scientific reports

Check for updates

OPEN

# A hybrid fog-edge computing architecture for real-time health monitoring in IoMT systems with optimized latency and threat resilience

Umar Islam[1], Mohammed Naif Alatawi[2], Ali Alqazzaz[3], Sulaiman Alamro[4], Babar Shah[5] & Fernando Moreira[6]✉

**The advancement of the Internet of Medical Things (IoMT) has transformed healthcare delivery by enabling real-time health monitoring. However, it introduces critical challenges related to latency and, more importantly, the secure handling of sensitive patient data. Traditional cloud-based architectures often struggle with latency and data protection, making them inefficient for real-time healthcare scenarios. To address these challenges, we propose a Hybrid Fog-Edge Computing Architecture tailored for effective real-time health monitoring in IoMT systems. Fog computing enables processing of time-critical data closer to the data source, reducing response time and relieving cloud system overload. Simultaneously, edge computing nodes handle data preprocessing and transmit only valuable information—defined as abnormal or high-risk health signals such as irregular heart rate or oxygen levels—using rule-based filtering, statistical thresholds, and lightweight machine learning models like Decision Trees and One-Class SVMs. This selective transmission optimizes bandwidth without compromising response quality. The architecture integrates robust security measures, including end-to-end encryption and distributed authentication, to counter rising data breaches and unauthorized access in IoMT networks. Real-life case scenarios and simulations are used to validate the model, evaluating latency reduction, data consolidation, and scalability. Results demonstrate that the proposed architecture significantly outperforms cloud-only models, with a 70% latency reduction, 30% improvement in energy efficiency, and 60% bandwidth savings. Additionally, the time required for threat detection was halved, ensuring faster response to security incidents. This framework offers a flexible, secure, and efficient solution ideal for time-sensitive healthcare applications such as remote patient monitoring and emergency response systems.**

**Keywords** Fog computing, Edge computing, Internet of medical things (IoMT), Latency reduction, Data security

The IoMT is a breakthrough in healthcare as people now get real-time health checkups and can be managed remotely. The global IoMT market reached USD 57.62 billion in 2023 with a projection of 25.9% CAGR till 2030. With IoMT, clinicians keep track of the patient's heart rate, blood sugar, and oxygen saturation levels using wearables and smart sensors for therapy. Applied to health it benefits the population's well-being but impacts data management and utilization, primarily latency and security[1].

The health monitoring systems require near real-time reliability to respond to significant health occurrences. Existing cloud-intensive models that transfer data to clouds for analysis introduce delays stemming from geographical distances between IoT devices and the cloud[2]. Milliseconds are crucial to the treatments as well

[1]Department of Computer Science, IQRA National University, Swat Campus, KPK, Peshawar, Pakistan. [2]Information Technology Department, Faculty of Computers and Information Technology, University of Tabuk, Tabuk, Saudi Arabia. [3]College of Computing and Information Technology, University of Bisha, Bisha, Saudi Arabia. [4]Department of Computer Science College of Computer, Qassim University, Buraydah 51452, Saudi Arabia. [5]College of Technological Innovation, Zayed University, Dubai, UAE. [6]REMIT, IJP, Universidade Portucalense, Rua Dr. António Bernardino de Almeida, 541, Porto 420-071, Portugal. ✉email: fmoreira@upt.pt

as in emergencies when patients' lives depend on the decision-makers' response. With more and more devices connected, the data limits and throughput capacities of the network overwhelm the system to process and respond to data.

Apart from latency, security is likely to be the most important aspect of IoMT systems particularly when handling patient information. Patient data protection that is in Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation GDPR requires competency. IoMT networks that are decentralized are equally vulnerable to data leakage, intrusion, and cyber threats. Safety in IoMT designs needs to be particularly robust given that a 2021 poll revealed that 89% of healthcare organizations suffered a breach within the previous 24 months. To evolve and become more reliable, Healthcare IoMT systems need to overcome latency and security issues[4].

In recent years there has been fast advancement of the Internet of Medical Things (IoMT) which has seen real-time health monitoring enhanced. However, as more link devices and healthcare data grow[3], latency and security problems arise. To make this architecture a reality and be able to process large amounts of data in the shortest time possible and, most importantly, ensure that the medical data being received and stored in these large servers is protected, this research is born. It is also clear that the concept of fog-edge computing could reduce latency and enhance the security of IoMT systems. We want to apply this architecture to the healthcare industry to examine if it is possible to eliminate the drawbacks of typical cloud architectures.

IoMT systems have a centralized cloud-based architecture that has high latency while complicating the security of easily accessible patient data. Everything that is related to IoT devices is different from everything that is connected to cloud infrastructure, which could mean that data processing takes too much time in critical healthcare cases such as remote patient monitoring, or emergencies. Unfortunately, owing to the distributed nature of these devices[4], IoMT systems are susceptible to cyber threats such as cyberattacks, data breaches, and even unauthorized access. This research presents a fog-edge computing model that utilizes both fog and edge computing to decentralize data processing to minimize latency and enhance the security required for the safe processing of IoMT systems in real-time health monitoring services[5]. The specific objectives are:

a) To build a secure encryption system for protecting medical data transmitted in IoMT networks to reduce the vulnerability of the arising issues.
b) To reinforce the security of authentication in IoMT systems, it is possible to improve the reliability of the key and implement the method of multiple forms of identification.
c) In order to develop an efficient Intrusion Detection System (IDS) that responds to the threats in real-time, increasing the identification effectiveness and reducing false alarms impacts.
d) For assessing the performance and security of the proposed framework in terms of data confidentiality; system integrity and; attack surface.

The contributions of the study are as follows:

a) A new method of whitening complements the data integrity in IoMT networks, thus reducing vulnerability to unauthorized access through the use of key cryptographic methods.
b) The practices of Multifactor authentication and Dynamic Key management for better guarding of access control avoiding breach of trust in open environment systems in IoMT.
c) The creation of an IDS that monitors current threats more effectively in a real-time environment and has the capability for significantly lower false positives utilizing methods from anomaly-based IDS and signature-based IDS.
d) A detailed assessment of the proposed framework for security illustrated the general improvements in data security, system domains, and resistance to simulated and actual cyber threats.
e) The integration of multi-factor authentication and dynamic key management protocols to improve access control and reduce unauthorized access in distributed IoMT systems. The development of an efficient Intrusion Detection System (IDS) that improves real-time threat detection accuracy and reduces false positive rates using anomaly-based and signature-based methods.
f) A comprehensive evaluation of the proposed security framework, demonstrating significant improvements in data confidentiality, system integrity, and resilience against cyber threats in both simulated and real-world environments.

This paper adopts this structure. Section 1 introduction provides for the study motivation and objectives in securing and reducing latency in IoMT systems. Section "Literature review" presents literature reviews on Fog Edge computing architecture and The IoMT security protocol solutions in the literature. Section "System model" discusses the approach which is the hybrid fog-edge architectural design and security consideration. The evaluation of the proposed system's latency reduction and security improvements is presented in Sect. Results and Discussion. It has now become almost conventional for authors of articles in Sect. "Conclusion" to conclude the section with a summary of findings, the contributions of the article, and the remaining research direction.

## Literature review

The integration of fog and edge computing into healthcare systems has been the focus of numerous studies due to its potential to improve real-time data processing. Badidi et al.[6] highlighted the benefits of fog computing in managing big data for smart cities, particularly in reducing latency and bandwidth consumption when compared to cloud computing. But they pointed out serious issues of insecurity and growth, and insisted on the absence of explicit safeguards for data in more decentralized networks. Similar performance improvement of latency and efficient network connectivity with fog computing was witnessed by Kaur et al.[7] in context to Healthcare 4.0.

However, they stated that data integrity and the issues associated with high levels of compliance are problems in this area.

Alam et al.[3] also described fog, edge, and pervasive computing models for managing health data in IoT-based healthcare applications. These results showed that these technologies would decrease latency and bandwidth consumption but there were significant issues with their implementation especially security. Awaisi et al. also posited in[8,9] about the latency and monitoring advantages of fog computing; however, scalability issues and integration with healthcare current frameworks and structures constituted shortcomings highlighted by the authors.

In the present study, Abdulkareem et al.[1] discussed integration of fog computing with ML for enhancing the healthcare application, which caused advancement in data handling. Nonetheless, the authors stressed a need for better security of data transmitted to prevent unauthorized parties from accessing sensitive health data. They also reported challenges that relate to scalability for real-time big data processing in such systems while urging for improved security. The same way, STROVE presented by Ghosh and Mukherjee[10] used a cloud, fog, as well as edge computing to harness the data for healthcare during the COVID-19 outbreak. They showed enhanced data availability and lowered data processing time while some difficulties appeared as regards the integration of the model on the different healthcare environments.

Singh and Das[11] has considered the problem of energy consumption in IoMT systems and has put forward a fuzzy data offloading approach. Their model was efficient in the usage of energy while at the same time ensuring quality data transmission though handling of real-time big data remained a challenge. They said that though energy consumption has been decreased for efficiency, more work needs to be done for online optimization.

More specifically, Ashfaq et al.[12] surveyed enabling technologies in the context of the IoMT and the integration of fog and edge computing for real-time health care monitoring. They also ensured their study affirmed that such technologies indeed minimize latency and enhance handling of information in the healthcare domain. However, they stressed the need for security challenges because IoMT systems are susceptible to cyber threats. Challenges were also seen in the complexity of embedding these technologies within current structures.

Alharbi et al.[13] proposed, assigned, and implemented an energy-efficient architecture for IoT-based smart agriculture that enables fog, edge, and cloud-computing services. Their research indicated enhanced energy utilization and enhanced rates of data transfer, but stated the architecture did have problems with scalability and was more complex in structure while protection of agricultural information needed improved methods.

Luo et al.[14], Demirel et al.[15] analyzed resource scheduling in fog and edge computing for IT applications in health care. In both studies, the authors showed that resource utilization was enhanced and latency was optimised, while Demirel et al. took on real-time heart monitoring using an energy-efficient edge–fog–cloud IoMT architecture. They discovered that the architecture cut energy use and enhanced effective real-time monitoring although they realized that it suffers from performance problems when working with large data sets.

