EasyChair Preprint

№ 7002

A Survey on Mobile Edge Computing: Joint Offloading and Resource Allocation Perspective

K Chakravarthy Antharaju, T Rama Reddy and
N Ramakrishnaiah

November 7, 2021

A survey on Mobile Edge Computing: joint offloading and resource allocation perspective

Antharaju K Chakravarthy
Aditya Engineering College, Surampalem
PIN 533437, Andhra Pradesh, India
Email:kalyan0501@gmail.com

Rama Reddy T
Aditya Engineering College ,Surampalem
PIN:533437, Andhra Pradesh, India
Email:ramatreddy@gmail.com

N Ramakrishnaiah
University College of Engineering (Autonomous),
Jawaharlal Nehru Technological University,
Kakinada PIN:533003, Andhra Pradesh, India.
Email:nrkrishna27@gmail.com

Abstract:

The exponential growth of distributed IoT devices and mobile applications are the driving forces for the smart world solutions. Cloud, Cloudlets and Mobile Edge Computing are the promising solutions for ubiquitous and on demand user applications at the Edge device. Edge computing is the current industry requirement to provide low latency services. Offloading techniques play a phenomenal role in Edge computing environment in terms of dynamic decision making on task execution location that depends on the delay, size of the task and, many other factors. Hence, offloading strategies in mobile edge computing is a vast research area which needs more focus. This paper focused on various offloading strategies and their comparison as this study gives a valuable start to fellow researchers who are interested in offloading techniques in mobile edge computing.

Keywords: Mobile Edge Computing, Offloading, Resource allocation, Internet of things.

I. Introduction

From the last decade, usage of smart devices and cellular networks increased exponentially, due to the enormous growth of sensors, smart devices and mobile device usage is a clear indication of the mobile IoT trend, which brings data, processes and people together to make human life with more comfort and valuable. According to CISCO report [1], it is estimated that the number of connected devices to IP networks will be more than three times the global population by 2023. The share of machine-to-machine connections will grow from 33 percentage in 2018 to 50 percentage in 2023. It is estimated that around the world there will be 14.7 billion Machine-to-machine connections by 2023. This gives us a new sign for wide number of business opportunities and challenges to the technical world. The sensor's prime objective is to generate data in the IoT environment, processing that data is a challenging operation at cloud data centre [19]. Massive growth of the data from those devices leads to computational and communicational bottleneck in device level and network level respectively. Applications like Autonomous vehicles and Aerial Unmanned vehicles need very low delay sensitivity. All these applications are emerged with lower delay and higher computing efficiency.

In the recent years there is a wide range of approaches available to fulfil the above requirement such as Cloud computing(CC), Mobile Cloud computing(MCC), cloudlets and Mobile Edge Computing(MEC). The conventional Cloud Computing solves the problem partially. In the cloud environment data needs to be processed at the centralized server, which is suffering with the delay sensitive applications. Even though there are enormous studies on the cloud computing still a research gap exists at this juncture. MCC came into the picture to come up with a virtualization concept, but that suffers with the offloading problem. MEC is the promising technology to address the above research gaps. Computing at the network's edge is referred to as MEC. MEC applications runs in Radio Access Networks (RAN) environment. MEC is defined by ETSI as any network site where computation, other resources, and services are available closer to the user than in a central data centre or cloud. Certain base stations, such as cell towers, data centres, Wi-Fi routers, and hotspots, serve as the network's edge. The edge's proximity decreases latency to milliseconds and ensures that users have a steady connection.The below prominent technologies are significant at delay sensitive application .

1.1 Various Technologies

A. Cloud Computing:

Cloud Computing technology enables run computations in a centralized data centre. Recent years cloud is the most desirable technology which addresses most of the problems. Domestic users and industrial users are benefited from the cloud properties such as availability and ubiquitous in nature. Cloud suffers from delay and resource allocation issue in emerging IoT environment, this leads to a scope for the MEC technology to come in front.

