A Stackelberg Game based Dynamic Resource Allocation in Edge Federated 5G Network

JARGIS AHMED¹, MD. ABDUR RAZZAQUE¹, (Senior Member, IEEE), MD. MUSTAFIZUR RAHMAN¹, SALMAN A. ALQAHTANI², AND MOHAMMAD MEHEDI HASSAN², (Senior Member, IEEE)

¹Green Networking Research Group, Department of Computer Science and Engineering, University of Dhaka, Bangladesh
²College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding author: Md. Abdur Razzaque (e-mail: razzaque@du.ac.bd).

This work was supported by King Saud University, Riyadh, Saudi Arabia, under Researchers Supporting Project number RSP-2021/18.

ABSTRACT Mobile Edge Computing (MEC) plays key role for providing fast-response and high interactivity in the emerging 5G network. Edge service providers (ESP) are responsible for serving edge users or IoT devices running latency-critical applications. An MEC server provides faster response to ESPs but has limited computation resources, hence, it can be overloaded due to extensive resource demand. Thus, federation of multiple MEC servers offers an opportunity for dynamic resource allocation in a distributed manner. The federation objective is to maximize the usage of underutilized edge resources and reduction of service provision time simultaneously. As all the ESPs and MECs act autonomously, it is quite impossible for all the individuals to achieve optimal behavior simultaneously. In this paper, we develop a Stackelberg Game (SBG) based dynamic resource allocation method to reach the expected performance. The SBG analyzes the pricing of MECs and the resource-purchasing problem of ESPs. We also develop a many-to-many matching algorithm for resource sharing among the MECs and a one-to-many matching algorithm for that between an MEC server and ESPs. The results from an extensive performance evaluation demonstrate effectiveness of the proposed system in increasing utilities for MECs and ESPs, reducing the turnaround time of application tasks, and ensuring fair resource distribution compared to state-of-the-art works.

INDEX TERMS Edge Federation, Mobile Edge Computing, Resource Allocation, Stackelberg Game, Utility Maximization.

I. INTRODUCTION

The cloud computing and wireless communication integration created enormous scope for faster real-time communication. According to the current Cisco report [1], the annual global IP traffic will reach 396 exabytes (EB) per month which is 4.8 zettabyte (ZB) per year by 2022. Also, Traffic from wireless and mobile devices will account for 71 percent of total IP traffic by 2022. This traffic is the result of the 5G era. There are typically 3 scenarios present in 5G. The first one is enhanced Mobile Broadband (eMBB), which is characterized by high data rate requirements as virtual reality (VR), augmented reality (AR), etc. Second is Ultra-Reliable and Low Latency Communication (URLCC), characterized by latency intolerant services like self-driving cars [2]. The third one is the massive Internet of Things (mIoT) [3] which can be depicted as smart city, smart agriculture. Therefore, different applications require heterogeneous resources and it is quite challenging to meet the ongoing resource demand for mobile devices and IoT present at the 5G edge network [4]–[6]. Thus, the cooperation of Mobile Edge Computing (MEC) in terms of sharing resources is a challenging issue.

Enhancing the user’s quality of experience (QoE) is very essential to ensure for diverse applications in 5G. Moreover, due to the increased computation requirement and limited computation and network resources, all tasks processing at local devices or offloading to the remote cloud for computation is inefficient and costly. In this case, the MEC will play a pivotal role [7], [8] for communication in 5G. The MECs provide not only computation capability and reduction of power consumption, but also shorten the computation and...
communication latency of IoT and mobile devices [9]. Most of the current research works investigate either to decrease the computation latency [9]–[11] for serving user tasks or computation resource optimization [12], [13]. However, the federation of MEC servers for resource sharing to serve edge user in the aid of resource shortage and fulfill resource demand at the edge still require further exploration.

In the scenario of resource shortage of any particular MEC server, the solution could be purchasing virtualized computational resource blocks (CRB) from different cloud servers [12]. However, in this case, the service latency may become intolerable, as ESPs are providing latency-critical applications, web services, etc. to edge users and IoT devices. Therefore, purchasing CRBs from other nearby MEC servers can ease the situation. Again, purchasing any amount of CRBs is not a practical approach to meet peak demand, as extra CRBs will remain unused. Hence, the MEC server must determine the optimal amount of CRB required by the ESPs as they are purchasing CRBs from MECs. Again, determining tolerable pricing of CRBs is required as ESPs will not buy CRBs at a higher price. When renting or federating resources, a buyer MEC server has to choose the best possible sellers for itself. In the same way, a particular seller can share resources with many buyers to increase its utility.

In this paper, we invade to overcome the above mentioned issues. We consider edge federation for sharing abundant unused resources to aid overloaded MEC servers in a distributed manner. The target is to maximize the usage of underutilized edge resources and reduction of service provision time simultaneously. The major contributions of this paper are listed as follows.

- We propose a distributed edge federation model for allocating resources dynamically between MEC servers and ESPs. The objective of each entity is to maximize its utility.
- We formulated the Stackelberg game to observe the interaction between ESP and MEC servers. In the game, each MEC acts as a leader and sets the renting price for ESPs. The ESPs act as a follower and purchase the optimal amount of CRBs from MEC.
- After knowing the optimal demand of CRBs, MEC determines if it has to purchase resources from other MEC servers or not. Thus we propose a many-to-many matching for resource sharing between buyer MEC and seller MEC servers.
- MEC server serves multiple ESPs and an ESP can avail service from one MEC server. Hence, we also proposed a one-to-many matching between MEC and ESPs.
- Extensive simulation done by using the Cloudsim [14]. The simulation results show the improvement in terms of the increased utility of MECs and ESPs, better service percentage of ESPs, and tasks turnaround time.

Rest of the paper is organized as follows. Section II is the literature review of the current scenario related to our work. Section III is the description of system model with assumptions and Section IV we formulated our problem. In Section V, we analyze the system based on the formulated problem and propose our algorithms. Section VI elaborates the performance evaluation and finally, Section VII summarizes the paper.

