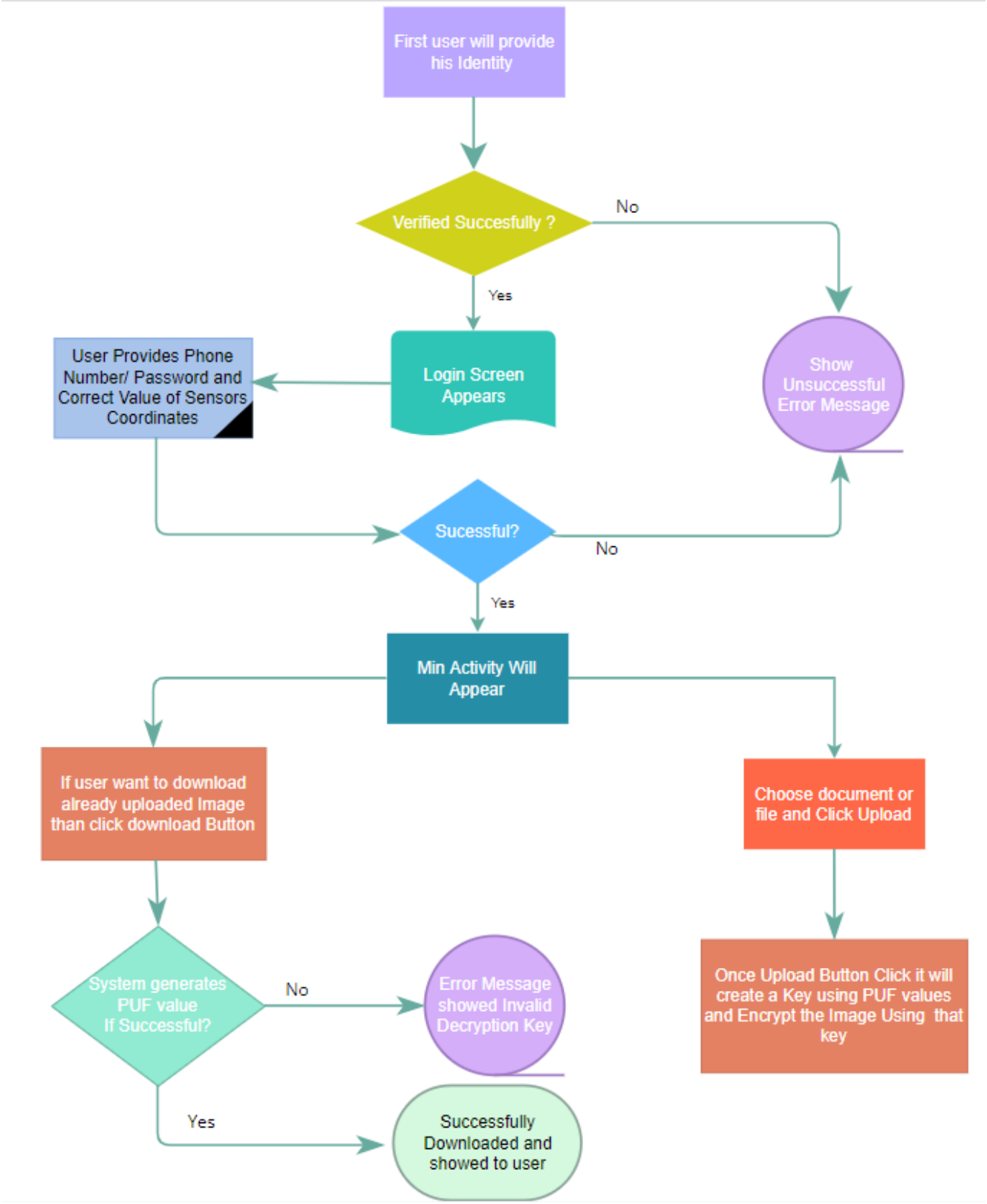

Figure 2 System Implementation Methodology

Design Goals

The research is focused on bringing latest and secure features to users to protect the security and privacy of their data. The application features can include following features in the android application framework in order to provide enhanced security protection mechanisms.

- Data Exfiltration attacks for mobile applications
- Using built-in sensors to generate key.
- Key theft solution by generating encryption key at runtime
- Behavioral based access to the application

Chapter 4

Implementation

4.1 Introduction:

This chapter discusses the implementation of secure mechanisms against data exfiltration and data siphoning so that data can be secure from malicious applications in android mobile devices. Everyday excessive security issues cause us trouble. On daily basis many malware attacks and dangerous applications in mobile extract unsecure documents and other useful data without our permission and thus ePHI and other sensitive data get breached, and companies has to pay huge fines imposed by department of Health and Human Services. We are going to implement one such application which will keep all the application documents in your mobile safely by encryption technique using physically unclonable functions. Provides your data high security while in memory or stored on devices. When you want to decrypt, just connect with your database, and save all your data with most easy way.