B. Cloudlets :

Cloudlets are well known for virtualization concept. The fundamental concept in cloudlets is, bringing the cloud services to the edge

network level by virtualizing the cloud data centre. Cloudlets are efficient at cloud applications by bringing the cloud services close to the user. Furthermore, cloudlets are a sub-optimal cloud solution since advanced applications such as autonomous vehicles and unmanned vehicles suffer from latency concerns.

C. Mobile Cloud Computing :

Cloud computing and mobile computing are combined in MCC. The concept behind this technology is to create a virtualized isolated environment in the cloud with various resources such as computing, storage, and communication that end users can access remotely. The MCC approach enables the app developer to build applications designed especially for mobile user, which can be used without being bound to the device Operating System or the storage capacity.

D. Mobile Edge Computing (MEC)

Most of the cloud stakeholders are very much concern on the following accepts such as network security, emergency efficiency, network manageability and bandwidth utilization. MEC address most of the above challenges within the radio access networks. The basic idea behind MEC is to bring the resources to RAN. This technique improves the communication and computational capabilities for delay sensitive applications.

1.2 Literature Study

MEC has emerged as a phenomenal technology in the most of the delay sensitive applications. There are numerous studies found in this field, few of them presented as below.

Quinqin et al[2] worked on Fog computing based IoT environment, they tried to address decentralized computation offloading. Fog computing environment can provide higher quality of experience (QoE) and battery life by introducing Energy harvesting [EH] technique. IoT devices make decisions in the suggested environment by using a Decentralised Partially Observable Markov Decision Process (Dec-POMDP). As this Dec-POMDP is a complex process, a learning based algorithm is proposed to simplify this process. This process needs an optimization algorithm to increase the performance. The optimal solution for the stated problem is found using a Lagrangian approach and the policy gradient method. It is observed that remaining pending tasks, battery life time and available fog resources are the factors which affect offloading technique.

Yang sun et al[3] In the MEC context, the authors has shed light on effective and efficient job assignment and computation offloading methodologies. With the Rapid development of mobile devices in recent years, multimedia services such as multimedia content delivery, video calling are the most significant areas that need to be addressed in wireless networks. In this paper the task offloading problem is divided in to several sub problems by using Lyapunov optimization technique, Stochastic optimization algorithms provide an alternative approach that permits less optimal local decisions to be made within the search procedure that may increase the probability of the procedure locating the global optima of the objective function.

Ishtiaque et al[4] focused on security breaches in MEC environment. There is a scope of malicious attacks and eavesdropping in multi user environment especially when they offload the task. Therefore assignment of security services in task offloading is essential. This problem is solved by design an resource constrained optimization model and jointly optimize energy consumption and security breach cost while maintaining delay condition. In the case of large number of servers involved in the computation delay calculation system complexity will increase, to solve this bottleneck a two-stage heuristic solution is implemented to overcome this limitation and achieve an acceptable solution within a reasonable time limit. According to numerical results 60% security breach cost reduced and 15% of energy consumption is reduced by comparing the current MEC technique. Mobility and incentive mechanisms are proposed as a future work for this study.

Zhiyong et al[5] concentrated offloading problem in MEC by introducing MNLP optimization technique. Mixed-integer Non Linear Programming (MNLP) addresses a very general class of optimization problems with nonlinearities in the objective and/or constraints as well as continuous and integer variables. Channel allocation, transmission power and computational resource allocation problems also addressed by subdividing original problem in to two sub problems. Intelligent algorithms, quasi-convex and convex optimization technology used as key algorithms to address the above sub problems.

