Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 6, December 2026, pp. 5486~5499
ISSN: 2302-9285, DOI: 10.11591/eei.v15i6.13705

Clinical risk-aware routing for secure IoMT data in edge-fog-cloud environments

Rayane ElRaba'a, Soha Rawas, Ali El-Zaart
Department of Mathematics and Computer Science, Faculty of Science, Beirut Arab University, Beirut, Lebanon

Article Info

Article history:

Received Jun 17, 2026
Revised Sep 14, 2026
Accepted Oct 4, 2026

Keywords:

Clinical risk-aware routing
Edge-fog-cloud computing
Elastic clinical penalty
Internet of medical things
Patient safety

ABSTRACT

The internet of medical things (IoMT) enables continuous, real-time monitoring of patients through wearable and bedside sensors, but conventional edge-fog-cloud routing strategies optimize only network-centric quality of service (QoS) metrics and ignore the clinical urgency of the transmitted data, creating a safety risk for high-risk patients. This paper proposes fog-aware patient-network scoring (FA-PNS), a continuous, clinical risk-aware routing framework that embeds a normalized patient risk score directly into the routing utility function through an elastic clinical penalty mechanism, allowing routing sensitivity to scale smoothly with patient risk rather than relying on fixed thresholds. FA-PNS is evaluated on a synthetic dataset of 10,000 IoMT patient instances against five baseline routing strategies (latency-based, shortest-path, congestion-aware, weighted multi-metric, and priority-based). Results show that FA-PNS reduces the unsafe routing rate for high-risk patients from 93.33% (latency-based routing) to 3.33%, while improving high-risk patient end-to-end latency by 15.01% compared with priority-based routing (68.91 ms versus 81.08 ms), with all improvements statistically significant (paired t-test, $p<0.01$). These findings demonstrate that clinical safety and network efficiency can be jointly optimized without a substantial performance trade-off, offering a practical foundation for risk-aware routing in next-generation IoMT healthcare networks.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Rayane ElRaba'a
Department of Mathematics and Computer Science, Faculty of Science, Beirut Arab University
Beirut, Lebanon
Email: rayan.rabaa@bau.edu.lb

1. INTRODUCTION

The internet of medical things (IoMT) has rapidly become a cornerstone of modern digital healthcare, enabling real-time, secure and adaptive patient monitoring via wearable and bedside sensors [1], [2]. Yet, these devices generate diverse physiological data streams, such as heart rate (HR), peripheral oxygen saturation (SpO_2), and electrocardiogram (ECG), that require both high transmission speed and processing reliability [3], [4]. IoMT deployments increasingly relied on distributed edge-fog-cloud architectures [4]. Edge nodes offered low latency but lacked computational reliability for high-risk cases [5]. Cloud provided reliability but introduced significant delays [6]. This created a trade-off where speed-optimized routing could compromise patient safety. A key gap in the existing literature is that most routing protocols treat medical data as generic network traffic and optimize only network-centric quality of service (QoS) metrics such as latency, throughput, and energy, while ignoring the clinical urgency of the data being transmitted [7]–[10]. To address this gap, we propose the FA-PNS framework, a continuous, clinically informed routing approach that dynamically adjusts the trade-off between latency and computational

reliability according to real-time patient risk [11]. Edge nodes had limited processing; cloud introduced >70 ms latency [12]. Fog nodes bridged this gap with moderate latency and higher reliability [13].

Routing in IoMT has attracted significant attention. Existing studies primarily optimize network-centric QoS metrics (latency, throughput, energy) but lack clinical awareness, as summarized in Table 1. Initial research on genetic algorithms exhibited multi-objective optimization capabilities for secure routing. Reliable routing using multiobjective genetic algorithm (RRMOGA) [14] achieved a 99.9% packet delivery ratio (PDR) and a 14% longer network lifetime. Embedding security risk levels into routing decisions is the approach taken by Aravind and Maddikunta [15] or probabilistic risk estimates by Cui *et al.* [16], yet, these methods are still treating medical data as regular data traffic.

Table 1. State of the art review

Ref	Methodology/proposed model	Key results/findings
[14]	RRMOGA: multi-objective genetic algorithm.	99.9% PDR; 14% longer network lifetime.
[15]	SA-JSO and ANFIS: geographic routing based with security risk.	Enhanced path optimality under clinical risk.
[16]	Uncertainty-aware risk-aware routing for emergency vehicles.	Multi-objective risk optimization.
[17]	WFQ and priority clustering: edge-IoMT.	40% delay reduction, 19.47% jitter improvement.
[18]	QoS multi-path routing: emergency classification.	High throughput, lower packet drops.
[19]	QoS-gateway: dynamic load balancing.	25% lower latency, 18% higher PDR.
[20]	CHBO: chaotic honey badger optimization.	Lower routing load, high PDR.
[21]	Adaptive AFL and hybrid PSO-GA.	92.5% PDR, 0.10s delay, 35% energy savings.
[22]	GWO-MARP: grey wolf optimization.	99.6% PDR, 1.175 ms delay, 44.30 kbps throughput.
[23]	EEMSRP: lattice-based cryptography.	30% energy reduction, 25% longer lifetime.
[24]	Ensemble ML for VANET security.	High malicious activity detection.
[25]	Heuristic-assisted feature fusion for IoT IDS.	96% accuracy, 40% feature reduction.
[26]	Hybrid ABE with blockchain.	35% lower decryption latency.
[27]	HRADMT: hybrid attack detection.	96.18% detection accuracy.
[28]	LPC-SNN-PCOA: blockchain+spiking neural network.	99.8% PDR, 0.005 J energy, privacy.
[29]	DA-DRL+blockchain for IoMT security.	Real-time AES key generation, intrusion detection.
[30]	Deep learning IDS (EmbedNet, ConvNet-SVM, DeepSVM-Net).	High detection accuracy, computational efficiency.
[31]	Multi-attention DeepCRNN+blockchain.	3.44% accuracy and 3.71% F1-score improvement.
[11]	Hybrid fog-edge: decentralized filtering.	70% latency reduction, 60% bandwidth savings.
[32]	SDN-enabled task allocation in 5G IoMT.	Reduced latency, deadline misses, energy consumption.
[33]	CV-STGCN for solar energy prediction.	23% lower prediction error than RNNs.

Note: Simulated annealing-jellyfish search optimizer (SA-JSO); adaptive neuro-fuzzy inference system (ANFIS); weighted fair queueing (WFQ); adaptive fuzzy logic (AFL); particle swarm optimization-genetic algorithm (PSO-GA); grey wolf optimization-based multi-adaptive routing protocol (GWO-MARP); energy-saving, layered, and safe routing protocol (EEMSRP); vehicular ad-hoc network (VANET); internet of things (IoT); intrusion detection system (IDS); attribute-based encryption (ABE); hybrid reliable adaptive data movement topology (HRADMT); linear predictive coding-spiking neural network-principal coordinates analysis (LPC-SNN-PCOA); domain-adversarial deep reinforcement learning (DA-DRL); embedding network (EmbedNet); convolutional network-support vector machine (ConvNet-SVM); deep support vector machine network (DeepSVM-Net); deep convolutional recurrent neural network (DeepCRNN); software-defined networking (SDN); advanced encryption standard (AES); cross-view spatial-temporal graph convolutional network (CV-STGCN); and recurrent neural networks (RNNs).