II. RELATED WORKS

Cloud computing brought revolutionary in the context of computing and service management. The federation between clouds is the concept of cooperation between different cloud servers, which is a concept of service aggregation based on the interoperability features. There are two basic federation types [15], one is the vertical federation which spans multiple levels of the cloud stack, other is the horizontal federation which spans a single level of the stack. Therefore the authors in [16] considered horizontal federation focusing on the smaller and medium cloud servers to federate along with large ones to enhance the economical gain of the server capabilities. Also, there are issues like resources availability, robustness, price of the federated resource, and the revenue gained from federation along with maintaining the high QoS. In [17] authors have addressed these issues and provided a cloud federation model where they consider the horizontal federation of Infrastructure as a Service (IaaS) layer of different cloud servers.

As the current network is growing rapidly for connecting new devices entering every day for computation and communication, the static resource provisioning will not be sufficient enough to cope-up with the same rate. The software defined networking (SDN) and network function virtualization enables flexible resource management and creates abstraction on the physical resource and internet services [18], [19]. Where NFV decouples network functions from dedicated hardware and SDN creates programmable centralized network control. The combination of SDN-NFV can be used as one of the key elements to make federation among clouds [20], [21]. In [22], the authors provided a centralized idea of SDN controller placed in MEC connected to the base station for the ultra-dense network. This SDN controller collects network status from a global perspective and advises mobile devices to perform tasks locally or offload tasks to a nearby small cell base station. An approach of distributed computation offloading and profit maximization is done in [13]. Where SDN controller placed in MEC and it controls only the resources required for different mobile virtual network operators (MVNOs) and offloading process from user to MEC with an extra element called virtualized user (VU). The joint optimization of the model considering the VU, MEC, and MVNOs are proposed by the authors and their proposed algorithm reduces the signaling overhead and complexity of the computation.

In [23] authors propose an auction based model for resource sharing between edge servers and the service providers to maximize the utility of each entity. Here, the allocation of resources is distributed and this process is also decoupled from service provisioning management at
the service provider end. Now, when one particular MEC has abundant resources, it may share them with the cloud server and other edge servers also. On the other hand, when the demand increases that particular MEC have to buy resources from cloud server as well as from MEC servers. In [12], authors presented an effective process of wholesaling abundant edge resources to cloud server and buyback again if required by a particular MEC. Here, the main objective was to increase profit for MEC by wholesaling its resource which is underutilized at a lower price and then buyback at a higher price, if required to serve tasks coming to that MEC. In [24] the authors discussed the issue of sharing the resources between different MEC servers. They proposed a cooperative game based framework, where the edge servers form a coalition for resource sharing. When one MEC can not meet the resource requirement for its native applications other MECs can aid this situation by providing resources after ensuring the fulfillment of their requirements. They provided Pareto optimal solution for allocating resources using the cooperation of the present MEC servers in the system.

Most of the existing studies do not focus on the impact and opportunity of the federation of edge servers to handle extensive edge resource demand. Few works focused on the sharing resources between edge to cloud servers [12] and edge to edge servers [24], although none of the works focus on the optimal resource requirement of the ESPs to serve edge users and the objective of maximizing the utility. When the ESPs resource demand information at MEC servers are properly available, then it eases the decision for MEC to buy or sell resources. Hence, this observation lead us to design a Stackelberg game based approach for finding the optimal behavior of ESPs and MEC servers and federate resources between preferred MEC servers.

### III. SYSTEM MODEL AND ASSUMPTIONS

In the edge network, the edge user runs different types of applications where most of which are latency-critical and requires real-time interactions. The Edge Service Provider(ESP) consists of User Equipment or IoT devices for smart-home. To serve the incoming tasks, ESP requests resources from one particular InP. The InP consists of a base station, MEC server, and SDN controller. For the case of simplicity, we use characterize InP as a MEC server. One MEC server is connected with other MEC servers. The figure (1) shows the federated environment of edge federation.

Let the set of ESP is $M = \{1, 2, ..., m\}$ and the number of ESP is $M$. The set of InPs containing the MEC server is $E = \{1, 2, ..., e\}$. For serving edge users or IoT devices, each ESP requests resources from its nearby MEC server. We define the unit amount of resource that can be purchased by ESP from the MEC server as a computational resource block (CRB). The service rate of each CRB is denoted as $\mu$. A particular MEC can serve multiple ESP by its resources. The aggregated tasks arrival rate to $m^{th}$ ESP is represented by $\lambda_m$. Now, if a large number of ESPs prefer to get CRBs from one particular MEC server to serve its incoming tasks, that MEC will face resource scarcity. Therefore, this MEC can now act as a buyer and buy CRBs by the federation process, from other MEC servers who have abundant CRBs to sell. $S_e$ is the set of sellers maintained by buyer MEC $e$, where the total number of sellers is $N$. A seller $n$ in practicality can't serve all its buyers. Therefore, it will maintain a preference list of the buyer MECs, denoted as $\mathcal{L}^n_i$. From the buyer set $B^n$ containing $B$ number of buyers. Also, a buyer MEC can buy resources from multiple sellers. Hence, it also maintains a preference list of sellers $\mathcal{L}^i_n$. Furthermore, $e^{th}$ MEC also has its preference to serve connected ESP, where the list is denoted by $\mathcal{L}_m^n$.

As already mentioned ESP serves UE and IoT devices running real-time applications, better QoS must be ensured for edge service. The measurement of QoS for $m^{th}$ ESP depends on the service delay $t_m$ which is calculated by

$$t_m = f_m + d_m$$
where \( f_m \) is the network delay and \( d_m \) is the queuing delay. According to [25] the queuing delay for serving ESP \( m \) can be calculated by the following equation

\[
 d_m = \frac{\lambda_m}{\mu - \frac{\lambda_m}{q_m}} \tag{2}
\]

where \( q_m \) is the amount of CRBs purchased by the ESP \( m \). The network delay can be influenced by the traffic situation of the network, distance for transmitting, and many other uncertain factors. However, for the case of simplicity, the dependency of latency for network delay is greatly dependent on the total distance \( l_m^r \) from ESP \( m \) to the physical service provider \( e \). Therefore \( f_m \) has a linear relationship with \( l_m^r \) and scalar \( \theta \) represented as

\[
 f_m = \theta l_m^r. \tag{3}
\]

To serve the edge users, ESP have to pay for purchasing CRBs from MEC, following the model presented in [26], the utility of ESP can be measured by the revenue received from the incoming workload of ESP and subtracting the cost for the payment of MEC and the cost incurred by the service delay, represented as following

\[
 U^m = \alpha_m \lambda_m - \omega_m^e \beta_m q_m r_e - \gamma_m t_m \tag{4}
\]

where \( \alpha_m, \beta_m, \gamma_m \) are the weight factors determined by each ESP for itself. \( \omega_m^e \) is the association indicator where \( \omega_m^e = 1 \) indicates ESP \( m \) is connected to MEC \( e \) and 0 otherwise. Also one ESP can avail service form one MEC only. \( r_e \) is the per unit price of virtualized CRBs set by MEC \( e \) and \( q_m \) is the amount of CRBs purchased by \( m \) form \( e \).

