**International Journal of Information and Cybersecurity**
**DLpress** is a publisher of scholarly books and peer-reviewed scientific research. With a dedication to academic excellence, DLpress publishes books and research papers on a diverse range of topics spanning various disciplines, including but not limited to, science, technology, engineering, mathematics, social sciences, humanities, and arts.
Published 15, February, 2024

# Developing a Multi-Level Security and Privacy-Preserved Data Model for Big Data in Healthcare: Enhancing Data Security Through Advanced Authentication, Authorization, and Encryption Techniques

**Ramya Avula** [1]

[1]**Business Information Developer Consultant, Carelon Research**

RESEARCH ARTICLE

**Abstract**

The healthcare industry faces an unprecedented challenge in managing the security and privacy of large volumes of sensitive clinical data, where breaches can compromise patient confidentiality and trust. This paper proposes a multi-level security and privacy-preserved data model adjusted for Big Data environments in healthcare. The framework integrates authentication protocols, robust authorization mechanisms, cutting-edge encryption techniques, and privacy-preserving data mining (PPDM) methods to protect sensitive healthcare information. Specifically, multi-factor authentication (MFA), role-based access control (RBAC), attribute-based access control (ABAC), homomorphic encryption, and differential privacy are discussed to create a resilient infrastructure capable of safeguarding data throughout its lifecycle. The framework incorporates real-time threat detection and response systems to ensure data integrity and availability in the face of cyber threats. This paper details the methodologies required to construct this multi-layered security architecture and highlights its efficacy in preserving privacy while allowing secure data analysis and sharing across healthcare platforms.

Keywords: attribute-based access control, differential privacy, healthcare data security, homomorphic encryption, multi-factor authentication, privacy-preserving data mining, role-based access control

OPEN ACCESS
**Reproducible Model**

***Edited by***
*Associate Editor*

***Curated by***
*The Editor-in-Chief*

***Submitted*** *18, October, 2023*

***Accepted*** *11, February, 2024*

***Citation***
*Ramya, R. (2024) Developing a Multi-Level Security and Privacy-Preserved Data Model for Big Data in Healthcare: Enhancing Data Security Through Advanced Authentication, Authorization, and Encryption Techniques*

## 1 Introduction

The explosive growth of Big Data in healthcare is fueled by advances in electronic health records (EHR), wearable devices, and genomics, introducing large data streams that reshape medical research and clinical practice.

EHRs aggregate detailed patient information, including structured data such as demographics, diagnostic codes, lab results, and medication histories, as well as unstructured data like clinical notes, imaging reports, and treatment plans. Structured data supports longitudinal analyses of patient outcomes, while unstructured text data captures nuanced clinical observations and decision rationales. The availability of this extensive historical and real-time data allows for more accurate modeling of patient trajectories, revealing trends and patterns in disease progression, treatment efficacy, and patient outcomes across diverse populations.

Wearable devices, including smartwatches, fitness trackers, and clinical-grade sensors, continuously monitor physiological parameters such as heart rate, blood pressure, activity levels, and glucose levels. These devices produce large volumes of time-series data that capture real-time physiological changes at high temporal resolution, enabling continuous monitoring of patients outside of clinical settings. Wearable data facilitates dynamic observations into chronic conditions,

| Data Type | Description | Examples |
|---|---|---|
| Structured Data | Organized in predefined formats | Demographics, lab results, medication histories |
| Unstructured Data | Free-text or non-standardized formats | Clinical notes, imaging reports, treatment plans |
| Real-time Data | Continuously updated data streams | Wearable sensor data, remote monitoring data |
| Genomic Data | High-dimensional sequence data | Whole-genome sequencing, exome sequencing |

**Table 1.** Types of Data in Healthcare

physical activity patterns, and recovery trends, offering a real-time feedback loop between patients and healthcare providers. The data generated from wearables contributes to personalized treatment plans and real-time adjustments in therapeutic approaches based on the patient's monitored physiological states.

Genomics introduces high-dimensional data into the healthcare through whole-genome sequencing (WGS), exome sequencing, and transcriptomic profiling. Each genome comprises billions of nucleotide sequences, generating a data-rich view of genetic variation across populations. This data allows for the identification of specific genetic markers associated with disease susceptibility, pharmacogenomic responses, and inherited conditions. High-throughput sequencing methods generate massive volumes of data, making it possible to conduct large-scale studies that link genetic variations to clinical outcomes. Genomics data, when combined with phenotypic data from EHRs, enables genotype-phenotype mapping, supports precision medicine, and offers observations into the molecular mechanisms underlying complex diseases like cancer, diabetes, and cardiovascular disorders.

| Security Measure | Application | Technology Used |
|---|---|---|
| Encryption | Securing data during transmission and storage | AES, TLS |
| Access Control | Managing user permissions for data access | RBAC, ABAC |
| Data Integrity | Ensuring that records are not tampered with | Cryptographic hash functions |
| Anonymization | Protecting sensitive information | De-identification, pseudonymization |

**Table 2.** Security Measures in Healthcare Data Management

Together, these advances have created a data ecosystem that allows for a more precise understanding of health, disease, and treatment response. Integrating EHR, wearable, and genomic data has transformed healthcare from a reactive discipline into a proactive and predictive science, with the potential to tailor interventions and therapies to individual patients based on a holistic view of their clinical history, real-time physiological metrics, and genetic profile. This convergence of data sources is reshaping diagnostics, personalized medicine, and population health management, providing new opportunities for identifying disease patterns, predicting outcomes, and optimizing therapeutic strategies.

This explosive growth of Big Data in healthcare necessitates a rigorous security and privacy framework. Traditional security mechanisms, often designed for small-scale systems, are insufficient for managing the vast and varied data streams in modern healthcare environments. The shift from localized storage to distributed cloud-based platforms, combined with the integration of real-time and high-dimensional data, introduces new vulnerabilities that must be addressed through advanced cryptographic techniques and distributed security models.

EHR systems, which store patient data, must ensure the confidentiality, integrity, and availability

| Device Type | Monitored Parameters | Examples of Applications |
|---|---|---|
| Smartwatches | Heart rate, activity levels | Fitness tracking, sleep monitoring |
| Clinical-grade Sensors | Blood pressure, glucose levels | Chronic condition management, remote patient monitoring |
| Wearable ECG Monitors | Electrical activity of the heart | Cardiac arrhythmia detection |
| Continuous Glucose Monitors | Glucose levels over time | Diabetes management, dietary feedback |

**Table 3.** Types of Wearable Devices in Healthcare

of sensitive information. The aggregation of data across multiple providers requires secure data sharing protocols that adhere to standards like HIPAA in the United States and GDPR in the European Union. Role-based access control (RBAC) and attribute-based access control (ABAC) mechanisms are critical in managing user permissions, allowing only authorized personnel to access or modify specific patient records. Data integrity must be protected using cryptographic hash functions, ensuring that any alterations to records are detectable and traceable. In addition, encryption protocols such as Advanced Encryption Standard (AES) and Transport Layer Security (TLS) are employed to secure data during storage and transmission, mitigating the risks of interception or unauthorized access [1].

Wearable devices, which continuously generate real-time physiological data, introduce additional security challenges due to their connection to personal devices and cloud services. The wireless nature of data transmission between wearables and mobile applications exposes this data to potential interception, making the use of end-to-end encryption critical. Secure authentication methods, such as multifactor authentication (MFA) and biometric authentication, help prevent unauthorized access to wearable data. However, the constraint of limited computational resources on wearable devices means that security protocols must be efficient, balancing the trade-offs between cryptographic complexity and power consumption. Secure data aggregation techniques, including homomorphic encryption and federated learning, allow data from multiple wearable devices to be processed collectively without exposing individual data points, maintaining privacy while enabling broader data analysis [2, 3].

