Computer Systems
Science and Engineering

Tech Science Press

DOI: 10.32604/csse.2023.039367
Article

Energy and Latency Optimization in Edge-Fog-Cloud Computing for the Internet of Medical Things

Hatem A. Alharbi[1], Barzan A. Yosuf[2], Mohammad Aldossary[3,*] and Jaber Almutairi[4]

[1]Department of Computer Engineering, College of Computer Science and Engineering, Taibah University, Madinah, Saudi Arabia
[2]School of Electronic and Electrical Engineering, University of Leeds, Leeds, U.K
[3]Department of Computer Science, College of Arts and Science, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia
[4]Department of Computer Science, College of Computer Science and Engineering, Taibah University, Madinah, Saudi Arabia

*Corresponding Author: Mohammad Aldossary. Email: mm.aldossary@psau.edu.sa
Received: 25 January 2023; Accepted: 17 April 2023; Published: 26 May 2023

Abstract: In this paper, the Internet of Medical Things (IoMT) is identified as a promising solution, which integrates with the cloud computing environment to provide remote health monitoring solutions and improve the quality of service (QoS) in the healthcare sector. However, problems with the present architectural models such as those related to energy consumption, service latency, execution cost, and resource usage, remain a major concern for adopting IoMT applications. To address these problems, this work presents a four-tier IoMT-edge-fog-cloud architecture along with an optimization model formulated using Mixed Integer Linear Programming (MILP), with the objective of efficiently processing and placing IoMT applications in the edge-fog-cloud computing environment, while maintaining certain quality standards (e.g., energy consumption, service latency, network utilization). A modeling environment is used to assess and validate the proposed model by considering different traffic loads and processing requirements. In comparison to the other existing models, the performance analysis of the proposed approach shows a maximum saving of 38% in energy consumption and a 73% reduction in service latency. The results also highlight that offloading the IoMT application to the edge and fog nodes compared to the cloud is highly dependent on the tradeoff between the network journey time saved *vs.* the extra power consumed by edge or fog resources.

Keywords: Internet of medical things (IoMT); e-healthcare; edge- fog-cloud computing; remote monitoring; energy consumption; computation offloading

1 Introduction

The Internet of Things (IoT) has given rise to a diversity of disruptive technology solutions, used in various industries. Many IoT applications are already in use, which improve our daily lives (e.g., smart supply chains, smart cities, smart agriculture, and smart health). The Internet of Medical Things (IoMT) is a network of sensors and actuators operating within the healthcare industry. It is estimated that there were 4.5 billion IoMT sensors in 2015, accounting for more than 30% of global IoT sensors. This number grew to around 30 billion IoMT sensors [1]. Thus, the demand for IoMT services is rising and is expected to increase substantially over the next decade.

Smart IoMT applications can be developed, and alert services are triggered based on monitoring and analyzing patients' data [2]. In hospitals, for example, doctors and nurses are continually moving around but must maintain patients' health monitoring throughout the day. Therefore, healthcare professionals can collect the readings they require anywhere by equipping medical devices with various sensors. The usage of IoMT would save time for healthcare professionals and help them to take life-saving actions. Also, based on IoMT application data, doctors decide how to provide effective and timely treatments for patients.

During COVID-19, which imposed a greater emphasis on the use of remote health monitoring tools (e.g., smart hospitals and home-based health monitoring services), the adoption of IoMT-enabled healthcare became a more widespread practice to enhance efficiency and productivity during the pandemic [3]. Experts are now developing adaptable systems based on Machine-to-Machine (M2M) interactions thanks to the usage of these IoMT sensors, saving time for patients and medical staff, and facilitating patient care at home [4]. Accordingly, IoMT sensors are important to support IoT-based healthcare systems that assist in lowering the number of patients that visit hospitals.

The IoMT-based system has a significant number of sensors and medical equipment that are connected. Therefore, the cost of adopting IoMT applications is becoming a key challenge that needs more consideration [5]. Thus, adding sensors that are cheap with reasonable costs and minimal maintenance requirements will encourage the growth of additional IoMT devices and increase their adoption.

Furthermore, multiple IoMT devices that gather patient data must be connected and integrated into a smart healthcare architecture. Due to the different communication technologies used in smart healthcare, data aggregation and management are also challenging [6]. Sensors produce enormous amounts of medical data, which ultimately affects how doctors make decisions. As a result, efficient communication becomes imperative for the flow of IoMT data transmission for accurate and quick diagnoses [7]. Since there are billions of medical smart devices, wireless connectivity solutions must support the extensive connectivity of IoMT devices [8].

Another obstacle to the deployment of IoMT applications is power consumption. The majority of IoMT devices rely on batteries, which have a finite lifetime. It is impossible to undertake data analytics since many healthcare monitoring systems comprise of wearable devices with constrained processing and storage capabilities [8,9]. Hence, the use of cloud services in these systems has generated a variety of options for completing these tasks. Due to the cloud's increased processing capacity, it can carry out these tasks effectively, more quickly, and with a higher degree of diagnostic accuracy. However, determining the data that must be handled locally from the data that needs to be sent to the cloud servers is critical, since IoMT devices and their applications that entail real-time interactions are a rising source of big data. Therefore, the efficient placement of IoMT applications that require intensive computational processing on edge-fog-cloud systems should be the primary focus in order to assimilate the IoMT system capabilities [10,11].

The processing of vast amounts of data can be done more effectively to assist big data analytics by using cloud services. Through the cloud, IoT subsystems do not have to deal with extensive computing, which will increase the battery lifetime of these devices [3]. The cloud-based framework's main shortcomings include network connections, the delay in getting the results, and even security concerns. Therefore, the cloud is an essential element of any IoT application, but it is problematic in systems that are sensitive to delay, such as health monitoring. For IoMT systems, distributed edge and fog computing technologies have replaced centralized cloud computing in the data processing architecture.

A hierarchy of layers is created by fog computing among the edge nodes and the cloud server, which aims to lower the amount of data transferred between the IoMT devices and the cloud server in order to reduce service latency [4]. Due to the fact that data is kept locally at the edge and in the fog rather than in the cloud, fog computing significantly improves data security and maintains system privacy. Modern fog concepts and edge computing technologies bring resources closer to the end-user and deliver energy efficiency and low service latency, as well as greatly enhance the system's reliability and quality of service. It also offers extra benefits over cloud computing due to its relatively high computing power, storage capacity, and real-time data analysis capabilities [2].

Edge computing is becoming an indispensable solution in the field of remote healthcare, due to its effective processing of healthcare data in real time. In edge computing, data is kept close to the edge of the network, or near the server where it was generated, rather than being sent and stored in the cloud [4]. This would reduce service latency by speeding up data streaming while processing IoMT data and giving medical devices an immediate response.

IoMT applications are empowered by edge-fog-cloud computing resources to store and process the heterogeneous data of IoMT sensors. Even though cloud layers can manage these data, some IoMT sensors produce a huge amount of data. This may cause a bottleneck in the network and therefore increase networking power consumption due to activating extra links for accommodating the traffic demands. Also, some IoMT applications are time-critical and therefore should be managed in a real-time manner with low latency. However, the decision to place IoMT applications in edge or fog layer leads to an increase in power consumption of processing nodes due to the higher associated Power Usage Effectiveness (PUE) compared to that of the cloud processing nodes [12]. The higher PUE of edge and fog nodes compared to the cloud results in a situation where offloading IoMT applications to edge or fog nodes for processing incurs extra power consumption. As a result, offloading IoMT data to the edge, fog, or cloud is highly dependent on the tradeoff between the power and journey time saved *vs.* the extra power consumed by edge or fog resources.

1.1 Motivation and Contribution

Motivated by the studies discussed in the previous section, we developed a four-tier IoMT-edge-fog-cloud architecture along with an optimization model that enables medical data to be offloaded in an efficient manner, while maintaining certain quality standards (e.g., energy consumption, service latency, network utilization). The major contributions of this work can be summarized as follows:

- The IoMT-edge-fog-cloud architecture is designed to support intelligent healthcare systems solutions using Mixed Integer Linear Programming (MILP) mathematical models.
- The IoMT application placement decisions over edge-fog-cloud architecture are optimized based on multiple factors, such as power consumption, service latency and network resource utilization.