According to the purchasing request from the serving ESP, each MEC calculates its utility from revenue gained from serving ESPs minus the energy cost for serving CRBs by itself. However, if \( e \) does not have a sufficient amount of CRBs present at the moment, it needs to buy CRBs from other sellers MECs \( n \). Therefore the payment for buying CRBs and cost for communication delay from different sellers also subtracted from the utility of \( e \), which is

\[
 U^e = \sum_{m=1}^{M} \omega_m^e q_m r_e - \sum_{n=1}^{N} (p_n^e q_n^e + \psi^e t_n^e) - \varepsilon_e q_e \tag{5}
\]

where \( q_m^e \) is the amount of CRBs purchased from seller \( n \), \( p_n^e \) is the price set by \( n \), \( \psi^e \) is the weight factor and \( t_n^e \) is the communication delay between \( e \) and \( n \). \( q_e^e \) is the number of CRBs is reserved by that particular MEC \( e \) and \( \varepsilon_e \) is the energy cost of serving each CRBs by itself.

Again from the perspective of a seller MEC the utility function will be the payment received from the buyers minus the maintenance cost for serving each buyer as follow

\[
 U^m = \sum_{i=1}^{B} q_n (p_i^e - c_n) \tag{6}
\]

where \( p_i^e \) is the price set by the seller \( n \) itself, \( c_n \) is the maintenance cost and \( q_n^e \) is the amount of CRBs rented by a buyer \( i \) from \( n \).

IV. PROBLEM FORMULATION

According to the federation architecture between MECs, the interactions of resource trading between ESP and MEC and between buyer MEC to seller servers, all the individual ESP, buyer MEC, and seller servers can’t achieve the maximal utility simultaneously. Therefore we need to observe a sequential decision making process. In this process, MEC predicts the total number of CRBs purchased by its serving ESPs and solves the pricing problem for ESPs using Stackelberg Game.

Observing the behavior of MEC, each ESP determines to buy resources from MEC. After that, a buyer MEC now knows the amount of CRBs it needs to buy from seller servers and requests accordingly. Finally, each seller will provide the required amount of CRBs to its buyer and compete with other seller servers.

The objective of ESP is to maximize its utility by serving the edge user and IoT devices. Therefore, we can write the optimization problem for each ESP as follows

\[
 \max_{q_m} U^m(q_m | r) \tag{7}
\]

subject to:

\[
 \sum_{e=1}^{E} \omega_m^e = 1 \tag{8}
\]

\[
 t_m \leq T \tag{9}
\]

\[
 \alpha_m \lambda_m \geq \beta_m q_m r_e + \gamma_m t_m \tag{10}
\]

\[
 q_m \geq 0 \tag{11}
\]

where the pricing profile \( r \) is the price value set by all the MEC servers nearby to the ESP \( m \). Constraint (8) ensures that ESP can avail service from only one MEC. Constraint (9) indicates tasks service delay should not exceed delay deadline. Constraint (10) indicates utility gained by edge users should be greater than the resource buying utility and constraint (11) ensures that ESPs received nonnegative resources from the MEC server.

After observing the behavior of ESPs, MEC now sets the serving price to serve ESPs and compete with other MECs. Also, all MEC will determine if it needs to buy CRBs from seller servers or not. Therefore the optimization problem for the buyer MEC is

\[
 \max_{r_e} U^e(r_e | Q_s^e, P, r_{s-e}) \tag{12}
\]

subject to:

\[
 \sum_{m=1}^{M} \omega_m^e q_m r_e \geq \sum_{n=1}^{N} (p_n^e q_n^e + \psi^e t_n^e) + \varepsilon_e q_e^e \tag{13}
\]

\[
 t_m + t_n^e \leq T \tag{14}
\]

\[
 r_e \geq 0 \tag{15}
\]

\[
 q_e^e \geq 0 \tag{16}
\]

\[
 q_e^e \geq 0 \tag{17}
\]

where \( Q_s^e \) denotes the optimal amount CRBs purchased by all the connected ESPs. \( P \) is the rent profile of all the seller servers. \( r_{s-e} \) is the optimal price profile, set by other MECs except \( e \). The constraint (13) ensures that utility gained by
serving ESPs should be greater or equal to the resource buying utility and energy utility. Service latency should not exceed the delay deadline of tasks, hence, the buyer will buy resources from the seller where the constraint (14) can be met. Constraints (15, 16, 17) indicates non negative price and non negative resource purchased from sellers and MEC allocates non negative resources to ESPs respectively.

A seller MEC also wants to maximize its utility, therefore, it chooses the most preferable buyer MECs. Also, it needs to compete to sell its CRBs to buyers. The optimization problem for the seller server is

\[
\text{maximize} \quad U^m(q^i_n | Q_{-i}, P^m_{-n}) \quad (18)
\]

subject to:

\[
p^i_n \geq c_n \quad (19)
\]

\[
g^i_n \leq q^{i-th} \quad (20)
\]

\[
\sum_{i=1}^{B} q^i_n \leq Q^s_n \quad (21)
\]

where \(Q_{-i}\) is the optimal amount of CRBs required by buyers other than \(i\). \(P^m_{-n}\) is the pricing profile set by other sellers except \(n\). The constraint (19) indicates pricing of CRBs should be greater or equal to the maintenance cost. Constraint (20) ensures allocation of resource to buyer \(i\) will be smaller or equal to the demand \(q^{i-th}\). Also, a seller will allocate CRBs to buyer MECs up to the available amount \(Q^s_n\), which is represented by constraint (21).