Genomic data presents unique privacy challenges due to its inherent link to individual identity and familial relationships. The storage and sharing of genomic sequences require specialized data protection mechanisms, as breaches can reveal highly sensitive information about an individual's predisposition to diseases. De-identification and pseudonymization techniques, often used to strip personally identifiable information (PII) from clinical data, must be adapted for genomic data to ensure that anonymization is robust against re-identification risks. Cryptographic methods like secure multiparty computation (SMPC) and differential privacy provide means for performing statistical analysis on genomic data without exposing the underlying sequences, enabling research while preserving privacy. Secure cloud environments and blockchain technology are increasingly used to manage the consent and access rights associated with genomic data, ensuring that patients retain control over how their genetic information is used [4].

The integration of EHR, wearable, and genomic data into large-scale, multi-modal databases creates a highly useful target for cyberattacks. The scale of data and the potential for cross-referencing information across different sources require advanced intrusion detection systems (IDS) and anomaly detection algorithms, which can identify patterns indicative of malicious activities or data breaches. Zero trust architecture (ZTA) principles are increasingly adopted, where the security model assumes that threats may originate from within the network and thus continuously validates every request for access to data or systems. Additionally, data provenance tracking helps ensure that any alterations or access to sensitive data can be audited, providing transparency and accountability in the handling of healthcare data. This layered approach to security and privacy is essential to protect the integrity and confidentiality of data in an era where the interconnectedness of digital health systems significantly broadens the attack surface [5].

To address this gap, we propose a multi-level security model that incorporates advanced cryptographic techniques, dynamic authorization mechanisms, and real-time threat detection strategies. This model is designed to enhance the security of patient data while enabling healthcare providers to leverage Big Data analytics for improved patient outcomes.

## 2 Authentication Protocols

Authentication is a fundamental component in securing sensitive information in environments like healthcare systems where the confidentiality of patient data is critical. The goal of authentication is to ensure that only authorized users have access to sensitive data, preventing unauthorized access and potential data breaches. A robust authentication model must be implemented to meet the stringent security requirements of healthcare systems, making multi-factor authentication (MFA) a key element. MFA enhances security by using a combination of independent authentication methods to verify the identity of users, thus reducing the risk of unauthorized access and ensuring a higher level of trust in the authentication process.

**Table 4.** Factors in Multi-Factor Authentication (MFA)

| Factor | Description | Examples |
|---|---|---|
| Knowledge-based | Something the user knows, used for verifying identity through memorized information. | Passwords, PINs |
| Possession-based | Relies on physical objects the user must possess during authentication. | Smart cards, Hardware tokens, USB keys |
| Inherence-based | Uses biological traits unique to the user for authentication. | Fingerprints, Iris scans, Facial recognition |

Multi-factor authentication (MFA) relies on three primary factors to verify a user's identity: knowledge-based factors, possession-based factors, and inherence-based factors. Knowledge-based factors involve something the user knows, such as a password or PIN. While this method is widely used due to its simplicity and ease of deployment, it remains vulnerable to various attacks like phishing, brute force, and social engineering, as passwords can be guessed or intercepted. The second factor is possession-based, which refers to something the user has, such as a smart card, hardware token, or USB key. These physical devices generate time-based one-time passwords (TOTPs) or hold cryptographic certificates that are used to validate user identity. By requiring possession of these devices, the authentication process becomes more secure, especially in preventing remote attacks. The third factor, inherence-based, is related to something the user is, involving biometric characteristics like fingerprints, iris patterns, or facial recognition. Biometric authentication is considered highly secure because it is tied to unique physical traits that are difficult to replicate or forge, providing a robust solution for environments where high assurance is needed, such as when healthcare professionals access patient records.

The combination of these factors in MFA provides layered security, making it significantly harder for attackers to gain access, even if one factor is compromised. For example, if an attacker obtains a user's password, they would still need the physical token or biometric data to successfully authenticate. The literature emphasizes that MFA's inclusion of biometric verification offers additional security layers critical for healthcare systems, as it allows rapid and secure access to patient data while ensuring compliance with regulatory requirements like HIPAA. However, the implementation of MFA comes with challenges, including the need for infrastructure to support tokens and biometric readers, as well as user acceptance and the complexity of managing multiple authentication factors [2].

In conjunction with MFA, the Extensible Authentication Protocol (EAP) serves as a vital framework to secure communication over both wireless and wired networks. EAP is defined in RFC 3748 and is specifically designed to provide a flexible structure for deploying various authentication methods without being tied to a specific encryption mechanism. This adaptability makes EAP

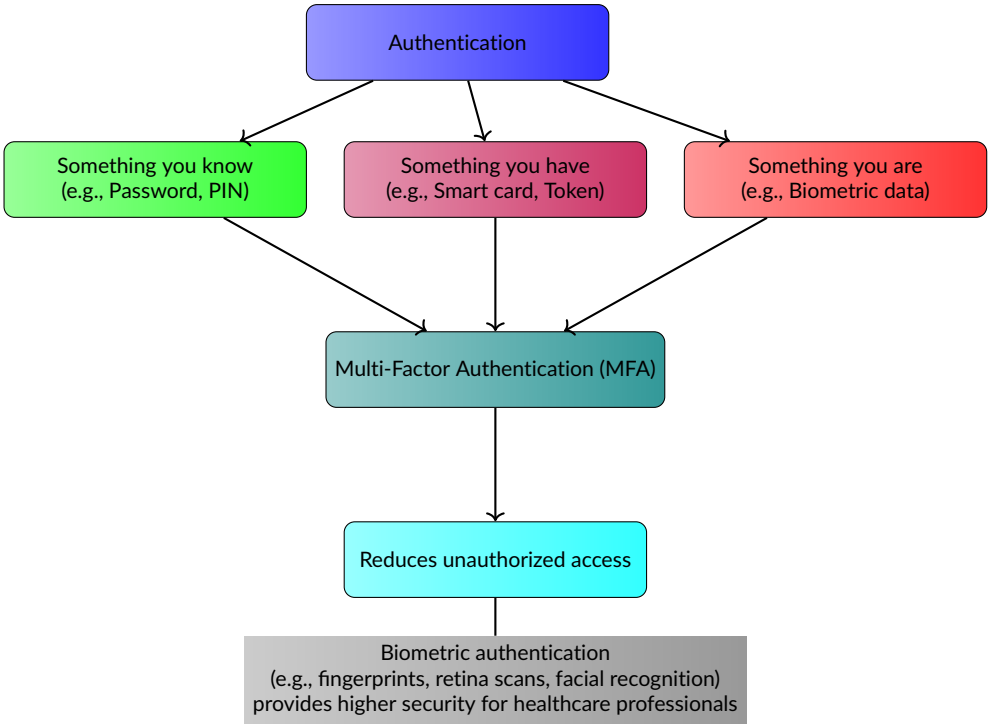


**Figure 1.** Model for Integrating Multi-Factor Authentication (MFA) in Healthcare Data Security.

suitable for diverse network environments found in healthcare systems. EAP operates at the data link layer, allowing it to be used even when IP connectivity is unavailable, thus making it ideal for securing access across wireless LANs and other networks where traditional protocols might face limitations.

**Table 5.** Comparison of EAP Variants

| EAP Variant | Security Level | Deployment Complexity | Use Cases |
|---|---|---|---|
| EAP-TLS | High, uses mutual authentication with digital certificates | High, requires PKI infrastructure for certificate management | Suitable for high-security environments like healthcare |
| PEAP | Moderate to High, uses server-side certificates to create secure tunnels | Moderate, simpler client-side configuration | Ideal for environments where managing client certificates is difficult |
| EAP-LEAP | Low, susceptible to dictionary attacks | Low, easy to deploy in resource-constrained environments | Generally not recommended for high-security needs |

Among the various EAP methods, EAP-TLS (Transport Layer Security) is one of the most secure and widely adopted methods. EAP-TLS uses the principles of public key infrastructure (PKI) to perform mutual authentication between the client and the server. Both entities are required to present digital certificates, which verify their identities before establishing a secure communication channel. This method is highly resistant to replay attacks, man-in-the-middle (MITM) attacks, and other common network threats. The use of certificates ensures that even if an attacker intercepts the communication, they cannot decrypt the data without possessing the correct private key. However, the deployment of EAP-TLS can be complex due to the need for managing certificates, which might require a robust PKI infrastructure [6].