- The result is analyzed and compared with the cloud-only approach (baseline) to evaluate and validate the performance of the proposed approach in terms of energy consumption, service latency and network resource utilization.

The remainder of this work is organized as follows: Section 2 presents some related works in the field of IoMT applications and their performance metrics. Section 3 presents the architecture of the IoMT-edge-fog-cloud. Next, Section 4 presents the mathematical modeling for energy consumption, service latency, and resource utilization of IoMT data offloading over the architecture of edge-fog-cloud. In Section 5, we describe how the model is designed and the input parameters. The performance metrics analysis of the defined architecture can be found in Section 6. In Section 7, the paper is concluded and future research is discussed.

2 Related Works

In the healthcare sector, computing layers (e.g., edge, fog, and cloud layers) provide different computing capabilities such as storage, processing, and communication links over the internet to fulfill IoMT application requirements. Therefore, an integrated edge-fog-cloud architecture offers scalable data analytics and trustworthy solutions to overcome IoMT application challenges (e.g., the problem of reducing service execution time and energy consumption of IoMT applications) [13]. A number of studies in the literature [10,11,14–16] have highlighted the importance of edge, fog, and cloud computing in terms of optimizing the placement of IoMT applications, considering several performance metrics such as energy consumption, service latency, resource usage, and security [17,18].

This section presents some of the related work for the placement of IoMT applications in edge-fog-cloud systems. For example, the authors in [14] developed an IoMT fog-based access-control determination (ACD) algorithm, which enables users to gain a broad scope of access to their applications by empowering cloud computing. Their approach has shown that enabling the fog layer could reduce execution time for IoMT applications while ensuring high-level privacy. In [19] an offline/online signature certificateless method has been proposed. In their study, the authors concluded that the proposed scheme provides enhanced security while being computationally and communicationally efficient. In [15], the authors developed an optimization conceptual fog computing framework that could reduce network communication delays and enhance fog resource utilization via the application of a genetic algorithm. To achieve better performance, authors in [11] conducted a simulation analysis for integrating fog computing with cloud computing paradigms. They found that offloading IoMT data to fog computing reduced response time by approximately 86% compared to the cloud layer.

Furthermore, an integrated edge-fog-cloud healthcare framework was developed in [10] to minimize the service latency and power consumption of using IoMT applications by approximately 28% and 27%, respectively. In [20], an effective framework based on the Remora Optimization algorithm was developed to jointly minimize latency and power consumption for IoMT applications. Also, a novel message exchange procedure with load balancing was proposed in [4] to offload IoMT applications via a cloud-fog architecture with the aim of reducing energy consumption and delay. Their energy-efficient fuzzy approach was able to achieve up to 77% and 60% reductions in energy consumption and delay, respectively. Also, the authors in [21] presented a genetic algorithm to provide a secure and energy-saving method to detect communicable infectious diseases using a wireless sensor network (WSN). In [22], the authors developed an optimization model (mixed-integer non-linear programming (MINLP) model) and applied an enhanced deep reinforcement learning approach to find an optimal strategy for resource allocation, computation offloading, and minimizing energy consumption. In addition, a hybrid energy-efficient model to monitor patients at their homes is developed in [23]. Using sensors, the proposed method captures and analyzes electrocardiograms (ECG). Through their proposed system, energy consumption, latency, and network utilization can be reduced.

In comparison to the works presented in this section, we propose an optimization model to minimize energy consumption, service latency, and network resource utilization to optimize the placement of IoMT applications over a four-tier IoMT-edge-fog-cloud architecture, considering different traffic loads and processing requirements. Table 1 summarizes the approach features, methodologies, and performance metrics of some related works.

Table 1: Related works

Reference	Approach Features	Methodologies	Performance metrics
[14]	Developed a fog-based access control model, which enables users to gain a broad scope of access to their applications by empowering cloud computing and Internet-based services.	• Access-control determination (ACD) algorithm	• Privacy protection • Energy efficiency • Reducing execution time
[24]	Developed a fog query-processing framework for IoT applications to lower network processing and dependency on the cloud.	• Query processing technique	• Response time • Latency • Energy consumption
[10]	Proposed a healthcare framework to reduce the latency and power consumption of end-user devices by ∼28% and ∼27%, respectively.	• Mobility pattern model • Prediction model-generative adversarial network (GAN)	• Reducing delay • Reducing energy consumption
[15]	Developed a conceptual fog computing framework that could reduce network communication delays and enhance fog resource utilization via the application of a genetic algorithm and an exact optimization method.	• Genetic algorithm	• Delay reduction in network communication •Utilizing fog resources effectively
[16]	Proposed an Artificial Intelligence (AI) paradigm to empower sixth-generation (6G) edge computing-based e-healthcare.	• Fuzzy-based Sustainable, Interoperable, and Reliable Algorithm (FSIRA)	• Reliability • Latency • Throughput
[11]	Integrated fog computing with cloud computing paradigms to achieve better performance.	• Cloud analyst simulation toolkit	• Response time • Processing time • Cost of transferring data to the cloud
[17]	Proposed an efficient and provable secure certificate-based combined signature, encryption, and signcryption (CBCSES) scheme.	• Detailed Security Analyses	• Security
[18]	Designed a lightweight and secure strategy for IoT-based WBAN.	• Certificateless signcryption and identity-based signcryption	• Security
[20]	Proposed an energy-efficient and secure strategy for IoMT.	• Meta-heuristic approach (e.g., Remora Optimization Algorithm (ESRO))	• Energy consumption • Security
[4]	Proposed an energy-efficient fuzzy approach to offload IoMT data using a novel message exchange procedure and load-balancing solution.	• Developed a four-level architecture using the MQTT protocol	• Energy consumption • Transmission delay • CPU utilization
[21]	Proposed a secure and energy-saving method for IoMT applications and used wireless sensor network (WSN) for communicable infectious diseases.	• Genetic algorithm	• Energy consumption • Security
[25]	Developed a clustering model for providing effective communication to IoMT-based applications.	• Cluster head selection technique	• Energy consumption • Network resource utilization
[26]	Developed an energy-aware clustering strategy for 5G-enabled edge computing-based IoMT framework.	• Optimization algorithm	• Energy efficiency
[27]	Proposed adaptive energy-efficient algorithm (EEA) to enhance system throughput and energy consumption.	• Algorithm based in MATLAB simulation	• Energy efficiency
[28]	Proposed an algorithm to network selection framework for IoMT.	• Reinforcement learning (RL)-based network selection scheme	• Energy efficiency • Delay
This Work	We developed a four-tier IoMT-edge-fog-cloud architecture along with an optimization model for efficient processing and placing IoMT applications.	• Mixed integer linear programming (MILP) model	• Energy consumption • Service latency • Network resource utilization

3 IoMT-Edge-Fog-Cloud Architecture

The IoMT architecture is a use case of the Internet of Things (IoT) paradigm, where multiple heterogeneous devices can interact and communicate with each other. Accordingly, the placement of IoMT applications is crucial since placing them on the edge, fog, or cloud is based on their energy consumption and service latency requirements. To allow such data-driven decision-making, a general and hierarchical architecture based on IoMT-edge-fog-cloud systems is designed.

