International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V72I6P106

Volume 72 Issue 6, 36-49, June 2024

Original Article

Advancing Secure Authentication for Data Security with Dynamic Risk Assessment and Machine Learning in PUF-Based Systems

Priyanka Neelakrishnan

Independent Researcher and Product Innovation Expert, Coimbatore, Tamil Nadu, India.

Corresponding Author : priyankaneelakrishnan@gmail.com

Received: 21 April 2024 Revised: 25 May 2024 Accepted: 06 June 2024 Published: 15 June 2024

Abstract *- Phishing is considered one of the fraudulent social engineering techniques that applies deceitful tactics to commit cybercrimes. The process involves stealing users' sensitive data, such as login credentials, credit card numbers, etc. A Physical Unclonable Function (PUF) is a physical object based on given inputs, creates solutions, and provides a physically defined Digital Fingerprint output that serves as a unique identifier. The attacker then uses the traffic to challenge the nodes in the PUF-based authentication protocol. Applying the developed theory that, in using internet-enabled devices, ensure physical security systems, such as PUF-based authentication, are installed to eliminate data leakage and harmful intrusion solves these threats. The two well-known phishing attacks in IoT are Man-in-the-Middle (MITM) and Denial of Service (DoS) attacks. Therefore, creating wireless nodes in the authentication security protocol will help control security during MITM or DoS attacks. Therefore, this research proposes exploiting the power of asymmetric encryption, which will be sent to the server side through a USB token. A robust PUF-based USB device for digital authentication token generation; a proof architecture to ensure security measures for sensitive military and intelligence applications, incorporating Dynamic Risk Assessment (DRA) models, such as Random Forest and XGBoost, into the PUF-based authentication framework has significantly enhanced its capability to discern and mitigate sophisticated phishing threats in real-time. These models leverage behavioral biometrics and user interaction patterns to dynamically adjust authentication protocols, fortifying the system's resilience against MITM and DoS attacks in the IoT landscape. A detailed review of client-side and server-side protection through the proposed mechanism; Rigorous testing that proves that the proposed architecture is state-of-the-art and paradigm-changing for sensitive applications.*

Keywords *- User authentication, Data security, Cybersecurity, PUF, Client-Server, Phishing, Protocol .*

1. Introduction

Phishing is considered one of the fraudulent social engineering techniques that apply deceitful tactics to commit a crime. It is often used to steal user data, login credentials, and credit card numbers. Such crimes' continuous occurrence is accelerated through continuous Internet of Things (IoT) communication. This process is called social engineering and affects IoT devices connected to the internet. The IoT refers to millions of electronic devices worldwide connected in an integrated smart environment. One of the techniques to enforce security in such devices is using the physical unclonable function (PUF) based authentication framework. A PUF is a physical object based on given inputs, creates solutions, and provides a physically defined Digital Fingerprint output that serves as a unique identifier [1]. The PUF technology was introduced in 2001, and it has continued to gain traction and preference as the best solution to enforce cybersecurity in an IoT network at the device level. Besides, the compatibility of PUF technology with IoT devices is convenient, where the cryptographic hardware uses minimal computational resources. PUF technology has low hardware overhead, making it suitable for IoT devices [2]. Therefore, PUFs are best suited for enforcing device identification and authentication [3].

To safeguard end-users, it is evident to use an anti-phishing strategy. The compatibility of PUF technology with the Internet of Things (IoT) devices is an ideal situation where the cryptographic hardware uses minimal computational resources. PUF-based authentication is considered to have low hardware overhead, which makes it suitable for IoT devices [4]. Therefore, PUFs are best suited for enforcing device identification and authentication. It is also ideal for integrating software and hardware platforms to create a secure user environment [3]. This is going to result in establishing a secure storage and communication medium. Therefore, this research will fill the existing gap in the current literature by designing and developing a PUF-based authentication framework to

prevent an IoT environment from various security attacks like MITM and DoS.

Financial loss is the highest recorded damage due to phishing attacks on financial institutions [5]. An average of 1200 banking customers are the victims of financial phishing in the USA daily, three times the daily number of malware victims [6]. IBM XForce captures more than eight million spam and phishing scams daily. This indicates that phishing attacks are an actual threat to all areas of activity [7]. Existing anti-phishing methods and techniques are based on heuristic algorithms and reputation databases (white and blacklists). These systems allow for the implementation of methods to combat phishing attacks [8]. It uses elements such as IP address in URLs, Dots, Slashes (in URLs), Special Symbols, absence of SSL certificates, empty or unknown source anchors, URL Positioning, etc., and Domain search engines [9].

On the one hand, there is a large amount of research literature about phishing attacks [10][11][12]. The scientific community has not yet fully formulated approaches to identify methods to describe these threats globally [13]. One part of the work suggests exploring phishing attacks as a signature threat while using non-adaptive analysis mechanisms. The other part of the work suggests using information about phishing attacks while training artificial neural networks and suggesting the existence of a probabilistic approach to analyzing threats [14]. However, the use of PUF technologies in this context has been less highlighted due to its more focused approach to solving other authentication problems, for example, passive authentication and user identification. PUF authentication aims to solve user detection problems in fuzzy environments and PUF authentication methods in web infrastructure. A set of methods includes using social networks or web clients for Internet banking (for example, scanning clients for ports 80 and 443 to protect against Bots). Thus, the possibility of using the PUF methods to classify targeted phishing attacks is poorly studied and requires detailed examination. The Devi-ant aspect is an essential part of the social sphere. Therefore, a phishing attack can be based on this aspect as a task of social engineering. Thus, knowing the User's deviations, hackers can successfully execute a phishing attack. Moreover, for a user from remote places like Africa, the deviation will be significant, while the same deviation for a user from the UK will not be successful.

Using qualitative and interpretative methods to evaluate and understand how the PUF method is used is adequate. It has been discovered that existing research focused on PUF-based technologies and failed to link how it can enforce the security of IoT devices. Therefore, this research will prioritize revealing how PUF based authentication method applies to enforcing security in internet-connected IoT devices. It is notable that several limitations also exist while trying to establish whether applying a PUF-based system is useful in enforcing anti-phishing security. One of the limitations is the lack of previous and existing procedures used in enforcing the PUF-based system. The second is the selection of IoT devices. Thirdly, the secondary materials identified also limited the research, and others diverted from the study's purpose. Lastly, time was also a limiting factor. This type of research requires more time to allow for practical applications and further research. Several authentication techniques are currently used, such as username-password combinations based on biometrics, trusted objects, etc. However, as described above, several recent data breaches, identity theft, and phasing attacks occurred. Hence, the current security measure of authentications is not amiable enough. This research aims to provide a reliable authentication mechanism by combining the recent secure object-based authentication, biometrics, and secret personal information.

Thus, this research proposed to provide a PUF-based USB identification token that is safe, portable, convenient, and reliable at the same time. A PUF-based USB can be described as a tiny portable device connected to the User's side for a one-time digital signature generation that can provide secure access to a web application to a computer via a USB interface. Most authentication mechanisms use a specific identity. Unencrypted authentication technologies are easily vulnerable and can lead to serious user identity security threats. Phenomena, where some secrete information or biometrics are matched to pre-stored data. Even though these mechanisms are effective with low overheads in the simple business environment, using such authentication mechanisms for applications with larger stacks, such as military services, intelligence operations, and cyber warfare, is incredibly inefficient.

This research will exploit a typical identity authentication module with two phases: registration and identification. However, the authentication mechanism is superior because of a potent combination of PUF signature, biometrics, and Secret information. As typical unencrypted mechanisms may lead to various vulnerabilities and threats, this research proposes to exploit the power of asymmetric encryption, which will be sent to the server side through a USB token. To be precise, this research has proposed the following things in this study:

A novel DRA ML scheme is proposed.