Another important EAP variant is Protected EAP (PEAP). PEAP is designed to address some of the deployment challenges of EAP-TLS by requiring a digital certificate only on the server side. It creates a secure TLS tunnel through which other EAP methods can be executed, such as

EAP-MSCHAPv2 or EAP-GTC (Generic Token Card). This tunnel protects user credentials from interception, ensuring that sensitive information like usernames and passwords are not exposed in plaintext during transmission. PEAP simplifies client-side configuration compared to EAP-TLS, making it a more practical solution in environments where managing client certificates is difficult. However, PEAP still offers robust security through the encrypted tunnel, making it a preferred choice in scenarios where both ease of deployment and strong security are needed [7].

EAP-LEAP (Lightweight EAP) is another method, initially developed by Cisco for environments with limited computational resources. LEAP uses dynamic Wired Equivalent Privacy (WEP) keys and supports mutual authentication between the client and the server. While LEAP offers the advantage of low computational overhead and ease of use, it is less secure than TLS-based methods due to its vulnerability to certain attacks, including dictionary attacks on captured credential exchanges. As such, LEAP is less frequently recommended for high-security environments like healthcare, where data protection is paramount.

The integration of EAP with MFA in healthcare systems provides a robust security architecture. EAP's flexibility allows it to support various MFA methods, ensuring that authentication can be adjusted to meet the specific security requirements of different devices and user roles within the healthcare environment. For instance, a healthcare professional might use a combination of EAP-TLS for secure network access and biometric verification for access to specific patient data. This layered approach not only strengthens security but also ensures that access controls are appropriately enforced based on the sensitivity of the data being accessed.

The use of public-key cryptography in conjunction with EAP enhances security further by ensuring that encryption keys used during communication are exchanged securely and cannot be intercepted or manipulated. This is important in healthcare, where data integrity and confidentiality are crucial. For example, EAP-TLS dynamically generates session keys after mutual authentication, ensuring that even if an attacker gains access to previous communication sessions, they cannot decrypt future communications [8, 9]. Similarly, the TLS tunnel used in PEAP ensures that user credentials remain secure during the authentication process.

The integration of multi-factor authentication (MFA) and the Extensible Authentication Protocol (EAP) provides a comprehensive approach to securing access within healthcare systems. By combining knowledge-based, possession-based, and biometric factors, MFA mitigates the risk of unauthorized access, while EAP offers a flexible and secure framework for network authentication. Together, these mechanisms ensure that only authorized users and devices can access sensitive patient data, maintaining the confidentiality, integrity, and availability of critical healthcare information. As healthcare systems continue to digitize, leveraging such robust authentication protocols will be essential for meeting both security demands and regulatory compliance requirements.

## 3 Dynamic Authorization Mechanisms

Authorization in healthcare systems must be adaptable to accommodate the varied roles of users, such as doctors, nurses, administrative staff, and other stakeholders, each with distinct access requirements. This flexibility is critical for maintaining the security and privacy of sensitive healthcare data while ensuring that authorized personnel can access the information they need. Two of the most prominent models for managing access in such environments are Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC) [10, 11].

### 3.1 Role-Based Access Control (RBAC)

Role-Based Access Control (RBAC) is a widely implemented model in healthcare due to its straightforward design and scalability. In RBAC, permissions are assigned to users based on predefined roles that reflect their job functions within the organization. Each role is associated with a set of permissions, which define the actions that users in that role can perform on various resources. For example, a doctor might be granted access to complete medical records, whereas a nurse may only have access to specific patient information pertinent to their assigned duties. This model ensures that users have access only to the information necessary to perform their tasks, adhering to the principle of least privilege [12].

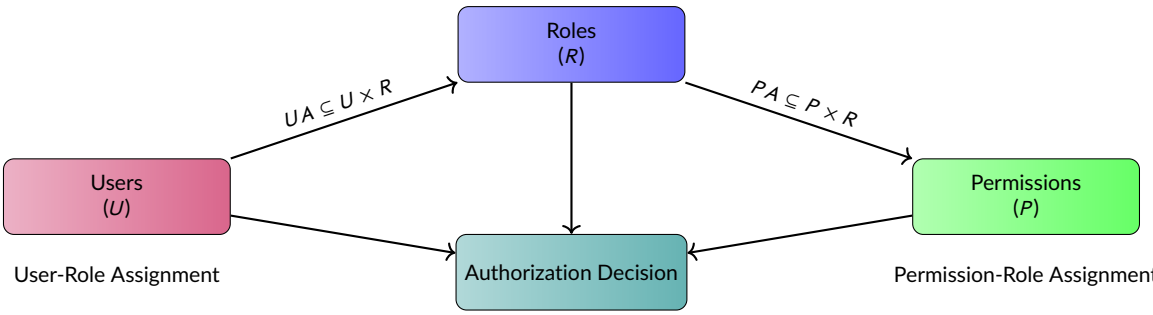


**Figure 2.** Role-Based Access Control (RBAC) in Healthcare Environments.

The underlying structure of RBAC allows for the efficient management of access rights in large healthcare environments by abstracting permissions into roles rather than assigning them directly to individual users. This abstraction simplifies the process of adding new users, modifying roles, or changing permissions as job functions change. RBAC can be described using a set of roles $R$, a set of users $U$, a set of permissions $P$, and mappings between these sets: $UA \subseteq U \times R$ represents the user-role assignments, and $PA \subseteq P \times R$ represents the permission-role assignments. The authorization decision is then derived based on these mappings, determining if a user $u \in U$ assigned to role $r \in R$ possesses the necessary permissions for a given action $p \in P$.

### 3.2 Attribute-Based Access Control (ABAC)

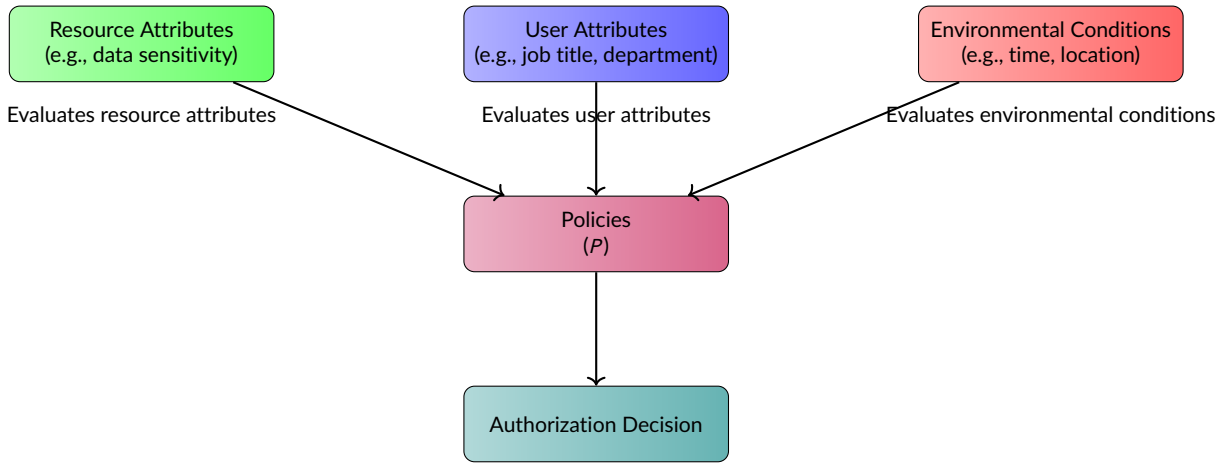


**Figure 3.** Simplified Diagram of Attribute-Based Access Control (ABAC) in Healthcare Environments.

Attribute-Based Access Control (ABAC) extends the capabilities of RBAC by introducing a more granular and context-aware approach to authorization. Instead of relying solely on roles, ABAC considers a variety of attributes associated with users, resources, and the environment. These attributes can include user attributes (e.g., job title, department, security clearance), resource attributes (e.g., sensitivity level of data), and environmental conditions (e.g., time of access, location of the request). This allows ABAC to dynamically evaluate access requests based on a combination of these attributes, offering a more fine-grained control over access to resources [13].