Several priority-based edge schemes [17] were implemented to cutting delay-like a 40% reduction [17] in the example-and also to increase throughput. On the downside, these schemes are using preset thresholds that do not consider changes in patient risk or network congestion [18], [19]. And, metaheuristic routing has caught attention in the research. CHBO [20] increases PDR and decrease the end-to-end delay. AFL combined with hybrid packet delivery ratio PSO-GA [21] is able to show 92.5% PDR, 0.10 s delay, and 35% energy saving. Along the same lines, GWO-MARP [22] reports 99.6% PDR and 1.175 ms delay. What is more, these techniques have not included patient-specific clinical risk scores which is a missed opportunity.

Security has been unstoppable issue. EEMSRP [23] enable post-quantum secure routing with 30% energy saving. Among the other features, ensemble machine learning (ML) [24], feature fuse [25], ABE combined with blockchain [26], attack detection [27], and privacy-preserving routing [28]. The new deep learning and blockchain architectures [29]–[31] give IoMT security a next level-e.g., multi-attention DeepCRNN increases accuracy by 3.44% and F1-score by 3.71% even beating state-of-the-art methods. These attempts to increase safety against attacks do not take clinical safety into account. On one hand, hybrid fog-edge [11], [32] have proved 70% latency reduction and better load balancing. However, predictive routing using CV-STGCN [33] can be taken as an effective way for congestion avoidance but it still does not have a clinical risk factor.

Despite these advances, existing IoMT routing methods largely ignore patient risk and clinical severity. FA-PNS is therefore proposed as the first unified, clinically informed routing framework for edge-fog-cloud IoMT environments.

The primary objective of this research is to design and evaluate a clinical risk-aware routing framework that jointly optimizes network performance and patient safety in edge-fog-cloud IoMT environments. Specifically, this study aims to: i) formulate a continuous, clinically informed utility function that embeds real-time patient risk into the routing decision; ii) develop an elastic clinical penalty mechanism that adaptively increases the cost of delay for high-risk patients; and iii) empirically evaluate the proposed patient-network scoring (FA-PNS) framework against conventional QoS-centric and priority-based routing baselines in terms of safety, latency, and clinical utility.

Unlike prior work that uses only QoS metrics [14], [15], [17], [18], static thresholds [19], or metaheuristics [20], [21], [23], FA-PNS introduces a continuous elastic clinical penalty, a three-tier utility function, and dynamic layer arbitration. This paper makes the following contributions: i) a risk-aware model that integrates patient medical risk directly into routing decisions, ii) a clinically informed routing that balances medical urgency, processing capability, queue congestion, and transmission latency, and iii) an evaluation framework for safety trade-offs under heterogeneous network conditions. The rest of this paper is organized as follows: section 2 describes the methodology, section 3 presents results and discussion, and section 4 concludes.

2. METHOD

This section describes the system architecture, IoMT dataset construction, the proposed FA-PNS framework and the underlying risk aware decision logic.

2.1. Internet of medical things architecture model

The considered system adopts a three-tier edge–fog–cloud architecture. Edge nodes generally offer the quickest access due to latency being low but they are limited on power to carry out complex processes. Fog nodes give out an amount of computation that is almost between and with a moderate delay. Whereas cloud servers are very reliable but the latency increases. Figure 1. presents the three-tier architecture and decision-making flow. IoMT sensors (wearables, bedside monitors) that emanate patient data streams are the first ones to encounter edge gateways (like local routers or smartphone gateways). The FA PNS decision engine, located at the fog layer, analyzes each packet's risk score (R_i), and the present state of the network (queue length, transmission delay) to decide the best target location: edge (when risk is low and congestion is minimal), fog (for an intermediate level of risk or moderate congestion), or cloud (for high risk or when fog resources have been fully utilized). The decision-making spots are represented with the dashed evaluation boxes.

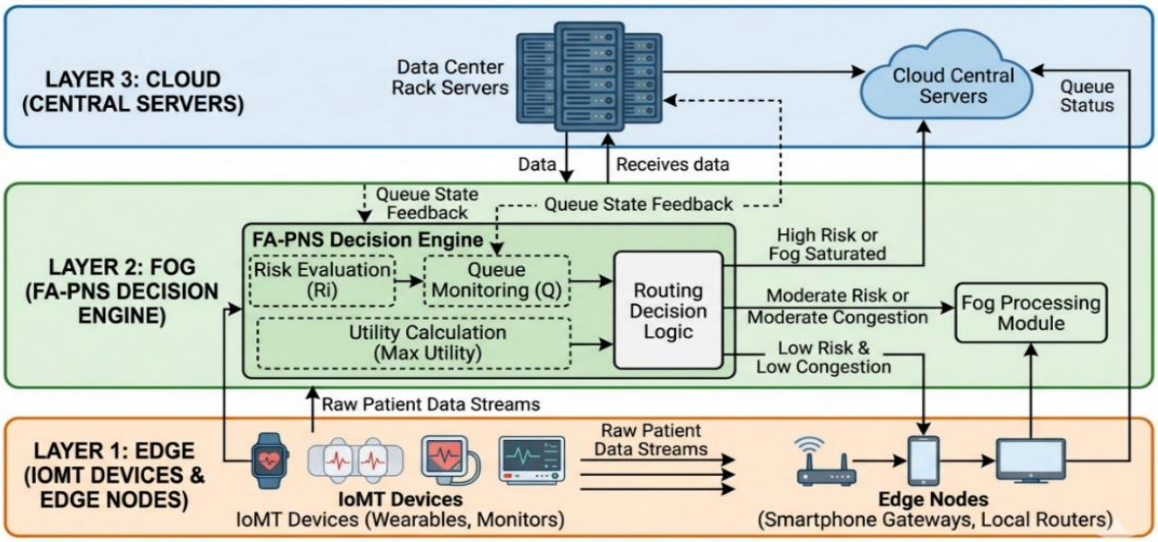


Figure 1. Three-tier IoMT architecture and data flow (FA-PNS system)

2.2. Synthetic internet of medical things dataset generation

A synthetic dataset of 10,000 IoMT patient instances has been utilized for framework efficiency evaluation while taking into consideration privacy and ethical constraints. Each patient is assigned a

normalized medical risk score (R), which is between 0 and 1, R is a composite feature derived from critical physiological indicators like HR fluctuation, blood oxygen saturation SpO_2, and the presence of ECG anomalies as shown in Table 2. Patients with R>0.5 indicates high-risk, requiring reliable processing [34].

The clinical ranges refer to the standards that the medical community has agreed upon: HRs from 40-180 bpm [35], temperature thresholds for fever/hypothermia [36], SpO_2<90% for hypoxemia [37], and blood pressure (BP) ranges for hypertension and shock [36]. A 15% chance of ECG abnormality is assumed [38]. The concept of data freshness (age of information) indicates that the medical relevance of data decreases with time [39]. Since no public dataset exists combining medical measurements, risk levels, and routing attributes, the synthetic IoMT data were generated. Validation is planned using Medical Information Mart for Intensive Care IV (MIMIC-IV) and hospital trials.

