A Learning-based Power Control Scheme for Edge-based eHealth IoT Systems

Haoru Su[1], Xiaoming Yuan[2], Yujie Tang[3], Rui Tian[1], Enchang Sun[1] and Hairong Yan[1,*]
[1] Faculty of Information Technology, Beijing University of Technology
Beijing, 100124 China
[e-mail: {suhaoru, rui.tian, ecsun, yanhairong}@bjut.edu.cn]
[2] Qinhuangdao Branch Campus, Northeastern University
Qinhuangdao, China
[e-mail: yuanxiaoming@neuq.edu.cn]
[3] Algoma University, Canada
[e-mail: yujie.tang@algomau.ca]
[*]Corresponding author: Hairong Yan

Abstract

The Internet of Things (IoT) eHealth systems composed by Wireless Body Area Network (WBAN) has emerged recently. Sensor nodes are placed around or in the human body to collect physiological data. WBAN has many different applications, for instance health monitoring. Since the limitation of the size of the battery, besides speed, reliability, and accuracy; design of WBAN protocols should consider the energy efficiency and time delay. To solve these problems, this paper adopt the end-edge-cloud orchestrated network architecture and propose a transmission based on reinforcement algorithm. The priority of sensing data is classified according to certain application. System utility function is modeled according to the channel factors, the energy utility, and successful transmission conditions. The optimization problem is mapped to Q-learning model. Following this online power control protocol, the energy level of both the senor to coordinator, and coordinator to edge server can be modified according to the current channel condition. The network performance is evaluated by simulation. The results show that the proposed power control protocol has higher system energy efficiency, delivery ratio, and throughput.

1. Introduction

World aged tendency of population and growth of chronic diseases result in rise of healthcare cost [1]. The challenge faced by the traditional healthcare services can be solved by technological advancements in microelectronics, wireless communications, and computing [2]. Technology of Wireless Body Area Networks (WBANs) can be applied in the Internet of Things (IoT) eHealth system, providing reliable, scalable, and robust health monitoring services [3]. Several sensor nodes and actuators that may deploy inside or around the human body, communicating to the coordinators (sink node) compose the Wireless Body Area Networks [4]. Physiological and environment data are collected by body and environment sensors. Different application systems gather different information, such as temperature, breath, blood pressure, glucose, SpO_2, heart-rate, ECG, and so on [5].

One of the vital challenges in WBANs is the efficiency of energy. The body and environment sensors of WBANs are implanted in or wore on human body, or attached to clothes. Thus the battery is inconvenient to recharge or replace. Due to the size constraints, the capacity of battery has strict limitation. Study indicates that the transmission of control and data frame is a major part of device energy consumption of the WBAN. The transmission power control (TPC) protocol is one of the crucial mechanisms in energy-restricted WBANs. The TPC protocol can moderate the energy consumption of sensor nodes and thus extend the network lifecycle. The transmission power control protocol adapts the transmission power level of sensor nodes and coordinators of WBAN according to the communication link state. The power level should not only maintain the reliability of the data transmission but also optimizing the utilization of energy.

The existing power control protocols for WBANs are based on the traditional cloud network architecture. We apply the edge-cloud network architecture, and propose a two-stage transmission power control (TPC) protocol on reinforce learning for eHealth Internet of Things systems. The Mobile Edge Computing (MEC) technology is adopt to the eHealth Internet of Things systems. We model for the energy utility of WBANs based on the transmission power level. The optimization problem is solved by the reinforcement learning algorithm. It is indicated by the simulation results that the network performance metrics in the system energy efficiency, delivery ratio, and throughput are improved by the proposed protocol.

2. Related Works

The transmission power control (TPC) protocol of the wireless body area networks (WBANs) can substantially moderate the system energy consumption. The TPC protocol attempts to adapt the transmission power level according to the real time link condition to improve energy efficiency in WBANs. To solve the energy limitation problem, numerous transmission power control mechanisms were proposed [6-18].

In [6], the authors used the information of periodic body movements in dynamic channel conditions. An accelerator-assisted transmission power control mechanism (AA-TPC) was proposed. It deployed additional local accelerators to adjust transmission power of nodes, and selected the best time to transmit data utilizing the relationship between channel states and accelerator signals. However, this protocol only utilized wrist-worn sensor devices. In [7], the authors designed a gait-cycle-driven transmission power control (G-TPC) mechanism applying the channel periodic fluctuation during walking. This mechanism employed a long-term average channel gain and a time-dependent channel model. The authors of [8] designed a relay-aided transmission power control protocol, which aims to alleviate the burden of

relaying of relay nodes on a premise of reliable transmission. In this protocol, transmission strategy is changed between relay aided transmission and direct transmission.

The authors of [9] proposed a PID method based on control theory to dynamically adjust transmission power. In [10], the authors proposed a time correlation model transmission power control scheme (TCM-TPC), which used a temporal correlation model to describe the channel condition. In [18], an optimized hybrid technique of Genetic algorithm with BAT algorithm (GABAT) is presented to achieve QoS metrics. The proposed scheme gives higher priority to the emergency packets than the normal packets taking into account the dynamic link constraint. In [19-20], the authors applied the learning-based algorithm to the botnet attack detection. The authors of [21] proposed a statistic Quality of Service (QoS) based fixed power allocation method to take the statistic QoS of the popular files into consideration and improve the energy efficiency.

The above-mentioned related works proposed divers power control protocols for WBANs. These mechanisms are based on the cloud network architecture. With the development of mobile edge computing, the edge-cloud architecture is applied to many different kinds of network systems. In the healthcare monitoring system based on edge computing, power control schemes are needed for the data transmission not only between the sensors and coordinator, but also between the coordinator and MEC servers. Thus two-stage power control scheme is required. Also, the tradeoff between energy consumption and link reliability need to be considered. In the real application, the link state changes rapidly due to mobility.

3. System Model

The traditional eHealth system employed WBANs is designed generally based on the cloud network architecture. Data is collected by sensor node and transmit to the coordinator, then relay to the remote cloud server [22]. Time delay is longer than local process. However, some emergencies such as heart attack need to be responded quickly. The data transmission delay between the user and remote cloud server can be quite long, which may procrastinate the rescue [23]. The mobile edge servers (MESs) deploy near nodes of WBANs in the edge-cloud network architecture. The mobile edge servers can reply the crucial request promptly and then contact the nearby ambulance if necessary. Compared with the traditional cloud-based architecture, it has many advantages, such as shorter latency, lower energy and higher Quality of Service (QoS).

Fig. 1 indicates an example of the edge-cloud Internet of Things eHealth systems. The system has three tiers: the cloud layer, the edge layer, and the user layer. The user layer consists of many WBANs. Usually, for each user, several body sensor nodes attach on or implanted in the user's body, collecting physiological information. The heart sensor, for instance, monitors heartbeat. The activity sensors is positioned on the human body to detect the posture and movement, like lying, sitting, walking, and running [24].

In one WBAN, the coordinator (or personal devices like smartphone) gathers the sensing data. It also has the function of data storage, process, fusion, analysis and display. The coordinator relays data to the wireless access point (AP). In some systems, there is no AP. The access point or coordinator may link to the Internet. The data is relayed the cloud data center. The reference nodes (RNs) are equipped by GPS, or preprogrammed with their locations, using for localization.

The second layer is edge layer, which is composed of the mobile edge computing (MEC) server and the wireless access point (AP). The coordinator transmits data and requirement to the MEC server and AP. Edge Computing is a technology aimed at offloading mobile devices

to nearby data process center, which can provide short-latency services. The MEC servers perform tasks in order of priority and data fusion. Then the MEC servers relay data to the cloud system and database. On account of data fusion, the data volume from sensor to cloud is significantly lessened. The utility efficiency of network communication and computing capability can be increased by applying edge computing architecture. Therefore, adopting edge computing architecture can improve the system Quality of Service (QoS) and Quality of Experience (QoE).
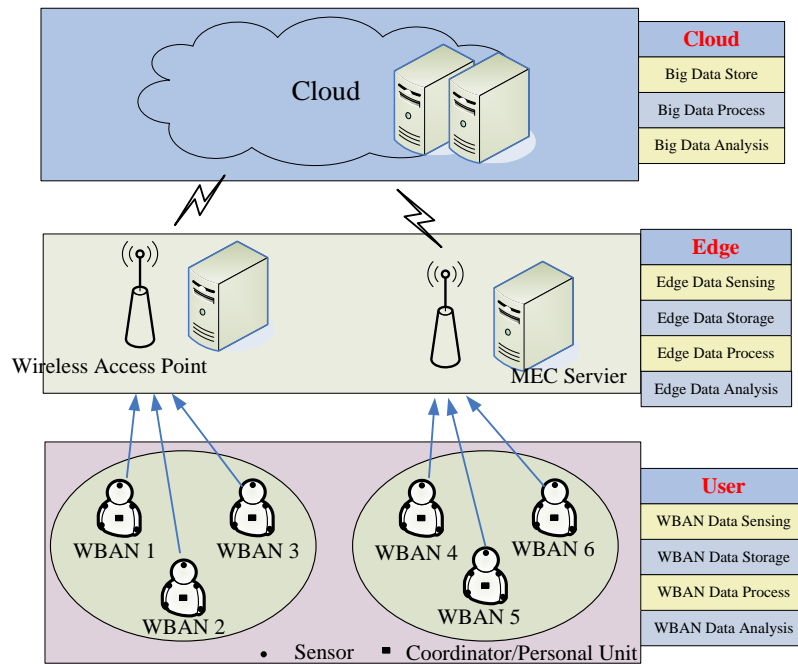

Fig. 1. Example of edge-cloud eHealth systems.

The third layer is cloud. Data of patients is stored in the database. The doctors or expert systems examine patients, then provide medical services. Through the Internet, remote experts from different places can conduct a consultation or collaboration. If emergency happens, such as heart attack, eHealth system will send the information to the nearest ambulance [25]. Also, the patients' health information can be stored in the cloud. Data analysis and statistics can be done in the long term.

The communication of WBANs is supported by IEEE 802.15.6 protocol stacks [26]. This standard specifies the network protocols of WBAN physical (PHY) layer and medium access control (MAC) layer. There are three kinds of physical layers: narrowband (NB) PHY, ultra wideband (UWB) PHY, and human body communication (HBC) PHY. The specification for the MAC layer divides the channel into superframes, which are held at both ends by the beacons in the beacon-enable mode. The superframe is composed by nine periods, including beacon, exclusive access phase 1 (EAP1), random access phase 1 (RAP1), managed access phase 1 (MAP1), exclusive access phase 2 (EAP2), random access phase 2 (RAP2), managed access phase 2 (MAP2), beacon 2, contention access phase (CAP). MAC layer controls user data to access wireless media. It is an essential part of how the network resources, such as channels and time slots, are used by the nodes.

4. Energy-Efficient Transmission Power Control Scheme

This section describes a power control algorithm, which is based on edge-cloud network architecture, using reinforce learning mechanism. It can effectively control the transmission power of sensor nodes and coordinator, thus improving the energy efficiency of the system without accurate channel state information.

Fig. 2 shows the communication architecture of body area networks in edge-cloud architecture. Health data are collected by the body sensors. Data is transmitted data from the body sensors to the BAN coordinator (or personal devices such as smartphone or PDA) after confusion. The coordinator transmit data to the access point.
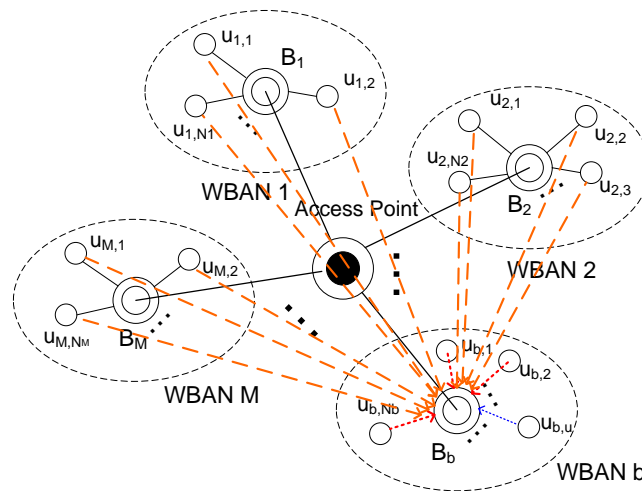

Fig. 2. Communication architecture of WBANs.

To model the system, we consider that the coverage area of mobile edge server is composed of several body area networks. It can be denoted as: $B=\{B_0, B_1,...B_M\}$, in which B_0 stands for the mobile edge server; $B_1, B_2,...B_M$ stand for the coordinator of WBAN. U_b stands for the set of bth wireless body area network. $U_b=\{u_{b,1}, u_{b,2},... u_{b,Nb}\}$. Specifically, $u_{b,\mu}$ is the μth sensor node in the bth WBAN. $p_{b,\mu}$ stands for the transmission power of $u_{b,\mu}$. $p^c_{b,\mu}$ stands for the constant power consumption of $u_{b,\mu}$. $SINR_{b,\mu}$ stands for the signal interference noise ratio of B_b, as shown in (1), in which $I_{b,\mu}$ is the interference signal, as shown in (2). $R_{b,\mu}$ is the throughput of the network, which is shown in (3). $g^b_{b,\mu}$ stands for the channel gain of $u_{b,\mu}$ and B_b.

$$SINR_{b,\mu} = \frac{p_{b,\mu}g^b_{b,\mu}}{I_{b,\mu}+Wn_0} \tag{1}$$

$$I_{b,\mu} = \sum_{\substack{i\in U_b \\ i\neq \mu}} p_{b,i}g^b_{b,i} + \sum_{\substack{m=0 \\ m\neq b}}^{M}\sum_{n\in U_m} p_{m,n}g^b_{m,n} \tag{2}$$

$$R_{b,\mu} = W\log_2(1+\frac{p_{b,\mu}g^b_{b,\mu}}{I_{b,\mu}+Wn_0}) \tag{3}$$

Wn_0 stands for the power of Gaussian white noise, which power spectral density is n_0. The throughput of unit transmission energy denote the transmission energy efficiency. $\eta_{b,\mu}$ stands for the energy utility of $u_{b,\mu}$.

$$\eta_{b,\mu} = \frac{R_{b,\mu}}{P_{b,\mu}} = \frac{W \log_2(1+\frac{p_{b,\mu}g_{b,\mu}^b}{I_{b,\mu}+Wn_0})}{p_{b,\mu}+P_{b,\mu}^c} \tag{4}$$

where $P_{b,\mu}$ stands for the energy consumption of $u_{b,\mu}$. We define sum of the energy utility of each sensor node in WBAN b as its energy utility.

$$\eta_b = \sum_{\mu \in U_b} \eta_{b,\mu} = \sum_{\mu \in U_b} \frac{W \log_2(1+\frac{p_{b,\mu}g_{b,\mu}^b}{I_{b,\mu}+Wn_0})}{p_{b,\mu}+P_{b,\mu}^c} \tag{5}$$

The optimization mathematical model of energy utility is:

$$\max_{p_{b,\mu}} \sum_{\mu \in U_b} \frac{W \log_2(1+\frac{p_{b,\mu}g_{b,\mu}^b}{I_{b,\mu}+Wn_0})}{p_{b,\mu}+P_{b,\mu}^c} \tag{6}$$

$$\text{s.t. } SINR_{b,\mu} \geq SINR_{b,\mu}^{\min}, \forall b \in B, \forall \mu \in U_b$$

$$p_{b,\mu} \leq p_{b,\mu}^{\max}, \forall b \in B, \forall \mu \in U_b$$

where $SINR^{min}_{b,\mu}$ stands for the minimum signal interference noise ratio (SINR) for signal receiving. $p^{max}_{b,\mu}$ stands for the maximum transmission power level of sensor node $u_{b,\mu}$. The two constraint conditions should be satisfied in the system.

The Q-Learning algorithm is a kind of reinforcement learning algorithm. It can be used to find the solution to the above optimization problem. In Q-learning, in certain system state, to maximize the objective function, system choose action. The Q-learning is used to allocate the transmission of both sensor nodes and the coordinator of WBANs. In Q-learning algorithm, there are three elements: status (state space) (S), action space (A), and rewards (R). It is shown in **Fig. 3** that the procedure of agent interaction with environment.

The entity may be in many different states. All these possible states compose the state space set. In a certain state, the entity can take various actions, which form the set of action space. The reward function is defined as the system reward value brought by executing the action a under the entity state s. $Q(s, a)$ is the "state-action" function. $P_b=[p_{b,1}, p_{b,2}, \dots p_{b,Nb}]$ is the vector of transmission power of each sensor. The adjusting of transmission power to P_b indicates to $a_b=P_b$.

By using (1) and $SINR^{min}_{b,\mu}$, it can be obtained that the range of transmission power of sensor node $u_{b,\mu}$ is $p^{max}_{b,\mu} \leq p_{b,\mu} \leq p^{max}_{b,\mu}$. In this range, there are $d_{b,\mu}$ transmission power levels. The set of transmission power levels of $u_{b,\mu}$ is expressed as, $\rho_{b,\mu} = \{p_{b,\mu}^1, p_{b,\mu}^2, ..., p_{b,\mu}^{d_{b,\mu}}\}$. The power control action space of the coordinator can be indicated as $A_b = (\rho_{b,1} \times \rho_{b,1} \times \cdots \times \rho_{b,N_b})$.

We define s_b as the state of the agent, where $s_b=[I_b\, l_b]$. The coordinator's state space set S (i.e $\forall s_b \in S$) include all the possible states. $I_b = [I_{b,1}, I_{b,2}, \dots I_{b,Nb}]$, which is the interference vector. l_b is the coefficient vector, where $l_b = [\lambda_{b,1}, \lambda_{b,2}, \dots \lambda_{b,Nb}]$. It is applied to estimate the relationship between $SINR^{min}_{b,\mu}$ and the actual $SINR_{b,\mu}$ as follows:

$$\lambda_{b,\mu} = \begin{cases} 1, SINR_{b,\mu} \geq SINR_{b,\mu}^{\min} \\ 0, \text{others} \end{cases} \tag{7}$$

The interference and *SINR* can be measured by the mobile edge server (MES).

   To maximize the energy efficiency (utility) of wireless body area network, we define the reward function as the energy efficiency (utility). The reward function of taking action a_b in the state S_b can be expressed as follows:

$$\Re_b(s_b, a_b) = \sum_{\mu \in U_b} \frac{\lambda_{b,\mu} W \log_2(1 + SINR_{b,\mu})}{p_{b,\mu} + P_{b,\mu}^c} \tag{8}$$

   In the Q-learning algorithm, the WBAN coordinator dynamically attunes the data transmission power of every sensor node based on the interference and *SINR*, so as to optimize the energy efficiency (reward function).
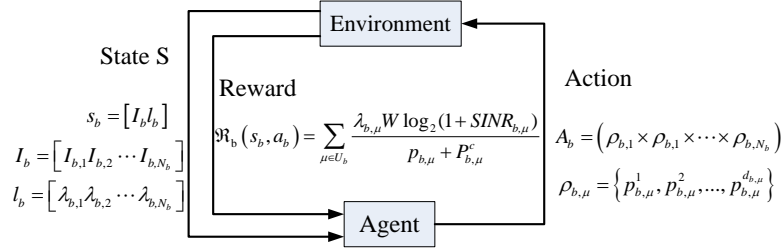

Fig. 3. Procedure of agent reinforcement learning.

   The interaction with the environment can be modeled as $(s_b^t, a_b^t, R_b^{t+1}, s_b^{t+1})$, where t is the discrete time counter. It represents that the node in state s_b^t will get a reward R_b^{t+1} by taking action d_b^t, and then the state will change to s_b^{t+1}. It means that when the coefficient vector of interference and SINR are I_b^t and l_b^t, respectively, after the agent alters the transmission power level of each node to $a_b^t=[p_{b,1}^t, p_{b,2}^t, \dots p_{b,Nb}^t]$, the total energy efficiency of WBAN b is $\eta_b = R_b^{t+1}(s_b^t, a_b^t)$. Since the new transmission power vector will result in the change of interference and SINR, the system will turn to a new state $s_b^{t+1}=[I_b^{t+1}, l_b^{t+1}]$. The Q-function related to state-action and power strategies π can be expressed as follows:

$$Q_\gamma^\pi(s_b, a_b) = E_{\pi_p}\left[\left(\sum_{t=0}^{+\infty} \gamma^t R_b^{t+1} \mid s_b^0 = s_b, a_b^0 = a_b\right)\right] \tag{9}$$

where γ is the discount factor. The optimal transmission power strategy π^* is obtained by maximizing the Q-function. Another expression of Q-function can be obtained by incremental summation:

$$Q_{t+1}^\pi(s_b, a_b) = Q_t^\pi(s_b, a_b) + \alpha\left(\Re_b(s_b, a_b) + \gamma \max_{a_b'}\left(Q_t^\pi(s_b', a_b')\right) - Q_t^\pi(s_b, a_b)\right) \tag{10}$$

where α is the learning efficiency. According to (8) and (10), the Q-function is updated to obtain the optimal transmission power.

   Fig. 3 indicate the process of agent reinforcement learning for transmission power control. It shows the interaction between the agent and environment. The environment is modeled by the interference vector I_b ($I_b=[I_{b,1}, I_{b,2}, \dots I_{b,Nb}]$) and the coefficient vector l_b ($l_b=[\lambda_{b,1}, \lambda_{b,2}, \dots \lambda_{b,Nb}]$). In a certain environment, the agent takes an action by arranging the transmission power

of sensor nodes and the coordinator in WBANs. By taking action, the agent can calculate the reward function of the current environment and action. The action also changes the system state and environment. To optimize the Q-value, the agent adjust the transmission power allocation. On this basis, the power control algorithm is designed. According to the execution effect and state change of power allocation strategy, the current optimal strategy is learned to realize centralized power control.

Fig. 4 shows the flow chart of power control mechanism base on reinforcement learning.

Step 1: Initialization. The channel model (bandwidth, gain, etc.) is set to the initial information. The number of learning times is set to M. The power control action space and state space are set to A and S, respectively.

Step 2: The state of agent is set to s.

Step 3: The transmission power of the sensor nodes are allocated by ε-greedy algorithm. It explores the action in A by probability ε, and applies the action by probability 1-ε. When exploring, it chooses an action randomly within space A. When applying, it selects the action maximizing the current Q-value.

Step 4: The sensor nodes transmit data using the allocated power. During transmission, the interference signals are detected. The reward function is calculated according to formula (8). The Q-function is calculated according to formula (10).

Step 5: According to the new Q-function, the transmission power levels are allocated to maximize the Q-function value. The state is updated.

Step 6: If the counter is smaller than M, the Q value is updated. The program will go to step 3 for next iteration of the loop. If the counter equals to M, the loop terminates.

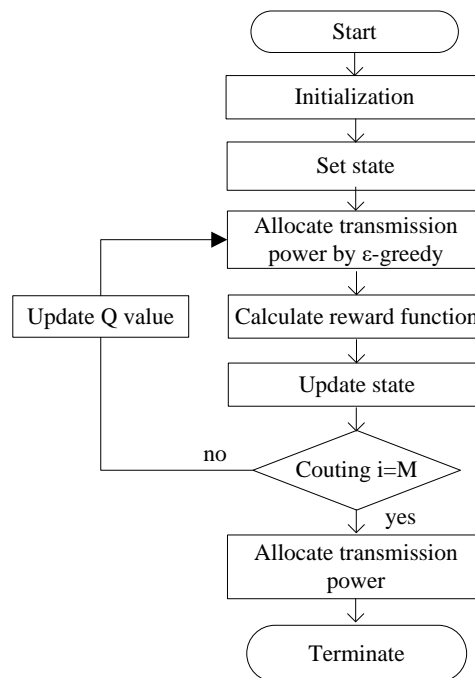Step 7: The transmission power of the sensor nodes are allocated according to the learning results.

Fig. 4. Procedure of agent reinforcement learning.

In this algorithm, the initial information determines the work of the agent. The system is in a new state because the new transmission power vector will lead to interference and signal interference to noise ratio change. The Q-functions related to state action pairs and power strategies can be expressed as the energy efficiency of data transmission. In the next stage, the transmission power is regulated referring to the system reward. After every cycle, the algorithm calculate the value of Q-function. Applying Q-learning algorithm, the optimal transmission power can be acquired.

4. Performance Evaluation

In WBANs, the applications are extremely heterogeneous. The traffic data rates vary greatly vary. Applications transmitting simple data need a few kbit/s rate. Video streams requires several Mbit/s rate. The transmission data rate may be significantly higher in a particular time period, which is called a burst. **Table 1** indicates the data rates requirement for some different applications. They are computed by the expected accuracy, the range, and the sampling rate [2-4]. Overall the user data levels cannot be seen to be high. Yet if the user ware many body sensors, for example, ECG, temperature, EEG, dozen motion sensors, the system assemble data rate can be several Mbps. It is higher than the commonly used radios.

Table 1. Data rate requirement of healthcare applications

Application	Data rate	Bandwidth
Blood saturation	16 bps	0-1 Hz
ECG (12 leads)	288 kbps	100-1000
ECG (6 leads)	71 kbps	100-500
EEG (12 leads)	43.2 kbps	0-150
EMG	320 kbps	0-10,000
Temperature	120 bps	0-1
Motion sensor	35 kpbs	0-500
Glucose monitoring	1600 bps	0-50
Cochlear implant	100 kbps	-
Voice	50-100 kbps	-
Audio	1 Mbps	-
Artificial retina	50-700 kbps	-

WBAN body sensor nodes of can be classified according to the data's priority. There are four categories: non-constrained traffic class (NTC), delay traffic class (DTC), reliability traffic class (RTC), critical traffic class (CTC). The non-constrained traffic class sensor nodes gather non-constrained data packets (NDP), which can tolerate certain degree of losses and have loose time delay requirement, such as Blood Pressure (BP) and temperature [12]. The delay traffic class sensor nodes collect delay data packets (DDP). These packets can tolerate some losses but have delay time-constraint, such as telemedicine video imaging. The reliability traffic class sensor nodes collect reliability data packets (RDP). This kind of data has strict requirement of packet loss ratio, but has no time delay constraint, such as heart rate (HR) and respiratory rate (RR). The critical traffic class sensor nodes collect critical data packets (CDP). These data packets have strict requirement of maximum loss and time-delay, like Electroencephalogram (EEG) and Electrocardiograph (ECG). **Table 2** shows the four classes and priorities.

Table 2. Sensor classes and priorities

Priority	Classification of Sensors	Traffic Class
0	Critical Data Packets (CDP)	Critical
1	Reliability Data Packets (RDP)	Reliability
2	Delay Data Packets (DDP).	Delay
3	Non-constrained Data Packets (NDP)	Non-constrained

The system energy efficiency, delivery ratio, and throughput of proposed reinforcement power control protocol is compared with above-mentioned TCM-TPC [10] and GABAT [18]. The system energy efficiency E_e is defined as the average energy consumption by each sensor.

$$E_e = \frac{E_t}{\sum_{i=1}^{n} T_s(i)} \tag{11}$$

where E_t stands for the system energy consumption, $T_s(i)$ stands for the ith successful transmission. The delivery ratio D_r is defined as the percentage of the transmitted data frames that are successfully received.

$$D_r = \frac{\sum_{i=1}^{n} T_s(i)}{\sum_{i=1}^{n} T(i)} \tag{12}$$

where $T(i)$ stands for the ith transmission. The system throughput is calculated by the total successful transmitted packets and the total time T:

$$Thp = \frac{\sum_{i=1}^{n} T_s(i)}{T} \tag{13}$$

We simulate the performance of proposed protocol in different scenarios with various node densities and environments. There are 100 sensors and 10 Coordinators in a 10m×10m square field. In one cycle, the first data frame is generated randomly. The application traffic is set to constant distribution with data rate 2 kbps. The channel rate is 250 kbps. The frequency band is in the range of 2400-2483.5 MHz. The MAC frame header size is 27 bytes. The payload length of data frame is 10 bytes.

The WBAN Coordinators stay static for 10 seconds after the simulation starts. Then the coordinators begin to move at a speed [0, 1.5] (m/s) randomly. The sensor nodes mobility velocity is a random variable uniformly distributed in [0, 0.5] (m/s), which emulates the walking speed of men. **Table 3** shows the energy consumption parameters applied in the simulation. The simulation program runs for 10 times. Every cycle of simulation executes independently. The average values are calculated according to the results.

Fig. 5 display the simulation results of the system energy efficiency. For WBAN protocols, due to power limitation, energy efficiency is one of the most important metrics. The lifetime of WBANs can be extended by lessening the power consumption of sensor nodes. The figure indicates that by applying the proposed Q-learning power control mechanism, the system energy efficiency substantially increases. The network architecture of edge-cloud diminishes

the amount of data delay, thus reduces the energy consumption. Moreover, the Q-learning power control protocol minimizes the energy consumption while maintains the link quality.

Table 3. Parameters of energy consumption

State	Energy Consumption
Transmission data	36.5 mW
Reception data	41.4 mW
Idle status	41.4 mW
Sleep status	42 μW

In the topology control mechanism, the transmission energy of sensor nodes and coordinators can be diminished, which not only reduces the energy consumption of the system, but also lower the probability of beacon and data collision. If collision happens, retransmission is needed, which cause abundant power waste. Lower transmission power shortens the transmission range. Therefore, the overlap area of transmission range can be reduced. The proposed scheme can moderate the power consumption and extend the WBAN lifetime.
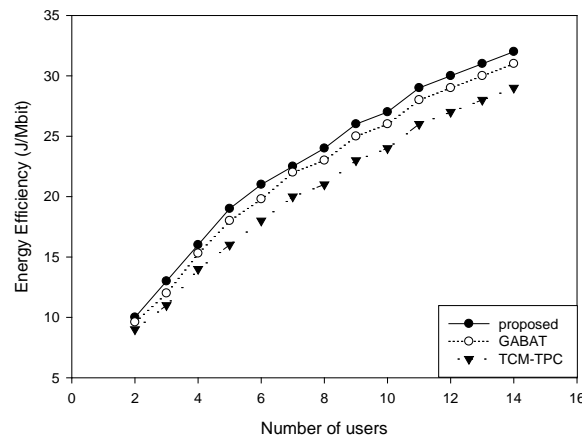

Fig. 5. Simulation results of energy efficiency.

Fig. 6 demonstrates the delivery ratio of WBAN, which is also an important measurement. With the increase of user numbers, the delivery ratio of WBAN gradually degrades. The figure illustrates that the delivery ratio using Q-learning transmission power control protocol is higher than TCM-TPC and GABAT. The proposed protocol can reduce the collisions of beacon or data frames. In the meanwhile, this scheme enhances the link reliability. Because of the smaller collision rate, the successful transmission rate enlarges.

Fig. 7 demonstrates the network throughput of WBAN. The system throughput reflects the WBANs data transmission capability, which is another critical measurement. With the increase of user numbers, the eHealth Internet of Things system throughput also increases. The figure illustrates that the system throughput using Q-learning power control protocol is higher than TCM-TPC and GABAT. Since the proposed protocol can reduce the collisions of beacon or data frames, thus decreasing the retransmission time. System throughput increases as there are less collision and the latency is shorter. Since the retransmission rate is lower, the data delay is shorter. The utilization of time slots is improved. It also shortens the average time delay.
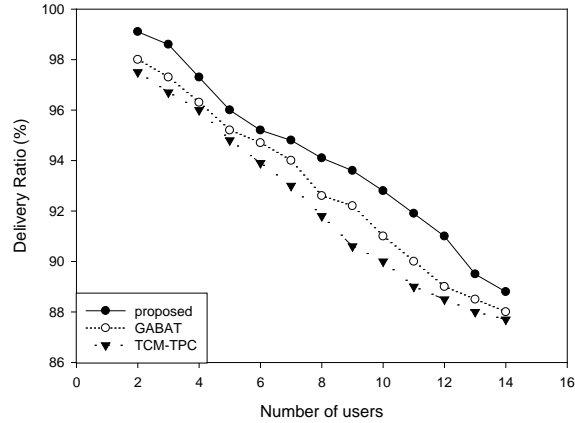
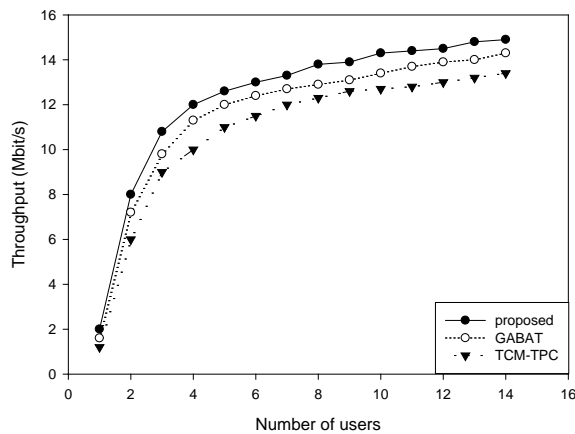**Fig. 6.** Simulation results of delivery ratio.

Fig. 7. Simulation results of throughput.

6. Conclusions

Wireless body area networks support the development of the Internet of Things eHealth systems. Because of the limitation of battery capacity, the energy efficiency of sensor nodes is an essential issue. Moreover, due to the mobility of sensor nodes, the rapidly changing link state, and human body shadowing, communication reliability has tradeoff of energy consumption. For shorter task process delay, this paper apply the edge-cloud network architecture to the eHealth systems. A reinforcement learning based transmission power control algorithm is proposed. The energy utility model and optimization problem is solved by learning algorithm. By evaluating its performance, the proposed protocol demonstrates enhancement of the system energy efficiency, delivery ratio, and throughput.

Acknowledgement

References

[1]   H. Zhang, J. Li, B. Wen, Y. Xun, and J. Liu, "Connecting intelligent things in smart hospitals using NB-IoT," *IEEE Internet Things Journal*, vol. 5, no. 3, pp. 1550–1560, Jun. 2018. Article (CrossRef Link)

[2]   S. M. R. Islam, D. Kwak, M. D. H. Kabir, M. Hossain, and K. S. Kwak, "The Internet of things for health care: a comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015. Article (CrossRef Link)

[3]   S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, and A. Jamalipour, "Wireless body area networks: a survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1658–1686, 2014. Article (CrossRef Link)

[4]   R. A. Khan and A. K. Pathan, "The state-of-the-art wireless body area sensor networks: a survey," *International Journal of Distributed Sensor Networks*, vol. 14, no. 4, pp. 1–23, 2018. Article (CrossRef Link)

[5]   IEEE Standard for Local and Metropolitan Area Networks-Part 15.6: Wireless Body Area Networks, IEEE Standard 802.15.6, 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6161600.

[6]   W.-L. Zang, S.-L. Zhang, and Y. Li, "An accelerometer-assisted transmission power control solution for energy-efficient communications in WBAN," *IEEE J. Sel. Areas. Commun.*, vol. 34, no. 12, pp. 3427–3437, Dec. 2016. Article (CrossRef Link)

[7]   W. Zang and Y. Li, "Gait-Cycle-Driven Transmission Power Control Scheme for a Wireless Body Area Network," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 3, pp. 697-706, 2018. Article (CrossRef Link)

[8]   Y. Zhang, B. Zhang, "A Relay-Aided Transmission Power Control Method in Wireless Body Area Networks," *IEEE Access*, no. 5, pp. 8408-8418, 2017. Article (CrossRef Link)

[9]   H. Moosavi, F. M. Bui, "Optimal Relay Selection and Power Control with Quality-of-Service Provisioning in Wireless Body Area Networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 8, pp. 5497-5510, 2016. Article (CrossRef Link)

[10]   S. Archasantisuk, T. Aoyagi, M. Kim, J.–I. Takada, "Temporal Correlation Model-Based Transmission Power Control in Wireless Body Area Network," *IET Wireless Sensor Systems*, vol. 8, no. 5, pp. 191-199, 2018. Article (CrossRef Link)

[11]   C. Zhang, Y. Wang, Y. Liang, et al., "An energy-efficient MAC protocol for medical emergency monitoring body sensor networks," *Sensors*, vol. 16, no. 385, pp. 1–19, 2016. Article (CrossRef Link)

[12]   S. Pandit, K. Sarker, M. A. Razzaque, et al., "An energy-efficient multiconstrained QoS aware MAC protocol for body sensor networks," *Multimedia Tools and Applications*, vol. 74, no. 14, pp. 5353–5374, 2015. Article (CrossRef Link)

[13]   H. Habibzadeh, K. Dinesh, O. R. Shishvan, A. Boggio-Dandry, G. Sharma, and T. Soyata, "A Survey of Healthcare Internet of Things (HIoT): A Clinical Perspective," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 53-71, 2020. Article (CrossRef Link)

[14]   A. Argyriou, A. C. Breva, and M. Aoun, "Optimizing data forwarding from body area networks in the presence of body shadowing with dual wireless technology nodes," *IEEE Trans. Mobile Comput.*, vol. 14, no. 3, pp. 632–645, Mar. 2015. Article (CrossRef Link)

[15]   S. Nepal, A. Pudasaini, J. Pyun, S. Hwang, C. G. Lee, and S. Shin, "A new MAC protocol for emergency handling in wireless body area networks," in *Proc. of IEEE Ubiquitous and Future Networks (ICUFN)*, Vienna, Austria, pp. 588–590, 2016. Article (CrossRef Link)

[16]   F. Masud, A.H. Abdullah, A. Altameem, et al., "Traffic class prioritization-based slotted-CSMA/CA for IEEE 802.15.4 MAC in intra-WBANs," *Sensors*, vol. 19, no. 3, pp. 466–488, 2019. Article (CrossRef Link)

[17]   Z. Liu, B. Liu, and C. W. Chen, "Joint Power-Rate-Slot Resource Allocation in Energy Harvesting-Powered Wireless Body Area Networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12152-12164, 2018. Article (CrossRef Link)

[18] R. Goyal, R. B. Patel, H. S. Bhaduria, and D. Prasad, "An Energy Efficient QoS Supported Optimized Transmission Rate Technique in WBANs," *Wireless Personal Communications*, vol. 117, no. 1, pp. 235-260, 2021. [Article (CrossRef Link)](#)

[19] S. I. Popoola, B. Adebisi, M. Hammoudeh, G. Gui, and H. Gacanin, "Hybrid Deep Learning for Botnet Attack Detection in the Internet-of-Things Networks," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4944-4956, 2021. [Article (CrossRef Link)](#)

[20] S. I. Popoola, B. Adebisi, M. Hammoudeh, H. Gacanin, and G. Gui, "Stacked recurrent neural network for botnet detection in smart homes," *Computers & Electrical Engineering*, vol. 92, p. 107039, 2021. [Article (CrossRef Link)](#)

[21] Y. Yin, M. Liu, G. Gui, H. Gacanin, H. Sari, and F. Adachi, "Cross-Layer Resource Allocation for UAV-Assisted Wireless Caching Networks with NOMA," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3428-3438, 2021. [Article (CrossRef Link)](#)

[22] F. Ullah, A. H. Abdullah, O. Kaiwartya, S. Kunar, and M. M. Arshad, "Medium access control (MAC) for wireless wody area network (WBAN): superframe structure, multiple access technique, taxonomy, and challenges," *Human-centric Computing and Information Sciences*, vol. 7, no. 1, pp.1–34, 2017. [Article (CrossRef Link)](#)

[23] A. K. Jacob and L. Jacob, "A green media access method for IEEE 802.15.6 wireless body area network," *Journal of Medical Systems*, vol. 41, no. 179, pp. 1–14, 2017. [Article (CrossRef Link)](#)

[24] V. Varga, G. Vakulya, A. Sample, and T.R. Gross, "Enabling interactive infrastructure with body channel communication," *ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, pp. 169:1–29, 2017. [Article (CrossRef Link)](#)

[25] G. Yang, L. Xie, M. Mantysalo, et al., "A health-IoT platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2180–2191, Nov. 2014. [Article (CrossRef Link)](#)

[26] S. Tennina, M. Santos, A. Mesodiakaki, et al., "WSN4QoL: WSNs for remote patient monitoring in e-health applications," in *Proc. of IEEE International Conference on Communications (ICC),* Kuala Lumpur, Malaysia, pp. 1–6, 2016. [Article (CrossRef Link)](#)

HAORU SU received the B.S. degrees in Computer Science and Engineering and M.S. in Computer Applied Technology from Northeastern University, China. She received the Ph.D. degree from the Department of Electrical and Computer Engineering in Korea University, South Korea. Currently, she is working in Faculty of Information at Beijing University, China. She was also a Visiting Scholar in the Broadband Communications Research Lab, Department of Electrical and Computer Engineering, University of Waterloo, Canada in 2019. Her research interests include Multiple Access Control protocol design and performance evaluation for Wireless Body Area Networks, beacon scheduling mechanisms for the IEEE 802.15.4/ZigBee Wireless Networks, and reader collision problem in dense RFID systems.

XIAOMING YUAN received her B.E. degree in Electronics and Information Engineering from Henan Polytechnic University, Jiaozuo, China, in 2012. She received the Ph.D. degree in Communication and Information System from Xidian University, Xi'an, China, in 2018. From September 2016 to August 2017, she was a Visiting Scholar with the Broadband Communications Research (BBCR) Group, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. She is now an assistant professor of Qinhuangdao Branch Campus, Northeastern University, Shenyang, China, from 2018. Her research interests include cloud/edge computing, medium access control and performance analysis for wireless body area networks and the Internet of Things.

YUJIE TANG received the B.E. degree in communications engineering from Lanzhou Jiaotong University, Lanzhou, China, the M.Sc. degree in information and communications engineering from Harbin Institute of Technology, Harbin, China, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada. She was a Postdoctoral Fellow with the Broadband Communications Research (BBCR) Group, University of Waterloo. She is currently an Assistant Professor with the School of Computer Science and Technology, Algoma University, Sault Ste. Marie, ON, Canada. Her research interests include internet of vehicles, software-defined networks, machine learning, cognitive radio networks, cooperative networks, and resource management in heterogeneous networks.

RUI TIAN received his Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences. He currently works in the Beijing Engineering Research Center for IoT Software and Systems, Information Department, Beijing University of Technology. His research interests include distributed computing algorithm, internet of things and vehicular networking algorithms.

ENCHANG SUN received the Ph.D. degree in Information and Telecommunications Engineering from Xidian University, China. He currently works in the Faculty of Information Technology, Beijing University of Technology. His research interests are communication and information theory with special emphasis on cognitive and collaborative communications, Internet of Things and data processing, distributed machine learning, and blockchain.
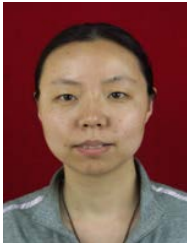
HAIRONG YAN received the Ph.D. degree in electrical and information engineering from Xi'an Jiaotong University, Xi'an, China, in 2006. She is currently an Associate Professor with the Faculty of Information Technology, Beijing University of Technology. She was a Researcher at Jonkoping University, Jonkoping, Sweden, in 2007, and was a Visiting Scholar at the University of Tsukuba, Tsukuba, Japan, in 2003. Her research interests include wireless communication, wireless sensor networks, and Industrial IoT architecture.