4.2 Implementation Workflow.

There are their different parts which are used to process data for overall scenario.

1. Database (Which is used to for storing credentials and sensitive data in secure manners.)
2. API (Application-Programing-Interface for interaction among database and application.)
3. Android Application (which provide users to upload and see data in secure manners)

The Webservice or API is present on the secure computer or server to get calls from different mobile applications. In this we have written all the functions so that malicious actor cannot get the actual implementation of the functions or modify it as per their needs.

The mobile application which preserves trail of user's behaviors using physically unclonable functions such as gyroscope and accelerometer. As well as who they are by verifying their fingerprints using fingerprint sensor in mobile devices.

The mobile application is responsible for generating secure Keys for encryption and decryption of data uploaded by it. Or Gives users a smart login which includes PUFs readings to get into the application. Further implementation is described below with screenshots attached.

4.3 Screenshots:

1. The application for android is built with following features. To get a successful authentication user has to prove his identity so that he could use application if the user failed to prove his identity, he cannot move further to get access to the documents.

Figure 3 Biometric Screen

2. After Fingerprint Authentication it moves to the Login verification form, which has been shown in Figure 2.2 below. At this step, it requires two more security checkpoints for making data secure and reliable.
    a. The unique identity in our case it phone number and secure password which includes NIST standard (minimum 8 characters with at least one uppercase, special character and number)
    b. Further It also substantiate all coordinates of PUFs I.e. accelerometer and gyroscope sensor average values.

    Successful verification of the above provided information leads users toward next activity where they can upload, and view images based on successful key generation. Below screen demonstrate the login screen.

Figure 4 Login Page Activity

3. After completing all verification processes, it shows phone number, enable smart security radio button, accelerometer values, update button, search image, take picture, upload image, and download image button in Figure 3.1.
    a. If you want to update your number, enter new number.
    b. When we enable smart security radio button for this phone number, it will enable fingerprint authentication.
    c. Remember all your accelerometer sensor values before update user details. Press Update button if you are doing any change in phone number.

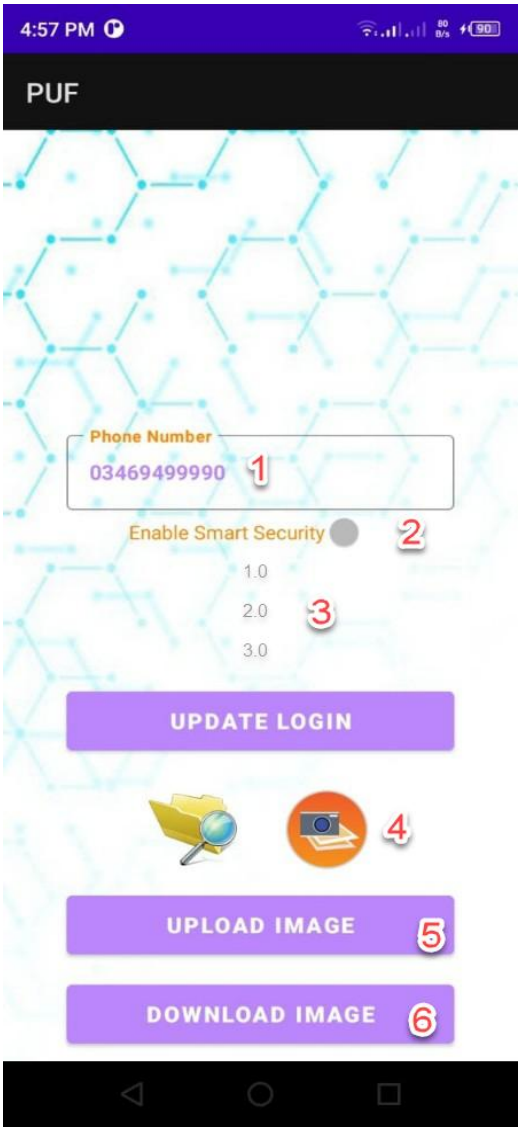

Figure 5 Upload & Download Activity Screen

4. For upload an image, press search a file button or take image as highlighted in above screenshot. After clicking for the first time, it asks for the permissions for taking access to gallery and upload image to database in encrypted form. After giving the required permission it takes us to the files where we can select the required document to be uploaded in the application as shown in figure 3.2.