The ABAC model can be formally defined using a policy-based approach, where access control policies are formulated as logical expressions over attributes. Let $A$ be a set of attributes, $P$ be a set of policies, and $R$ be the set of resources. Each policy $p \in P$ is expressed as a Boolean function over the attributes: $p(a_1, a_2, \ldots, a_n) \rightarrow \{true, false\}$, where $a_1, a_2, \ldots, a_n \in A$. An access request is allowed if the policy associated with a resource evaluates to $true$ for the given attributes. For example, a policy might specify that a doctor can access patient records only if they are on duty and the request is made from within the hospital network. This approach allows the system to

adjust access control decisions in real-time based on the current context, making it suited for dynamic healthcare environments.

ABAC's flexibility makes it ideal for scenarios where users frequently change roles or require temporary access to sensitive data. It enables healthcare systems to adapt to varying conditions without the need for constant reconfiguration of roles and permissions, as is required with RBAC. For example, if a healthcare worker temporarily moves to a different department, ABAC can adjust their access rights automatically based on the change in their department attribute, without manual role reassignment. This dynamic capability ensures that access remains tightly controlled even as the system's operational environment changes [14].

### 3.3 Comparative Analysis of RBAC and ABAC in Healthcare

RBAC is often preferred for its simplicity and ease of implementation in scenarios where user roles are well-defined and change infrequently. Its structure allows for rapid deployment and straightforward management of permissions, making it suitable for many healthcare applications with stable user-role mappings [15].

ABAC, on the other hand, offers a more sophisticated and adaptive framework that is well-suited for environments with complex access control needs. It allows for dynamic adjustment of permissions based on real-time factors, offering finer control over who can access data and under what conditions. This is important in scenarios where users may need varying levels of access depending on the situation, such as emergency access to patient records during critical situations.

From a technical perspective, the choice between RBAC and ABAC often involves trade-offs between performance and flexibility. The evaluation of ABAC policies can be computationally more intensive, especially as the number of attributes and policy conditions increases. Conversely, RBAC can provide faster authorization decisions because role assignments and permissions are typically precomputed. However, ABAC's ability to adapt to context and user attributes can justify the additional complexity in scenarios where access control needs to adapt dynamically to changing conditions [16].

## 4 Advanced Encryption Techniques

Encryption is a critical component for safeguarding sensitive healthcare data both at rest (stored data) and in transit (data being transmitted). Ensuring the confidentiality, integrity, and availability of data requires advanced encryption methods that can provide robust security while allowing for efficient data processing. In the proposed framework, state-of-the-art encryption techniques such as homomorphic encryption and differential privacy are employed to achieve high levels of security and functionality, especially when handling large-scale healthcare data for analytics and machine learning applications [17, 18].

### 4.1 Homomorphic Encryption

Homomorphic encryption is a form of encryption that allows computations to be performed on ciphertexts, producing an encrypted result that, when decrypted, matches the result of operations as if they had been performed on the plaintext. This property is useful in healthcare settings where sensitive patient data must be kept confidential even during analysis. By enabling operations like addition and multiplication directly on encrypted datasets, homomorphic encryption supports advanced analytics, such as machine learning and statistical computations, without exposing the underlying data.

Homomorphic encryption schemes are defined as follows: Let $E$ be an encryption function and $D$ be the corresponding decryption function. For a given plaintexts $m_1$ and $m_2$ and operations $\circ$ (e.g., addition, multiplication), a homomorphic encryption scheme satisfies:

$$D(E(m_1) \circ E(m_2)) = m_1 \circ m_2$$

This property enables computations to be performed directly on encrypted data. Common types of homomorphic encryption include:

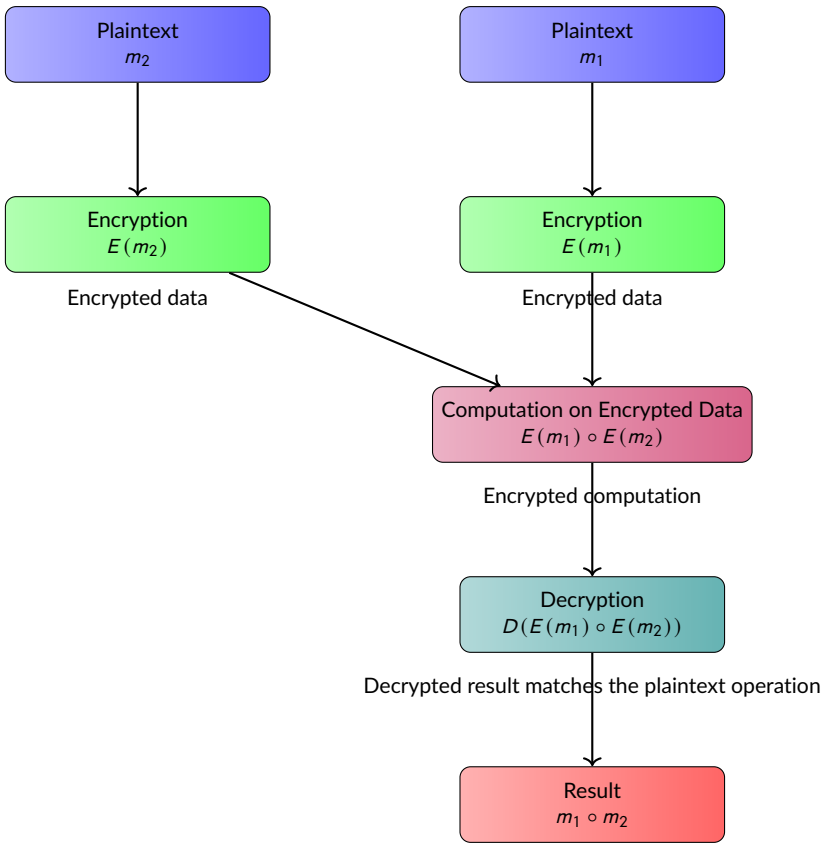


**Figure 4.** Simplified Diagram of Homomorphic Encryption Process.

- **Partially Homomorphic Encryption (PHE)**: Supports a limited set of operations, such as addition or multiplication. For example, RSA and ElGamal encryption schemes allow specific arithmetic operations on encrypted data but do not support a combination of operations.
- **Somewhat Homomorphic Encryption (SHE)**: Allows a limited number of both addition and multiplication operations but becomes impractical with increasing complexity due to noise growth during encryption.
- **Fully Homomorphic Encryption (FHE)**: Enables both addition and multiplication operations on encrypted data without limit, supporting arbitrary computation. FHE schemes, such as the Gentry scheme, utilize lattice-based cryptography and enable complex data processing while preserving data confidentiality.

In healthcare, FHE allows encrypted patient data to be used for training machine learning models without the need to decrypt the data, thereby ensuring compliance with regulations like HIPAA. Despite its theoretical strengths, FHE is computationally intensive, requiring optimizations such as bootstrapping and ciphertext packing to be practical in real-world applications [17, 18].

### 4.2 Differential Privacy

Differential privacy is a technique designed to provide privacy guarantees when analyzing datasets by adding controlled noise to the output of queries. This ensures that the presence or absence of a single individual's data in a dataset does not significantly alter the output, thereby making it difficult for attackers to infer information about any specific individual. Differential privacy is especially important in healthcare scenarios involving the sharing of data for research and analytics, where protecting patient privacy is paramount.