After observing the interactions between ESPs, MECs and seller servers, we can focus on the following problems

1) **Amount of resource purchased by ESP from MEC:**

ESP s are serving latency-critical applications of edge users and IoT devices. Again, the demand for CRBs is influenced by the service delay and pricing of the MEC server. Hence, the number of CRBs required by the ESPs should be optimal to increase utility.

2) **Determining the price for ESP by MEC:**

When MEC provides CRBs to ESPs, the pricing of the CRBs should not be intolerable. Because at higher price ESPs will buy less resource from that MEC server and will give up if it cannot gain positive utility. Also, MEC has to bear the serving cost and resource buying cost if it acts as a buyer. Hence, MEC has to ensure the optimal pricing of CRBs.

3) **Resource allocation problem between buyer-seller:**

When a MEC needs to buy resources from seller servers, it does have some preferences for selecting sellers. In the same way, each seller wants to allocate CRBs to its buyers to increase utility value. To maximize the utility, each seller will also have a preference list for selecting its buyers.

4) **Resource allocation problem for MEC-ESP:**

MEC will allocate resources to its serving ESPs after matching between MEC and all its ESPs. MEC will prefer the ESPs form which the total utility will be maximized based on the pricing to ESP.

According to the above formulated problem and discussion, we observe that each ESPs, MECs are autonomous in their actions. Each individual observes other actions and takes the optimal decisions to improve their utility. Therefore to reach the optimal behavior, we first formulate the Stackelberg game where the MEC server acts as a leader and ESPs acts as followers. The purpose of the game is to solve the pricing problem for MECs and the resource purchasing problem for the ESPs. After calculating the optimal amount of CRBs purchased by ESPs, if MEC needs to buy external resources to meet the demand, then a many-to-many matching between buyers and sellers servers is proposed. Finally, to serve ESPs by a particular MEC server a one-to-many matching algorithm is proposed.

**Lemma IV.1.** The stackelberg game between MEC and the ESP, when MEC sets the service price \(r^e\), the optimal amount of CRBs required to maximize utility of a ESP is

\[
q^e_m = \frac{\lambda_m}{\beta} \left(1 + \sqrt{\frac{\gamma_m}{\beta_m r^e}}\right) \quad (22)
\]

**Proof.** Based on the utility equation of ESPs in eq. (4) the first derivative of \(U^m\) with respect to \(q^m\) is

\[
\frac{\partial U^m}{\partial q^m} = -r^e \beta_m + \gamma_m \left(\frac{\lambda_m}{\mu q^m - \lambda_m}\right)^2 \quad (23)
\]

Also the second derivative of \(U^m\) with respect to \(q^m\) is as follows

\[
\frac{\partial^2 U^m}{\partial q^m} = -2 \gamma_m \mu^2 \lambda^2_m \quad (24)
\]

Now as the \(\frac{\partial^2 U^m}{\partial q^m} < 0\), which indicates that \(U^m\) follows a decreasing relationship with respect to \(q^m\). Hence if we consider the first derivative of \(U^m\) with respect to \(q^m\) to be 0, we will get the optimal number of CRBs or \(q^*\) purchased by each ESP \(m\).  

As the actions of ESPs regarding the optimal CRBs are known to MEC server, the adjusted optimization function for a particular MEC server \(e\) is as follows

\[
\text{maximize} \quad \tilde{U}^e(r^e | Q^e, P, r^s) \quad (25)
\]

subject to:

\[
\sum_{m=1}^{M} \frac{\lambda_m}{\mu} \left(1 + \sqrt{\frac{\gamma_m}{\beta_m r^e}}\right) \omega_m r^e \geq \sum_{n=1}^{N} (p^e_n q^e_n + \psi^e_n r^e_n) + \varepsilon e q^e_n. \quad (26)
\]

and constraints 14-17, where the value of \(U^e\) is updated to \(\tilde{U}^e\) as follows,

\[
\tilde{U}^e = \sum_{m=1}^{M} \frac{\lambda_m}{\mu} \left(1 + \sqrt{\frac{\gamma_m}{\beta_m r^e}}\right) \omega_m r^e - \sum_{n=1}^{N} (p^e_n q^e_n + \psi^e_n r^e_n) - \varepsilon e q^e_n. \quad (28)
\]

We observed that the first derivative of \(\tilde{U}^e\) with respect to \(r^e\) is greater than 0, where

\[
\frac{\partial \tilde{U}^e}{\partial r^e} = \frac{\lambda_m}{\mu} \left(1 + \frac{1}{2} \sqrt{\frac{\gamma_m}{\beta_m r^e}}\right) \quad (29)
\]
which indicates that the utility of a MEC server $e$ is positive co-related with the pricing set by $e$ to its ESPs. Again the as the service delay for any particular ESPs can not cross $\mathcal{T}$, therefore the low bound for purchasing CRBs by any ESPs given as follows

$$q_m \geq \frac{\lambda_m \mathcal{T}}{\mu \mathcal{T} - \lambda_m} \quad (30)$$

Finally following the low bound and eq. (22), the optimal pricing by MEC $e$ to its ESPs for maximizing its utility is

$$r_e = \frac{\gamma_m}{\beta_m} \left( \frac{\mu \mathcal{T} - \lambda_m}{\lambda_m} \right)^2 \quad (31)$$

After calculating the optimal demand at the edge and pricing for ESPs, now we have to allocate resource dynamically.

V. DISTRIBUTED ALGORITHM FOR RESOURCE FEDERATION AT THE EDGE

After observing the behavior of ESPs, one particular MEC knows the amount of required CRBs ($Q_e$) it needs to purchase from nearby server servers. To find out the best possible seller, each buyer needs to make their preference list. Each buyer MEC wants to buy resources from the seller considering lower renting price and minimal communication latency. Therefore the preference list of a particular MEC $e$ who needs to buy CRBs form sellers is $\mathcal{L}_n^e = \{\mathcal{L}_1^e, \mathcal{L}_2^e, \ldots, \mathcal{L}_m^e\}$ considering,

$$\mathcal{L}_n^e = -(\zeta^e \mathcal{P}_n^e + \eta^e \tau^e_n), \quad (32)$$

where $\zeta^e$ and $\eta^e$ are weight factors set by each buyer for itself, $\mathcal{P}_n^e$ and $\tau^e_n$ are the normalized price and communication delay, respectively.

