Master's Thesis

Three-Tier Computation Offloading Game in a Multi-User Environment

SuHwan Jung

Department of Electrical Engineering

Graduate School of UNIST

2019

Three-Tier Computation Offloading Game in a Multi-User Environment

SuHwan Jung

Department of Electrical Engineering

Graduate School of UNIST

Three-Tier Computation Offloading Game in a Multi-User Environment

A thesis/dissertation
submitted to the Graduate School of UNIST
in partial fulfillment of the
requirements for the degree of
Master of Science

SuHwan Jung

07. 04. 2019 of submission

Approved by

Advisor
Hyoil Kim

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Three-Tier Computation Offloading Game in a Multi-User Environment

SuHwan Jung

This certifies that the thesis/dissertation of SuHwan Jung is approved.

07/04/2019

signature

Advisor: Prof. Hyoil Kim

signature

typed name: Prof. HyunJong Yang

signature

typed name: Prof. Kyunghan Lee

signature

typed name: Thesis Committee Member #3

signature

typed name: Thesis Committee Member #4;
three signatures total in case of masters

Abstract

Mobile edge cloud computation offloading is proposed as a solution to satisfy mobile devices' desires of processing resource intensive computational tasks with low latency and energy usage. Many previous researches about MEC computation offloading consider only one of a macro cell or a small cell as a choice of offloading. We consider both of cells for the computation offloading.

For considering a practical implementation scenario, different communication schemes are applied to each cell, millimeter wave (mmWave) for communication with a small cell base station and 4G/5G cellular network for communication with a macro cell base station. Because the centralized optimal decision-making problem has several disadvantages, we introduce a decentralized decision-making problem. For finding a mutually satisfactory decision set for all decentralized decision makers, we adopt an ordinal potential game theoretic approach which ensures the existence of a Nash equilibrium and finite time convergence to an equilibrium.

For proving that our game is an ordinal potential game, we analyze dynamics of the game and then find the potential function. Then, we model an asynchronous decision update algorithm for guaranteeing the property of convergence in finite time. Through analyzing the potential function, we find the maximum number of decision slots required for the convergence.

Finally, we propose numerical simulations considering multi-user environments. Our numerical evaluations show that the three-tier offloading game converges in finite time and makes the system-wide overhead decreased.

Keywords - Mobile edge computing (MEC), Vehicular edge computing (VEC), millimeter wave (mmWave), computation offloading, game theory, macro cell MEC, small cell MEC, multi-tier offloading architecture, potential game, potential function, convergence analysis.

Contents

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

List of Figures

I. Introduction

As the computing power of mobile devices is growing at an astonishing speed, the devices are required to process more heavy and complicated applications, such as image processing, artificial intelligence, virtual reality, and so on. The growth, however, has several limitations coming from its physical size constraint and slow evolution of mobile battery technology. Therefore, efficiently processing resource-intensive tasks has become one of the most important problems to be solved by the mobile device technology [1].

Nowadays, computation offloading is viewed as an effective solution to the aforementioned problem. Through offloading resource-intensive tasks to the resource–rich remote cloud servers via wireless access, mobile devices can process heavy tasks with their energy much saved. In particular, mobile cloud computing (MCC) is believed to be the technology to implement computation offloading in a practical manner, which utilizes the remote public cloud servers like Amazon Elastic Compute Cloud (EC2), Google Amazon Web Services (AWS), and Microsoft Azure. Nevertheless, high latency in data transmission between a mobile user and its remote server should be properly taken care of to make the approach viable [2], [3].

To overcome the latency limitation of MCC, Mobile Edge Computing (MEC) has been recently proposed. In MEC, a computing server is located at the edge of the cellular network so that users should not suffer from high latency thanks to the cloud computing resources located near mobile devices. In previous research on MEC-based computation offloading, a MEC server is connected to a 3G or 4G Long-Term Evolution (LTE) macro cell base station through a wired connection [4].

In the 3rd Generation Partnership Project (3GPP)'s fifth generation (5G) of cellular mobile communications project, small cell (a lower power cellular radio access node that operates in both licensed and unlicensed spectrum) is considered as another location of a MEC server. Small cells had not been considered as a computing node until the arise of 5G because of its physical constraint. But recently it has been proposed that the small cell cloud (SCC) can provide enough computation power for mobile devices, especially for services and applications having stringent requirements on latency. Especially, 5G proposes that small cells can provide extremely low latency by utilizing the extremely high frequency (EHF) ranging from 24 to 300 GHz, known as millimeter wave (mmWave), as the spectrum resources for uplink/downlink communications [5].

In this thesis, we propose a three-tier computation offloading mechanism in a multi-user environment. Previous research on computation offloading considered that users offload to an MEC server located either at a macro cell or at a small cell [6]. On the contrary, our work considers a novel offloading decision model where each user can choose either local computing, offloading to a macro cell, or offloading to a small cell, as depicted in Fig. 1. Each choice has its own tradeoff between communication overhead and computation overhead as follows. Local computing has the weakest computing power but doesn't incur any communication overhead. When a user decides to offload, however, the communication overhead of small cell offloading is usually smaller than that of macro cell offloading due to the closer location of a small cell, while the small cell MEC server offers inferior computing power than the macro cell MEC server. Therefore, the key challenge is what offloading option to choose for the minimal combined overhead of computation and communication.

There are two ways to make an offloading decision, centralized vs. decentralized. In centralized decision, a central unit has all the information about users, calculates all user-related overheads, and assigns an optimal action (i.e., compute locally or offload to a specific server) to each user. By contrast, in decentralized decision, each user makes its own decision. The centralized approach has several limitations --- the overhead of gathering all information increases exponentially with the number of users, and making optimal decision for all users is known to be NP-hard [7]. Hence, we adopt the decentralized decision-making in this thesis.

The decentralized decision, however, should confront its own challenge: each user cannot have complete offloading-related information of others due to privacy issues and inter-user communication overhead. To this end, this thesis adopts game theory to model strategic interactions between rational decision makers, using which the users find the common solution from which no user has the incentive to deviate unilaterally. Through this game, each user makes his/her own decision which strikes a balance between user-centric utility and inter-user dynamics.

The main contributions of this thesis are three-fold:

- We present a three-tier computation offloading game in a multi-user environment. Each user can compute his computational task locally, offload to a small cell MEC server of offload to a macro cell MEC server. The decision will be made through a game theoretic approach.

- For the practicality, we implement different communication schemes for a small cell and a macro cell communication through reflecting the latest 5G implementation scenarios. The communication between users and a small cell base station is performed through 5G New Radio (NR) mmWave. For the macro cell communication, each user should be assigned channel resources through the 4G / 5G LTE resource allocation scheme.

- We then formulate the multi-user computation offloading game through using a potential game. For showing the game is a potential game, we need to prove the existence of a potential function. We will introduce the way of proof and show our game should have a potential function. Then numerical results show that the game admits the finite improvement property which proves the existence of a Nash equilibrium and each user achieves the proper optimal decision.

The rest of the thesis is organized as follows. Section II introduces related work, and then Section III describes our system model. In Section IV, we formulate the three-tier computation offloading game and find the potential function. And then we will prove that the potential function reflects the dynamics of our game well. In Section V, we formulate a finite improvement algorithm, and analyze the convergence of our game with mathematical approach. In Section VI, we analyze the performance of the game using numerical simulations and show the convergence. Finally, we conclude our work in Section VII.

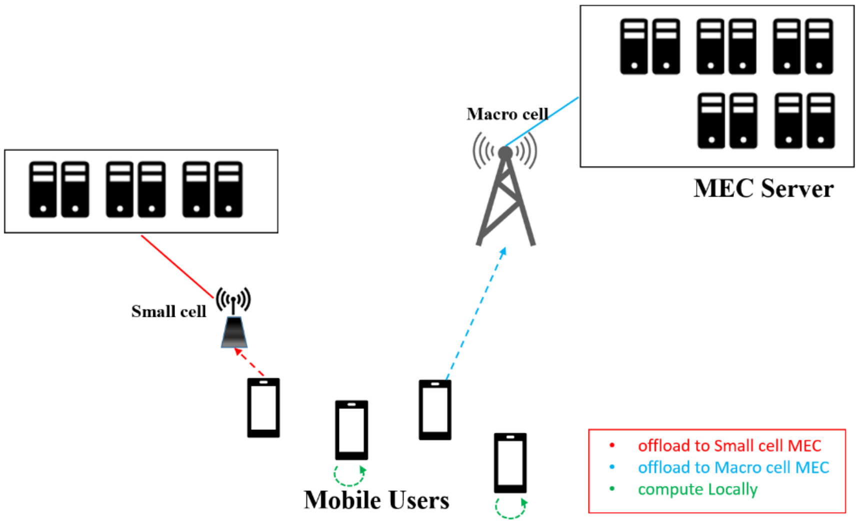

Fig 1. Three-tier computation offloading architecture in a multi-user environment

II. Related Work

Many previous researches have treated the resource allocation problem such as computational resource of MEC server, communication channel resource and so on at the MEC offloading [3],[6]. They have solved the centralized optimization problems for maximizing the revenue of the service operator or minimizing users' overheads at the offloading to a macro cell MEC server or a small cell MEC server. But Chen has proposed that the centralized manner optimization problem at the computation offloading has several problems [7], [10], [11]. As the solution, he suggests the decentralized manner computation offloading decision making mechanism. Different with the centralized manner which a central unit (i.e., a macro cell BS) solve the optimization problem and allocates resources to users, each user who wants to offload to a MEC server decides whether to offload or not according to his own utility function. The goal of the decentralized offloading problem is maximizing each user's utility (i.e., minimizing the overhead occurred by the offloading process) considering latency and energy consumption. Because it is impossible to know other users' information accurately at the decentralized manner, game theory has been suggested as the method for getting a mutually satisfactory decision set for users [8],[9],[12] – [17]. Among many kinds of game theory techniques, researches utilize potential games at the computation offloading decision problems. If a game is a potential game, then it ensures the existence of a Nash equilibrium and the convergence of a game in finite time. Because latency is an important factor at the offloading problem, the finite time convergence is treated as the key factor to ensure the practicality of offloading decision problems [18], [19].