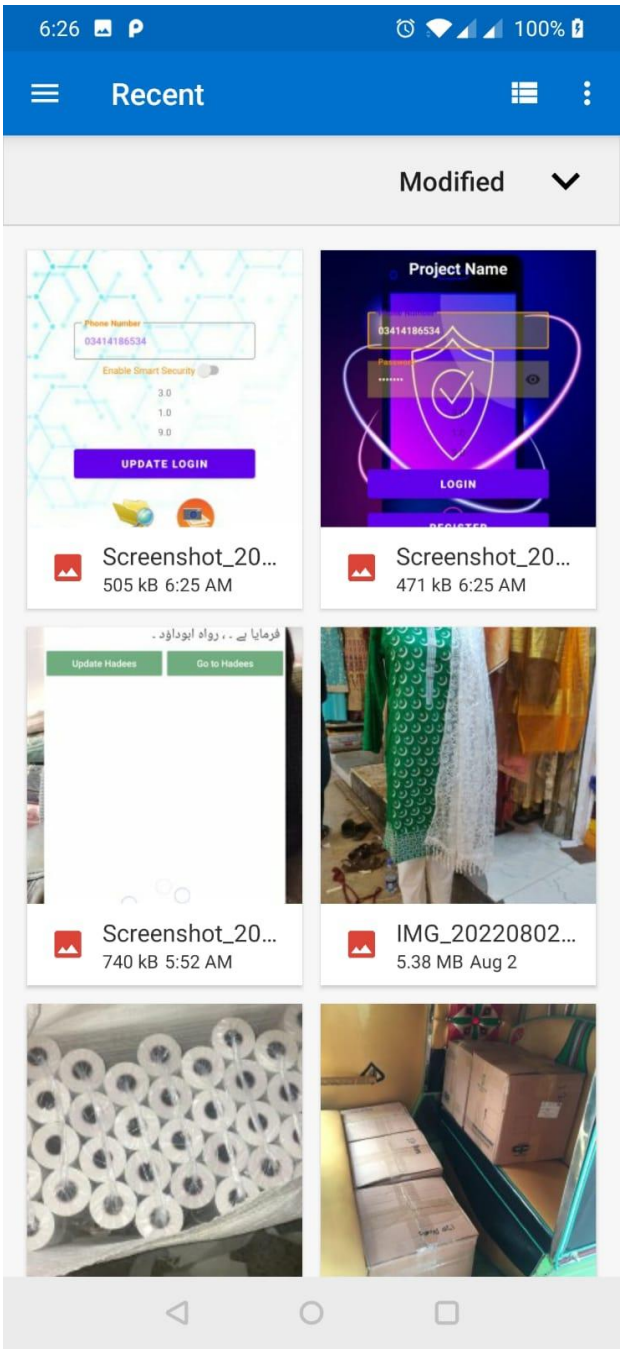

Figure 6 Data Upload Screen

5. Select required image or document and click the upload image button. After select the image for upload in database, our image will show in a small image view where we can verify the selected image. Press upload button for uploading it to the database in a secure way. When user click the upload image the function first creates a key using unique identity and PUFs values and then

encrypt the image with the key generated your document/picture secure in database. We can also see this uploading process in Figure 3.5