Formally, a randomized algorithm $\mathcal{A}$ is $\epsilon$-differentially private if, for any two neighboring datasets

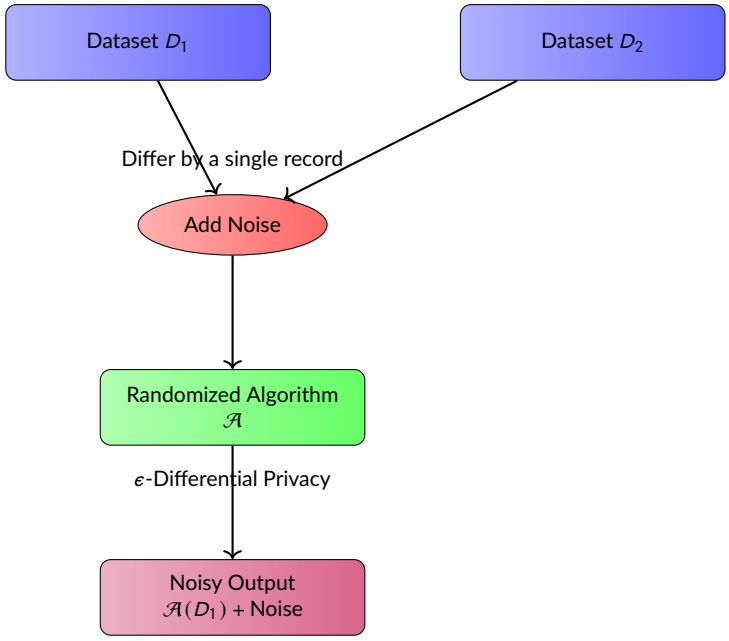


**Figure 5.** Simplified Diagram of Differential Privacy in Data Analysis.

$D_1$ and $D_2$ that differ by a single record, and for any possible output $S$ of the algorithm:

$$\Pr[\mathcal{A}(D_1) \in S] \leq e^{\epsilon} \cdot \Pr[\mathcal{A}(D_2) \in S]$$

where $\epsilon$ is a parameter that controls the privacy level—smaller values of $\epsilon$ provide stronger privacy guarantees. The noise added is typically drawn from a Laplace or Gaussian distribution and is calibrated to the sensitivity of the query, which measures how much a single individual's data can affect the query's output.

Differential privacy is well-suited for scenarios such as federated learning, where models are trained across multiple healthcare institutions without sharing raw data. By applying differential privacy, each institution can contribute to a global model while ensuring that patient-specific information remains protected. This makes it possible to perform accurate statistical analysis and machine learning while adhering to regulatory requirements like the General Data Protection Regulation (GDPR) and HIPAA [19, 20].

### 4.3 Comparison and Challenges

The choice between homomorphic encryption and differential privacy often depends on the specific requirements of the application. Homomorphic encryption provides a stronger form of data confidentiality since the data remains encrypted throughout the processing, making it suitable for applications where the data needs to remain private even from the processing party. However, the high computational overhead of FHE schemes limits their practicality, necessitating the use of optimizations like packing multiple data items into a single ciphertext to reduce computational load.

Differential privacy, on the other hand, introduces noise to achieve privacy, which allows for more efficient computations compared to FHE. However, it requires careful tuning of the noise parameter $\epsilon$ to balance privacy and utility. High privacy guarantees can reduce the accuracy of the results due to increased noise, making it crucial to select an appropriate level of privacy based on the sensitivity of the data and the specific analytic tasks.

In practice, a hybrid approach is often employed where differential privacy is used to protect data during collaborative analysis, while homomorphic encryption is applied to ensure that computations can be performed on encrypted datasets. This combination enables healthcare

providers to leverage the strengths of both techniques, maintaining high levels of security while allowing for the efficient processing of large datasets.

The integration of advanced encryption techniques such as homomorphic encryption and differential privacy into healthcare systems provides a robust solution for securing sensitive patient data. These methods ensure that data can be analyzed and shared securely, supporting both operational needs and regulatory compliance.

## 5 Privacy-Preserving Data Mining (PPDM)

Data mining is an essential process in healthcare, enabling the extraction of useful observations from large datasets. However, traditional data mining methods can expose sensitive patient information, leading to privacy concerns. Privacy-preserving data mining (PPDM) techniques are designed to mitigate these risks by ensuring that privacy is maintained during the data analysis process. Key techniques include $k$-anonymity, $l$-diversity, $t$-closeness, and randomization methods. These techniques help secure sensitive data while preserving the utility required for meaningful analytics [21].

### 5.1 $k$-Anonymity and $l$-Diversity

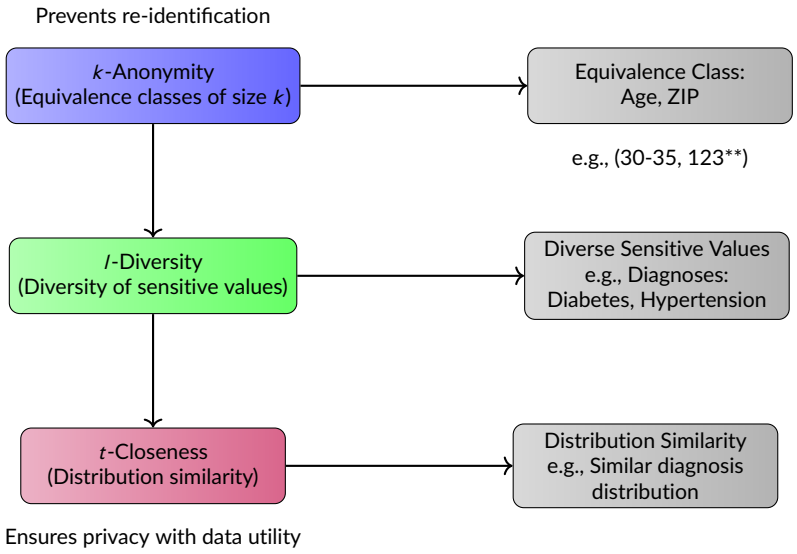


**Figure 6.** Simplified Diagram of $k$-Anonymity, $l$-Diversity, and $t$-Closeness in PPDM.

**$k$-Anonymity** is a widely used technique in PPDM that aims to prevent re-identification of individuals within a dataset. It ensures that each record is indistinguishable from at least $k-1$ other records based on a set of quasi-identifiers (QIDs). QIDs are attributes that, while not directly identifying, can be used in combination with external data to identify individuals. By grouping records into equivalence classes of size $k$, $k$-anonymity limits the ability of an attacker to associate a given record with a specific individual [21, 22]. Formally, a dataset satisfies $k$-anonymity if every combination of values for the QIDs appears in at least $k$ records.

However, $k$-anonymity does not protect against attribute disclosure, where an attacker could infer sensitive information from the homogeneity of sensitive attributes within an equivalence class. To address this, **$l$-Diversity** extends $k$-anonymity by requiring that sensitive attributes within each equivalence class have at least $l$ diverse values. This ensures that even if an attacker can identify the class to which a record belongs, the probability of inferring the sensitive attribute is reduced. $l$-Diversity can be implemented in different ways, such as:

- **Distinct $l$-Diversity**: Requires at least $l$ distinct values for the sensitive attribute in each equivalence class.

- **Entropy $l$-Diversity**: Requires the entropy of the sensitive attribute's distribution in each class to be at least $\log(l)$. This approach provides a measure of how evenly the sensitive values are distributed.

The choice of $l$ affects the balance between privacy and data utility; higher values provide greater privacy but may reduce the dataset's usefulness due to increased generalization or suppression of data.

### 5.2 $t$-Closeness

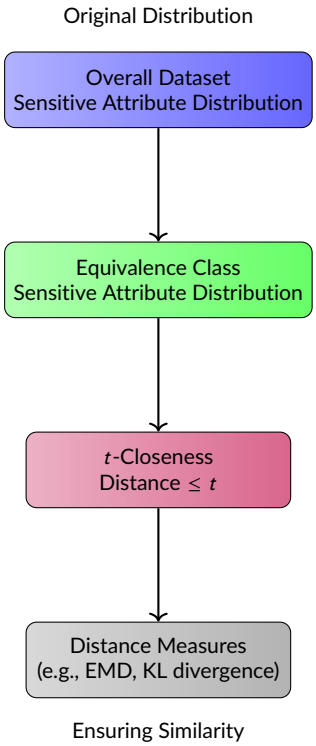


**Figure 7.** Simplified Diagram of $t$-Closeness for Privacy Preservation.

While $l$-diversity addresses some of the weaknesses of $k$-anonymity, it can still be vulnerable to attacks when sensitive attribute distributions are skewed. **$t$-Closeness** is designed to address this issue by ensuring that the distribution of sensitive attributes in each equivalence class is similar to the distribution of those attributes in the overall dataset [23]. Formally, an equivalence class satisfies $t$-closeness if the distance between the distribution of a sensitive attribute in the class and its distribution in the overall dataset is no more than a threshold $t$.