As 5G NR mmWave scheme has been proposed, the abundant bandwidth of 24 ~ 300 GHz is considered as the solution for extremely high data rate and low latency such as 10 Gbps peak data rate at uplink and 1 ~ 5ms end-to-end latency. Because this extremely low end-to-end latency fits well to computation offloading, many researches consider mmWave as the communication scheme for offloading. Furthermore, 5G aims to manage multiple communication scenarios such as Device-to-device (D2D), Vehicle-to-vehicle (V2V) and Vehicle-to-infra (V2I) [20], [21]. However, 5G mmWave has several critical issues for practical implementations. Because 5G mmWave utilizes extremely high frequency, it is too sensitive to blockage issues. The path loss model of mmWave is completely different according to the Line-of-sight (LOS) and Non-line-of-sight (NLOS) situations. So, there are many researches for modeling proper mmWave channel models considering many environments such as rural, dense urban, inside and so on [22].

To the best of our knowledge, there are only few researches which treat multi-tier structure considering both of macro cell and small cell for offloading. And those works consider the centralized resource allocation problems [23]. Different with existing works, our thesis proposes the three-tier computation offloading decision game in a multi-user environment. We treat the multi-tier structure in the decentralized manner. Each user calculates the best decision which minimizes the overhead occurred by the computation among three choices, local computing, small cell offloading and macro cell offloading. Through utilizing game theory, we will show that all users should get mutually satisfactory decisions without harming other users' utility. For the practicality, we will utilize different communication schemes for the small cell communication and the macro cell communication. Because there is no standard way to find the potential function, many researches treat potential games don't propose the potential function. But we will propose the potential function which shows the dynamics of our game perfectly.

III. System Model

3.1 System Environment

We consider a three-tier computation offloading system, consisting of a set $\mathcal{N} = \{1,2,\ldots,N\}$ of mobile devices with computationally intensive tasks, an MEC server located at a small cell base station (BS), and an MEC server located at a macro cell BS. The position of users should change dynamically in principle, and thus we need to consider the factors related to user mobility like handover, live migration of Virtual Machines (VMs) on a MEC server, the change of channel state, and so on. But if we consider all those factors, it is hard to get a tractable analysis. Aligned with existing works, we consider a quasi-static scenario where $\mathcal{N}$ doesn't change during an offloading period, leaving the consideration of mobility as our future work.

Among the technologies that treat computational offloading as a core part, autonomous driving is a representative case since it handles a lot of computationally intensive tasks with limited computing power. Because driving has a direct impact on people's safety, autonomous driving processors need to make accurate decisions through analyzing sensor data. Hence, the unit cost of production of an autonomous driving car is too high if all tasks should be performed locally. Also, it needs to consider the practical road environment such as multiple mobile devices. In fact, our three-tier offloading system dovetails nicely with these needs, where $\mathcal{N}$ becomes a set of autonomous driving vehicles [24].

3.2 Communication Model

We consider two different communication schemes for each type of base station according to several researches and white papers which treat the practical use of computation offloading. The communication between a UE and the small cell base station s is achieved using 5G millimeter wave (mmWave). The other case considering the macro cell base station m is performed through 4G/5G cellular network. We assume that mobile devices are equipped with multiple network interfaces which consists of mmWave and cellular interfaces. In addition, the downlink communication time is assumed negligible because the computation result in many applications is much smaller than the input data size.

We denote by $a_n \in \{0,1,2\}$ the computation offloading decision of user n. $a_n = 0$ means that the user n decides to compute its task locally. If $a_n = 1$, the user n offloads its task to the small cell MEC server. Finally, $a_n = 2$ means that the user n offloads its task to the macro cell MEC server. Given the decision profile of all users, denoted by $\boldsymbol{a} = (a_1, a_2, \ldots, a_N)$, we can calculate the uplink transmission rate between a user and a specific base station.

3.2.1 Communication between a UE and the small cell base station s

mmWave is the wireless operating system at the millimeter wave spectrum, ranging from 24GHz to 300GHz [5]. Because customers require fast transmission everywhere, wireless devices and technologies has grown greatly to meet the need. But the wireless spectrum below 6 GHz is not enough to meet the requirements of users. So, the 3GPP has a goal of improvement from LTE communication schemes (i.e., LTE, LTE advanced, and LTE pro) by utilizing the spectrum above 6 GHz. While mmWave's abundant channel bandwidth seems to solve the problem of spectrum shortfall below 6 GHz, its high path loss affected by the environment requires more complex infrastructure than the legacy mobile environment. In this vein, beamforming has been proposed as a solution for eliminating high interference, where highly directed signals enhance the signal to interference ratio (SIR) so that each user can fully utilize the channel bandwidth.

Because the signal between a UE and the small cell base station should be beamformed, user n can fully utilize the channel bandwidth provided by the small cell base station. Therefore, the uplink transmission rate to the small cell is not affected by other users' decision profile, and we can formulate the uplink transmission rate r_n^s of user n with $a_n = 1$ as

$$r_n^s = W^s \cdot log_2(1 + SNR_n^s) \qquad (1)$$

where W^s is the channel bandwidth. We assume that there is no inter-cell interference. For calculating the signal to noise ratio (SNR) between a user n and the base station s, we need to know the path loss model based on the mmWave environment. According to a previous research [25] the omnidirectional large-scale path loss model for a 28 GHz Line of Sight (LOS) channel is given as

$$\mathrm{PL}_{28\mathrm{GHz}}(LOS)[dB](d) = (61.4\,dB - 49\,dB) + 21\log(d) + \chi_\sigma\ [\sigma = 3.6\,dB] \qquad (2)$$

Here, $d(m)$ is the distance between a user n and the station s, and χ_σ is a zero mean Gaussian random variable with a standard deviation of σ in decibels. The meaning of $(-49\,dB)$ in the equation is came from 24.5 dBi Tx and Rx antenna gains. Then, we can calculate r_n^s using (1) and (2).

3.2.2 Communication between a UE and the macro cell base station m

To treat 4G/5G cellular network in a multi-user environment, we need to consider the bandwidth allocation algorithm which fits in our system model. Because the uplink transmission scheme of 5G is similar to the LTE uplink transmission scheme, we can implement the LTE uplink scheme for calculating the data rate. The LTE transmission scheme is based on SC-FDMA (Single Carrier Frequency Division Multiple Access) with cyclic prefix. In LTE, users obtain uplink transmission resources according to eNodeB's scheduling, where the resources are a multiple of physical resource blocks (PRBs). A PRB includes 12 subcarriers in the frequency domain and a 0.5 ms long slot in the time domain [26].

In our thesis, we transform the way of channel resource allocation from the PRB allocation to the channel bandwidth allocation. Because the channel bandwidth provided by m is limited, we need to perform a proper allocation algorithm. We assume each user wants to finish their uplink transmission between the station m within 1 ms, where 1 ms corresponds to the LTE transmission time interval (TTI). Accordingly, we can calculate the required amount of bandwidth W_n by each user for completing the uplink transmission within 1 ms as

$$W_n = \frac{10^3 \cdot b_n}{log2(1 + SNR_n^m)} \tag{3}$$

Here, b_n means the input data size of user n and SNR_n^m denotes the signal to noise ratio between user n and the base station m. W_n should be smaller than W^m, the total bandwidth provided by the base station m. Differing from the small cell uplink transmission, the channel bandwidth W^m is shared with other users who offload to the macro cell MEC server. For fair bandwidth allocation, the amount of channel bandwidth allocated to each user is decided proportionally to the bandwidth demand such as

$$\frac{W^m \cdot W_n}{W_n + \sum_{i \neq n} W_i \cdot I_{\{a_i=2\}}} \tag{4}$$

The aforementioned bandwidth allocation algorithm considers how the actual offloading service is provided. We assume that the actual service will be served by Mobile Network Operators (MNOs), and that all users who use the plan pay the same cost, are allocated similar data rates, but can use a different amount of data. According to this service, we assume that W_m is reserved for the offloading service and users share the limited resources without any priority. Note that in (4), $\sum_{i \neq n} W_i \cdot I_{\{a_i=2\}}$ implies the users who decide to offload to the macro cell MEC server. Finally, we can calculate the uplink transmission rate of user n who has chosen the macro cell for offloading as

$$r_n^m(\boldsymbol{a}) = \frac{W^m \cdot W_n}{W_n + \sum_{i \neq n} W_i \cdot I_{\{a_i=2\}}} \cdot log_2(1 + SNR_n^m) \tag{5}$$

Note that if too many mobile devices choose to offload to the macro cell MEC server, they lead to very low uplink transmission rates.

3.3 Computation Model

We consider that mobile device n has a computationally intensive task with an input data size of b_n $(bits)$ and a required amount of CPU cycles of c_n $(cycles)$. Using these two variables and the uplink transmission rates calculated earlier, we next discuss the overhead of local computing, small cell offloading, and macro cell offloading.

3.3.1 Local Computing

When a user chooses local computing, there are two kinds of overhead involved, the computing time and the computing energy consumption at its own device. Set f_n^l (i.e., CPU cycles per second) as the computing power of mobile device n. Then, we can calculate the local computing time as

$$t_n^l = \frac{c_n}{f_n^l} \tag{6}$$

For the energy consumption, we need another parameter ϵ_n which denotes the energy consumption per CPU cycle. Then, mobile device n's energy consumption becomes [27]

$$e_n^l = \epsilon_n \cdot c_n \tag{7}$$

According to (6) and (7), we can formulate the local computing overhead in terms of computation time and energy consumption as

$$Z_n^l = e_n^l + \lambda \cdot t_n^l \tag{8}$$

$$= c_n \cdot (\epsilon_n + \frac{\lambda}{f_n^l})$$

where $\lambda \in [0,1]$ is the weighting factor between computing time and energy.