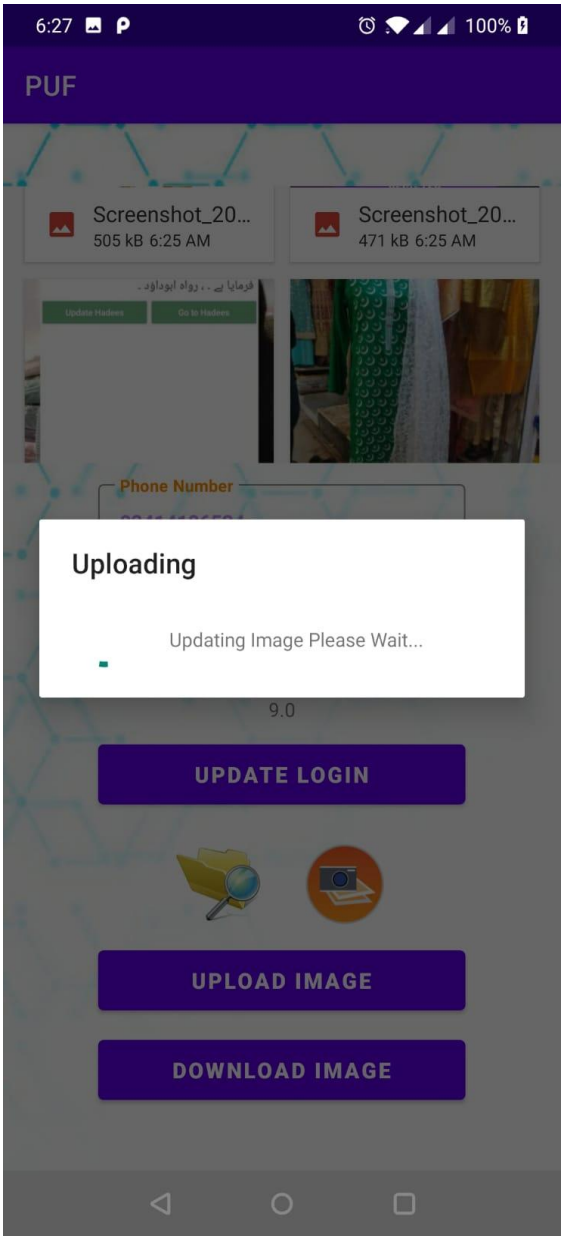

Figure 7 Key Generation While Uploading

6. When we required to access our encrypted image or document login back into the application and this time click the download image. After the successful verification of unique identity and PUFs values generate the same key again and decrypt the image back into the user readable format. The below figure shows the download function of the application.

```
btnDownloadImage.setOnClickListener(new View.OnClickListener() {
    @Override
    public void onClick(View v)
    {
        ProgressDialog progress = new ProgressDialog( context: MainActivity.this);
        progress.setTitle("Verification");
        progress.setMessage("Downloading Image Please Wait...");
        progress.setCancelable(false); // disable dismiss by tapping outside of the dialog
        progress.setIndeterminate(true);
        //progress.setIndeterminateDrawable(loadingView);
        progress.show();

        downloadImage(phoneNumber.getText().toString(),txtX.getText().toString(),txtY.getText().toString(),txtZ.getText().toString(),

    }
```

Figure 8 Encryption Functionality Using PUF values

Chapter 5

Results and Analysis

5.1 Sensors Unique Values

In our thesis we used different mobile sensors such as gyroscope and accelerometer values to encrypt the documents and files in mobile devices. To start with these sensors, I have performed different test in order to get the validity of values against same brand as well as different brands.

1. When validating the values on same model in my case **Oppo A35** we get the different values of sensors on same place this indicates that every device has unique PUF sensors values based on their ICs. As seen in the screenshot below that we have to check for the values of sensors up to 2 decimal values so that it will give unique results for each coordinate. For reference screenshots are attached below.