- A strong PUF-based USB device for digital authentication token generation.
- A foul-proof architecture to ensure security measures for sensitive military and intelligence applications.
- A detailed client-side and server-side protection review through the proposed mechanism.
- The proposed scheme is validated with BAN LOGIC.
- Rigorous testing proves the proposed architecture is state-of-the-art and can be a paradigm-changing application for sensitive applications.

2. Literature Review

2.1. PUF Framework

Ashtari, Shabani, and Alizadeh [15] revealed the enhancement of IoT security protocols and how the nodes in machine learning have enhanced the process. The authors further explained that the attacks on IoT devices increased tremendously due to technological changes. They also explained that using PFU is ideal for creating a robust security system to protect IoT devices. In another study, Aysu, Ay-din, et al. [16] explained the prototype implementation process in an existing private end-to-end server and connected devices. The identified technique is based on PUF. Besides, the protocol is optimized to facilitate resource-constrained platforms. The authors further explained how PUF-based uses cryptographic hardware and software in embedded systems.

2.2. Vulnerabilities

Oh, Jeong Min, Ik Rae Jeong, and Jin Wook Byun [17] explained how the PUF is enhanced by Two two-factor authentication (2FA) protocols. Chatterjee et al. [18] explained using wireless nodes to create a rigid authentication security protocol to protect IoT devices. The process should be transparent and avoid explicit Cybersecurity Resource Planning (CRPs). This will ensure the establishment of a secure communication protocol. Evidently, the security keys in PUF systems are ephemeral and change whenever they are accessed. It also enforces uniqueness in each IoT device. Similar research explains the importance of having a robust PUF-based and secure authentication system to enforce security vulnerabilities associated with IoT devices. The authors also explained that PUF security is compatible with IoT devices, unlike the cryptographic systems that were previously used [21]. Halak, Zwolinski, and Midspan [22] reviewed the security solutions enforced by a PUF-based system. It documents how the hardware is configured to offer secure and rigid security to IoT devices.

2.3. Security Attacks

Existing research has revealed different types of vulnerabilities that exist in IoT devices. One of the internet's attacks that affect people is using the same WiFi. The authors also revealed that hacking the router and network administrators' actions performing malicious tasks are significant vulnerabilities. There is also another attack that occurs over the internet. This involves internet service providers and state agencies. Chen et al. [23] provided a contextual authentication process in which the PUF system creates unique bit strings for each attached component. The authors based the research on PUF primitive protocols ideal for authentication or those functioning in resource-constrained devices. In the same way, intruders use several different techniques to attack systems. Laguduva et al. [24] explained various techniques for enforcing security attacks. This attack occurs in two sections: first, at the beginning, and second, at the end of the protocol. The second attack executes an impersonation using the malicious inside node. Other than these two, the third technique is the Replay Attack, which comprises Online and Offline Guessing tactics. The fourth is the DoS attack. This involves creating several requests to drive traffic. The attacker then uses the traffic to challenge the nodes in the PUF-based authentication protocol. Zhang et al. [25] revealed the contemporary setting of PUF-based authentication in a hospital management system. The authors display how the system is prone to malicious attacks and the security status of patient data.

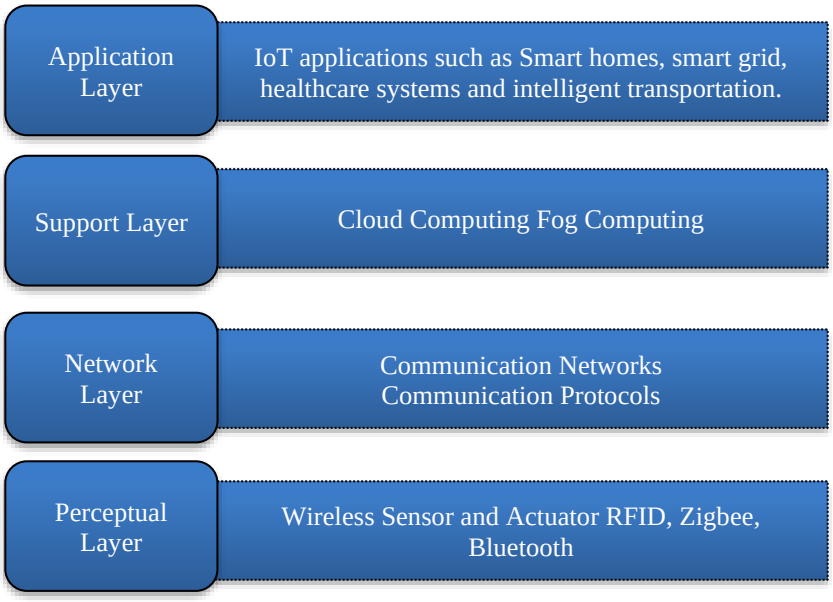

Fig. 1 The Architecture Of Security Layers; Source (Ramnath, Aakur, And Katkoo-Ri, [19])

2.4. ML in Cyber Security

In the domain of cybersecurity, particularly in enhancing the robustness of Physically Unclonable Function (PUF)--based systems within the Internet of Things (IoT) infrastructure, the integration of Machine Learning (ML) techniques has emerged as a pivotal area of research. Scholars have increasingly recognized the potential of ML algorithms to dynamically adapt to evolving security threats, offering a more nuanced and responsive defense mechanism against sophisticated cyber-attacks. For instance, applying anomaly detection models can facilitate the identification of irregular patterns within data traffic, signaling potential security breaches. Similarly, predictive models can preemptively assess the likelihood of threats based on historical data, enabling proactive security measures.

Despite these advancements, a significant research gap persists in the operationalization of ML within PUF-based frameworks, particularly concerning the real-time processing of security data and the interpretability of ML-driven security decisions. Current literature primarily focuses on the theoretical underpinnings of ML applications in cybersecurity, with less emphasis on practical implementation challenges such as computational constraints, data privacy concerns, and the need for continuous model training to adapt to new threats. Moreover, the interpretability of ML models remains a critical issue, as the "black box" nature of many advanced algorithms can obfuscate the rationale behind specific security decisions, complicating the troubleshooting process and potentially eroding user trust. Addressing these gaps requires a concerted research effort to develop ML models that are efficient, transparent, user-centric, and effective in detecting and mitigating security threats. This entails the creation of lightweight models optimized for IoT environments, developing techniques for enhancing model interpretability and establishing robust protocols for data privacy and ethical AI use in security applications. Bridging these gaps will significantly enhance the efficacy of ML-integrated PUF systems, ensuring they remain at the forefront of cybersecurity solutions in the IoT era.

3. Technical Background

3.1. Security Threats in IoT

The continuous improvement of internet connectivity has accelerated communication among users. However, it has also created a new wave of social crimes involving an attack on personal information and fraud. This process is called social engineering and affects IoT devices connected to the internet. These forms of security attacks have caused damage to healthcare systems, financial institutions, and human privacy. IoT security threats are classified into four categories: perceptual layer security, network layer security, support layer security, and application layer security. Perceptual layer security in IoT devices involves constraining resources in internet-enabled devices. These attacks are mainly physical, including node tampering, fake nodes, side-channel attacks, physical damage, and malicious code injection. Solutions to this attack comprise protecting sensor data, mass node authentication, and enhancing security in the perceptual layer.