The four-tier of our proposed architecture (as shown in Fig. 1) consists of the IoMT layer, an edge layer, a fog layer, and a cloud layer. First, the IoMT layer sends real-time data through IoMT devices such as IoMT sensors (body sensors). Next, the edge layer gathers and receives IoMT data for processing highly critical and latency-sensitive tasks. Then, the fog layer is used to reduce service latency and make it suitable for real-time analysis. Finally, the cloud server is used to handle any compute-intensive tasks and big data analytics. The proposed system architecture's layers are described as follows:

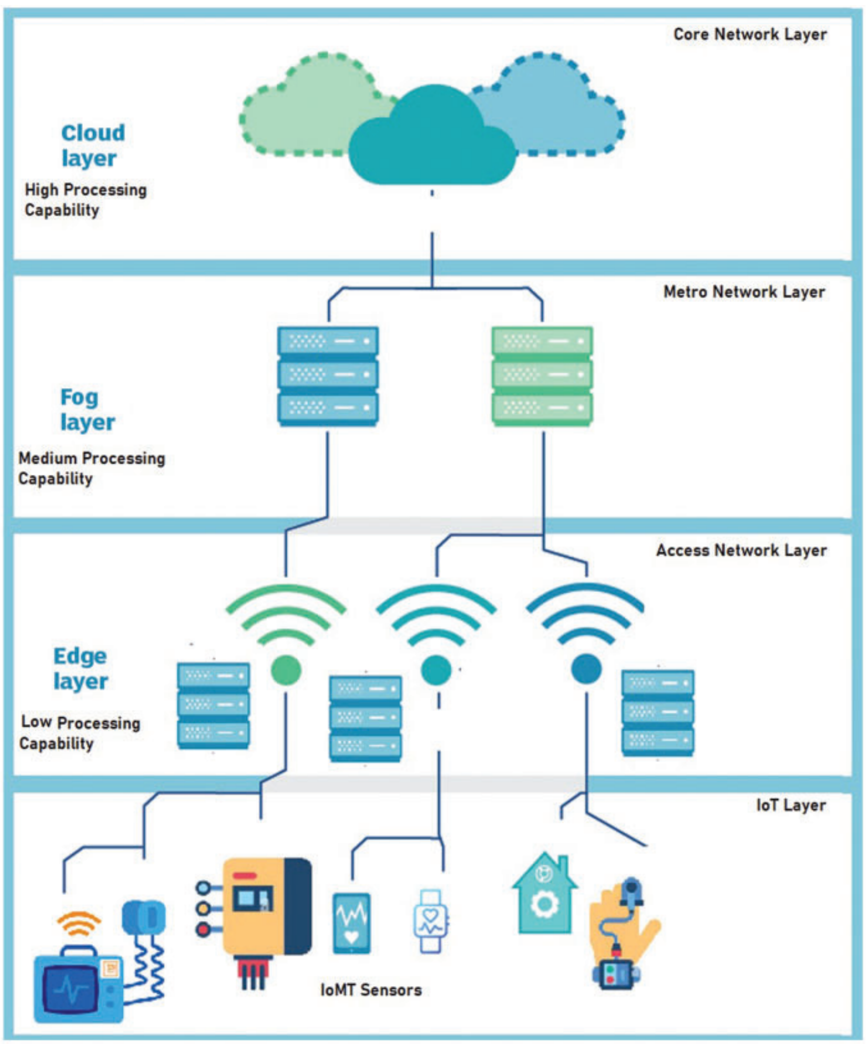

Figure 1: The IoMT-edge-fog-cloud system architecture

3.1 Internet of Medical Things (IoMT) Layer

The IoMT layer is composed of different devices such as remote patient monitoring sensors, including (diabetes sensors, heart failure sensors, body temperature sensors, glucose level sensors, heart rate sensors, connected inhalers sensors, and many more). The IoMT sensors monitor and gather data from wearable health devices that are used by patients to notify emergency medical services or patients who require quick diagnosis and health checks. Aged people and patients with serious medical issues are the main players in this layer.

In addition, these IoMT devices generate heterogeneous massive data, which exceeds their processing and storage capabilities. Therefore, various solutions are suggested to overcome this problem. For example, some of the data can be processed locally at the device level, and some other data that require more processing and storage capabilities can be offloaded to the edge, the fog, or the cloud level via wireless communication networks [9].

Various wireless communication technologies, including NB-IoT, Zigbee, LoRa, Bluetooth, 4G/5G, and Wireless Fidelity (Wi-Fi) are utilized to transport patient data to higher layers (edge/fog/cloud) for processing in real-time [2]. Although these communication technologies have unique features (e.g., low/long-range communications, power efficiency, high transmission rate, etc.), this may assist IoMT-based systems in achieving optimal performance.

3.2 Edge-Fog-Cloud Layers

These layers' main responsibilities include processing and analyzing the data transferred from sensors related to patient care. The IoMT sensors collect a massive amount of patients' data (healthcare data), which needs to be immediately processed. This introduces high bandwidth usage as a result of sending this massive data to the cloud and it may lead to inefficient computation and increased service latency. Therefore, processing some or all of the data at the edge or fog layers is the preferred choice [6].

The edge layer consists of different computing and networking equipment, such as gateway devices, etc. The edge layer offers wireless access to IoMT devices through Bluetooth, 4G/5G, Wi-Fi, etc. Essentially, the edge layer provides more powerful processing and storage capabilities than the IoMT layer. Thus, it receives and processes the data offloaded from the IoMT layer in order to provide a rapid service response with low communication overhead.

The fog layer's objective is to reduce network usage and energy consumption for processing IoMT applications. Also, services in the fog layer can be offloaded to the cloud layer for better performance, since the fog and the cloud are complementary solutions. Fog computing brings computing resources like data management, networking, processing, and storage closer to the "*Things*", which is the main advantage of integrating it into an IoMT system. Furthermore, fog computing eliminates the need for a drawn-out journey through the cloud and allows for real-time applications and quick responses.

The cloud layer has more powerful processing and storage capabilities compared to the edge and fog layers in order to handle some difficult tasks (e.g., big data analytics, predictions, and diagnoses). While the edge and fog layers are aimed to reduce the service latency of the processed data, the cloud layer can also reduce energy consumption due to its highly efficient processing and storage capabilities. At this layer, the processed IoMT sensor data and patient data are permanently stored in the cloud (e.g., patient records, medical images, image diagnostics, alerts, and notifications). Using this method, doctors can investigate certain medical problems whenever they want and from any remote location. Also, patients having cloud-based data can conveniently consult with medical professionals who have been approved by the system.

To allow for the effective use of resources in the edge-fog-cloud architecture, virtualization concept [29] should be adopted. Also, the resource allocation management technique offered by virtualization abstracts the physical resources of the data center into several logical units known as Virtual Machines (VMs) [30]. Each isolated VM running on a physical server has access to its own resources, including CPU, memory, network bandwidth, and storage, to run its applications. In edge-fog-cloud computing infrastructure (distributed data centers), placing or relocating VMs is a key activity where the best server is selected to host the VM based on (e.g., network resource utilization, workload balancing, and energy efficiency).

3.3 Communication Networks

The core network, the metro network, and the access network are the three primary layers of the conventional IP network structure employed by Internet Service Providers (ISPs). As it connects important regions and cities, the core network serves as the basis of any telecom network. The core network uses IP over WDM technology extensively because of its high scalability and the data rates it offers on its links. Each core node is connected to a metro network, which serves a metropolitan area, and offers a direct connection between end-users in access networks and core network nodes. In optical access networks, ONU and OLT devices are the main used components [31].

4 Mathematical Model

This section defines the structure of the Mixed Integer Linear Programming (MILP) model to optimize the usage of network and computing resources in order to minimize total energy consumption, service latency, and network utilization for offloading IoMT applications over edge-fog-cloud systems. The parameter sets and variables of the MILP model are defined in the next subsection.

4.1 Internet of Medical Things (IoMT), Edge, Fog, and Cloud Layers

Tables 2 and 3 show the parameters and variables that describe the IoMT layer.

Table 2: The IoMT parameters

Parameter	Description
$IoMT_s^{(number)}$	The number of IoMT sensors distributed in a geographical location s.
U_i	The data rate of IoMT sensor i distributed in a single geographical location.
$OLT^{(power)}$	Optical line terminal (OLT) power consumption.
$ONU^{(power)}$	Optical network unit (ONU) power consumption.
$UI_{i,s,d}$	Offloaded data from IoMT i in node *s* to VM located in either cloud, fog or edge in node d given as: $\sum_{d \in N} UI_{i,s,d} = IoMT_s^{(number)} U_i$
$OLT^{(bitrate)}$	The maximum bitrate of an OLT.
$ONU^{(bitrate)}$	The maximum bitrate of an ONU.
$PUE^{(network)}$	Network power usage effectiveness (PUE).
ω	Large enough positive number.

Table 3: The IoMT variables

Variable	Description
$ONU_s^{(number)}$	The number of ONU terminals.
$OLT_s^{(number)}$	The number of OLT.

The defined parameters and variables (in Tables 4 and 5) provide the IoMT VMs that will be migrated/replicated in the three computing layers along with data power consumption of communication and computation devices.