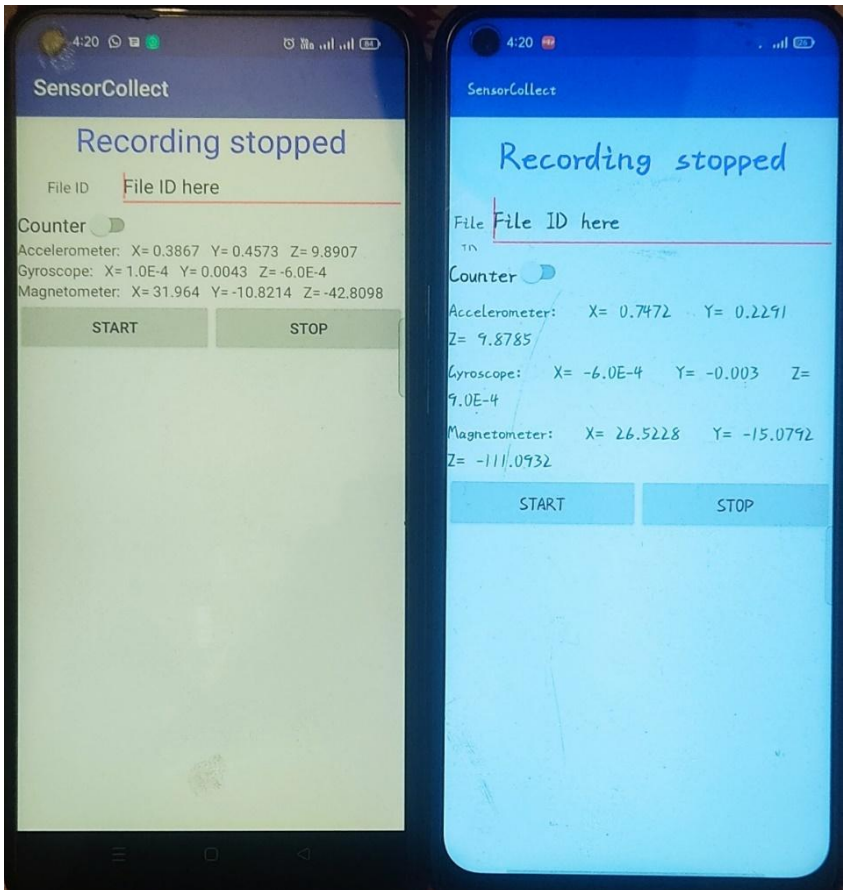

Figure 9 Same Device PUF value Comparison

2. When comparing sensor value with different model devices such as **Oppo A35** and **Redmi Note 9s** in my scenario and we get the desired results. Both mobile devices got different values of gyroscope and accelerometer values on same surface. Results can be seen in below images which varies from device to device. Accelerometer and Magnetometer values are used in order to get unique key generation.

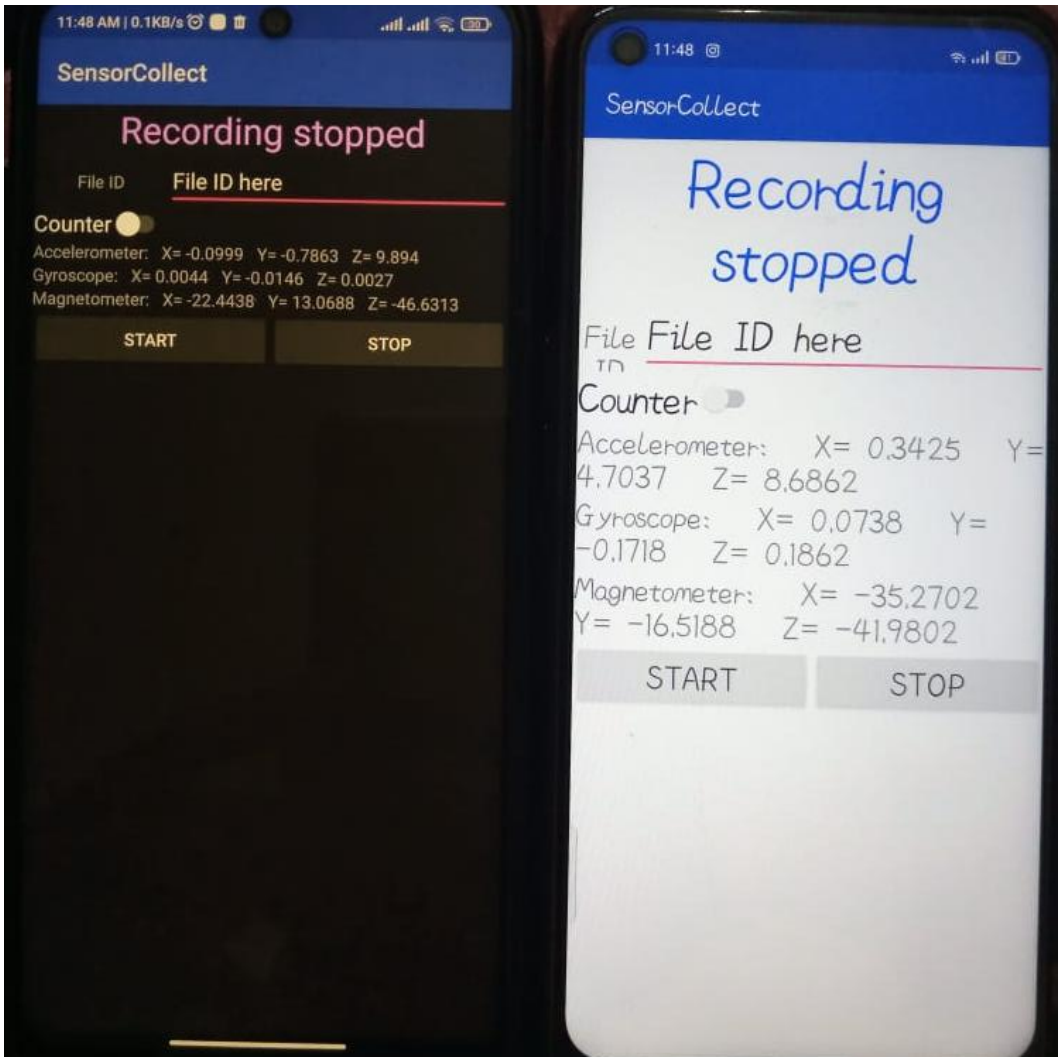

Figure 10 Different Android Models PUF value comparison

5.2 Key Entropy

As we have used above mentioned sensors values in key generation for documents or files encryption and decryption mechanisms. The reason we used these PUFs is that mobile devices have less computational resources and hence cannot do computations to generate strong symmetric keys for encryption. To validate the key strength, we need some parameters such as entropy values of different generated keys. The screenshot below illustrates the key entropy which shows the key randomness and non-guessable.

