1-1-2019

An Optimal Dynamic Admission Control Policy and Upper Bound Analysis in Wireless Sensor Networks

Xiaolu Zhang
Donghua University

Demin Li
Donghua University

Wei W. Li
Texas Southern University

Wei Zhao
American University of Sharjah

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 8, 2019, accepted April 2, 2019, date of publication April 22, 2019, date of current version May 2, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2912396

An Optimal Dynamic Admission Control Policy and Upper Bound Analysis in Wireless Sensor Networks

XIAOLU ZHANG[1,2], DEMIN LI[1], WEI W. LI[2], (Senior Member, IEEE), AND WEI ZHAO[3], (Fellow, IEEE)

[1]Engineering Research Center of Digitized Textile and Fashion Technology of Ministry of Education, College of Information Science and Technology, Donghua University, Shanghai 201620, China
[2]NSF Center for Research on Complex Networks, Department of Computer Science, Texas Southern University, Houston, TX 77004, USA
[3]American University of Sharjah, Sharjah 27272, United Arab Emirates

Corresponding author: Demin Li (deminli@dhu.edu.cn)

This work was supported in part by the NSF of China under Grant 71171045, Grant 61772130, and Grant 61301118, in part by the Innovation Program of Shanghai Municipal Education Commission under Grant 14YZ130, in part by the International S&T Cooperation Program of Shanghai Science and Technology Commission under Grant 15220710600, in part by the Fundamental Research Funds for the Central Universities under Grant CUSF-DH-D-2015058, in part by the Shanghai Science and Technology Committee under Grant 18DZ1200500, and in part by the US National Science Foundation under Grant HRD-1137732 and Grant CNS-1827940.

ABSTRACT Data admission control is an important issue in wireless sensor networks (WSNs) because of the limited transmission coverage area and the limited battery capability of each sensor node. To achieve the optimum benefit of the sensor network, a reward when a data package arriving at a sensor is accepted (not rejected) for transmission is considered, but a holding cost per unit time for the accepted data package waiting in the sensor for transmission is also incurred. For the sensor with sleep and active phases, a dynamic data admission control model is developed in this paper. By constructing a suitable Markov decision process (MDP), we verify that the optimal admission control policy of when to admit or reject an arriving data packet is a control limit policy, in order to achieve the maximum expected a discounted reward in the sleep and active phase, respectively. Furthermore, we provide a formula for how to calculate the upper bound of these threshold values. For the identified model with the optimal expected discount reward, the energy consumption of the sensors in active and sleep phases, as well as the energy consumption switching from active to sleep per unit time and vice versa is investigated. Extensive simulation is implemented. The results show that the problem is effectively solved by an optimal scheme with high energy efficiency. The results of this paper can be applied to designing optimal sensor nodes in wireless sensor networks.

INDEX TERMS Admission control, Markov decision process, control limit policy, upper bound, energy consumption.

I. INTRODUCTION

Wireless sensor networks are increasingly popular as sensing systems that can interact with the surrounding environment's dynamics and objects. For example, kinematic sensors can be used to remotely supervise elderly patients and humidity sensors can be deployed to control field irrigation for more sustainable agriculture [1]. Data gathering is the main task of WSNs. In these networks, a sensor node can be either in active or sleep mode. Packets arrive into sensor nodes' queues according to a stationary arrival process. Rewards are obtained when a data packet is accepted into the sensor, while there is also a storage cost when data in a sensor are waiting for transmission. In addition, the storage cost increases significantly with an increase in the data [2], [3]. Therefore, developing an optimal admission control policy in the data gathering process is meaningful in wireless sensor nodes.