Table 2. Synthetic IoMT dataset key attributes and clinical boundary constraints

Vital sign	Simulation logic (distribution)	Operational range	Clinical condition
HR	Average 85 bpm (SD±20)	40–180 bpm	Tachycardia/bradycardia
SpO_2	Average 94% (SD±3)	70–100%	Hypoxia monitoring
Temperature	Average 37.2 °C (SD±0.7)	35–41 °C	Fever/hypothermia
Systolic BP	Average 125 mmHg (SD±20)	80–220 mmHg	Hypertension/shock
ECG anomaly	15% chance of abnormality	0 (no)/1 (yes)	Cardiac arrhythmia
Data freshness	Random age between 0.1–1.0 s	0.1–1.0 (age)	Data timeliness

2.2.1. Formal definition of the clinical risk score

The normalized medical risk score $R_i \in [0,1]$ for patient instance i is computed as weighted sum of five clinically relevant vital signs. each first normalized to a [0,1] scale using established clinical thresholds. Let $v(i,k)$ denote the raw value of vital sign k for patient i, and let min_k, max_k be the operational bounds:

$$\tilde{v}_{(i,k)} = \frac{v_{i,k} - min_k}{max_k - min_k} \quad (1)$$

For ECG anomaly (binary), $\tilde{v}_{(i,ECG)}$=1 if anomaly present, else 0. The risk score is then:

$$R_i = \sum_{K=1}^{5} w_k \cdot \tilde{v}_{(i,k)}, \text{ with } \sum_{k=1}^{5} w_k = 1 \quad (2)$$

Based on clinical literature [34]–[38], the weights are set as: HR (0.25), SpO_2 (0.30), temperature (0.10), systolic BP (0.20), and ECG anomaly (0.15). SpO_2 receives the highest weight due to its strong association with acute deterioration. A threshold of Ri>0.5 classifies high-risk patients, consistent with mid-point risk stratification used in early warning scores, e.g., National Early Warning Score 2 (NEWS2).

2.3. Development of fog-assisted patient network selection framework

Figure 2 illustrates the methodology. Two parallel tracks feed the FA-PNS engine: synthetic dataset generation with clinical risk scoring, and network state monitoring (queue and latency). For each packet, the engine computes the utility based on patient risk, processing delay, and network latency, then selects the optimal edge/fog/cloud node. This fuses clinical and network data at decision time.

2.4. Mathematical formulation for fog-assisted patient network selection routing model

Traditionally, multi-metric routing relied on static weighted-sum models [40]. FA-PNS jointly considers clinical priority, computational complexity and queue induced delay within a unified framework. First, clinical priority, let $R_i \in [0,1]$ be the normalized medical risk score for patient i, as defined in (2) where $R_i = 0$ indicates no risk (healthy) and $R_i = 1$ indicates critical condition.

The queue delay at node n is modeled as stochastic congestion process of L_n in accordance to the result of difference between service rate μ_n and arrival rate λ_n.

$$Q_{i,n} = \frac{L_n}{\mu_n - \lambda_n}, \text{ where } \mu_n > \lambda_n \ for\ stability \quad (3)$$

As for computational complexity, it is modeled as a composite processing delay calculated by reflecting computational burden incurred when an IoMT packet is processed at node n (edge, fog, and cloud). Processing delay $D_{i,n}$ accounts for queue delay Q_i, clinical risk R_i, node complexity C_n, and data freshness F_i. Higher risk increases processing load, especially on resource-constrained edge nodes, while cloud nodes mitigate this. Stale data (F_i low) also increases delay. For packet ii at node n:

$$D_{i,n} = \beta_1 Q_i + \beta_2 R_i \cdot C_n + \beta_3 (1 - F_i) \quad (4)$$

where C_n is the node complexity and $\beta_1 + \beta_2 + \beta_3 = 1$.

Weighted sum models are frequently used in multi-criteria decision-making and network optimization capacity to represent proportional influence among independent factors without introducing nonlinear dominance effects [41]. In the resulting formulations, priority R_i is used as a dominant term, and queue delay $L_{i,n}$ and complexity $D_{i,n}$ directly influence routing decision through the FA-PNS function:

$$U_{i,n} = w_1 R_i + w_2 \frac{1}{D_{i,n}} + w_3 \frac{1}{L_{i,n}} \quad (5)$$

The routing decision or optimal node is then:

$$n^* = \arg \max_n U_{i,n} \quad (6)$$

The inverse formulations $\frac{1}{D_{i,n}}$ and $\frac{1}{L_{i,n}}$ are used to convert minimization goals into utility components that are consistent with maximizing. FA-PNS dynamically reflects processing constraints while preserving clinical relevance.

In plain terms, (5) favors nodes that combine high clinical priority with low processing delay and low queueing delay, so that, for each packet, FA-PNS selects the node offering the best overall balance between patient urgency and network responsiveness.

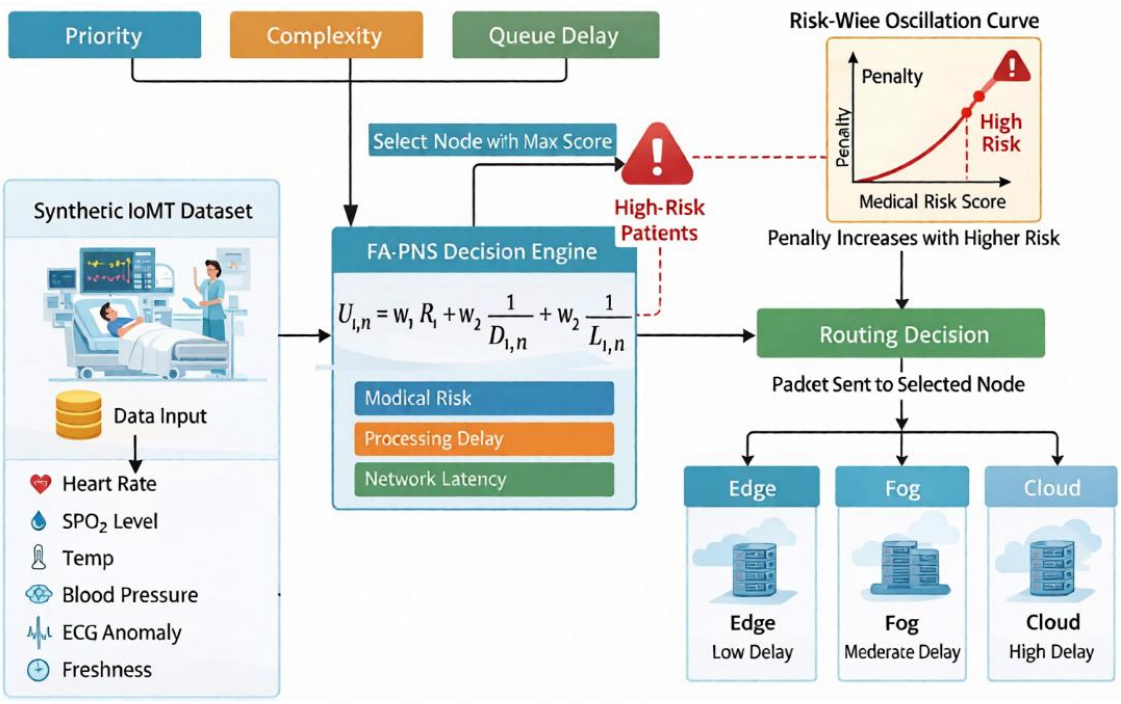


Figure 2. Proposed method for FA-PNS

2.5. Fog-assisted patient network selection routing decision workflow

Rather than selecting routes solely based on network factors like delay or queue length, the FA-PNS algorithm considers clinical context when deciding on routes. The workflow, see Figure 3, starts with the calculation of the risk score. Delays in processing and queueing are predicted for each candidate node. Next, the utility function identifies the one with the highest value. FA-PNS, compared to static priority schemes, re-assesses each packet because of this it can adjust to the changing patient condition or network congestion.