3.3.2 Offloading to the small cell MEC server

When a user decides to offload his task to the small cell base station, the task is transmitted using the mmWave band and then performed at the MEC server connected to s. Because the scale of a small cell is limited, we need to consider weaker computing power than today's usual cloud servers.

We need to consider the additional overhead came from wireless transmission different with local computing. Using (1), we can calculate the time overhead incurred by uplink transmission to the base station s as

$$t_{n,o}^s = \frac{b_n}{r_n^s} \tag{9}$$

Set P_n^s as the transmit power of device n to the base station s. Then, the transmission energy consumption can be derived as

$$e_{n,o}^s = P_n^s \cdot \frac{b_n}{r_n^s} \tag{10}$$

Finally, we need to calculate the computing time at the small cell MEC server. Setting f^s as the computing power of the small cell MEC server, the computing time is given as

$$t_{n,c}^s = \frac{c_n}{f^s} \tag{11}$$

According to (9), (10) and (11), we can formulate the overhead of time and energy consumption as

$$Z_n^s = e_{n,o}^s + \lambda \cdot (t_{n,o}^s + t_{n,c}^s) \tag{12}$$

$$= \frac{b_n}{r_n^s} \cdot (P_n^s + \lambda) + \lambda \cdot \frac{c_n}{f^s}$$

3.3.3 Offloading to the macro cell MEC server

Like the small cell offloading case, the overhead of the macro cell offloading consists of the transmission time consumption, the transmission energy consumption, and the computing time consumption at the macro cell MEC server.

We denote by f^m the computation capability of the macro cell MEC server and by P_n^m the transmission power from mobile device n to the base station m. Then, we can formulate the overhead functions of macro cell offloading using (3) and (5) as

$$t_{n,o}^m = \frac{b_n}{r_n^m(\boldsymbol{a})} \tag{13}$$

$$e_{n,o}^m = P_n^m \cdot \frac{b_n}{r_n^m(\boldsymbol{a})} \tag{14}$$

$$t_{n,c}^m = \frac{c_n}{f^m} \tag{15}$$

Finally, we can compute the overhead of time and energy consumption using (13), (14), and (15), as

$$\mathrm{Z_n^m} = e_{n,o}^m + \lambda \cdot \left(t_{n,o}^m + t_{n,c}^m\right) \tag{16}$$

$$= \frac{\mathrm{b_n}}{r_n^m(\boldsymbol{a})} \cdot (P_n^m + \lambda) + \lambda \cdot \frac{c_n}{f^m}$$

IV. Game Formulation

In this section, we consider how each user can make the optimal decision at the three-tier computation offloading scenario. We first formulate the game as a strategic game form. The key point at formulating the game is that each user wants to minimize his overhead. Then, we will introduce potential games which is the powerful tool for ensuring the existence of an equilibrium and finite time convergence to an equilibrium. Through analyzing the dynamics of our game, we should find the potential function of the three-tier computation offloading game in a multi-user environment. Then, we will provide proofs that the suggested potential function reflects the dynamics of the game well.

From the communication model in Section III, we can observe that the local computing and the small cell offloading overheads of a user are not affected by other users. So, each user can individually decide a better one between them. A user's macro cell offloading overhead, however, is influenced by others. If the number of users to offload to the macro cell increases, the per-user channel bandwidth allocation by the macro cell BS tends to decrease, and accordingly macro cell users should achieve lower data rates. In Eqs. (13), (14), and (16), the decrease of a data rate causes the increase of the transmission time and energy overhead so that the total overhead of the macro cell offloading becomes increased. In this situation, users who consider the macro cell MEC server should reconsider their decisions. So, the most important thing for making an optimal decision is to predict other users' decisions precisely.

According to [10], solving the optimization problem for minimizing system-wide computation overhead through centralized manner is NP-hard, since it consists of a combinatorial optimization over the multi-dimensional discrete space (i.e., $\{0,1,..,M\}^N$). It means that the centralized offloading management is very complex. As the number of offloading users increases, the offloading service provider should collect massive amount of data related to offloading from mobile device users. Because users have different properties (i.e., input data size, local computing power), the information gathering process causes huge overhead for centralized management. As a solution, we consider the decentralized manner decision making process such that each user makes his own decision. But, giving other users' information related to offloading directly to a specific user faces the privacy problem. Game theory is very powerful tool for this situation. It analyzes interactions between decision makers who have different interests and finding a mutually satisfactory decision set for a system. In some decentralized scenarios, game theory helps to formulate tractable and analytical models so that we can find the optimal decision set for users. For applying a game theoretical approach, we need to formulate our game with explanations about several fundamental concepts related to a game theory.

4.1 Game Formulation as a strategic game

The most basic form of a non-cooperative decentralized game is the strategic game [18]. For formulating a game as a strategic form, we need to define three fundamental elements clearly, the set of players, the strategies of each player (i.e., all actions that a player can select) and the utility functions for players that show change of the payoffs (or overheads) that a player will be awarded for taking a certain action given the other players' strategy profiles. In our game, the set of players can be defined as $\mathcal{N} = \{1,2, \dots, N\}$. Then, we can define each user's decision as a_n. For user n, we need to define the set of strategies $a_n \in \mathcal{A}_n \triangleq \{0,1,2\}$ where $a_n = 0$ means the local computing, $a_n = 1$ means the offloading to the small cell MEC server and finally $a_n = 2$ means the offloading to the macro cell MEC server. By the definition [19], the strategy space $\mathbb{S}$ is defined as the Cartesian products of all individual strategy sets, i.e.,

$$\mathbb{S} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_N \tag{17}$$

Each element (i.e., $(\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n) \in \mathbb{S}$) is said to be a strategy profile. Set $a_{-n} = (a_1, a_2, \dots, a_{n-1}, a_{n+1}, \dots, a_N)$ is the decisions of all other players except a player n. Using these definitions, we can formulate the strategies of each user and the overhead functions considering the interaction with other users. First, the strategies of each user can be formulated as

$$\min_{a_n \in \mathcal{A}_n \triangleq \{0,1,2\}} Z_n(a_n, a_{-n}) \tag{18}$$

Each user wants to minimize their overhead considering given other users' decision profiles. According to (8), (12) and (16), we can formulate Z_n which is the overhead function showing interactions with other players as

$$Z_n(a_n, a_{-n}) = \begin{cases} Z_n^l, & \text{if } a_n = 0 \\ Z_n^s, & \text{if } a_n = 1 \\ Z_n^m, & \text{if } a_n = 2 \end{cases} \tag{19}$$

Finally, we can formulate our game as a strategic game denoted as $\mathcal{G} = [\mathcal{N}, \mathbb{S}, \{Z_n\}_{n \in \mathcal{N}}]$. For analyzing the outcomes for a game, the concept of Nash equilibrium should be introduced. By the definition [18], a Nash equilibrium is a strategy profile such that if other users' strategies remain unchanged, no player will not change his current strategy. The strategy profile $a^* = (a_1^*, a_2^*, \dots, a_n^*)$ is a pure-strategy Nash equilibrium if and only if

$$Z_n(a_n^*, a_{-n}^*) \le Z_n(a_n, a_{-n}^*) \quad \forall a_n \in \boldsymbol{a}, \forall n \in \mathcal{N} \tag{20}$$

At a Nash equilibrium, no user can improve, so it is also called as a stable operating point. Also, it is called as a mutually satisfactory decision set. By these reasons, the goal of a certain game is the proof of Nash equilibrium existence and find Nash equilibriums.

4.2 Potential Game

We next introduce the method of proving the existence of Nash equilibrium of the three-tier computation offloading game. In game theory, a game is considered as a potential game if it has a potential function which maps each user's change of strategies to change of real numbers with the form of a single global function. A potential game is a very powerful tool for showing the existence of Nash equilibrium because of two properties it has. If a game is a potential game, it always ensures the existence of Nash equilibrium. Besides, any asynchronous improvement (i.e., one user makes a decision at each given decision time) leads to convergence in finite time if it is a potential game. So, we can state that our game has a Nash equilibrium and can achieve it in finite time if our game is a potential game. For proving this, we need to define the definition of a potential game first.

By the definition [19], $\mathcal{G}$ is an ordinal potential game if and only if a potential function ϕ with the given decision profile $\boldsymbol{a}$ $\phi(\boldsymbol{a}): \mathbb{S} \mapsto \mathbb{R}$ exists such that, $\forall n \in \mathcal{N}$:

$$Z_n(a_n, a_{-n}) - Z_n(a'_n, a_{-n}) > 0 \Leftrightarrow \phi(a_n, a_{-n}) - \phi(a'_n, a_{-n}) > 0 \qquad (21)$$

$$\forall a_n, a'_n \in \mathcal{A}_n$$

An ordinal potential game is established when the sign of overhead difference caused by the change of a user's decision is same with the sign of potential function difference caused by the strategy change. In other words, if a player changes his decision in direction of decreasing his overhead, it should increase the output of a potential function and vice versa. Moreover, there is another important condition for ordinal potential games. By the Voorneveld [32], the game $\mathcal{G}$ is an ordinal potential game if and only if $\mathbb{S}$ has no week improvement cycle. In other words, an ordinal potential game is established if there is no cycle through the process of users' decision update and users should decide towards decreasing their overhead.