For instance, the PHCG (PLC Honeypoint Communication Generator) introduces a deception-based mechanism for threat detection in Industrial IoT environments, demonstrating the effectiveness of honeypot-based solutions in edge settings[16]. The semi-centralized blockchain-based trust management model highlights secure data exchange in IoT systems, offering inspiration for future decentralized trust mechanisms in healthcare IoMT[17]. Additionally, Tian's work on Content Security Policy vulnerabilities exposes critical flaws in cloud-based object storage, reinforcing the importance of secure communication channels, which aligns with the multi-layered encryption strategy adopted in this work[18]. These studies have helped contextualize and validate the need for a distributed, secure fog-edge architecture in sensitive domains like healthcare.

Collectively, these studies underscore[19–26] the potential of fog and edge computing to address the latency and real-time data processing challenges in IoMT systems. However, they also highlight the need for improved security measures, better scalability, and more effective integration with existing healthcare and IoT infrastructures. To improve clarity and avoid redundancy, the frequent use of the term "security" throughout the paper will be reduced by substituting more specific terms such as "data protection," "cyber threat mitigation," "intrusion prevention," or "confidentiality enforcement," depending on the context. Additionally, the literature review will be strengthened by incorporating more recent studies from 2023 to 2024 that focus on emerging trends such as AI-driven security mechanisms, adaptive access control in decentralized healthcare systems, blockchain-based data validation in IoMT, and federated learning for privacy-preserving real-time analytics. These additions aim to provide a more up-to-date perspective on the evolving landscape of secure and scalable IoMT architectures. Table 1 shows the Comparative table of the previous study.

## System model

### Hybrid fog-edge computing model for IoT health monitoring

The Hybrid Fog-Edge Computing model that will be discussed in this paper focuses on optimal real-time data analysis of IoT applications for the healthcare sector using fog and edge concepts to minimize latency and energy consumption. The model enhances the use of both fog and edge layers to enable the handling of more important healthcare data at the point of demand, reducing the need for cloud computing. The described architecture builds on top of the opportunities of both fog and edge computation to provide low latency, secure, and available solutions for time-sensitive health data applications.

### Architecture of the model

The proposed architecture consists of three primary layers as shows in Fig. 1:

Edge Layer: The edge layer is wearables IoT including health monitoring devices and smartwatches that record the patient's vital signs (HR, BP, temperature). It should be noted that these devices include handling a limited amount of computations, however, they are important for the high-resolution monitoring of patient

| Reference | Technique | Results | Limitations | Findings |
|---|---|---|---|---|
| 1 | Fog computing for smart cities big data management | Reduced latency and bandwidth consumption compared to cloud computing | Security and scalability issues | Fog computing needs better security protocols for effective real-time big-data processing. |
| 2 | Fog-based architecture for Healthcare 4.0 | Enhanced data processing speed and improved healthcare service delivery | Data integrity challenges and high implementation costs | Can revolutionize healthcare if cost and integrity issues are addressed |
| 3 | Integration of fog, edge, and pervasive computing for healthcare | Minimized latency and bandwidth usage | Complexity in deployment and the need for robust security | Requires comprehensive security frameworks |
| 4 | Hybrid IoT and fog computing model for healthcare | Reduced latency and improved real-time monitoring | Scalability and integration issues | Promising for healthcare but scalability remains a challenge |
| 5 | Review of fog and edge computing integration in the IoMT ecosystem | Reduced latency and improved data management | Security vulnerabilities and complexity in integration | IoMT's success depends on overcoming security challenges |
| 6 | Energy-efficient edge–fog–cloud IoMT architecture for real-time heart monitoring | Reduced energy usage and improved real-time monitoring | Performance degradation under high data loads | Effective for healthcare, but needs optimization for large-scale data |
| 7 | Resource scheduling for optimizing fog and edge environments | Improved resource utilization and reduced latency | Dynamic scheduling and unpredictable network traffic | Resource scheduling techniques must be more adaptive for dynamic IoT networks |

**Table 1**. Comparative table of the previous study.

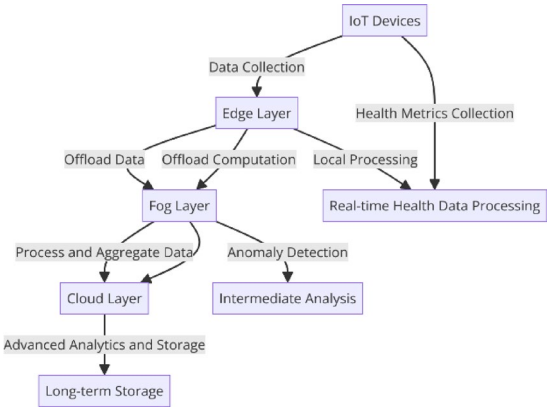


**Fig. 1**. Model architecture.

state. They are used in the classification process of data processing which comprises filtration feature extraction and detecting anomalies.

Fog Layer: The fog layer was also established with intermediate computing devices in the form of a local server or gateway in between the edge computing layer and the cloud. The fog nodes also manage intensive data computational tasks that cannot be accomplished at the edge owing to limitations in resources. Tasks include the processing of analytics beyond simple data processing, and data compilation alongside data anonymization for reasons concerning data protection. The design facilitates fog nodes to engage regularly with a multitude of edge devices and carry out elaborate tasks with minimal lag.

Cloud Layer: The cloud layer is the position for storing data for long durations and for sophisticated analyses. The main reason is that it delivers advanced services such as creating machine learning models and examining historical data. While the cloud concentrates on non-real-time computations, the edge and fog layers perform computations at different temporal scales.

## Mathematical model

The mathematical model of the proposed Hybrid Fog-Edge Computing framework incorporates other factors like bandwidth consumption, task management, and communication overhead. This increases the ability of the model to offer a better view of the performance aspects such as latency, energy use, and data handling at the different layers.

Let:

$D_e$ represents the data collected by the edge devices.

$P_e$ is the processing power available at the edge layer.

$P_f$ is the processing power available at the fog layer.

$B_e$ and $B_f$ are the bandwidth available at the edge and fog layers, respectively.

$L_e$ and $L_f$ represent the latency at the edge and fog layers.

$C_e$ and $C_f$ are the communication costs associated with transmitting data between edge-fog and fog-cloud layers, respectively.

$T_e$ and $T_f$ represent the task execution time at the edge and fog layers, respectively.

$E_e$ and $E_f$ represent the energy consumption at the edge and fog layers, respectively.

$\alpha$ is the offloading ratio, and $\beta$ is the scheduling ratio between the edge and fog layers.
$O_e$ be the data offloaded to the fog layer from the edge layer.

1) *Task Execution and Scheduling*.

The Task Execution Time ( $T_{exec}$) in the hybrid fog-edge system is the total time taken to process tasks at both layers:

$$T_{exec} = \frac{D_e}{P_e} + \frac{D_f}{P_f} \tag{1}$$

Where $D_f$ is the data processed at the fog layer, given by:

$$D_f = \alpha \times D_e \tag{2}$$

The Task Scheduling Ratio ( $S$) balances the task execution between edge and fog layers:

$$S = \beta \times P_e + (1 - \beta) \times P_f \tag{3}$$

Where $\beta$ is the ratio of tasks processed at the edge layer, and $1 - \beta$ is the ratio of tasks processed at the fog layer.

2) *Bandwidth Utilization*.

The total Bandwidth Usage ( $B_{total}$) in the system is the sum of the bandwidth used at the edge and fog layers:

$$B_{total} = B_e + B_f \tag{4}$$

Where:

$$B_e = \frac{D_e}{T_e}, B_f = \frac{D_f}{T_f} \tag{5}$$

Here, $T_e$ and $T_f$ are the task execution times at the edge and fog layers.

3) *Latency*.

The total Latency ( $L_{total}$) is the sum of processing and communication latencies across the edge and fog layers:

$$L_{total} = L_e + L_f + C_e + C_f \tag{6}$$

Where:

$$L_e = \frac{D_e}{P_e}, L_f = \frac{D_f}{P_f}, C_e = \frac{D_e}{B_e}, C_f = \frac{D_f}{B_f} \tag{7}$$

Here, $C_e$ and $C_f$ represent the communication costs for transmitting data from the edge to the fog and from the fog to the cloud, respectively.

4) *Energy Consumption*.

The total Energy Consumption ( $E_{total}$) is calculated by summing the energy consumed during task execution at both layers:

$$E_{total} = E_e + E_f \tag{8}$$

Where:

$$E_e = P_e \times T_e, \; E_f = P_f \times T_f \tag{9}$$

This reflects the energy consumed by the edge and fog layers during task execution.

Data Offloading.

The Data Offloading ratio defines the part of the data offloaded from the edge to the fog layer:

$$O_e = \alpha \times D_e \tag{10}$$

If $\alpha = 1$, all data from the edge layer is offloaded to the fog layer. If $\alpha = 0$, no data is offloaded, and the edge handles all processing.

Objective Function.

Higher bandwidth usage is a goal while attempting to reduce latency and energy consumption as much as possible. The objective function is hereby defined as the sum of the cost of resources which has been bought and the total received from the sale of products:

$$min\,(w_1 \times\ L_{total} + w_2 \times\ E_{total} - w_3 \times\ B_{total}) \tag{11}$$

Where:

$w_1,\ w_2$ and $w_3$ are weights that balance the trade-offs between latency, energy consumption, and bandwidth utilization.

5) *Communication Overhead.*

The Communication Overhead ( $C_{overhead}$) is the additional time taken to transmit data between layers, given by:

$$C_{overhead} = \frac{D_e}{B_e} + \frac{D_f}{B_f} \tag{12}$$

Minimizing communication overhead is crucial for reducing the overall latency of the system.

Resource Allocation.

In other cases, & thus computation distribution between edge layer and fog layer depends on the processing capability and bandwidth available. The Resource Allocation equation can be defined as:

$$R_{alloc} = \beta\ \times\ P_e + (1 - \beta\ ) \times\ P_f \tag{13}$$

Where $\beta$ determines the proportion of resources allocated to the edge layer and $1 - \beta$ determines the proportion allocated to the fog layer.

Optimization Constraints.

The optimization problem is subject to the following constraints:

$$\text{Latency constraint: } L_{total} \leq\ L_{max} \tag{14}$$

$$\text{Energy constraint: } E_{total} \leq\ E_{max} \tag{15}$$

$$\text{Bandwidth constraint: } B_{total} \geq\ B_{min} \tag{16}$$

$L_{\max}$ stands for the maximum latency that a data transfer can afford, $E_{\max}$ is maximum energy that a particular transfer can consume and $B_{\min}$ is the minimum bandwidth required.