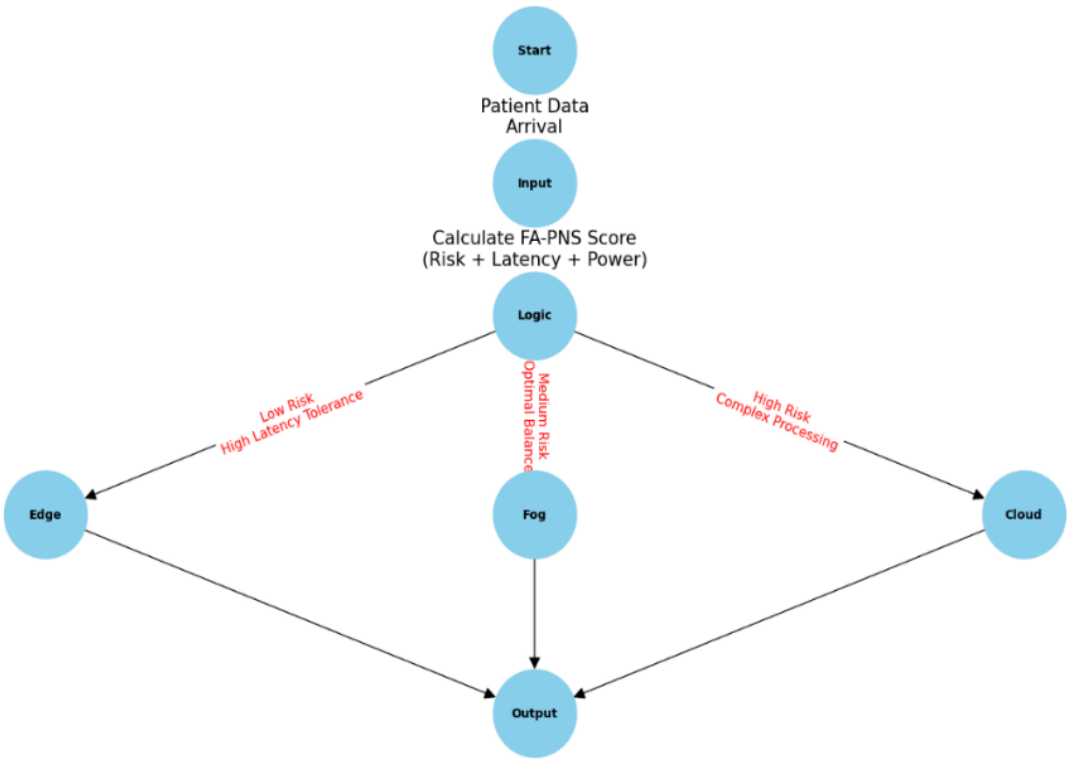


Figure 3. FA-PNS routing decision workflow

Figure 4 illustrates how FA-PNS continuously shifts traffic from edge to fog (peaking at 0.4–0.7) and cloud (dominating above 0.8) as medical risk increases, adapting to clinical urgency without hard thresholds.

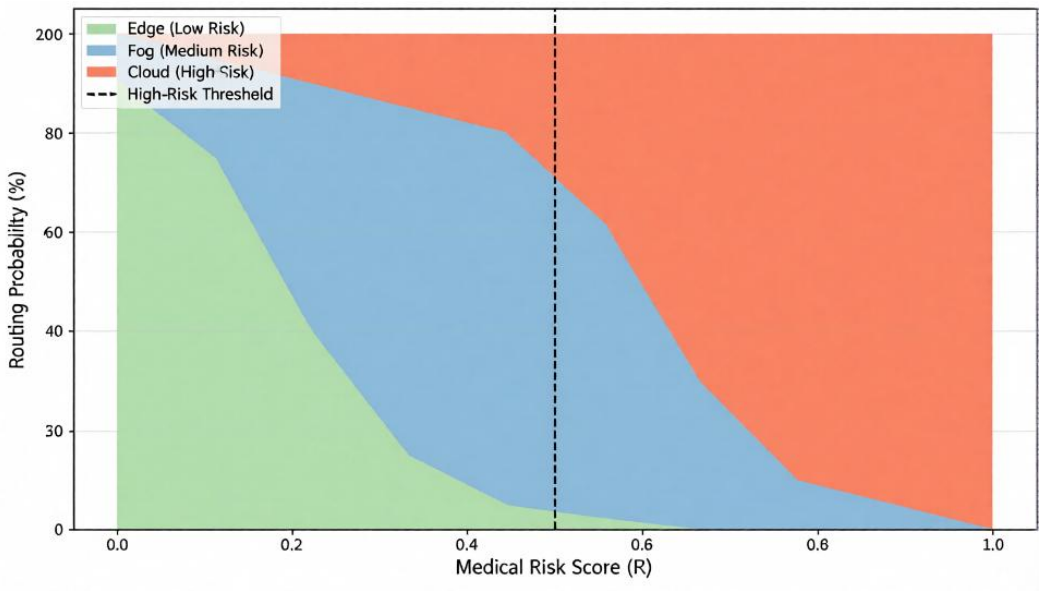


Figure 4. FA-PNS decision logic: node selection vs. medical risk level

Algorithm 1. FA-PNS routing decision

Input: Patient data packet i with vital signs $v_{(i,k)}$

Output: Selected node n* ∈ {Edge, Fog, Cloud}

1: Compute normalized vital signs $\tilde{v}_{(i,k)}$
2: Compute risk score R_i (weights from Sec 2.2.1)
3: For each node n in {Edge, Fog, Cloud}:
    4: Obtain current queue length L_n and service rate μ_n, arrival λ_n
    5: Compute $Q_{i,n}$
    6: Retrieve node complexity C_n and packet freshness F_i
    7: Compute $D_{i,n}$
    8: Compute utility $U_{i,n}$
9: End for
10: calculate routing decision n^*
11: Route packet i to n^*
12: Update queue length L_{n*} (increment by 1)

2.6. Elastic clinical penalty mechanism

The elastic clinical penalty adjusts the routing utility's sensitivity to delay per the risk level of the patient. Going beyond a binary penalty (which is given only above a threshold), we introduce a continuous penalty factor $P(R_i)$ which multiplicatively modifies the delay components of the utility function:

$$P(R_i) = 1 + \gamma \cdot \left(\frac{R_i - 0.5}{0.5}\right)^3, \text{for } R_i > 0.5 \quad (7)$$

where $P(R_i)$=for $R_i \leq 0.5$, where γ=2.0 is a tunable sensitivity parameter. The cubic exponent ensures that penalty grows slowly near the threshold but accelerates for very high-risk patients (e.g., at R_i=0.9, penalty≈2.0, doubling the effective cost of delay).
This penalty is embedded into the processing delay term of (4):

$$D_{i,n}^{penalized} = P(R_i) \cdot D_{i,n} \quad (8)$$

Consequently, the utility function becomes:

$$U_{i,n} = w_i R_i + w_2 \frac{1}{P(R_i) \cdot D_{i,n}} + w_3 \frac{1}{L_{i,n}} \quad (9)$$

Figure 5 visualizes the penalty curve as a function of R_i. The cubic penalty remains 1.0 for low/medium risk then grows slowly until R≈0.7, after which it accelerates sharply, reaching 2.0 at R=0.9.