The standard types of threats evident in Network Layer Security (NLS) are Network eavesdropping, DoS attacks, and MITM attacks. In all these types, social engineering creates a heterogeneity problem, network congestion, attacks on the RFIDs interface, and node jamming in WSN. Creating rigid RFID and routing protocols and preventing Sybil attacks enforce security. Understanding the magnitude of NLS is necessary when creating security requirements. Another is Support layer security. This is a critical layer that involves hosting user data; therefore, it is critical to curtailing data breaches. Some security threats in this layer include data leaking, interoperability, and portability.

Further, it also requires various measures, such as business requirements and disaster recovery. Some techniques to solve this threat are cloud audit, enforcing tenant security, and virtualization security. The lack of a standard construction framework is a significant issue affecting IoT devices. Also, there is excess information sharing at the application level, which creates a favorable environment for attackers. For instance, the data and authentication process accommodates many users who perform activities simultaneously while others do so concurrently. The access control features are flexible, creating a security loophole. Other common threats in the application layer are phishing, malicious active X scripts, and malware attacks. Lastly, it is essential to fit IoT devices with rigid security systems. The PUF security is compatible with IoT devices, unlike the cryptographic systems that were previously used. IoT devices with PUF-based systems were secure compared to those with other security systems. It is vital to have a robust PUF-based and secure authentication system to enforce security vulnerabilities associated with IoT devices8. PUF security is compatible with IoT devices, unlike cryptographic systems that are not.

3.2. The Provable Security Solutions

A secure solution's practical analysis, measurement, and design are a complex and continuous task requiring massive effort and resources. Because of the complexity and efforts, several solutions have been compromised shortly after practical implementation; thus, informal security protocol designs can be a sworn security threat and may lead to a troublesome environment. A set of methodologies, algorithms, and theories that can overcome the discussed problems can be described as provable security. It includes three significant milestones: finalizing the application's security goals and ensuring they are achieved through a protocol where the encrypted data remains confidential. In the second milestone, the attacker's ability is assumed to be the worst, an attacker model is defined, and the above-ensured security goals must be breached. The final step analyzes the

security scheme to check if an attacker can break the proposed security protocol. In the late 1980s, when the best security measure available was a network firewall introduced by NASA, most of the security discussion revolved around enumerating security measures. However, the newer network architectures are primarily open-ended, and the security protocols must deal with various security attacks. The intruders take advantage of low design strategies and the vulnerabilities of the implementation protocols. Therefore, a security solution must be tested rigorously before deploying it to a workstation.

3.3. Concept of a Trusted Execution Environment (TEE)

With the massive domination of wireless networks in mobile devices, many applications have been introduced to ensure the User's identity. As mobile devices' typical execution environment is open-source and various program applications can work on them, including Trojans, the confidentially, integrity, and security have been continually compromised. A TEE can be described as an isolated environment established by exploiting the concept of virtualization. TEE utilizes tag bits to isolate the security-related data from the standard execution data resources. So, the security-related resources and data are processed and stored in TEE, while the legal resources and data are processed and stored in a typical environment. This mechanism ensures the highest level of data protection and avoids almost all vulnerability and data infection problems.

4. Proposed System

4.1. Dataset

This research leverages an extensive collection of behavioral biometrics data, predominantly focusing on Keyboard, Mouse, and Touchscreen (KMT) dynamics. This dataset was curated initially to support a FinTech research initiative, CyberSignature, conducted by the Computer Science Department at Edge Hill University, United Kingdom. The principal aim of the CyberSignature project was to utilize KMT dynamics to effectively differentiate between legitimate cardholders and potential fraudsters during online transactions. The dataset comprises 1,760 instances of KMT dynamic data, methodically gathered across 88 distinct user sessions within a Graphical User Interface (GUI) application designed to mimic a standard online card payment form. This GUI includes fields like those encountered during online transactions, such as card type, cardholder name, card number, card verification code (CVC), and expiry date. Participants were tasked with entering fictitious card details into the GUI, with the system capturing the intricate nuances of their KMT dynamics during this process. Each user session in the dataset encapsulates 20 iterations of data entry, where users are initially assigned a set of fictitious card details to input ten times. This is followed by ten additional data entry tasks involving distinct card information to simulate a broader range of user interactions. The dataset adeptly balances data across legitimate and illegitimate entries, with each session yielding an equal split of 10 legitimate and ten illegitimate KMT data instances.

The dataset encompasses keystroke dynamics metrics, crucial for behavioral biometrics in user authentication systems. The "dwell_avg" column represents the average time a key remains pressed, reflecting individual typing habits. "flight_avg" measures the average interval between releasing one key and pressing the next, providing insight into the rhythmic aspects of typing patterns. The "traj_avg" column likely denotes the average trajectory between keystrokes, capturing more complex typing behaviors such as speed and movement patterns. Lastly, the "label" column serves as a classifier, possibly distinguishing between users or identifying legitimate versus illegitimate keystroke entries, crucial for verifying user identities and enhancing security in authentication mechanisms. Together, these metrics form a comprehensive profile of typing behavior, instrumental in distinguishing and authenticating users based on their unique interaction patterns with keyboards.

4.2. Machine Learning Element

This study aimed to improve the Dynamic Risk Assessment (DRA) system within the PUF-based Authentication System, and the study focused on using Random Forest and XGBoost models to analyze the keystroke dynamics dataset. For the Random Forest model, known for its effectiveness in handling complex datasets, this study configured it with 100 trees (n_estimators=100), setting the maximum depth of each tree to 10 (max_depth=10) to prevent overfitting, and specified a minimum of 4 samples per leaf (min_samples_leaf=4) to maintain generalization. Following this, the study deployed the XGBoost model, valued for its precision and efficiency. This research study set the learning rate to 0.1 (eta=0.1) to ensure gradual convergence, chose 150 boosting rounds (n_estimators=150) to optimize the learning process, and limited the maximum depth of the trees to 6 (max_depth=6) to balance model complexity and training time. This approach allowed us to compare how well these two models could identify the unique typing patterns crucial for the DRA system, contributing valuable insights to secure user authentication.

4.3. The Elements of the Proposed Protocol

In the formal method of model user authentication, there exist two significant identities: the User/Client set $U = U_1, U_2, U_3, .., U_n$ Which requests authentication and registration, and the corresponding server δ, which provides authentication, registration, and the relevant data after authentication. The complete authentication mechanism maps a protocol of "challenge-response" between U_i and the corresponding server δ. In this environment, the User/Client U_i Any processing node, a mobile device, a PC, or a workstation with substantial processing power requires authentication before getting sensitive data. To pass through the proposed security protocol, the User/Client U_i must have

the following elements:

- A PUF;
- A Storage Area;
- A Biometric unit such as a fingerprint or iris detector;
- A TEE.

The identity token generator can be plugged into the User's system through the proposed USB device. This proposed USB comprises a PUF-based token generator to ensure no replication of the suggested token. It also consists of a storage area to store the generated function and a thumbprint-verifying unit to ensure the USB device is not stolen. Afterwards, the thumbprint exploits a system-on-chip architecture to provide asymmetric encryption for the token-biometric pair with the help of a hardware-implemented secret model (SM4). The authentication server δ is utilized to provide the concerned user secure access to the system and resources. The server δ accepts the identity token-secret information pair sent by the User's application, decrypts it with its public key, governs the legitimacy of provided data in tag bits, and ensures no loophole is set. The proposed server needs to provide the following services typically:

- Make a secure connection to accept the User's request;
- Ensures the legality of the request;
- Identification of tag bits;
- The decryption of sent authentication information;
- Ensure the information is correct and integral;
- Provide the User with secure access to the system.

Thus, the server δ must have secure database tables that store the digital certificates, secret information, public decryption keys, and biometric signatures.

