UNIVERSITI TEKNIKAL MALAYSIA MELAKA

MULTI-OBJECTIVE OPTIMIZATION FOR INTEGRATED TASK SCHEDULING AND LOAD BALANCING IN FOG COMPUTING FOR IOT APPLICATIONS

MAJID RAFIQUE

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

2025

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Faculty of Electrical Technology and Engineering

Multi-objective Optimization for Integrated Task Scheduling and Load Balancing in Fog Computing for IoT Applications

MAJID RAFIQUE

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

2025

Multi-objective Optimization for Integrated Task Scheduling and Load Balancing in Fog Computing for IoT Applications

MAJID RAFIQUE

A thesis submitted
in fulfillment of the requirements for the degree of
Master of Science in Electrical Engineering

Faculty of Electrical Technology and Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2025

DECLARATION

I declare that this thesis, entitled "Multi-objective Optimization for Integrated Task Scheduling and Load Balancing in Fog Computing for IoT Applications" is the result of my research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature for any other degree.

Signature :...

Name :...............Majid Rafique...............

Date :...........23 September 2025...........

APPROVAL

I hereby declare that I have read this thesis and in my opinion, this thesis is sufficient in terms of scope and quality for the award of Master of Science in Electrical Engineering.

Signature : ....

Supervisor Name : Ts. Dr. Nur Ilyana Anwar Apandi

Date : 23 September 2025

DEDICATION

To my beloved family, my SV, and my friends.

ABSTRACT

Fog Computing acts as a bridge between cloud infrastructure and the Internet of Things (IoT) devices, including mobile devices, sensors, and smart technologies. By locating itself closer to edge devices, fog computing enhances data processing. This proximity reduces latency, energy consumption, and communication costs while improving real-time response capabilities. Even though 5G technology has significantly improved connectivity and data transmission rates, the increasing number of IoT devices creates processing challenges. This study focuses on a Fog Optimized Computing System (FOCS) algorithm, which is developed to handle network congestion and processing challenges imposed by the increasing number of IoT devices. By using a task scheduling and offloading method, the FOCS algorithm arranges data according to size and sends it to the appropriate fog nodes. High-capacity fog nodes get large data packets, while lower-capacity nodes receive smaller packets. During times of network congestion, FOCS uses load-balancing mechanisms to ensure that data is transmitted to the closest accessible fog nodes. The FOCS algorithm seeks to improve system performance through the reduction of latency, stabilization of energy consumption and communication costs. By utilizing the Eclipse IDE for implementation and the Cloudsim Toolkit for analysis, the efficiency of the FOCS algorithm will be evaluated, focusing on how effectively it optimizes latency, energy consumption, and cost. FOCS outperforms current techniques in comparative studies conducted under the same simulated conditions, further validating its effectiveness. The results demonstrate that the proposed FOCS algorithm considerably boosts performance through effective task distribution among fog nodes, reduces latency, energy conservation, and utilization cost.

PENGOPTIMUMAN PELBAGAI OBJEKTIF UNTUK PENJADUALAN TUGASAN DAN PENGIMBANGAN BEBAN BERSEPADU DALAM PENGKOMPUTERAN FOG BAGI APLIKASI IOT