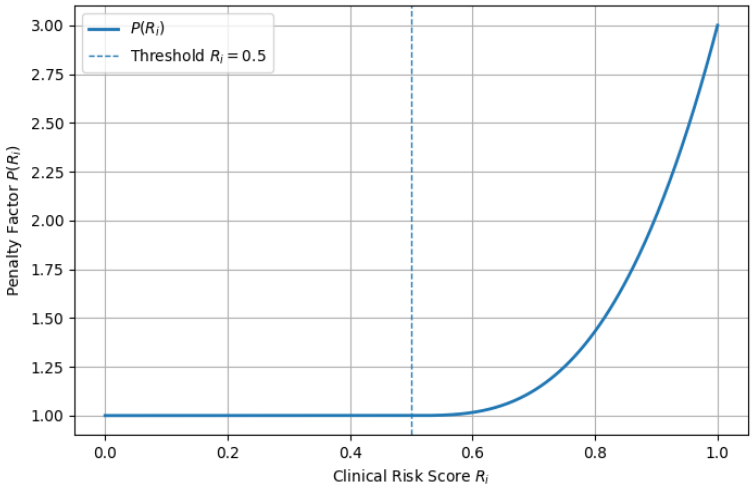


Figure 5. Elastic clinical penalty curve

3. RESULTS AND DISCUSSION

In this section, a thorough experimental evaluation of the proposed clinically conscious, risk-adaptive routing system is performed in an Edge-Fog-Cloud IoMT scenario. The assessment aims at not only

the clinical safety implications but also the system-level performance with a special emphasis on how predictive risk awareness changes routing decisions under realistic network constraints. The research concentrates mainly on the system's ability to reduce risky routing of high-risk situations, shorten end-to-end latency for critical data, and increase overall productivity against conventional routing methods. The analysis starts with a repeatable experimental setup, followed by quantitative performance comparisons, risk analysis with a focus on safety, and ultimately, an assessment of clinical utility in emergency medical conditions.

3.1. Experimental setup and reproducibility

All simulations were run in Google Colaboratory (Python 3) on a synthetic dataset of 10,000 IoMT patient instances. Five baselines were implemented:

- Latency-based routing: selects the node with minimum network transmission delay (favors edge).
- Shortest-path routing: minimizes hop count (similar to latency-based).
- Congestion-aware routing: selects the node with the shortest queue.
- Weighted multi-metric routing: fixed-weight combination of latency and congestion.
- Priority-based routing: sends all high-risk packets to the cloud regardless of network state.

3.2. Evaluation metrics

The following metrics were used (all baselines and FA-PNS evaluated identically):

- End to end (E2E) delay was defined as the sum of network transmission latency and processing delay at the selected node:

$$D_{E2E} = D_{process}(n) + D_{network}(n) \tag{10}$$

- Unsafe routing rate: as the percentage of high-risk packets routed to edge nodes:

$$Unsafe\ routing\ rate = \frac{\text{high risk packets routed to Edge}}{\text{Total high risk packets}} \times 100 \tag{11}$$

- Clinical utility score: penalized delay by the unsafe routing rate F:

$$CU = D_{E2E} \times (1 + F) \tag{12}$$

- Improvement over baseline: quantify performance gains in network optimization and routing:

$$improvement\ \% = \frac{D_{process} - D_{FA-PNS}}{D_{baseline}} \times 100 \tag{13}$$

- Statistical significance: paired t-tests with confidence score of 95% were assessed.

3.3. High-risk safety analysis (unsafe routing rate)

Table 3. reveals that latency-based and shortest-path routing deliver 93.33% of high-risk packets to the edge. Priority routing pushes all critical data to the cloud (0% unsafe) but creates bottlenecks. FA-PNS achieves only 3.33% unsafe routing.

Table 3. Unsafe routing percentages

Algorithm	Unsafe routing (%)
FA_PNS	3.33
Latency	93.33
Priority	0.00
Shortest_Path	93.33
Congestion	41.67
Weighted	90.00

3.4. End-to-end delay and the efficiency gap

Table 4 presents the end-to-end delays. FA-PNS increases average latency by only 1.26 ms compared to latency-based method but reduces high-risk latency by 15.01% (68.91 ms vs 81.08 ms for priority routing). The cumulative distribution, see Figure 6, shows that FA-PNS has a steeper, left-shifted curve for high-risk packets, indicating lower variance and fewer long delays clinically preferable for time-critical data.

Table 4. Average end to end delays (ms) with standard deviation

Algorithm	Average E2E delays	
	All packets (mean±std)	High-risk patients (mean±std)
FA-PNS	34.01±4.23	68.91±7.56
Latency	32.75±8.91	74.89±12.34
Priority	33.41±5.67	81.08±9.87
Shortest_Path	32.75±8.91	74.89±12.34
Congestion	53.15±14.22	76.93±11.45
Weighted	33.07±6.34	74.14±10.21

Note: High-risk patient count=60; Paired t-tests (95% confidence) between FA-PNS and each baseline for high-risk patient latency was statistically significant ($p<0.01$) in all comparisons

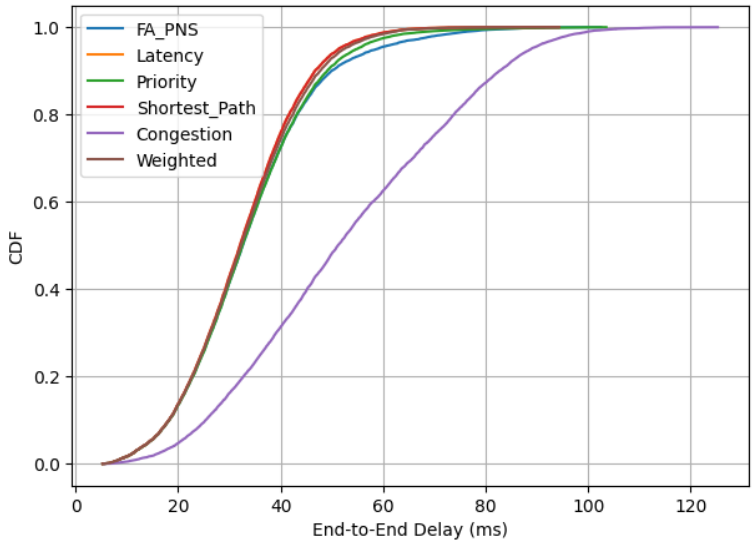


Figure 6. Cumulative distribution of end-to-end delay for all evaluated routing strategies

3.5. Improvement over baselines and clinical utility evaluation

Table 5 shows that FA-PNS improved high-risk patient latency by 15.01% over priority routing, 7.98% over latency-based and shortest-path routing, and 10.42% over congestion-aware routing. FA-PNS achieves the best trade-off: <5% unsafe rate and <70 ms high-risk latency, see Figure 7.

Table 5. FA-PNS improvement (%) compared to baseline methods

Algorithm	Improvement %
Latency	+7.9830
Priority	+15.0118
Shortest_Path	+7.9830
Congestion	+10.4194
Weighted	+7.04700

3.6. Comparison with related routing approaches

To contextualize the performance of FA-PNS relative to the broader IoMT routing literature, Table 6 qualitatively compares FA-PNS with representative metaheuristic-based routing methods (e.g., CHBO [20], GWO-MARP [22], hybrid PSO-GA [21]) and blockchain-based secure routing methods (e.g., ABE combined with blockchain [26], LPC-SNN-PCOA [28], DA-DRL with blockchain [29]).