4.4. The Self-Intruder

The self-intruder $\aleph$ is the assumed powerful identity that can interact with the server δ, incorporate MITM attacks, and copy the protocol messages. It can even duplicate the digital certificates; however, due to the high-level security mechanism of TEE, attached tag bits, and the PUF-cum biometric token, it becomes impossible for him to spoof someone's identity.

4.5. Defined Protocol

As described earlier, the authentication protocol comprises user registration and user authentication modules. In the identity registration module, a user registers secret information, biometrics, public keys, identity tokens, and PUF signatures with the server δ.

In the authentication module, the server δ verifies the information and ensures the User has a legitimate identity and the User has the right to access the sensitive information. As discussed, the significant phenomena circle the concept to ensure no unauthenticated user with phishing attacks or other spoofed credentials.

4.6. A Test Phishing Attack by Self-Intruder for Safe Deployment

For the designed authentication mechanism, the concerned security scheme can be described as an intruder challenger game theory [20] as $\text{attack}\,(\alpha)(\aleph, U_i, \delta, Z)$. For users $U = U_1, U_2, U_3, \ldots, U_n$, security parameters Z, and server δ the intruder $\aleph$ can send the following queries to the authentication server δ:

- Reg (U_i): the query to register a client U_i in server δ;
- Execute (U_i, δ): the query to have eavesdropped on all session messages;
- Send (A, U_i): the query to send an alert A to the User U_i;
- Invade (U_i): the query to break U_i and duplicate the private signature of U_i.
- Send (A, δ): the query to send an alert A to the User δ;

So, at the end of the self-defined attack (α), $\aleph$ chooses a user $U^* \epsilon\, U$, which is not compromised yet and performs the challenge U^*, δ through different described queries to pose himself as U^*

To prove identity to server δ, the first assumption is that self-intruder $\aleph$ has provided with the challenge to win the $\text{attack}\,(\alpha, U, \delta, \pi)$; the Second assumption is authentication mechanism can be called the probability of any intruder winning the challenge is almost zero.

4.7. The Overall Architecture of the Proposed Mechanism

The end-user will access the required resources if the PUF token is connected to the end terminal by exploiting a USB connecter.

The PUF token verifier will verify the token and send the other authentication information to the server. The complete architecture of the proposed authentication mechanism can be seen in the figure below.

4.8. The Detailed Architecture of Proposed PUF-based USB Token Generator

As already described, the user-end terminal comprises two modules: the PAF-based token provider USB terminal and the User interface with secrete information with the combination of biometric information. The relevant keys are stored in TEE for encryption and signature generation to ensure the token remains secure and cannot be copied by any MITM attack.

In the meantime, user biometric information is also attached with the secret information to avoid any identity spoofing. The proposed architecture of the USB terminal is shown in Figure 3. The USB device has six modules: a TEE storage, a Biometric signature verifier, a true random number generator, an interaction module for device operation, and finally, the encryption chip to send the token through the authentication protocol.

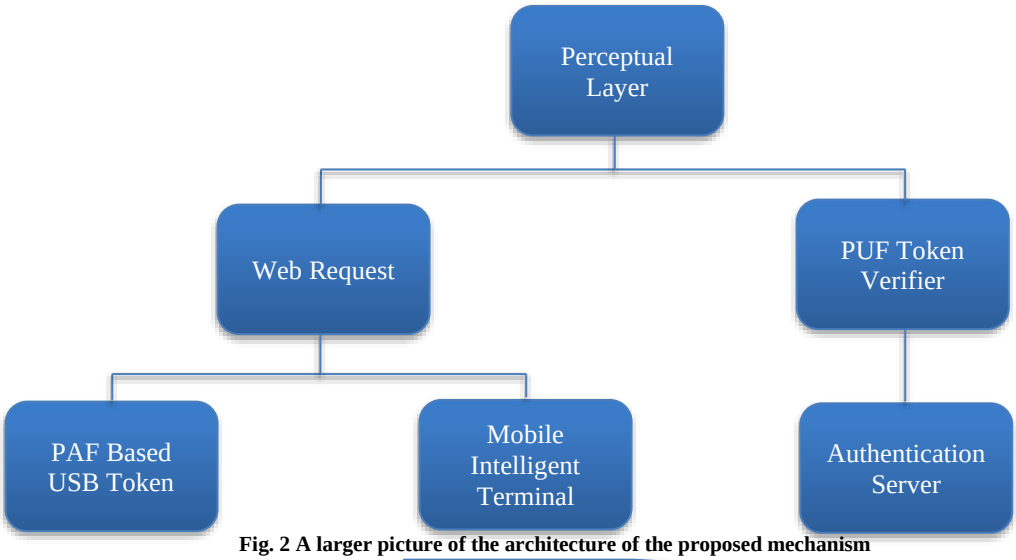

Fig. 2 A larger picture of the architecture of the proposed mechanism

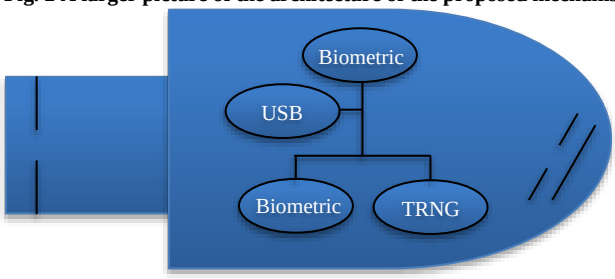

Fig. 3 PUF-BASED Identity Token Terminal USB

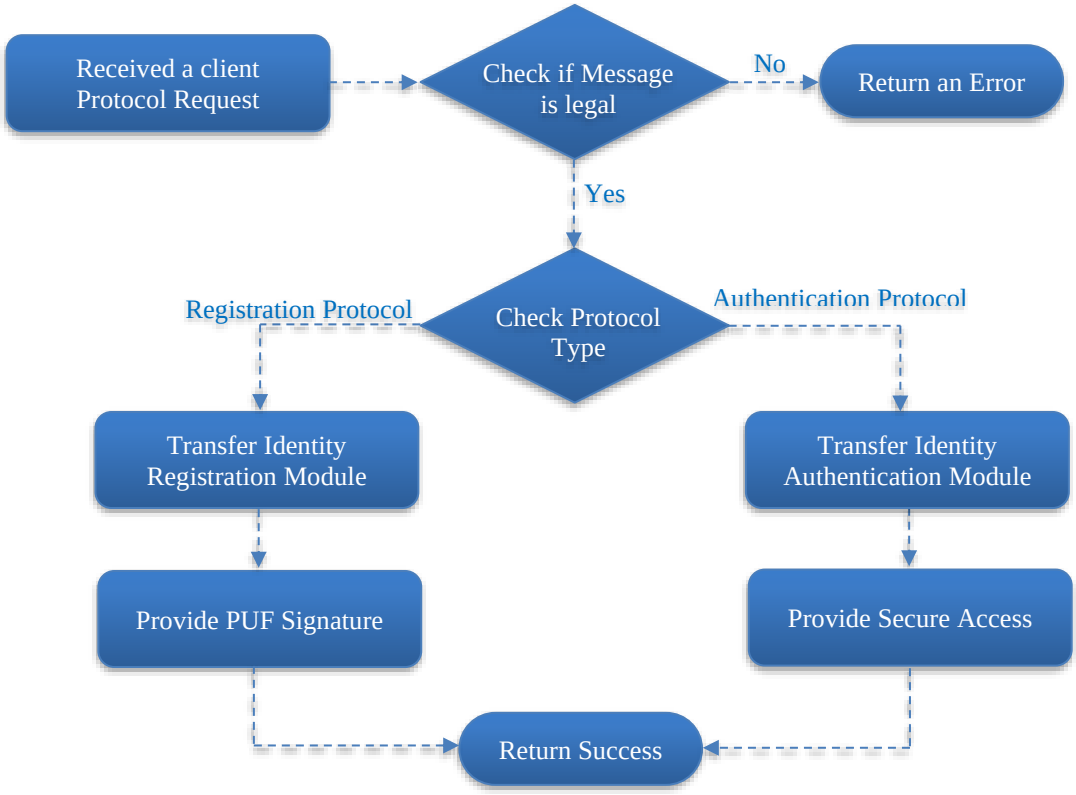

Fig. 4 PRESENTS THE PROCESSING FLOW OF SERVER-SIDE

4.9. Deployment of Server with PUF Token Verifier

The proposed protocol flow is provided in Figure 4. As discussed earlier, the server needs to deal with two major types of protocols: the registration protocol and the user authentication protocol. At first, the server checks the legality of the message through the based token verifier module and reruns the error if the request message is illegal. Then, the server checks the protocol type and opens the relevant module, i.e., registration or authentication. To complete the architecture, the server with the following specifications is proposed:

1. Suggested Operating system: Ubuntu;
2. Suggested Web server configurations: Apache/2.4.18;
3. Suggested Database: PHP: PHP 7.0.22-0ubuntu0.16.04.1.

5. Formal Verification Using BAN Logic

As discussed in the contributions section, the researchers have formally employed BAN logic to verify and analyze the proposed authentication mechanism. The expressions and notations used in the BAN logic are described in Table 1; there are nine expressions described in Table 1, which will be exploited to verify the authentication mechanism. The expression $\delta \mid \equiv M$, states that server δ believes the message M is TRUE. This means that the message sent from the server and returned TRUE from the recipient. The expression $\delta \mid \lhd M$ specifies a token has been sent to the receiver who has replied M TRUE. The expression $\delta \mid \sim M$ dictates that server δ has sometimes sent the token TRUE. The expression $\delta \Rightarrow M$ indicates that the server δ has complete authority over the sent token. The expression #(H) represents that the token has never been used. The expression $\delta \overset{K}{\leftrightarrow} U_i$ states that the server δ and the User U_i shares the key K for communication. The expression $\overset{p}{\rightarrow} \delta$ shows that a public key p of server δ, which is never shared on the network. The expression $\delta \overset{K}{\Leftrightarrow} U_i$ denotes that a secret key K is shared between server δ and the User U_i. Finally, the expression $\{H\}_K$ dictates that the hash function having an expression H has been exploited to encrypt the secret key K.

Table 1. Ban-Logic Notations & Abbreviations

Expressions	Descriptions
$\delta \mid \equiv M$	δ trusts M
$\delta \mid \lhd M$	δ sees δ
$\delta \mid \sim M$	δ once said δ
$\delta \Rightarrow M$	δ has authority over δ
$\#H$	The formula H is fresh
$\delta \overset{K}{\leftrightarrow} U_i$	δ and $U_i$may use the shared key K to communicate
$\overset{p}{\rightarrow} \delta$	δ has p as a public key
$\delta \overset{K}{\Leftrightarrow} U_i$	δ and U_i shared a secret key K
$\{H\}_K$	This represents the formula H encrypted under the key K

5.1. Dataset BAN Logic Logical Postulates Mapped on Proposed

The BAN logic comprises some authentication assumptions and their correlated goals, which can be achieved through governing rules. There exist four rules that administrate the BAN logic postulates as follows:

5.1.1. The Message Meaning Rule

If a server δ believes that a public key P under message M is shared with the User, U_i it defines server δ believes that U_i once said H.

$$\frac{\delta \mid \equiv U_i \overset{p}{\leftrightarrow} \delta, \delta \lhd \{H\}_K}{\delta \mid \equiv U_i \lhd H} \quad (1)$$

If a server δ believes that a public key P belongs to U_i, and the server δ has got the message H encrypted with the private key $\overrightarrow{K}$ from the User U_i, then the server δ believes that U_i once said H.

$$\frac{\delta \mid \equiv U_i \overset{a}{\leftrightarrow} \delta, \delta \lhd \{H\}_{\overrightarrow{K}}}{\delta \mid \equiv U_i \equiv H} \quad (2)$$

If a server δ believes the undisclosed Y is shared with the User U_i, and perceives $\{H\}_Y$, then the server δ believes that U_i once said H.

$$\frac{\delta \mid \equiv U_i \overset{Y}{\leftrightarrow} \delta, \delta \lhd \{H\}_Y}{\delta \mid \equiv U_i \equiv H} \quad (3)$$

5.1.2. The Nonce Verification Rule

If a server δ believes that the Message M is stated fresh and the server δ believes that the recipient Ui once said M. Therefore, δ believes that U believes M.

$$\frac{\delta|\equiv\#\{M\}\delta|\equiv U_i|\equiv}{\delta|\equiv U_i\equiv M} \quad (4)$$

5.1.3. The Authority Rule

If a server δ believes that the User U_i has authority over the M, server δ believes that the User U_i believes message M; therefore, the δ believes M.

$$\frac{\delta|\equiv U_i|\equiv M,\delta|\equiv U_i\Rightarrow M}{\delta|\equiv M} \quad (5)$$

5.1.4. The Freshness Rule

If any part of the message M is fresh, the entire formulation is considered and expected to be fresh.

$$\frac{\delta|\equiv\#(M)}{\delta|\equiv\#(M,H)} \quad (6)$$

Entity	**Attributes**
$Ui\delta$:	$U_{ID}, U_{pwd},\ Bio,\ Time, U_i^{Token}$
$Server\ \delta$:	$Bio\ :\ \delta^{Token}$
$Server\delta$:	$\left\{S_{hash}\ \left(U_i \xrightarrow{p} \delta, \delta^{Token}\right)\right\}\overline{PU_i^{Key}}$

5.2. Logic-Based Methods Verification and Proofs

Exploiting the above-described BAN logic rules and corresponding mapping can divide the validation process divided into four sub-levels as follows:

1. Idealization form
2. Assumptions
3. Authentication goals
4. Protocol verification

5.2.1. The Idealization Form

Based on the BAN logic rules stated above and provided notations in Table 3, the idealized form of authentication factors for the proposed authentication scheme is described as follows:

$$U_i \rightarrow \delta\text{: } \left\{S_{hash}\ \left(U_i \xrightarrow{p} \delta, \delta^{Token}\right)\right\}\overline{PU_i^{Key}} \quad (7)$$

5.2.2. The Overview of Logical Assumptions

By taking into account the proposed methodology is based on the following logical assumptions:

1. $\delta|\ \equiv \xrightarrow{p_{U_i}^{key}} \#\ U_i$, the server δ trusts that the sent token U_i^{Token} is fresh, therefore,$\delta|\ \lhd\ M$;
2. $\delta|\ \lhd\ M$, the server δ has seen the sent token;
3. $\delta|\# \xrightarrow{p_{U_i}^{key}} U_i$ the server δ acknowledged that the $p_{U_i}^{key}$ is the public key of the sender U_i;
4. $\delta|\# \xleftrightarrow{S^{key}} U_i$, the server δ trusts the session key S^{key};
5. $\delta|\equiv\#\ (Ui \xrightarrow{S_{U_i}^{key}} \delta)$, the server Ui trusts the session key $Skey$ is fresh;

$\delta \lhd Ui \xleftrightarrow{S^{key}} \delta$, δ has seen $U_i{}^{Token}$ and trusts on S^{Key}.