5.3 Key Security

Different mobile applications are vulnerable to key theft attacks as malicious actors continuously looks for keys in mobile application because mobile application storage areas lack security and thus vulnerable to key theft attacks. In our application we are generating key on runtime and perform further operations. The Key generation is based on device fingerprint value plus unique attributes. As the Image below shows that when download image button click it generates the key at runtime and decrypt data.

```
0 references
public DownloadImageViewModel DownloadImage(String PhoneNumber, String AccelerometerX, String AccelerometerY, String AccelerometerZ,
{
    SHA256 sha256 = SHA256.Create();
    byte[] inputBytes = Encoding.ASCII.GetBytes($"{PhoneNumber}-{AccelerometerX}--{AccelerometerY}-{AccelerometerZ}-{GyroscopeX}--{(
    byte[] outputBytes = sha256.ComputeHash(inputBytes);
    string hash = Convert.ToBase64String(outputBytes);
    string shastr = Convert.ToBase64String(sha256);
    string Encrypt = AES(outputBytes, hash);
    // 8u390mLzfP8rqqoJGYsbwQW6NbROrLvHTgYFXF0opp4=
    // 8u390mLzfP8rqqoJGYsbwQW6NbROrLvHTgYFXF0opp4=

    DownloadImageViewModel downloadImage = new DownloadImageViewModel();
    repo = new DataBaseRepo();
    var user=repo.downloadImage(PhoneNumber);

    if (user.Hash.Equals(hash))
    {

        downloadImage.Message = "Done!!";
        downloadImage.Image = user.Image;

    }
    else
    {
        downloadImage.Message = "Security Error Unable To Decrypt Image!!";
    }
    return downloadImage;
```

Figure 11 Decryption using PUF values to extract Image in Presentable Format

5.4 Insecure Data Storage

Mobile applications are more prone to data theft attacks as SQLite and storage preference in mobile applications are used to store data which lacks security components and solely rely on users to make it secure. All the healthcare reports and data that kept secure from attackers need to be encrypted and stored in such a way that it should not be accessed by unauthorized users. We use secure storage of data, documents, files, or images in encrypted database. So, in case if any bad actor gets the data, it still of no use for them as its encrypted with strong encryption mechanism.

Chapter 6

Conclusion & Future Work

6.1 Conclusion

The overall design goals mentioned in the research objectives as well as in the proposed methodology are fulfilled in this research. Such as key theft in mobile application are mitigated through in-built sensors such as gyroscope and accelerometer fingerprints to generate symmetric key for encryption and decryption at runtime. Security features which allow the secure execution of application and which protect application and personal data stored on a device [37].

As we have seen that our scope is limited to healthcare application. This solution can further be implemented to other sensitive applications which stores user data in secure manner by generating encryption keys at runtime. The enhance Security features includes PUFs to generate secret key includes sensor coordinates values. Our approach will help industry to implement new security features in login mobile applications that will include user behavior and unique attributes such as fingerprint to get access to sensitive data to only authorized users.

The biometric Identifying a person based on their behavioral and biological qualities in an automated manner is called biometrics. The authentication system substituting traditional password and token for authentication and relies gradually on biometric authentication methods for verification of the identity of an individual. This proves the fact that society has started depending on biometric-based authentication systems. Security of biometric authentication needs to be reviewed and discussed as there are multiple points related to integrity and public reception of biometric-based authentication systems with integration of physically unclonable function [40].

Users employing a smartphone or laptop as BYOD can be integrated into such devices as a data loss prevention technique. Organizations such as banks, law enforcement, and military can take advantage of the mechanism and internally manage all smart phones etc. that while acting as access control devices, will also facilitate higher functions (such

as accessing online reports and commenting) that are currently not possible with access control cards. This is already in practice but without the essentially required security and auditing trail functionality being offered by the suggested mechanism.[25]**.**

6.2 Future Work

For future the implementation can further enhanced by adding additional PUF sensors to enhance the security of key generation technologies for encrypting documents and files. Based upon requirements this implementation can be used by other security concern organizations to enhance their authentication management. Further it can also help BYOD to secure data using this technique. This implementation can also be applicable to IoTs for authentication and transferring data in a secure manner. The same can be implemented in IOs applications to enhance security of key theft and other authentication scenarios.