ABSTRAK

Pengkomputeran Kabus (Fog Computing) berfungsi sebagai penghubung antara infrastruktur awan dan peranti Internet of Things (IoT), termasuk peranti mudah alih, sensor, serta teknologi pintar. Dengan kedudukan yang lebih hampir kepada peranti tepi, pengkomputeran kabus meningkatkan kecekapan pemprosesan data. Kedekatan ini membolehkan pengurangan kependaman, penggunaan tenaga, dan kos komunikasi di samping memperkukuh keupayaan tindak balas masa nyata. Walaupun teknologi 5G telah membawa peningkatan ketara dari segi kesambungan dan kadar penghantaran data, pertumbuhan pesat bilangan peranti IoT terus menimbulkan cabaran pemprosesan. Kajian ini menumpukan kepada algoritma Fog Optimized Computing System (FOCS) yang dibangunkan untuk menangani kesesakan rangkaian dan cabaran pemprosesan yang berpunca daripada peningkatan jumlah peranti IoT. Melalui pendekatan penjadualan tugasan dan pemindahan beban, algoritma FOCS menyusun data mengikut saiz sebelum mengagihkannya kepada nod kabus yang sesuai; di mana nod berkapasiti tinggi memproses data bersaiz besar manakala nod berkapasiti rendah mengendalikan data bersaiz kecil. Semasa berlakunya kesesakan rangkaian, FOCS menggunakan mekanisme pengimbangan beban bagi memastikan data dihantar ke nod fog terdekat yang boleh dicapai. Algoritma FOCS bertujuan meningkatkan prestasi sistem melalui pengurangan kependaman, penstabilan penggunaan tenaga serta kos komunikasi. Dengan menggunakan Eclipse IDE untuk pelaksanaan dan CloudSim Toolkit untuk analisis, kecekapan algoritma FOCS dinilai dengan menumpukan kepada sejauh mana ia dapat mengoptimumkan kependaman, penggunaan tenaga, dan kos. FOCS terbukti mengatasi teknik semasa dalam kajian perbandingan yang dijalankan di bawah keadaan simulasi yang sama, sekali gus mengesahkan keberkesanannya. Hasil kajian menunjukkan bahawa algoritma FOCS yang dicadangkan berjaya meningkatkan prestasi secara signifikan melalui pengagihan tugas yang berkesan antara nod fog, pengurangan kependaman, penjimatan tenaga, serta pengurangan kos penggunaan sumber.

ACKNOWLEDGEMENT

Starting with the Name of Allah, the most Kind and the Most Merciful. First, I would like to extend my sincere thanks to my supervisor, Ts. Dr. Nur Ilyana Anwar Apandi, and my Co-supervisor, Ir. Dr. Zamani Bin Md. Sani, for all the valuable guidance, encouragement and constructive feedback at every stage of this research. I gratefully acknowledge the financial support by the Ministry of Higher Education Malaysia (MOHE) under the Fundamental Research Grant Scheme (FRGS), grant number FRGS/1/2022/FKE/F00502. I am very thankful to the Faculty of Electrical Technology and Engineering, Centre for Research and Innovation Management (CRIM) and specifically PVSG Lab, whose assistance helped create a conductive environment for my research. My heartfelt appreciation goes to my family and friends for their support, patience and motivation during all the challenging phases of this whole journey. Finally, I acknowledge all the researchers whose work contributed to this thesis and inspired the direction of my study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

UTeM	-	Universiti Teknikal Malaysia Melaka
IoT	-	Internet of Things
5G	-	Fifth Generation
FC	-	Fog Computing
FOCS	-	Fog Optimized Computing System
IDE	-	Integrated Development Environment
CC	-	Cloud Computing
RQ	-	Research Question
RO	-	Research Objective
IBM	-	International Business Machine Corporation
SaaS	-	Software as a Service
IaaS	-	Infrastructure as a Service
PaaS	-	Platform as a Service
EC	-	Edge Computing
BS	-	Base Station
TS	-	Task Scheduling
PSO	-	Particle Swarm Optimization
GA	-	Genetic Algorithms
FCFS	-	First-Come-First-Serve
SDN	-	Software-Defined Networks
TO	-	Task Offloading
DRL	-	Deep Reinforcement Learning

DL	-	Deep Learning
FN	-	Fog Node
MIPS	-	Million instructions per second
LB	-	Load Balancing
QoS	-	Quality of Service
VNF	-	Virtual Network Function
SJF	-	Shortest Job First
OFA	-	Optimal Fog Algorithm
ms	-	milliseconds
kJ	-	kilo joules

LIST OF SYMBOLS

I	-	Set of tasks from the IoT devices
J	-	Set of fog nodes
F_j	-	jth fog node
i	-	number of tasks
j	-	number of fog nodes
T^{up}	-	Uploading time of the task
T^{proc}	-	Processing time of the task
T^{down}	-	Downloading time of the task back to the end devices
D_i	-	Data size of the ith task
R_{ij}^{up}	-	Uploading rate of the task
R_{ij}^{proc}	-	Processing rate of the task
R_{ij}^{down}	-	Downloading rate of the task
L^{Total}	-	Total latency
E_c	-	Current consumed energy
E^{up}	-	Energy consumption at uploading a task
E^{proc}	-	Energy consumption at processing a task
E^{down}	-	Energy consumption at downloading a task
E^{Total}	-	Total Consumed Energy
P_j^{up}	-	Power required by the fog nodes during uploading
P_j^{proc}	-	Power required by the fog nodes during processing
P_j^{down}	-	Power required by the fog nodes during downloading