Yinghao et al[6] studied on computational offloading strategy in single user multiple antennas where data transmission applications are come into the picture. Problem statement figured In MEC environment and multiple antennas are considered. In this study they proposed intelligent offloading strategy for MEC networks assisted by array signal processing. Array signal processing is centred on the ability to fuse temporal and spatial information captured via sampling signals emitted from a number of sources at the sensors of an array in order to carry out a specific estimation task. Two kinds of array signal processing are employed i.e., Maximum Ratio Transmission (MRT) and Selection Combining (SC).

Hong et al[7] worked on offloading technique in small cell MEC environment. In a small cell, mobile devices attached to small base stations, these base stations are attached to Micro Base Station (MBS). Small cell mobile computation offloading decision depends on two factors, i.e., communication and computation resources. Small base stations collaboratively take offloading decision, this is a complex task at the network level. First problem is formulated as a mixed binary non liner programming problem then converted as General Binder Decomposition (GBD). A Heuristic solution is proposed in order to invoke more MBS to involve in decision making.

Kehao et al[8] focused of MEC with 5G network, which gives a great breakthrough in transmission time and execution delay. In the practical scenario MEC has limited resources, in order to efficient utilization of the resources we need an optimized computation offloading, resource allocation and data cache services. The proposed problem is NP-hard since it is formulated as a Mixed Integer Programming (MIP) problem. The original problem was broken into two sub-problems: downlink resource allocation combined offloading decision and computation resource allocation. These two sub problems can be solved with the help of convex and non-convex optimization techniques. Convex optimization problems are more generic than linear programming problems, although they share some of the same desired characteristics: Up to a very larger frame, they can be addressed rapidly and consistently. Proposed iterative algorithm gives better results than the bench mark algorithms.

Xinchen et al[9] worked on MEC offloading techniques along with energy minimization technique. High reliability and low latency are the desirable features in IoT environment, MEC will address the above issue by introducing verities of offloading techniques, this may give phenomenal improvement in storage and computational limitations. In this study they proposed a three tire architecture by incorporating MEC, Cloud and IoT devices. A Light weight request and admission control scalability issues. Various parameters like input size, required cpu cycles, uplink and downlink capabilities, offloading decisions varies based on these parameters.

Long et al[10] proposed a joint offloading and resource allocation strategies. Age of Information [AoI] is much significant in decision making on video surveillance applications. Mobile edge computing addresses the above issue by offloading the computational intensive task to nearby edge server. In the MEC environment it is possible to solve the AoI by optimize task generation, offloading, communication and computation resource allocation. To handle the above issues it is necessary to convert Lyapunov optimization technique to the long-term stochastic optimization problem into a per-time slot deterministic optimization problem. Based on the numerical results this algorithm reaches close to the optimal results.

Nahida Kiran et al [11] focused on Offloading and resource allocation problem in wireless Mobile Edge computing environment. By introducing SDN in MEC, they tried to increase QoS, solve offloading and resource allocation problem with the help of Reinforcement Learning (RL). According to their study Machine Learning models play significant role in delay prediction on large amount of data. Two major Machine learning algorithms are observed here such as Reinforcement algorithm and Cooperative Q learning methods. A Q-Learning algorithm is a value-based reinforcement learning algorithm that uses a Q function to discover the best action-selection policy.

1.3 Organization of the paper

The paper organized as follows. Section II addresses MEC architecture. Section III describes need of offloading. Section IV addresses types of offloading. Section V focuses on Problem formulation and optimization techniques. Section VI addresses various parameters which affects offloading. Section VII describes various simulators and Section VIII covers conclusion.

II. MEC Architecture

Figure: 1 describes a typical MEC architecture. MEC servers are placed close to the user equipment. In this scenario, edge devices are connected to local base station or a Wi-Fi router. With the advancement of 5G technology and Edge computing it is possible to fill the communication gap between heterogeneous devices. In this case MEC servers play a vital role to coordinate edge devices. Generated data at edge level being processed at MEC server with no time delay. This enables MEC to support delay intensity applications such as autonomous vehicles and IoT applications. This paradigm can be applicable to most of the delay effective applications.