There is no definite formula to perform a potential function from a game. The only way for finding a potential function is a heuristic formulation from the system model of a game. To formulate the potential function for our game, we need to consider the dynamics of our system model. We should treat 3 cases, local computing vs. small cell offloading, local computing vs. macro cell offloading and small cell offloading vs. macro cell offloading for observing the dynamics. Through catching the conditions when each user changes his decision, we can find the form of our potential function.

4.2.1 Local computing vs. Small cell offloading

Because users want to minimize their overhead for computation, each user should choose the smaller one between local computing overhead ($a_n = 0$) and small cell offloading overhead ($a_n = 1$). We can formulate the dynamics considering local and small cell MEC for user n as

$$Z_n(a_n, a_{-n}) \begin{cases} Z_n^l, & \text{if } Z_n^l < Z_n^s \\ Z_n^s, & \text{if } Z_n^l > Z_n^s \end{cases} \qquad (22)$$

The local computing overhead and small offloading overhead are deterministic values not affected by other users' decisions. So, we can easily compare those two overheads. Using Eq.(4), (8) and (12), $Z_n^l < Z_n^s$ becomes

$$c_n \cdot \left(\epsilon_n + \frac{\lambda}{f_n^l}\right) < \frac{b_n}{W^s \cdot log_2(1 + SNR_n^s)} \cdot (P_n^s + \lambda) + \lambda \cdot \frac{c_n}{f^s} \tag{23}$$

4.2.2 Local computing vs. Macro cell offloading

Using Eq. (5), (8) and (16), we need to compare the user n's local computing overhead and the macro cell offloading overhead. If $Z_n^l < Z_n^m$,

$$c_n \cdot \left(\epsilon_n + \frac{\lambda}{f_n^l}\right) < \frac{\left(W_n + \sum_{i \neq n} W_i \cdot I_{\{a_i=2\}}\right) \cdot b_n}{W^m \cdot W_n \cdot log_2(1 + SNR_n^m)} \cdot (P_n^m + \lambda) + \lambda \cdot \frac{c_n}{f^m} \tag{24}$$

$$c_n \cdot \left(\epsilon_n + \frac{\lambda}{f_n^l} - \frac{\lambda}{f^m}\right) < \frac{\left(W_n + \sum_{i \neq n} W_i \cdot I_{\{a_i=2\}}\right) \cdot b_n}{W^m \cdot W_n \cdot log_2(1 + SNR_n^m)} \cdot (P_n^m + \lambda) \tag{25}$$

$$\frac{W^m \cdot W_n \cdot c_n \cdot (\epsilon_n + \frac{\lambda}{f_n^l} - \frac{\lambda}{f^m})}{b_n \cdot (P_n^m + \lambda)} \cdot \log_2(1 + SNR_n^m) < W_n + \sum_{i \neq n} W_i \cdot I_{\{a_i=2\}} \tag{26}$$

Using the left side of equation (26), we should define the local computing vs. macro cell offloading threshold $T_n^{(l,m)}$ as

$$T_n^{(l,m)} = \frac{W^m \cdot W_n \cdot c_n \cdot (\epsilon_n + \frac{\lambda}{f_n^l} - \frac{\lambda}{f^m})}{b_n \cdot (P_n^m + \lambda)} \cdot log_2(1 + SNR_n^m) \tag{27}$$

Then we can express Eq. (26) as

$$T_n^{(l,m)} < W_n + \sum_{i \neq n} W_i \cdot I_{\{a_i=2\}} \tag{28}$$

The equation (28) states that if the sum of required bandwidth by users who has chosen macro cell offloading is smaller than the threshold $T_n^{(l,m)}$, user n should choose macro cell offloading $(a_n = 2)$.

If $Z_n^l > Z_n^m$,

$$T_n^{(l,m)} > W_n + \sum_{i \neq n} W_i \cdot I_{\{a_i=2\}} \tag{29}$$

4.2.3 Small cell offloading vs. Macro cell offloading

At previous subsection, we find the threshold considering local vs. macro cell. Now, we need to find the threshold considering small cell vs. macro cell case using Eq. (1), (5), (12) and (16). For user n, if $Z_n^s < Z_n^m$,

$$\frac{b_n}{W^s \cdot \log_2(1+SNR_n^s)} \cdot (P_n^s + \lambda) + \lambda \cdot \frac{c_n}{f^s} < \frac{\left(W_n + \sum_{i\neq n} W_i \cdot I_{\{a_i=2\}}\right) \cdot b_n}{W^m \cdot W_n \cdot \log_2(1+SNR_n^m)} \cdot (\mathrm{P}_n^m + \lambda) + \lambda \cdot \frac{c_n}{f^m} \tag{30}$$

$$\frac{W^m \cdot W_n}{b_n \cdot (P_n^m + \lambda)} \cdot \log_2(1+SNR_n^m) \cdot (\frac{b_n}{W^s \cdot \log_2(1+SNR_n^s)} \cdot (P_n^s + \lambda) + \lambda \cdot c_n \cdot \left(\frac{1}{f_s} - \frac{1}{f^m}\right))$$

$$< W_n + \sum_{i\neq n} W_i \cdot I_{\{a_i=2\}} \tag{31}$$

We can define the small cell offloading vs. macro cell offloading threshold $T_n^{(s,m)}$ using the left side of Equation (31).

$$T_n^{(s,m)} = \frac{W^m \cdot W_n}{b_n \cdot (P_n^m + \lambda)} \cdot \log_2(1+SNR_n^m) \cdot (\frac{b_n}{W^s \cdot \log_2(1+SNR_n^s)} \cdot (P_n^s + \lambda) + \lambda \cdot c_n \cdot \left(\frac{1}{f_s} - \frac{1}{f^m}\right)) \tag{32}$$

Finally, we can express the equation (31) as

$$T_n^{(s,m)} < W_n + \sum_{i\neq n} W_i \cdot I_{\{a_i=2\}} \tag{33}$$

If $Z_n^s > Z_n^m$,

$$T_n^{(s,m)} > W_n + \sum_{i\neq n} W_i \cdot I_{\{a_i=2\}} \tag{34}$$

4.2.4 Finding the potential function

As mentioned previously, we need to understand the dynamics of our game for formulating the potential function. For user k, there are 6 cases of possible decision change. Consider that user k changes his decision from a_k to a'_k for decreasing his overhead. At first, user k changes his decision from local computing ($a_k = 0$) to small cell offloading ($a'_k = 1$) when $T_k^{(l,m)}$ is larger than

$T_k^{(s,m)}$.

$$a_k = 0 \rightarrow a_k' = 1 \text{ when } T_k^{(l,m)} - T_k^{(s,m)} > 0 \tag{35}$$

On the other hand, user k changes his decision from small cell offloading ($a_k = 1$) to local computing ($a'_k = 0$) because the local computing overhead is smaller than small cell offloading overhead.

$$a_k = 1 \rightarrow a_k' = 0 \text{ when } T_k^{(s,m)} - T_k^{(l,m)} > 0 \tag{36}$$

Using Eq. (28) and (29), we can observe the dynamics between local computing decision and macro cell offloading decision like

$$a_k = 0 \rightarrow a_k' = 2 \text{ when } T_k^{(l,m)} - \left(W_k + \sum\nolimits_{i \neq k} W_i \cdot I_{\{a_i=2\}}\right) > 0 \tag{37}$$

$$a_k = 2 \rightarrow a_k' = 0 \text{ when } \left(W_k + \sum\nolimits_{i \neq k} W_i \cdot I_{\{a_i=2\}}\right) - T_k^{(l,m)} > 0 \tag{38}$$

Using Eq. (33) and (34), we can find the dynamics between small cell offloading decision vs. macro cell offloading decision in the same context with Eq. (37) and (38).

$$a_k = 1 \rightarrow a_k' = 2 \text{ when } T_k^{(s,m)} - \left(W_k + \sum\nolimits_{i \neq k} W_i \cdot I_{\{a_i=2\}}\right) > 0 \tag{39}$$

$$a_k = 2 \rightarrow a_k' = 1 \text{ when } \left(W_k + \sum\nolimits_{i \neq k} W_i \cdot I_{\{a_i=2\}}\right) - T_k^{(s,m)} > 0 \tag{40}$$

Using these relationships, we can show that the three-tier computation offloading decision game is a potential game by constructing the potential function ϕ with given decision profile $\boldsymbol{a}$ as

$$\begin{aligned}\phi(\boldsymbol{a}) = &\sum_{i=1}^{N} W_i \cdot T_i^{(s,m)} \cdot I_{\{a_i=1\}} \\ &+ \sum_{i=1}^{N} W_i \cdot T_i^{(l,m)} \cdot I_{\{a_i=0\}} \\ &+ \sum_{i=1}^{N} W_i^2 \cdot I_{\{a_i=2\}} \\ &+ \frac{1}{2}\sum_{i=1}^{N}(\sum_{j \neq i} W_j \cdot I_{\{a_j=2\}}) W_i \cdot I_{\{a_i=2\}}\end{aligned} \tag{41}$$

4.2.5 Proof

Finally, if we can prove that our potential function shows the dynamics according to each user's decision change well, then we can say that our game is a potential game. We need to consider 3 cases. Case 1 is local computing vs. small cell offloading ($a_k = 0 \rightarrow a'_k = 1$ & $a_k = 1 \rightarrow a'_k = 0$). Case 2 is local computing vs. macro cell offloading ($a_k = 0 \rightarrow a'_k = 2$ & $a_k = 2 \rightarrow a'_k = 0$). Lastly, case 3 is small cell offloading vs. macro cell offloading ($a_k = 1 \rightarrow a'_k = 2$ & $a_k = 2 \rightarrow a'_k = 1$).

