Dynamic Fog-Cloud Task Allocation Strategy for Smart City Applications

Mikail Mohammed Salim*, Jungho Kang**, Jong Hyuk Park*
*Department of Computer Science and Engineering, Seoul National University of Science and Technology, Seoul
** Department of Information Protection, Baewha Women's University, Seoul
mikail@seoultech.ac.kr, kangsamm@baewha.ac.kr, jhpark1@ seoultech.ac.kr

Abstract

Smart cities collect data from thousands of IoT-based sensor devices for intelligent application-based services. Centralized cloud servers support application tasks with higher computation resources but introduce network latency. Fog layer-based data centers bring data processing at the edge, but fewer available computation resources and poor task allocation strategy prevent real-time data analysis. In this paper, tasks generated from devices are distributed as high resource and low resource intensity tasks. The novelty of this research lies in deploying a virtual node assigned to each cluster of IoT sensor machines serving a joint application. The node allocates tasks based on the task intensity to either cloud-computing or fog computing resources. The proposed Task Allocation Strategy provides seamless allocation of jobs based on process requirements.

1. Introduction

The rapid expansion of the urban population is expected to reach 9.7 billion people [1]. The rapid development of smart cities worldwide requires efficient real-time management of data requests from millions of sensors deployed to support the growing population. The expansion of smart cities enabled critical and intelligent applications such as Healthcare, Transportation, Traffic Management, and Smart Grid [2]. Traditionally, data computation tasks are offloaded to cloud networks due to the availability of high computing resources for managing complex tasks. Complicated and time-critical applications are constrained by limited bandwidth and high latency and are therefore unsuitable for future wireless networks. The Fog-Edge computing environment enables intelligent applications to offload time constraint tasks for quick and real-time results. The introduction of the Fog-Edge computing paradigm enables support for the complex architecture of all future smart cities within the provided bandwidth capabilities [3].

Task Scheduling facilitates smart city networks to allocate tasks to Fog-Edge-Cloud based on the requirements of the computational job. An inefficient task offloading strategy results in mismanagement of available resources where local data centers are crowded with tasks queues resulting in increased latency and thus affecting the overall expected Quality of Service. A practical and robust task scheduling strategy scheme analyzes the available resources, each task's requirements and distributes them for computational operations at either the cloud or Fog-Edge-based datacenters.

Recent research on efficient task scheduling strategies focuses on optimizing task allocation for increasing device data and expanding intelligent applications' requirements at the Fog-Edge layer. Adhikari et al. [4, 14] proposed a task scheduling strategy based on a two-layer algorithm for a combined path planning optimization and efficient task scheduling. The study's objective is to minimize the task allocation time at the local edge datacenter before the previous job completes its computational operation. Sharma et al. [5, 15] present a neural network-based task scheduler for cloud computing environments. The study analyzes and predicts the ideal computing resource each job requires before allocating it to the cloud resource. Ding et al. [6, 17] proposed a task scheduler framework where individual assigned jobs are based on task priority and task life. A reward-based function remunerates the scheduler for allotting each task a virtual machine for processing with the least waiting time. Wu et al. [7, 20] presented a task scheduling algorithm that allocated tasks based on their requirements on the Cloud, Edge, or the local IoT devices. An optimization technique depends on each job's communication and computation tasks and assigns them accordingly on the Cloud-Edge-IoT layer.

In this paper, the contributions of the proposed Task Allocation Strategy (TAS) are as follows,

1. A virtual node is allocated to each intelligent application responsible for building a profile of each application process based on their requirements.
2. Each processing job is identified based on task intensity, which is dependent on the computation resources required to perform the job and the delay tolerance value.
3. Task intensity-based division of jobs enables optimum allocation of cloud and Fog-Edge resources.

The remainder of this paper is as follows. Section II presents the proposed Task Allocation Strategy. Finally, Section III concludes the paper and suggests future directions of this research.