Table 4: Communications and computing parameters

Parameter	Description
$NP^{(cloud)}$	Power consumption of internal cloud layer network (Calculated as energy/bit).
$NP^{(fog)}$	Power consumption of internal fog layer network (Calculated per energy/bit).
$NP^{(edge)}$	Power consumption of internal edge layer network (Calculated per energy/bit).
$PCS^{(power)}$	Peak power consumed per cloud server.
$PFS^{(power)}$	Peak power consumed per fog server.
$ES^{(power)}$	Peak power consumed per edge server.
$CS^{(MIPS)}$	Number of MIPS a cloud server can serve.
$FS^{(MIPS)}$	Number of MIPS a fog server can serve.
$ES^{(MIPS)}$	Number of MIPS an edge server can serve.
$PPMIPS^{(cloud)}$	Power consumed by cloud server per MIPS operation, where $PPMIPS^{(cloud)} = \frac{CS^{(power)}}{CS^{(MIPS)}}$.
$PPMIPS^{(fog)}$	Power consumed by fog server per MIPS operation, where $PPMIPS^{(fog)} = \frac{FS^{(power)}}{FS^{(MIPS)}}$.
$PPMIPS^{(edge)}$	Power consumed by edge server per MIPS operation, where $PPMIPS^{(edge)} = \frac{ES^{(power)}}{ES^{(MIPS)}}$.
$PUE^{(cloud)}$	Cloud layer PUE.
$PUE^{(fog)}$	Fog layer PUE.
$PUE^{(edge)}$	Edge layer PUE.
N	A number of heterogeneous nodes in the IoMT-edge-fog-cloud system architecture.
c	Cloud layer-nodes set.
f	Fog layer-nodes set.
e	Edge layer-nodes set.
i	A number of nodes in IoMT layer.
s and d	An IoMT-edge-fog-cloud system architecture source and destination indices.
$\mathbb{I}$	The number of VMs hosting applications of IoMT.
S_i	The data rate uploaded by IoMT devices to VM i.
U_i	Quantity of IoMT devices served by VM i.
P_i	MIPS requirements of IoMT application hosted in VM i.
$MR^{(power)}$	Peak power consumed per metro router.
$MS^{(power)}$	Peak power consumed per metro switch.

(Continued)

Table 4: Continued

Parameter	Description
$MR^{(bitrate)}$	Peak traffic rate per metro router.
$MS^{(bitrate)}$	Peak traffic rate per metro switch.
m and n	Source/destination indices of the nodes *s*, m ∈ N of the IoMT-edge-fog-cloud system.
N_m	The neighboring node of node m in the IoMT-edge-fog-cloud system.
$\mathbb{r}^{(power)}$	Power consumption by router port in the core network.
$\mathbb{t}^{(power)}$	Power consumption by the transponder in the core network.
$\mathbb{E}^{(power)}$	Power consumption by the EDFAs in the core network.
$\mathbb{S}^{(power)}$	An optical switch's power consumption in the core network.
B	A bit rate for each wavelength.
$\mathbb{S}$	Distance between two amplifiers.
$\mathbb{D}_{m,n}$	Distance between two connected nodes in the core network (m, n) ∈ $\mathbb{c}$ in kilometers.
$\mathbb{A}_{m,n}$	Quantity of amplifiers two connected nodes (m, n) ∈ $\mathbb{c}$. $\mathbb{A}_{m,n} = \left\lfloor \frac{\mathbb{D}_{m,n}}{S} - 1 \right\rfloor$, $\mathbb{S}$ is the maximum distance an amplifier can reach.

Table 5: Communications and computing variables

Variable	Description
Θ_d	$\begin{cases} \text{set to 1 if server in node d } \in \text{ N is switched on} \\ \text{set to 0 if f server in node d } \in \text{ N is switched off} \end{cases}$
$\Psi_{i,d}$	$\begin{cases} \text{set to 1 if an IoMT VM , i} \in \mathbb{I} \text{ is placed in a server d } \in \text{ N} \\ \text{set to 0, if not} \end{cases}$
$MIPS_{i,s}^{iot}$	Processing requirements in MIPS for VM i placed in either edge, fog, or cloud s.
$MIPS_{s}^{iot}$	The sum of processing requirements in MIPS for VM i placed in either edge, fog, or cloud s.
$T_{i,s,d}^{iot}$	Data offloaded from IoMT devices in node d to VM i to either edge, fog, or cloud s.
$TU_{s,d}$	Data offloaded from IoMT devices in node d to either edge, fog, or cloud s.
$MR_{s}^{(number)}$	The sum of routers utilized in a metro network in node *s*.
$MS_{s}^{(number)}$	The sum of switches utilized in a metro network in node *s*.
$\mathbb{r}_d$	Quantity of router ports count in core node d ∈ $\mathbb{c}$
T_s	The sum of offloaded data in each node s ∈ N.
$\mathbb{F}_{m,n}$	Quantity of fibers on the connection (m, n) ∈ $\mathbb{c}$.
$\mathbb{L}_{m,n}^{s,d}$	Offloaded data traverse nodes (s, d) ∈ $\mathbb{c}$ communicating through the physical link (m, n) ∈ $\mathbb{c}$.
$\Gamma_{m,n}^{s,d}$	$\begin{cases} \text{set to 1 if IoMT device sends traffic (s, d)} \in \mathbb{c} \text{ using the physical link (m, n)} \in \mathbb{c} \\ \text{set to 0, if not} \end{cases}$

An edge-fog-cloud system architecture consumes the following power:

The cloud computing layer (Cloud):

$$PUE^{(cloud)}\left(\sum_{s\in N} MIPS_{i,s}^{iot} PPMIPS^{(cloud)} + \sum_{s\in N} NP^{(cloud)} TU_{s,d}\right) \forall s \in \mathbb{C} \quad (1)$$

The core network layer (Core):

$$PUE^{(network)}\left(\sum_{d\in N} \mathbb{r}^{(power)} \mathbb{r}_d + \sum_{m\in N} \sum_{n\in Nm_m:n\neq m} \sum_{s\in N} \sum_{d\in N:s\neq d} \Gamma_{m,n}^{s,d} \mathbb{t}^{(power)} + \sum_{m\in N} \sum_{n\in Nm_m:n\neq m} \mathbb{E}^{(power)} \mathbb{F}_{m,n} \mathbb{A}_{m,n} + \sum_{d\in N} \mathbb{S}_d^{(power)}\right) \quad (2)$$

The cloud layer's power consumption is calculated using function (1), taking into consideration the processing nodes and the internal networking components with the Power Usage Effectiveness (PUE) factor. Function (2) defines the power consumption of the core network, taking into consideration core switches, amplifiers, transponders, and router ports as well as the networking PUE.

The fog computing layer (Fog):

$$PUE^{(fog)}\left(\sum_{s\in N} MIPS_{i,s}^{iomt} PPMIPS^{(fog)} + \sum_{s\in N} NP^{(fog)} TU_{s,d}\right) \forall s \in \mathbb{f} \quad (3)$$

The metro area network (Metro):

$$PUE^{(network)} \left(\left(MR_s^{(number)} MR_s^{(power)}\right) + \left(MS_s^{(number)} MS_s^{(power)}\right)\right) \forall s \in N \quad (4)$$

The fog layer's power consumption is calculated using function (3), through processing nodes and internal networking components and the associated fog PUE. Based on function (4), several factors are considered when calculating a network's power consumption like router ports, switch devices, and their PUEs.

The edge layer (Edge):

$$PUE^{(edge)}\left(\sum_{s\in N} MIPS_{i,s}^{iomt} PPMIPS^{(edge)} + \sum_{s\in N} NP^{(edge)} TU_{s,d}\right) \forall s \in \mathbb{e} \quad (5)$$

The access network (Access):

$$PUE^{(network)}\left(\sum_{s\in N} ONU_s^{(number)} ONU^{(power)}\right) + \left(\sum_{s\in N} OLT_s^{(number)} OLT^{(power)}\right) \quad (6)$$

Function (5) defines the power consumption of the edge layer, taking into consideration processing nodes and internal networking components, with the associated edge PUE. The access network's power consumption is defined in function (6), involving ONU terminals and OLT equipment's and the associated networking PUE.