The edge devices are connected to local Access Point (AP) or BTS (Base Transceiver systems). BTS are equipped with edge server to run user application. The advantage of the below architecture is to reduce the burden on the cloud by pre-processing the data at the edge server.

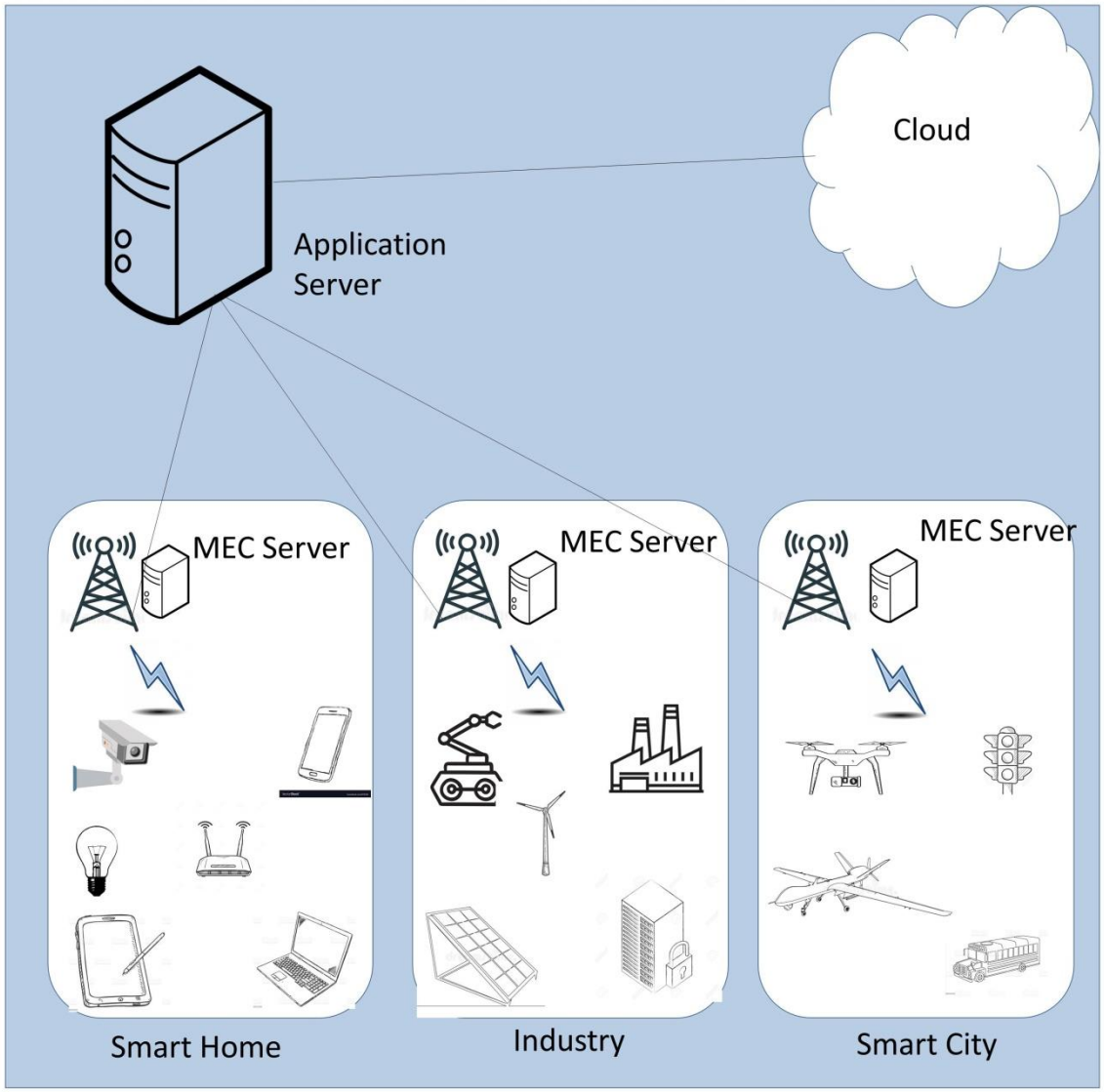

Figure 1 : Illusion of Mobile Edge Computing

III. Need of Offloading

Future world will be densities with IoT and smart devices, this situation leads to a tremendous load on cloud with respect to computation and communication accepts. Especially delay sensitive applications needed more attention such as autonomous cars, AR&VR applications, gaming and Arial unmanned vehicles. At this juncture it is necessary to reduce burden on cloud by introducing offloading technique at edge devices. The basic approach in offloading depends on various factors such as current network capacity, Computation(Size and number of pending tasks), Transmission, Traffic, Power-delay Trade off, Energy, Communication, Caching, Number of available edge resources, Computational resource allocation, Channel allocation and many more. Depends on the above factor an efficient offloading technique need to decide where to send computational intensive task ,either to nearby edge server Or cloud data centre, so that result may not be effected with the delay factors. There are wide range of approaches are available to address the above issue discussed as below.

IV. Types of Offloading

1. Centralized offloading

Based on the below literature offloading can be classified as Centralized and distributed. Most of the cloud computing applications follow centralized offloading technique by executing business logic at the centralized data centre. There are numerous kick stoppers can be found in this approach among that Network latency and security are the most desirable issues.

2. Distributed offloading

Introducing distributed mechanism in cloud environment we can overcame the above problem. MEC is most well known for a distributed mechanism where data pre processing takes place at the data

generation point only by introducing distributed environment in MEC. Offloading decision process executes at mobile edge servers which respond with no time delay due to its proximity from the edge node.

V. Problem Formulation and Optimization techniques

Problem formulation is a key factor to find optimal solution in offloading. There are different researchers adopt different techniques but it is observed that most of the researcher convinced as problem is NP-hard, Game based approach , Convex and non-convex problems. All the above problems that are solved in a polynomial time interval. The below table helps us to summarize various problem formulation approaches to obtain optimal output with respect to offloading and resource allocation strategies.