$MIPS_A$	-	Allocated MIPS in the fog nodes
$MIPS_T$	-	Total MIPS in the fog nodes
$MIPS_R$	-	Remaining MIPS in the fog nodes
C_c	-	Current Cost (past value)
C^{Total}	-	Total Utilization Cost
R_{MIPS}	-	Rate per MIPS
F_c	-	Fog node capacity
F_l	-	Load on the fog node (fog load)
O_l	-	Overload

LIST OF APPENDICES

LIST OF PUBLICATIONS

The following is a list of publications related to the work on this thesis:

Rafique, M., Anwar Apandi, N. I., Nor Azmi, S. N. L. K., Md. Sani, Z., Muhammad, N. A., & Sanaullah, S. (2025). Joint Task Scheduling and Offloading in Fog Optimized Computing System (FOCS) Algorithm for IoT based Network Applications. ASEAN Engineering Journal, 15(3), 167–174.

CHAPTER 1

INTRODUCTION

1.1 Background

Internet of Things (IoT) is a major technological advancement. It connects smart objects and sensors to the internet, allowing them to gather data from their environment, which can then be stored in the cloud and processed automatically or on demand in real-time. Cisco says that there will be more than 50 billion Internet-connected devices soon, with about seven devices per person. John T. Chambers, former CEO of Cisco, thinks that there will be a whopping 500 billion devices connected to the Internet by 2025 (Jonathan Camhi, 2015). Expanding the capabilities and reach of IoT devices has been made possible by the resource-rich environment of Cloud Computing (CC). Through cloud-based resource allocation and administration, customers can assign tasks related to hardware and software. Despite the advantages of CC, there are some issues with IoT applications that have not yet been fixed.

From the IoT environment, a large volume of data is generated in each instant from sensors, messaging systems, mobile phones, and social media; hence, the various technical challenges revolve around the IoT environment, due to its distributed, complicated, and dynamic character. These challenges include latency, connectivity, capacity, cost, power, scalability, and reliability. Typically, the use of the Cloud is able for processing data that is created in the IoT context. However, there are circumstances under which it is not feasible to transfer all created data to the Cloud for processing and storage. Transferring all created data to the Cloud consumes a certain amount of network bandwidth. The Cloud infrastructure, though, is not able to handle latency-demanding applications because of the comparatively high response times needed by these time-sensitive applications (Vermesan and Friess, 2014).

In 2012, Cisco introduced the idea of Fog Computing (FC), which involves processing some workload and services locally on fog devices like routers and switches, rather than sending everything to the cloud. FC, sometimes referred to as fog networking or fogging, is the most complete and appropriate approach to increase the efficiency of data transferred to the cloud. Fog can be defined as "cloud closer to the ground," which provides an understanding of how it works (Bonomi et al., 2012). It lies between the device and cloud layer, as shown in Figure 1.1.