According to the MILP model, the objective is to minimize the total power consumption of IoMT-edge-fog-cloud architecture by calculating all functions (1–6) as follows:

$$\text{Cloud} + \text{Core} + \text{Fog} + \text{Metro} + \text{Edge} + \text{Access} \quad (7)$$

As mentioned above, the purpose is to reduce the power consumption, where function (7) provides the total power consumption of IoMT-edge-fog-cloud architecture as the sum of the different processing and communication layers.

Subject to the following constraints:

Data uploaded by IoMT devices:

$$\sum_{s,d\in N} UI_{i,s,d} = \sum_{s,d\in N} T^{iot}_{i,s,d} \quad \forall\, i \in \mathbb{I} \tag{8}$$

Constraint (8) ensures that all the data uploaded by IoMT devices are handled by an edge, fog, or cloud node.

IoMT VM in edge-fog-cloud system architecture constraints:

$$\sum_{s\in N} T^{iomt}_{i,s,d} \geq \Psi_{i,d} \quad \forall\, d \in N\,, i \in \mathbb{I} \tag{9}$$

$$\sum_{s\in N} T^{iot}_{i,s,d} \leq \omega\, \Psi_{i,d} \quad \forall\, d \in N\,,\ i \in \mathbb{I} \tag{10}$$

As a result of constraints (9) and (10), the binary variable $\Psi_{i,d}$ is set to 1 if the server in node $d \in N$ is turned on to serve the IoMT VM $i \in \mathbb{I}$, otherwise $\Psi_{i,d}$ it is set to 0.

Physical communication link:

$$\mathbb{L}^{s,d}_{m,n} \geq \Gamma^{s,d}_{m,n} \quad \forall\, s, d, m, n \in N \tag{11}$$

$$\mathbb{L}^{s,d}_{m,n} \leq \Gamma^{s,d}_{m,n} \quad \forall\, s, d, m, n \in N \tag{12}$$

When data traverses between nodes $(s, d) \in \mathbb{c}$ using the physical link $(m, n) \in \mathbb{c}$, constraints (11) and (12) verify that the physical communication link $(m, n) \in \mathbb{c}$ is switched on.

Edge-fog-cloud processing requests:

$$MIPS^{iomt}_{i,d} = \Psi_{i,d}\, MIPS^{iomt}_{i,d} \quad \forall\, d \in N\,, i \in \mathbb{I} \tag{13}$$

$$MIPS^{iomt}_{d} = \sum\nolimits_{i\in\mathbb{I}} MIPS^{iomt}_{i,d} \quad \forall\, d \in N \tag{14}$$

Constraint (13) defines the processing requests of IoMT application hosted in VM $i \in \mathbb{I}$ in either edge, fog, or cloud layer. Constraint (14) calculates the sum of processing requests in either edge, fog, or cloud layer $d \in N$.

Data in core network:

$$TU_{s,d} = \sum_{i\in\mathbb{I}} T^{iomt}_{i,s,d} \quad \forall\, s,\ d \in \mathbb{c} \tag{15}$$

Constraint (15) defines the data traverse between core network nodes due to the IoMT VMs placed in the clouds.

Flow conservation:

$$\sum_{m\in N: m\neq n} \mathbb{L}^{s,d}_{m,n} - \sum_{n\in N: m\neq n} \mathbb{L}^{s,d}_{m,n} = \begin{cases} L_{s,d} & i = s \\ -L_{s,d} & i = d \\ 0 & \text{otherwise} \end{cases} \quad \forall\, s,\ d\ \in N\ : s \neq d \tag{16}$$

The flow conservation of the core network is defined by constraint (16). It guarantees equality in arriving/leaving data in all core networks; except the source/destination nodes.

Physical communication link size:

$$\sum_{s\in N}\sum_{d\in N:i\neq j}\mathbb{L}_{m,n}^{s,d}\leq W\mathbb{B}F_{m,n}\quad\forall\, m,\ n\in N \tag{17}$$

Constraint (17) defines the physical communication size by ensuring that the data traversing the physical communication link will not overreach its capability.

The number of core network router ports:

$$\mathbb{r}_d\geq\frac{\sum_{s\in\mathbb{c}}TU_{s,d}}{\mathbb{B}}\quad\forall\, d\in\mathbb{c} \tag{18}$$

Constraint (18) defines the number of network router ports in each core node.

The number of ONU terminals:

$$ONU_s^{(number)}\geq\frac{\sum_{i\in\mathbb{i}}\sum_{d\in N}UI_{i,s,d}}{ONU^{(bitrate)}}\quad\forall\, s\in N \tag{19}$$

The number of ONU terminals access node is calculated by constraint (19).

The number of OLT:

$$OLT_s^{(number)}\geq\frac{\sum_{i\in\mathbb{I}}\sum_{d\in N}UI_{i,s,d}}{OLT^{(bitrate)}}\quad\forall\, s\in N \tag{20}$$

The number of OLTs in each access node is calculated by constraint (20).

The number of metro routers:

$$MR_s^{(number)}\geq 2\frac{\sum_{i\in\mathbb{i}}\sum_{d\in(\mathbb{f}\cap\mathbb{c})}UI_{i,s,d}}{MR^{(bitrate)}}\quad\forall\, s\in N \tag{21}$$

Constraint (21) calculates the number of routers in each metro node.

The number of metro switches:

$$MS_s^{(number)}\geq\frac{\sum_{i\in\mathbb{i}}\sum_{d\in(\mathbb{f}\cap\mathbb{c})}UI_{i,s,d}}{MS^{(bitrate)}}\quad\forall\, s\in N \tag{22}$$

Constraint (22) calculates the number of switches in each metro node.

The sum of offloaded data in the communication network:

$$T_d=\sum_{i\in\mathbb{i}}\sum_{d\in N}UI_{i,s,d}\quad\forall\, d\in N \tag{23}$$

The aggregated traffic data in each destination node *d* is determined by constraint (23).

5 Model Design and Input Parameters

The IoMT, edge, fog, and cloud layers are the four layers that make up the intelligent healthcare system in the proposed MILP model. The type of tasks (based on the required MIPS) demanded by

IoMT devices at the IoMT layer determine how the edge, fog, and cloud layers are configured. Table 6 contains all the input parameters for the model's various layers.

Table 6: Model input parameters [6,32]

Parameter	Value
Number of IoMT sensors ($IoMT_s^{(number)}$)	100 million sensors (high-critical sensors 60%, medium-critical sensors 30%, low-critical sensors 10%
Processing requests of each VM (P_i).	- 500 MIPS for high critical data which require low processing requirements. - 2000 MIPS for medium critical data which require medium processing requirements. - 5000 MIPS for low critical data which require high processing requirements.
The data rate of each IoMT sensor (U_i).	100 Kbps
Server power consumption in cloud ($CS^{(power)}$).	630 Watts
Server power consumption in fog ($FS^{(power)}$).	126 Watts
Server power consumption in edge ($ES^{(power)}$).	63 Watts
Size in MIPS of each cloud server ($CS^{(MIPS)}$).	18000 MIPS
Size in MIPS of each fog server ($FS^{(MIPS)}$).	3600 MIPS
Size in MIPS of each edge server ($ES^{(MIPS)}$).	1800 MIPS
Power consumption of internal networking of cloud layer ($PPbits^{(cloud)}$).	2.48 W/Gbps
Power consumption of internal networking of fog layer ($PPbits^{(cloud)}$).	2.57 W/Gbps
Power consumption of internal networking of edge layer ($PPbits^{(cloud)}$).	2.70 W/Gbps
Cloud layer PUE ($PUE^{(cloud)}$)	1.1
Fog layer PUE ($PUE^{(fog)}$)	1.9
Edge layer PUE ($PUE^{(edge)}$)	2.5
Core router port power consumption ($\mathbb{r}^{(power)}$)	37.1 W
Core router port bandwidth (B)	40 Gbps
Transponder power consumption ($\mathbb{t}^{(power)}$)	129 W
Optical switch power consumption ($\mathbb{S}_d^{(power)}$)	85 W
Amplifier power consumption ($\mathbb{E}^{(power)}$)	11 W
Maximum distance between two amplifiers ($\mathbb{S}$)	80 KM
PUE of network ($PUE^{(network)}$)	1.5
Metro router port power consumption ($MR^{(power)}$)	30 W
Metro switch power consumption ($MS^{(power)}$)	470 W
Metro router port bitrate ($MR^{(bitrate)}$)	40 Gbps
Metro switch data rate ($MS^{(bitrate)}$)	0.5 Tbps

(Continued)

Table 6: Continued

Parameter	Value
ONU terminal power consumption ($ONU^{(Power)}$)	5 Watts
ONU size ($ONU^{(bitrate)}$)	2.4 Gbps
OLT size ($OLT^{(bitrate)}$)	1.280 Tbps
Power consumption of OLT devices ($OLT^{(Power)}$)	1.842 kW

IoMT applications have special processing and traffic patterns compared to traditional IoT applications. There are three types of IoMT applications: high-critical, medium-critical, and low-critical. Each requires a different data rate, different analysis and processing requirements, and a different frequency of uploading data. All these characteristics make IoMT applications distinct from IoT applications. In this work, we used 100 million IoMT devices distributed across the AT&T core network architecture [6] as a basis for our model. Also, we consider three categories of IoMT sensor task requirements, which are highly critical sensing (accounting for 60% of the number of IoMT devices), medium critical sensing (30%), and low critical sensing (10%).