Common measures for this distance include the Earth Mover's Distance (EMD) and Kullback-Leibler (KL) divergence. EMD, for example, calculates the minimum effort required to transform one probability distribution into another. A lower $t$ value provides stricter privacy guarantees but can lead to greater loss of data utility, as more generalization is needed to maintain the distribution similarity.

### 5.3 Randomization Techniques

**Randomization** is a PPDM technique that introduces noise into the dataset to mask the original values, making it difficult to recover specific data points. Randomization can be applied at the data collection stage or during data processing. A simple randomization technique can be represented as:

$$Z = X + Y$$

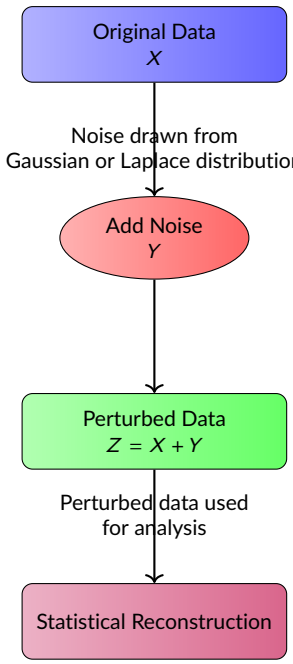


**Figure 8.** Simplified Diagram of Randomization Technique in Privacy-Preserving Data Mining.

where $X$ represents the original data, $Y$ is the noise added to each data point, and $Z$ is the resulting perturbed data. The noise $Y$ is typically drawn from a known distribution, such as a Gaussian or Laplace distribution.

Randomization can be further categorized into:

- **Additive Noise**: Adds a random value to each data point, allowing statistical properties of the dataset to be reconstructed without revealing individual values.
- **Multiplicative Noise**: Modifies data by scaling each value with a random factor, which can be more effective in some privacy scenarios.

After perturbation, the original data distribution can be estimated using statistical reconstruction techniques, allowing for the recovery of aggregate information while keeping individual records confidential. Randomization is useful in privacy-preserving machine learning, where it enables training on large-scale data while maintaining privacy.

### 5.4 Comparative Analysis and Challenges

Each PPDM technique offers unique strengths and trade-offs in terms of privacy, data utility, and computational complexity:

- **Privacy vs. Utility Trade-off**: Techniques like $k$-anonymity and $l$-diversity often require generalization or suppression, which can degrade data quality. $t$-closeness offers better control over attribute disclosure but can further reduce the granularity of data. Randomization maintains the utility of the data distribution but may result in inaccurate individual-level data [20].
- **Computational Complexity**: Implementing $t$-closeness is computationally more intensive than $k$-anonymity due to the calculation of distance measures like EMD. Randomization requires balancing the noise level to maintain privacy while preserving data accuracy, which can be challenging for real-time applications.
- **Adversarial Models**: These techniques must be designed considering potential adversarial models, such as attackers with background knowledge that could be used to correlate quasi-identifiers or sensitive attributes with external datasets. Advanced models like $t$-closeness

address some of these threats but require careful parameter tuning.

In practice, a combination of techniques may be used to balance privacy and utility, depending on the sensitivity of the data and the specific requirements of the data mining application. For instance, $k$-anonymity might be used to anonymize data at the publishing stage, while randomization techniques could be applied during data analysis to protect privacy during computation.

Privacy-preserving data mining techniques such as $k$-anonymity, $l$-diversity, $t$-closeness, and randomization play a crucial role in enabling secure analytics on sensitive datasets. These techniques allow for the extraction of observations without compromising individual privacy, making them essential tools for the secure management of healthcare data [24]. Each method involves a trade-off between privacy, data utility, and computational cost, necessitating careful consideration during implementation to meet privacy regulations like HIPAA and GDPR while ensuring that the data remains useful for analysis.

## 6 Real-Time Threat Detection and Response

As cyber threats become more sophisticated, healthcare systems must adopt real-time monitoring and anomaly detection to protect sensitive data and maintain system integrity. Real-time threat detection enables rapid identification of potential security breaches by analyzing network activity and user behavior as events occur. This section describes the integration of machine learning-based threat detection systems, anomaly-based intrusion detection systems (IDS), and their role in identifying and mitigating cyber threats in healthcare environments.

### 6.1 Machine Learning-Based Threat Detection

Machine learning algorithms are a crucial component in modern threat detection systems due to their ability to analyze large volumes of network and user activity data and identify patterns that deviate from normal behavior. These systems are typically trained on historical data to recognize the characteristics of legitimate user behavior and common network traffic. By learning from past events, machine learning models can identify patterns indicative of potential security incidents, such as unauthorized access attempts, data exfiltration, or denial-of-service (DoS) attacks.

**Algorithm 1:** Machine Learning-Based Threat Detection

**Input:** Features $X = \{x_1, x_2, \ldots, x_n\}$ from network logs or user actions
**Output:** Anomaly detection and alerts
**Initialize:** Prepare labeled dataset $D = \{(x_i, y_i)\}$, choose model $M$ (e.g., SVM, Random Forest, $k$-means, LSTM).
**Train:** Train $M$ on $D$ to learn function $f : X \to Y$.
**foreach** *new input* $x \in X$ **do**
    $y \leftarrow M(x)$ ; `// Predict behavior: normal or anomalous`
    **if** $y = 1$ **then**
        Trigger alert;
        Initiate countermeasures (e.g., block IP, require MFA);
    **end**
**end**

Let $X = \{x_1, x_2, \ldots, x_n\}$ represent a set of features extracted from network logs or user actions, where each $x_i$ is a feature vector representing a specific event or behavior. The model is trained to learn a function $f : X \to Y$, where $Y$ is a binary or multi-class label indicating normal ($y = 0$) or anomalous ($y = 1$) behavior. Common machine learning methods used for threat detection include:

- **Supervised Learning**: Techniques such as Support Vector Machines (SVM), Random Forests, and Neural Networks are trained on labeled datasets, where known attack patterns and benign behaviors are provided. These models can achieve high accuracy in identifying known threats but may struggle with novel attack vectors.

- **Unsupervised Learning**: Techniques such as clustering (e.g., $k$-means clustering) and autoencoders can identify deviations from normal behavior without requiring labeled data. These methods are effective in identifying zero-day attacks, where the threat patterns are not known in advance.
- **Anomaly Detection Models**: Models like Gaussian Mixture Models (GMM) and Long Short-Term Memory (LSTM) networks are used for detecting anomalies in time-series data, such as login attempts, network flows, and system access logs. These models can learn temporal patterns and flag deviations that may indicate suspicious activity.

Once an anomaly is detected, the system can automatically trigger an alert or initiate pre-defined countermeasures, such as temporarily blocking access from a suspicious IP address or requiring multi-factor authentication (MFA) for further verification. This real-time response capability is essential in minimizing the impact of potential security incidents, as it allows for rapid intervention before a threat escalates [22].

### 6.2 Anomaly-Based Intrusion Detection Systems (IDS)

**Algorithm 2:** Anomaly-Based Intrusion Detection System (IDS)

**Input:** Set of network packets or events $N = \{n_1, n_2, \ldots, n_m\}$
**Output:** Detection of anomalies
**Initialize:** Build model $M = \mathcal{F}(N)$, where $\mathcal{F}$ captures normal traffic patterns (e.g., mean, variance).
**foreach** *incoming data point* $n$ **do**
    Compute distance $D(n, M)$;
    **if** $D(n, M) > \theta$ **then**
        Flag $n$ as anomaly;
        Trigger alert or response action;
    **end**
    **else**
        Mark $n$ as normal;
    **end**
**end**

Anomaly-based intrusion detection systems (IDS) play a critical role in identifying both known and unknown threats by monitoring network traffic and system activities for deviations from established baselines. Unlike signature-based IDS, which relies on predefined signatures of known attacks, anomaly-based IDS can detect zero-day attacks—previously unknown vulnerabilities exploited by attackers—by identifying unusual patterns of behavior [25].