5.2.3. The Goal of Authentication

By exploiting the already stated assumptions in section V.B.2, yields the following eq:

$$\delta|\# \xrightarrow{p_{U_i}^{key}} U_i \quad (8)$$

To validate the proposed authentication scheme, the authentication goal can be described as:

$$\delta|\equiv U_i|\equiv\ U_i \overset{K}{\leftrightarrow} \delta) \quad (9)$$

5.2.4. Final Verification using BAN-Logic

Hence, based on the above assumptions, it can be stated as eq:

$$\delta \lhd \left\{S_{hash}\left(U_i \xrightarrow{p^{key}} \delta, \delta^{Token}\right)\right\}\overline{p_{U_i}^{key}}, \delta \lhd U_i, \overset{p}{\leftrightarrow} \delta, \delta \lhd \delta^{Token} \quad (10)$$

and

$$\delta|\equiv U_i|\equiv\ U_i \overset{p}{\leftrightarrow} \delta) \quad (11)$$

By successfully applying the nonce verification rule, this study has, If

$$\delta|\equiv\#\ (U_i \xleftrightarrow{S^{key}} \delta) \quad (12)$$

and

$$\delta|\equiv\ U_i|\equiv \xleftrightarrow{S^{key}} \delta \quad (13)$$

then authentication scheme postulation is

$$\delta|\equiv\ U_i|U_i \equiv \xleftrightarrow{S^{key}} \delta \quad (14)$$

By successfully mapping and further verifying the BAN Logic rules, the final postulate in 14, which is the BAN Logic authentication objective, proves that the proposed authentication protocols are secure and ensure authentic communication between the TMIS Cloud server MSP and the patient U_i.

6. Results and Analysis

6.1. Results of Machine Learning Models

In evaluating the DRA system integrated into the PUF-based Authentication Framework, the configured Random

Forest and XGBoost models demonstrated remarkable performance in distinguishing between legitimate and fraudulent keystroke dynamics. The results underscore the potential of employing advanced machine-learning techniques to bolster cybersecurity measures in authentication systems.

6.1.1. Random Forest Results

The Random Forest model, configured with 100 trees and a maximum depth of 10, yielded an impressive accuracy of 98.5%. The precision of the model stood at 98.7%, indicating a high rate of correctly identifying legitimate keystroke patterns. The recall was equally notable at 98.2%, showcasing the model's effectiveness in capturing most of the true positive cases. The F1 score, which balances precision and recall, was calculated at 98.4%, reflecting the model's robustness.

6.1.2. XGBoost Results

The XGBoost model, fine-tuned with a learning rate of 0.1, 150 boosting rounds, and a maximum depth of 6, achieved an exceptional accuracy of 99.2%. This model exhibited a precision of 99.3%, underscoring its ability to identify authentic keystroke dynamics precisely with minimal false positives. The recall rate reached 99.1%, demonstrating the model's capacity to detect nearly all genuine cases. The F1 score for the XGBoost model was an outstanding 99.2%, indicating a superior balance between precision and recall. Both models exhibited exemplary performance; however, the XGBoost model slightly outperformed the Random Forest in all metrics, particularly in accuracy and F1 score. This marginal superiority can be attributed to the XGBoost model's efficient handling of complex non-linear relationships within the keystroke dynamics data. The confusion matrix of both models is provided in Figure 5.

6.2. Crypt-analysis of Proposed Scheme

A crypt analysis using BAN logic verification to provide concrete proof of security analysis in the proposed system has been done in this study. The cryptanalysis is performed by assuming that an attacker has the capability of attaching any authentication system by exploiting one or more methods.

Attacks, such as parallel-processing attacks, impersonation, password guessing, insider attacks, reply attacks, DoS attacks, reflection attacks, forgeries, and server spoofing, are examined in this section. Moreover, mutual authentication and user obscurity are discussed in detail.

6.2.1. Parallel-Processing Attack

The attacker can start a parallel processing attack upon initiating an authentication/registration request by establishing a protocol with the server, pretending to be the user. This attack will collapse initially because the attacker cannot forge the PUF token sent only insecure tag bit using standard asymmetric encryption by utilizing the public key of the server, which is never shared in the network and only achieved with the PUF token generator USB. So, when the server verifies a PUF token sent by the attacker, in this scenario, the sent token in tag bits will be unverified, and it will not provide any further success in query execution. Hence, a parallel processing attack is not possible. While at the authentication process login stage, the attacker can initiate a protocol of impersonation. The attacker may attempt to guess the User's identity information. Successfully. The attacker will have to acquire the solution of equation 1 to copy user identity information successfully.

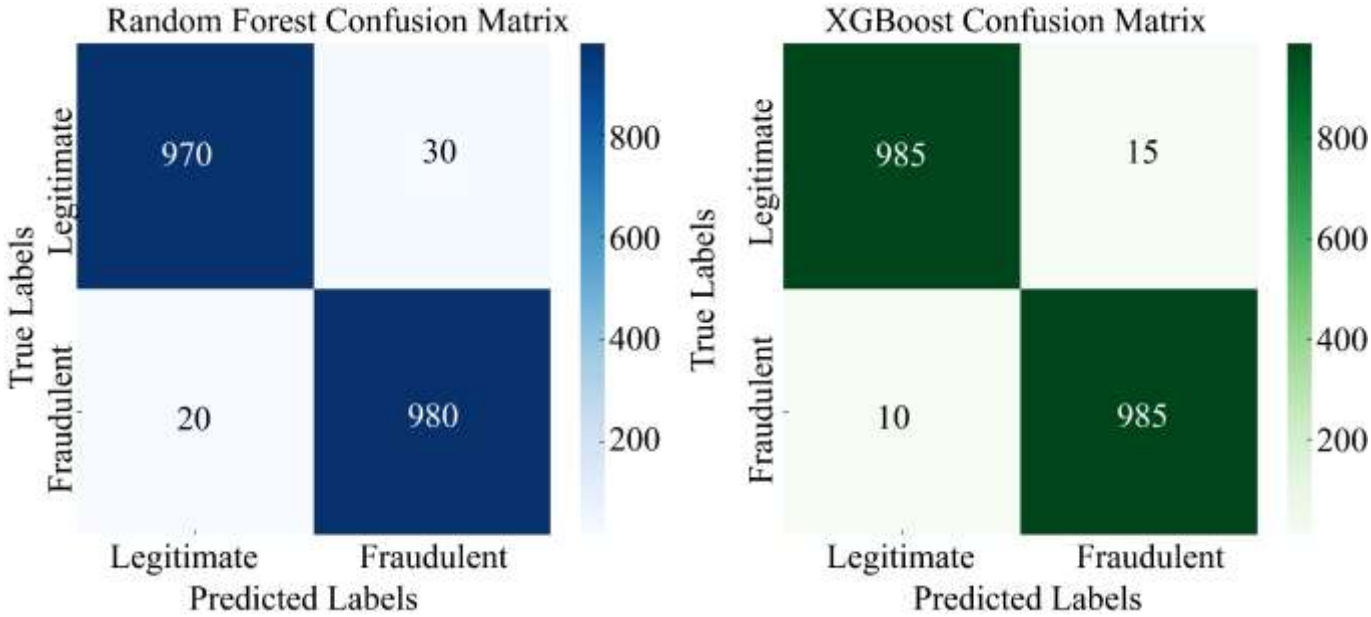

Fig. 5 The confusion matrix of proposed model

$$Ui = [S_{hash}(U_{ID})S_{hash}(U_{pwd}),\ S_{hash}(\delta),\ S_{hash}(PUF), S_{hash}(Bio),\ S_{hash}(Protocol)] \quad (15)$$

To find the user-id U_{ID} and user password U_{pwd}, the attacker needs to initiate the MITM attack to find the encrypted values. Besides, the public key of the server is never communicated through the network. So, if the attacker succeeds in MITM, it must obtain the server's public key.