For modeling simplicity, the following scenarios are considered. First, the IoMT sensors offload high-critical sensing data (e.g., heart failure sensors, body temperature sensors, glucose sensors, heart-rate sensors, and connected inhaler sensors) that require intensive periodic processing i.e., every minute, are assumed to be offloaded to the edge. Second, the medium-critical sensing devices (e.g., to notify emergency medical services or patients who require quick diagnosis and health checks), require less frequent uploading of data (every 15 min). This group of tasks is assumed to be offloaded to the fog. Third, the low-critical sensing tasks (e.g., hospital environment data which require big data analytics techniques, predictions, and diagnoses) are far more relaxed in terms of the frequency of uploads. This group of tasks is assumed to be offloaded to the cloud. All these scenarios would be associated with a cloud-only approach (can also be referred to as the baseline) when all the data are transported for processing in the cloud.

6 Experiments and Analysis

This section covers the model setup and performance metrics for evaluating the results. The simulation environment is used to measure the performance of the proposed approach in terms of optimizing energy efficiency, service latency, and network resource utilization.

6.1 Model Setup

As a mathematical model for solving complex optimization tasks, Mixed Integer Linear Programming (MILP) determines the objectives' function within a set of linear constraints and bounds. Therefore, MILP problems are commonly solved using a linear programming-based branch-and-bound algorithm, where only a few variables are integers, while other variables can be non-integers.

MILP is often used for systems analysis and optimization as it presents a flexible and effective method for solving large and complex problems. Also, MILP can be used in many application areas, including but not limited to economics, scheduling, energy system optimization, UAV guidance, and network design problems.

In terms of the experiment environment, the MILP model is solved using the IBM ILOG CPLEX 12.5 optimization solver [33] on a PC with an i7 CPU and 32 GB of RAM. CPLEX provides a high-performance, powerful, and trustworthy mathematical solver to find an optimal solution to the problem. In this work, one cloud server, 1 fog node in each city, 2 edge nodes in each city, and 100 million IoMT sensors across the USA have been used for performance evaluation.

6.2 Performance Metrics

The following presents the proposed edge-fog-cloud system to support IoMT applications and compares it to the existing traditional approach, where data is offloaded only to a central cloud node. The performance of the proposed approach has been assessed by analyzing several metrics, namely, energy efficiency, service latency, and network resource utilization, as shown in the following subsections.

6.2.1 Energy-Efficiency

Figs. 2–4 show the total power consumption of the edge-fog-cloud approach *vs.* cloud-only approach, considering high critical IoMT applications (in Fig. 2), medium critical IoMT applications (in Fig. 3), and all applications combined (in Fig. 4). In these figures, we investigated the power consumption under 10 workloads which ranged between 10% and 100%. Note that different workloads refer to the percentage of devices sending live data to the edge-fog-cloud-architecture (e.g., in the case of high critical IoMT, 10% workloads mean that 6 million devices from a total of 60 million devices are active and sending live data to the edge-fog-cloud-architecture).

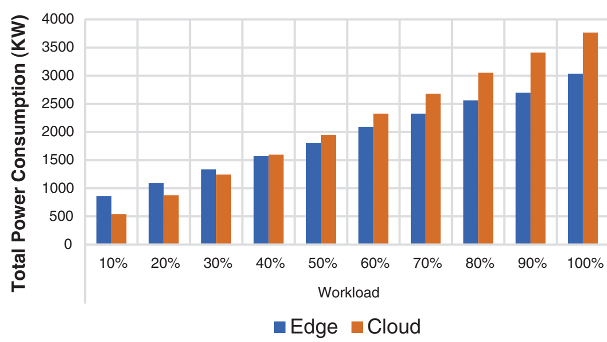

Figure 2: The power consumption of high-critical IoMT applications under different workloads

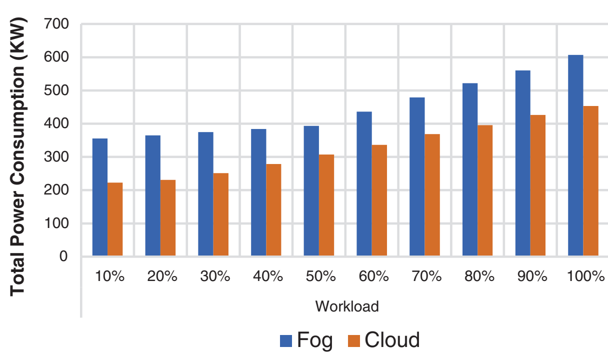

Figure 3: The power consumption of medium-critical IoMT applications under different workloads

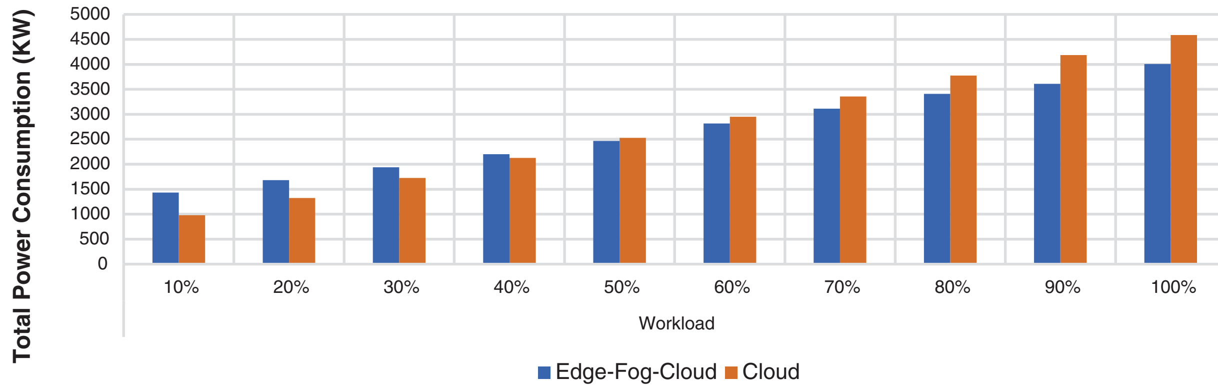

Figure 4: The power consumption of combined IoMT applications under different workloads

In Fig. 2, we observe the power consumption of high critical IoMT applications between 10%–100% workloads. We compare the power consumption of cloud placement *vs.* edge placement. Under 10%–30% workload, cloud placement shows more energy efficiency than edge placement (energy savings are 38%, 20%, and 7%, respectively). While, at workloads 40% and higher, edge placement shows more energy efficiency than cloud placement. Energy savings from placing IoMT applications range between 1% and 21%. The decision to place IoMT applications in edge or cloud is determined by the trade-off between the network power saved by placing VMs in edge closer to where data is created, and the rise in power consumption due to processing requirements that results from replicating VMs to the edge.

In Fig. 3, we investigated the power consumption of medium-critical IoMT applications under different workloads. We compare the power consumption of cloud placement *vs.* fog placement. Due to the lower upload frequency of data from IoMT devices compared to high critical IoMT applications (every 15 min upload rate *vs.* every 1-min upload rate, respectively), cloud placement shows more energy efficiency than fog placement under all data workloads scenarios. The power savings range between 22% and 37% under different scenarios.