Accordingly a seller server $n$ will make its preference list of buyers $\mathcal{L}_i^n = \{\mathcal{L}_1^n, \mathcal{L}_2^n, \ldots, \mathcal{L}_m^n\}$. Each seller prefers the buyers from which it can gain more profit thus,

$$\mathcal{L}_i^n = p_i^n - c^n. \quad (33)$$

After the construction of the preference list by each buyer, it will try to buy resources from its preferred sellers. In Algorithm 1 we set a pointer for each buyer MEC to point its preferred seller from the preference list. At the starting phase, the pointer will point to the first seller. In each iteration, the buyer will propose its preferred seller for sharing resources and wait for confirmation. We also set a flag $f_e$ for each buyer MEC, which indicates wheather the seller $n$ allocates its resource to that buyer or not. Initially, the value is set to 0 and when a buyer proposes a seller $n$, the value of $f_e$ will be 0. Now, if one particular buyer is rejected by seller $n$, it will update its pointer to point next most preferred seller form $\mathcal{L}_n^e$ and propose in the same way. This process of requesting seller continued until the demand of total amount of required CRBs $Q_e$ met.

From the perspective of a seller in Algorithm 2, based on the proposals from buyers, each seller constructs buyer preference list $\mathcal{L}_i^n$, and allocate its resource for rent to the most preferred buyer. After allocating CRBs to one buyer, the pointer of the seller updates to point next most preferable buyer in the next iteration. Selection of buyer will happen until the total available amount of CRBs $Q_n^s$ of the seller $n$ is not 0. If there are no more available CRBs to share by the seller to its buyer, it simply rejects the requests.

Algorithm 1 Many-to-Many Matching Resource Sharing for Buyer MEC

1: Buyer $e$ constructs preference list of sellers $\mathcal{L}_n^e$ according to Eq. (32)
2: To act as an indicator one pointer is set to point at the first seller in $\mathcal{L}_n^e$
3: Set flag $f_e$ to indicate whether $e$ has been chosen by $n$, initially $f_e = 1$
4: while all seller from $\mathcal{L}_n^e$ not scanned by $e$ do
5: if $f_e = 1$ then
6: Pointer stays current position in $\mathcal{L}_n^e$
7: else
8: Pointer iterates next preferred seller in $\mathcal{L}_n^e$
9: end if
10: propose to pointed seller $n$ according to preference for purchasing CRBs
11: $f_e \leftarrow 0$
12: end while

Algorithm 2 Many-to-Many Matching Resource Sharing for Seller MEC

1: Seller $n$ constructs preference list of buyers $\mathcal{L}_i^n$ according to Eq. (33)
2: while $Q_n^s \neq 0$ do
3: if buyer $i \in S_i$, selected according to the $\mathcal{L}_i^n$ then
4: if $Q_n^s > q_i$ then
5: Allocate $q_i$ to $i^{th}$ buyer
6: $Q_n^s \leftarrow Q_n^s - q_i$
7: else
8: Allocate $Q_n^s$ to $i^{th}$ buyer
9: $Q_n^s \leftarrow 0$
10: end if
11: else
12: Reject buyer $i$ and set flag $f_i \leftarrow 1$
13: end if
14: Pointer moves to next preferred buyer
15: end while

On the other hand the MEC will choose its serving ESPs based on the preferences. A particular MEC server will construct its own preference list as $\mathcal{L}_m^e = \{\mathcal{L}_1^e, \mathcal{L}_2^e, \ldots, \mathcal{L}_m^e\}$ where,

$$\mathcal{L}_m^e = r_e \quad (34)$$

After the federation process of renting resources from different sellers, each MEC $e$ will allocate its currently available number of CRBs $Q_e$ to its serving ESPs. The allocation of CRBs by MEC server to ESPs depicted in algorithm 3. MEC will construct the preference list of

On the other hand the MEC will choose its serving ESPs based on the preferences. A particular MEC server will construct its own preference list as $\mathcal{L}_m^e = \{\mathcal{L}_1^e, \mathcal{L}_2^e, \ldots, \mathcal{L}_m^e\}$ where,

$$\mathcal{L}_m^e = r_e \quad (34)$$

After the federation process of renting resources from different sellers, each MEC $e$ will allocate its currently available number of CRBs $Q_e$ to its serving ESPs. The allocation of CRBs by MEC server to ESPs depicted in algorithm 3. MEC will construct the preference list of...
Table 2. Simulation Parameters

| Parameters             | Value                  |
|------------------------|------------------------|
| Number of ESPs         | 200 - 320              |
| Number of MECs         | 3 - 9                  |
| Tasks arrival rate at ESP ($\lambda_m$) | 200 - 800 per second |
| Service rate of CRB ($\mu$) | 100 tasks per second |
| Available CRB at MEC   | 200 - 400              |
| Size of CRBs, tasks size | 1000 MIPS, 2000 MIPS  |
| Maximum delay deadline ($T$) | 60 ms                  |
| $\alpha_m$, $\beta_m$, $\gamma_m$ | 20, 0.01, 0.001       |
| Distance of ESP from MEC | 0 ~ 100 km            |
| Announced price of MEC for ESP | 5 ~ 10                |
| Announced price for buyer MEC | 0 ~ 5                 |
| Transmission speed     | 50 km/ms               |

B. PERFORMANCE METRICS

1) *Average utility of MECs* is defined as the average utility achieved by MECs, providing the virtualized CRBs for serving ESPs.

2) *Average utility of ESPs* is the average utility gained by ESPs as it serves edge users or IoT devices.

3) *Percentage of served ESPs* is the measurement of the ESPs, which receives the required amount of CRBs from the requesting MEC.

4) *Average turnaround time of tasks* is calculated for successful completion of tasks, which arrived at each ESPs.

C. SIMULATION RESULTS

We study the performance of our proposed SBG method by varying the number of ESPs, changing the task arrival rate at ESPs, and varying the number of MECs.