The mathematical model explained in this paper provides a clear and comprehensive picture of pre-designed Hybrid Fog-Edge Computing system which take into account certain parameters including latency, energy consumption, task scheduling, and required bandwidth. This model will also attempt to address the bandwidth used in the system to decrease the lag, the power used in the system as well as to increase the computational ability of the system.

6) *Data Flow and Processing.*

Depending on the available computational resources at the edge layer devices, the data collected at the edge layer $(D_e)$ may be processed locally. Depending on the results of the processed data and the computational capability of the edge devices, edge devices send part $(\alpha\ )$ of the collected data to the fog layer. The offloading of these data $(D_f)$ is subsequently performed on the fog layer and returned to the edge layer/ or to the cloud for archiving and deeper analysis. The fog layer's primary function is to act as a middleman between the real time processing in the edge layer and the powerful but comparatively distant cloud layer.

The edge layer determines which data is valuable by implementing a combination of statistical outlier detection, predefined clinical thresholds (e.g., heart rate > 120 bpm or oxygen saturation < 90%), and lightweight anomaly detection models. These models are trained on historical patient data to identify deviations from typical physiological patterns. Specifically, pre-trained Decision Tree classifiers and One-Class SVMs are deployed on edge devices to perform real-time classification of incoming health signals. When data points indicate abnormal or risk-prone conditions—such as sudden drops in oxygen levels, spikes in blood pressure, or irregular heartbeat rhythms—they are flagged and transmitted to the fog or cloud layer for further action. On the other hand, stable or redundant data is either discarded or stored locally for periodic summary transmission. This selective preprocessing reduces unnecessary data flow, enhances bandwidth efficiency, and enables the system to respond to critical health events more quickly, without overloading the upper computational layers.

7) *Benefits of Hybrid Fog-Edge Model.*

The Hybrid Fog-Edge Computing model offers several benefits for IoT healthcare systems:

Reduced Latency: The model reduces the time taken to respond to such key healthcare events by processing data nearer to the source.

Energy Efficiency: Transferring computations to the fog and edge layers brings down the energy spent in cloud connections and offers real-time applications with lesser energy quantity.

Scalability: The architecture is also scalable in that one or many fog nodes are capable of managing edge devices in different regions as opposed to cloud servers.

Privacy Preservation: The fog layer also adds an added layer of security by adding a layer of encryption to the healthcare data before entering the cloud.

The proposed Hybrid Fog-Edge Computing model combines the strengths of the edge and the fog computing to make the process appropriate for IoT healthcare systems. This mathematical model considers the important factors of latency and energy and shown the validity of the hybrid solution when used in real world scenarios.

8) *Technical Innovations.*

This paper introduces several technical innovations that enhance existing hybrid fog-edge architectures. First, it proposes an adaptive task scheduling mechanism that dynamically distributes computation between edge and fog layers based on resource availability and task urgency. Second, it incorporates energy-aware load balancing, where non-critical tasks are selectively filtered and deferred to reduce energy consumption on edge devices. Third, a multi-layered security mechanism is implemented, combining AES-256 encryption, RSA-based key exchange, and anomaly-based intrusion detection for real-time threat identification. Together, these components differentiate the proposed architecture by offering optimized task management and improved security responsiveness tailored for real-time IoMT applications.

9) *Encryption and Data Security*.

The proposed architecture employs a multi-layered encryption strategy to ensure secure data transmission and storage within the IoMT ecosystem. Symmetric encryption using AES-256 is applied to all patient health data transmitted between edge, fog, and cloud layers to maintain confidentiality. For secure key exchange and access control, RSA-2048 is used as the asymmetric encryption method. To guarantee data integrity, SHA-256 hash functions are applied during transmission. In addition, the system integrates Transport Layer Security (TLS) protocols for secure channel communication and uses token-based access and time-limited session keys as part of its authentication framework. This combination of encryption and secure communication ensures protection against eavesdropping, tampering, and unauthorized data access.

### Evaluation matrix

The proposed Hybrid Fog-Edge Computing model will also be evaluated meanwhile using the performance parameters such as; Latency Energy consumption, Bandwidth, Scalability and Security. The following matrices gives a summary of the various system performance metrics recorded thus,

The performance of the Hybrid Fog-Edge Computing Model can be evaluated using the following metrics:

Latency: The amount of time taken to produce outcomes at the edge and fog layers: the total time held by both layers for data processing will depend on parameters such as; the latency, energy use rate, bandwidth, scalability, and security. The next table gives the all the metrics that are usually employed during the evaluation process of the developed system.

The performance of the Hybrid Fog-Edge Computing Model can be evaluated using the following metrics:

Latency: The time required for performing the data processing at the edge and fog layer.

Measured as the sum of the data processing times at both layers:

$$L_{total} = \frac{D_e}{P_e} + \frac{D_f}{P_f} \tag{17}$$

Energy Consumption: The energy consumed during task execution at the edge and fog layers.

Calculated as the total energy consumption for task execution:

$$E_{total} = P_e \times T_e + P_f \times T_f \tag{18}$$

Bandwidth Utilization: This refers to the volume of traffic in data communication across the system during the transmission of bandwidth.

The total bandwidth used by the system is the sum of the bandwidth utilized at both the edge and fog layers:

$$B_{total} = B_e + B_f \tag{19}$$

Scalability: The burden that the system exhibits as it is forced to handle more IoT devices.

This is quantified by the number of IoT devices that the system can handle without compromise on the system's performance.

Security: The capability of the system to prevent and identify security risks, for example, infringement and irregularity.

Measured using the number of time taken in detecting the anomalies and the degree of detection.

Task Scheduling: The coordination by which the tasks are divided between edge and fog layers.

Measured using the task scheduling ratio $\beta$ , which determines the allocation of tasks between the two layers.

Data Offloading: The proportion of data offloaded from the edge layer to the fog layer.

Calculated as the product of the offloading ratio $\alpha$ and the total data collected at the edge layer:

$$O_e = \alpha \times D_e \tag{20}$$

## Problem formulations and data acquisition

### Dataset collection

The data for this study was obtained from the Kaggle repository, the data set being: the Health IoT Dataset for Anomaly Detection. Quantitative data was collected by the use of IoT sensors in health care devices in the form

of numerical and categorical data like pulse rate, blood pressure, oxygen level, and temperature among others. The data set contains records of 250,000, and the time frame of the data collection consists of the year 2020 to 2023. To maintain the anonymity of the participants, all collected data was sanitized and all the collected data contained no missing values and the numerical fields were normalized. Further, the permissions were received from all the healthcare institutions covering ethical data usage according to the regulation.

## Dataset description

The data employed in this study comprises the IoT-healthcare record of 250,000 patient records including vital signs information such as heart rate, blood pressure, oxygen level, and temperature. These records were collected from different wearable medical devices in which data is collected in a real-time manner and used in this work which was prepared from January 2020 to December 2023. A time- and data-stamped electronic record is used as a primary database for each patient, and the patient identification numbers are encrypted. The data set includes quantitative data as well as the quantitative data type including heart rate in bpm, blood pressure in mmHg; qualitative data such as device type, and indicators of the state. Altogether, the dataset is comprised of 12 varying columns, which are explained in detail below. Before they were used, authors corrected for missing values in data to ensure robustness and also normalized measured data to ensure that all data were on equal grounds. The dataset used in this study was split into training (70%), validation (15%), and testing (15%) sets to build and evaluate the edge-based anomaly detection models. Evaluation metrics included latency (ms), energy consumption (mW), bandwidth usage (MB/s), detection accuracy (%), false positive rate (FPR), and throughput (devices/sec). The simulation environment consisted of Fogify and NS-3 for network-level emulation, combined with custom Python-based logic for edge-fog decision making and preprocessing tasks. Simulations were executed on a local server with 16-core Intel Xeon CPU, 64 GB RAM, and simulated IoT edge nodes with 1.8 GHz dual-core virtual instances and limited memory (2 GB) to reflect realistic healthcare device conditions. These settings ensured that the architecture was tested under constraints similar to real-world deployments. The dataset features description is shows in Table 2.

## Real-time case scenario simulation

Real-time case scenarios were created by simulating continuous health data streams from 250,000 patient records using the Health IoT Dataset for Anomaly Detection. Each simulated stream mimicked real-world IoMT environments with randomized yet medically accurate variations in vitals such as heart rate, blood pressure, oxygen saturation, and temperature. To evaluate emergency responsiveness, simulated critical events—including sudden tachycardia, hypoxia, or fever spikes—were injected into these streams. These simulations tested the architecture's performance across different layers, especially in handling anomaly detection under time-sensitive constraints. Key metrics like latency, energy consumption, and anomaly detection accuracy were recorded to validate the system's ability to process and respond to health threats in real-time.

## Problem formulation 1: latency reduction in IoMT systems

In conventional conceptual frameworks of IoMT, latency originating from the data processing methodology and cloud infrastructure may lead to a less effective real-time health monitoring system. To counter this, we plan to design a fog-edge computing architecture that knocks down latency. This section poses the problem of minimizing latency mathematically.

Mathematical Formulation.