In Fig. 4, we observe the power consumption of combined IoMT applications under different workloads. We compare the power consumption of cloud placement *vs.* hybrid edge-fog-cloud placement. After combining the three different IoMT scenarios, we notice the impact of high-critical IoMT applications on optimal placement. Under workloads of 50% and higher, edge-fog-cloud placement shows savings over cloud placement. The total savings range between 2% and 14%. While cloud placement shows power savings of up to 32% under workloads of 50% and lower.

6.2.2 Service Latency

The decision of offloading IoMT live data to either cloud, fog, or edge is highly driven by the criticality of the application. Edge and fog computing process data from IoMT devices, close to the data sources, which highly reduces latency created by core and metro network components. In Fig. 5, we observe the latency of offloading IoMT live data to edge, fog, or cloud computation layers. It can be observed that under 100% workload, the latency of offloading live data to the edge layer provides 47% and 73% less latency than fog and cloud placement, respectively.

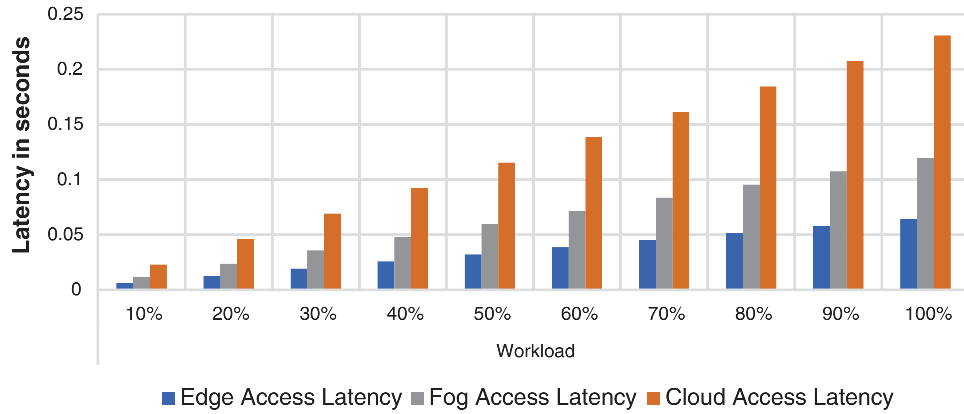

Figure 5: Latency of offloading IoMT data to edge, fog, or cloud computation layers

6.2.3 Network Resource Utilization

Offloading IoMT data to the edge saves high networking space on metro and core networks and provides the IoMT with the same processing capabilities as the cloud layer. Similarly, the fog layer will save traffic from traversing through the core network by exploiting computing resources in the metro network. Fig. 6 illustrates the data offloaded to different computing placement scenarios which traverse the different communication layers.

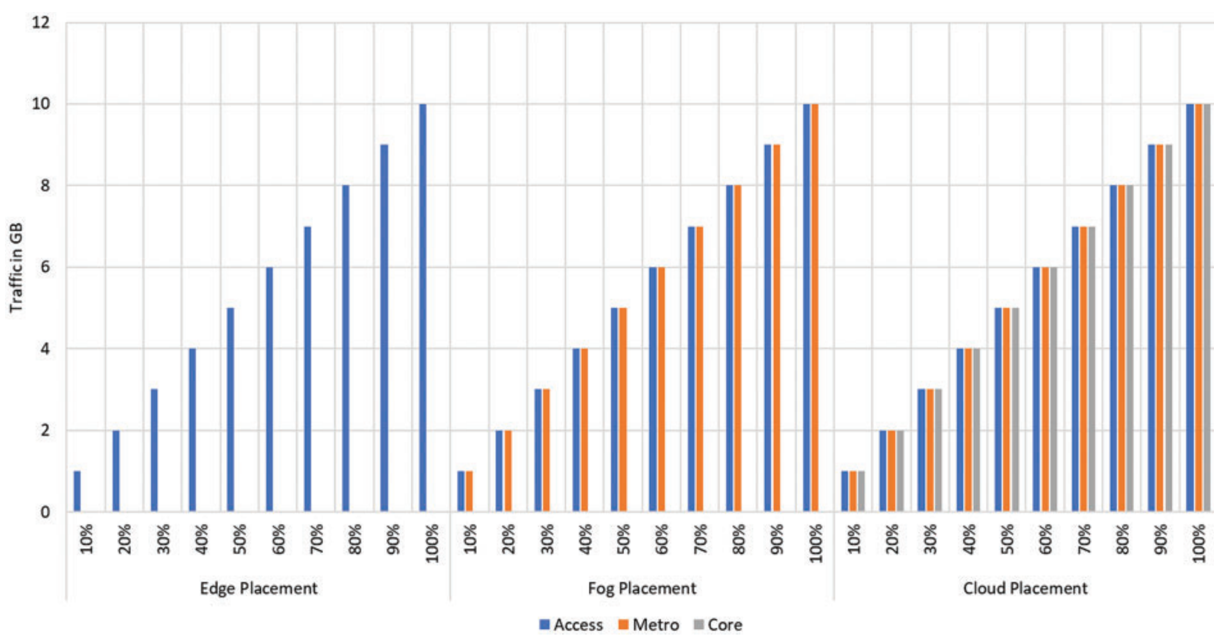

Figure 6: Total traffic traverse through different network layers under the edge, fog, or cloud placements

6.2.4 Discussion and Limitations

There are several challenges when it comes to offloading IoMT applications that require numerous efforts to resolve. This section discusses some of the limitations and challenges associated with offloading IoMT tasks in edge-fog-cloud environments.

- IoMT task dependency: Conducted works on offloading tasks did not consider the dependency between tasks, making them unreliable. Also, this might result in poor quality of service (QoS) for IoMT applications if tasks (that are contingent on the results of other tasks) are allocated to various resources in the edge-fog-cloud architecture [30]. As a result, it is necessary to understand how application components interact with one another. When this factor is considered, overall system performance can be enhanced and the QoS of IoMT applications may be improved.
- A high degree of mobility is required by IoMT applications: Portable medical devices, such as wearable activity trackers, blood pressure cuffs, personal digital assistants, and other mHealth (mobile health) devices, require offloading IoMT tasks to edge, fog, or cloud nodes based on their requirements. For instance, when the IoMT application processes tasks while moving between covered areas in the edge-fog-cloud system, this may result in high network latency or failure of a process. This issue remains a challenge, despite the efforts of several researchers.
- Prediction of IoMT application workload: IoMT tasks are constantly changing, so their procedures may take irregular amounts of time to complete. In some locations, the number of IoMT devices may increase due to the mobility nature of these devices, increasing the workload on connected edge nodes. This may result in dynamic changes in IoMT workload across the edge-fog-cloud system, resulting in service degradation [34]. Consequently, workload prediction modeling is needed to conserve the performance of the edge-fog-cloud system and maintain the QoS of IoMT applications.

7 Conclusion and Future Works

The four-tier IoMT-edge-fog-cloud architecture presented in this work was designed using a Mixed Integer Linear Programming (MILP) mathematical model, which aimed to optimize the end-to-end IoMT architecture. This approach considered multiple performance metrics, such as energy consumption, service latency, and network resource utilization. The results showed that the proposed approach can optimize energy consumption and service latency by up to 21% and 73% respectively, compared to the cloud-only approach. The results also showed that offloading the IoMT application to the edge and fog nodes compared to the cloud is highly dependent on the tradeoff between the network journey time *vs.* the extra power consumed by edge or fog resources. The future scope of this work can be expanded to include the security of IoMT data as well as Artificial Intelligence (AI)-based adaptive modeling. The focus would also be on monitoring and diagnosing patients in remote areas by implementing an IoMT framework on Unmanned Aerial Vehicles (UAVs). Moreover, mathematical models are computationally difficult to solve, hence designing heuristic algorithms to approximate such models is of interest in future works.

Acknowledgement: The authors would like to extend their appreciation to Taibah University for its supervision support. They also would like to thank the Deanship of Scientific Research, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia, for supporting this research.