Ref	Objective	Problem formulation	Optimization technique	Contributions
[2]	Computation Offloading	Decentralizedpartially observable Markov decision process (Dec-POMDP)	Lagrangian approach and the policy gradient method	To discover the approximate optimal solution, a learning-based decentralised offloading mechanism with low complexity is given.
[3]	Computation Offloading	Stochastic optimization	Lyapunov optimization technique	Under various settings, the algorithm may achieve reasonable energy consumption and delay performance.
[4]	Computation Offloading	Constrained non-linear program (NLP) optimization problem	two-stage heuristic algorithm	Achieved optimal solution by TCO,CMEC,STO algorithms .
[5]	Computation Offloading	Mixed integer nonlinear programming problem (MINLP)	Immune algorithm (IA), Quasi-convex optimization and Convex optimization techniques	Achieved better results on different constraints
[7]	Computation Offloading and Resource allocation .	Mixed binary nonlinear programming problem	Heuristics algorithm using general bender decomposition	Results in significantly reduced worst-case complexity while maintaining high energy performance across a wide range of system parameters
[8]	Computational offloading	mixed integer programming (MIP) one which is NP-hard	convex and non-convex optimization techniques	Obtained better results comparatively with the bench mark algorithms
[12]	Computational offloading	Multi-user computation offloading game	Distributed computation offloading algorithm that can achieve a Nash equilibrium,	This algorithm performs well when number of users increases.
[13]	Computational offloading	Mixed-integer nonlinear programming problem that is NP-hard	novel heuristic algorithm	Reduce mobile device energy consumption while meeting application completion time limitations.
[15]	Computation Offloading	physical resource block (PRB) allocation	Graph colouring method	Reduce mobile device energy consumption while meeting application completion time limitations.
[17]	Computation Offloading	Combines local computing and data offloading	Gradient descent method	Shows various trade-offs based on different parameters
[18]	Computation Offloading	Weighted sum latency-minimization problem	convex optimization problem	It performs better than other bench mark algorithms

Table 1 :Problem formulation and optimization methods

VI. Parameters

There are number of parameters are observed in the study which is tabulated below.