For case 1, we need to find two output values of potential function according to user k's decisions. When his decision is local computing ($a_k = 0$),

$$\begin{aligned}\phi(a_k, a_{-k}) = &\sum_{i \neq k} W_i \cdot T_i^{(s,m)} \cdot I_{\{a_i=1\}} \\ &+ W_k \cdot T_k^{(l,m)} + \sum_{i \neq k} W_i \cdot T_i^{(l,m)} \cdot I_{\{a_i=0\}} \\ &+ \sum_{i \neq k} W_i^2 \cdot I_{\{a_i=2\}} \\ &+ \frac{1}{2} \sum_{i \neq k} \Big(\sum_{\substack{j \neq i \\ j \neq k}} W_j \cdot I_{\{a_j=2\}} \Big) W_i \cdot I_{\{a_i=2\}}\end{aligned} \tag{42}$$

In the same context, we can calculate the potential function output when his decision is $a_k = 1$ as

$$\begin{aligned}\phi(a_k, a_{-k}) = W_k \cdot T_k^{(s,m)} &+ \sum_{i \neq k} W_i \cdot T_i^{(s,m)} \cdot I_{\{a_i=1\}} \\ &+ \sum_{i \neq k} W_i \cdot T_i^{(l,m)} \cdot I_{\{a_i=0\}} \\ &+ \sum_{i \neq k} W_i^2 \cdot I_{\{a_i=2\}} \\ &+ \frac{1}{2} \sum_{i \neq k} \Big(\sum_{\substack{j \neq i \\ j \neq k}} W_j \cdot I_{\{a_j=2\}} \Big) W_i \cdot I_{\{a_i=2\}}\end{aligned} \tag{43}$$

Then, we need to consider two cases, $a_k = 0 \rightarrow a'_k = 1$ & $a_k = 1 \rightarrow a'_k = 0$. When $a_k = 0 \rightarrow a'_k = 1$, it means that user k's small cell offloading overhead is smaller than local computing overhead, so user k changes his decision. Using Eq. (42) and (43), we can observe that

$$\begin{aligned}\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) &= W_k \cdot T_k^{(l,m)} - W_k \cdot T_k^{(s,m)} \\ &= W_k \cdot (T_k^{(l,m)} - T_k^{(s,m)})\end{aligned} \tag{44}$$

We know that W_k is positive. Then by the given condition Eq. (35), we can state that Eq. (44) is positive ($\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) > 0$). So, it satisfies the definition of ordinal potential games.

User k changes his decision from local computing to small cell offloading because it causes the decrease at his overhead (i.e., $Z_k^l - Z_k^s > 0$). And we can observe it through the potential function $(\text{i.e.}, \phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) > 0)$. Similarly, when $a_k = 1 \rightarrow a'_k = 0$, we can find that

$$\begin{aligned}\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) &= W_k \cdot T_k^{(s,m)} - W_k \cdot T_k^{(l,m)} \\ &= W_k \cdot (T_k^{(s,m)} - T_k^{(l,m)}) \qquad (45)\end{aligned}$$

By the given condition (36), we can state that $\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) > 0$.

For case 2, we can use the same proof method with case 1. First, we should calculate two potential function outputs when $a_k = 0$ and $a_k = 2$. When user k's decision is macro cell offloading, the potential function output becomes

$$\begin{aligned}\phi(a_k, a_{-k}) = &\sum_{i \neq k} W_i \cdot T_i^{(s,m)} \cdot I_{\{a_i=1\}} \\ &+ \sum_{i \neq k} W_i \cdot T_i^{(l,m)} \cdot I_{\{a_i=0\}} \\ &+ W_k^2 + \sum_{i \neq k} W_i^2 \cdot I_{\{a_i=2\}} \\ &+ W_k \cdot \left(\sum_{i \neq k} W_i \cdot I_{\{a_i=2\}} \right) \\ &+ \frac{1}{2} \sum_{i \neq k} \Big(\sum_{\substack{j \neq i \\ j \neq k}} W_j \cdot I_{\{a_j=2\}} \Big) W_i \cdot I_{\{a_i=2\}} \qquad (46)\end{aligned}$$

When user k changes his decision from local computing to macro cell offloading ($a_k = 0 \rightarrow a'_k = 2$), we know that $T_k^{(l,m)} - \left(W_k + \sum_{i \neq k} W_i \cdot I_{\{a_i=2\}}\right) > 0$ by the given condition Eq. (37). Using Eq. (42) and (46), we can observe that

$$\begin{aligned}\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) &= W_k \cdot T_k^{(l,m)} - W_k^2 - W_k \cdot \left(\sum_{i \neq k} W_i \cdot I_{\{a_i=2\}} \right) \\ &= W_k \cdot \left(T_k^{(l,m)} - W_k - \sum_{i \neq k} W_i \cdot I_{\{a_i=2\}} \right) > 0 \qquad (47)\end{aligned}$$

By the given condition Eq. (38), we can also state that the potential function ϕ shows the dynamics of our game well when $a_k = 2 \rightarrow a'_k = 0$ as

$$\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) = W_k^2 + W_k \cdot \left(\sum_{i \neq k} W_i \cdot I_{\{a_i=2\}}\right) - W_k \cdot T_k^{(l,m)}$$

$$= W_k \cdot \left(W_k + \sum_{i \neq k} W_i \cdot I_{\{a_i=2\}} - T_k^{(l,m)}\right) > 0 \qquad (48)$$

For case 3, by the similar context with case 1 and case 2, when $a_k = 1 \rightarrow a'_k = 2$, we can show that $\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) > 0$ through using the given condition Eq. (39) and potential function outputs Eq. (43) & (46).

$$\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) = W_k \cdot T_k^{(s,m)} - W_k^2 - W_k \cdot \left(\sum_{i \neq k} W_i \cdot I_{\{a_i=2\}}\right)$$

$$= W_k \cdot \left(T_k^{(s,m)} - W_k - \sum_{i \neq k} W_i \cdot I_{\{a_i=2\}}\right) > 0 \qquad (49)$$

When user k changes his decision from macro cell offloading to small cell offloading ($a_k = 2 \rightarrow a'_k = 1$), we can observe that $\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) > 0$ by the given condition Eq. (40).

$$\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) = W_k^2 + W_k \cdot \left(\sum_{i \neq k} W_i \cdot I_{\{a_i=2\}}\right) - W_k \cdot T_k^{(s,m)}$$

$$= W_k \cdot \left(W_k + \sum_{i \neq k} W_i \cdot I_{\{a_i=2\}} - T_k^{(s,m)}\right) > 0 \qquad (50)$$

In conclusion, the three-tier computation offloading game is an ordinal potential game because the potential function can show the dynamics of the game well. So, we can state that our game has a Nash equilibrium and it will converge in finite time by the definition of an ordinal potential game.

V. Offloading Algorithm and Convergence Analysis

By the definition of ordinal potential games mentioned at the previous section, a game should be asynchronous improvement process for ensuring finite time convergence to a Nash equilibrium [18], [19]. It means that only one user should change his decision for decreasing his overhead at each decision slot. In this section, we propose the three-tier computation offloading game algorithm showing asynchronous improvement property. And then we analyze the convergence of our game through mathematical approach using the potential function.

Algorithm: Three-tier Computation Offloading Decision Game

---

initialization:
each mobile device user n chooses the initial offloading decision as local computing $a_n(0) = 0$
end initialization
repeat for each user n and each decision slot t in parallel:
    if $a_n(t) == 2$
        transmit the information of W_n to the macro cell BS
    end if
    receive the information of $\sum W_i \cdot I_{\{a_i=2\}}$ from the macro cell BS
        compute $\sum_{i \neq n} W_i \cdot I_{\{a_i=2\}}$
    compute $a'_n(t) = \underset{a_n \in \mathcal{A}_n \triangleq \{0,1,2\}}{\arg\min} Z_n(a_n, a_{-n})$
    if $a_n(t) \neq a'_n(t)$ **then**
        send Change Decision message to the macro cell BS
        if receive the approval message form the macro cell BS **then**
            update $a_n(t+1) = a'_n(t)$
        else keep the original decision $a_n(t+1) = a_n(t)$
        end if
    else keep the original $a_n(t+1) = a_n(t)$
    end if
until CONVERGENCE message is received from the macro cell BS

5.1 Three-tier computation offloading decision game algorithm

There are two important points we need to talk about. The first point is how each user can know or predict other users' decision and the other point is how the game can guarantee that only one user can change his decision at each decision slot. Through the three-tier computation offloading decision game algorithm, we will explain those points.

Through the communication model (i.e., Eq. (4)), we explained that the interaction with other users is occurred at the communication between mobile devices and the macro cell BS using 4G/5G cellular network. Because the cellular channel bandwidth is limited, the allocated bandwidth to each user who wants to offload to the macro cell MEC server is determined according to the number of users who select macro cell offloading. If the number of macro cell offloading users increases, the allocated data rate to each user should decrease which causes the increase of the communication overhead. Then some users may change their decision from macro cell offloading to local computing or small cell offloading. So, users need to know the information about how many users select the macro cell offloading. But notifying each user's decision to other users has several problems such as privacy. For solving this problem, the macro cell BS broadcasts $\sum W_i \cdot I_{\{a_i=2\}}$ which is the aggregated value of the bandwidth demand from users who select macro cell offloading. At 4G/5G cellular network, the uplink and downlink resources are allocated according to their usages [28], [29]. Using this structure, users who want to offload to the macro cell MEC server should transmit W_n which is the bandwidth demand to the macro cell BS using the allocated uplink subframe. Then, the macro cell BS broadcasts $\sum W_i \cdot I_{\{a_i=2\}}$ to mobile device users using the allocated downlink resource. For this structure, we need to define the period of a game properly. Users can achieve the information about other users' decision without the privacy problem such as specifying a certain user's decision.

