Learning-Based Privacy-Aware Offloading for Healthcare IoT With Energy Harvesting

Minghui Min, *Student Member, IEEE*, Xiaoyue Wan, Liang Xiao, *Senior Member, IEEE*, Ye Chen, Minghua Xia, *Member, IEEE*, Di Wu, *Senior Member, IEEE*, and Huaiyu Dai, *Fellow, IEEE*

***Abstract*—Mobile edge computing helps healthcare Internet of Things (IoT) devices with energy harvesting provide satisfactory quality of experiences for computation intensive applications. We propose a reinforcement learning (RL)-based privacy-aware offloading scheme to help healthcare IoT devices protect both the user location privacy and the usage pattern privacy. More specifically, this scheme enables a healthcare IoT device to choose the offloading rate that improves the computation performance, protects user privacy, and saves the energy of the IoT device without being aware of the privacy leakage, IoT energy consumption, and edge computation model. This scheme uses transfer learning to reduce the random exploration at the initial learning process and applies a Dyna architecture that provides simulated offloading experiences to accelerate the learning process. A post-decision state learning method uses the known channel state model to further improve the offloading performance. We provide the performance bound of this scheme regarding the privacy level, the energy consumption, and the computation latency for three typical healthcare IoT offloading scenarios. Simulation results show that this scheme can reduce the computation latency, save the energy consumption, and improve the privacy level of the healthcare IoT device compared with the benchmark scheme.**

***Index Terms*—Energy harvesting (EH), healthcare systems, Internet of Things (IoT), mobile edge offloading, privacy, reinforcement learning (RL).**

Manuscript received May 22, 2018; revised July 15, 2018; accepted October 6, 2018. Date of publication October 15, 2018; date of current version June 19, 2019. The work of L. Xiao was supported in part by the National Natural Science Foundation of China under Grant 61671396 and Grant 91638204 and in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University under Grant 2018D08. The work of M. Xia was supported by the National Natural Science Foundation of China under Grant 61671488. The work of D. Wu was supported in part by the National Natural Science Foundation of China under Grant 61572538, in part by the Guangdong Special Support Program under Grant 2017TX04X148, and in part by the Fundamental Research Funds for the Central Universities under Grant 17LGJC23. The work of H. Dai was supported by the U.S. National Science Foundation under Grants ECCS-1444009 and CNS-1824518. *(Corresponding author: Liang Xiao.)*

M. Min, X. Wan, L. Xiao, and Y. Chen are with the Department of Communication Engineering, Xiamen University, Xiamen 361005, China, and also with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China (e-mail: lxiao@xmu.edu.cn).

M. Xia is with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China (e-mail: xiamingh@mail.sysu.edu.cn).

D. Wu is with the School of Data and Computer Science, Sun Yat-Sen University, Guangzhou 510006, China, and also with the Guangdong Province Key Laboratory of Big Data Analysis and Processing, Guangzhou 510006, China (e-mail: wudi27@mail.sysu.edu.cn).

H. Dai is with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695 USA (e-mail: huaiyu_dai@ncsu.edu).

Digital Object Identifier 10.1109/JIOT.2018.2875926

I. Introduction

INTERNET of Things (IoT) devices in healthcare systems, such as remote health monitoring, fitness programs, chronic diseases and eldly care measure, and evaluate healthcare data, such as the blood pressure, body temperature, electrocardiograms, and oxygen saturation of users to provide healthcare reports and alarms [1]. Healthcare IoT devices can apply energy harvesting (EH) technique to use the energy from the environment, such as the ambient radio-frequency (RF) and the body motion to extend the battery life [2]. Mobile edge computing (MEC) also saves energy for healthcare IoT devices by processing the sensed healthcare data at the edge devices, such as the serving base stations, access point (AP), and laptops that have more computation and energy resources [3]–[9]. For instance, an edge device can help an IoT device evaluate the cardiac measurement and make healthcare diagnosis.

Healthcare IoT devices with EH have to resist eavesdroppers that analyze the sensing data via radio channels to reveal the user location and habits, such as the usage pattern privacy [1], [10], [11]. More specifically, the user location privacy can be inferred from the offloading data size, e.g., a user is very likely to stay in the outage locations or far away from the edge device if the IoT device locally computes all the sensing data under severe radio channel condition connecting to the edge device. An attacker can estimate the size of the sensing data newly generated and thus evaluate the usage pattern if the IoT device offloads all the sensing data to the edge device under good radio channel state. Therefore, the IoT device has to protect both the user location and usage pattern privacy in the mobile offloading.

Current steganography and homomorphic encryption techniques are not always applicable for healthcare IoT devices with limited computation resources during the edge computing [12], [13] and most existing mobile offloading schemes, such as [14]–[16] ignore user privacy. The seminar work on the privacy-aware mobile computing as presented in [17] allows mobile devices choose the offloading policy and formulates a constrained Markov decision process (CMDP) to ensure the prespecified privacy level for simplified offloading scenario with reduced computation latency and energy consumption. This scheme suffers from a slow learning speed and the offloading performance degrades in practical healthcare IoT devices with EH.

In this paper, we propose a privacy-aware offloading scheme to improve the privacy level, reduce the computation latency,

and save the energy consumption of healthcare IoT devices. In this scheme, the offloading rate and the local processing rate of an IoT device is chosen based on the current radio channel state, the size and priority of the healthcare sensing data or computation tasks, the estimated EH state and the battery level. For instance, more sensing data are offloaded to the edge device under good radio channel state. Otherwise, the IoT device locally processes the sensing data in the rare cases with narrow radio bandwidth to save the computation latency. Therefore, the proposed offloading scheme optimizes the offloading rate according to the radio channel state to improve the computation performance of the IoT device. This scheme analyzes the difference between the amount of the sensing data and the size of the offloading data under different channel power gains to avoid privacy leakage.

The optimal offloading policy depends on the accurate knowledge of the privacy leakage, the IoT energy consumption, and the edge computation model in each time slot, which is challenging to determine especially in a dynamic IoT system. As the future state observed by a healthcare IoT device is independent of the previous states for a given current state and offloading policy in a repeated offloading process, we have a Markov decision process (MDP) and thus a healthcare IoT device can apply reinforcement learning (RL) techniques, such as Q-learning to achieve the optimal offloading policy via trial-and-error without being aware of the underlying models [18]. We propose an RL-based privacy-aware offloading algorithm for a healthcare IoT device to choose the offloading and local computing policy. This offloading algorithm uses the model learning method that exploits the IoT offloading experiences to build a Dyna architecture and generate simulated experiences accordingly to update the value function of the RL technique. A post-decision state (PDS) method as investigated in [19] is also applied to use the known radio channel model to accelerate the learning process. A transfer learning method as developed in [20] is used to exploit the offloading experiences in similar scenarios for the initialization of the learning parameters and thus save the initial exploration in the offloading process.

We prove that the proposed scheme achieves the optimal offloading policy after long enough time slots in the dynamic game. The offloading performance bound is provided in terms of the privacy level, the total computation latency, and the energy consumption of the healthcare IoT device with EH. This offloading algorithm can improve the privacy level of the IoT device, which depends on the amount of the computation tasks. Both the computation latency and the energy consumption of the IoT device linearly increase with the size of the healthcare sensing data. Simulation results show that this scheme decreases the computation latency, saves the energy consumption, improves the privacy level, and increases the utility of the healthcare IoT device, compared with the benchmark scheme CMDP as proposed in [17].

The main contributions of this paper are summarized as follows.