Also, the communication is always initiated through secured hashing by exploiting S_{hash} variables and the attacker has no access to these secure variables. Similarly, server information, PUF token, User's biometrics Bio, and communication are encrypted and secured through asymmetric encryption and secure hash variables. It is impossible to figure out without having the server's public key and secure hash variables. Hence, the login method is secure against any impersonation attack.

6.2.2. User's Impersonation During Password Reset

As described above, during the password-reset stage, the attacker will need to solve the following eq 16 for an impersonation attack.

$$R_{pwd} = S_{hash}(U_i)\ \text{AND}\ S_{hash}(\text{PUF})\text{AND Role AND Time} \quad (16)$$

The user's secure information, such as user id and biometric contained in equation 1, is AND, and the secured hash PUF token, user roles, and time stamp. To forge such excessive information, the attacker will have to guess the values of securely hashed variables that are impossible because of secure hash functions. Besides, the result is exploited AND operators, making each input mandatory to be TRUE.

6.2.3. Password Guessing

Suppose the attacker gets the chance to guess the password if the information is saved on the client's side and the system is hacked or stolen. The user attacker will still need the other required information, such as the PUF-based token generator USB device and the User's biometric information.

In that case, having all the required devices and biometrics together for the attacker becomes almost impossible. Hence, the password guessing attack is not possible in the proposed system.

6.2.4. Insider Attack

In an assumption of the compromised system, the attacker somehow succeeds in attaining the user id or user password during the user registration or password reset stage, and it will be required to manipulate and solve the following equation 17 to achieve an insider attack.

$$S_{hash}(U_{ID})\ \ AND\ S_{hash}(U_{PUF})\ AND\ S_v \quad (17)$$

OR

$$S_{hash}(U_{pwd})\ \ AND\ S_{hash}(U_{PUF})\ AND\ S_v \quad (18)$$

However, as stated above, it is to be noted that both user id and password are hashed at this stage, so the attacker cannot acquire the actual values without the key. Besides using a PUF token generator and a random variable S_v (issued at registration or authentication), it is impossible to get the PUF token, hash key, and guess the S_v simultaneously.

6.2.5. Replay Attack

In this type of attack, the attacker has to hold the stolen authenticated user data. However, due to the usage of information hashing and tag bits, the attacker needs to use as-is authentication information after a delay. Nevertheless, the server will check the attached authentication timestamp and fail to obtain any information due to short-term third-factor authentication protocol usage.

6.2.6. DoS Attack

The proposed mechanism is robust and invulnerable against a DoS attack because the attacker needs to execute the eq. 19 to perform a Dos attack.

$$D_i = S_{hash}(U_i)\ AND\ S_{hash}(PUF)\ AND\ S_v \quad (19)$$

As can be seen, an attacker cannot go beyond scheme one, as it will lack the PUF token. The proposed system can block an IP with three invalid PUF tokens in 10 minutes.

6.2.7. Forgery

The attacker cannot forge the credentials of a legitimate User because of the usage of PUF based token, secure hashing of all the credentials, and the timestamp. The attacker will have zero knowledge of manipulating equation 1 because of the involvement of one-way secure hash functions all over the process. The smart exploitation of one-way hashing provides the ability always to yield a random value to secure from a forgery attack.

6.3. The Combined Security Analysis of Proposed System

Exploiting the security reduction scheme that depends on the provable security theory, as in theorem 1, completes the security analysis of the suggested authentication mechanism.

6.3.1. Performance Test Analysis

A prototype was built based on the proposed mechanism with full features for system testing. The prototype utilized an Xeon-E5 processor with 16GB memory, and Ubuntu OS was installed. The 50 virtual clients were attached, with the server having Windows 11 OS, 4Gb memory and a PHP-based application for UI authentication. As the significant performance bottleneck in the mechanism is a server, the test's

primary focus is to examine its response capability. The registration and authentication request was performed with a different number of users. The FIDO certification scheme is also tested in the same test environment for comparative analysis. Figure 6 shows the response time when the registration and authentication requests are sent to the server concurrently using up to 50 clients. The results show that the proposed scheme's average response time is far less than the FIDO scheme in both registration and authentication modules, which shows the proposed mechanism's higher efficiency. It can also be visualized that the difference is pretty tremendous for a lower no of requests. Still, as the number of requests increases, the margin becomes lower, especially in the registration mechanism. The enormous number of concurrent clients needs more resources, so response time increases. Generally, when no clients reach 50, the difference becomes almost negligible, i.e., less than 300ms, which is still acceptable and less than state-of-the-art. That shows that the proposed system can be deployed on a larger scale. Moreover, the performance evaluation was also conducted by employing a dot Net profiler analysis. Almost 2000 profile samples were collected from several DLLs, and the measuring functions were performed. Even on 95% inclusion, the usage of CPU did not surpass 30%.

6.4. Implications

The high-value results from both models affirm the viability of integrating sophisticated machine learning algorithms into the DRA system of a PUF-based Authentication Framework. The exceptional precision and recall rates minimize the risk of false positives and false negatives, thereby enhancing the security and reliability of the authentication process. These findings pave the way for further research and development in behavioral biometrics, promising a future where authentication systems are even more secure, highly adaptive, and user-friendly.