For ensuring the asynchronous improvement property, we need to guarantee that only one user can change his decision at each decision slot [19]. Users who have different decisions with previous decision slot send Change Decision messages to the macro cell BS. Then, the macro cell BS will select one user randomly and transmit the approval message which allows the selected user's decision change. The selected user will change his decision and the other users who are not selected keep their decision to next decision slot. When the macro cell BS doesn't get any Change Decision messages from uses through several slots, it will broadcast the CONVERGENCE message meaning that users' decisions are in state of a Nash equilibrium. Then, the game should be finished. Finding the proper game period considering the uplink/downlink resource allocation for information exchange (i.e., W_n, $\sum W_i \cdot I_{\{a_i=2\}}$, Change Decision, approval and CONVERGENCE) between mobile devices and the macro cell BS is our future work.

5.2 Convergence analysis

We already mentioned that potential games ensure the finite time convergence. For showing that this property works well, we find the maximum number of decision slots required for CONVERGENCE using the potential function. Define M as the maximum decision slot for the termination of a game. For finding M, we need to find two factors, the maximum potential value denoted as P_{max} which our potential function has, and the minimum amount of potential required denoted as P_{min} when each user changes his decision. Then, M can be calculated through dividing the maximum potential by the minimum potential requirement. At first, we calculate the maximum potential value. Denote $WT_{max}^{(s,m)} \triangleq \max_{n\in\mathcal{N}}\{W_n \cdot T_n^{(s,m)}\}$, $WT_{max}^{(l,m)} \triangleq \max_{n\in\mathcal{N}}\{W_n \cdot T_n^{(l,m)}\}$ and $W_{max} \triangleq \max_{n\in\mathcal{N}} W_n$. Then, we can state that P_{max} becomes

$$P_{max} = \begin{cases} N \cdot WT_{max}^{(s,m)}, & \text{if } \max\left\{WT_{max}^{(s,m)}, WT_{max}^{(l,m)}, \left(1+\frac{1}{2}N\right)W_{max}^2\right\} == WT_{max}^{(s,m)} \\ N \cdot WT_{max}^{(l,m)}, & \text{if } \max\left\{WT_{max}^{(s,m)}, WT_{max}^{(l,m)}, \left(1+\frac{1}{2}N\right)W_{max}^2\right\} == WT_{max}^{(l,m)} \\ \left(N+\frac{1}{2}N^2\right)W_{max}^2, & \text{if } \max\left\{WT_{max}^{(s,m)}, WT_{max}^{(l,m)}, \left(1+\frac{1}{2}N\right)W_{max}^2\right\} == \left(1+\frac{1}{2}N\right)W_{max}^2 \end{cases} \quad (51)$$

This maximum potential is not the optimal value, so finding the optimal maximum potential value is our future work. Next, we need to consider all 6 possible decision changes (i.e., local computing ⇄ small cell offloading) for finding the P_{min}. The minimum amount of potential required when each user changes his decision means that the minimum output difference at the potential function output when consider all 6 cases (i.e., $\phi(a_k, a_{-k}) - \phi(a'_k, a_{-k}) \geq P_{min}$). For calculating P_{min}, denote $W_{min} \triangleq \min_{n\in\mathcal{N}} W_n$, $j = \arg\min_{j\in\mathcal{N}} W_j$. Then we can state that

$$P_{min} = W_{min} \cdot \min\left\{\left|T_j^{(s,m)} - T_j^{(l,m)}\right|, \left|T_j^{(l,m)} - \sum_{i=1}^{N} W_i\right|, \left|T_j^{(s,m)} - \sum_{i=1}^{N} W_i\right|\right\} \quad (52)$$

Each element of the MIN function states the potential function output difference occurred by decision changes, local computing ⇄ small cell offloading, local computing ⇄ macro cell offloading, and small cell offloading ⇄ macro cell offloading. For ease of explanation, think about the case local computing ⇄ small cell offloading. At first, we need to remind that P_{min} must be positive because of the definition of ordinal potential games. Users are change their decisions when their overhead is decreased by decision changes. So, a decision change occurs when a user's overhead is decreased (i.e., $Z_k^l - Z_k^s > 0$). Ordinal potential games state that the sign of a user's overhead difference must be same with the sign of the potential function output difference. Therefore, P_{min} must be positive. Next, we have two kinds of decision change such as $a_k = 0 \to a'_k = 1$ & $a_k = 1 \to a'_k = 0$. At the equation (44) and (45), W_k, $T_k^{(l,m)}$ and $T_k^{(s,m)}$ are deterministic values which are calculated with given parameters. If $T_k^{(l,m)} - T_k^{(s,m)}$ is positive, then $T_k^{(s,m)} - T_k^{(l,m)}$ should be negative and vice versa. But the absolute value for these two cases are same. So, we can utilize the absolute value for finding the potential difference of local computing $\rightleftarrows$ small cell offloading like $\left|T_k^{(s,m)} - T_k^{(l,m)}\right|$. Other cases such as local computing $\rightleftarrows$ macro cell offloading, and small cell offloading $\rightleftarrows$ macro cell offloading can be formulated in the same vein. Finally, we can calculate the maximum decision slot M as

$$P_{min} \le \phi(\boldsymbol{a}) \le P_{max} \tag{53}$$

$$M = \lceil P_{max}/P_{min} \rceil \tag{54}$$

Because the P_{max} and P_{min} values are real number in practical, we need to use the ceiling function for finding the proper M. In the simulation & evaluation section, we will analyze M in detail.

VI. Simulation and Evaluation

In this section, we propose the numerical evaluations for the three-tier computation offloading game in a multi-user environment. We introduce simulation settings at first. And then we will show that our game converges to a Nash equilibrium in finite time.

For the number of mobile device users N, we assume two cases, $N = 4$ and $N = 8$. The reason why we choose these small number of users is related to the practical implementation of mmWave antenna on small cell BSs. Because the mmWave beam should be beamformed, we need to consider the phase of antennas and the size of a small cell BS. With the consideration about those factors, we decide the number of mobile device users for the practicality. According to previous researches, we need to consider the blockage issues for implementing the mmWave. We suppose that all mobile devices perform a Line-Of-Sight (LOS) mmWave communication with a small cell BS. We assume that the coverage of a small cell BS is 100m and the coverage of a macro cell BS is 1km. Then users should be deployed randomly over the coverages. Also, we assume that users have different amount of input data size and computation resource requirement. Based on the previous researches, we consider the face recognition as the users' offloading task [1]. For showing the property that different users have their own characteristics, we assume that the input data b_n is deployed randomly between 3000 ~ 7000 KB. Also, the required CPU cycles for the task c_n is determined as $c_n = 200 \times b_n$. The magnification factor (i.e., 200) is decided by the property of an application. For example, a simple task which requires simple computation should has a less magnification factor and a complicate task requires a higher magnification factor. For mobile devices, we assume the computing capability f_n^l as 1 GHz. For a small cell BS, the computing capability f^s is determined as 2.5 GHz. Finally, the macro cell computing capability f^m is 10 GHz. The weighting factor λ is determined by each user's comparison between the time and energy overhead. If a mobile device user considers the communication time is more important factor, then the λ should be set close to 1 and vice versa. With these simulation settings, we first show the users' decision change with 3D graphs. Next, we will provide the per-user overhead graphs. Through the graphs, we can observe that the three-tier computation offloading decision game converges in finite time although a user's decision change affects to other users' overheads and decisions.

We will analyze M which is the maximum number of decision slot required for convergence with simulation results. Finally, we will propose the system-wide overhead graph which shows that the minimizing each user's overhead leads the decrease and convergence of the system overhead.

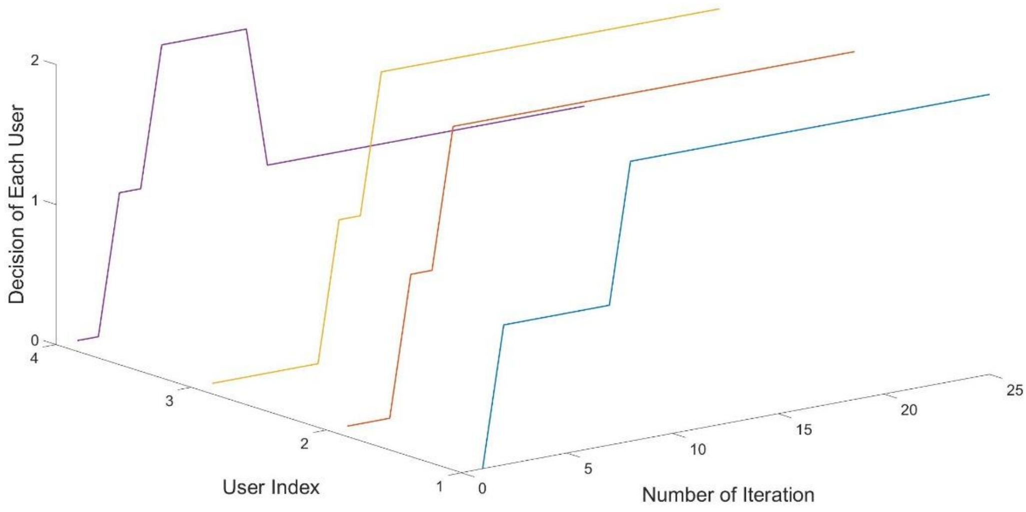

Fig 2. Each user's decision change converged at 11th slot ($N = 4$)