Table 2:Parameters

Parameters	References
Data Size	[2],[4],[5][7]
CPU Cycles	[2]
Number of Devices	[2],[3],[4],[5],[7]
Task Load	[5]
Average Delay	[2]
Average Energy Consumption	[2]
Number of time slots	[3]
Available CPU cores in MEC server	[3]
Channel Band width	[4]

VII. Tools and Simulators

The researchers have recommended a few key tools, which are summarised in the table below.

Theme	Ref	Simulator
Offloading and resource allocation	[2][5]	MAT LAB
Offloading	[4]	IBM ILOG CPLEX & MATLAB R 2016b
Offloading and resource allocation	[7]	MATLAB & CVX
Offloading and resource allocation	[11]	OpenAI Gym & Python

VIII. Conclusions

MEC is becoming a popular solution as the Internet of Things (IoT) and smart gadgets grow in popularity. Many researchers from worldwide contributed papers on offloading , still there exist research gaps at this juncture. In this paper we have surveyed existing literature on Offloading strategies and resource allocation .Comparison of various strategies of offloading is presented. To conclude our research, we looked into the architecture of edge computing for IoT, the performance targets, job offloading schemes, and accompanying edge computing countermeasures, as well as typical offloading algorithms as examples. Through this study, it is anticipated that research efforts and outputs in this exciting new field will skyrocket in the coming decades due to various smart IoT applications.

REFERENCES:

[1] CISCO(2020)Cisco Annual Internet Report (2018-2023).White Paper.Source:https://www.cisco.com/c/en/us/solutions/ollateral/executive-perspectives/annual internet-report/white-paper-c11741490.pdf

[2]Q. Tang, R. Xie, F. R. Yu, T. Huang and Y. Liu, "Decentralized Computation Offloading in IoT Fog Computing System With Energy Harvesting: A Dec-POMDP Approach," in IEEE Internet of Things Journal, vol. 7, no. 6, pp. 4898-4911, June 2020, doi: 10.1109/JIOT.2020.2971323.

[3]Y. Sun, T. Wei, H. Li, Y. Zhang and W. Wu, "Energy-Efficient Multimedia Task Assignment and Computing Offloading for Mobile Edge Computing Networks," in IEEE Access, vol. 8, pp. 36702-36713, 2020, doi: 10.1109/ACCESS.2020.2973359.

[4] M. I. A. Zahed, I. Ahmad, D. Habibi and Q. V. Phung, "Green and Secure Computation Offloading for Cache-Enabled IoT Networks," in IEEE Access, vol. 8, pp. 63840-63855, 2020, doi: 10.1109/ACCESS.2020.2982669.

[5] Z. Li, C. Du and S. Chen, "HIQCO: A Hierarchical Optimization Method for Computation Offloading and Resource Optimization in Multi-Cell Mobile-Edge Computing Systems," in IEEE Access, vol. 8, pp. 45951-45963, 2020, doi: 10.1109/ACCESS.2020.2977988.

[6] Y. Guo et al., "Intelligent Offloading Strategy Design for Relaying Mobile Edge Computing Networks," in IEEE Access, vol. 8, pp. 35127-35135, 2020, doi: 10.1109/ACCESS.2020.2972106.

[7] H. Chen, D. Zhao, Q. Chen and R. Chai, "Joint Computation Offloading and Radio Resource Allocations in Small-Cell Wireless Cellular Networks," in IEEE Transactions on Green Communications and Networking, vol. 4, no. 3, pp. 745-758, Sept. 2020, doi: 10.1109/TGCN.2020.2976932.