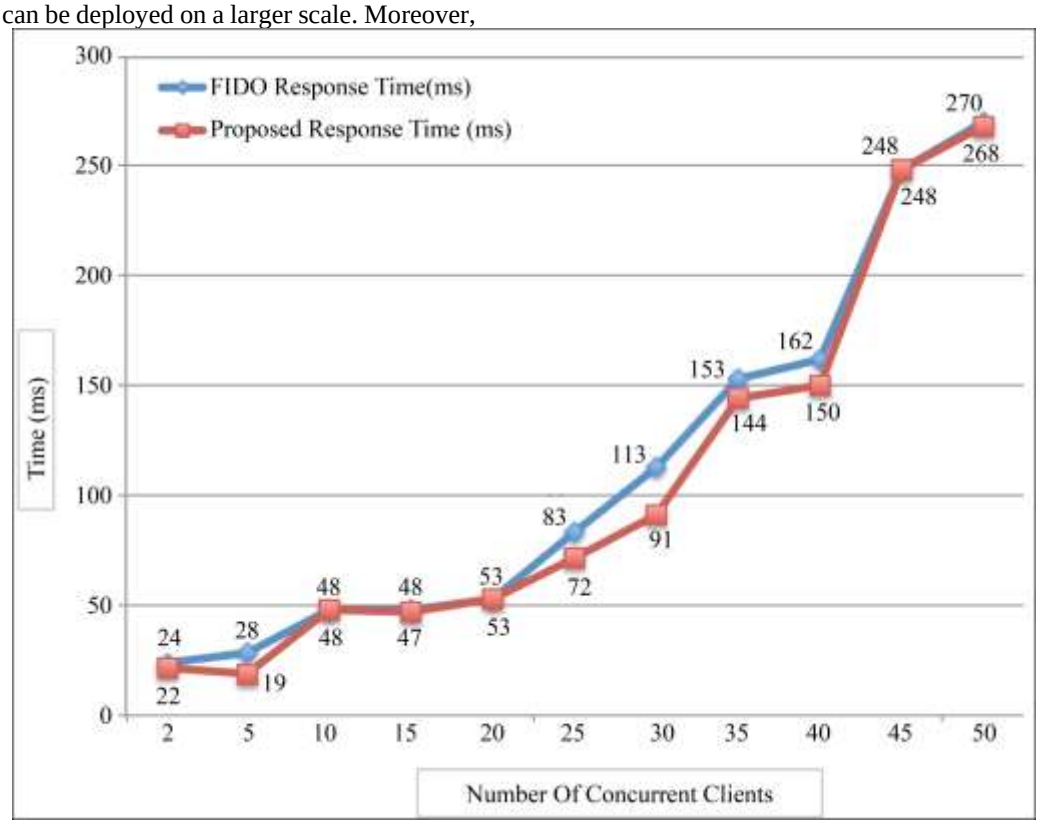

Fig. 6 Comparison of registration protocol performance

7. Conclusion

Phishing is considered a fraudulent social engineering technique that applies deceitful tactics to commit cybercrimes. The process involves stealing user-sensitive data, such as login credentials, credit card numbers, etc. A Physical Unclonable Function is a physical object based on given inputs, creates solutions, and provides a physically defined Digital Fingerprint output that serves as a unique identifier. The attacker then uses the traffic to challenge the nodes in the PUF-based authentication protocol. Applying the developed theory that, in using internet-enabled devices, ensure physical security systems, such as PUF-based authentication, are installed to eliminate data leakage and harmful intrusion solves these threats. The two well-known phishing attacks in IoT are MITM and DoS attacks. Therefore, creating wireless nodes in the authentication security protocol will help control security during MITM or DoS attacks. Therefore, this research proposed a robust, stringent, PUF-based authentication

framework with various security protocols to address IoT network security threats.

Moreover, the security and performance testing showed that the system is reliable and can manage itself in less than 300ms response time even at 50 concurrent requests, making it a viable business solution for secure organizations, mostly military and intelligence data organizations. Incorporating Dynamic Risk Assessment models, such as Random Forest and XGBoost, into the PUF-based authentication framework has significantly enhanced its capability to discern and mitigate sophisticated phishing threats in real-time. These models leverage behavioral biometrics and user interaction patterns to dynamically adjust authentication protocols, fortifying the system's resilience against MITM and DoS attacks in the IoT filed. The future directions will make the system less dependent on the client side, i.e., automatic authentication using PUF-based USB and blockchain technology to develop a foolproof security application during this research endeavour.

Acknowledgments

Sincere gratitude is extended to Loshni P. for her invaluable support in navigating the challenges encountered during this research endeavour.

References

[1] Charles Herder et al., "Physical Unclonable Functions and Applications: A Tutorial," *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1126–1141, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[2] Zhangqing He et al., "A Highly Reliable Arbiter PUF with Improved Uniqueness in FPGA Implementation using Bit-Self-Test," *IEEE Access*, vol. 8, pp. 181751–181762, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Mario Barbareschi et al., "Enforcing Mutual Authentication and Confidentiality in Wireless Sensor Networks Using Physically Unclonable Functions: A Case Study," *Quality of Information and Communications Technology*, vol. 1439, pp. 297–310, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Zhao Huang, and Quan Wang, "A PUF-Based Unified Identity Verification Framework for Secure IoT Hardware via Device Authentication," *World Wide Web*, vol. 23, no. 2, pp. 1057–1088, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Richard G. Brody et al., "Pharming and Identity Theft," *Academy of Accounting & Financial Studies Journal*, vol. 11, no. 3, pp. 43-56, 2007. [Google Scholar] [Publisher Link]

[6] Lance James, *Phishing Exposed*, Elsevier, 2005. [Google Scholar] [Publisher Link]

[7] John Thompson Okpa, Benjamin Okorie Ajah, and Joseph Egidi Igbe, "Rising Trend of Phishing Attacks on Corporate Organizations in Cross River State, Nigeria," *International Journal of Cyber Criminology*, vol. 14, no. 2, pp. 460–478, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[8] Kang Leng Chiew, Kelvin Sheng Chek Yong, and Choon Lin Tan, "A Survey of Phishing Attacks: Their Types, Vectors and Technical Approaches," *Expert Systems with Applications*, vol. 106, pp. 1–20, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[9] Zane Zheng Ma, "*Understanding the Trust Relationships of the Web PKI*," Ph.D Thesis, University of Illinois at Urbana-Champaign, 2021. [Google Scholar] [Publisher Link]

[10] Ayesha Arshad et al., "A Systematic Literature Review on Phishing and Anti Phishing Techniques," *arXix*, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Jingguo Wang et al., "Research Article Phishing Susceptibility: An Investigation into the Processing of a Targeted Spear Phishing Email," *IEEE Transactions on Professional Communication*, vol. 55, no. 4, pp. 345–362, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[12] Jukka Komulainen, Abdenour Hadid, and Matti Pietikäinen, "Context Based Face Anti-Spoofing," *IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems*, Arlington, VA, USA, pp. 1-8, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[13] Guang Xiang, and Jason I. Hong, "A Hybrid Phish Detection Approach by Identity Discovery and Keywords Retrieval," *Proceedings of the 18th International Conference on World Wide Web*, pp. 571-580, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[14] Xiaohai Tian et al., "Spoofing Detection from a Feature Representation Perspective," *IEEE International Conference on Acoustics, Speech and Signal Processing*, Shanghai, China, pp. 2119–2123, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[15] Amir Ashtari, Ahmad Shabani, and Bijan Alizadeh, "A Comparative Study of Machine Learning Classifiers for Secure RF-PUF-Based Authentication in Internet of Things," *Microprocessors and Microsystems*, vol. 93, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Aydin Aysu et al., "End-to-End Design of a PUF-Based Privacy Preserving Authentication Protocol," *Cryptographic Hardware and Embedded Systems*, vol. 9293, pp. 556–576, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[17] Jeong Min Oh, Ik Rae Jeong, and Jin Wook Byun, "An Enhanced Scheme of PUF-Assisted Group Key Distribution in SDWSN," *Journal of the Korea Institute of Information Security & Cryptology*, vol. 29, no. 1, pp. 29–43, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[18] Urbi Chatterjee et al., "Building PUF Based Authentication and Key Exchange Protocol for IoT without Explicit CRPS in Verifier Database," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 3, pp. 424–437, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[19] Vishalini Laguduva Ramnath, Sathyanarayanan N. Aakur, and Srinivas Katkoori, "Latent Space Modeling for Cloning Encrypted PUF-Based Authentication," *IFIP International Internet of Things Conference,* pp. 142-158, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] K. Nimmy, Sriram Sankaran, and Krishnashree Achuthan, "A Novel Lightweight PUF based Authentication Protocol for IoT without Explicit CRPs in Verifier Database," *Journal of Ambient Intelligence and Humanized Computing,* vol. 14, pp. 6227-6242, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[21] Armin Babaei, Gregor Schiele, and Michael Zohner, "Reconfigurable Security Architecture (RESA) Based on PUF for FPGA-Based IoT Devices," *Sensors,* vol. 22, no. 15, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[22] Basel Halak, Mark Zwolinski, and M. Syafiq Mispan, "Overview of PUF-Based Hardware Security Solutions for the Internet of Things," *IEEE 59th International Midwest Symposium on Circuits and Systems,* Abu Dhabi, United Arab Emirates, pp. 1–4, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[23] Yuanjun et al., "Single-Atom Catalysts: Synthetic Strategies and Electrochemical Applications," *Joule,* vol. 2, no. 7, pp. 1242–1264, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[24] Vishalini Laguduva et al., "Machine Learning Based IoT Edge Node Security Attack and Countermeasures," *IEEE Computer Society Annual Symposium on VLSI,* Miami, FL, USA, pp. 670–675, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[25] Li-Jun Zhang et al., "The Impact of Deep-Tier Burrow Systems in Sediment Mixing and Ecosystem Engineering in Early Cambrian Carbonate Settings," *Scientific Reports,* vol. 7, no. 1, pp. 1–9, 2017. [CrossRef] [Google Scholar] [Publisher Link]