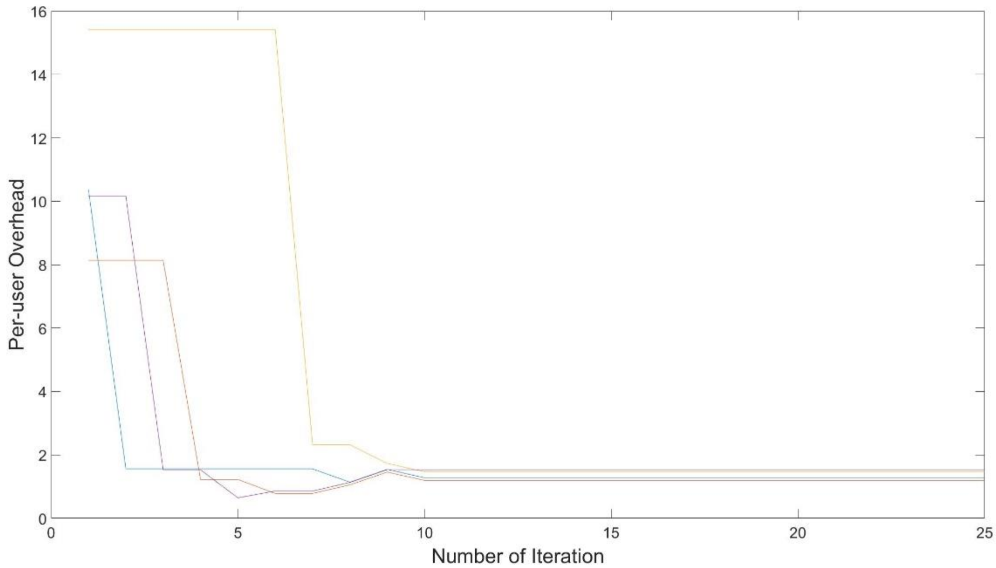

Fig 3. Per-user overhead converged at 11th slot ($N = 4$)

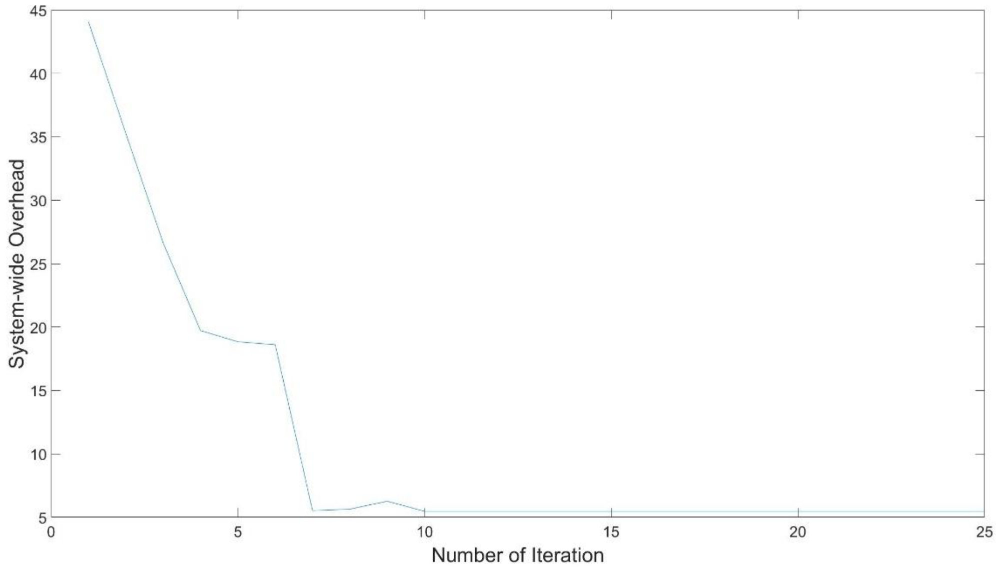

Fig 4. System-wide overhead converged at 11th slot ($N = 4$)

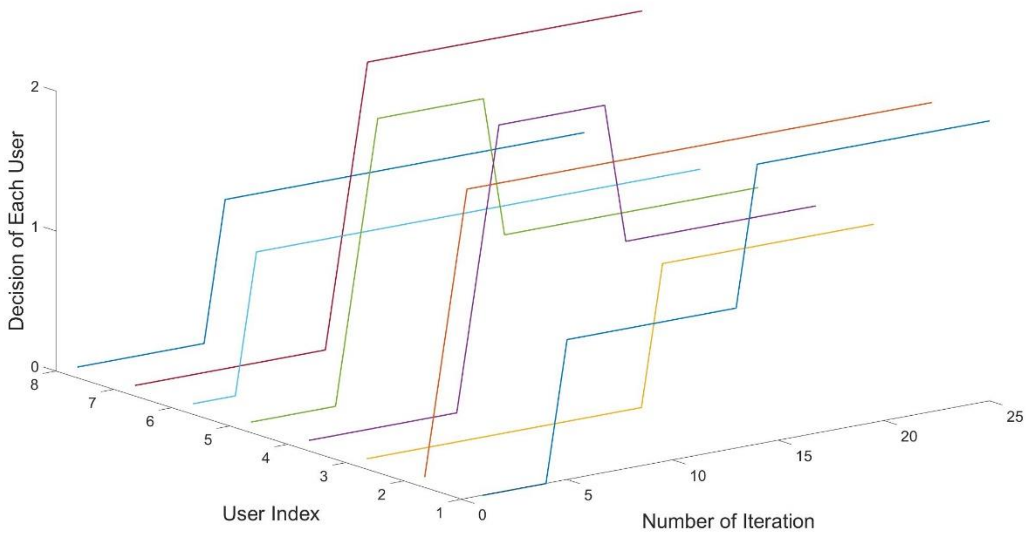

Fig 5. Each user's decision change converged at 17th slot ($N = 8$)

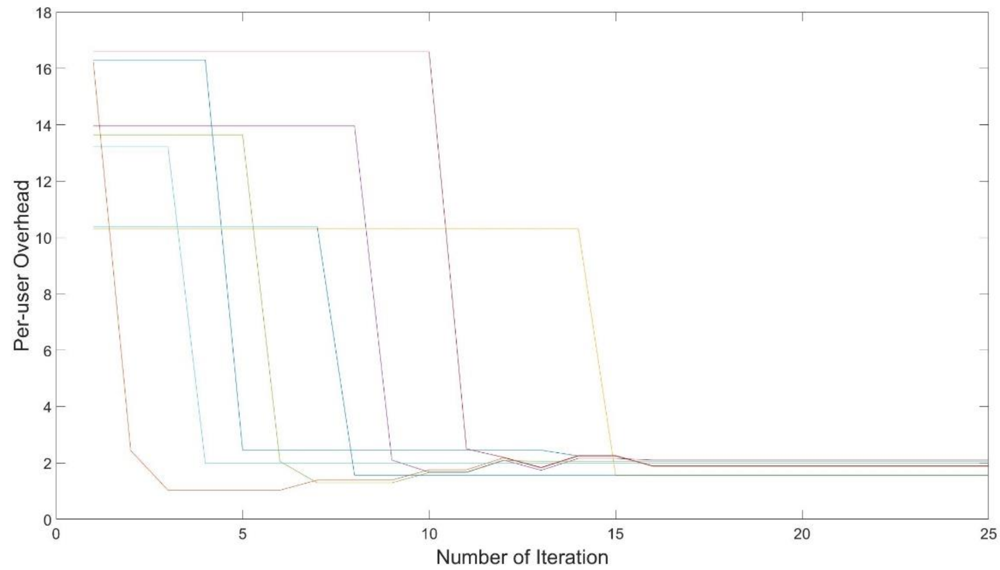

Fig 6. Per-user overhead converged at 17th slot ($N = 8$)

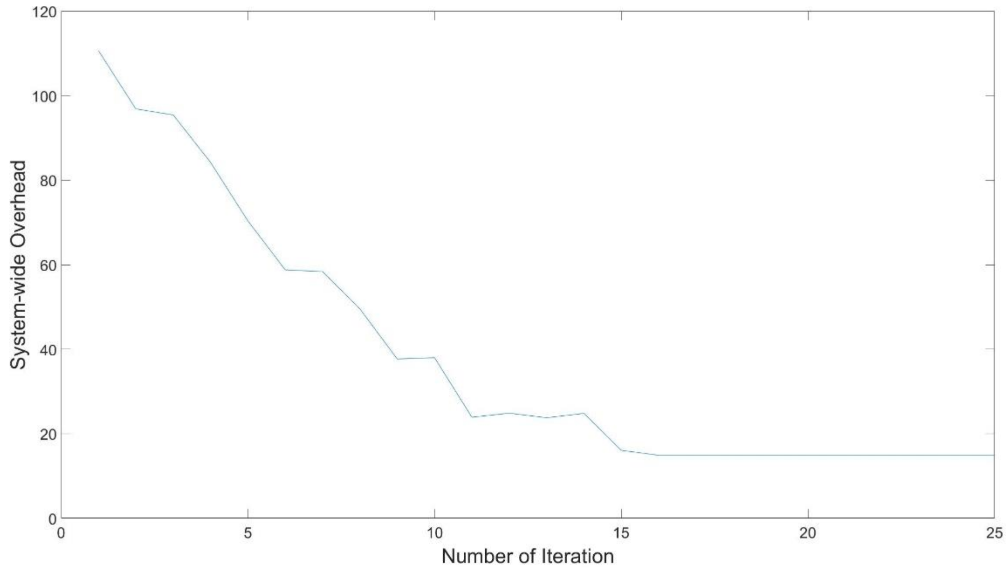

Fig 7. System-wide overhead converged at 17th slot ($N = 8$)

Through Fig 2. ~ Fig 7., we can observe that our three-tier computation offloading decision games in multi-user environment considering 4 and 8 users converge in finite decision slot. Through observing Fig 2. & 3. and Fig 5. & 6., we can see how interactions between users affect to decision changes. Some users change their decisions from macro cell offloading to small cell offloading because the allocated bandwidths to those users are decreased by the increase of macro cell offloading users. At Fig 4. and Fig 7., the system-wide overhead is decreased and converged in finite decision slots.