[8] K. Wang et al., "Joint Offloading and Charge Cost Minimization in Mobile Edge Computing," in IEEE Open Journal of the Communications Society, vol. 1, pp. 205-216, 2020, doi: 10.1109/OJCOMS.2020.2971647.

[9] X. Lyu et al., "Selective Offloading in Mobile Edge Computing for the Green Internet of Things," in IEEE Network, vol. 32, no. 1, pp. 54-60, Jan.-Feb. 2018, doi: 10.1109/MNET.2018.1700101.

[10]L. Liu, X. Qin, Z. Zhang and P. Zhang, "Joint Task Offloading and Resource Allocation for Obtaining Fresh Status Updates in Multi-Device MEC Systems," in IEEE Access, vol. 8, pp. 38248-38261, 2020, doi: 10.1109/ACCESS.2020.2976048.

[11] N. Kiran, C. Pan, S. Wang and C. Yin, "Joint resource allocation and computation offloading in mobile edge computing for SDN based wireless

networks," in Journal of Communications and Networks, vol. 22, no. 1, pp. 1-11, Feb. 2020, doi: 10.1109/JCN.2019.000046.
[12]X. Chen, L. Jiao, W. Li and X. Fu, "Efficient Multi-User Computation Offloading for Mobile-Edge Cloud Computing," in IEEE/ACM Transactions on Networking, vol. 24, no. 5, pp. 2795-2808, October 2016, doi: 10.1109/TNET.2015.2487344.
[13]Y. Geng, Y. Yang and G. Cao, "Energy-Efficient Computation Offloading for Multicore-Based Mobile Devices," IEEE INFOCOM 2018 - IEEE Conference on Computer Communications, 2018, pp. 46-54, doi: 10.1109/INFOCOM.2018.8485875.
[14] J. Zhao, Q. Li, Y. Gong and K. Zhang, "Computation Offloading and Resource Allocation For Cloud Assisted Mobile Edge Computing in Vehicular Networks," in IEEE Transactions on Vehicular Technology, vol. 68, no. 8, pp. 7944-7956, Aug. 2019, doi: 10.1109/TVT.2019.2917890.
[15]C. Wang, F. R. Yu, C. Liang, Q. Chen and L. Tang, "Joint Computation Offloading and Interference Management in Wireless Cellular Networks with Mobile Edge Computing," in IEEE Transactions on Vehicular Technology, vol. 66, no. 8, pp. 7432-7445, Aug. 2017, doi: 10.1109/TVT.2017.2672701.
[16] S. Bi and Y. J. Zhang, "Computation Rate Maximization for Wireless Powered Mobile-Edge Computing With Binary Computation Offloading," in IEEE Transactions on Wireless Communications, vol. 17, no. 6, pp. 4177-4190, June 2018, doi: 10.1109/TWC.2018.2821664.
[17] H. Sun, F. Zhou and R. Q. Hu, "Joint Offloading and Computation Energy Efficiency Maximization in a Mobile Edge Computing System," in IEEE Transactions on Vehicular Technology, vol. 68, no. 3, pp. 3052-3056, March 2019, doi: 10.1109/TVT.2019.2893094.
[18] J. Ren, G. Yu, Y. Cai, Y. He and F. Qu, "Partial Offloading for Latency Minimization in Mobile-Edge Computing," GLOBECOM 2017 - 2017 IEEE Global Communications Conference, 2017, pp. 1-6, doi: 10.1109/GLOCOM.2017.8254550.
[19] Vanathi Arunachalam, Nagamalleswara Rao Nallamothu," Load Balancing in RPL to Avoid Hotspot Problem for Improving Data Aggregation in IoT " in International Journal of Intelligent Engineering & Systems, Vol.14, No.1, 2021 DOI: 10.22266/ijies2021.0228.49