2. Proposed Task Allocation Strategy

The objective of this research is to present an efficient TAS for an IoT-supported Smart City network. In this research, the proposed virtual node-based TAS includes a smart city network supporting several intelligent applications, such as Healthcare, Transportation, Parking, Manufacturing, and Logistics. Each application is divided into a cluster that includes several IoT-based sensor devices, and a set of virtual nodes are assigned per cluster. Two entities that perform task computation are the Core Cloud and Fog layer-based datacenters.

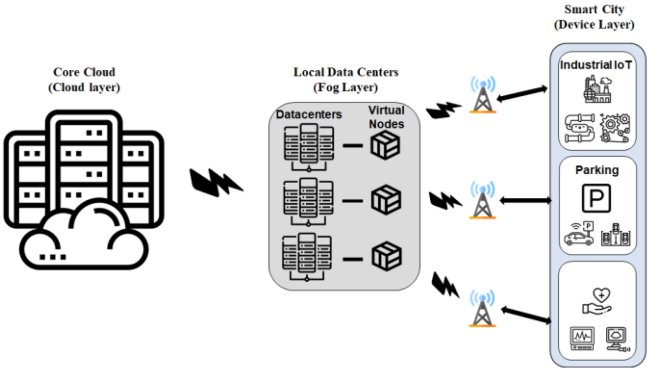

Figure 1. Task Allocation Strategy Overview

Figure 1 illustrates the overview of the proposed TAS where The IoT devices are part of the Device Layer. A fog computing-based server responsible for data computation is placed at the Fog layer. The third layer, the Cloud layer, performs task operations that require a higher amount of resources. Data computation is assigned and delegated to the Fog or Cloud layer based on the resources required and the acceptable delay tolerance. Virtual nodes exist at the Fog layer that behaves as devices representatives to ensure the most optimized network performance for intelligent applications. Each application has several processes which are part of its daily sub-routine to complete the task. Each task is allocated a task intensity level which indicates the urgency required to complete the process. In this research, we study the role of a single virtual node managing a select few IoT device data computation requirements for a smart application.

The Virtual nodes have the following objectives in this research:

1. Monitor all task requests received from the intelligent applications and IoT devices.
2. Log all processes included in the application and build a profile for each process. Each profile consists of the expected data offloading required to perform the task, the expected CPU resources to complete the job, and the task intensity of the process.
3. Divide each task based on their Task intensity rank. A High Task Intensity (HTI) ranking process implies that the expected data offloading is large, requires high CPU resources, and is a delay-tolerant task. A Low Task Intensity (LTI) refers to processes that require fewer CPU resources to process, have reduced data offloading size, and are not delay tolerant.
4. Assign each HTI task for computation at the core cloud network and LTI tasks at the local Fog layer-based data centers.

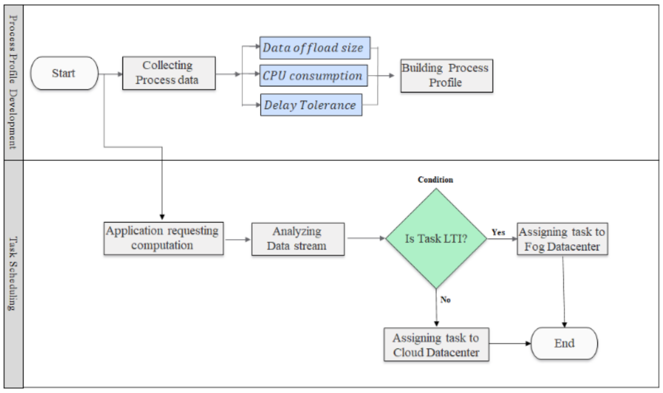

Figure 2. Task Allocation Strategy Workflow

The TAS scheme workflow implements a 2-phase approach, Process Profile development and Task Scheduling. In the first phase, the Process Profile development workflow is as follows:

1. The virtual node builds a profile for each individual process part of the application (a) when it is first initialized. Here each application process is represented as (p_i) and a set of several processes in an application are represented as,

$$a = \{p_1, p_2, p_3 \dots \dots p_n\}$$

2. Each p_i is a profile that consists of offloading data size (s_i), expected CPU consumption (c_i), and delay tolerance level (d_i). The virtual nodes build each process profile which is represented as,

$$p_i = \{s_i, c_i, d_i\}$$

In the second phase, Task Scheduling, all application query requests are sent to the local fog-based virtual node for processing before being forwarded to the local fog-based or cloud-based datacenters for computation. The Task Scheduling process flow is as follows:

1. Each data stream transmitted from IoT devices for a common process-based task request is considered part of a single application request.
2. The application process computation request requires the data size of the task, the CPU resources needed, and the delay tolerance level.
3. A task with low delay tolerance, lower CPU resources required, and fewer data offloading size are termed as LTI. The TAS identifies them as processes requiring immediate real-time results with minimum latency

tolerance. Therefore the local fog-based data centers are assigned for computation operations.

4. The TAS schedules tasks requiring high CPU resources, high data offload size, and high delay tolerance as HTI jobs and schedules them for computation operations at the cloud datacenter.

Results from the computation are transmitted from the Datacenter to the application. The existence of a pre-identified profile of each application process enables a seamless transfer of tasks to the appropriate resource. As task requests arrive, the TAS analyzes the p_i of each computation request and schedules the computation task.

3. Conclusion

In this paper, an efficient Task Allocation Strategy is proposed for IoT-enabled smart cities. A profile for each application process is designed to enable a seamless transfer of tasks to the appropriate computation resource. LTI jobs are assigned to Fog-based data centers due to their low delay tolerance, and HTI jobs are delegated to the core cloud computing resources.

Our future work is to extend the TSA further to include Blockchain for safeguarding the selected computation processing strategy from cyberattacks. Furthermore, cloud and local Fog-based datacenters profiles would measure which computing resource has an empty queue to process the next task. Finally, a complete evaluation of the TSA will be based on measuring the computation cost and energy efficiency compared to existing studies.

Acknowledgement

This work was supported by the Survey and analysis of quantum information technology trends and regional strategy establishment research services project funded by the Gangwon Technopark.

References

[1] Anna J. Benites and Andre F. Simoes "Assessing the urban sustainable development strategy: An application of a smart city services sustainability taxonomy" Ecological Indicators, vol. 127, pp. 107734, 2021.

[2] Abdul R. Javed, Muhammad U. Sarwar, Mirza O. Beg, Muhammad Asim, Thar Baker and Hissam Tawfik "A collaborative healthcare framework for shared healthcare plan with ambient intelligence" Human-centric Computing and Information Sciences, vol.10, pp. 1-10, 2020.

[3] Hamoud Alshammari, Sameh A. El-Ghany and Abdulazziz Shehab "Big IoT Healthcare Data Analytics Framework Based on Fog and Cloud Computing" Journal of Information Processing Systems, vol. 16, no.6,pp. 1238-1259, 2020.

[4] Rong Cong, Zhiwei Zhao, Geyong Min, Chenyuan Feng and Yuhong Jiang "EdgeGO: A Mobile Resource-sharing Framework for 6G Edge Computing in Massive IoT Systems" IEEE Internet of Things Journal, pp. 1-9, 2021.

[5] Mohan Sharma and Ritu Garg "An artificial neural network based approach for energy efficient task scheduling in cloud data centers" Sustainable Computing: Informatics and Systems, vol. 26, pp. 100272, 2020.

[6] Ding Ding, Xiaocong Fan, Yihuan Zhao, Kaixuan Kang, Qian Yin and Jing Zeng "Q-learning based dynamic task scheduling for energy-efficient cloud computing" Future Generation Computer Systems, vol. 108, pp. 361-371, 2020.

[7] Huaming Wu, Katinka Wolter, Pengfei Jiao, Yingjun Deng, Yubin Zhao and Minxian Xu "EEDTO: An Energy-Efficient Dynamic Task Offloading Algorithm for Blockchain-Enabled IoT-Edge-Cloud Orchestrated Computing" IEEE Internet of Things Journal, vol. 8, no. 4, pp. 2163-2176, 2021.