Metaheuristic approaches such as GWO-MARP achieve very high packet delivery ratios (99.6%) and low delay (1.175 ms) by optimizing purely network-centric objectives, but, like the other QoS-only baselines evaluated in subsections 3.1-3.3, they do not distinguish high-risk from low-risk patients and would therefore be expected to exhibit an unsafe routing rate comparable to the latency-based and shortest-path baselines (93.33%) rather than the 3.33% achieved by FA-PNS. Blockchain-based methods such as ABE-blockchain [26] and LPC-SNN-PCOA [28] substantially improve data confidentiality and integrity (e.g., 35% lower decryption latency, 99.8% PDR) but add cryptographic processing overhead at each hop and, similarly, do not incorporate a clinical risk signal into the routing decision itself.

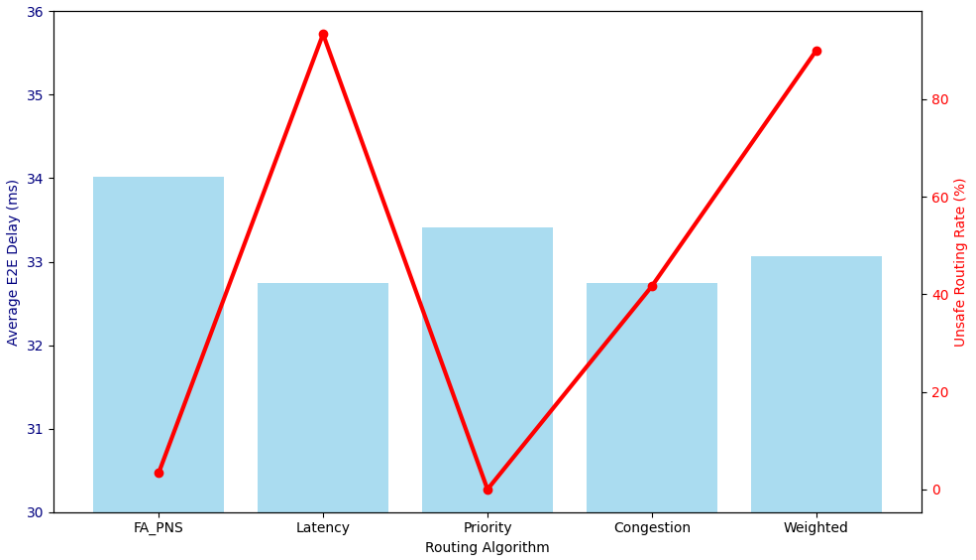


Figure 7. Safety–latency trade-off

Table 6. Qualitative comparison of FA-PNS with representative IoMT routing paradigms

Routing paradigm	Representative method(s)	Primary optimization focus	Clinical risk awareness	Reported performance highlights
QoS-centric (latency/shortest-path)	[14], [17]-[19]	Latency, throughput, packet delivery ratio	No	Up to 99.9% PDR; 40% delay reduction
Metaheuristic	CHBO [20]; GWO-MARP [22], hybrid PSO-GA [21]	Multi-objective network optimization	No	99.6% PDR; 1.175 ms delay; 35% energy savings
Blockchain/security-centric	ABE-blockchain [26]; LPC-SNN-PCOA [28]; DA-DRL+blockchain [29]	Data confidentiality and integrity	No	35% lower decryption latency; 99.8% PDR
Priority-based	Priority baseline (this paper)	Binary clinical urgency (threshold)	Partial (threshold-based)	0% unsafe routing but highest high-risk latency (81.08 ms)
FA-PNS (proposed)	This paper	Joint clinical risk+network efficiency	Yes (continuous)	3.33% unsafe routing; 15.01% latency improvement over priority routing

FA-PNS is therefore complementary rather than competing with these approaches: its elastic clinical penalty could, in principle, be layered on top of a metaheuristic path-selection engine or a blockchain-secured transport layer to combine network efficiency or security with clinical risk-awareness. A direct quantitative comparison of FA-PNS with these approaches under an identical experimental protocol is identified as an important direction for future work.

3.7. Time complexity and overhead analysis

Time complexity for each packet is O(N), where N=3 (edge fog cloud), which means practically it is O(1) for each packet. The overheads were measured: 14 more bytes/packet (2.7% for a 512-byte packet) and 0.12 ms Also processing time negligible when compared to the network transmission delays (1050 ms).

3.8. Limitations

Although the results obtained with FA-PNS are promising, several limitations must be acknowledged: i) synthetic data: the evaluation relies entirely on synthetic patient and network data and does not model real-world artefacts such as sensor noise, missing values, packet loss, or variable transmission delays; the reported percentages are therefore comparative rather than absolute, and no real-world clinical validation has yet been performed—calibration against real deployments would be required before clinical use; ii) fixed weights: the utility parameters were set empirically; optimal values could differ across hospital policies or patient populations, motivating reinforcement-learning-based weight adaptation as a next step; iii) scalability: the topology considered only a single, homogeneous fog layer; large-scale deployments may require hierarchical fog or multi-region cloud coordination, which would add complexity; iv) security: FA-PNS addresses clinical safety only and does not account for tampering or node compromise;

cryptographic safeguards would be required to protect against adversarial manipulation of the routing decision; and v) computational overhead: calculating the utility per packet incurs negligible delay (0.12 ms), but energy-constrained edge devices could still experience a non-trivial cumulative burden at high packet rates.

3.9. Practical deployment and policy relevance

In practice, FA-PNS is designed to run as a lightweight decision engine at the fog layer, sitting between hospital edge gateways (bedside monitors, wearable hubs) and cloud-based electronic health record systems. As each IoMT packet arrives, the engine computes the patient risk score and node utilities in real time (subsection 2.4) and forwards the packet to the edge, fog, or cloud tier accordingly, without requiring any change to the underlying sensors or network hardware. For a hospital or remote-monitoring provider, this means high-risk patients (e.g., patients with unstable vitals on a step-down unit, or those enrolled in a remote cardiac-monitoring program) automatically receive prioritized, more reliable routing paths, while stable patients continue to be served by low-latency edge paths, reducing unnecessary load on fog and cloud resources. Because the computational overhead per packet is negligible (0.12 ms, subsection 3.5), FA-PNS can be integrated into existing hospital IoMT gateways as a software-only update.

Beyond technical performance, FA-PNS is also relevant to medical-data governance. Its risk-stratified routing logic aligns with the intent of clinical data-protection regulations such as Health Insurance Portability and Accountability Act (HIPAA) and the European Union General Data Protection Regulation (EU GDPR), which require that health data be handled in a manner proportionate to its sensitivity and urgency, and it is compatible with IoMT interoperability standards such as Health Level Seven Fast Healthcare Interoperability Resources (HL7 FHIR) and the Institute of Electrical and Electronics Engineers (IEEE) 11073 for structuring the vital-sign data used to compute the risk score. Embedding a standards-compliant risk score into the routing layer offers healthcare providers a mechanism to demonstrate that time-critical patient data is handled under a documented, auditable prioritization policy, which is increasingly expected by hospital IT-governance and medical-device interoperability frameworks.