1) We propose a privacy-aware offloading scheme for an EH powered healthcare IoT device to select the offloading rate and the local computation rate to process the healthcare sensing data. This scheme considers the current radio channel state, the size and the priority of the new healthcare sensing data, the estimated EH state, the battery level and the task computation history to decrease the computation latency, save the energy consumption of the IoT device, and improve the privacy level.
2) We propose a RL-based offloading algorithm for an IoT device to achieve the optimal offloading policy via trial-and-error without being aware of the privacy leakage, IoT energy consumption, and edge computation model. This algorithm uses the transfer learning technique, the PDS method and the Dyna architecture to accelerate the learning process of an IoT device.
3) We provide the performance bound of the RL-based offloading scheme in terms of the privacy level, the energy consumption, and the computation latency and prove its convergence to the optimal performance in the dynamic offloading process.

The remainder of this paper is organized as follows. We review the related work in Section II, and present the system model in Section III. We propose an RL-based privacy-aware offloading scheme for IoT devices in Section IV, and analyze its performance in Section V. We provide simulation results in Section VI and conclude this paper in Section VII.

II. Related Work

Edge computing helps IoT devices support computational-intensive and latency-sensitive applications with reduced energy consumption and computation latency. For instance, the binary offloading as proposed in [21] chooses the data transmission rate under stochastic wireless channel with a single edge to reduce the computation overhead for resource-constrained mobile devices. The partial offloading scheme as proposed in [22] uses the time-division and the orthogonal frequency-division multiple access to reduce the energy consumption under a latency constraint in a multiuser MEC network. The mobile offloading scheme as proposed in [14] uses the Lyapunov optimization to reduce the execution latency and the task failure rate for the case with a single known MEC server, assuming that both the transmission delay model and the local execution model are known.

EH is a promising technique to prolong the battery lifetime and provide satisfactory experience of IoT devices [2]. For instance, a renewable-powered MEC system as investigated in [15] combines the value iteration with the RL technique to improve the edge computing performance of the mobile device for delay-sensitive applications with intermittent and unpredictable renewable energy. The wireless powered multiuser MEC system as proposed in [16] jointly improves the AP beamforming and the user time allocation to save the AP energy consumption subject to the users' latency constraints.

Privacy is critical for edge computing of IoT applications [10], [11], [23], [24]. For instance, the IoT computation offloading scheme as proposed in [12] and [13] uses steganography and homomorphic encryption to hide the image privacy and save energy and protect privacy, respectively.

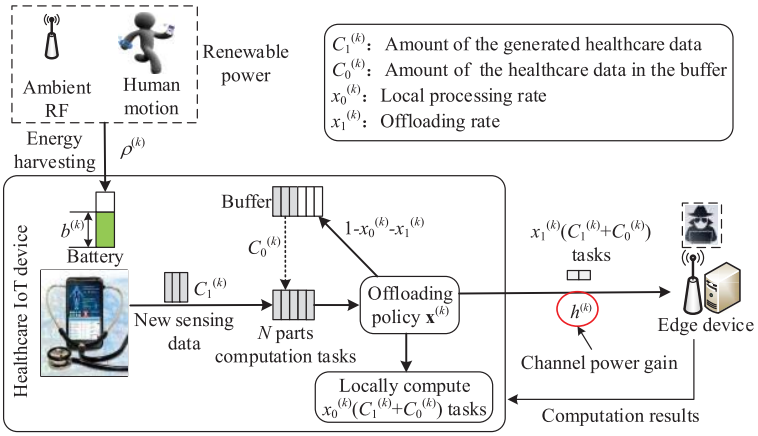


Fig. 1. Illustration of the privacy-aware offloading of an EH powered healthcare IoT device, in which the IoT device offloads $x_1^{(k)}$ of the sensing data to the edge device, locally process $x_0^{(k)}$ of the data at time slot k, and stores the rest data in the buffer to process in the future.

A privacy-preserving opportunistic computing framework for m-Healthcare emergence as proposed in [25] exploits the attribute-based access control and the privacy-preserving scalar product computation technique to reduce the medical data privacy disclosure.

RL techniques have been applied for offloading in MEC. For instance, a Q-learning-based traffic offloading scheme as presented in [26] makes tradeoff between the energy consumption and the quality-of-service for mobile devices in heterogeneous cellular networks. An online learning-based resource management algorithm in [15] uses the PDS to choose the on-the-fly workload offloading rate to both the centralized cloud and the edge server to reduce both the service delay and the operational cost. The computation offloading strategy proposed in [27] uses Q-learning to help IoT devices choose the offloading rate and reduce the attack rate of smart attackers without being aware of the channel model.

III. System Model

We consider a healthcare IoT device that uses multiple sensors to measure and evaluate the healthcare data, such as the blood pressure and electrocardiograms to provide the emergency care and telehealth advice. Powered with both the battery and the EH module, the IoT device can locally process some computation tasks, offload some tasks to the edge device and save the others to process in the next time slot, as shown in Fig. 1.

The IoT device at time slot k is assumed to generate new sensing data of size $C_1^{(k)}$ and has to process the previous sensing data stored in the buffer of size $C_0^{(k)}$. The time index k in the superscript is omitted if no confusion incurs. By applying the computation partition scheme proposed in [28], the IoT device divides the sensing data of size $C_1^{(k)} + C_0^{(k)}$ into N equivalent computation tasks for simplicity. The priority of such sensing data denoted by $\chi^{(k)}$ can be estimated according to the data analysis algorithm, such as [29].

The IoT device offloads $x_1^{(k)}$ of the computation tasks to the edge device over the radio channel with the radio channel power gain $h^{(k)}$, locally processes $x_0^{(k)}$ of the sensing data with the local CPU at the computation speed of f bits per second and stores the rest tasks in the buffer to process in the future, with $\{x_0^{(k)}, x_1^{(k)}\} \in \{l_0/N, l_1/N\}_{0\le l_0, l_1 \le N}$. The radio channel power gain $h^{(k)}$ is formulated as a Markov chain model with

$$\Pr\left(h^{(k+1)} = m | h^{(k)} = n\right) = h_{mn} \quad \forall\, m, n \in \mathbf{H} \tag{1}$$

where $\mathbf{H}$ is the radio channel state set. The IoT device consumes ς energy to process one bit sensing data and uses P energy to send one bit sensing data to the edge device.

The edge device sends the computation results to the IoT device and some attacker might be curious about the user privacy, such as the user location and the usage pattern of the IoT device. An edge device can infer the location privacy and the usage pattern of the IoT device based on the offloading history under different channel states that depends on the distance of the user to the edge node [17]. The privacy level is associated with the size of sensing data and the offloading rate.

The IoT device applies the privacy metric similar to [17] to evaluate the privacy level denoted by $R^{(k)}$, estimate the queuing cost denoted by $W^{(k)}$ and measure the computation latency denoted by $T^{(k)}$, and the energy consumption denoted by $E^{(k)}$. The computation latency is the maximum of the local processing latency $T_0^{(k)}$ and the processing delay of offloading $T_1^{(k)}$, i.e.,

$$T = \max\left\{T_0^{(k)}, T_1^{(k)}\right\}. \tag{2}$$

The energy consumption of the IoT device consists of the local processing energy consumption $E_0^{(k)}$ and the transmission cost $E_1^{(k)}$.

The RF energy harvester and piezoelectric materials enable the IoT device convert renewable energy (such as ambient RF signals and human motion) to electricity [30]–[33]. The IoT device obtains $\rho^{(k)}$ harvested energy at time slot k to support the local data processing and offloading. The battery level at the beginning of time slot k denoted by $b^{(k)}$ is related to the previous battery level, the energy consumption, and the harvested energy and given by

$$b^{(k)} = b^{(k-1)} - E^{(k-1)} + \rho^{(k-1)}. \tag{3}$$

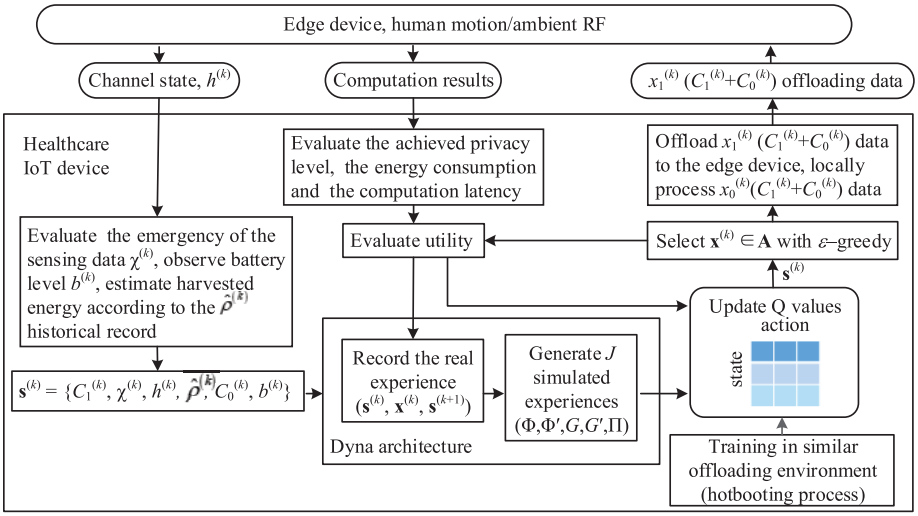


Fig. 2. Illustration of the RL-based privacy-aware offloading for healthcare IoT devices.

TABLE I
LIST OF NOTATIONS

Symbol	Description
$C_1^{(k)}$	Amount of the healthcare data newly generated at time slot k
$\chi^{(k)}$	Priority of the healthcare data
$h^{(k)}$	Channel power gain between the IoT device and the edge device
$\rho^{(k)}$	Amount of the harvested energy
$C_0^{(k)}$	Amount of the healthcare data in the buffer
$b^{(k)}$	Battery level
$\mathbf{x}^{(k)} \in \mathbf{A}$	Offloading strategy
P	Energy consumption per bit for offloading the healthcare data to the edge device
ς	Energy consumption for the IoT device to process a bit data
f	Computation capability of the IoT device
$R^{(k)}$	Achieved privacy level
$W^{(k)}/\hat{W}^{(k)}$	Actual/measured queuing cost
$E^{(k)}/\hat{E}^{(k)}$	Actual/measured energy consumption
$T^{(k)}/\hat{T}^{(k)}$	Actual/measured computation latency

Note that the IoT device drops the computation tasks at time slot k, if its energy is insufficient, i.e., $b^{(k+1)} < 0$ [14]. Important symbols are summarized in Table I.

IV. RL-Based Privacy-Aware Offloading Scheme

We propose an RL-based privacy-aware offloading scheme as shown in Fig. 2 for a healthcare IoT device to choose both the offloading rate and the local processing rate. More specifically, the offloading policy is chosen based on the expected discounted long-term utility or Q-function denoted by Q for the current state. The offloading policy is chosen based on the current state $\mathbf{s}^{(k)}$ that consists of the size and priority of the new healthcare sensing data, the current radio channel state, the estimated renewable energy generated in the time slot, the current battery level of the IoT device, and the computation history. This scheme applies the known radio channel model and generates simulated experiences to reduce the time required to learn the optimal policy.

Algorithm 1 RL-Based IoT Offloading Algorithm

Initialize α, γ and δ
Hotbooting process as [34]
Set $\mathbf{Q} = \bar{\mathbf{Q}}$, $\Phi = 0$, $\Phi' = 0$, $G' = 0$, $G = 0$, $\Pi = 0$
for $k = 1, 2, 3, \ldots$ **do**
  Observe $C_1^{(k)}$, $C_0^{(k)}$ and $b^{(k)}$
  Evaluate $\chi^{(k)}$
  Estimate $h^{(k)}$ and $\hat{\rho}^{(k)}$
  $\mathbf{s}^{(k)} = \{C_1^{(k)}, \chi^{(k)}, h^{(k)}, \hat{\rho}^{(k)}, C_0^{(k)}, b^{(k)}\}$
  Divide the healthcare data with size of $C_0^{(k)} + C_1^{(k)}$ into N equivalent computation tasks
  Choose $\mathbf{x}^{(k)} = [x_0^{(k)}, x_1^{(k)}] \in \mathbf{A}$ with ε-greedy policy
  Offload $x_1^{(k)}(C_1^{(k)} + C_0^{(k)})$ healthcare data to the edge device, process $x_0^{(k)}(C_1^{(k)} + C_0^{(k)})$ of the data locally, and store the rest in the buffer
  Evaluate the achieved privacy $\hat{R}^{(k)}$, the total energy consumption $\hat{E}^{(k)}$, and the computation latency $\hat{T}^{(k)}$
  Evaluate $u^{(k)}$ via (5)
  Estimate $\tilde{\mathbf{s}}^{(k)}$ via (6)
  Evaluate $b^{(k+1)}$ via (3)
  Update $Q(\tilde{\mathbf{s}}^{(k)}, \mathbf{x}^{(k)})$ via (7)
  Update $Q(\mathbf{s}^{(k)}, \mathbf{x}^{(k)})$ via (8)
  Formulate the real experience $(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, u^{(k)}, \mathbf{s}^{(k+1)})$
  Update $\Phi'\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \mathbf{s}^{(k+1)}\right)$ via (9)
  Update $\Phi\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right)$ via (10)
  Update the state transition probability function $\Pi\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \mathbf{s}^{(k+1)}\right)$ via (11)
  Calculate the reward record $G'\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \Phi\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right)\right)$ via (12)
  Update the reward function $G\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right)$ via (13)
  for $j = 1$ to J **do**
    Randomly select $\left(\bar{\mathbf{s}}^{(j)}, \bar{\mathbf{x}}^{(j)}\right)$
    Select $\bar{\mathbf{s}}^{(j+1)}$ based on $\Pi\left(\mathbf{s}^{(j)}, \mathbf{x}^{(j)}, \mathbf{s}^{(j+1)}\right)$
    Calculate $\hat{u}^{(j)} = G\left(\bar{\mathbf{s}}^{(j)}, \bar{\mathbf{x}}^{(j)}\right)$ via (14)
    Update $Q\left(\bar{\mathbf{s}}^{(j)}, \bar{\mathbf{x}}^{(j)}\right)$ via (7)
  end for
end for

As shown in Algorithm 1, upon measuring the healthcare data of size $C_1^{(k)}$ at time slot k, an IoT device briefly

evaluates the priority of the healthcare data denoted by $\chi^{(k)}$ and estimates the channel power gain to the edge device $h^{(k)}$. According to the historical record and the offloading experiences, the IoT device estimates the amount of the harvested energy $\hat{\rho}^{(k)}$ and observes the current battery level of the IoT device $b^{(k)}$. The state is chosen as $\mathbf{s}^{(k)} = \{C_1^{(k)}, \chi^{(k)}, h^{(k)}, \hat{\rho}^{(k)}, C_0^{(k)}, b^{(k)}\}$. Let Λ be the state space.

The new and buffered healthcare data with size of $C_0^{(k)}+C_1^{(k)}$ are divided into N equivalent computation tasks based on the computation partition method [28]. The offloading policy $\mathbf{x}^{(k)} = [x_0^{(k)}, x_1^{(k)}] \in \mathbf{A}$ is chosen according to the ϵ-greedy policy to make a tradeoff between exploration and exploitation [18]. More specifically, the offloading policy that maximizes $Q(\mathbf{s}^{(k)}, \mathbf{x})$ is chosen with $1-\epsilon$ and other feasible offloading policies are randomly selected with a small probability. The IoT device offloads $x_1^{(k)}(C_1^{(k)} + C_0^{(k)})$ healthcare data to the edge device, processes $x_0^{(k)}(C_1^{(k)} + C_0^{(k)})$ of the data locally, and stores the rest in the buffer to process in the future.

After receiving the computation report from the edge device and finishing the local processing, the IoT device analyzes the difference between the size of the sensed data and the size of the offloading data as well as the current channel states to evaluate the achieved privacy level $R^{(k)}$.

More specifically, the IoT device tends to offload all the sensed data to the edge device if it has good radio channel compared with the good channel index $\hat{h}$, and processes all the sensed data locally if it has the bad radio channel compared with the bad channel index $\check{h}$. Let ω denote the importance of the location privacy over the usage pattern privacy. The indicated function denoted by $\mathbb{I}$ equals 1 if the statement is true and 0 otherwise. The achieved privacy level consists of the achieved usage pattern privacy and the location privacy. The former is modeled with $|C_1^{(k)} - x_1^{(k)}(C_1^{(k)} + C_0^{(k)})| \cdot \mathbb{I}(h^{(k)} \geq \hat{h})$, and the latter is $\omega \cdot \mathbb{I}(x_1^{(k)}(C_1^{(k)} + C_0^{(k)}) > 0) \cdot \mathbb{I}(h^{(k)} \leq \check{h})$. The IoT device achieves the privacy level ξ if $\check{h} < h^{(k)} < \hat{h}$. Similar to [17], the privacy level $R^{(k)}$ is estimated with

$$\begin{aligned} R^{(k)} = &\left|C_1^{(k)} - x_1^{(k)}\left(C_1^{(k)} + C_0^{(k)}\right)\right| \cdot \mathbb{I}\left(h^{(k)} \geq \hat{h}\right) \\ &+ \omega \cdot \mathbb{I}\left(x_1^{(k)}\left(C_1^{(k)} + C_0^{(k)}\right) > 0\right) \cdot \mathbb{I}\left(h^{(k)} \leq \check{h}\right) \\ &+ \xi \mathbb{I}\left(\check{h} < h^{(k)} < \hat{h}\right). \end{aligned} \tag{4}$$

An IoT device deliberately reduces the offloading rate under good channel state and increases the offloading rate under low channel power gains to protect privacy. The usage pattern privacy as indicated in the first term of (4) represents the difference between the actual sensing data size and the offloading data size under high radio channel power gains. The location privacy as indicated in the third term of (4) shows whether the IoT device stays in some specific locations with severe radio channel degradations.

The utility of the IoT device depends on the queuing cost $\hat{W}^{(k)}$, the computation latency $\hat{T}^{(k)}$, and the energy consumption $\hat{E}^{(k)}$. Let ψ represent the loss to the device due to the failure to carry out a computation task in time and ν be the queuing weight. Let β and μ denote the importance of the energy saving and the fast computation, respectively. The

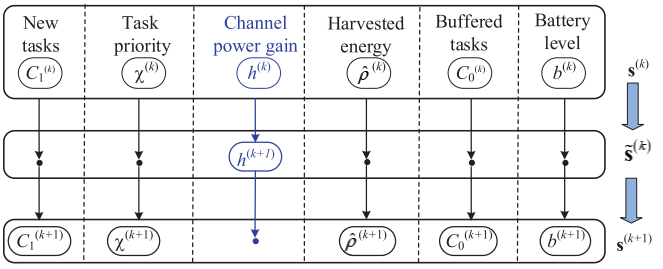


Fig. 3. State transition of the IoT device based on PDS-learning in the dynamic healthcare system, where $\tilde{\mathbf{s}}^{(k)}$ is the immediate state between $\mathbf{s}^{(k)}$ and the next state $\mathbf{s}^{(k+1)}$.

IoT device evaluates its utility $u^{(k)}$ by

$$u^{(k)} = R^{(k)} - \psi \mathbb{I}\left(b^{(k+1)} < 0\right) - \beta \hat{E}^{(k)} - \mu \hat{T}^{(k)} - \nu \hat{W}^{(k)}. \tag{5}$$

The IoT device estimates the next channel power gain $h^{(k+1)}$ based on the known channel model given by (1). As shown in Fig. 3, an immediate state $\tilde{\mathbf{s}}^{(k)} = [C_1^{(k)}, \chi^{(k)}, h^{(k+1)}, \hat{\rho}^{(k)}, C_0^{(k)}, b^{(k)}]$ with a known transition probability h_{mn}, i.e.,

$$\Pr\left(\tilde{\mathbf{s}}^{(k)} \middle| \mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right) = h_{mn}. \tag{6}$$

The IoT device estimates intermediate utility $u(\tilde{\mathbf{s}}^{(k)}, \mathbf{x}^{(k)})$ via (5). Based on the received offloading reports, the estimated energy consumption and the computation latency, the IoT device obtains the next state $\mathbf{s}^{(k+1)}$. The Q-function $Q(\tilde{\mathbf{s}}, \mathbf{x})$ is then updated based on the immediate state $\tilde{\mathbf{s}}^{(k)}$ and utility $u(\tilde{\mathbf{s}}^{(k)}, \mathbf{x}^{(k)})$ according to the iterated Bellman equation as follows:

$$\begin{aligned} Q\left(\tilde{\mathbf{s}}^{(k)}, \mathbf{x}^{(k)}\right) \leftarrow &(1-\alpha) Q\left(\tilde{\mathbf{s}}^{(k)}, \mathbf{x}^{(k)}\right) \\ &+ \alpha\left(u\left(\tilde{\mathbf{s}}^{(k)}, \mathbf{x}^{(k)}\right) + \gamma \max_{\mathbf{x}' \in \mathbf{A}} Q\left(\mathbf{s}^{(k+1)}, \mathbf{x}'\right)\right) \end{aligned} \tag{7}$$

where the learning rate $\alpha \in (0, 1]$ weighs the current offloading experience and the discount factor $\gamma \in [0, 1]$ indicates the myopic view of the IoT device regarding the future reward. The quality function is then updated as follows:

$$Q(\mathbf{s}, \mathbf{x}) \leftarrow \sum_{\tilde{\mathbf{s}} \in \Lambda} \Pr(\tilde{\mathbf{s}}|\mathbf{s}, \mathbf{x}) Q(\tilde{\mathbf{s}}, \mathbf{x}). \tag{8}$$

The offloading experience $(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, u^{(k)}, \mathbf{s}^{(k+1)})$ is used to build the Dyna architecture and generate J simulated experiences in each time slot. More specifically, the model learning depends on an occurrence count vector of the next state denoted by Φ', which is updated by

$$\Phi'\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \mathbf{s}^{(k+1)}\right) \leftarrow \Phi'\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \mathbf{s}^{(k+1)}\right) + 1. \tag{9}$$

The occurrence count vector in the simulated experience denoted by Φ is updated in each real offloading experience by

$$\Phi\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right) \leftarrow \sum_{\mathbf{s}' \in \Lambda} \Phi'\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \mathbf{s}'\right). \tag{10}$$

The transition probability to reach the state $\mathbf{s}^{(k+1)}$ from the state-action pair $(\mathbf{s}^{(k)}, \mathbf{x}^{(k)})$ is denoted by $\Pi(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \mathbf{s}^{(k+1)})$

and updated by

$$\Pi\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \mathbf{s}^{(k+1)}\right) \leftarrow \frac{\Phi'\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \mathbf{s}^{(k+1)}\right)}{\Phi\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right)}. \tag{11}$$

The reward record denoted by G' is the utility of the IoT device from the real offloading experience $u^{(k)}$, i.e.,

$$G'\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \Phi\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right)\right) = u^{(k)}. \tag{12}$$

The average reward function over all the occurrence realizations denoted by G is updated by

$$G\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right) = \frac{1}{\Phi\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right)} \sum_{\kappa=1}^{\Phi\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}\right)} G'\left(\mathbf{s}^{(k)}, \mathbf{x}^{(k)}, \kappa\right). \tag{13}$$

The J simulated experiences are then generated from the Dyna architecture model (Π, G) via search control. Each simulated experience at time slot k leads to an additional Q-function update. More specifically, in the jth update, the IoT device first randomly chooses a state-action pair $(\bar{\mathbf{s}}^{(j)}, \bar{\mathbf{x}}^{(j)})$ and selects the next state $\bar{\mathbf{s}}^{(j+1)}$ based on the state transition probability Π given by (11). The modeled reward $\hat{u}^{(j)}$ depends on the reward function G in (13) with the state-action pair $(\bar{\mathbf{s}}^{(j)}, \bar{\mathbf{x}}^{(j)})$ as follows:

$$\hat{u}^{(j)}\left(\bar{\mathbf{s}}^{(j)}, \bar{\mathbf{x}}^{(j)}\right) = G\left(\bar{\mathbf{s}}^{(j)}, \bar{\mathbf{x}}^{(j)}\right). \tag{14}$$

The Q-function for $(\bar{\mathbf{s}}^{(j)}, \bar{\mathbf{x}}^{(j)})$ is updated via the iterated Bellman equation again with (7).

As shown in Fig. 2, this scheme uses a transfer learning method named hotbooting as developed in [34] to initialize the Q-values with the computation offloading experiences in similar environments to save the random explorations. More specifically, the Q-values after training with ξ offloading experiences randomly selected from an offloading experience pool are denoted by $\overline{\mathbf{Q}}$ and used to initialize the Q-values in Algorithm 1. To sum up, the proposed offloading scheme is formalized in Algorithm 1.

V. Performance Evaluations

We analyze the performance of the proposed RL-based privacy-aware offloading scheme regarding the privacy level, energy consumption, computation latency, and utility. Similar to [14] and [35], we focus on the delay-sensitive applications, i.e., both the local execution and the task offloading can be completed in a time slot. For simplicity, both the computation latency of the edge device and the transmission latency of the computation results is assumed to be negligible, while this algorithm works in the other scenarios as well.

At time slot k, the IoT device offloads $x_1^{(k)}(C_1^{(k)} + C_0^{(k)})$ healthcare data to the edge device, locally processes $x_0^{(k)}(C_1^{(k)} + C_0^{(k)})$ data, and stores the rest in the buffer to process in the future. According to [21] and [36], the IoT device consumes $\varsigma x_0^{(k)}(C_1^{(k)} + C_0^{(k)})$ energy to compute the sensing data locally and takes $Px_1^{(k)}(C_1^{(k)} + C_0^{(k)})$ energy in the offloading process. Thus, we have

$$E = Px_1(C_1 + C_0) + \varsigma x_0(C_1 + C_0). \tag{15}$$

The data rate in the offloading can be modeled with $\log_2(1 + Ph)$. The IoT device takes $T_0^{(k)} = x_0^{(k)}(C_1^{(k)} + C_0^{(k)})/f$ to compute the local sensing data at time slot k [36], and the offloading latency is $T_1^{(k)} = x_1^{(k)}(C_1^{(k)} + C_0^{(k)})/\log_2(1 + Ph^{(k)})$. Thus by (2), the total computation latency of the IoT device denoted by $T^{(k)}$ is given by

$$T = \max\left\{\frac{x_0(C_1 + C_0)}{f}, \frac{x_1(C_1 + C_0)}{\log_2(1 + Ph)}\right\}. \tag{16}$$

According to [37], the average queuing delay linearly increases with the average queue length and the priority of the generated healthcare data $\chi^{(k)}$. Thus, the queuing cost denoted by $W^{(k)}$ is defined as

$$W = \chi(1 - x_0 - x_1)(C_1 + C_0). \tag{17}$$

The offloading selection over multiple time slots can be viewed as an MDP, as the future state is independent of the previous states, for a given current state and offloading policy. Therefore, the RL-based offloading scheme in Algorithm 1 can achieve the optimal policy via trial-and-error without being aware of the privacy leakage, IoT energy consumption, and edge computation model.

Theorem 1: The healthcare IoT device using Algorithm 1 in the dynamic offloading game can achieve the optimal policy given by $\mathbf{x}^* = [0, 1]$, and the privacy level is C_0. The computation latency, the energy consumption and the utility are given, respectively, by

$$T = \frac{C_1 + C_0}{\log_2(1 + Ph)} \tag{18}$$

$$E = P(C_1 + C_0) \tag{19}$$

$$u = C_0 - \left(\beta P + \frac{\mu}{\log_2(1 + Ph)}\right)(C_1 + C_0) \tag{20}$$

if

$$h > \hat{h} \tag{21}$$

$$\mu < (\nu\chi + 1 - \beta P)\log_2(1 + Ph) \tag{22}$$

$$\nu\chi < \beta\varsigma \tag{23}$$

$$b + \rho > P(C_1 + C_0). \tag{24}$$

Proof: By (5), if (21) and $b + \rho - \varsigma(C_1 + C_0)x_0 > P(C_1 + C_0)x_1$, we have

$$\begin{aligned} u(\mathbf{x}) &= \left(\nu\chi - \beta P + 1 - \frac{\mu}{\log_2(1 + Ph)}\right)(C_1 + C_0)x_1 \\ &\quad + (\nu\chi - \beta\varsigma)(C_1 + C_0)x_0 - C_1 - \nu\chi(C_1 + C_0) \\ &\quad - \psi\mathbb{I}(b + \rho - \varsigma(C_1 + C_0)x_0 - P(C_1 + C_0)x_1) \\ &= \left(\nu\chi - \beta P + 1 - \frac{\mu}{\log_2(1 + Ph)}\right)(C_1 + C_0)x_1 \\ &\quad + (\nu\chi - \beta\varsigma)(C_1 + C_0)x_0 - C_1 - \nu\chi(C_1 + C_0). \end{aligned} \tag{25}$$

If (22), (23), $\forall\ \mathbf{x} \in \mathbf{A}$

$$\frac{\partial u}{\partial x_0} = (\nu\chi - \beta\varsigma)(C_0 + C_1) < 0 \tag{26}$$

indicating that the utility decreases with x_0

$$\frac{\partial u}{\partial x_1} = \left(\nu\chi - \beta P + 1 - \frac{\mu}{\log_2(1 + Ph)}\right)(C_0 + C_1) > 0 \tag{27}$$

indicating that the utility with x_1. As $x_0 \in [0, 1]$ and $x_1 \in [0, 1]$, we have $\arg\max_{\mathbf{x}\in\mathbf{A}} u = [0, 1]$.

According to [18], this RL-based scheme can achieve the optimal policy $\mathbf{x}^* = [0, 1]$ in the MDP after a sufficient long time. Therefore, this algorithm can achieve $\mathbf{x}^* = [0, 1]$. If (24), by (4), (16), and (15), we have $R = C_0$ and prove (18)–(20). ■

Remark 1: A healthcare IoT device applies the RL-based privacy-aware offloading algorithm to achieve the optimal policy without being aware of the privacy leakage, IoT energy consumption, and edge computing model in the dynamic offloading process. If the IoT device has good radio channel to the edge device as shown in (22), the local processing energy overhead is high as shown in (23), and the offloading energy consumption is low as shown in (24), the IoT device will offload all the computation tasks to the edge device. In this case, the privacy level of the IoT device equals the size of the buffered tasks, and both the computation latency and IoT energy consumption increase linearly with the size of the total computation tasks as indicated in (18) and (19).

Theorem 2: The healthcare IoT device using Algorithm 1 in the dynamic offloading game can achieve the optimal policy given by $\mathbf{x}^* = [1, 0]$, and the privacy level is C_1. The computation latency, the energy consumption, and the utility are given, respectively, by

$$T = \frac{C_1 + C_0}{f} \tag{28}$$

$$E = \varsigma(C_1 + C_0) \tag{29}$$

$$u = C_1 - \left(\beta\varsigma + \frac{\mu}{f}\right)(C_1 + C_0) \tag{30}$$

if

$$h > \hat{h} \tag{31}$$

$$\nu\chi < \beta P + 1 \tag{32}$$

$$\mu < f(\nu\chi - \beta\varsigma) \tag{33}$$

$$b + \rho > \varsigma(C_1 + C_0). \tag{34}$$

Proof: The proof is similar to that of Theorem 1. ■

Remark 2: If the healthcare sensing data seem normal as shown in (32), the offloading energy overhead is high as shown in (33), and the IoT device has powerful computation resources as shown in (34), the IoT device will process all the computation tasks locally. In this case, the privacy level equals the size of the new sensing data, and both the IoT energy consumption and the computation latency increase with the total computation tasks size as indicated in (28) and (29).

Theorem 3: The healthcare IoT device using Algorithm 1 in the dynamic offloading game can achieve the optimal policy given by $\mathbf{x}^* = [0, C_0/(C_1 + C_0)]$, and the privacy level is ω. The computation latency, the energy consumption, and the utility are given, respectively, by

$$T = \frac{C_0}{\log_2(1 + Ph)} \tag{35}$$

$$E = PC_0 \tag{36}$$

$$u = \omega - \psi - \left(\beta P - \frac{\mu}{\log_2(1 + Ph)}\right)C_0 \tag{37}$$

if

$$h < \check{h} \tag{38}$$

$$\mu < (\nu\chi - \beta P)\log_2(1 + Ph) \tag{39}$$

$$\nu\chi < \beta\varsigma \tag{40}$$

$$b + \rho < PC_0. \tag{41}$$

Proof: The proof is similar to that of Theorem 1. ■

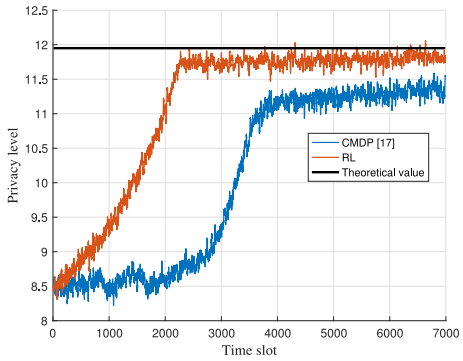
Remark 3: If the offloading energy overhead is low as shown in (39), the local processing energy overhead is high as shown in (40), and the IoT device has insufficient computation resources as shown in (41), the IoT device will offload some sensing data to the edge device and store the rest tasks to the buffer to protect the user privacy. In this case, the IoT device can achieve a privacy level given by ω, and both the computation latency and the IoT energy consumption increase linearly with the size of the buffered tasks as indicated in (35) and (36).

VI. Simulation Results

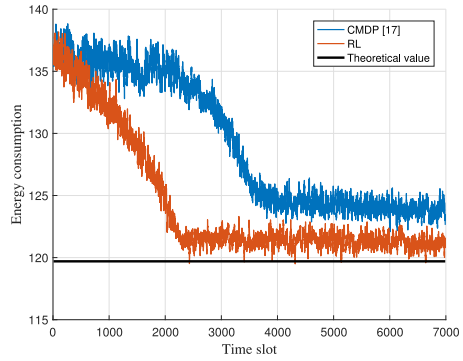
Simulations have been performed to evaluate the RL-based privacy-aware offloading scheme in dynamic healthcare systems. In the simulations, each time slot lasts 1 s, and the IoT device generates new healthcare data of 30 kb. According to [14], the IoT device consumes 10^{-4} J energy to locally process one bit sensing data and uses 0.2 J energy to send one bit sensing data to the edge device. The queuing delay, the location privacy over usage pattern privacy, the energy consumption, and the computation latency are weighted with 40, 5, 2.5, and 5, respectively. If not specified otherwise, the learning rate is 0.8, the discount factor is 0.7 and ϵ is 0.1 according to [38]. The CMDP-learning-based offloading scheme in [17] has been evaluated in the simulations as a benchmark.

As shown in Fig. 4, the RL-based offloading scheme converges to the performance bound given by Theorem 2. This scheme exceeds the CMDP-based offloading scheme as proposed in [17] with a higher privacy level. This scheme also saves the energy consumption of the IoT device, reduces the computation latency, and increases the utility of the IoT device. For instance, this scheme improves 36.63% of the privacy level, saves 9.63% of the energy consumption, and decreases 68.79% of the computation latency, compared with the CMDP-based scheme at the 2200th time slot. Consequently, as shown in Fig. 4(d), the utility of the healthcare IoT device increases about two times compared with that of the CMDP-based scheme. Fig. 4 shows that the RL-based scheme accelerates the learning speed, e.g., this scheme saves 40% of time slots to reach the privacy level of 11 compared with the CMDP. This is due to the fact that the transfer learning technique, a PDS method and a Dyna architecture are used to accelerate the learning speed of the healthcare IoT device with the extended state space.

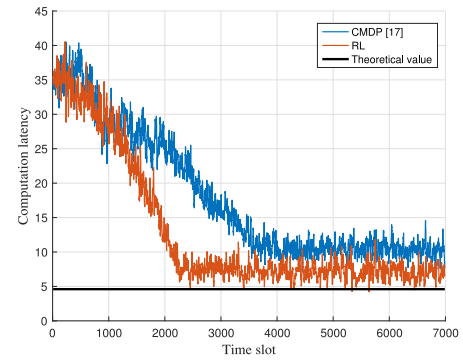
The offloading performance averaged over the first 4500 time slots in the dynamic offloading game is shown in Fig. 5. In the simulations, an IoT device has to compute 10 to 50 kb new healthcare sensing data in each time slot. The privacy level of the healthcare IoT device increases, as the amount of the sensing data changes from 10 to 50 kb. For instance, the privacy level, the energy consumption, the computation



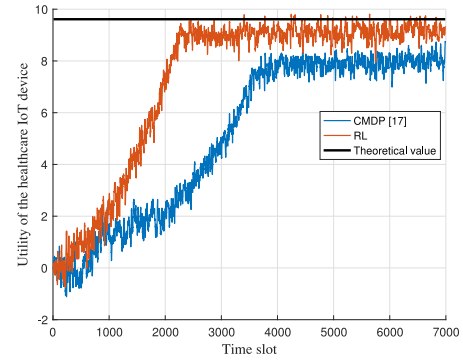
(a)



(b)

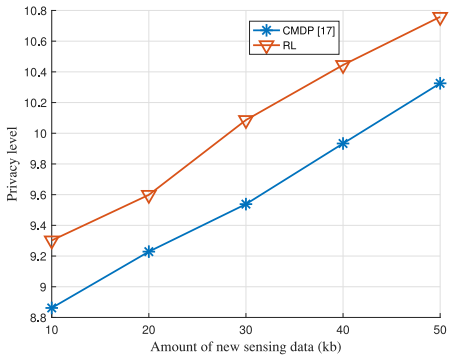


(c)

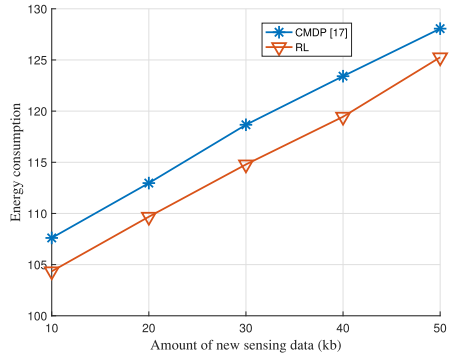


(d)

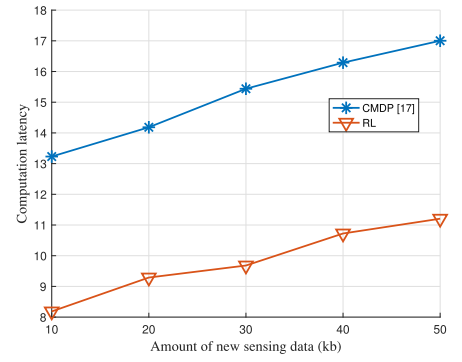
Fig. 4. Performance of the privacy-aware offloading scheme in a healthcare IoT device with EH. (a) Achieved privacy level. (b) Energy consumption of the IoT device. (c) Computation latency. (d) Utility of the IoT device.



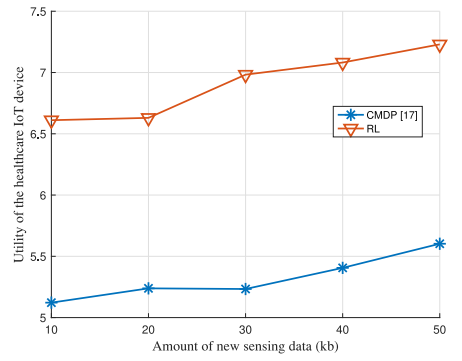
(a)



(b)



(c)



(d)

Fig. 5. Performance of the privacy-aware offloading scheme in a healthcare IoT device with EH that generates a amount of the sensing data in each time slot. (a) Achieved privacy level. (b) Energy consumption of the IoT device. (c) Computation latency. (d) Utility of the IoT device.

latency, and the utility of the healthcare IoT device using RL-based offloading scheme increase by 28.93%, 40.78%, 100% and 28.79%, respectively, if the amount of the IoT healthcare data increases from 10 kb to 50 kb. If the healthcare IoT device has to process 50 kb healthcare data in each time slot as shown in Fig. 5, the RL-based offloading scheme exceeds

the benchmark CMDP scheme with 12.39% higher privacy level, 5.38% lower energy consumption, and 32.35% shorter computation latency.

VII. Conclusion

In this paper, we have proposed an RL-based privacy-aware offloading scheme for an EH powered healthcare IoT device to choose the offloading rate and the local computing rate without being aware of the privacy leakage, IoT energy consumption, and edge computation model. This scheme evaluates the privacy level, the energy consumption, and the computation latency to choose the offloading policy to the edge device in each time slot. The RL-based offloading scheme uses the transfer learning technique, a known radio channel model and a Dyna architecture to accelerate the learning speed for dynamic healthcare IoT systems. We prove that this scheme can achieve the optimal offloading policy in the dynamic offloading process and provide its performance upper bound in terms of the privacy level, the computation latency, and the energy consumption. Simulation results show that this scheme improves the privacy level by 36.63%, saves 9.63% of the energy consumption and decreases 68.79% of the computation latency compared with the benchmark CMDP scheme.

References

[1] S. M. R. Islam, D. Kwak, M. 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.

[2] S. Ulukus *et al.*, "Energy harvesting wireless communications: A review of recent advances," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360–381, Mar. 2015.

[3] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile edge computing: A survey," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 450–465, Feb. 2018.

[4] M. Chiang and T. Zhang, "Fog and IoT: An overview of research opportunities," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 854–864, Dec. 2016.

[5] J. Ren, H. Guo, C. Xu, and Y. Zhang, "Serving at the edge: A scalable IoT architecture based on transparent computing," *IEEE Netw.*, vol. 31, no. 5, pp. 96–105, Aug. 2017.

[6] X. Peng *et al.*, "BOAT: A block-streaming app execution scheme for lightweight IoT devices," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1816–1829, Jun. 2018.

[7] D. Zhang, R. Shen, J. Ren, and Y. Zhang, "Delay-optimal proactive service framework for block-stream as a service," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 598–601, Aug. 2018, doi: 10.1109/LWC.2018.2799935.

[8] W. Ji, Z. Li, and Y. Chen, "Joint source-channel coding and optimization for layered video broadcasting to heterogeneous devices," *IEEE Trans. Multimedia*, vol. 14, no. 2, pp. 443–455, Apr. 2012.

[9] W. Ji, B.-W. Chen, Y. Chen, and S.-Y. Kung, "Profit improvement in wireless video broadcasting system: A marginal principle approach," *IEEE Trans. Mobile Comput.*, vol. 14, no. 8, pp. 1659–1671, Aug. 2015.

[10] S. Yi, Z. Qin, and Q. Li, "Security and privacy issues of fog computing: A survey," in *Proc. Int. Conf. Wireless Algorithms Syst. Appl. (WASA)*, Qufu, China, Aug. 2015, pp. 685–695.

[11] J. Ni, A. Zhang, X. Lin, and X. S. Shen, "Security, privacy, and fairness in fog-based vehicular crowdsensing," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 146–152, Jun. 2017.

[12] J. Liu, K. Kumar, and Y.-H. Lu, "Tradeoff between energy savings and privacy protection in computation offloading," in *Proc. ACM/IEEE Int. Symp. Low Power Electron. Design (ISLPED)*, Austin, TX, USA, Aug. 2010, pp. 213–218.

[13] J. Liu and Y.-H. Lu, "Energy savings in privacy-preserving computation offloading with protection by homomorphic encryption," in *Proc. Int. Conf. Power Aware Comput. Syst. HotPower*, vol. 10. Vancouver, BC, Canada, 2010, pp. 1–7.

[14] Y. Mao, J. Zhang, 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.

[15] J. Xu, L. Chen, and S. Ren, "Online learning for offloading and autoscaling in energy harvesting mobile edge computing," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 3, pp. 361–373, Sep. 2017.

[16] F. Wang, J. Xu, X. Wang, and S. Cui, "Joint offloading and computing optimization in wireless powered mobile-edge computing systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1784–1797, Mar. 2018.

[17] X. He, J. Liu, R. Jin, and H. Dai, "Privacy-aware offloading in mobile-edge computing," in *Proc. IEEE Glob. Commun. Conf. (GlobeCom)*, Singapore, Dec. 2017, pp. 1–6.

[18] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 1998.

[19] X. He, H. Dai, and P. Ning, "Improving learning and adaptation in security games by exploiting information asymmetry," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Hong Kong, Aug. 2015, pp. 1787–1795.

[20] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[21] W. Zhang *et al.*, "Energy-optimal mobile cloud computing under stochastic wireless channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4569–4581, Sep. 2013.

[22] C. You, K. Huang, H. Chae, and B.-H. Kim, "Energy-efficient resource allocation for mobile-edge computation offloading," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1397–1411, Mar. 2017.

[23] C. Xu, J. Ren, Y. Zhang, Z. Qin, and K. Ren, "DPPro: Differentially private high-dimensional data release via random projection," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 12, pp. 3081–3093, Dec. 2017.

[24] L. Xiao *et al.*, "Security in mobile edge caching with reinforcement learning," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 116–122, Jun. 2018.

[25] R. Lu, X. Lin, and X. Shen, "SPOC: A secure and privacy-preserving opportunistic computing framework for mobile-healthcare emergency," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 3, pp. 614–624, Mar. 2013.

[26] X. Chen, J. Wu, Y. Cai, H. Zhang, and T. Chen, "Energy-efficiency oriented traffic offloading in wireless networks: A brief survey and a learning approach for heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 4, pp. 627–640, Apr. 2015.

[27] L. Xiao, C. Xie, T. Chen, H. Dai, and H. V. Poor, "A mobile offloading game against smart attacks," *IEEE Access*, vol. 4, pp. 2281–2291, 2016.

[28] W. Liu *et al.*, "AppBooster: Boosting the performance of interactive mobile applications with computation offloading and parameter tuning," *IEEE Trans. Parallel Distrib. Syst.*, vol. 28, no. 6, pp. 1593–1606, Jun. 2017.

[29] N. Golrezaei, A. F. Molisch, A. G. Dimakis, and G. Caire, "Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution," *IEEE Commun. Mag.*, vol. 51, no. 4, pp. 142–149, Apr. 2013.

[30] M. Xia and S. Aissa, "On the efficiency of far-field wireless power transfer," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2835–2847, Jun. 2015.

[31] S. Wang, M. Xia, and Y.-C. Wu, "Space-time signal optimization for SWIPT: Linear versus nonlinear energy harvesting model," *IEEE Commun. Lett.*, vol. 22, no. 2, pp. 408–411, Feb. 2018.

[32] D. Zhang *et al.*, "Two time-scale resource management for green Internet of Things networks," *IEEE Internet Things J.*, to be published, doi: 10.1109/JIOT.2018.2842766.

[33] D. Zhang *et al.*, "Utility-optimal resource management and allocation algorithm for energy harvesting cognitive radio sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3552–3565, Dec. 2016.

[34] L. Xiao, Y. Li, C. Dai, H. Dai, and H. V. Poor, "Reinforcement learning-based NOMA power allocation in the presence of smart jamming," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3377–3389, Apr. 2018.

[35] O. Muñoz, A. Pascual-Iserte, and J. Vidal, "Optimization of radio and computational resources for energy efficiency in latency-constrained application offloading," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4738–4755, Oct. 2015.

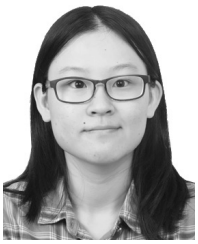
[36] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," *IEEE/ACM Trans. Netw.*, vol. 24, no. 5, pp. 2795–2808, Oct. 2016.

[37] N. Mastronarde and M. Van der Schaar, "Fast reinforcement learning for energy-efficient wireless communication," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6262–6266, Dec. 2011.

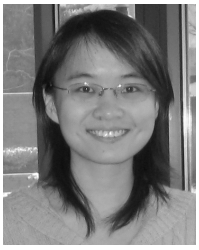
[38] L. Xiao, Y. Li, X. Huang, and X. J. Du, "Cloud-based malware detection game for mobile devices with offloading," *IEEE Trans. Mobile Comput.*, vol. 16, no. 10, pp. 2742–2750, Oct. 2017.

Minghui Min (S'17) received the B.S. degree in automation from Qufu Normal University, Rizhao, China, in 2013, the M.S. degree in control theory and control engineering from Shenyang Ligong University, and joint training with the Shenyang Institute of Automation, Chinese Academy of Sciences, Shengyang, China, in 2016. She is currently pursuing the Ph.D. degree at the Department of Communication Engineering, Xiamen University, Xiamen, China.

Her current research interests include network security and wireless communications.



Xiaoyue Wan received the B.S. degree in communication engineering from Xiamen University, Xiamen, China, in 2016, where she is currently pursuing the M.S. degree at the Department of Communication Engineering.



Liang Xiao (M'09–SM'13) received the B.S. degree in communication engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2000, the M.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 2003, and the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 2009.

She is currently a Professor with the Department of Communication Engineering, Xiamen University, Xiamen, China. She was a Visiting Professor with Princeton University, Princeton, NJ, USA, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, and the University of Maryland at College Park, College Park, MD, USA.

Dr. Xiao was a recipient of the Best Paper Award of 2016 INFOCOM Big Security WS and 2017 ICC. She has served as an Associate Editor for the IEEE TRANSACTIONS INFORMATION FORENSICS AND SECURITY and as a Guest Editor for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.

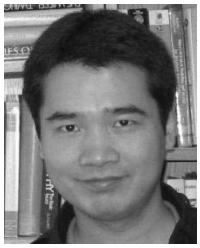
Ye Chen received the B.S. degree in communication engineering from Xiamen University, Xiamen, China, in 2017, where he is currently pursuing the M.S. degree at the Department of Communication Engineering.

His current research interests include network security and wireless communications.

Minghua Xia (M'12) received the Ph.D. degree in telecommunications and information systems from Sun Yat-sen University, Guangzhou, China, in 2007.

Since 2015, he has been a Professor with Sun Yat-sen University. From 2007 to 2009, he was with the Electronics and Telecommunications Research Institute, Daejeon, South Korea, and the Beijing Research and Development Center, Beijing, China, where he was a Member and then a Senior Member of Engineering Staff and participated in projects on the physical layer design of 3GPP LTE mobile communications. From 2010 to 2014, he was a Post-Doctoral Fellow with the University of Hong Kong, Hong Kong, the King Abdullah University of Science and Technology, Jeddah, Saudi Arabia, and the Institut National de la Recherche Scientifique, University of Quebec, Montreal, QC, Canada. He holds two patents granted in China. His current research interests include the general area of 5G wireless communications, and in particular the design and performance analysis of multiantenna systems, cooperative relaying systems and cognitive relaying networks, and recently a focus on the design and analysis of wireless power transfer and/or energy harvesting systems, as well as massive MIMO and small cells.

Dr. Xia was a recipient of the Professional Award of IEEE TENCON, Macau, China, in 2015. He was also recognized as an Exemplary Reviewer by the IEEE TRANSACTIONS ON COMMUNICATIONS in 2014, IEEE COMMUNICATIONS LETTERS in 2014, and IEEE WIRELESS COMMUNICATIONS LETTERS in 2014 and 2015.



Di Wu (M'06–SM'17) received the B.S. degree from the University of Science and Technology of China, Hefei, China, in 2000, the M.S. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2003, and the Ph.D. degree in computer science and engineering from the Chinese University of Hong Kong, Hong Kong, in 2007.

He was a Post-Doctoral Researcher with the Department of Computer Science and Engineering, Polytechnic Institute of New York University, Brooklyn, NY, USA, from 2007 to 2009, advised by Prof. K. W. Ross. He is currently a Professor and the Assistant Dean with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. His current research interests include cloud computing, multimedia communication, Internet measurement, and network security.

Prof. Wu was a co-recipient of the IEEE INFOCOM 2009 Best Paper Award. He has served as an Editor for the *Journal of Telecommunication Systems* (Springer), the *Journal of Communications and Networks*, *Peer-to-Peer Networking and Applications* (Springer), *Security and Communication Networks* (Wiley), and the *KSII Transactions on Internet and Information Systems*, and as a Guest Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He has also served as the MSIG Chair of the Multimedia Communications Technical Committee of the IEEE Communications Society from 2014 to 2016. He served as the TPC Co-Chair of the IEEE Global Communications Conference—Cloud Computing Systems, and Networks, and Applications in 2014, the Chair of the CCF Young Computer Scientists and Engineers Forum—Guangzhou from 2014 to 2015, and as a member of the Council of China Computer Federation.

Huaiyu Dai (F'17) received the B.E. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1996 and 1998, respectively, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2002.

He was with Bell Labs, Lucent Technologies, Holmdel, NJ, USA, in 2000, and AT&T Labs-Research, Middletown, NJ, USA, in 2001. He is currently a Professor of electrical and computer engineering with North Carolina State University, Raleigh, NC, USA. His current research interests include communication systems and networks, advanced signal processing for digital communications, and communication theory and information theory, networked information processing and crosslayer design in wireless networks, cognitive radio networks, network security, and associated information-theoretic, and computation-theoretic analysis.

Dr. Dai was a co-recipient of the Best Paper Award of the 2010 IEEE International Conference on Mobile Ad-Hoc and Sensor Systems, the 2016 IEEE INFOCOM BIGSECURITY Workshop, and the 2017 IEEE International Conference on Communications. He has served as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He is currently an Area Editor in charge of wireless communications for the IEEE TRANSACTIONS ON COMMUNICATIONS. He co-edited two special issues of EURASIP journals on distributed signal processing techniques for wireless sensor networks and on multiuser information theory and related applications, respectively. He co-chaired the Signal Processing for Communications Symposium of IEEE Globecom 2013, the Communications Theory Symposium of IEEE ICC 2014, and the Wireless Communications Symposium of IEEE Globecom 2014.