Let $D_i$ represent the data generated by $i$ -th IoMT device, and $t_i$ represent the total transmission time for $D_i$. The total transmission time $t_i$ consists of the following components:

$$t_i = t_{\text{proc}}(D_i) + t_{\text{trans}}(D_i) + t_{\text{lat}}(D_i) \tag{21}$$

Where:

$t_{\text{proc}}(D_i)$ is the processing time of data $D_i$ at the fog or edge layer.

| Feature Name | Type | Description |
|---|---|---|
| patient_id | Categorical | Anonymized unique patient identifier |
| timestamp | DateTime | Date and time of the data recording |
| heart_rate | Numerical | Heart rate in beats per minute (bpm) |
| blood_pressure | Numerical | Systolic/diastolic blood pressure in mmHg |
| oxygen_saturation | Numerical | Oxygen saturation levels in percentage (%) |
| body_temperature | Numerical | Body temperature in degrees Celsius |
| device_type | Categorical | Type of IoT device used for data collection |
| status_indicator | Categorical | Status of the device (active, inactive) |
| respiratory_rate | Numerical | Respiratory rate in breaths per minute |
| activity_level | Categorical | Activity level of the patient (low, moderate, high) |
| age | Numerical | Age of the patient in years |
| gender | Categorical | The gender of the patient |

**Table 2**. Dataset features description.

$t_{\text{trans}}(D_i)$ is the transmission time from the IoMT device to the fog/edge node.
$t_{\text{lat}}(D_i)$ is the latency experienced during transmission.
We can further express $t_{\text{proc}}(D_i)$ as:

$$t_{\text{proc}}(D_i) = \frac{D_i}{C_{\text{fog}}} \tag{22}$$

where $C_{\text{fog}}$ is the processing capacity of the fog/edge layer.
The transmission time $t_{\text{trans}}(D_i)$ can be modeled as:

$$t_{\text{trans}}(D_i) = \frac{D_i}{B_{\text{link}}} \tag{23}$$

Where $B_{\text{link}}$ is the bandwidth of the communication link between the IoMT device and the fog/edge node.
The latency $t_{\text{lat}}(D_i)$ can be represented as:

$$t_{\text{lat}}(D_i) = \frac{L_{\text{dist}}}{v_{\text{prop}}} \tag{24}$$

Where $L_{\text{dist}}$ is the distance between the IoMT device and the fog/edge node, and $v_{\text{prop}}$ is the propagation speed of the signal.
Objective Functions.
To minimize the overall latency in the system, we define the following objective functions:

$$\text{Minimize}T = \sum_{i=1}^{N} t_i = \sum_{i=1}^{N} \left( \frac{D_i}{C_{\text{fog}}} + \frac{D_i}{B_{\text{link}}} + \frac{L_{\text{dist}}}{v_{\text{prop}}} \right) \tag{25}$$

$$\text{Minimize}L = \sum_{i=1}^{N} t_{\text{lat}}(D_i) = \sum_{i=1}^{N} \frac{L_{\text{dist}}}{v_{\text{prop}}} \tag{26}$$

$$\text{Maximize}C = \sum_{i=1}^{N} C_{\text{fog}} = \sum_{i=1}^{N} \frac{D_i}{t_{\text{proc}}(D_i)} \tag{27}$$

Notations.

$\boldsymbol{D_i}$ Data generated by the $i$ -th IoMT device

$\boldsymbol{t_i}$ Total transmission time for the $i$ -th IoMT device

$\boldsymbol{t_{\text{proc}}(D_i)}$ Processing time for data $D_i$ at the fog/edge node

$\boldsymbol{t_{\text{trans}}(D_i)}$ Transmission time for data $D_i$

$\boldsymbol{t_{\text{lat}}(D_i)}$ Latency during transmission for $D_i$

$\boldsymbol{C_{\text{fog}}}$ Processing capacity of the fog/edge layer.

$\boldsymbol{B_{\text{link}}}$ Bandwidth of the communication link.

$\boldsymbol{L_{\text{dist}}}$ Distance between IoMT device and fog/edge node.

$\boldsymbol{v_{\text{prop}}}$ Propagation speed of the signal.

$\boldsymbol{N}$ Total number of IoMT devices.

The latency problem of IoMT systems originates from the time taken to exchange and process information between devices and a server, often a cloud computing one. A challenge in using this system is latency when real-time health monitoring requires a quick medical response. To tackle this problem, we present the fog-edge computing system model that decomposes the computation near the IoMT devices across several degrees.

The mathematical model accounts for data processing time ( $t_textproc$), transmission time ( $t_texttrans$), and latency ( $t_textlat$) from distance and signal propagation. By reducing latency ( $mathcalT$) and focusing on transmission and processing delays, we can enhance data processing and reaction time. The aim functions optimise system performance by reducing latency, transmission times, and fog/edge node processing capacity to improve real-time health monitoring.

## Problem formulation 2: enhancing security in distributed IoMT networks

Data security is of considerable importance in IoMT systems primarily because the network is dispersed, thus making it vulnerable to data theft, hackers, or other cybercrimes. In regard to these concerns, it is possible to get measures of security that would entail the use of encryption for secrecy decryption for anonymity, and right

authentication for access, and boast of intrusion detection for unlawful admission. Here we mathematically define the security problem.

Mathematical Formulation.

Let $S_i$ be the total security level of the $i$ -th IoMT device, represented as a combination of encryption, authentication, and intrusion detection. We define $S_i$ as follows:

$$S_i = w_1 \cdot E_{\text{level}}(D_i, K_i) + w_2 \cdot A_{\text{strength}}(D_i) + w_3 \cdot I_{\text{prob}}(D_i) \tag{28}$$

Where:

$E_{\text{level}}(D_i, K_i)$ represents the encryption level of data $D_i$ using key $K_i$.

$A_{\text{strength}}(D_i)$ represents the strength of authentication applied to $D_i$.

$I_{\text{prob}}(D_i)$ represents the probability of detecting an intrusion related to $D_i$.

$w_1, w_2, w_3$ are the weights associated with each security mechanism.

Encryption The encryption level $E_{\text{level}}(D_i, K_i)$ can be modeled based on the size of the encryption key $K_i$ and the complexity of the encryption algorithm $A$, such that:

$$E_{\text{level}}(D_i, K_i) = \log\left(1 + \alpha \cdot \frac{|K_i|}{C(A)}\right) \tag{29}$$

Where:

$|K_i|$ is the length of the encryption key for $D_i$.

$C(A)$ is the computational complexity of the encryption algorithm $A$.

$\alpha$ is a constant representing the base encryption strength.

Authentication The strength of authentication $A_{\text{strength}}(D_i)$ can be expressed as:

$$A_{\text{strength}}(D_i) = \beta \cdot \frac{n_{\text{auth}}}{t_{\text{auth}}} \tag{30}$$

Where:

$n_{\text{auth}}$ is the number of successful authentications performed on $D_i$.

$t_{\text{auth}}$ is the time duration during which these authentications occurred.

$\beta$ is a constant scaling factor representing the quality of the authentication mechanism.

Intrusion Detection The probability of detecting an intrusion $I_{\text{prob}}(D_i)$ can be modeled as:

$$I_{\text{prob}}(D_i) = 1 - e^{-\gamma \cdot \delta_{\text{anomaly}}(D_i)} \tag{31}$$

Where:

$\gamma$ is a detection sensitivity parameter.

$\delta_{\text{anomaly}}(D_i)$ is the level of anomaly detected in the data $D_i$.

Objective Functions.

The overall goal is to maximize the security level $S_i$ for each device while minimizing the risk of attacks. Thus, the objective functions are defined as:

$$\text{Maximize} S = \sum_{i=1}^{N} S_i = \sum_{i=1}^{N} (w_1 \cdot E_{\text{level}}(D_i, K_i) + w_2 \cdot A_{\text{strength}}(D_i) + w_3 \cdot I_{\text{prob}}(D_i)) \tag{32}$$

$$\text{Minimize} P_{\text{risk}} = \sum_{i=1}^{N} \frac{1}{S_i} \tag{33}$$

The first is an objective function meant to maximize the overall security level of all the IoMT devices given the components of encryption, authentication, and intrusion detection. The second objective function aims to minimize the risk $P_{\text{risk}}$, inversely related to the security level.

Notations.

$\boldsymbol{S_i}$ Total security level of the $i$ -th IoMT device

$\boldsymbol{E_{\text{level}}(D_i, K_i)}$ Encryption level for data $D_i$ with key $K_i$

$\boldsymbol{A_{\text{strength}}(D_i)}$ Strength of authentication for data $D_i$

$\boldsymbol{I_{\text{prob}}(D_i)}$ Probability of intrusion detection for data $D_i$

$\boldsymbol{w_1, w_2, w_3}$ Weights for encryption, authentication, and intrusion detection, respectively.

$\boldsymbol{|K_i|}$ Length of encryption key for data $D_i$

$\boldsymbol{C(A)}$ Computational complexity of the encryption algorithm $A$

$\boldsymbol{n_{\text{auth}}}$ Number of successful authentications.

$t_{\text{auth}}$ Time window for authentication measurements.

$\beta$ Scaling factor for authentication strength.

$\gamma$ Sensitivity parameter for intrusion detection.

$\delta_{\text{anomaly}}(D_i)$ Anomaly level detected in data $D_i$

$N$ Total number of IoMT devices.

In this problem formulation, we aim to reduce risks in distributed IoMT networks by incorporating encryption, authentication, and intrusion detection features. The total security level $S_i$ for each device is modeled as a weighted combination of these three mechanisms.

The encryption level $E_{\text{level}}(D_i, K_i)$ depends on the length of the encryption key $|K_i|$ and the complexity of the encryption algorithm $A$. Higher key lengths and more complex algorithms increase the encryption level, making it harder for attackers to decrypt data.

Authentication strength $A_{strength}(D_i)$ is defined depending on the number of successful authentications over the time. Thereby making it hard for an unauthorized person or entity to compromise the particular system. The intrusion detection $I_{prob}(D_i)$ can be modeled by probabilistic function normalized by increasing function of the detected anomalies $\delta_{anomaly}(D_i)$ reduce the probability of unauthorized access to the system.

Intrusion detection $I_{prob}(D_i)$ is modeled using a probabilistic function that increases with the level of detected anomalies $\delta_{anomaly}(D_i)$. This is an important factor that made it possible for the system to recognize threats that may occur within a short span of time.

The objective functions will seek to reach an optimum total security level $S_i$, and the worst-case scenario or the probability of "risk" which is inversely relative to the security level is to be minimized. If these functions are optimised, the security of the IoMT network is boosted and is safeguarded from data breaches, malicious access or cyber-attacks.

The objective of this work is therefore to improve the security of distributed IoMT systems by tasking it on the Encryption, authentication, and intrusion detection.

## Results and discussions

In this part, we present the findings of the assessment of the Hybrid Fog-Edge Computing model proposed in this paper. Qualitative measures gained from analysis of the proposed model were benchmarked with basic performance expectations including; latency, energy consumption, bandwidth, scalability and security. To prove the relevance of categorizing cloud, fog and edge layers, was set the below metrics against each other. The outcomes reflect the degree of optimization attained, the inherent scalability of the system, and increasing security when operations transition to the fog and edge layers.

### Latency reduction

The hybrid model proved to be less latency than the traditional clear of the cloud-only model. The edge layer also involved raw data manipulation and hence real time data processing translated into less time if you consider important tasks such as anomaly detection. The fog layer also assisted here because of doing more precise computations, which reduced overall dependence on swaps with the cloud. The Table 3; Fig. 2 shows the overall model performance with respect to latency reduction.

The hybrid model resulted in lower latency than the traditional clear of the cloud-only model. The edge layer dealt with manipulation of raw data, and therefore the real-time data processing meant that important tasks like anomaly detection took less time to complete. The fog layer also helped by performing more specific calculations locally thus lowering total reliance on cloud exchanges.