4. CONCLUSION

This paper presented FA-PNS, a clinical risk-aware adaptive routing framework for the IoMT that embeds a continuous patient risk score directly into routing decisions through an elastic clinical penalty mechanism. In terms of contributions, this work delivered: i) a risk-aware routing model that replaces purely network-centric routing criteria with clinical priority, ii) a clinically informed utility function that jointly balances medical urgency, processing capability, queue congestion, and transmission latency, and iii) an evaluation framework for quantifying safety-latency trade-offs under heterogeneous network conditions. Experiments on a synthetic dataset of 10,000 patient instances showed that FA-PNS reduced unsafe routing of high-risk patients from 93.33% (latency-based routing) to 3.33%, while improving high-risk patient latency by 15.01% relative to priority-based routing (68.91 ms versus 81.08 ms, $p<0.01$). These results indicate that clinical safety and network efficiency can be jointly optimised without a substantial performance penalty. Nonetheless, as discussed in the limitation's subsection, these findings are subject to important caveats, most notably the reliance on synthetic data, fixed utility weights, and the absence of cryptographic safeguards against tampering or node compromise; the reported improvements should therefore be interpreted as comparative evidence of feasibility rather than as absolute, clinically validated performance guarantees.

Future work will accordingly prioritize: i) real-clinical validation using MIMIC-IV and prospective hospital testbeds; ii) AI-based dynamic weight adaptation (e.g. deep Q-networks) informed by clinical-expert input; and iii) integration with 5G/6G network slicing and lightweight cryptographic protection to jointly support prioritization, security, and scalability of high-risk IoMT data. Pursuing these directions will help FA-PNS evolve from a validated research prototype into a deployable clinical decision-support component of next-generation IoMT networks.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Rayane ElRaba'a	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
Soha Rawas	✓	✓		✓	✓		✓			✓		✓	✓	
Ali El-Zaart				✓	✓					✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, RE, upon reasonable request.

REFERENCES

[1] P. Matthew *et al.*, "A Review of the State of the Art for the Internet of Medical Things," *Sci*, vol. 7, no. 2, p. 36, 2025, doi: 10.3390/sci7020036.

[2] P. Dubey, P. Dubey, and P. N. Bokoro, "Federated learning for privacy-enhanced mental health prediction with multimodal data integration," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, vol. 13, no. 1, 2025, doi: 10.1080/21681163.2025.2509672.

[3] S. F. Ahmed *et al.*, "Transformative impacts of the internet of medical things on modern healthcare," *Results in Engineering*, vol. 25, p. 103787, Mar. 2025, doi: 10.1016/j.rineng.2024.103787.

[4] R. Buenrostro-Mariscal, P. C. Santana-Mancilla, O. A. Montesinos-López, M. Vazquez-Briseno, and J. I. Nieto-Hipolito, "Prioritization-Driven Congestion Control in Networks for the Internet of Medical Things: A Cross-Layer Proposal," *Sensors*, vol. 23, no. 2, p. 923, 2023, doi: 10.3390/s23020923.

[5] S. I. Loutfi, I. Shayea, U. Tureli, A. A. El-Saleh, and W. Tashan, "An overview of mobility awareness with mobile edge computing over 6G network: Challenges and future research directions," *Results in Engineering*, vol. 23, p. 102601, 2024, doi: 10.1016/j.rineng.2024.102601.

[6] A. Rancea, I. Anghel, and T. Cioara, "Edge Computing in Healthcare: Innovations, Opportunities, and Challenges," *Future Internet*, vol. 16, no. 9, p. 329, 2024, doi: 10.3390/fi16090329.

[7] C. Huang, J. Wang, S. Wang, and Y. Zhang, "Internet of medical things: A systematic review," *Neurocomputing*, vol. 557, p. 126719, 2023, doi: 10.1016/j.neucom.2023.126719.

[8] S. H. Chiew *et al.*, "Multiobjective Optimization and Network Routing with Near-Term Quantum Computers," *IEEE Transactions on Quantum Engineering*, vol. 5, pp. 1–19, 2024, doi: 10.1109/TQE.2024.3386753.

[9] L. Ekstrom, H. Wang, and S. Schmitt, "Variational quantum multiobjective optimization," *Physical Review Research*, vol. 7, no. 2, p. 023141, 2025, doi: 10.1103/PhysRevResearch.7.023141.

[10] K. Deb, "Multi-Objective Optimization Using Evolutionary Algorithms: An Introduction," Wiley, Kanpur, Uttar Pradesh, 2011.

[11] U. Islam, M. N. Alatawi, A. Alqazzaz, S. Alamro, B. Shah, and F. Moreira, "A hybrid fog-edge computing architecture for real-time health monitoring in IoMT systems with optimized latency and threat resilience," *Scientific Reports*, vol. 15, no. 1, p. 25655, Jul. 2025, doi: 10.1038/s41598-025-09696-3.

[12] S. F. Ahmed, S. S. Shawon, S. Afrin, S. J. Rafa, M. Hoque, and A. H. Gandomi, "Optimising Internet of Things (IoT) Performance Through Cloud, Fog and Edge Computing Architecture," *IET Wireless Sensor Systems*, vol. 15, no. 1, 2025, doi: 10.1049/wss2.70016.

[13] P. R. Kumar and S. Goel, "Empowering Smart Cities with Fog Computing: A Versatile Framework for Enhanced Healthcare Services and Beyond," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 9, pp. 335–341, 2023, doi: 10.17762/ijritcc.v11i9.8363.

[14] S. Javid and A. Mirzaei, "Presenting a Reliable Routing Approach in IoT Healthcare Using the Multiobjective-Based Multiagent Approach," *Wireless Communications and Mobile Computing*, 2021, doi: 10.1155/2021/5572084.

[15] K. Aravind and P. K. R. Maddikunta, "Multiobjectives for Optimal Geographic Routing in IoT Health Care System," *Complexity*, 2022, doi: 10.1155/2022/7568804.

[16] Q. Cui *et al.*, "Uncertainty-aware traffic prediction and routing for emergency vehicles: a multi-objective risk-aware framework," *Connection Science*, vol. 38, no. 1, 2026, doi: 10.1080/09540091.2026.2658934.

[17] L. M. Tseng, P. F. Chen, and C. Y. Wen, "Design of Edge-IoMT Network Architecture with Weight-Based Scheduling," *Sensors (Basel, Switzerland)*, vol. 23, no. 20, p. 8553, 2023, doi: 10.3390/s23208553.

[18] G. Mehmood, M. Z. Khan, A. K. Bashir, Y. D. Al-Otaibi, and S. Khan, "An Efficient QoS-Based Multi-Path Routing Scheme for Smart Healthcare Monitoring in Wireless Body Area Networks," *Computers and Electrical Engineering*, vol. 109, p. 108517, 2023, doi: 10.1016/j.compeleceng.2022.108517.

[19] F. Ouakasse, A. Mosaif, and S. Rakrak, "A QoS-Gateway-Based Framework for Prioritizing Emergency Data in IoMT Applications," *Ingenierie des Systemes d'Information*, vol. 29, no. 1, pp. 369–375, 2024, doi: 10.18280/isi.290136.

[20] B. N. V, A. Kumar, and J. J. Thomas, "To design and implement the QoS -Aware Energy Efficient Routing Mechanism for the BAN-IoT networks in Smart Health care Applications," *Journal of Smart Internet of Things*, no. 2, pp. 55–78, 2024, doi: 10.2478/jsiot-2024-0012.

