Hong Quy Le, Hussein Al-Shatri, Anja Klein, ”Efficient Resource Allocation in Mobile-edge Computation Offloading: Completion Time Minimization,” in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Aachen, Germany, 25-30 June 2017.

Efficient Resource Allocation in Mobile-edge Computation Offloading: Completion Time Minimization

Hong Quy Le, Hussein Al-Shatri, Anja Klein
Communications Engineering Lab
Technische Universität Darmstadt, Merckstrasse 25, 64283 Darmstadt, Germany
{h.le, h.shatri, a.klein}@nt.tu-darmstadt.de

***Abstract*—Mobile-edge computation offloading (MECO) is a promising solution for enhancing the capabilities of mobile devices. For an optimal usage of the offloading, a joint consideration of radio resources and computation resources is important, especially in multiuser scenarios where the resources must be shared between multiple users. We consider a multi-user MECO system with a base station equipped with a single cloudlet server. Each user can offload its entire task or part of its task. We consider parallel sharing of the cloudlet, where each user is allocated a certain fraction of the total computation power. The objective is to minimize the completion time of users' tasks. Two different access schemes for the radio channel are considered: Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA). For each access scheme, we formulate the corresponding joint optimization problem and propose efficient algorithms to solve it. Both algorithms use the bisection-search method, where each step requires solving a feasibility problem. For TDMA, the feasibility problem has a closed-form solution. Numerical results show that the performance of offloading is higher than of local computing. In particular, MECO with FDMA outperforms MECO with TDMA, but with a small margin.**

I. Introduction

Today's mobile devices are equipped with advanced technologies, for example, high resolution cameras and integrated sensors. With the improving capabilities of the devices and the increasing interest in mobile applications and services for daily purposes, the functionality of mobile devices has made the applications which require data collection and data processing possible, for example, augmented reality, speech-to-text, image processing [1]. One of the key challenges of these applications is the high computation power requirement. However, with current technology, the computation capabilities of mobile devices are still limited. To overcome this limitation, mobile edge computing [1][2][3] is proposed as a promising solution. In the solution, small-scale computing clouds - also known as *cloudlets* - are deployed at the edges of wireless networks, e.g., at wireless access points [1][2][3]. The mobile devices can offload intensive computation tasks to the nearby cloudlets - also refered to as *mobile edge computation offloading* (MECO) [2], [4], [5]. In comparison with other cloud computing solutions, one of the key advantages of cloudlets is low latency [1] due to the short distance from the mobile devices to the cloudlet.

Two critical limiting factors of MECO are radio resources of the wireless links and computation resources at the cloudlet. Both factors are playing important roles. For example, for shortening a job's completion time, both radio transmission time and cloudlet processing time must be reduced. In a multiuser scenario with multiple users using the same cloudlet to offload their tasks, the resources must be shared between multiple users. Therefore, efficient resource allocation algorithms are critical [5].

Many researchers have worked on resource allocation for MECO. A common scenario is multi-user MECO with the objective of minimizing the total energy consumption of all user nodes [4][6][7][8][9]. However, often the researchers focused on one type of resources, either radio resources or computation resources. For example, [6] and [7] focused on designing radio resources allocation algorithms with the objective of minimizing weighted sum energy consumption for predefined execution delay deadline. The execution delay at the cloudlet was modeled as a constant and can be subtracted from the deadline constraint. Under that assumption, the offloading decision problem can be modeled as a pure radio resource allocation problem. Both works applied game theory to solve the radio resource allocation problem. Some other works focused on designing computation resources allocation and power control algorithms where they assumed that the radio resource had been pre-allocated for each user [8][9]. Both [8] and [9] assumed the same channel bandwidth for each user and proposed joint power control and computation resource allocation algorithms. [8] focused on the energy minimization problem and [9] focused on minimizing the weighted sum of energy and delay. In [4], the authors considered both channel time allocation and computation resource allocation. However, they assume that the computation resource allocated to each user is propotional to the offloaded tasks' size of that user.