1) Impact of varying the number of ESPs

The Fig. 2 depicts the impact of varying the number of ESPs in the system. To study the impact we change the number of ESPs from 200 to 320 and assume the tasks arrival rate at each ESP is 500 per second and 5 MECs for providing resources to the ESPs.

In Fig. 2a, we can observe that as the number of ESP increases the average utility of MECs increases. The utility increases rapidly for MECs when the number of ESPs is 200 to 280, after that the growth of utility gradually decreased to the nearly equal utility for the last few points. However, the saturation point is reached when the number of ESP is 300 to 320. The reason behind this is when all MEC have available resources, increasing ESP helps MECs to gain more utility by providing its resources to the demanding ESPs according to their requirements. After when the ESPs number reached 300 to 320, all MECs are now saturated by demanding ESPs and 5 MECs for providing resources to the ESPs.

In Fig. 2a, we can observe that as the number of ESP increases the average utility of MECs increases. The utility increases rapidly for MECs when the number of ESPs is 200 to 280, after that the growth of utility gradually decreased to the nearly equal utility for the last few points. However, the saturation point is reached when the number of ESP is 300 to 320. The reason behind this is when all MEC have available resources, increasing ESP helps MECs to gain more utility by providing its resources to the demanding ESPs according to their requirements. After when the ESPs number reached 300 to 320, all MECs are now saturated by demanding ESPs and MECs have no more resources available to provide ESPs.

We can also observe that our proposed SBG and the method CG have a very close average utility of MECs compared to the NF method. The reason behind this is in SBG a buyer MEC and buy CRBs from other sellers and in CG all the MECs which have abundant resources can share resources ESPs other than its serving ESPs. On the contrary in the

Algorithm 3 One-to-Many matching algorithm for an MEC

1. MEC $e$ constructs $\mathcal{L}^m_m$ for all ESPs following Eq. (34).
2. To act as an indicator one pointer is set at the first ESP in $\mathcal{L}^m_m$.
3. Set flag $f_m$, $\forall m \in \mathcal{M}$ to indicate whether $m$ has been chosen by $e$, initially $f_m = 1$.
4. While all $m$ is not scanned form $\mathcal{L}^m_m$ do
5. if $Q_e \geq q_m$ then
6. Allocate $q_m$ to ESP $m$
7. $Q_e \leftarrow Q_e - q_m$
8. $f_m \leftarrow 0$
9. else
10. Allocate $Q_e$ to ESP $m$
11. $Q_e \leftarrow 0$
12. $f_m \leftarrow 0$
13. end if
14. Pointer moves to next preferred ESP where $f_m = 1$
15. end while

ESP and set a pointer for each MEC to point ESPs in $\mathcal{L}^m_m$. Initially, the pointer points first ESP. We also set flag $f_m$ to indicate if ESP $m$ is chosen or not. Initially, the pointer value is 1. Each MEC selects its serving ESPs based on the preference and allocate resource according to the requirement and set $f_m = 0$. If one particular MEC can not fulfill all ESPs resource demands then it prefers to serve only the most preferable ESPs and low preferred ESP will not get the desired amount of CRBs.

VI. RESULT ANALYSIS

In this section, we discuss the simulation setup and the analysis of the proposed edge federation process in detail. We used CloudSim [14] for the implementation and simulation result analysis. For simplification, we denoted our proposed method as Stackelberg Game (SBG). We compare our SBG with standalone non-federated (NF) MECs and another state-of-the-art-work Cooperative Game(CG) [24].

A. SIMULATION ENVIRONMENT

We considered there are 3-9 MEC servers, and 200 - 320 ESPs randomly in a circular distance of 100km. By following the practical environmental settings in [25] we set the tasks arrival rate of each ESP $\lambda_m$ as a random number having the average of 0.5 per millisecond (ms) which is 500 per second (s). The CRB service rate is $0.1 (ms)^{-1}$. The maximum delay deadline for all ESPs is 60 ms. Based on the usage of its computing resources and its service cost, we set the announced rent as a random number satisfying the uniform distribution between (5, 10) also the announced price for buyer MEC is between (0, 5), and the amount of available CRB as a random number satisfying uniform distribution between (200, 400). The weight factors $\alpha_m$, $\beta_m$, $\gamma_m$ are set as 20, 0.01 and 0.001, respectively. Performance parameters can have values as summarized in Table 2.
In Fig. 2b, the average utility of ESPs is higher when the number of ESPs was low and decreased drastically according to the increasing number of ESPs. The reason behind this is when the number of ESPs increases MECs are providing available CRBs to ESPs and can satisfy the total amount of resource demand up to a certain point. After that increasing ESPs actually increases the demand of CRBs and now it will cause resource scarcity, therefore the number of fully served ESPs will decrease. In this scenario, each ESPs have to wait more and the competition of getting resource is higher for all the described methods. As usual, the NF method performs worst and our proposed SBG outperforms the other two methods of resource provisioning.

As already mentioned that, an increasing number of ESPs will cause resource starvation and eventually decrease the average utility of ESPs; the percentage of served ESPs will also decrease. In Fig. 2c, we observe that when the ESPs number in the system is low link 200 to 220 both SBG and CG can serve 100% of ESPs according to the demand. When ESPs increases the percentage of served ESPs decreases accordingly. The NF performs worst as in this approach, MECs can not buy or federate resources from other MECs hence the percentage of ESPs getting served is very low compared to the other methods and SBG outperform CG.

In Fig. 2d, we observe the average turnaround time (TAT) of successfully completed tasks. We can find that the increasing number of ESPs increases the TAT of tasks in all the described methods. We observe that the NF method has the lowest TAT whereas the CG method has the highest TAT value. The reason behind this is, in the NF method the MECs do not communicate with other MECs to share resources that’s why the burst time is less than other methods. Whereas in CG ESPs have to wait until another MEC who have abundant resources decides to share its CRBs to that particular ESP which are not getting resources. In contrast, in the SBG method if a MEC faces resource shortage, that particular MEC by itself request other MECs to share...
resources, and thus SBG requires less time for finding resources and satisfying it’s demanding ESPs requests. We also can observe that when the number of ESPs much higher like 300 or 320 the TAT is much better in our proposed SBG method. Although it seems NF is better than the other two approaches, but we can surely say that is not true when we observe in Fig. 2c.