[21] R. Manchanda *et al.*, "Energy-efficient clustering and routing for IoT-enabled healthcare using adaptive fuzzy logic and hybrid optimization," *Scientific Reports*, vol. 15, no. 1, p. 34619, 2025, doi: 10.1038/s41598-025-18243-z.
[22] S. C. Swarna, V. R. Kolluru, and B. Potti, "Optimized cluster based routing protocol for IoT enabled healthcare data networks," *Scientific Reports*, vol. 16, no. 1, p. 17549, 2026, doi: 10.1038/s41598-026-46552-4.
[23] N. Shanmugapriya, "An Energy-Efficient Multi-Layer Secure Routing Protocol for Post-Quantum Cryptographic Networks in Internet of Medical Things (IoMT)," *International Journal of Multidisciplinary Sciences and Technology*, vol. 1, no. 1, pp. 1–8, 2025, doi: 10.64137/31079911/ijmst-v1i1p101.
[24] A. B. Kathole, K. Vhatkar, S. A. Ubale, V. V. Kimbahune, A. Dhumane, and A. Goyal, "Enhanced security mechanism in vehicular networks using ensemble machine learning to detect malicious activity in VANETs," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 7, pp. 2005–2014, 2024, doi: 10.47974/JDMSC-2075.
[25] A. B. Kathole, K. Vhatkar, G. Dharmale, S. Chiwhane, V. V. Kimbahune, and A. Goyal, "A novel approach to IoT security for intrusion detection system using ensemble network and heuristic-assisted feature fusion," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 7, pp. 2207–2217, 2024, doi: 10.47974/JDMSC-2092.
[26] A. B. Kathole *et al.*, "Secure Federated Cloud Storage Protection Strategy Using Hybrid Heuristic Attribute-Based Encryption With Permissioned Blockchain," *IEEE Access*, vol. 12, pp. 117154–117169, 2024, doi: 10.1109/ACCESS.2024.3447829.
[27] A. Sakthivel, P. Aruchamy, and B. Perumal, "A QoS-aware hybrid adaptive routing attack detection and mitigation approach for secure IoT-LLN environment," *Computer Journal*, vol. 69, no. 6, pp. 975–993, 2026, doi: 10.1093/comjnl/bxag005.
[28] A. Punitha, M. K. Vanteru, K. Manojkumar, and T. V. S. Reddy, "Blockchain-enabled spiking neural network for optimal routing and location privacy protection in IoT healthcare networks," *Iran Journal of Computer Science*, vol. 9, no. 1, p. 31, 2026, doi: 10.1007/s42044-025-00362-5.
[29] N. Sharma and P. G. Shambharkar, "Enhancing internet of medical things security: A multi-layered approach using dynamic adaptive deep reinforcement learning and blockchain," *Computers and Electrical Engineering*, vol. 129, p. 110808, 2026, doi: 10.1016/j.compeleceng.2025.110808.
[30] N. Sharma and P. G. Shambharkar, "Transforming security in internet of medical things with advanced deep learning-based intrusion detection frameworks," *Applied Soft Computing*, vol. 180, p. 113420, 2025, doi: 10.1016/j.asoc.2025.113420.
[31] N. Sharma and P. G. Shambharkar, "Multi-attention DeepCRNN: an efficient and explainable intrusion detection framework for Internet of Medical Things environments," *Knowledge and Information Systems*, vol. 67, no. 7, pp. 5783–5849, 2025, doi: 10.1007/s10115-025-02402-9.
[32] R. Mohammadi, "Optimized task allocation in SDN-enabled 5G IoMT networks using fog computing," *BMC Medical Informatics and Decision Making*, vol. 26, no. 1, 2026, doi: 10.1186/s12911-026-03603-0.
[33] A. B. Kathole, D. Jadhav, K. N. Vhatkar, S. Amol, and N. Gandhewar, "Solar energy prediction in IoT system based optimized complex-valued spatio-temporal graph convolutional neural network," *Knowledge-Based Systems*, vol. 304, p. 112400, 2024, doi: 10.1016/j.knosys.2024.112400.
[34] G. J. Escobar, M. N. Gardner, J. D. Greene, D. Draper, and P. Kipnis, "Risk-adjusting hospital mortality using a comprehensive electronic record in an integrated health care delivery system," *Medical Care*, vol. 51, no. 5, pp. 446–453, 2013, doi: 10.1097/MLR.0b013e3182881c8e.
[35] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000, doi: 10.1161/01.CIR.101.23.e215.
[36] P. A. Mackowiak, "A critical appraisal of 98.6 degrees F, the upper limit of the normal body temperature, and other legacies of Carl Reinhold August Wunderlich," *JAMA: The Journal of the American Medical Association*, vol. 268, no. 12, pp. 1578–1580, 1992, doi: 10.1001/jama.268.12.1578.
[37] World Health Organization, *Oxygen therapy for children*. Geneva, Switzerland: WHO Document Production Services, 2016.
[38] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001, doi: 10.1109/51.932724.
[39] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?," in *Proceedings - IEEE INFOCOM*, 2012, pp. 2731–2735, doi: 10.1109/INFCOM.2012.6195689.
[40] S. Khan, Y. V. Singh, P. Singh, and R. S. Singh, "An Optimized Artificial Intelligence System Using IoT Biosensors Networking for Healthcare Problems," *Computational Intelligence and Neuroscience*, pp. 1–14, 2022, doi: 10.1155/2022/2206573.
[41] R. L. Keeney, H. Raiffa, and D. W. Rajala, "Decisions with Multiple Objectives: Preferences and Value Trade-Offs," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 7, pp. 403–403, Jul. 1979, doi: 10.1109/TSMC.1979.4310245.

BIOGRAPHIES OF AUTHORS



Rayane ElRaba'a received her master's degree in Computer Science in 2024 and is currently a Ph.D. candidate in the Department of Mathematics and Computer Science, Faculty of Science, at Beirut Arab University (BAU), Lebanon. Alongside her doctoral studies, she serves as an Internet and Computing Core Certification (IC3) instructor. She has experience in being a teaching assistant for Computer Science courses and continues to teach various technology courses at the university's Continuing Education Center. Her research interests lie at the intersection of artificial intelligence, management information systems (mis), and information technology. She can be contacted at email: rayan.rabaa@bau.edu.lb.



Soha Rawas holds a Ph.D. in Mathematics and Computer Science from Beirut Arab University (2019). Her research spans artificial intelligence, deep learning, IoMT, cloud computing, medical image processing, secure intelligent systems, and AI in education, with applications across healthcare, learning technologies, and smart systems. She is a faculty member in the Department of Mathematics and Computer Science, Faculty of Science at BAU, and serves as the Director of the Continuing Education Center (CEC), leading strategic development, digital transformation, and innovative training programs across diverse professional fields. She can be contacted at email: soha.rawas2@bau.edu.lb.



Ali El-Zaart received his M.Sc. degree in 1996 and his Ph.D. degree in Computer Science in 2001 from the University of Sherbrooke, Sherbrooke, Canada. He worked as a senior software developer at the Research and Development Department of Semiconductor Insights Inc., Ottawa, Canada. In 2001, he joined King Saud University, Saudi Arabia, as an assistant professor in the Department of Biomedical Technology, College of Applied Medical Sciences and was promoted in 2010 to the rank of associate professor in computer science at the Department of Computer Science, College of Computer and Information Sciences, King Saud University. Since 2011, has been serving as a faculty member at Beirut Arab University (BAU), Lebanon and promoted in March 2016 to the rank of full professor in computer science. He held several administrative positions at BAU, including director of the Center for Continuing and Professional Education, director of the Entrepreneurship Center, and supervisor for the Maintenance of Agricultural Areas in Debbieh Campus. In 2023, he served as chairman of the Department of Computer Science at Beirut Arab University. He can be contacted at email: elzaart@bau.edu.lb.