References

[1] F. Ullah, M. Edwards, R. Ramdhany, R. Chitchyan, M. A. Babar, and A. Rashid, “Data exfiltration: A review of external attack vectors and countermeasures,” Journal of Network and Computer Applications, vol. 101, pp. 18–54, 2018.

[2] A. Shabtai, Y. Elovici, and L. Rokach, A survey of data leakage detection and prevention solutions. Springer Science & Business Media, 2012.

[3] C. S. Kruse, B. Smith, H. Vanderlinden, and A. Nealand, “Security techniques for the electronic health records,” Journal of Medical Systems, vol. 41, no. 8, p. 127, 2017.

[4] Z. El Ouazzani and H. El Bakkali, “A classification of noncryptographic anonymization techniques ensuring privacy in big data,” International Journal of Communication Networks and Information Security, vol. 12, pp. 142–152, 2020.

[5] D. Preethi, N. Khare, and B. K. Tripathy, “Security and privacy issues in blockchain technology,” in Blockchain Technology and the Internet of :ings, Apple Academic Press, New Jersey, NJ, USA, 2020.

[6] Risk Based Security, “2021 Mid year data breach QuickView report,” Risk Based Security, Richmond, VA, USA.

[7] M. Verdonck and G. Poels, “Comparison of the categories of MIoT and the privacy-by-design principles. Security and Communication Networks International Conference on Business Process Management, Rome, Italy, September 2021.

[8] H. Zhong, Y. Zhou, Q. Zhang, and Y. Xu, J. Cui, An efficient and outsourcing-supported attribute-based access control scheme for edge-enabled smart healthcare," Future Generation Computer Systems, vol. 115, pp. 486–496, 2021.

[9] W. S. Janzen, "Iterated password hash systems and methods for preserving password entropy," ed: Google Patents, 2014.

[10] R. Arenburg, S. Chawla, A. Mathur, and C. Skawratananond, "Method, apparatus and program storage device for providing a secure password manager," ed: Google Patents, 2007.

[11] A. Vijayakumar, V. Patil, and S. Kundu, "On improving reliability of SRAM-based physically unclonable functions," Journal of Low Power Electronics and Applications, vol. 7, no. 1, p. 2, 2017.

[12] M. Mohammadinodoushan, B. Cambou, C. R. Philabaum, and N. Duan, "Resilient password manager using physical unclonable functions," IEEE Access, vol. 9, pp. 17060 - 17070, 2021.

[13] L. Lamport, "Password authentication with insecure communication," Communications of the ACM, vol. 24, no. 11, pp. 770-772, 1981.

[14] A. Pinto and R. Costa, "Hash-chain-based authentication for IoT," ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, vol. 5, no. 4, p. 43, 2016.

[15] S. Assiri and B. Cambou, "Homomorphic password manager using multiple-hash with PUF," Cham, 2021: Springer International Publishing, in Advances in Information and Communication, pp. 772- 792.

[16]. Gustav L. (2016). Bypassing modern sandbox Technologies. Masters Thesis, Department of Electrical and Information Technology, Faculty of Engineering, LTH, Lund University. (pp. 1-94)

[17]. Nikolaos V. (2015). Detecting Advanced Persistent Threats through Deception Techniques. Phd Thesis, Information Security and Critical Infrastructure Protection (INFOSEC) Laboratory. Department of Informatics. (pp. 1- 174).

[18] Shabtai, A., Y. Elovici, and L. Rokach, A survey of data leakage detection and prevention solutions. 2012: Springer Science & Business Media.

[19] Raman, P., H.G. Kayacık, and A. Somayaji. Understanding data leak prevention. in 6th Annual Symposium on Information Assurance (ASIA'11). 2011.

[20] Yang, Y., M. Manoharan, and K.S. Barber. Modelling and Analysis of Identity Threat Behaviors through Text Mining of Identity Theft Stories. in Intelligence and Security Informatics Conference (JISIC), 2014 IEEE Joint. 2014. IEEE.

[21] Trostle, J. Applying network address encryption to anonymity and preventing data exfiltration. in MILCOM 2008- 2008 IEEE Military Communications Conference. 2008. IEEE.

[22] van't Noordende, G., F.M. Brazier, and A.S. Tanenbaum. Guarding security sensitive content using confined mobile agents. in Proceedings of the 2007 ACM symposium on Applied computing. 2007. ACM.