Formally, let $N = \{n_1, n_2, \ldots, n_m\}$ represent a set of network packets or system events, where each $n_i$ is characterized by a vector of features, such as source and destination IP addresses, port numbers, and packet payload size. The anomaly-based IDS builds a statistical model $M$ of normal network behavior using metrics like mean, variance, and frequency distributions:

$$M = \mathcal{F}(N)$$

where $\mathcal{F}$ is a function that captures the normal distribution of network traffic or user actions. During real-time monitoring, incoming data points $n$ are compared against the model $M$ using a distance metric $D(n, M)$. If $D(n, M)$ exceeds a predefined threshold $\theta$, the IDS flags $n$ as an anomaly:

$$\text{Anomaly} = \begin{cases} \text{True} & \text{if } D(n, M) > \theta \\ \text{False} & \text{otherwise} \end{cases}$$

Anomaly detection methods used in IDS include:

- **Statistical Methods**: These include Z-score analysis and Principal Component Analysis (PCA), which detect outliers in network data based on statistical deviations.
- **Machine Learning-Based Methods**: These include Isolation Forests and One-Class SVMs, which can identify outliers based on their distance from the majority of data points in high-dimensional feature spaces.
- **Deep Learning-Based Methods**: Autoencoders and LSTM networks are used for identifying temporal anomalies in time-series data, making them suitable for detecting complex attacks that unfold over time.

### 6.3 Integration of Real-Time Detection in Healthcare Systems

The integration of real-time threat detection mechanisms into healthcare systems provides several benefits, including proactive security measures and the ability to respond to changing threats. The system can detect anomalies such as:

- **Unusual Access Patterns**: For example, a user accessing patient records outside of normal working hours or from an unfamiliar location.
- **Brute Force Login Attempts**: Repeated failed login attempts from a single IP address or user account, indicating a potential brute force attack.
- **Malware or Ransomware Activity**: Unusual spikes in network traffic or encrypted outbound connections, which may indicate the presence of malware attempting to exfiltrate data.

Upon detection, the system can automatically initiate actions such as blocking the offending IP address, logging the event for further investigation, or requiring MFA for access to sensitive data. This automation is critical for minimizing the response time and reducing the window of opportunity for attackers [26, 27].

Additionally, anomaly-based IDS is suited for detecting zero-day attacks in healthcare systems, where attackers exploit previously unknown vulnerabilities. By identifying deviations from normal behavior, the IDS can detect these threats even in the absence of predefined signatures, providing an additional layer of defense.

### 6.4 Considerations

Real-time threat detection systems also present certain challenges:

- **False Positives**: Machine learning models and anomaly-based IDS can produce false positives, flagging legitimate activities as threats. This requires careful tuning of thresholds and regular updates to the training data to maintain accuracy.
- **Computational Overhead**: Real-time analysis of large volumes of network data can be resource-intensive, requiring robust hardware and efficient algorithms to maintain system performance without impacting user experience.
- **Data Privacy Concerns**: Real-time monitoring involves collecting and analyzing user activity data, which must be handled in compliance with privacy regulations such as HIPAA. This requires implementing privacy-preserving techniques to ensure that monitoring does not expose sensitive patient information.

Enabling rapid detection and response, these systems can prevent potential breaches from causing significant damage and ensure that patient data remains secure. The integration of machine learning-based threat detection and anomaly-based IDS into healthcare systems enhances security by enabling real-time monitoring and rapid response to cyber threats. These systems can identify deviations from normal behavior, detect zero-day attacks, and automatically initiate countermeasures to protect sensitive data [26].

## 7 Conclusion

The proposed multi-level security model is designed to safeguard sensitive patient data while enabling healthcare providers to use Big Data analytics for improved clinical decision-making and patient outcomes. This model employs a layered approach, combining advanced cryptographic techniques, dynamic authorization mechanisms, and real-time threat detection strategies.

A core aspect of this model is the integration of advanced authentication protocols, which serve as the first line of defense against unauthorized access. Authentication is critical in ensuring that only verified and authorized users can access sensitive healthcare data. Multi-factor authentication (MFA) is a fundamental component of this framework, requiring multiple independent factors for user verification. MFA relies on three primary authentication factors: knowledge-based factors like passwords or PINs, possession-based factors such as smart cards or tokens, and inherent factors, including biometric data like fingerprints or retina scans. By leveraging a combination of these elements, MFA significantly mitigates the risks of unauthorized access, as each factor provides an independent barrier to potential attackers. Biometric authentication is useful in healthcare, where high levels of security are required for accessing sensitive patient information, as highlighted by studies emphasizing the robustness of methods like facial recognition and retina scans.

To further enhance secure access across various network types, this model integrates the Extensible Authentication Protocol (EAP), which supports flexible authentication processes. EAP provides a framework that accommodates multiple authentication methods, including Transport Layer Security (TLS), Protected EAP (PEAP), and Lightweight EAP (LEAP). This adaptability allows the healthcare system to support secure communication across both wired and wireless networks, accommodating a range of devices and user scenarios. EAP's integration with public-key cryptography ensures that only authenticated users and devices can interact with the system, thus reinforcing data protection at the network level [28].

Dynamic authorization mechanisms complement these authentication protocols by controlling user access to healthcare data based on their roles and contextual attributes. The use of Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC) is central to this approach, each offering distinct advantages in managing access rights. RBAC is widely adopted due to its simplicity, as it assigns permissions based on predefined user roles. For instance, a doctor might have access to comprehensive patient records, while a nurse's access might be restricted to specific patient information directly relevant to their duties. This role-specific access helps to limit exposure of sensitive data and minimizes the potential attack surface. Its straightforward structure also makes RBAC highly scalable, making it suitable for large healthcare organizations with numerous roles and responsibilities.

In contrast, ABAC introduces greater flexibility by considering various attributes beyond user roles, such as time of access, user location, and security clearances. This allows the system to dynamically adjust permissions according to contextual factors, such as the user's department or the device they are using. ABAC's ability to incorporate multiple attributes provides a fine-grained control over access, making it especially suitable for dynamic healthcare environments where users frequently shift roles or access data from different locations. By adapting permissions in real-time, ABAC ensures that access controls remain relevant and secure even as user needs and environmental conditions change.

The security of data during storage and transmission is further bolstered by advanced encryption techniques, ensuring confidentiality throughout the data lifecycle. Homomorphic encryption and differential privacy are two advanced methods integrated into this framework, allowing secure data analytics without compromising patient privacy. Homomorphic encryption enables computations on encrypted data without requiring decryption, making it useful for healthcare analytics. This approach allows organizations to conduct complex analyses, such as machine learning, on encrypted datasets, ensuring that raw patient data remains hidden from potential threats during processing. Homomorphic encryption thus supports the dual goals of data privacy and analytical utility, facilitating secure healthcare research and data-driven observations.

Differential privacy complements this by adding carefully calibrated noise to datasets, ensuring that individual data points cannot be traced back to specific patients. This technique is crucial when healthcare data is shared for collaborative research or public health studies, as it maintains the statistical integrity of the data while protecting patient identities. Differential privacy's balance between data utility and privacy compliance makes it an effective solution for adhering to regulations like HIPAA and GDPR, which mandate stringent data protection measures in healthcare. By obscuring individual records within the broader dataset, differential privacy ensures that analytics can proceed without compromising the confidentiality of sensitive health information.

In addition to encryption, privacy-preserving data mining (PPDM) techniques play a vital role in extracting useful observations from large healthcare datasets while maintaining patient confidentiality. Traditional data mining methods pose risks of re-identification and privacy breaches, which PPDM methods such as k-anonymity, l-diversity, and randomization address effectively. k-Anonymity works by ensuring that each individual record in a dataset is indistinguishable from at least k-1 other records, reducing the likelihood that a specific individual can be identified. This technique is often enhanced by l-diversity, which ensures that sensitive attributes within anonymized groups have sufficient diversity, further protecting against re-identification attacks. These approaches ensure that even if data is accessed by unauthorized parties, the likelihood of linking records to specific individuals remains low, thus safeguarding patient privacy.