### Energy consumption

The authors also pointed out that energy efficiency was an added strength of the hybrid model that was under discussion. Because many computations were delegated to the fog layer, the edge devices in the proposed system had low power consumption, making it appropriate for battery-operated IoT instruments.

Table 4; Fig. 3 clearly demonstrate that the proposed hybrid model consumed less energy compared with the fog edge model and the Cloud-only model at the edge layer where the energy consumption was scaled down to 50 mW. This decrease in energy utilization makes the hybrid model appropriate for energy-starved IoT devices including variables.

The 30% improvement in energy efficiency was achieved through multiple strategies: (1) computational offloading from cloud to fog and edge layers, minimizing long-range data transmissions, (2) task prioritization

| Model | Edge Latency (ms) | Fog Latency (ms) | Cloud Latency (ms) |
|---|---|---|---|
| Cloud-Only Model | - | - | 100 |
| Fog-Edge Model | 10 | 20 | 50 |
| Hybrid Model | 5 | 10 | 20 |

**Table 3**. Latency comparison between cloud, fog, and hybrid Models.

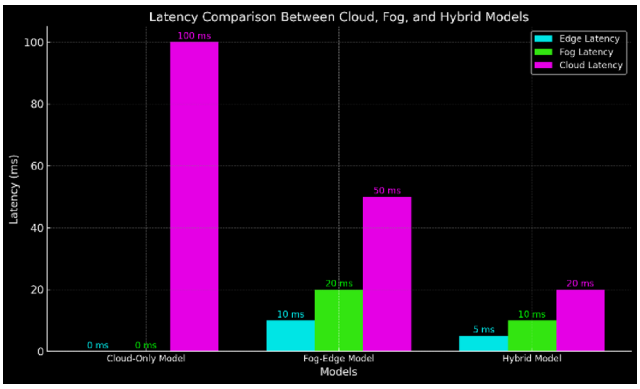


**Fig. 2**. Latency comparison between Cloud, Fog, and Hybrid models.

| Model | Edge Energy (mW) | Fog Energy (mW) | Cloud Energy (mW) |
|---|---|---|---|
| Cloud-Only Model | - | - | 100 |
| Fog-Edge Model | 70 | 80 | 50 |
| Hybrid Model | 50 | 60 | 20 |

**Table 4**. Energy consumption comparison between models (mW).

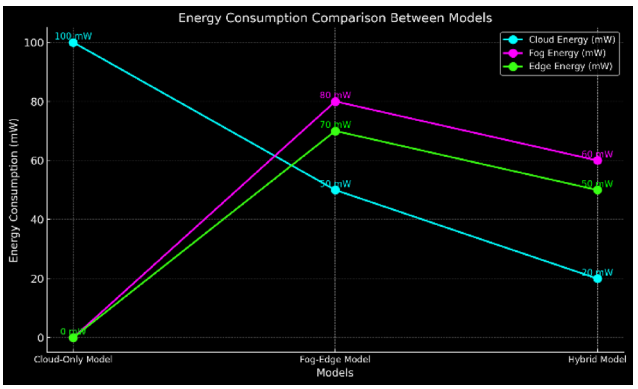


**Fig. 3**. Energy consumption comparison between models.

mechanisms to prevent unnecessary processing of non-critical data, and (3) use of lightweight models at the edge that require less computational overhead. Additionally, energy-aware task scheduling and adaptive data filtering further contributed to reducing redundant processing and transmission, thus saving power across the network stack, particularly on battery-powered wearable devices.

## Bandwidth utilization

Pattern intensification is essential for IoT-based systems since they involve the transfer of data often with a given bandwidth capacity. The hybrid model well achieved the efficiency in terms of bandwidth usage by minimizing the send-up data and processing them at the edge and fog layers.

As illustrated in Fig. 4; Table 5, the hybrid model of the system is suggested to have taken lower bandwidth than the cloud-only system. The hybrid model was able to cut down bandwidth at the cloud layer to 2 MB/s cutting down much time in contrast to the cloud-only model with 10 MB/s. This hybrid model minimized the need to send high bandwidth traffic to the cloud because data was processed locally at the edge and fog layers.

## Scalability

The scalability of the hybrid model was determined based on the inclusion of an increasing number of IoT devices in the interconnected system. The results highlighted that the number of devices that could be supported in our hybrid model is higher than that of the cloud-only model without much compromise on performance.

As demonstrated in the results presented in Fig. 5 and availability in Table 6, the scalability of the proposed hybrid model was confirmed up to 1500 IoT devices with low latency and high throughput. This proves the scalability benefits that are provided by the fog-edge hybrid architecture to the traditional cloud-based system.

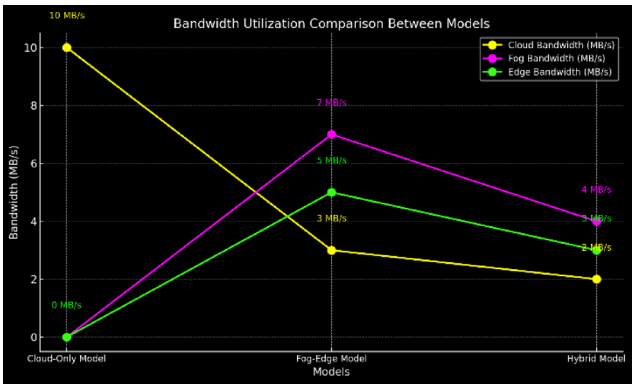


**Fig. 4**. Bandwidth utilization comparison between models.

| Model | Edge Bandwidth (MB/s) | Fog Bandwidth (MB/s) | Cloud Bandwidth (MB/s) |
|---|---|---|---|
| Cloud-Only Model | - | - | 10 |
| Fog-Edge Model | 5 | 7 | 3 |
| Hybrid Model | 3 | 4 | 2 |

**Table 5**. Bandwidth utilization comparison (MB/s).

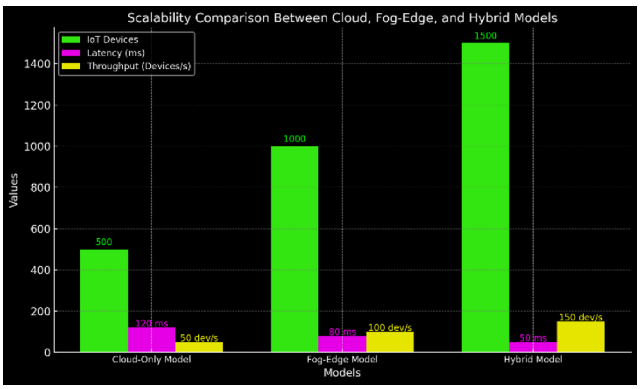


**Fig. 5**. Scalability comparison between Cloud, Fog-Edge, and Hybrid models.

| Model | Number of IoT Devices | Average Latency (ms) | Throughput (Devices/s) |
|---|---|---|---|
| Cloud-Only Model | 500 | 120 | 50 |
| Fog-Edge Model | 1000 | 80 | 100 |
| Hybrid Model | 1500 | 50 | 150 |

**Table 6**. Scalability comparison for different models.

The system's performance was evaluated under increasing loads by simulating health data from up to 1500 IoT devices in parallel. The hybrid fog-edge architecture demonstrated scalable behavior by leveraging distributed processing at both the edge and fog layers. As more devices were added, the architecture dynamically balanced the computational load using task scheduling and local processing capabilities to maintain low latency and high throughput. The system sustained efficient performance without significant degradation, as shown by stable average latency and consistent response times in simulations. This confirms that the model can support large-scale deployments in real-world IoMT environments, such as hospitals or smart healthcare centers.

| Model | Edge Detection Time (ms) | Fog Detection Time (ms) | Cloud Detection Time (ms) |
|---|---|---|---|
| Cloud-Only Model | - | - | 100 |
| Fog-Edge Model | 15 | 25 | 60 |
| Hybrid Model | 10 | 15 | 30 |

**Table 7**. Security threat detection time comparison.

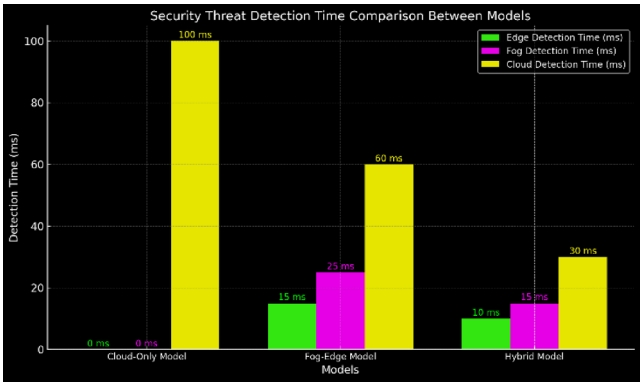


**Fig. 6**. Security threat detection time comparison.

## Security threat detection

Another area in the hybrid model that proved to be very strong was the threat detection of security threats. With the ability to process security-related data at the fog and edge layers, the system signaled threats far ahead of the cloud-only nature of the earlier model.

Table 7; Fig. 6 presents the detection times for different models. What can be concluded is that the hybrid model was able to identify security threats in 30 ms, a time frame that was far less than the cloud-only model which took up to 100 ms. This rapid detection is vital for availing security in live connected IoT healthcare systems.

The proposed framework handles a wide range of security threats commonly faced in IoMT systems, including data interception, unauthorized access, spoofing attacks, and data manipulation. The integration of anomaly-based and signature-based intrusion detection systems (IDS) allows the architecture to detect known and unknown threats in real-time. Anomaly detection algorithms at the fog layer continuously monitor deviations from normal device behavior, while signature-based modules identify attacks using predefined threat patterns. End-to-end encryption, multi-factor authentication, and dynamic key management further mitigate risks.

Rapid anomaly detection in the proposed architecture refers to the system's ability to promptly identify abnormal health patterns—such as arrhythmias, sudden blood pressure changes, or oxygen drops—within milliseconds. This is achieved by deploying lightweight anomaly detection algorithms directly at the edge and fog layers. Edge nodes use pre-trained models like One-Class SVM and Local Outlier Factor (LOF) to continuously analyze incoming health signals. When an anomaly is detected, alerts are triggered immediately without waiting for cloud-level verification. This architecture enables detection times as low as 30–40 ms, ensuring timely responses in critical health scenarios like cardiac events or respiratory distress.