[23] Guo, Yunxi, and Akhilesh Tyagi. "Voice-based user-device physical unclonable functions for mobile device authentication." *Journal of Hardware and Systems Security* 1, no. 1 (2017): 18-37

[24] Maes, Roel. "PUF-based entity identification and authentication." In *Physically unclonable functions*, pp. 117-141. Springer, Berlin, Heidelberg, 2013.

[25] Rahim, Kashif, Hasan Tahir, and Nassar Ikram. "Sensor based PUF IoT authentication model for a smart home with private blockchain." In *2018 International Conference on Applied and Engineering Mathematics (ICAEM)*, pp. 102-108. IEEE, 2018.

[26] Chaterjee, Urbi, Debdeep Mukhopadhyay, and Rajat Subhra Chakraborty. "3PAA: A private PUF protocol for anonymous authentication." *IEEE Transactions on Information Forensics and Security* 16 (2020): 756-769.

[27] Price, Nathan. *How to generate repeatable keys using physical unclonable functions: Correcting PUF errors with iteratively broadening and prioritized search*. University of Maryland, Baltimore County, 2014.

[28] Yamasaki, Norikazu. "Evaluation of Software PUF Based on Gyroscope." In *Information Security Practice and Experience: 15th International Conference, ISPEC 2019, Kuala Lumpur, Malaysia, November 26–28, 2019, Proceedings*, vol. 11879, p. 232. Springer Nature, 2019.

[29] Almuhaideb, Abdullah M., and Shikah J. Alsunaidi. "Sensor-based identification to detect counterfeit smartphones using Blockchain." *Journal of Ambient Intelligence and Humanized Computing* (2022): 1-18.

[30] Roy, Sourav, Dipnarayan Das, Anindan Mondal, Mahabub Hasan Mahalat, Suchismita Roy, and Bibhash Sen. "PUF based Lightweight Authentication and Key Exchange Protocol for IoT." In *SECRYPT*, pp. 698-703. 2021.

[31] Maes, R., A. Herrewege, and I. PUFKY Van Verbauwhede. "A Fully Functional PUF-Based Cryptographic Key Generator." *Proceedings of the BT—Cryptographic Hardware and Embedded Systems—CHES* (2012).

[32] Khalafalla, Mahmoud, and Catherine Gebotys. "PUFs Deep Attacks: Enhanced modeling attacks using deep learning techniques to break the security of double arbiter PUFs." In *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 204-209. IEEE, 2019.

[33] Liang, Wei, Bo Liao, Jing Long, Yan Jiang, and Li Peng. "Study on PUF based secure protection for IC design." *Microprocessors and Microsystems* 45 (2016): 56-66.

[34] Mahmood, Khalid, Salman Shamshad, Minahil Rana, Akasha Shafiq, Shafiq Ahmad, Muhammad Arslan Akram, and Ruhul Amin. "PUF enable lightweight key-exchange and mutual authentication protocol for multi-server based D2D communication." *Journal of Information Security and Applications* 61 (2021): 102900.

[35] Maes, Roel, Vincent van der Leest, Erik van der Sluis, and Frans Willems. "Secure key generation from biased PUFs: extended version." *Journal of Cryptographic Engineering* 6, no. 2 (2016): 121-137.

[36] Xu, Kun, Weidong Zhang, and Zheng Yan. "A privacy-preserving mobile application recommender system based on trust evaluation." *Journal of computational science* 26 (2018): 87-107.

[37] Posegga, Joachim, and Daniel Schreckling. "Next generation mobile application security." In *IT-Sicherheit zwischen Regulierung und Innovation*, pp. 181-199. Vieweg+ Teubner Verlag, 2011.

[38] Shahriar, Hossain, Md Arabin Talukder, Hongmei Chi, Mohammad Rahman, Sheikh Ahamed, Atef Shalan, and Khaled Tarmissi. "Data protection labware for mobile security." In *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*, pp. 183-195. Springer, Cham, 2019.

[39] Alanda, Aide, Deni Satria, H. A. Mooduto, and Bobby Kurniawan. "Mobile application security penetration testing based on OWASP." In *IOP Conference Series: Materials Science and Engineering*, vol. 846, no. 1, p. 012036. IOP Publishing, 2020.

[40] Sarkar, Arpita, and Binod K. Singh. "A review on performance, security and various biometric template protection schemes for biometric authentication systems." *Multimedia Tools and Applications* 79, no. 37 (2020): 27721-27776.

[41] Mathis, Florian, Hassan Ismail Fawaz, and Mohamed Khamis. "Knowledge-driven biometric authentication in virtual reality." In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-10. 2020.