Randomization techniques add another layer of privacy by introducing noise to individual records, making it more difficult for attackers to reverse-engineer the original values. This method is suitable for privacy-preserving machine learning, where large-scale datasets are required for training predictive models without exposing raw data. By altering data points through randomization, the framework ensures that while the dataset remains useful for analysis, the underlying patient information is protected from potential breaches. These privacy-preserving techniques enable healthcare organizations to utilize advanced analytics while maintaining strict confidentiality standards.

The proposed model incorporates machine learning-based anomaly detection systems, which continuously monitor user behavior and network activities to identify suspicious activities. These systems analyze patterns, such as access times, login locations, and the volume of data accessed, to detect deviations from normal behavior. For instance, an unusual login attempt from an unfamiliar location or device could trigger an immediate alert, prompting further investigation or automated countermeasures. This real-time analysis allows healthcare organizations to quickly respond to potential security incidents, reducing the risk of data breaches before they can escalate into larger issues.

Moreover, the integration of anomaly-based intrusion detection systems (IDS) within this model provides an additional layer of security against zero-day attacks. Zero-day attacks exploit previously unknown vulnerabilities, making them difficult to detect using signature-based methods. Anomaly-based IDS addresses this challenge by identifying patterns that deviate from established norms, even if the specific threat signature is not known. By leveraging machine learning algorithms, these systems can adapt to new attack patterns over time, providing a robust defense against emerging threats. This capability is critical in healthcare environments, where the consequences of data breaches can include not only financial losses but also harm to patient trust and safety.

The proposed multi-level security model thus integrates advanced cryptographic techniques, dynamic authorization mechanisms, and real-time threat detection to create a comprehensive framework for securing healthcare data. By combining robust authentication methods, flexible access control, cutting-edge encryption, privacy-preserving analytics, and adaptive threat detection, the model offers a layered approach to protecting sensitive patient information. This enables healthcare providers to leverage the power of Big Data analytics for improving patient care while maintaining the highest standards of data privacy and security.

Homomorphic encryption is computationally demanding. The process requires significant process-

ing power and time compared to traditional encryption methods. This increased computational burden can lead to slower response times, which may be problematic in time-sensitive healthcare applications where real-time data access is critical for patient care, such as emergency services or telemedicine consultations. The requirement for high-performance computing resources can also make it challenging for smaller healthcare providers or those with limited IT budgets to fully adopt such a system, creating a potential disparity in the level of security that can be implemented across different institutions.

Although differential privacy adds noise to the datasets to protect individual privacy, this noise can impact the accuracy of the analysis when dealing with smaller datasets or when the noise must be calibrated to a high degree of privacy. In healthcare, where precise data observations are often required for critical decision-making, such as diagnosing diseases or assessing the effectiveness of treatments, the trade-off between privacy and data accuracy can become a significant issue. Overly aggressive noise addition can degrade the quality of analytical results, potentially leading to less effective or inaccurate predictions and analyses. As a result, healthcare providers may need to carefully balance the privacy requirements and analytical needs, which can complicate the deployment of this privacy-preserving framework in practical scenarios.

## References


[1] Breaux T, Antón A. Analyzing regulatory rules for privacy and security requirements. IEEE transactions on software engineering. 2008;34(1):5-20.

[2] Benaloh J, Chase M, Horvitz E, Lauter K. Patient controlled encryption: ensuring privacy of electronic medical records. In: Proceedings of the 2009 ACM workshop on Cloud computing security; 2009. p. 103-14.

[3] Chan W. Development and Evaluation of Deep Learning-Based Diagnostic Framework for Accurate Differentiation Between Benign and Malignant Breast Tumors Using Histopathological Imaging Data. Journal of Advanced Analytics in Healthcare Management. 2023;7(1):229-46.

[4] Abouelmehdi K, Beni-Hessane A, Khaloufi H. Big healthcare data: preserving security and privacy. Journal of big data. 2018;5(1):1-18.

[5] Yue X, Wang H, Jin D, Li M, Jiang W. Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. Journal of medical systems. 2016;40:1-8.

[6] Meingast M, Roosta T, Sastry S. Security and privacy issues with health care information technology. In: 2006 international conference of the IEEE engineering in medicine and biology society. IEEE; 2006. p. 5453-8.

[7] Martínez-Pérez B, De La Torre-Díez I, López-Coronado M. Privacy and security in mobile health apps: a review and recommendations. Journal of medical systems. 2015;39:1-8.

[8] Price WN, Cohen IG. Privacy in the age of medical big data. Nature medicine. 2019;25(1):37-43.

[9] Zhang R, Liu L. Security models and requirements for healthcare application clouds. In: 2010 IEEE 3rd International Conference on cloud Computing. IEEE; 2010. p. 268-75.

[10] Maisel WH. Improving the security and privacy of implantable medical devices. The New England journal of medicine. 2010;362(13):1164.

[11] Lohr KN, Donaldson MS. Health data in the information age: use, disclosure, and privacy. 1994.

[12] Rindfleisch TC. Privacy, information technology, and health care. Communications of the ACM. 1997;40(8):92-100.

[13] Schwartz PM. Privacy and the economics of personal health care information. Tex L Rev. 1997;76:1.

[14] Singleton P, Kalra D. Trust and Privacy in Healthcare. In: Future of Trust in Computing: Proceedings of the First International Conference Future of Trust in Computing 2008. Springer; 2009. p. 111-21.

[15] Smith E, Eloff JH. Security in health-care information systems—current trends. International journal of medical informatics. 1999;54(1):39-54.

[16] Itani W, Kayssi A, Chehab A. Privacy as a service: Privacy-aware data storage and processing in cloud computing architectures. In: 2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing. IEEE; 2009. p. 711-6.

[17] Fernández-Alemán JL, Señor IC, Lozoya PÁO, Toval A. Security and privacy in electronic health records: A systematic literature review. Journal of biomedical informatics. 2013;46(3):541-62.

[18] Halperin D, Heydt-Benjamin TS, Fu K, Kohno T, Maisel WH. Security and privacy for implantable medical devices. IEEE pervasive computing. 2008;7(1):30-9.

[19] Esposito C, De Santis A, Tortora G, Chang H, Choo KKR. Blockchain: A panacea for healthcare cloud-based data security and privacy? IEEE cloud computing. 2018;5(1):31-7.

[20] Dagher GG, Mohler J, Milojkovic M, Marella PB. Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. Sustainable cities and society. 2018;39:283-97.

[21] Esmaeilzadeh P. The effects of public concern for information privacy on the adoption of health information exchanges (HIEs) by healthcare entities. Health communication. 2019.

[22] Chen D, Zhao H. Data security and privacy protection issues in cloud computing. In: 2012 international conference on computer science and electronics engineering. vol. 1. IEEE; 2012. p. 647-51.

[23] Dwivedi AD, Srivastava G, Dhar S, Singh R. A decentralized privacy-preserving healthcare blockchain for IoT. Sensors. 2019;19(2):326.

[24] Coventry L, Branley D. Cybersecurity in healthcare: A narrative review of trends, threats and ways forward. Maturitas. 2018;113:48-52.

[25] Tawalbeh L, Muheidat F, Tawalbeh M, Quwaider M. IoT Privacy and security: Challenges and solutions. Applied Sciences. 2020;10(12):4102.

[26] Wang C, Zhang B, Ren K, Roveda JM, Chen CW, Xu Z. A privacy-aware cloud-assisted healthcare monitoring system via compressive sensing. In: IEEE INFOCOM 2014-IEEE Conference on Computer Communications. IEEE; 2014. p. 2130-8.

[27] Zhang Y, Zheng D, Deng RH. Security and privacy in smart health: Efficient policy-hiding attribute-based access control. IEEE Internet of Things Journal. 2018;5(3):2130-45.

[28] Zhang M, Chen Y, Susilo W. PPO-CPQ: a privacy-preserving optimization of clinical pathway query for e-healthcare systems. IEEE Internet of Things Journal. 2020;7(10):10660-72.