Finally, we evaluate M which is defined at the convergence analysis section. According to the simulation results, $M_{N=4}$ is 346 and $M_{N=8}$ is 132. But the 4-user game is converged at 11th slot and the 8-user game is converged at 17th slot. Through the numerical evaluations, we can find that why theses gaps are occurred. When we calculate M values, the maximum amount of potential is related to the maximum value among $WT_{max}^{(s,m)}$, $WT_{max}^{(l,m)}$ and $\left(1+\frac{1}{2}N\right)W_{max}^{2}$. By the simulations, the $WT_{max}^{(s,m)}$ value is measured as the maximum value. It means that when all users select small cell offloading, then the potential function has the maximum amount of potential. But we can observe that lots of users select macro cell offloading for minimizing their overhead rather than small cell offloading. The same circumstance can be observed at the P_{min}.For these reasons, the game is converged faster than calculated M slot.

For proving that our game fits well for practical situations, think about the urgent situation such as a car accident. Because the task processed at each decision slot is simple overhead calculations, we can assume the decision slot size as a few milliseconds (i.e., 1 ms). According to the Korea Road Safety Act, the safety distance for the 50km/h vehicle is 35m. Then, we can calculate that the car needs 2.69s for the break. Through using the mentioned M values ($M_{N=4} = 346$ & $M_{N=8} = 132$), we can observe that a game requires hundreds of milliseconds for the convergence. So, we can state that our game should terminate fast enough, not take a second. In conclusion, we state that the three-tier computation offloading decision game in a multi-user environment converges fast enough and should work well in practice situations.

VII. Conclusion

Different with previous researches which consider only one between a small cell and a macro cell as the location of a MEC server used for computation offloading, we propose the three-tier computation offloading in a multi-user environment with ordinal potential games approach. For the practicality, we consider different communication schemes of mmWave for the communication with the small cell BS and 4G/5G cellular network for the communication with the macro cell BS. Using the properties of ordinal potential games, we state that our game always has Nash equilibriums and converges in finite time. For ensuring that our game is an ordinal potential game, we find the potential function which reflects the dynamics of the three-tier computation offloading game.

Then, we perform the three-tier computation offloading decision game algorithm for showing the asynchronous improvement property. For analyzing the convergence of our game, we calculate the maximum amount of potential that our potential function has, and the minimum amount of potential required for a user's decision change. Using these factors, we calculate the maximum number of decision slot required for the convergence. Through numerical simulations, we show that our game converges to a Nash equilibrium in finite time.

For the future work, we will analyze the Price of Anarchy (PoA) of our game [30]. The PoA is important concept for game theory that it evaluates the efficiency of Nash equilibriums calculated by a game. By the definition, the PoA measures the efficiency of system through quantifying the ratio between a centralized optimal solution (i.e., minimizing the system-wide overhead) and the worst-case decentralized optimal solution. And then, we will treat the mobility of game participants. In this thesis, we assume a quasi-static scenario, so we don't care about the mobility of users. When the mobility is considered, we need to care about LOS blockage model, handover problem between base stations and so on [31]. For the LOS blockage model, we need to model the state change between LOS and NLOS by the velocity properly. Also, the positions of mobile devices change consequently, so we need to consider the overhead occurred by the handover between bases stations.

REFERENCES

[1] T. Soyata, R. Muraleedharan, C. Funai, M. Kwon and W. Heinzelman, "Cloud-Vision: Real-time face recognition using a mobile-cloudlet-cloud acceleration architecture," *2012 IEEE Symposium on Computers and Communications (ISCC)*, Cappadocia, 2012, pp. 000059-000066.

[2] L. Yang, J. Cao, H. Cheng and Y. Ji, "Multi-User Computation Partitioning for Latency Sensitive Mobile Cloud Applications," in *IEEE Transactions on Computers*, vol. 64, no. 8, pp. 2253-2266, 1 Aug. 2015.

[3] E. Meskar, T. D. Todd, D. Zhao and G. Karakostas, "Energy Aware Offloading for Competing Users on a Shared Communication Channel," in *IEEE Transactions on Mobile Computing*, vol. 16, no. 1, pp. 87-96, 1 Jan. 2017.

[4] P. Mach and Z. Becvar, "Mobile Edge Computing: A Survey on Architecture and Computation Offloading," in *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1628-1656, thirdquarter 2017.

[5] Niu, Y., Li, Y., Jin, D. et al., "A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges" in *Wireless Netw* (2015) 21: 2657.

[6] C. You, K. Huang, H. Chae and B. Kim, "Energy-Efficient Resource Allocation for Mobile-Edge Computation Offloading," in *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1397-1411, March 2017.

[7] X. Chen, "Decentralized Computation Offloading Game for Mobile Cloud Computing," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 4, pp. 974-983, 1 April 2015.

[8] Meng, S., Wang, Y., Miao, Z. et al., "Joint optimization of wireless bandwidth and computing resource in cloudlet-based mobile cloud computing environment" in *Peer-to-Peer Netw. Appl.* (2018) 11: 462.

[9] S. Ahn, J. Lee, S. Park, S. H. S. Newaz and J. K. Choi, "Competitive Partial Computation Offloading for Maximizing Energy Efficiency in Mobile Cloud Computing," in *IEEE Access*, vol. 6, pp. 899-912, 2018.

[10] 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.

[11] L. Tang and X. Chen, "An Efficient Social-Aware Computation Offloading Algorithm in Cloudlet System," *2016 IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, 2016, pp. 1-6.

[12] H. Cao and J. Cai, "Distributed Multiuser Computation Offloading for Cloudlet-Based Mobile Cloud Computing: A Game-Theoretic Machine Learning Approach," in *IEEE Transactions on*

Vehicular Technology, vol. 67, no. 1, pp. 752-764, Jan. 2018.

[13] J. Zheng, Y. Cai, Y. Wu and X. S. Shen, "Stochastic computation offloading game for mobile cloud computing," *2016 IEEE/CIC International Conference on Communications in China (ICCC)*, Chengdu, 2016, pp. 1-6.

[14] M. Messous, H. Sedjelmaci, N. Houari and S. Senouci, "Computation offloading game for an UAV network in mobile edge computing," *2017 IEEE International Conference on Communications (ICC)*, Paris, 2017, pp. 1-6.

[15] M. Deng, H. Tian and X. Lyu, "Adaptive sequential offloading game for multi-cell Mobile Edge Computing," *2016 23rd International Conference on Telecommunications (ICT)*, Thessaloniki, 2016, pp. 1-5.

[16] Cardellini, V., De Nitto Personé, V., Di Valerio, V. et al., "A game-theoretic approach to computation offloading in mobile cloud computing" in *Math. Program.* (2016) 157: 421.

[17] S. Kim. "Nested game-based computation offloading scheme for mobile cloud IoT systems. " EURASIP Journal on Wireless Communications and Networking, 2015(1):1–11, 2015.

[18] D. Monderer and L. S. Shapley, Potential games, Games and Economic Behavior, 14 (1996), pp. 124–143.

[19] Quang Duy L, Yong Huat Chew, and Boon-Hee Soong. 2016. Potential Game Theory: Applications in Radio Resource Allocation (1st ed.). Springer Publishing Company, Incorporated.

[20] T. E. Bogale and L. B. Le, "Massive MIMO and mmWave for 5G Wireless HetNet: Potential Benefits and Challenges," in *IEEE Vehicular Technology Magazine*, vol. 11, no. 1, pp. 64-75, March 2016.

[21] K. Zhang, Y. Mao, S. Leng, S. Maharjan and Y. Zhang, "Optimal delay constrained offloading for vehicular edge computing networks," *2017 IEEE International Conference on Communications (ICC)*, Paris, 2017, pp. 1-6.

[22] T. S. Rappaport *et al*., "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," in *IEEE Access*, vol. 1, pp. 335-349, 2013.

[23] C. Wang, C. Liang, F. R. Yu, Q. Chen and L. Tang, "Computation Offloading and Resource Allocation in Wireless Cellular Networks With Mobile Edge Computing," in *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 4924-4938, Aug. 2017.

[24] A. Tassi, M. Egan, R. J. Piechocki and A. Nix, "Modeling and Design of Millimeter-Wave Networks for Highway Vehicular Communication," in *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 10676-10691, Dec. 2017.

[25] T. S. Rappaport, R. W. Heath, Jr., R. Daniels, and J. Murdock, Millimeter Wave Wireless Communications. Upper Saddle River, NJ, USA: Prentice-Hall, 2015.

[26] N. Abu-Ali, A. M. Taha, M. Salah and H. Hassanein, "Uplink Scheduling in LTE and LTE-Advanced: Tutorial, Survey and Evaluation Framework," in *IEEE Communications Surveys & Tutorials*,

vol. 16, no. 3, pp. 1239-1265, Third Quarter 2014.

[27] Y. Wen, W. Zhang and H. Luo, "Energy-optimal mobile application execution: Taming resource-poor mobile devices with cloud clones," *2012 Proceedings IEEE INFOCOM*, Orlando, FL, 2012, pp. 2716-2720.

[28] J. Lim, H. G. Myung, K. Oh and D. J. Goodman, "Channel-Dependent Scheduling of Uplink Single Carrier FDMA Systems," *IEEE Vehicular Technology Conference*, Montreal, Que., 2006, pp. 1-5.

[29] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia and P. Camarda, "Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and a Survey," in *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678-700, Second Quarter 2013.

[30] Perakis, G., G. Roels. 2007. The price of anarchy in supply chains: Quantifying the efficiency of price-only contracts. Management Sci. 53(8) 1249–1268.

[31] V. Va, T. Shimizu, G. Bansal and R. W. Heath, "Beam design for beam switching based millimeter wave vehicle-to-infrastructure communications," *2016 IEEE International Conference on Communications (ICC)*, Kuala Lumpur, 2016, pp. 1-6.

[32] M. Voorneveld, “Potential Games and Interactive Decisions with Multiple Criteria,” PhD Dissertation, Tilburg University, Netherlands, 1996.