2) Impact of varying task arrival rates

In Fig. 3 we measure the impacts of changing tasks arrival rate at each ESPs. We change the arrival rate from 200 to 800 tasks per second and maintain the number of ESPs present in the system is 200 also the number of MECs are 5, and observe the impacts.

In Fig. 3a, depicts that when the average tasks arrival rate of each ESP increases the average utility of MECs increases. The reason behind this is now MECs can increase prices for CRBs and ESPs have to buy them at a higher price. The utility value is increasing up to tasks arrival rate 700 for all the NF, CG, and SBG methods. However, when the number of tasks arrival increases further, the utility values are not increasing for MECs as all MECs are now at saturation point. Therefore when MECs reach the saturation point of resource availability, it is not possible to allocate more resources to demanding ESPs. Here we also can observe that MECs in the NF method have the lowest utility value than other methods. Also, the SBG and CG ensure very close utility values for the MECs.

In Fig. 3b, we observe that the average utility of ESPs also increases when the tasks arrival rate is increasing. The utility values are nearly similar for tasks arrival rate 200 to 400 for all described methods. The reason behind this is, at this tasks arrival rate, nearly all the demand of resources of all ESPs can be served in flexible manner as the resources are highly available due to low demand. However then the increasing rate of the utility of ESPs decreases when the tasks arrival rate is higher like 500 to 800. Also, when the tasks arrival rate is 700 - 800 as MECs are highly saturated therefore ESPs utility also reach a saturation level.

In Fig. 3c, we can see that increasing the rate of the arrival of the tasks decreases the average served percentage of ESPs. As already mentioned when more tasks are arriving,
the demand of resources of ESPs also increases as they have to serve the increasing edge demand. However, the MECs are resource-constrained and all the MECs will face resource scarcity at the higher tasks arrival rate. For that particular reason more ESPs will face resource shortages and the number of ESPs getting service from MECs will drastically decrease. However as usual NF method performed worst in this case, CG, and our proposed method SBG performs moderately well to handle the increased resource demand.

In Fig. 3d, we observe that the NF method has the lowest TAT for successful tasks completion and CG has the highest. Our proposed SBG performs moderately than CG. Though NF has the lowest TAT it does not mean that NF is better than the other two methods as we already observed in Fig. 3c the percentage of served ESPs are very low. Both CG and SBG require extra communication time or burst time for searching and providing abundant resources of different MECs.

3) Impact of varying the number of MECs

In figure 4 we change the number of MECs in the system from 3 to 9 and maintain the number of ESPs present in the system is 320. Also, we consider the tasks arrival rate at ESPs to 500 tasks per second and observe the impacts.

In Fig. 4a, increasing the number of MECs decreases the average utility value for all the MECs in all the described methods. When we increase the number of MECs it means we increase the total amount of resources to serve edge load. Now all the MEC servers will have fewer ESPs to serve, which means ESPs are now distributed between MECs more evenly. This will eventually decrease the utility of MECs. The NF method faces a drastic decrease in utility value, whereas the CG and our proposed SBG method behave more similarly when the number of MEC increases. We can see that when the number of MECs is 9 both CG and our proposed SBG have are nearly equal average utility for MECs because the need to spend less communication time and the demand of ESPs also nearly equally distributed among serving MECs.

In Fig. 4b, we observe that the average utility value of
ESP's increases when the number of MEC's increases. As already mentioned, we run this simulation by keeping the number of ESPs to 320, which is an overloaded situation for all MECs. That's why the utility value of ESPs is quite low when the MECs number is 3. After, when we increase the MEC count, which indicates we increase the resource availability. When the number of MECs is much higher like 7 to 9 we can see the gained utility of ESPs are nearly equal for both CG and SGB methods. This indicates that the system has more resources than the demand of ESPs and the utility value of ESPs is now at saturation point. However, the NF still performs poorly although we increase the number of MECs.

In Fig. 4c, we observe the percentage of served ESPs is increasing when the number of MECs increases. As already stated when the MECs number was low the scarcity of resources was very high. Which causes more than half of the ESPs not to get the required amount of CRBs. This situation eases when we increase resource availability by increasing the number of MECs in the system. All the three methods NF, CG, and SGB are follow the same intuition. We also observe that when the number of MECs is between 7 to 9, the percentage of served ESPs are close to 100% for both CG and our proposed SGB methods. That is all the ESPs are getting their required resources to fulfill the edge demand as the resource are abundant for all the MECs.

In Fig. 4d, we observe the decrease in average TAT concerning the increasing number of MECs. We see that the NF method has the lowest TAT value than the other two methods. The reason behind this is, in NF there is no possibility of renting resources from other MECs, thus this method serves its ESPs faster than rest two methods, although the served percentage is not satisfactory as we can see in Fig. 4c. However, our proposed SGB method has less TAT for successful tasks completion than CG. In this approach, the burst time or the communication time is lower than the CG method as finding resources is done by the MEC server itself which reduces the burst time. On the other hand in the CG method, MECs allocate resources to their own ESPs and then find other ESPs to provide abundant resources, this increases the burst time for tasks getting served. When the number of MEC servers reaches 8 to 9, we can also observe that all three methods have nearly equal average TAT for serving their ESPs.

 VII. CONCLUSION

In this paper, we proposed the resource sharing approach or the federation process between the MEC servers. Our target was to aid overloaded MEC servers and federate abundant resources of other MEC servers. The objective was to find dynamic resource requirements and the allocation of resources in a distributed manner. Hence, our proposed SGB method first calculates the optimal resource demand for each ESP, and based on that we categorized buyer and seller MEC servers. We observed the interactions of buyer and seller servers and the interaction of MEC with ESPs. We developed a many-to-many matching algorithm between buyer and seller MEC servers for resource sharing or federation. Then, we also represented a one-to-many matching between MEC and ESPs. The proposed method showed improvement in a comparative study with the state-of-the-art-work and non-federated standalone MEC approach.

References

[1] V. Cisco, “Cisco visual networking index: Forecast and trends, 2017–2022 white paper,” Porto Salvo, Lisboa. Disponível em:< https://www.cisco.com/c/en/us/solutions/collateral/service-provider/pt_pubs/press/news-archive-2018/20181217. html>[Accessed: 20-Sep.-2021]. Acesso em, vol. 17, 2019.