Funding Statement: The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education in Saudi Arabia for funding this research work the project number (442/204).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Cogniteq, "Internet of medical things (IoMT): Innovative future for healthcare," 2023. [Online]. Available: https://www.cogniteq.com/blog/internet-medical-things-iomt-innovative-future-healthcare-industry. [Accessed: 27-Feb-2023].

[2] R. Dwivedi, D. Mehrotra and S. Chandra, "Potential of internet of medical things (IoMT) applications in building a smart healthcare system: A systematic review," *Journal of Oral Biology and Craniofacial Research*, vol. 12, no. 2, pp. 302–318, 2022.

[3] S. Ghosh and A. Mukherjee, "STROVE: Spatial data infrastructure enabled cloud–fog–edge computing framework for combating COVID-19 pandemic," *Innovations in Systems and Software Engineering*, pp. 1–17, 2022. https://doi.org/10.1007/s11334-022-00458-2

[4] N. Singh and A. K. Das, "Energy-efficient fuzzy data offloading for IoMT," *Computer Networks*, vol. 213, pp. 109127, 2022.

[5] Z. Ashfaq, A. Rafay, R. Mumtaz, S. M. H. Zaidi, H. Saleem *et al.,* "A review of enabling technologies for internet of medical things (IoMT) ecosystem," *Ain Shams Engineering Journal*, vol. 13, no. 4, pp. 101660, 2022.

[6] H. A. Alharbi and M. Aldossary, "Energy-efficient edge-fog-cloud architecture for IoT-based smart agriculture environment," *IEEE Access*, vol. 9, pp. 110480–110492, 2021.

[7] Q. Luo, S. Hu, C. Li, G. Li and W. Shi, "Resource scheduling in edge computing: A survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2131–2165, 2021.

[8] B. U. Demirel, I. A. Bayoumy and M. A. Al Faruque, "Energy-efficient real-time heart monitoring on edge–fog–cloud internet of medical things," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12472–12481, 2022.

[9] A. D. Aguru, E. S. Babu, S. R. Nayak, A. Sethy and A. Verma, "Integrated industrial reference architecture for smart healthcare in internet of things: A systematic investigation," *Algorithms*, vol. 15, no. 9, pp. 309, 2022.

[10] A. Mukherjee, S. Ghosh, A. Behere, S. K. Ghosh and R. Buyya, "Internet of health things (IoHT) for personalized health care using integrated edge-fog-cloud network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 943–959, 2021.

[11] M. L. H. Kazem, "Efficient resource allocation for time-sensitive IoT applications in cloud and fog environments," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 3, pp. 2356–2363, 2019.

[12] A. Shehabi, "United States data center energy usage report|energy technology area," 2023. [Online]. Available: https://eta.lbl.gov/publications/united-states-data-center-energy. [Accessed: 27-Feb-2023].

[13] M. Aldossary, "Multi-layer fog-cloud architecture for optimizing the placement of IoT applications in smart cities," *Computers, Materials & Continua*, vol. 75, no. 1, pp. 633–649, 2023.

[14] X. Wang, L. Wang, Y. Li and K. Gai, "Privacy-aware efficient fine-grained data access control in internet of medical things based fog computing," *IEEE Access*, vol. 6, pp. 47657–47665, 2018.

[15] O. Skarlat, M. Nardelli, S. Schulte, M. Borkowski and P. Leitner, "Optimized IoT service placement in the fog," *Service Oriented Computing and Applications*, vol. 11, no. 4, pp. 427–443, 2017.

[16] A. H. Sodhro and N. Zahid, "AI-Enabled framework for fog computing driven e-healthcare applications," *Sensors*, vol. 21, no. 23, pp. 8039, 2021.

[17] I. Ullah, N. U. Amin, M. A. Khan, H. Khattak and S. Kumari, "An efficient and provable secure certificate-based combined signature, encryption and signcryption scheme for internet of things (IoT) in mobile health (M-health) system," *Journal of Medical Systems*, vol. 45, no. 1, pp. 1–14, 2021.

[18] I. Ullah, S. Zeadally, N. U. Amin, M. Asghar Khan and H. Khattak, "Lightweight and provable secure cross-domain access control scheme for internet of things (IoT) based wireless body area networks (WBAN)," *Microprocessors and Microsystems*, vol. 81, pp. 103477, 2021.

[19] M. A. Khan, S. R. Rehman, M. I. Uddin, S. Nisar, F. Noor *et al.,* "An online-offline certificateless signature scheme for internet of health things," *Journal of Healthcare Engineering*, vol. 2020, pp. 1–10, 2020.

[20] C. Mangla, S. Rani and N. Herencsar, "An energy-efficient and secure framework for IoMT: An application of smart cities," *Sustainable Energy Technologies and Assessments*, vol. 53, pp. 102335, 2022.

[21] S. Singh, A. S. Nandan, G. Sikka, A. Malik and A. Vidyarthi, "A secure energy-efficient routing protocol for disease data transmission using IoMT," *Computers and Electrical Engineering*, vol. 101, pp. 108113, 2022. https://doi.org/10.1016/j.compeleceng.2022.108113

[22] H. Zhou, Z. Zhang, Y. Wu, M. Dong and V. C. M. Leung, "Energy efficient joint computation offloading and service caching for mobile edge computing: A deep reinforcement learning approach," *IEEE Transactions on Green Communications and Networking*, pp. 1, 2022. https://doi.org/10.1109/TGCN.2022.3186403

[23] K. Alatoun, K. Matrouk, M. A. Mohammed, J. Nedoma, R. Martinek *et al.,* "A novel low-latency and energy-efficient task scheduling framework for internet of medical things in an edge fog cloud system," *Sensors*, vol. 22, no. 14, pp. 5327, 2022.

[24] N. Tomar and R. Matam, "Optimal query-processing-node discovery in IoT-fog computing environment," in *Proc. Int. Conf. on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, India, pp. 237–241, 2018.

[25] T. Han, L. Zhang, S. Pirbhulal, W. Wu and V. H. C. de Albuquerque, "A novel cluster head selection technique for edge-computing based IoMT systems," *Computer Networks*, vol. 158, pp. 114–122, 2019.

[26] J. K. Samriya, M. Kumar, M. Ganzha, M. Paprzycki, M. Bolanowski *et al.,* "An energy aware clustering scheme for 5G-enabled edge computing based IoMT framework," in *Proc. Int. Conf. Computational Science (ICCS)*, London, UK, pp. 169–176, 2022.

[27] A. H. Sodhro, M. S. Al-Rakhami, L. Wang, H. Magsi, N. Zahid *et al.,* "Decentralized energy efficient model for data transmission in IoT-based healthcare system," in *Proc. IEEE 93rd Vehicular Technology Conf.*, Helsinki, Finland, pp. 1–5, 2021.

[28] A. Abo-Eleneen, A. A. Abdellatif, A. Mohamed and A. Erbad, "RLENS: RL-based energy-efficient network selection framework for IoMT," in *Proc. Wireless Telecommunications Symp. (WTS)*, Pomona, CA, USA, pp. 1–6, 2022.

[29] M. C. Silva Filho, C. C. Monteiro, P. R. M. Inácio and M. M. Freire, "Approaches for optimizing virtual machine placement and migration in cloud environments: A survey," *Journal of Parallel and Distributed Computing*, vol. 111, pp. 222–250, 2018.

[30] M. Aldossary, "A review of dynamic resource management in cloud computing environments," *Computer Systems Science and Engineering*, vol. 36, no. 3, pp. 461–476, 2021.

[31] Y. Zhang, P. Chowdhury, M. Tornatore and B. Mukherjee, "Energy efficiency in telecom optical networks," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 4, pp. 441–458, 2010.

[32] M. Aldossary and H. A. Alharbi, "An eco-friendly approach for reducing carbon emissions in cloud data centers," *Computers, Materials & Continua*, vol. 72, no. 2, pp. 3175–3193, 2022.

[33] IBM, "Downloading IBM ILOG CPLEX optimization studio V12.9.0," 2023. [Online]. Available: https://www.ibm.com/support/pages/downloading-ibm-ilog-cplex-optimization-studio-v1290. [Accessed: 27-Feb-2023].

[34] M. Aldossary, "A review of energy-related cost issues and prediction models in cloud computing environments," *Computer Systems Science and Engineering*, vol. 36, no. 2, pp. 353–368, 2021.