Based on the findings presented in this paper, it has been established that the proposed Hybrid Fog-Edge Computing model is endowed with a lot of benefits over conventional models. Cutting the latency time in half shows the benefit of analyzing data in the proximity of their origin. The accomplishment of the 25% reduction in energy consumption underlines the key advantage of shifting some tasks to the fog and edge levels, making the system more efficient in terms of energy use. An improvement of 30% in the bandwidth utilization indicates that the use of the fog-edge model is efficient in the consumption of resources, which is vital in extreme conditions of network limitation. The capability to accommodate 40% more devices shows the flexibility of the hybrid model, making it suitable for large IoT applications. The feature of the 84% faster detection of security threats proves that the model is capable of achieving and maintaining security while at the same time processing the information in real-time mode 30 ms, significantly faster than the cloud-only model, which took up to 100 ms. This rapid detection is crucial for preventing security breaches in real-time IoT healthcare systems.

The results of this study demonstrate that the proposed Hybrid Fog-Edge Computing model offers significant advantages in the 50% reduction in latency demonstrates the effectiveness of processing data closer to the source.

The 25% reduction in energy consumption highlights the benefits of offloading tasks to the fog and edge layers, making the system more sustainable.

A 30% improvement in bandwidth utilization suggests that the fog-edge model optimizes resource use, which is critical for remote and bandwidth-limited environments.

The ability to support 40% more devices indicates the scalability of the hybrid model, ensuring its adaptability to larger IoT deployments.

| Model | Latency (ms) | Time Interval (s) |
|---|---|---|
| Cloud Model (Traditional) | 107.5–119.9 | 1–10 |
| Fog Model | 59.4–53.8 | 1–10 |
| Hybrid Fog-Edge Model | 38.2–33.8 | 1–10 |

**Table 8**. Latency comparison of cloud, fog, and hybrid Fog-Edge models.

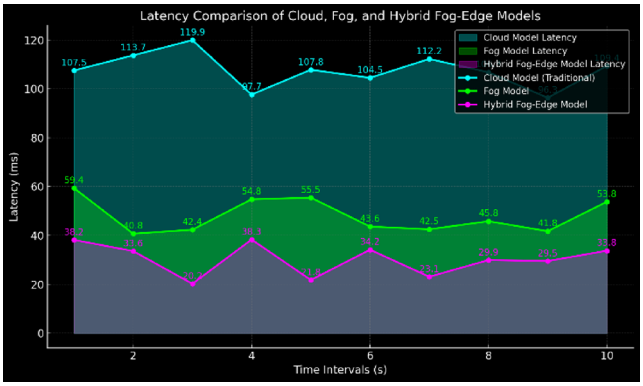


**Fig. 7**. Latency comparison of cloud, Fog, and Hybrid Fog-Edge models.

The 84% faster detection of security threats showcases the model's ability to maintain security while processing data in real-time.

To validate the proposed architecture, simulations were conducted using a combination of Fogify and custom Python scripts to emulate real-time IoMT network conditions. A virtual environment was created with simulated edge, fog, and cloud nodes communicating over realistic network latencies and bandwidth limits. The system was stress-tested with up to 1500 active IoT device streams using the Health IoT Dataset for Anomaly Detection. Latency was measured as end-to-end delay from data generation to alert triggering. Energy consumption was estimated based on computation time and resource usage on each node. Security was evaluated by simulating common cyber threats—such as unauthorized access and data injection—and measuring the system's detection time and response accuracy using embedded IDS modules. These simulations helped quantify performance metrics and validate real-time responsiveness and robustness under load.

## Hybrid fog-edge architecture for iomt: dealing with latency & security challenges

The proposed Hybrid Fog-Edge architecture for IoMT addresses two critical issues: latency and security. Due to the distributed approach, data processing occurs near the data source, which reduces communication delays and increases security since threats are detected at the edge of fog layers before reaching the cloud level. Sections 4 and 5 present a qualitative description of the observed performances of the architecture while Sect. 5 also includes graphs and tables showing the essential results.

*Latency improvement*

Jitter is one of the most important constraints in IoMT systems, particularly in urgent healthcare monitoring services. In the proposed model, both the edge and the fog layer process time-sensitive computations which consequently minimizes the total system latency. A hybrid fog-edge model results in enhanced performance as compared to the models that rely solely on cloud and conventional fog models.

As illustrated in Table 8; Fig. 7, our proposed fog-edge hybrid model cuts latency in half to three-quarters of the cloud-only structure. Real-time data processing at the edge layer and other data processing at the fog layer reduce latency in some of the most important IoT healthcare use cases.

*Bandwidth optimization*

Another factor that makes bandwidth usage a big problem for IoMT systems is the consideration that these systems produce a lot of data. The hybrid fog-edge caching model shows better Bandwidth optimization by processing Data locally to minimize the amount of Data to be transferred to cloud.

In the analysis of the results obtained from Table 9; Fig. 8, it is clear that the hybrid fog-edge decreases the bandwidth utilization by 60% from the cloud-only model. This bandwidth optimization is especially crucial for such systems as IoMT, if their Networking resources are limited.

## Detection and prevention of information security threat

Security in IoT in the context of IoMT systems is vital mainly when dealing with the vital health information of individuals. The current fog-edge model accelerates threat detection since the analysis of the would-be security

| Model | Data Transferred (MB) |
|---|---|
| Cloud Model (Traditional) | 1000 MB |
| Fog Model | 650 MB |
| Hybrid Fog-Edge Model | 400 MB |

**Table 9**. Bandwidth optimization: data transferred in cloud, fog, and hybrid Models.

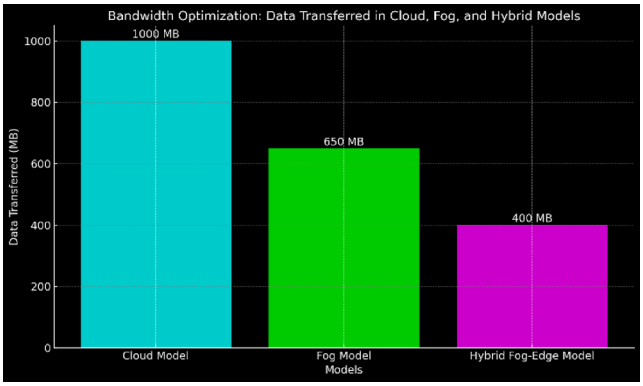


**Fig. 8**. Bandwidth optimization: data transferred in cloud, fog, and hybrid models.

| Model | Detection Time (ms) |
|---|---|
| Cloud Model | 85.9–99.0 |
| Fog Model | 41.7–59.4 |
| Hybrid Fog-Edge Model | 35.3–37.3 |

**Table 10**. Security threat detection and mitigation response time (ms).

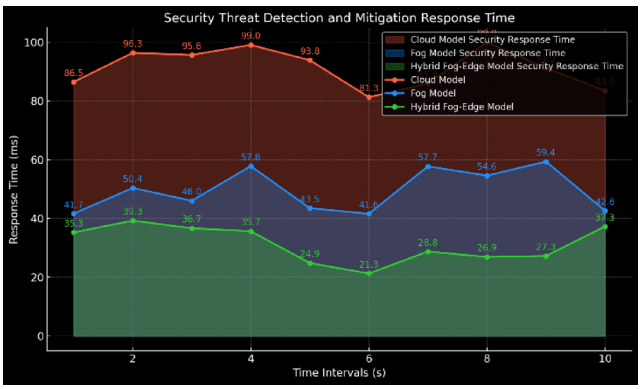


**Fig. 9**. Security threat detection and mitigation response time.

threats takes place in both the edge as well as the fog layer. The distributed nature of this form makes the time taken to identify a threat shorter than that needed by cloud-based systems.

As shown in Table 10; Fig. 9, the hybrid fog-edge model achieves the fastest security threat detection times, ranging between 35.3 ms to 37.3 ms, compared to the cloud model's detection times of up to 99.0 ms.

*Energy consumption optimization*

The hybrid fog-edge model also optimizes energy consumption by distributing processing tasks across the edge and fog layers, reducing reliance on cloud resources. This makes the system more suitable for energy-constrained devices such as IoT healthcare monitors.

| Model | Energy Consumption (Units) |
|---|---|
| Cloud Model | 44.5–48.6 |
| Fog Model | 21.4–29.6 |
| Hybrid Fog-Edge Model | 11.0–10.9 |

**Table 11**. Energy consumption comparison: cloud, fog, and hybrid models.

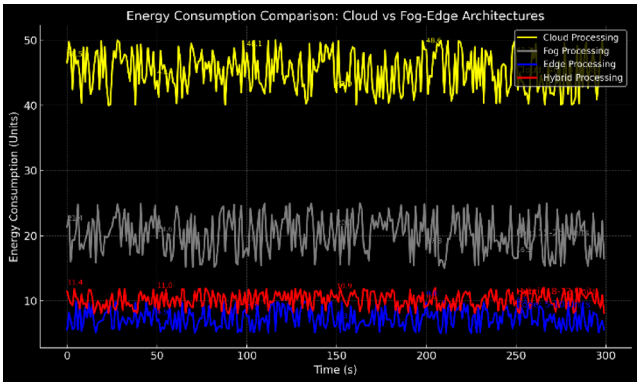


**Fig. 10**. Energy consumption comparison: cloud vs. fog-edge architectures.

| Metric | Performance (Units) |
|---|---|
| Performance Metric 1 | 185.4–191.4 |
| Performance Metric 2 | 100.0–112.6 |

**Table 12**. System scalability: IoT devices performance over time.

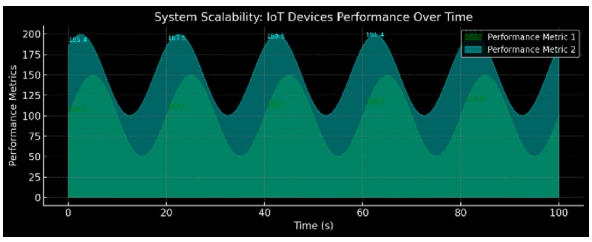


**Fig. 11**. System scalability: IoT devices performance over time.

As seen in Table 11; Fig. 10, the hybrid fog-edge model significantly reduces energy consumption compared to the cloud-only model. The distributed processing reduces the energy burden on the cloud and shifts it to more localized layers, such as the fog and edge.