In this work, we consider a MECO system with a base station and a single cloudlet server. Each user has one task. Each task can be split into two parts, one for local computation and one for offloading. We consider the problem of minimizing the tasks' completion delay including the time for data transmission and the time for computing. We aim at developing joint algorithms for the allocation of radio resources (including power control) and computation resources. As in [8], we

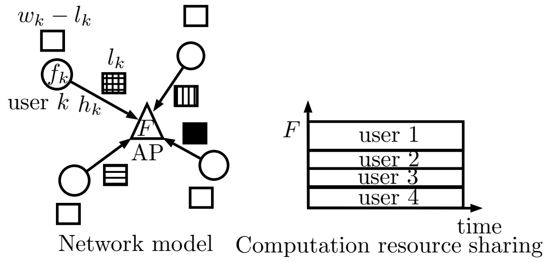

Fig. 1. Network model and computation resource sharing model

assume that cloudlet resources can be allocated as percentages of the total computation power to each user. This enables parallel processing of jobs from different users. For radio channel resource allocation, we consider two different multiple access schemes: Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA). We formulate the joint resources allocation as an optimization problem and propose efficient algorithms to solve it.

The rest of this paper is organized as follows: Section II presents the system model where task model, transmission, computation delay, as well as multiple access schemes are introduced. The problem of optimal offloading with TDMA is considered in section III and FDMA is considered in section IV. The performance of the proposed algorithms is then investigated via numerical simulation in Section V. Finally, we conclude our work in Section VI.

II. System Model

We consider a multiuser system consisting of K single-antenna mobile users (MUs) and a single-antenna wireless access point (AP), as shown in Fig. 1. A cloudlet with finite computation capability is deployed at the AP to provide computing services. We consider a snapshot when the CPU of the cloudlet is available. Let $\mathcal{C} = \{1, 2, \ldots, K\}$ denote the K users, each with a task to execute. The AP schedules a subset of users for complete/partial offloading. The users with partial or no offloading compute a fraction of or all input data, respectively, using their local CPU. The users with partial or complete offloading offload a fraction of or all input data, respectively, to the cloudlet. We consider a frequency flat channel model. For multiple access, we consider two different schemes: TDMA and FDMA. The AP is assumed to have perfect knowledge of all the channel gains, local computation capability of the user nodes, and the sizes of the input data at all users. In addition, the channel gains are assumed to remain constant within the considered snapshot duration. Using this information, the AP selects and allocates the resources to the users: the transmit power of the nodes, and the fraction of cloudlet computation power for each node will be determined together with the fraction of channel time for the TDMA case and the fraction of channel bandwidth for the FDMA case for each user.

A. Data rate with multiple access model

Let B denote the total channel bandwidth of the system and $N_0/2$ denote the power spectral density of the complex white Gaussian channel noise. Let h_k denote the channel gain of user k to the AP and p_k denote the transmission power for mobile k. We assume that the user uses only one transmit power level in each snapshot.

1) TDMA: Each user will be assigned a fraction of time to use the channel. Let $x_k \geq 0$ denote the fraction of time allocated to user k. Then the data rate of user k is

$$r_k^{\text{TDMA}} = x_k R_k, \quad \text{and} \quad \sum_{k=1}^{K} x_k = 1, \tag{1}$$

where

$$R_k = B \log_2 \left(1 + \frac{p_k |h_k|^2}{B N_0} \right) \tag{2}$$

is the Shannon channel capacity of user k.

2) FDMA: Each user will be allocated a fraction of the system bandwidth. Let $z_k \in [0, 1]$ denote the fraction of bandwidth allocated to user k. Then the data rate of the user k is

$$r_k^{\text{FDMA}} = z_k B \log_2 \left(1 + \frac{p_k |h_k|^2}{z_k B N_0} \right) \quad \text{and} \quad \sum_{k=1}^{K} z_k = 1. \tag{3}$$

B. Task model and execution time model

We follow the splittable task model used in [4]. Each task $\mathcal{T}_k$ is described by its input data size w_k in bits, and a known constant β_k in CPU cycles per bit, which describes the number of CPU cycles required to process one bit of input data. Eask task $\mathcal{T}_k$ can be divided into two jobs with l_k and $w_k - l_k$ bits of input data, respectively, see Fig. 1. The first job with l_k bits will be offloaded to the cloudlet. It will be called offloaded job. The second job with $w_k - l_k$ bits will be computed locally by the local CPU. It will be called local job. How the tasks should be split is one of the subjects in our joint algorithms and will be presented later.

1) Execution time of local job: For user k, the frequency of the local CPU is f_k. The size of the local job is $w_k - l_k$ bits. The processing time of the local job of user k is

$$T_k^{\text{local}} = \frac{\beta_k (w_k - l_k)}{f_k}. \tag{4}$$

2) Execution time of offloaded job: The execution time of the offloaded job consists of the data transmission time and the job processing time at the cloudlet. User k offloads l_k bits of data to the cloudlet. Let F in CPU cycles per second denote the computation capability of the cloudlet. The total computation power is split among the users, each with a fraction of the total capability, see Fig. 1. Let $y_k \in [0, 1]$ denote the fraction of computation power allocated for the offloaded job of user k. With the data rate r_k, the total execution time of the offloaded job of user k is

$$T_k^{\text{offload}} = \frac{l_k}{r_k} + \frac{\beta_k l_k}{y_k F}, \tag{5}$$

where the first term is the data transmission time and the second term is the job processing time. We do not consider the time spent for sending back the result from the AP to the users. This amount of time is often very short compared with the total data offloading time and task execution time [4]. Therefore, the completion time of user k is defined as the time when both the local job and the offloaded job are completed, i.e.

$$T_k^{\text{compl}} = \max\{T_k^{\text{local}}, T_k^{\text{offload}}\}. \tag{6}$$

III. Minimizing completion time with TDMA

In this section, resource allocation for multiuser MECO is formulated as an optimization problem for the TDMA case. The objective is to minimize the completion time of all the users, i.e., minizing

$$T = \max_{1\le k\le K} T_k^{\text{compl}}.$$

Under the constraints on the total channel access time and total CPU time, the resource allocation problem can be formulated as

$$\begin{aligned}
\min_{T,\{l_k\},\{x_k\}} \quad & T && \text{(P-1)}\\
\text{subject to} \quad & \frac{\beta_k(w_k - l_k)}{f_k} \le T,\ \forall k && (7)\\
& \frac{l_k}{x_k R_k} + \frac{\beta_k l_k}{y_k F} \le T,\ \forall k && (8)\\
& \sum_{k=1}^{K} x_k = 1, \sum_{k=1}^{K} y_k = 1 && (9)\\
& 0 \le l_k \le w_k, 0 \le x_k, 0 \le y_k. && (10)
\end{aligned}$$

The objective T represents the completion time. The constraints in (7) mean that the execution time of the local jobs should not exceed the completion time. The constraints in (8) mean that the execution time of the offloaded jobs (including data transmission time) should not exceed the completion time. The constraints in (9) are the sum constraints of channel time and total computation power. It is worth to mention that constraints (8) consider the TDMA using the variables $x_k, \forall k$ with $\sum_{x=1}^{K} = 1$ which determine the fractions of transmission time of each user.

Our approach to solve the problem (P-1) is to use bisection search on T. For each fixed T, we must solve a feasibility problem for constraints (7) - (10). Due to constraint (7), we have

$$l_k \ge l_k^{\min} := \max\left\{0, w_k - \frac{T f_k}{\beta_k}\right\}. \tag{11}$$

Because the equality in (11) holds only for the offloaded job with minimum number $l_k^{\min}$ of bits, it is sufficient to ensure the feasibility of the smallest of the offloaded jobs, i.e., $l_k = l_k^{\min}$. We must solve the following feasibility problem:

$$\begin{aligned}
\min_{\{x_k\},\{y_k\}} \quad & 0 && \text{(P-1A)}\\
\text{subject to} \quad & \frac{l_k^{\min}}{x_k R_k} + \frac{\beta_k l_k^{\min}}{y_k F} \le T && (12)\\
& \sum_{k=1}^{K} x_k = 1,\ \sum_{k=1}^{K} y_k = 1 && (13)\\
& 0 \le x_k, 0 \le y_k. && (14)
\end{aligned}$$

With $a_k = \frac{l_k^{\min}}{R_k}$ and $b_k = \frac{\beta_k l_k^{\min}}{F}$, we have the following lemma:

Lemma 1. *The necessary and sufficient conditions for the feasibility problem* (P-1A) *are*

$$\sum_{k=1}^{K} a_k \le T, \quad \sum_{k=1}^{K} b_k \le T \tag{15}$$

$$\left(\sum_{k=1}^{K} \sqrt{a_k b_k}\right)^2 \le \left(T - \sum_{k=1}^{K} a_k\right)\left(T - \sum_{k=1}^{K} b_k\right). \tag{16}$$

Proof: See Appendix. ■

The algorithm for achieving minimum completion time with TDMA is given in Algorithm 1:

Algorithm 1 Min completion time with TDMA

1) Initialize: $T_{\text{low}} = 0$, $T_{\text{high}} = \max_{1\le k\le K} \frac{\beta_k w_k}{f_k}$, set ϵ.
2) If $T_{\text{hight}} - T_{\text{low}} < \epsilon$, terminate the algorithm.
3) Set $T = \frac{T_{\text{high}} + T_{\text{low}}}{2}$. Calculate $l_k^{\min}$, a_k, b_k. Check the feasibility conditions (15) and (16). If feasible, then set $T_{\text{high}} = T$, else set $T_{\text{low}} = T$. Go to step 2.

IV. Minimizing Completion Time With FDMA

We have the following optimization problem for the FDMA:

$$\begin{aligned}
\min_{T,\{l_k\},\{x_k\}} \quad & T && \text{(P-2)}\\
\text{subject to} \quad & \frac{\beta_k(w_k - l_k)}{f_k} \le T && (17)\\
& \frac{l_k}{z_k B \log_2\left(1 + \frac{p_k|h_k|^2}{z_k B N_0}\right)} + \frac{\beta_k l_k}{y_k F} \le T && (18)\\
& \sum_{k=1}^{K} z_k = 1, \sum_{k=1}^{K} y_k = 1 && (19)\\
& 0 \le l_k \le w_k, 0 \le z_k, 0 \le y_k && (20)
\end{aligned}$$

This problem is almost the same as the problem for the TDMA case. The only difference are the constraints (18), where the time for data offloading is calculated based on the rate achieved with FDMA.

Similar to the TDMA case, we use bisection search method. For a fixed T, we need to solve the feasibility problem for constraints (17)-(20). With the same $l_k^{\min}$ as in (11), we must only check the feasibility when $l_k = l_k^{\min}$. The feasibility problem becomes

$$\min_{\{z_k\},\{y_k\}} \quad 0 \tag{P-2A}$$

$$\text{subject to} \quad \frac{l_k^{\min}}{z_k B \log_2\left(1+\frac{p_k|h_k|^2}{z_k B N_0}\right)} + \frac{\beta_k l_k^{\min}}{y_k F} \le T \tag{21}$$

$$\sum_{k=1}^{K} z_k = 1, \ \sum_{k=1}^{K} y_k = 1 \tag{22}$$

$$0 \le z_k, 0 \le y_k. \tag{23}$$

Let $c_k = \frac{l_k^{\min}}{B}$, $d_k = \frac{p_k|h_k|^2}{BN_0}$, and $e_k = \frac{\beta_k l_k^{\min}}{F}$. From constraints (21), we have

$$y_k \ge \frac{e_k}{T - \frac{c_k}{z_k \log_2\left(1+\frac{d_k}{z_k}\right)}}. \tag{24}$$

Combined with the constraint (22), the constraints on y_k can be formulated as constraints for z_k:

$$z_k \log_2\left(1+\frac{d_k}{z_k}\right) \ge \frac{c_k}{T} \Leftrightarrow z_k \ge z_k^{\min} \tag{25}$$

$$\sum_{k=1}^{K} \frac{e_k}{T - \frac{c_k}{z_k \log_2\left(1+\frac{d_k}{z_k}\right)}} \le 1, \tag{26}$$

where $z_k^{\min}$ satisfies $z_k^{\min} \log_2\left(1+\frac{d_k}{z_k^{\min}}\right) = \frac{c_k}{T}$. Thus, in order to check the feasibility problem (P-2A), we have to solve the following optimization problem:

$$\min_{\{z_k\}} \quad \sum_{k=1}^{K} \frac{e_k}{T - \frac{c_k}{z_k \log_2\left(1+\frac{d_k}{z_k}\right)}} \tag{P-2B}$$

$$\text{subject to} \quad \sum_{k=1}^{K} z_k = 1, \ z_k \ge z_k^{\min}. \tag{27}$$

If the minimum value is smaller than or equal to 1, then (P-2A) is feasible.

Lemma 2. *The problem* (P-2B) *is a convex optimization problem.*

Proof: Because the function $\frac{e}{T-\frac{c}{x}}$ is decreasing for $x > \frac{c}{T}$, and the function $z\log_2\left(1+\frac{d}{z}\right)$ is a concave function, the function $u(z) = \frac{e}{T-\frac{c}{z\log_2\left(1+\frac{d}{z}\right)}}$ is a convex function for z such that $z\log_2\left(1+\frac{d}{z}\right) \ge \frac{c}{T}$. As the result, the optimization problem is a convex optimization problem. ■

Algorithm 2 Min completion time with FDMA

1) Initialize: $T_{\text{low}} = 0$, $T_{\text{high}} = \max_{1\le k\le K} \frac{\beta_k w_k}{f_k}$, set ϵ.
2) If $T_{\text{hight}} - T_{\text{low}} < \epsilon$, terminate the algorithm.
3) Set $T = \frac{T_{\text{high}}+T_{\text{low}}}{2}$. Calculate $l_k^{\min}$, c_k, d_k, e_k, $z_k^{\min}$. Solve the problem (P-2B). If it is infeasible or the min value is greater than 1, then set $T_{\text{low}} = T$, otherwise set $T_{\text{high}} = T$. Go to step 2.

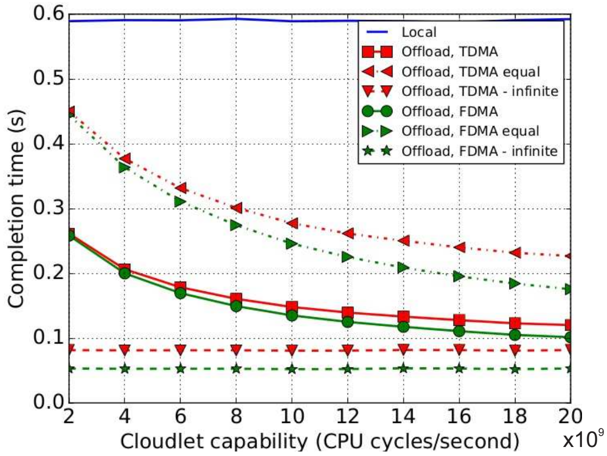

Fig. 2. Minimum completion time for varying cloudlet capability

The problem (P-2B) is convex and has a single linear constraint, thus, it can be solved efficiently, e.g. with bisection-search method.

The algorithm for achieving the minimum completion time with FDMA is given in Algorithm 2.

V. Numerical Results

We use the following paramters for our simulations. The system consists of 10 mobile users. We take the following parameters from [4]. The channel gains are modeled as independent Rayleigh fading with average power loss set to 10^{-6}. In addition, the power spectral density of the complex white Gaussian noise is $N_0/2 = 10^{-13}$W/Hz and the channel bandwidth $B = 10$ MHz. For each user k, the speed of the local CPU f_k is randomly selected from the set $\{0.5, 0.6, 07, 0.8, 0.9, 1.0\} \times 10^9$ CPU cycles/second. For the computation tasks, the data size follows a uniform distribution with $w_k \in [100, 300]$kbits and the number of CPU cycles per bit is $\beta_k \in [500, 1500]$. The random variables are generated independently for different users. The computation capability F of the the cloudlet varies between 2×10^9 and 20×10^9 CPU cycles per second.

We compare the performance of the optimized offloading schemes with TDMA and FDMA with the following reference schemes:

- local computation scheme,
- equal resources allocation (computation and radio resources) with TDMA and FDMA, respectively, and
- the extreme case when the cloudlet has infinite computation capabilities.

Fig. 2 shows the curves of the completion time for varying cloudlet capability. The performance gain increases with increasing computation capability of the cloudlet. With offloading, the completion time is reduced greatly compared to local computing. The optimized TDMA and FDMA schemes outperform the corresponding equal resource allocation schemes. Moreover, there is a certain limit on the performance gain. We also show the situation when the capability of the cloudlet is infinite. In this case, the performance of the offloading schemes depends only on the computation capability of the local CPUs and the data transmission. Another observation is that the performance of FDMA is better than that of TDMA. This is due the different effective noise bandwidth in FDMA and TDMA, see (2) and (3). Moreover, the performance gap increases with increasing computation capability of the cloudlet.

VI. Conclusion

We focused on the problem of designing optimal algorithms for solving the joint radio resources and computation resources problem in a multiuser MECO system for minimizing completion time. We formulated optimization problems that can be solved efficiently. High performance gain can be obtained using offloading. In particular, the performance of MECO with FDMA is higher than of MECO with TDMA, but with a small margin.

Acknowledgement

This work has been performed in the project B3 within the Collaborative Research Center (CRC) 1053-MAKI.

Appendix

A. Proof of lemma 1

From the constraint (12), we have $\frac{a_k}{x_k} + \frac{b_k}{y_k} \leq T$. Thus,

$$x_k \geq \frac{a_k}{T} \quad \text{and} \quad y_k \geq \frac{b_k x_k}{T x_k - a_k}. \tag{28}$$

Combining this with the constraint (13), we obtain

$$\sum_{k=1}^{K} \frac{b_k x_k}{T x_k - a_k} \leq 1. \tag{29}$$

From (28) and (13), we must have $\sum_{k=1}^{K} \frac{a_k}{T} \leq 1$, thus one condition for feasibility is

$$\sum_{k=1}^{K} a_k \leq T. \tag{30}$$

In addition, (29) is satisfied if and only if the minimum value of $\sum_{k=1}^{K} \frac{b_k x_k}{T x_k - a_k}$ with constraints $x_k \geq \frac{a_k}{T}$ and $\sum_{k=1}^{K} x_k = 1$ does not exceed 1. First, we solve the problem

$$\min_{\{x_k\}} \quad \sum_{k=1}^{K} \frac{b_k x_k}{T x_k - a_k} \tag{31}$$

$$\text{subject to} \quad \sum_{k=1}^{K} x_k = 1, \; x_k \geq \frac{a_k}{T}. \tag{32}$$

It is easy to check that the function $\frac{bx}{Tx-a}$ is a convex function for $x > \frac{a}{T}$ with $a \geq 0$, $T > 0$, $b \geq 0$. Therefore, the above optimization problem is a convex problem with one linear constraint. The corresponding Lagrangian function is

$$\mathcal{L}(\{x_k\}, \lambda) = \sum_{k=1}^{K} \frac{b_k x_k}{T x_k - a_k} - \lambda \left(\sum_{k=1}^{K} x_k - 1 \right). \tag{33}$$

We have

$$\frac{\partial \mathcal{L}}{\partial x_k} = -\frac{a_k b_k}{(T x_k - a_k)^2} + \lambda \tag{34}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \sum_{k=1}^{K} x_k - 1. \tag{35}$$

Thus,

$$x_k^* = \frac{a_k + \sqrt{\frac{a_k b_k}{\lambda}}}{T} \quad \text{and} \quad \sqrt{\lambda} = \frac{\sum\limits_{k=1}^{K} \sqrt{a_k b_k}}{T - \sum\limits_{k=1}^{K} a_k} \tag{36}$$

Substituting x_k^* in the objective function, we obtain the following minimum:

$$\text{min value } = \frac{1}{T} \left[\frac{\left(\sum\limits_{k=1}^{K} \sqrt{a_k b_k} \right)^2}{T - \sum\limits_{k=1}^{K} a_k} + \sum_{k=1}^{K} b_k \right]. \tag{37}$$

Comparing the minimum value with 1, we obtain the feasibility conditions

$$\left(\sum_{k=1}^{K} \sqrt{a_k b_k} \right)^2 \leq \left(T - \sum_{k=1}^{K} a_k \right) \left(T - \sum_{k=1}^{K} b_k \right) \tag{38}$$

$$\sum_{k=1}^{K} a_k \leq T, \; \sum_{k=1}^{K} b_k \leq T. \tag{39}$$

References

[1] M. Satyanarayanan, Z. Chen, K. Ha, W. Hu, W. Richter, and P. Pillai, "Cloudlets: at the leading edge of mobile-cloud convergence," in *IEEE 6th International Conference on Mobile Computing, Applications and Services (MobiCASE)*, 2014, pp. 1–9.

[2] M. Patel, B. Naughton, C. Chan, N. Sprecher, S. Abeta, A. Neal *et al.*, "Mobile-edge computing introductory technical white paper," *White Paper, Mobile-edge Computing (MEC) industry initiative*, 2014.

[3] S. Barbarossa, S. Sardellitti, and P. Di Lorenzo, "Communicating while computing: Distributed mobile cloud computing over 5G heterogeneous networks," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 45–55, 2014.

[4] C. You and K. Huang, "Multiuser resource allocation for mobile-edge computation offloading," *arXiv preprint arXiv:1604.02519*, 2016.

[5] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "Mobile edge computing: Survey and research outlook," *arXiv preprint arXiv:1701.01090*, 2017.

[6] X. Chen, "Decentralized computation offloading game for mobile cloud computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 4, pp. 974–983, 2015.

[7] E. Meskar, T. D. Todd, D. Zhao, and G. Karakostas, "Energy efficient offloading for competing users on a shared communication channel," in *IEEE International Conference on Communications (ICC)*, 2015, pp. 3192–3197.

[8] S. Barbarossa, S. Sardellitti, and P. Di Lorenzo, "Joint allocation of computation and communication resources in multiuser mobile cloud computing," in *IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2013, pp. 26–30.

[9] X. Lyu, H. Tian, P. Zhang, and C. Sengul, "Multi-user joint task offloading and resources optimization in proximate clouds," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3435–3447, 2016.