[2] C. Campolo, A. Molinaro, A. Iera, and F. Menichella, “5G networking for vehicle-to-everything services,” IEEE Wireless Communications, vol. 24, no. 6, pp. 38–45, 2017.

[3] V. W. Wong, Key technologies for 5G wireless systems. Cambridge university press, 2017.

[4] Z. Chen, W. Hu, J. Wang, S. Zhao, B. Amos, G. Wu, K. Ha, K. Elgazzar, P. Pillai, R. Klatzyk et al., “An empirical study of latency in an emerging class of edge computing applications for wearable cognitive assistance,” in Proc. of the 2nd ACM/IEEE Symposium on Edge Computing, 2017, pp. 1–14.

[5] T. Taleb, A. Ksentini, M. Chen, and R. Janti, “Coping with emerging mobile social media applications through dynamic service function chaining,” IEEE Trans. on Wireless Commun., vol. 15, no. 4, pp. 2859–2871, 2015.

[6] D. Tian, J. Zhou, Z. Sheng, M. Chen, N. Qi, and V. C. Leung, “Self-organized relay selection for cooperative transmission in vehicular ad-hoc networks,” IEEE Trans. on Vehicular Tech., vol. 66, no. 10, pp. 9534–9549, 2017.

[7] T. X. Tran, A. Hajisami, P. Pandey, and D. Pompili, “Collaborative mobile edge computing in 5G networks: New paradigms, scenarios, and challenges,” IEEE Commun. Mag., vol. 55, no. 4, pp. 54–61, 2017.

[8] S. Kekki, W. Featherstone, Y. Fang, P. Kuure, A. Li, A. Ranjan, D. Purkayastha, F. Jiang, D. Dyfman, G. Verin et al., “MEC in 5G networks,” ETSI white paper, vol. 28, pp. 1–28, 2018.

[9] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, “A survey on mobile edge computing: The communication perspective,” IEEE Commun. Surveys & Tutorials, vol. 19, no. 4, pp. 2322–2358, 2017.

[10] T. G. Rodrigues, K. Suto, H. Nishiya, and N. Kato, “Hybrid method for minimizing service delay in edge cloud computing through vm migration and transmission power control,” IEEE Trans. on Computers, vol. 66, no. 5, pp. 810–819, 2016.

[11] J. Ahmed, M. M. Ashhab, M. A. Razzaque, and M. M. Rahman, “Execution delay-aware task assignment in mobile edge cloud and internet cloud,” in 2019 Int. Conf. on Sustainable Tech. for Industry 4.0 (STI). IEEE, 2019, pp. 1–6.

[12] H. Zhang, X. Lan, Y. Li, L. Cai, and J. Pan, “Efficient computation resource management in mobile edge-cloud computing,” IEEE Internet of Things Journal, vol. 6, no. 2, pp. 3455–3466, 2018.

[13] M. Liu, Y. Mao, S. Leng, and S. Mao, “Full-duplex aided user virtualization for mobile edge computing in 5G networks,” IEEE Access, vol. 6, pp. 2996–3007, 2017.

[14] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. De Rose, and R. Buyya, “Cloudsim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms,” Software: Practice and experience, vol. 41, no. 1, pp. 23–50, 2011.

[15] T. Kurze, M. Klem, D. Bermbach, A. Lenk, S. Tai, and M. Kunze, “Cloud federation,” Cloud Computing, vol. 2011, pp. 32–38, 2011.

[16] A. Celesti, F. Tusa, M. Villari, and A. Pailiauto, “How to enhance cloud architectures to enable cross-federation,” in 2010 IEEE 3rd Int. Conf. on Cloud Computing. IEEE, 2010, pp. 337–345.

[17] A. K. Das, T. Adhikary, M. A. Razzaque, E. J. Cho, and C. S. Hong, “A QoS and profit aware cloud confederation model for IaaS service providers,” in Proc. of the 8th Int. Conf. on Ubiquitous Info. Management and Commun., 2014, pp. 1–7.

[18] Y. Li and M. Chen, “Software-defined network function virtualization: A survey,” IEEE Access, vol. 3, pp. 2542–2553, 2015.

[19] Q. Duan, N. Ansari, and M. Toy, “Software-defined network virtualization: An architectural framework for integrating sdn and nfv for service provisioning in future networks,” IEEE Network, vol. 30, no. 5, pp. 10–16, 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3144960, IEEE Access
[20] C. Gaul, M. Körner, and O. Kao, “Design and implementation of a cloud-federation agent for software defined networking,” in 2015 IEEE Int. Conf. on Cloud Eng., pp. 323–328.

[21] I. Petri, M. Zou, A. R. Zamani, J. Diaz-Montes, O. Rana, and M. Parashar, “Integrating software defined networks within a cloud federation,” in 2015 15th IEEE/ACM Int. Symposium on Cluster, Cloud and Grid Computing, pp. 179–188.

[22] M. Chen and Y. Hao, “Task offloading for mobile edge computing in software defined ultra-dense network,” IEEE Journal on Selected Areas in Commun., vol. 36, no. 3, pp. 587–597, 2018.

[23] J. Xu, B. Palanisamy, H. Ludwig, and Q. Wang, “Zenith: Utility-aware resource allocation for edge computing,” in 2017 IEEE Int. Conf. on edge computing (EDGEC), pp. 47–54.

[24] F. Zafari, K. K. Leung, D. Towsley, P. Basu, A. Swami, and J. Li, “Let’s share: A game-theoretic framework for resource sharing in mobile edge clouds,” IEEE Trans. on Network and Service Management, vol. 18, no. 2, pp. 2107–2122, 2020.

[25] Z. Liu, M. Lin, A. Wierman, S. H. Low, and L. L. Andrew, “Geographical load balancing with renewables,” ACM SIGMETRICS Performance Evaluation Review, vol. 39, no. 3, pp. 62–66, 2011.

[26] H. Zhang, Y. Xiao, S. Bu, D. Niyato, R. Yu, and Z. Han, “Fog computing in multi-tier data center networks: A hierarchical game approach,” in 2016 IEEE Int. Conf. on Commun. (ICC). IEEE, 2016, pp. 1–6.

***