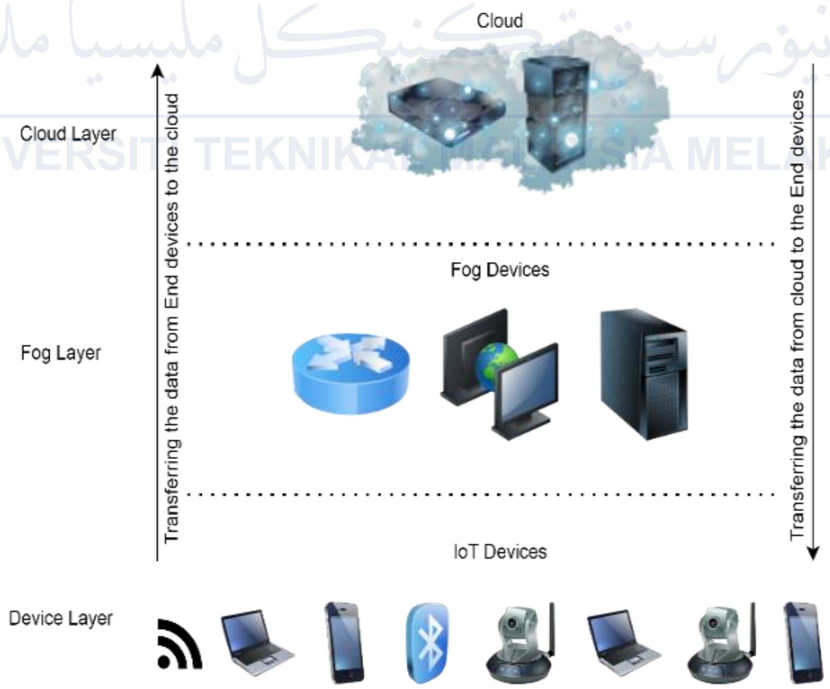

Figure 1.1 Fog Computing Architecture.

FC is a decentralized computing infrastructure that seeks to facilitate data management, networking, storage, and computation not only within the cloud but also across the cloud-to-IoT devices path as data moves from the network to the cloud (Bonomi et al., 2012; Abdali et al., 2021).

It encompasses all the intermediate computing nodes and devices. After processing the data streams originating from sensors on IoT devices in the device layer, the fog layer initiates requests from actuators for appropriate actions for tasks. The end devices are responsible for actuation and sensing activities. Non-time-sensitive processing can take place on the cloud layer, whereas time-sensitive applications should process on the fog layer. By deciding what information should and shouldn't be delivered to the cloud, the fog layer serves as a gatekeeper. End users can obtain resources from the cloud and fog layers according to their needs.

1.2 Problem Statement

The rapid growth of IoT sensors and devices in various applications, including Healthcare, Smart Vehicles, Smart Cities, Smart Industries, generates vast amount of data, leading to delayed decision-making due to high latency, energy consumption and cost (Rahbari and Nickray, 2019; Al-khafajiy et al., 2019; Jamil et al., 2020). When these IoT devices communicate to the cloud over long distances, it takes too long to get a response. Additionally, the huge amount of data from large number of IoT devices cause high energy consumption, and, as a result, the utilization cost also become high.

Fog computing offers great research potential, especially in the areas of load balancing and task scheduling, which are most prominent in its successful deployment and functioning. However, efficient task scheduling and fair distribution of workloads is a challenging issues because of the constrained computation resources of fog nodes and the low-latency demands of end-user applications. Improper scheduling mechanisms lead to underutilization or overloading of resources, whereas improper load balancing leads to higher latency and lower system reliability. All these problems need to be resolved to realize the potential of fog computing in dynamic and heterogeneous environments. The researchers have developed many algorithms and strategies to optimize, but it is still hard to manage and optimize most of the challenges at once.

1.3 Research Questions

i) How to optimize the processing challenges of the IoT-based network applications by FC?

ii) What is an optimized solution in FC through joint task scheduling and load balancing?

iii) How much will the impact of the proposed system be improved in terms of latency, energy, and cost performance?

1.4 Research Objectives

The main objectives of this research are as follows:

i) To investigate the approaches and algorithms in Fog Computing for IoT-based network applications. (RQ1)

ii) To develop a Fog Optimized Computing System (FOCS) algorithm based on a joint task scheduling and load balancing simulation approach. (RQ2)

iii) To analyze the proposed FOCS algorithm in terms of latency, energy consumption, and utilization cost in network applications. (RQ3)

1.5 Scope of Research

This study provides an optimization solution for IoT networks between the IoT devices and the cloud with less energy consumption and utilization cost, with better performance.

The focus of this work is the analysis and comparison of current algorithms and approaches in FC environments. The scope of the analysis is restricted to algorithmic solutions for performance metrics that are latency, energy consumption, and utilization cost in dynamic heterogeneous fog infrastructures. Emphasis is given to gap identification in current methodologies and comparison of shortlisted algorithms on the basis of input parameters and performance metrics in simulated fog environments.

The study assumes the computation is done within the processor layer with perfect operation at the hardware-level architectures, fog device setups, or security issues in fog computing. Cloud and edge computing are addressed only for comparison. The evaluation primarily comprises theoretical and simulation-based analysis with emphasis on operational efficiency of algorithms under dynamic workloads and fog network topologies.

1.6 Thesis Outline

In line with the intended objectives and the research strategy adopted, this thesis is structured in five chapters, which are briefly elaborated as follows: