IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

SPECIAL SECTION ON TOWARDS SERVICE-CENTRIC INTERNET OF THINGS (IOT): FROM MODELING TO PRACTICE

Received September 12, 2018, accepted October 11, 2018, date of publication October 16, 2018, date of current version November 9, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2876311

Wireless Body Area Network Mobility-Aware Task Offloading Scheme

YANGZHE LIAO[1], YI HAN[1], QUAN YU[1], QINGSONG AI[1], QUAN LIU[1], AND MARK S. LEESON[2], (Senior Member, IEEE)
[1]School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China
[2]School of Engineering, University of Warwick, Coventry CV4 7AL, U.K.

Corresponding authors: Yi Han (hanyi@whut.edu.cn) and Quan Yu (yuquan@whut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51475342, Grant 51675389, Grant 61801341, and Grant 61801342, and in part by the Fundamental Research Funds for the Central Universities under Grant 2018IVA098 and Grant 2018IVA099.

ABSTRACT The increasing amount of user equipment (UE) and the rapid advances in wireless body area networks bring revolutionary changes in healthcare systems. However, due to the strict requirements on size, reliability and battery lifetime of UE devices, it is difficult for them to execute latency sensitive or computation intensive tasks effectively. In this paper, we aim to enhance the UE computation capacity by utilizing small size coordinator-based mobile edge computing (C-MEC) servers. In this way, the system complexity, computation resources, and energy consumption are considerably transferred from the UE to the C-MEC, which is a practical approach since C-MEC is power charged, in contrast to the UE. First, the system architecture and the mobility model are presented. Second, several transmission mechanisms are analyzed along with the proposed mobility-aware cooperative task offloading scheme. Numerous selected performance metrics are investigated regarding the number of executed tasks, the percentage of failed tasks, average service time, and the energy consumption of each MEC. The results validate the advantage of task offloading schemes compared with the traditional relay-based technique regarding the number of executed tasks. Moreover, one can obtain that the proposed scheme archives noteworthy benefits, such as low latency and efficiently balance the energy consumption of C-MECs.

INDEX TERMS WBANs, C-MEC, task offloading, mobility-aware.

I. INTRODUCTION

Due to the rapid developments in wireless communication technologies, wireless body area networks (WBANs) are becoming of increasing interest to research and industry. The main objective of WBANs is to facilitate communication inside or near the human body to measure different bodily attributes [1]–[3]. Typically, WBANs are divided into two categories, in-body and on-body WBANs, which are made feasible by taking advantage of small-size, low power consumption and intelligent body sensors or user equipment (UE). Deployment of a list of UE devices to continuously monitor patients' physiological signals can significantly reduce medical expenditures and improve quality of life [4]. However, there exist several technical challenges in WBANs. Firstly, due to the technical constraints of sensor batteries, the power supply becomes a major bottleneck for long-term health monitoring [5]. Secondly, the improvement of UE computing capacity to handle low service latency medical applications is still challenging. In addition, the question of how to effectively investigate the patient's mobility plays a vital role in the accuracy of WBAN performance [6].

In order to provide communication and computation cooperation, cloud computing has been proposed to enhance UE experience and the computation capacity [7], [8]. Numerous cloud-based platforms have been recommended over the last decade such as Amazon Elastic Compute Cloud and Google Cloud Platform [9]. By taking advantage of the cloud computation resources, UE can offload tasks to the cloud server and then receive the results via the downlink. However, the cloud server is generally placed in a remote place, possibly a thousand miles from the users, and consequently the remote transmission route degrades the quality of service regarding service time and energy waste.

To relieve the disadvantage of remote data transmission, an innovative paradigm of mobile computing has been recommended, known as mobile edge computing (MEC) [10], [11]. This scheme allows a shift in the employment of computing resources from the core network to network edge such as base stations or femto clouds [12]. In comparison with the traditional cloud computing facilities MEC can reduce the transmission distance and lower the energy consumption of UEs. This scheme provides cloud-like computing services at the edge of wireless communication networks such as access points. However, this technique suffers from various drawbacks when considering healthcare applications [13]. For example, with the increasing number of tasks offloaded to the MEC server, the average service time will increase severely. Moreover, since every piece of information is of great importance, UE task offloading decision-making strategy should be carefully studied. One practical solution to handle the aforementioned technical difficulties is to deploy distributed small-cloud computing servers, which are smaller than the traditional MEC as reported in [14] and can be placed in the hospital. Each small size MEC can receive the UE offloading tasks within a certain distance. In this way, the offloaded tasks can be distributed to the coordinator-based MEC (C-MEC) servers to execute promptly.

Practically, it is worth noticing that smart devices are designed to have active and idle periods to reduce energy waste [15], [16]. The idle-based scheme can potentially provide a nearby UE with unused computation resources to help with other UE execute tasks.. The data transmission between two UE entities forms a device to device (D2D) link, which can effectively decrease the network congestion when designing dense networks. Moreover, healthcare applications require reliable data transmission and rapid response once an abnormal condition is detected [1].

In this paper, we investigate a collection of UEs distributed in the network to monitor patients' health status considering the task offloading technique and mobility scenarios. The primary aim is to provide a different aspect to handle the mobility-aware resource intense applications by taking advantage of the task offloading strategy. Firstly, the system architecture and the task offloading model are introduced. Furthermore, we model patient mobility in detail by adopting a Markov model. A series of key performance metrics are selected and discussed. Three different transmission mechanisms are considered to investigate the system performance regarding the number of executed tasks. Moreover, the task offloading based schemes are further analyzed regarding service time, the percentage of failed tasks and the energy consumption of each C-MEC server.

The rest of this paper is organized as follows: In section II, the background to WBANs, edge computing and related topics are presented. Section III illustrates the proposed system architecture and Section IV presents the detailed information of the transmission mechanisms considered. Section V summarizes and discusses the system performance. Section VI concludes the paper and lists potential topics for future research.

II. RELATED WORK

Generally, WBANs are recognized as a promising technique that can provide data transmission between different UEs or connect to a coordinator within a certain range [1]. Ahmed *et al.* [17] demonstrated a people-centric healthcare monitoring system by employing a series of wearable devices; as a result, abnormal medical conditions could be detected promptly. However, the authors did not consider patient mobility and the percentage of failed tasks. One efficient solution for data transmission for WBANs is employing efficient routing schemes. Liao *et al.* [18] proposed an incremental relay-based cooperative routing protocol, which can forward data collected from smart devices to the external medical server. However, the outcomes showed that this technique consumes additional energy caused by the long transmission distances. In real-world patient-related healthcare applications such as medical images, data transfer is significant in the process of healthcare services and managing these media data from various WBANs is vital for multiple uses. Zigbee is commercially available for WBANs, all collected data from body sensors can be transmitted to a coordinator [19]. However, this technique cannot support large volumes of e-health media data derived from different resource-constrained body sensors in terms of data transmission, analysis and storage.

Numerous mobility models have been proposed in [20]–[22]. However, they were designed for wireless sensor networks and are not appropriate for WBANs because of the higher degree of difficulty in predicting human movement. Specifically, with high human mobility, choosing a global coordinator to manage the high data rate applications is not applicable. Dong and Smith [23] have presented a mechanism that applies to mitigate interference and coexistence by employing on-body relays on the WBANs to improve the network reliability. Moreover, Nabi *et al.* [22] proved that body sensors' movement could be treated as independent single mobility instances and there is no need to consider the correlation between different body sensors within the same WBAN.

Another research challenge is that UE capabilities struggle to keep up with the development of resource-intensive applications due to limitations in batteries, computation capacity and data storage. To address these challenges, several cloud-based healthcare system prototypes have been proposed, which can enhance media healthcare services at a low cost [24], [25]. However, the cloud server is usually placed several thousand miles away from the users and cannot support latency-sensitive healthcare data transmission. Typically, when one piece of UE requests task offloading to the cloud, this necessitates the use of the Internet for data transmission [10], [26]. A collection of research challenges exists when designing task offloading enabled WBANs to satisfy rigorous healthcare monitoring requirements [27]–[29]. As stated in Chen [29] proposed a game-theoretic model that can advise UE to choose where to offload the tasks. However, this approach is only applicable to the Amazon elastic compute cloud. Moreover, Nakamura *et al.* [30] found that small cell networks that can improve the network capacity and combat interference. However, such small cell networks cannot handle the necessary signal processing in the strict

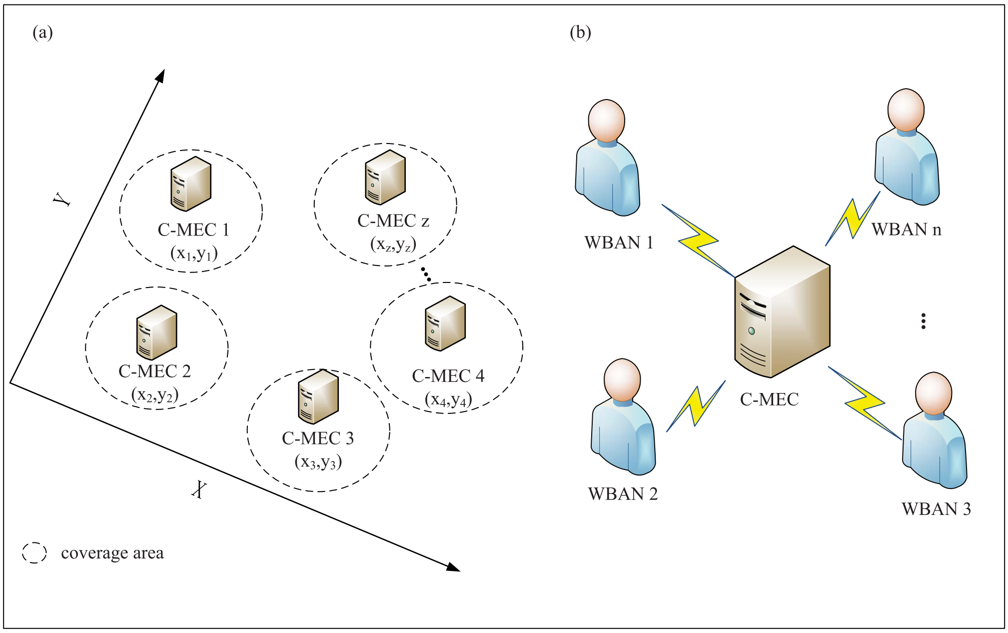

FIGURE 1. The proposed WBAN-based task offloading system.

time allowance, and therefore fail to meet the demands of healthcare applications.

III. SYSTEM ARCHITECTURE

Inspired by the WBAN coordinators proposed to gather collected data from all body sensors, power-charged high computation capacity smart devices are employed as the coordinator-based MEC (C-MEC), to which UE can offload tasks when necessary. The system configuration is shown in Fig.1. We consider a hospital-based healthcare monitoring scenario where a set of C-MECs, each C-MEC i with a coordinator (x_z, y_z) for all $1 \leq i \leq z$. Moreover, there are N_c coexisting patients that are denoted by $\{w_n | n = 1, 2, \ldots, N_c\}$ and w_n indicates the n-th WBAN. Each WBAN consists of a UE set $\{r_k | i = 1, 2, \ldots, N_s\}$ and serves one patient. The proposed system allows UE to offload tasks to the C-MEC located in the floor, which provides computation resource when UE's computation ability is insufficient.

A. TASK OFFLOADING MODEL

Definition 1: A task U_i can be defined as $U_i = (F_i, P_i, \alpha_i)$ where F_i and P_i represent the required computation resource (i.e., CPU cycles per second) and communication resource allocated to the task U_i, respectively, for all $i \in \mathcal{N}, j \in \mathcal{N}$. $\alpha_i = 1$ denotes that the smart device i decides to offload the task U_i while $\alpha_i = 0$ means that UE i decides not to offload the task.

We assume that each UE can offload the task either to the C-MEC or to the rest of the smart devices, depending on the communication and computation resource of its own, of other devices and of the C-MEC. If the UE i decides to offload its task D_i to the UE j or the C-MEC via D2D link, the maximum data rate can be given as

$$\mathcal{R}_{ij} = B_{ij} log_2 \left(1 + \frac{h_{ij} p_i^T}{\sigma^2}\right), \quad i \in \mathcal{N}, \; j \in \mathcal{N}, \tag{1}$$

where we ignore the interference caused by other devices and all channels are assumed to be orthogonal. h_{ij} is the channel state information from UE i to UE j, and p_i^T denotes the transmission power of the UE i. B_{ij} represents the allocated bandwidth to the UE j and σ^2 describes the variance of the white Gaussian noise. When UE i decides to execute the task itself, we let f_i be the computation capacity of the UE i so the execution time of local computing can be expressed as

$$T_i = \frac{F_i}{f_i}, \quad i \in \mathcal{N}. \tag{2}$$

If the UE i decides to execute the task locally, the corresponding computing power consumption can be expressed as

$$p_i = \kappa_i (f_i)^{\nu_i}, \quad i \in \mathcal{N}, \tag{3}$$

where κ_i^l and ν_i^l are the pre-configured parameters and both are positive constants as reported in [31]. Realistic measurements of those parameters are $\kappa_i^l = 10^{-11}$ and $2 \leq \nu_i^l \leq 3$. The energy consumption of UE i to execute the task locally then can be obtained as

$$E_i = p_i \cdot T_i = F_i \cdot \kappa_i \cdot (f_i)^{\nu_i - 1}, \quad i \in \mathcal{N}. \tag{4}$$

When UE i cannot execute the task by itself due to lack of computation resource f_i or the execution time $T_i \geq T_i^r$, where T_i^r is the predetermined execution time threshold, it can offload the task either to other UEs or the C-MEC, one can obtain the offloading time as

$$T_i^o(U_i) = \sum_{j \in \mathcal{N} \setminus \{i\}} \alpha_i(\frac{D_i}{\mathcal{R}_i} + \frac{F_i}{f_j}), \quad i \in \mathcal{N},\ j \in \mathcal{N},\ i \neq j. \quad (5)$$

The energy consumption of the task offloading can be calculated as

$$E_{ij}(U_i) = p_i^T \cdot T_i^o(U_i), \quad i \in \mathcal{N},\ j \in \mathcal{N},\ i \neq j. \quad (6)$$

In accordance with [32] and [33], we neglect the overhead for the C-MEC to send the computation results back to the UE because the size of the computation results is much smaller than the size of the computation input data.

B. MOBILITY ANALYSIS

Mobility is an essential aspect that cannot be ignored when investigating the network performance [6]. The patients are moving within the hospital from one place to other places randomly and this is difficult to predict. In this paper, we consider a scenario where patients are moving to different locations within the floor based on a nomadic mobility model. Moreover, we evaluate the mobility within a WBAN then present the global movement for the network.

Definition 2: We define a mobility scenario as $L = (A, \tau)$ where A represents the attractiveness levels, and τ denotes the dwell time of different attractiveness locations.

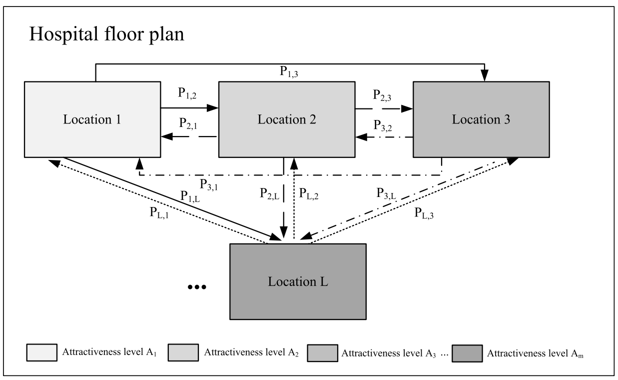

FIGURE 2. The Markov model of the one patient's global mobility.

Assume there are M locations with the corresponding attractiveness levels $A_1, A_2, \ldots, A_M$, each location with the dwell time $\tau_1, \tau_2, \ldots, \tau_M$. The probability of moving from one location to another is investigated employing the Markov model shown in Fig.2. In particular, a high level of location attractiveness means the dwell time is also higher. By using the Markov model, the current location with attractiveness level A_M and dwell time τ_M is considered for selecting the next location. When the probability that a patient moves from location i to j is P_{ij}, one can state that

$$\sum_{i=1}^{M} P_{ij} = 1, \quad 1 \leq i \leq M,\ i \in N,\ j \in N, \quad (7)$$

where $i = j$ means that the patient stays in the same place. Moreover, the locations of all patients are updated based on the probability of different attractiveness locations. The probability of changing location can be determined by real human mobility traces. The patient is considered to be continuously moving among different locations, and stays at one location for a specific period of time (dwell time) according to the attractiveness level of the location.

C. LINK QUALITY ANALYSIS

In the hospital-based healthcare monitoring scenario, one significant factor is characterizing the transmission energy attenuation between the UE and the C-MEC. At a distance d this can be mathematically expressed as

$$PL_{dB}(d) = PL_{dB}(d_0) + 10n\, log_{10}\left(\frac{d}{d_0}\right) + S, \quad d \geq d_0 \quad (8)$$

where d_0 and $PL_{dB}(d_0)$ mean the reference distance and the PL(path loss) value at d_0, respectively. S denotes the shadow fading effect, which follows a Normal distribution [1], [18]. n is the PL coefficient, which depends on the transmission environment. As a result, the transmission distance d can significantly affect the link quality. Moreover, one should note that efficiently allocating the number of offloaded tasks for each C-MEC to balance the computing and communication costs is of great significance to realize low latency.

IV. PROBLEM FORMULATION

In this section, we first demonstrate numerous existing selected decentralized task offloading techniques regarding local computation and task offloading schemes. Moreover, a mobility-aware cooperative task offloading scheme is proposed in detail.

The all tasks locally executed scheme has been widely used in the literature due to its simplified structure [1], [35]. This architecture allows UE to execute all tasks locally. As analyzed in (2) and (6), task failure reasons for this strategy can be summarized as follows: 1) when the execution time $T_i^l > T_i^r$, which represents that UE i cannot provide enough computation capacity to execute the large length tasks. 2) the residual energy status of the UE i becomes $E_i^{res} < E_{ij}(U_i)$.

Another effective technique is the cooperative task offloading scheme. All tasks are initially checked by UE and if one UE cannot handle the task, this will offload to another UE within the same patient. If the selected UE executes the task unsuccessfully, the task will offload to the C-MEC. The advantage of this transmission technique is that the task is expected to be accomplished within the same WBAN where the medical notices or warnings can be delivered promptly to the patient. This is a useful solution to solve the scenarios when all C-MECs are far away from the patient or when there are too many UE simultaneous offloading requests [13], [36].

The critical idea of the proposed approach in this paper is to decrease service time and achieve load balancing in comparison with the approaches above, which makes it suitable for the offloading of latency sensitive healthcare applications. The mobility-aware cooperative task offloading scheme is proposed to investigate *patient-on-the-go* conditions where patients move randomly within the hospital. In addition, unlike the work in [7] and [34], we consider tasks generated with random task lengths and thus the UE needs to decide whether to offload the task or not after each task is generated. One algorithm regarding location selection and task generation is given in Algorithm 1. We assume that tasks are randomly generated according to a Poisson distribution and patients move according to the nomadic mobility model given in Section III. UE i will decide whether to locally execute or offload to another UE after one task generated. A failed task execution is as follows:

- When the patients move outside the C-MECs' coverage area, the task cannot offload to the C-MEC. One should note that commercially available platforms, such as Sensium, cover the range from a few millimeters to a few meters [37]–[39].
- There exists a large number of patients (dense network) producing many task offloading requests, which leads to longer service times.
- Some healthcare applications require strict latency requirements such as medical video transmission. If one task cannot be accomplished within T_i^r, the task can be regarded as an uncompleted task.

Due to the resource-limited nature of UEs, we consider the transmission strategies of UEs within the same WBAN. One effective transmission strategy is proposed if the overhead $V(U_i)$ can be minimized, the problem formulation can be expressed as

$$\min_{\alpha_i \in \{0,1\}} V(U_i), \quad i \in \mathcal{N}, \tag{9}$$

where $V(U_i)$ is the overhead that consists of processing time and energy consumption. Recalling (2) and (4), one can obtain the following equation when $\alpha_i = 0$

$$V_i(U_i) = T_i + \lambda_i E_i(U_i), \quad i \in \mathcal{N}, \tag{10}$$

where T_i and E_i represent the time and energy consumption for local computing, respectively. λ_i is the weight factor, which depends on the medical requirements. Similarly, when $\alpha_i = 1$, $V(U_i)$ can be rewritten as

$$V_i^o(U_i) = T_i^o + \lambda_i^o E_{ij}(U_i), \quad i \in \mathcal{N},\ j \in \mathcal{N},\ i \neq j, \tag{11}$$

where T_i^o and $E_{ij}(U_i)$ have been explained in (5) and (6), respectively. λ_i^o is the weight factor when the UE decides to offload the task.

A. TASK GENERATION AND LOCATION SELECTION

In this paper, we assign each static location with an attractiveness level that determines the dwell period that a patient would stay. At the time of t, one patient can move from location i (with the corresponding attractiveness level A_t and dwell time τ_t) to location j according to the probability P_{ij}. The patient stays at the updated location for a period of time τ^t. Tasks $\mathcal{T}_1 \ldots \mathcal{T}_n$ are generated by the UEs on a patient that is bound to one location depending on the active period, the idle period and Poisson interarrival configurations.

Algorithm 1 Proposed Iterative Method for Task Offloading Scheme

Input: predeterminate time threshold T_i^r
residual energy status E_i^{res}
Output: ρ_e,ρ_t
1: **if** $E_i^{res} \geq E_i^{thr}$ **then**
2: calculate the execution time T_i^C for task $\mathcal{T}_i$;
3: **if** $T_i^C \leq T_i^r$ **then**
4: $a_i^* = 0$; // task $\mathcal{T}_i$ is executed on UE i
5: UE i updates its energy status E_i^{res} after the task executed;
6: **else**
7: $a_i^* = 1$;
8: task $\mathcal{T}_i$ fails to execute;
9: $\rho_t = \rho_t + 1$;
10: **end**
11:**else** //all UEs do not have sufficient residual energy for the task;
12: generate tasks $\mathcal{T}$ at location l^{t+1} from starting time τ^t to
τ^{t+1}, depending on the active period, the idle period and Possion interarrival;
13: task $\mathcal{T}_i$ fails to execute;
14: $\rho_e = \rho_e + 1$;
14:**end**

B. TASK OFFLOADING DECISION PROCESS

We consider the resource-limited UE i with a task D_i where the length of D_i is randomly generated.

Lemma 1: Consider the strategies of UE i, one can follow the task execution strategy a_i^* to maintain the network reliability.

$$a_i^* = \begin{cases} 1, & if \ \sum_{i \in \mathcal{N}} E_i \leq E_i^{thr}, \\ 0, & \text{otherwise}, \end{cases} \tag{12}$$

where E_i^{thr} means the predeterminated threshold of the residual energy of UE i and can be expressed as

$$E_i^{thr} = (T_i^o + \lambda_i^o E_i(U_i) - T_i)/\lambda_i \tag{13}$$

Proof: For UE i, $i \in \mathcal{N}$, the overhead can be expressed as

$$\begin{aligned} V_h &= V_i + V_i^o \\ &= (T_i + \lambda_i E_i) + (T_i^o + \lambda_i^o E_{ij}(U_i)), \end{aligned} \tag{14}$$

Consider to decrease the overhead of UE i, one can obtain that $V_i^o \geq V_i^l$. Therefore, (13) can be rewritten as

$$T_i^o + \lambda_i^o E_i(U_i) \geq T_i + \lambda_i E_{ij}(U_i) \tag{15}$$

Then one can obtain that

$$E_i^{thr} \leq T_i^o(U_i) + \lambda_i^o E_{ij}(U_i) - T_i)/\lambda_i \tag{16}$$

The advantage of the proposed scheme is that all residual UE energy is considered as a whole resource for all tasks generated by different sources. As a result, the UE energy can be balanced and the percentage of failed tasks can be decreased. Detailed information regarding the proposed iterative method for task offloading scheme is demonstrated in Algorithm 1. According to the predeterminate time threshold T_i^r and the residual energy status E_i^{res} of UE i, the task offloading strategy can be executed in each round to decrease the overhead and maintain the network stability. Moreover, the number of failed tasks caused by lack of residual energy ρ_e and exceeding the allowance time threshold ρ_t can be obtained.

C. TASK OFFLOADING RECEIPT SELECTION

Considering that one C-MEC covers an area with a radius of R_c, when the UE i decides to offload the task to the C-MEC, the selection function can be given as

$$C_j(w) = \frac{d_u(w)}{E_u(w)}, \quad R_c \geq d_u(w) \tag{17}$$

where w is the C-MEC's ID, $d_u(w)$ is the distance between the UE i and the potential C-MEC and $E_u(w)$ means the residual energy of the C-MEC. $E_u(w)$ is obtained by subtracting the currently consumed energy of the offloading task from the previous residual energy. After the task offloading request the UE obtains the locations and energy status of all potential C-MECs, and then selects the one with the minimum $C_j(w)$ as the task offloading receiver, which can efficiently achieve load balancing. One should note that when $R_c < d_u(w)$, the number of failed tasks due to mobility ρ_c will increase as demonstrated in Algorithm 2.

Algorithm 2 C-MEC Selection

Initialization: C-MEC is assigned an unique ID w;
C-MEC coverage area with a radius of R_w;
number of failed tasks due to coverage is set to ρ_c;
1: **for** $i \in \mathcal{N}$ **do**
2: calculate $d_u(w)$;
3: calculate $E_u(w)$;
4: **if** $R_c \geq d_u(w)$ **then**
5: $\rho_c = \rho_c + 1$;
6: **else**
7: find $\underset{j\in\mathcal{N}}{\operatorname{argmin}} C_j(w)$;
8: **end**
9: **end for**
10:return w, ρ_c;

D. DATA SCHEDULING AND TRANSMISSION

In this phase, the offloaded tasks are executed on a first-come-first-serve manner by the corresponding C-MEC. We define the number of failed tasks due to lack of residual energy E_i^{res} and exceeding the predetermined time threshold of T_i^r as ρ_t and ρ_e, respectively.

After the active period, the UE then switches to the idle period to decrease energy waste. The percentage of failed tasks r can be expressed as the ratio between the number of failed tasks ρ_{failed} and the total number of generated tasks ρ_{total} thus

$$r = \frac{\rho_{failed}}{\rho_{total}} * 100\%, \tag{18}$$

where

$$\rho_{failed} = \rho_c + \rho_t + \rho_e. \tag{19}$$

The detailed information regarding calculating ρ_{failed} can be found in Algorithm 2. UE i will update its energy status and select the task offloading strategy after one task is generated. In this way, the energy consumption of UE i can be significantly reduced and thereby the network lifetime can be prolonged.

V. PERFORMANCE EVALUATION AND DISCUSSION

A. THE SELECTED KEY PERFORMANCE METRICS

Some key parameters are proposed as follows:

Executed task: One successfully executed task can be seen as implementation by the UE population or offloading receipts when the requirements proposed in Algorithm 2 are met.

Failed task: One failed task can be defined as a task that cannot satisfy the requirements mentioned in Section IV. In this paper, uncompleted tasks are seen as failed ones.

Average service time: This is an important metric to represent the total time of the task execution services. Generally, the service time will significantly increase when a large amount of patient congestion occurs in the same place(s).

B. PERFORMANCE EVALUATION AND DISCUSSION

In this section, theoretical analysis is employed for evaluating the performance of the proposed schemes. Due to the strict requirements of the human body safety requirements, the transmission power is set as −12 dBm as this is maximum value regulated by the IEEE 802.15.6 technical standard [1], [40]. The parameters related to the simulation are summarized in Table 1. The patients are moving to different attractiveness locations randomly as mentioned in Section III. The decision weights $\lambda_i = \lambda_i^o = 1$ as the same with [29]. We investigate the network performance of the proposed schemes regarding the selected key performance

TABLE 1. Simulation parameters.

Parameters	Value (unit)
Poisson interarrival time	25 s
Simulation time	1 hour
Computation capacity	1500
Bandwidth	1 Mbps
T_i^r threshold	10 ms
Transmission power	0.5 mW
Number of attractiveness locations L	5
Average task length (low)	250 MI
Average task length (medium)	350 MI
Average task length (heavy)	500 MI
The active period of the UE	15 s
The idle period of the UE	5 s
Probability of P_{L_k}	(0.1,0.15,0.2,0.25,0.3)
Attractiveness average waiting time	(300,200,100,50,20)

metrics; the number of executed tasks, the percentage of the failed task, average service time and energy consumption of all MECs. Also, each patient is provided with five pieces of UE fixed at predetermined locations to monitor the physiological signals. The average task size k proposed as the low, medium and the heavy payload is defined as 250, 350 and 500 million instructions (MI), respectively. In the same manner as [41], we ignore the effects of channel interference on the data transmission. Also, we have assumed there are 25 C-MECs fixed in the predeminated positions with different attractiveness values as shown in Fig. 3, as shown in Fig. 3, each C-MEC covers an area of a circular region with a radius R_c of 5 meters by default. The system performance investigation was implemented using JAVA in the 64-bit Windows 10 Professional operating system using an Intel Xeon(R) E5-1630@3.70GHz processor with 16 GB RAM. We ran the simulation five times and obtained the 90% confidence interval for further analysis. The detailed explanations of the crucial results are given in detail below.

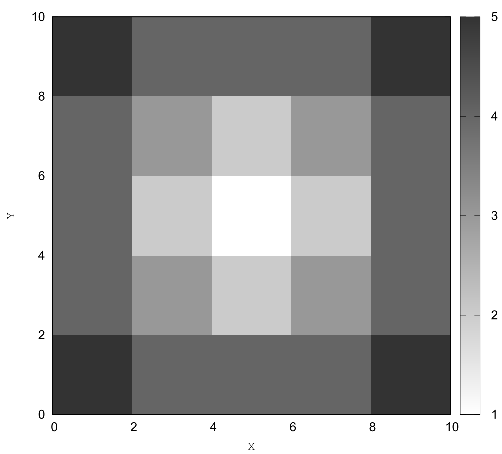

FIGURE 3. **The different locations with multiple attractiveness.**

Figs. 4-6 show the comparison between a series of transmission techniques under different average task lengths. One should note that since all tasks are randomly generated and the tasks for UE to execute or offload thus have random lengths. It can be seen that as the number of patients increases, the number of executed tasks increases as one would expect. The task offloading scheme executes approximately 41500, 41300 and 44500 tasks for average task lengths of 250, 350 and 500 MI respectively. When considering patients' mobility, the corresponding figures are 38300, 37900 and 37600. This is because many tasks cannot offload to a C-MEC when patients are moving out of the coverage area of all C-MECs. The total numbers of executed tasks with local execution are roughly 18100, 14800 and 12300 for the average task lengths of 250, 350 and 500 MI, respectively. It should also be noted that with local execution, the number of executed tasks plateaus when the number of patients increases from 180 to 200 as shown in Fig. 5. This is because the number of generated tasks keeps increasing but reaches the limit of the total UE computational capabilities, resulting in significant task failure. Therefore, this strategy is not suitable for computation-intensive applications.

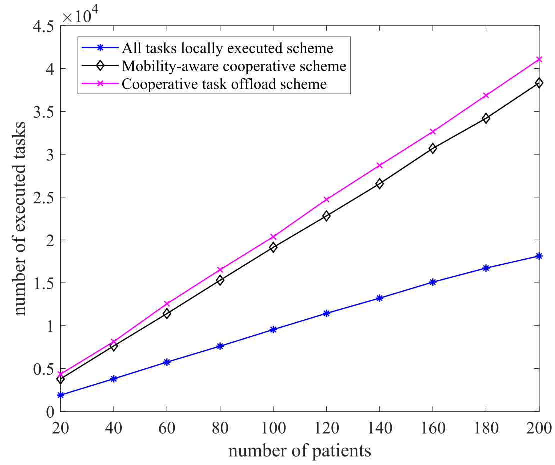

FIGURE 4. **Comparison of the number of executed tasks when $k = 250$ MI.**

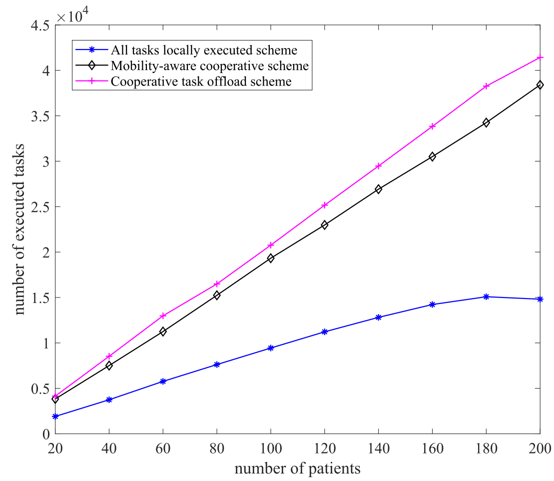

FIGURE 5. **Comparison of the number of executed tasks when $k = 350$ MI.**

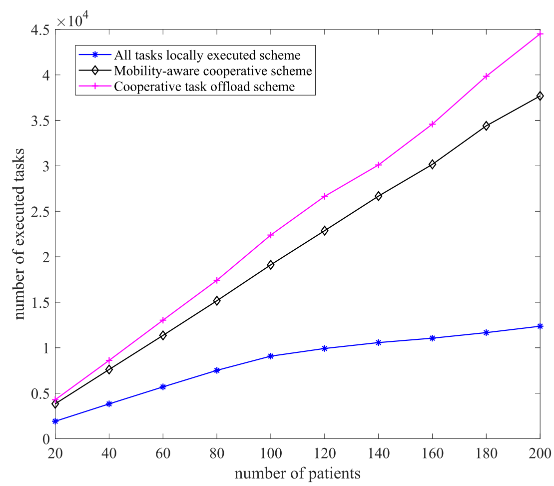

FIGURE 6. **Comparison of the number of executed tasks when $k = 500$ MI.**

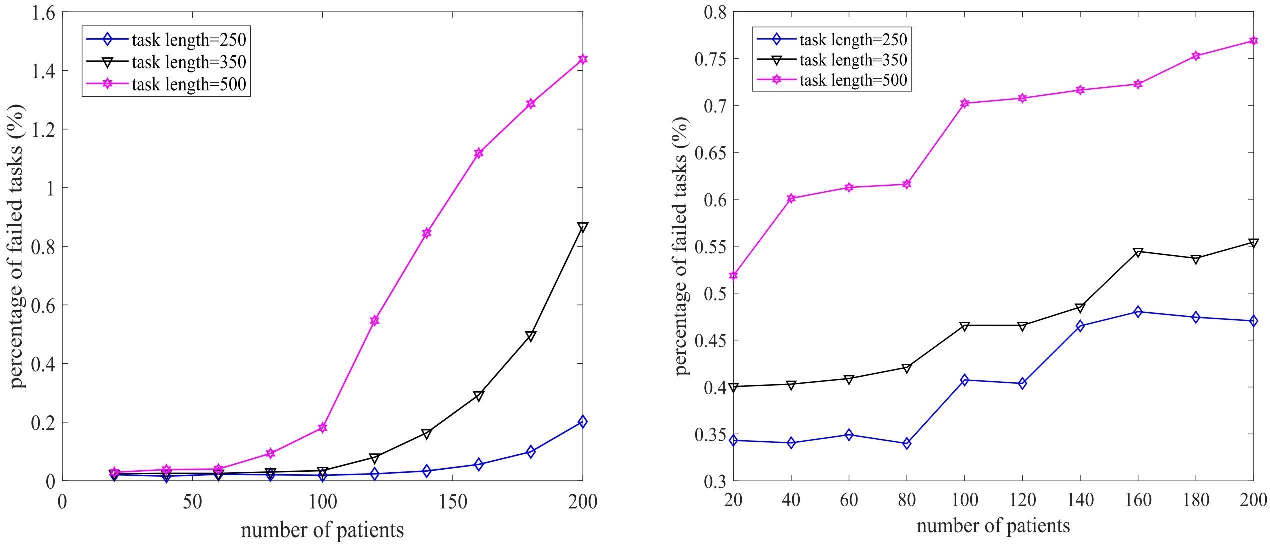

FIGURE 7. The average percentage of failed tasks for different task lengths. (left) cooperative task offload scheme; (right) the proposed scheme.

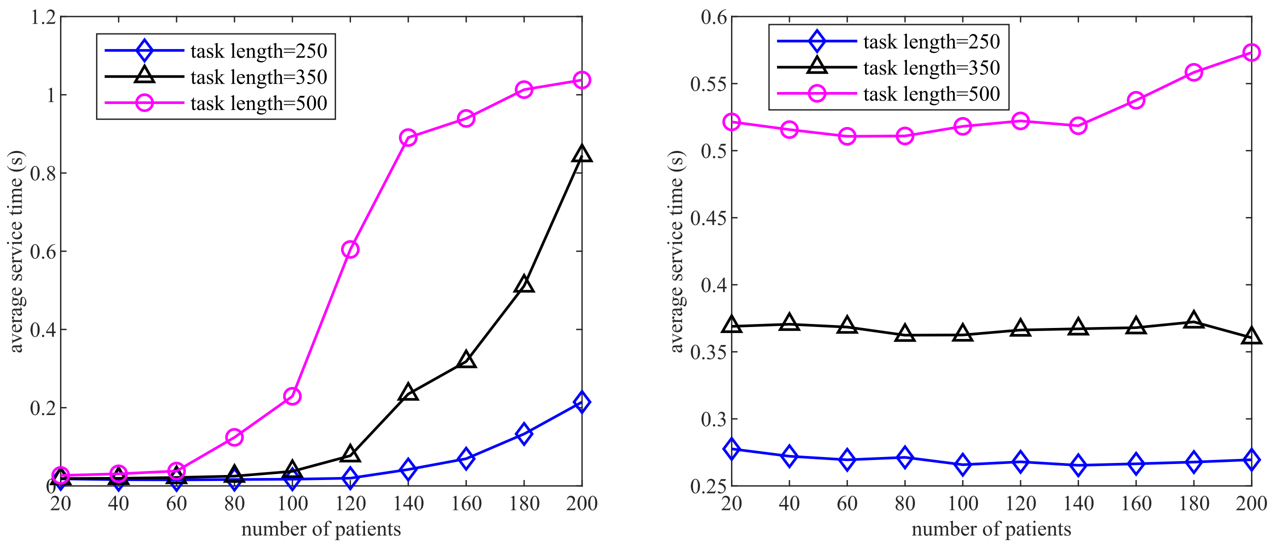

FIGURE 8. The behavior of the average service time. (left) cooperative task offload scheme; (right) the proposed scheme.

We now pursue the impact of human mobility further. In Fig. 7, the percentage of failed tasks under different task length conditions for multiple schemes is given. The portion of failed tasks increases with the number of patients and larger task lengths result in higher failed percentages. For length $k = 500$ MI, the task offloading scheme reaches a failure rate of 1.45% with 200 patients compared to our proposed scheme delivering just 0.75%.

The performance of the average service time under task offloading scenarios is illustrated in Fig. 8. Considering the heavy payload condition, the service time is nearly 0.57 s when used for 200 patients while only 0.27 s when the task length is 250 MI. Moreover, the task offloading scheme achieves the highest service time when task length is 500 MI and there are 200 patients at around 1.05 s. One can obtain that the proposed mobility-aware cooperative task offloading scheme realizes stable average service time performance when considering low and medium payload scenarios. This is because each UE can select the best task offloading recipient according to the selection function proposed in (17).

Figs. 9-11 summarize the energy consumption of each MEC under various conditions. These show that the some of the C-MEC servers consume significant energy such as 8, 12, 13, 14 and 18. As for our proposed approach, the power consumption of all MECs is nearly the same. This is because of the UEs' MEC selection scheme proposed in Algorithm 2, which means that the load can be balanced across MECs. The energy consumption of all three conditions is approximately 7800 Joule (J), 9100 J and 13500 J for the average task lengths of 250 MI, 350 MI and 500 MI, respectively.

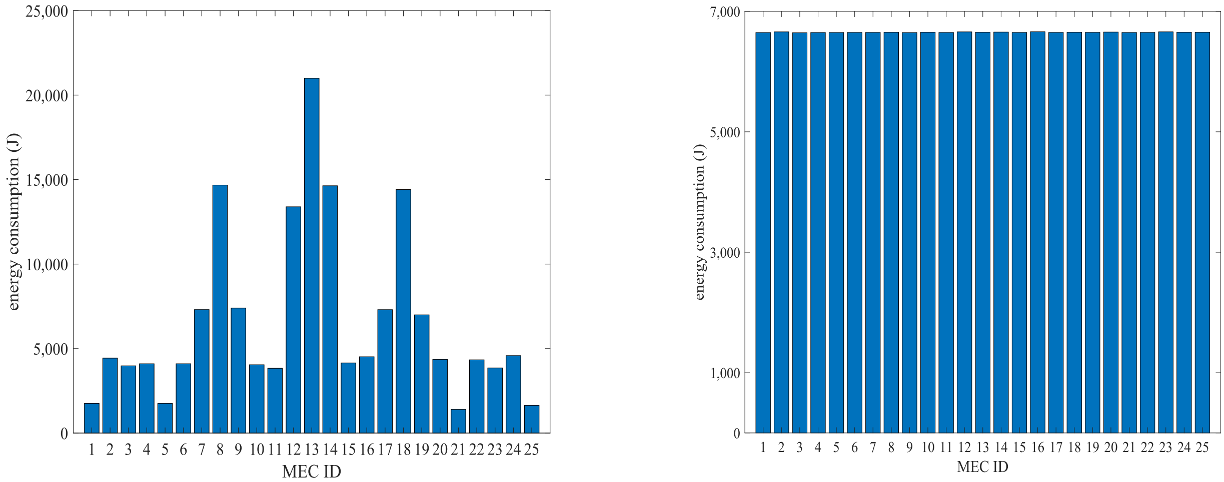

FIGURE 9. The average energy consumption when task length $k = 250$ MI. (left) cooperative task offload scheme; (right) the proposed scheme.

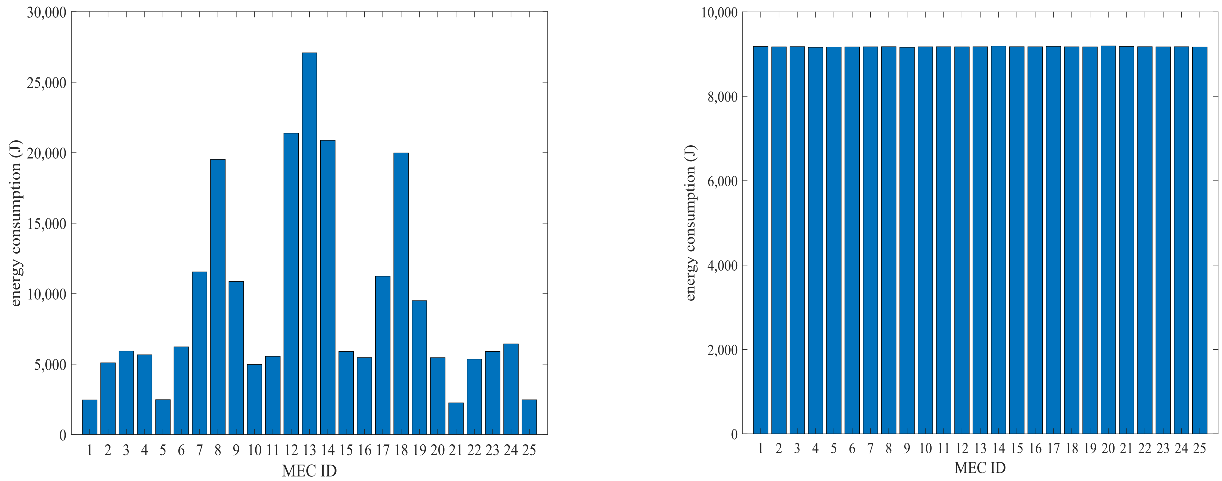

FIGURE 10. The average energy consumption when task length $k = 350$ MI. (left) cooperative task offload scheme; (right) the proposed scheme.

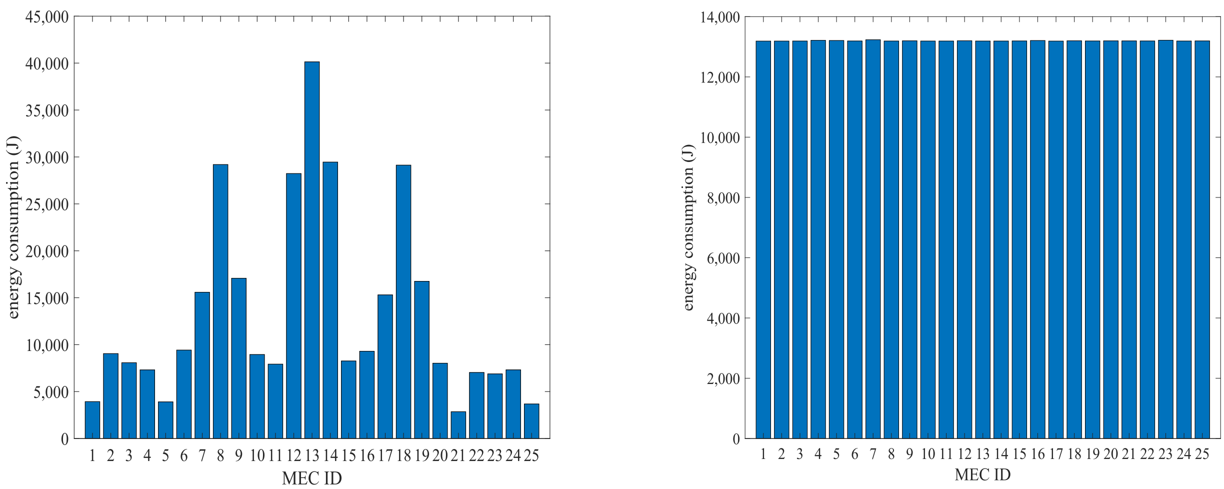

FIGURE 11. The average energy consumption when task length $k = 500$ MI. (left) cooperative task offload scheme; (right) the proposed scheme.

VI. CONCLUSION

Edge computing provides a great opportunity to reduce latency in WBAN-based healthcare where limitations in UE computation, communication and energy storage capabilities set performance limits. In this paper, a mobility-aware cooperative task offloading scheme is proposed, which employs C-MECs as computation platforms and WBANs as the communication interface. We first present the system architecture and the patient mobility model. Moreover, existing approaches have been analyzed and compared with the proposed one. Algorithms regarding decision making and transmission mechanism have been proposed in detail. The results show that the traditional relay-based transmission scheme achieves poor performance in terms of the number of executed tasks. When comparing the previously published task offloading scheme with the proposed one, it is seen that the new mobility-aware cooperative task offloading approach delivers better performance in term of several aspects such as average service time, the percentage of failed tasks and energy consumption balancing of all C-MECs.

Future work includes the implementation of the proposed protocol on a realistic experimental testbed. Moreover, since the hospital-based healthcare monitoring technique is becoming a promising candidate to decrease medical costs, joint energy minimization and resource allocation in the UEs and C-MECs, femto cloud-based fault tolerance and task offloading system design are worthy of further investigation [7], [42].

REFERENCES

[1] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, and A. Jamalipour, "Wireless body area networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1658–1686, 3rd Quart., 2014.

[2] T. Hayajneh, G. Almashaqbeh, S. Ullah, and A. V. Vasilakos, "A survey of wireless technologies coexistence in WBAN: Analysis and open research issues," *Wireless Netw.*, vol. 20, no. 8, pp. 2165–2199, 2014.

[3] G. V. Crosby, T. Ghosh, R. Murimi, and C. A. Chin, "Wireless body area networks for healthcare: A survey," *Int. J. Ad Hoc Sensor Ubiquitous Comput.*, vol. 3, no. 3, pp. 1–26, 2012.

[4] M. Chen, Y. Ma, J. Song, C. F. Lai, and B. Hu, "Smart clothing: Connecting human with clouds and big data for sustainable health monitoring," *Mobile Netw. Appl.*, vol. 21, no. 5, pp. 825–845, 2016.

[5] Y. Liao, M. S. Leeson, and M. D. Higgins, "Flexible quality of service model for wireless body area sensor networks," *Healthcare Technol. Lett.*, vol. 3, no. 1, pp. 12–15, 2017.

[6] B.-S. Kim, K. H. Kim, and K.-I. Kim, "A survey on mobility support in wireless body area networks," *Sensors*, vol. 17, no. 4, p. 797, 2017.

[7] K. Wang, K. Yang, H.-H. Chen, and L. Zhang, "Computation diversity in emerging networking paradigms," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 88–94, Feb. 2017.

[8] J. Oueis, E. C. Strinati, and S. Barbarossa, "The fog balancing: Load distribution for small cell cloud computing," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, Glasgow, U.K., May 2015, pp. 1–6.

[9] N. Sultan, "Cloud computing for education: A new dawn?" *Int. J. Inf. Manage.*, vol. 30, pp. 109–116, Apr. 2010.

[10] Y. C. Hu, M. Patel, D. Sabella, D. Sprecher, and N. Young, "Mobile edge computing—A key technology towards 5G," Eur. Telecommun. Standards Inst., Sophia Antipolis, France, White Paper 11, 2015, pp. 1–16.

[11] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, Fog *et al.*: A survey and analysis of security threats and challenges," *Future Gener. Comput. Syst.*, vol. 78, pp. 680–698, Jan. 2016.

[12] M. Aazam, S. Zeadally, and K. A. Harras, "Offloading in fog computing for IoT: Review, enabling technologies, and research opportunities," *Future Gener. Comput. Syst.*, vol. 87, pp. 278–289, Oct. 2018.

[13] W. Shi and S. Dustdar, "The promise of edge computing," *Computer*, vol. 49, no. 5, pp. 78–81, 2016.

[14] B. R. Chang, H.-F. Tsai, C.-M. Chen, Z.-Y. Lin, and C.-F. Huang, "Assessment of hypervisor and shared storage for cloud computing server," in *Proc. 3rd Int. Conf. Innov. Bio-Inspired Comput. Appl.*, Kaohsiung, Taiwan, 2012, pp. 67–72.

[15] S. Pathak, M. Kumar, A. Mohan, and B. Kumar, "Energy optimization of ZigBee based WBAN for patient monitoring," *Procedia Comput. Sci.*, vol. 70, pp. 414–420, Nov. 2015.

[16] D.-T. Huynh and M. Chen, "An energy efficiency solution for WBAN in healthcare monitoring system," in *Proc. Int. Conf. Syst. Inform.*, Nov. 2017, pp. 685–690.

[17] S. Ahmed *et al.*, "Co-LAEEBA: Cooperative link aware and energy efficient protocol for wireless body area networks," *Comput. Hum. Behav.*, vol. 51, pp. 1205–1215, Oct. 2015.

[18] Y. Liao, M. S. Leeson, M. D. Higgins, and C. Bai, "An incremental relay based cooperative routing protocol for wireless in-body sensor networks," in *Proc. IEEE Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, New York, NY, USA, Oct. 2016, pp. 1–6.

[19] Y. Kim, S. Lee, and S. Lee, "Coexistence of ZigBee-based WBAN and WiFi for health telemonitoring systems," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 1, pp. 222–230, Jan. 2016.

[20] S. Movassaghi, A. Majidi, A. Jamalipour, D. Smith, and M. Abolhasan, "Enabling interference-aware and energy-efficient coexistence of multiple wireless body area networks with unknown dynamics," *IEEE Access*, vol. 4, pp. 2935–2951, 2016.

[21] M. M. Sandhu *et al.*, "Modeling mobility and psychological stress based human postural changes in wireless body area networks," *Comput. Hum. Behav.*, vol. 51, pp. 1042–1053, Oct. 2015.

[22] M. Nabi, M. Geilen, and T. Basten, "MoBAN: A configurable mobility model for wireless body area networks," in *Proc. Int. ICST Conf. Simulation Tools Techn.*, 2011, pp. 168–177.

[23] J. Dong and D. Smith, "Joint relay selection and transmit power control for wireless body area networks coexistence," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 5676–5681.

[24] M. Quwaider and Y. Jararweh, "Cloudlet-based efficient data collection in wireless body area networks," *Simul. Model. Pract. Theory*, vol. 50, pp. 57–71, Jan. 2015.

[25] G. Almashaqbeh, T. Hayajneh, A. V. Vasilakos, and B. J. Mohd, "QoS-aware health monitoring system using cloud-based WBANs," *J. Med. Syst.*, vol. 38, p. 121, Oct. 2014.

[26] M. Altamimi, A. Abdrabou, K. Naik, and A. Nayak, "Energy cost models of smartphones for task offloading to the cloud," *IEEE Trans. Emerg. Topics Comput.*, vol. 3, no. 3, pp. 384–398, Sep. 2015.

[27] E. H. Cherkaoui and N. Agoulmine, "Context-aware mobility management with WiFi/3G offloading for ehealth WBANs," in *Proc. IEEE Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Natal, Brazil, Oct. 2014, pp. 472–476.

[28] Z. Li, H. Wang, M. Daneshmand, and H. Fang, "Secure and efficient key generation and agreement methods for wireless body area networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–6.

[29] X. Chen, "Decentralized computation offloading game for mobile cloud computing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 4, pp. 974–983, Apr. 2015.

[30] T. Nakamura *et al.*, "Trends in small cell enhancements in LTE advanced," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 98–105, Feb. 2013.

[31] A. P. Miettinen and J. K. Nurminen, "Energy efficiency of mobile clients in cloud computing," in *Proc. USENIX Conf. Hot Topics Cloud Comput.*, Boston, MA, USA, 2010, p. 4.

[32] C. Wang, F. R. Yu, C. Liang, Q. Chen, and L. Tang, "Joint computation offloading and interference management in wireless cellular networks with mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7432–7445, Aug. 2017.

[33] X. Lyu *et al.*, "Selective offloading in mobile edge computing for the green Internet of Things," *IEEE Netw.*, vol. 32, no. 1, pp. 54–60, Jan./Feb. 2018.

[34] H. Mei, K. Wang, and K. Yang, "Multi-layer cloud-RAN with cooperative resource allocations for low-latency computing and communication services," *IEEE Access*, vol. 5, pp. 19023–19032, 2017.

[35] N. Javaid, A. Ahmad, Q. Nadeem, M. Imran, and N. Haider, "iM-SIMPLE: iMproved stable increased-throughput multi-hop link efficient routing protocol for wireless body area networks," *Comput. Hum. Behav.*, vol. 51, pp. 1003–1011, Oct. 2015.

[36] M. M. Hassan, K. Lin, X. Yue, and J. Wan, "A multimedia healthcare data sharing approach through cloud-based body area network," *Future Gener. Comput. Syst.*, vol. 66, pp. 48–58, Jan. 2016.

[37] M. Patel and J. Wang, "Applications, challenges, and prospective in emerging body area networking technologies," *IEEE Wireless Commun.*, vol. 17, no. 1, pp. 80–88, Feb. 2010.

[38] P. J. Soh, G. A. E. Vandenbosch, M. Mercuri, and D. M. M.-P. Schreurs, "Wearable wireless health monitoring: Current developments, challenges, and future trends," *IEEE Microw. Mag.*, vol. 16, no. 4, pp. 55–70, May 2015.

[39] K. S. Kwak, S. Ullah, and N. Ullah, "An overview of IEEE 802.15.6 standard," in *Proc. Int. Symp. Appl. Sci. Biomed. Commun. Technol. (ISABEL)*, Rome, Italy, 2010, pp. 1–6.

[40] Y. Mao, J. Zhang, Z. Chen, and K. B. Letaief, "Dynamic computation offloading for mobile-edge computing with energy harvesting devices," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3590–3605, Dec. 2016.

[41] C. Sonmez, A. Ozgovde, and C. Ersoy, "Performance evaluation of single-tier and two-tier cloudlet assisted applications," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Paris, France, 2017, pp. 302–307.

[42] K. Wang, K. Yang, and C. S. Magurawalage, "Joint energy minimization and resource allocation in C-RAN with mobile cloud," *IEEE Trans. Cloud Comput.*, vol. 6, no. 3, pp. 760–770, Sep. 2018.

YANGZHE LIAO received the B.S. degree in measurement and control technology from Northeastern University, China, in 2013, and the Ph.D. degree from The University of Warwick, U.K., in 2017. He is currently a Lecturer at the School of Information Engineering, Wuhan University of Technology. His research interests include wireless body area networks, mathematical modeling, and mobile cloud computing.

YI HAN received the B.Eng. degree from the International School of Software, Wuhan University, China, in 2010, the M.S. degree in telecommunication from Dublin City University, in 2011, and the Ph.D. degree from the Performance Engineering Laboratory, University College Dublin, Ireland. He is currently a Lecturer at the School of Information Engineering, Wuhan University of Technology. His research interests include QoE assessment and performance-aware adaptive multimedia deliveries.

QUAN YU received the B.S. degree in electronic information engineering from Central China Normal University, China, in 2009, and the Ph.D. degree from the City University of Hong Kong in 2014. She then progressed through the post-doctoral fellow position for two years at the City University of Hong Kong before undertaking the Lecturer at the School of Information Engineering, Wuhan University of Technology. Her research interests include distributed storage systems, cloud storage, and network coding.

QINGSONG AI received the M.S. and Ph.D. degrees in information engineering from the Wuhan University of Technology, Wuhan, China, in 2006 and 2008, respectively. From 2006 to 2007, he was a Visiting Researcher at the University of Auckland, Auckland, New Zealand, where he was involved in medical robots. He is currently a Professor at the Wuhan University of Technology and a Senior Editor of *Cogent Engineering*. He has authored over 50 technical publications, proceedings, and editorials. In recent years, he has directed over 10 research projects. His research interests include signal processing, rehabilitation robots, and advanced manufacturing technology.

QUAN LIU received the Ph.D. degree in mechanical engineering from the Wuhan University of Technology, Wuhan, China, in 2003. She is currently a Professor at the Wuhan University of Technology. During recent years, she authored over 60 technical publications, proceedings, editorials, and books. She has directed over 20 research projects. Her research interests include signal processing, embedded systems, and robots and electronics. He received two national awards and three provincial and ministerial awards. She received the National Excellent Teacher in 2007. She is a Council Member of the Chinese Association of Electromagnetic Compatibility and the Hubei Institute of Electronics.

MARK S. LEESON (SM'08) received the B.Sc. and B.Eng. degrees (Hons.) in electrical and electronic engineering from the University of Nottingham, U.K., in 1986, and the Ph.D. degree in engineering from the University of Cambridge, U.K., in 1990. From 1990 to 1992, he was a Network Analyst with National Westminster Bank, London. After holding academic posts in London and Manchester, he joined the School of Engineering, Warwick, in 2000, where he is currently a Reader. He has over 230 publications and has supervised 15 successful research students. His major research interests are coding and modulation, nanoscale communications, and evolutionary optimization. He is a fellow of both the Institute of Physics and the U.K. Higher Education Academy.