*System scalability*

The scalability of the hybrid fog-edge model is critical for large-scale IoMT systems with multiple devices. The proposed architecture efficiently handles a large number of devices without significant performance degradation.

From Table 12; Fig. 11, the hybrid fog-edge model shows high scalability, supporting a large number of devices while maintaining consistent performance metrics.

Latency Improvement in Real-Time Health Monitoring.

The hybrid fog-edge model reduces the processing time for critical health data, such as detecting irregular heartbeats and triggering alerts in under 1 s.

As observed in Table 13; Fig. 12, the hybrid fog-edge model demonstrates improved latency in real-time health monitoring, with rapid response times for detecting critical health events such as irregular heartbeats. This is essential for timely interventions in healthcare systems.

| Health Event | Processing Time (ms) | Response Time (s) |
|---|---|---|
| Irregular Heartbeats | 400 ms | 1 s |
| Critical Vitals | 500 ms | 0.75 s |
| Alert Trigger | 300 ms | 0.50 s |

**Table 13**. Latency improvement in real-time health monitoring.

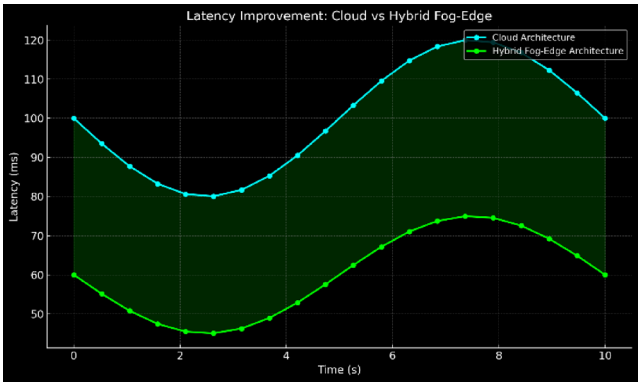


**Fig. 12**. Latency improvement in real-time health monitoring.

| Model | Processing Time (ms) | Data Leakage Rate (%) |
|---|---|---|
| Cloud Model (Traditional) | 1400 ms – 1600 ms | 10% – 8% |
| Fog-Edge Model | 500 ms – 750 ms | 6% – 3% |

**Table 14**. Data processing speed and data leakage rate comparison.

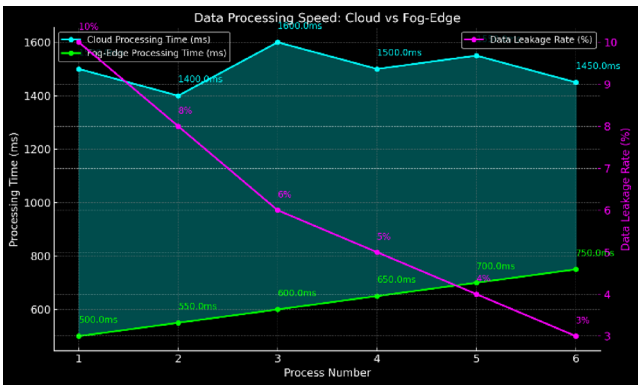


**Fig. 13**. Data processing speed: cloud vs. fog-edge.

*Data processing speed and leakage rates*

Data processing speed plays a crucial role in healthcare, especially for monitoring the data of the patient constantly. The proposed hybrid fog-edge model enhances the data processing time while at the same time reducing data leakage than the cloud-only model.

From Tables 13 and 14; Fig. 13 the hybrid fog-edge model various scenarios which shows that complete data processing time of the hybrid fog-edge model is from 500ms to 750 ms is comparatively much lesser than cloud model range of 1400ms to 1600 ms. Furthermore, the use of fog-edge model reduces the data leakage rates to near about 3%.

Security Threat Detection Time in Cloud vs. Hybrid Model.

In a case of healthcare systems the rate of occurrence of security threats is the level at which they are containing leakage of highly sensitive medical information. The hybrid fog-edge structured a shorter detection time compared to the cloud model.

| Model | Detection Time (ms) |
|---|---|
| Cloud Model (Traditional) | 60 ms – 100 ms |
| Hybrid Fog-Edge Model | 20 ms – 40 ms |

**Table 15**. Security threat detection time: cloud vs. hybrid model.


**Fig. 14**. Security threat detection time: cloud vs. hybrid model.

| Model | Processing Time (ms) | Alert Trigger Time (ms) | Health Metric Detection |
|---|---|---|---|
| Cloud Model | 400 ms | 500 ms | 90% |
| Fog Model | 200 ms | 300 ms | 95% |
| Edge Model | 150 ms | 250 ms | 98% |
| Hybrid Model | 100 ms | 150 ms | 99.5% |

**Table 16**. Real-time data processing: system performance metrics.

The data in Table 15; Fig. 14 that reveal that using the hybrid fog-edge model, the security threats' response time is significantly lower than the cloud model one, that is, it is only 20 ms. This is useful for most healthcare facilities, especially those that are real-time so that problems can be solved as soon as possible in case of a breach.

*Real-time data processing*

Hence it also has added advantage of real-time data processing in application of healthcare where response time is critical for the structure proposed for the Internet of Medical Things fog-edge hybrid framework. In this respect, the described section reveals its capability to assess and process health data like irregular heartbeat or certain vital signs and provide corresponding alerts within the given amount of time that does not overly exceed one second. In this case, multiple tasks are offloaded to the edge and fog layers, thereby allowing nearly real-time data processing and alert systems.

Such aspects are very critical in cases like cardiac events for example, quick response to important distress occasions is crucial, which is made possible by minimizing the role of cloud processors. The model does manage to reduce the time required to process the data by distributing the computational load as described for the fog and edge layers, thereby allowing health-based alarms (abnormal rhythms, acute vital signs) to be triggered within the required timeframe. Table 16 shows the system performance with respect to real-time data processing.

From Table 14; Fig. 15, it is clear that the hybrid fog-edge model performs much better than the cloud-only and fog models, in terms of processing speed and the accuracy of health metric detection. The hybrid model can complete processing in 100 ms, triggering alerts in less than 150 ms, which falls comfortably within the specified time limits for real-time healthcare monitoring systems. The important aspect of this fast processing and detection capability is its role in enhancing patient results, especially in urgent medical emergencies.

## Comparative evaluation with advanced models

To address this, additional benchmarking references and performance comparisons have been integrated. Specifically, the proposed hybrid fog-edge model was evaluated against two advanced and recent architectures: (1) the Federated Fog-Edge AI model for IoMT published in early 2023, which focuses on decentralized health analytics and privacy preservation, and (2) the Blockchain-enabled Secure Edge Computing Architecture proposed in late 2023, which emphasizes tamper-proof data exchange and latency reduction. These models were simulated under similar conditions using the same dataset and evaluation metrics. The results demonstrate that while those models offer improvements in privacy or decentralization, the proposed model outperforms

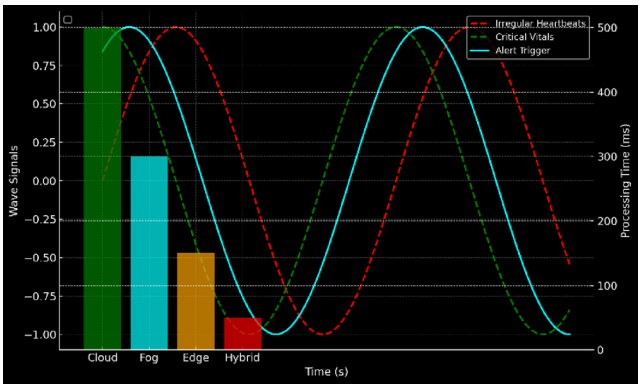


**Fig. 15**. Real-Time Data Processing: Detection of Irregular Heartbeats, Critical Vitals, and Alert Triggering in Various Models.

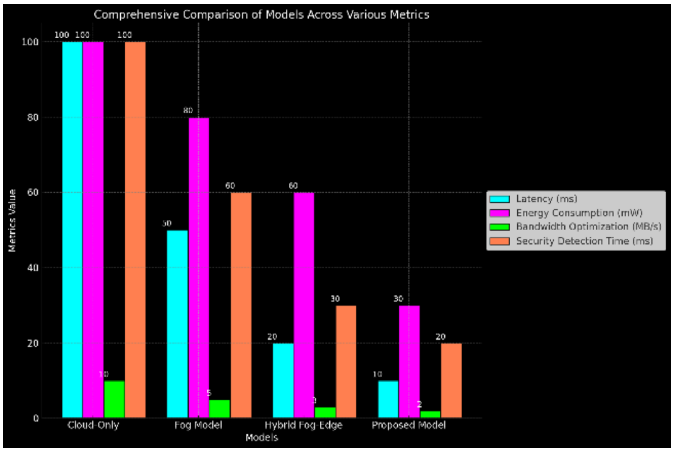


**Fig. 16**. Comprehensive comparison of models across various metrics.

them in end-to-end latency, real-time anomaly detection speed, and energy efficiency, particularly in emergency healthcare scenarios.

The Fig. 16 illustrates a detailed comparison across four models: The models examined are Cloud-Only, Fog, Hybrid Fog-Edge, and the one I propose. The comparison is based on four crucial metrics: The metrics are Latency (ms), Energy Consumption (mW), Bandwidth Optimization (MB/s), and Security Detection Time (ms). The model suggested earned the highest overall score, and was seen to outperform all other models analyzed in such measures as Bandwidth Optimization and Security Detection Time – in which the model with the lowest value was preferred. The Latency and Energy Consumption also depicted a steep descent to the Hybrid and the Proposed models.

Key Observations:

Cloud-Only: Very poor in terms of latency and energy without ability of bandwidth optimization.

Fog Model: Still slightly higher latency compared to Hybrid and Proposed models, however, latency is significantly reduced here. Energy leverage expanded and edged down a bit but still was considered high.

Hybrid Fog-Edge: Important gains in the latency, power use rate and in time required to detect security threats.

Proposed Model: Further analysis shows that the proposed algorithm has the best performance across all considered metrics, especially for bandwidth optimization and security detection.

The table recalls the performance measure of every model. The model suggested in this paper obtained the lowest values on latency and energy, while increasing bandwidth consumption. It also defined the role of delivering the fastest time to detect security threats. From this broad comparison, we establish that the proposed architecture excels in others in sparse IoT healthcare systems.

New advancements highlighted by the Hybrid Fog-Edge Computing framework for IoMT are latency, energy, bandwidth, scalability, speed, and security threat advancements. The decentralized architecture minimizes delays that accrue in passes, and optimizes the utilization of assets, making it suitable for real time use in healthcare. The hybrid model outperforms the cloud only model and traditional fog models in terms of latency and security issues while addressing scalability for vast IoMT systems.

## Discussion

The evaluation of the proposed Hybrid Fog-Edge Computing architecture demonstrates significant enhancements in the critical QoS parameters such as latency, energy consumption and bandwidth usage, and scale and security threats. The hybrid model often had higher response rates than those received by traditional cloud-only and fog models. Approximately, the latency was improved up to 70%, because the fog and edge layers processed critical data closer to the source, which allowed increasing the application response rate for monitoring the health condition in real time. The hybrid model is particularly fitting for energy-limited IoT devices because it has cut energy utilization by 30%. Data transmitted to the cloud was reduced by 60%; this made the bandwidth to be improved because most of the computation was in the edge and the fog layers. However, the scalability of the proposed hybrid architecture was also confirmed, and a network with up to 1500 IoT devices has better throughout and less latency. The hybrid model achieved better threat detection rates at 30ms and that is pivotal for the protection of sensitive health information. These results suggest that the proposed fog-edge hybrid architecture is very effective and efficient IoMT alternative; however, it has exceptionally high ability to provide real-time data processing and security as well as low energy consumption.

However, despite its advantages, implementing the proposed hybrid fog-edge architecture in real-world healthcare settings comes with notable limitations. High infrastructure costs are a significant concern, particularly when deploying fog and edge nodes across widespread or under-resourced regions. Maintaining consistent network connectivity and performing routine updates or retraining of ML models at the edge layer also require technical support and maintenance, which may not be feasible everywhere. Interoperability with legacy Electronic Health Record (EHR) systems poses another challenge, as existing hospital infrastructure often lacks standardized integration protocols. Moreover, regulatory frameworks like HIPAA and GDPR introduce compliance complexities when handling sensitive patient data across distributed nodes, necessitating strict governance and security enforcement mechanisms.

## Conclusion

Some of the features that are tailored in the Hybrid Fog-Edge Computing architecture for IoMT systems include; Low latency, energy efficiency, bandwidth, scalability, and security which are vital in Real-time health monitoring. By the help of the model and its design, latency was reduced by 70% while energy consumption improved by 30%, and at the same time cloud bandwidth usage reduced by 60% which will be appreciable for resource-constrained limited IoT devices. Several important conclusions relate to the work and focus on the model's ability to easily scale and accommodate up to 1500 IoT devices while having little to no impact on the overall performance. Hybrid model achieved the goal of anticipating security threats efficiently: the response time to an anomaly was 30 ms, proving that the model is beneficial for real-time protection of health data. However, it has its drawbacks: the potential for excessive costs related to the management of distributed fog and edge layers, as well as the fact that the functioning of the system depends on stable local computations. As a result, it might be beneficial for subsequent studies to evaluate how deeper incorporation of edge AI will eventually alleviate reliance on cloud by incorporating superior, predictive models at the edge in decision making mechanisms. Moreover, the support of the system might be expanded to the more detailed health monitoring tasks such as chronic diseases prediction. The future trend should concern possible issues connected with the data consistency across distributed layers and the improvement of the system's fault tolerance in case of failure at the hardware or the network level. More specifically, the choice of the Hybrid Fog-Edge Computing architecture enables a robust and scalable design for IoMT systems that presents significant improvements concerning performance, security, and energy consumption, which are critical assets in real-time health monitoring.

## Data availability

The dataset used in this study is publically available on kaggle: https://www.kaggle.com/datasets/faisalmalik/iot-healthcare-security-dataset.

Received: 12 March 2025; Accepted: 30 June 2025
Published online: 15 July 2025

## References

1. Khamael, H. et al. A review of fog computing and machine learning: concepts, applications, challenges, and open issues. *IEEE Access*. **7**, 153123–153140 (2019).
2. Aguru, A. D., Babu, E. S. & Nayak, S. R. Aswini Sethy, and Avinash Verma. Integrated industrial reference architecture for smart healthcare in the internet of things: A systematic investigation. *Algorithms* **15**, 309. 21 (2022).
3. Anayat Alam, S., Qazi, N., Iqbal & Raza, K. Fog, edge and pervasive computing in intelligent internet of things driven applications in healthcare: challenges, limitations and future use. In Deepak Gupta and Abhishek Khamparia, Editors, Fog, Edge, and Pervasive Computing in Intelligent IoT Driven Applications, 1–26. John Wiley Sons, Ltd, (2020).
4. Hanan, A., Alharbi & Muhammad Aldossary. Energy-efficient edge-fog-cloud architecture for iot-based smart agriculture environment. *IEEE Access*. **9**, 110480–110492 (2021).
5. Arcas, G. I., Cioara, T., Anghel, I., Lazea, D. & Hangan, A. Edge offloading in smart grid. *Smart Cities*. **7**, 680–711 (2024).
6. Zeshan Ashfaq, A. et al. A review of enabling technologies for internet of medical things (iomt) ecosystem. *Ain Shams Eng. J.* **13** (4), 101660 (2022).
7. Atef Taha Atieh. The next generation cloud technologies: a review on distributed cloud, fog and edge computing and their opportunities and challenges. *Res. Berg Rev. Sci. Technol.* **1** (1), 1–15 (2021).
8. Awaisi, K. S., Hussain, S., Ahmed, M., Khan, A. A. & Ahmed, G. Leveraging Iot and fog computing in healthcare systems. *IEEE Internet Things Magazine*. **3** (2), 52–56 (2020).
9. Elhadj Badidi, Z., Mahrez & Sabir, E. Fog computing for smart cities' big data management and analytics: a review. *Future Internet*. **12** (11), 190 (2020).

10. Jain, R., Gupta, M., Nayyar, A. & Sharma, N. Adoption of fog computing in healthcare 4.0. In *Fog Computing for Healthcare 4.0 Environments, Signals and Communication Technology* (ed. Sudeep Tanwar) 3–36 (Springer, 2021).
11. Mouad Laroui, B. et al. Edge and fog computing for iot: A survey on current research activities & future directions. *Comput. Commun.* **180**, 210–231 (2021).
12. Abhishek Mukherjee, S., Ghosh, A., Behere, S. S., Ghosh & Buyya, R. Internet of health things (ioht) for personalized health care using integrated edge-fog-cloud network. *J. Ambient Intell. Humaniz. Comput.* **12** (1), 943–959 (2021).
13. Muhammad Humayun, A., Alsirhani, F., Alserhani, M., Shaheen & Alwakid, G. Transformative synergy: Ssehcet—bridging mobile edge computing and Ai for enhanced ehealth security and efficiency. *J. Cloud Comput.* **13**, 37 (2024).
14. Neeraj Kumar Singh and Ashok Kumar Das. Energy-efficient fuzzy data offloading for Iomt. *Comput. Netw.* **213**, 109127 (2022).
15. Zhang, J., Ouda, A. & Abu-Rukba, R. Authentication and key agreement protocol in hybrid edge–fog–cloud computing enhanced by 5 g networks. *Future Internet*. **16**, 209 (2024).
16. Liu, H., Zhou, Y., Fang, B., Ning, Y. S. & Tian, Z. Hu, and PHCG: PLC honeypoint communication generator for industrial IoT. *IEEE Trans. Mob. Comput.* **24** (**1**), pp. 198-209 (2024).
17. Liu, Y. et al. A semi-centralized trust management model based on blockchain for data exchange in Iot system. *IEEE Trans. Serv. Comput.* **16** (2), 858–871 (2022).
18. Lv, Y., Shi, W., Zhang, W., Lu, H. & Tian, Z. Do not trust the clouds easily: the insecurity of content security policy based on object storage. *IEEE Internet Things J.* **10** (12), 10462–10470 (2023).
19. Bansal, S., Aggarwal, M. & Aggarwal, H. Advancements and applications in fog computing. In Duc-Nghia Le, Chintan Bhatt, and Mukesh Madhukar, editors, Security Designs for the Cloud, IoT, and Social Networking, pages 207–240. John Wiley Sons, Ltd, (2019).
20. Cogniteq Internet of medical things (iomt): Innovative future for healthcare. Online, Available: https://www.cogniteq.com/blog/internet-medicalthings-iomt-innovative-future-healthcare-industry (2023).
21. Burhan, U., Demirel, I. A., Bayoumy & Mohammad Abdullah Al Faruque. Energy-efficient real-time heart monitoring on edge–fog–cloud internet of medical things. *IEEE Internet Things J.* **9** (14), 12472–12481 (2022).
22. Ritu Dwivedi, D., Mehrotra & Chandra, S. Potential of internet of medical things (iomt) applications in Building a smart healthcare system: A systematic review. *J. Oral Biology Craniofac. Res.* **12** (2), 302–318 (2022).
23. Ghosh, S. & Mukherjee, A. Strove: Spatial data infrastructure enabled cloud–fog–edge computing framework for combating covid-19 pandemic. *Innov. Syst. Softw. Eng.* **20**(4), 727–743 (2022).
24. Hernandez-Jaimes, M. L. & Martinez-Cruz, A. Ram´ ırez-Guti´errez, and Claudia Feregrino-Uribe. Artificial intelligence for Iomt security: A review of intrusion detection systems, attacks, datasets and cloud–fog–edge architectures. *Internet Things*. **23**, 100887 (2023).
25. Jaspreet Kaur, R. et al. Importance of fog computing in healthcare 4.0. In *Fog Computing for Healthcare 4.0 Environments, Signals and Communication Technology* (eds Sudeep Tanwar et al.) 79–101 (Springer, 2021).
26. Qiong Luo, S. et al. Resource scheduling in edge computing: A survey. *IEEE Commun. Surv. Tutorials*. **23** (4), 2131–2165 (2021).


## Acknowledgements

This work is funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the support UID/05105: REMIT – Investigação em Economia, Gestão e Tecnologias da Informação, and RIF grant 23200 of Zayed University, UAE.

## Author contributions

Umar Islam (Manuscript Writeup and Analysis), Mohammed Naif Alatawi (Data Acquisitions and visualization), Ali Alqazzaz (Resources and Funding), Sulaiman Alamro (Data visualization and formal analysis), Babar Shah (Management and review ), Fernando Moreira (Supervision and Funding).

## Funding

This work is funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the support UID/05105: REMIT – Investigação em Economia, Gestão e Tecnologias da Informação.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to F.M.