Markov decision processes (MDPs), referred to as stochastic control problems, are models for sequential decision making when outcomes are uncertain. MDPs have been applied to many areas, including finance, logistics, manufacturing, and health-care [4]. Also, numerous applications and research studies have proven the MDP to be an effective technique for

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Omer Farooq.

solving optimization problems [5]. In this paper, we consider the data admission control problem of sensor nodes in the data gathering process with MDP models based on the sleep/active scheme. The contributions of this paper are as follows:

- Based on the Markov decision process, we propose a data admission control model with a sleep/active scheme in sensors. For sensor nodes, immediate rewards are obtained once a data packet is accepted into the sensor. However, a storage cost is incurred when data in the sensor are awaiting for transmission. Thus, the admission control model with MDPs is presented to obtain the optimal discounted reward. We model the process clearly, which is necessary and helpful in solving the admission control problem in the WSN' sleep/active scheme.
- Results of the admission control problem to act as a control limit policy are verified with the proposed model. We also provide an upper bound of our obtained optimal $(M; N)$ policy to indicate that when more than M data packets arrive at the sensor in the active status or when more than N data packets arrive at the sensor in the sleep status, the data should be rejected. The solution of the $(M; N)$ policy can be implemented based on a reference table, which can be stored in the sensor node's memory for online operations with minimal complexity.
- For the above identified optimal $(M; N)$ policy model-scheme, performance on energy consumption (energy consumption switching from sleep status to active status and switching from active status to sleep status, average energy consumption in the sleep status, and average energy consumption in the active status) is presented as the sensor dynamics in the sleep/active mode change.

The rest of this paper is organized as follows. In Section II, related works are presented. Section III describes the system model and formulates the problem. Section IV introduces the optimal admission control of our proposed model and presents the main component of this paper. Section V provides the performance evaluation on energy consumption and Section VI provides the performance evaluation by numerical simulation. Finally, Section VII summarizes the findings.

II. RELATED WORK

To prolong the life of a sensor node, the sleep/active mechanism is typically utilized. For this kind of mechanism, the MAC operation is divided into cycles, and in each cycle interval, each sensor node periodically cycles between an awake state and a sleep state and the energy consumed during the awake state and the sleep state [6]. During these cycles, a sensor node admits and serves data for different purposes. There is abundant literature related to the data admission control during the sleep/active mechanism in WSNs from different metrics, such as transmission delay, energy consumption, and expected network congestion [5]. Meanwhile, different MDP models are employed to optimize the sleep/active mechanism.

Some research is devoted to studying the tradeoff between data gathering and transmission delay. In [7], with respect to delay guarantees, a fully distributed algorithm based on a constrained MDP is proposed during the data admission process to maximize data quality from sensor fusion. In that paper, the authors achieve soft delay guarantees and good data quality compared with other schemes. In [8], two real time solutions with a Semi Markov decision process (SMDP) are provided, one based on dynamic programming and the other based on Q-learning, to solve the problem of the trade-off between energy consumption and data delivery latency within local data exchanging. And in [9], the authors put forward an MDP based on a geographic routing protocol. The nodes select their actions of "immediate data forwarding" or "wait to collect more data samples" based on local network states to control the data admission process. Simulation results show that the adaptive routing protocol enhances successful packet delivery ratios under end-to-end delay constraints.

WSNs operate under limited energy resource during data gathering [5]. Thus, researches concerning energy saving during data gathering have been attractive with last several years also. In [10], to improve the performance of data gathering while saving the energy of the network, the authors formulate the moving process of data collectors as a Markov chain and determine the moving path using MDP. In [11], a new sleep-scheduling algorithm, referred to as the energy consumed uniformly-connected K-neighborhood (EC-CKN) algorithm, is proposed based on MDP to prolong the network life. The algorithm EC-CKN takes the nodes' residual energy information as the parameter to decide whether a node should be active or asleep.

The problem of precise target tracking in resource-constrained WSNs has been investigated. In [12], the minimum number of active nodes is selected by cluster heads based on an MDP to optimize energy consumption of the network. In this model, a sensor node can be in either sleep, partially active, or fully active mode. A cluster head knows about an object's existence and in the partially active mode, the cluster head can transmit an activation request to the node to switch it to the full active mode.

From the existing research above, it is shown that most research on data admission control is based on the sleep/active scheme for different metrics (energy saving, lower transmission delay, etc.). And for the existing proposed algorithms above, they usually assume that during the active status, a sensor node admits arrival data and processes data in this status and during the sleep status, it only performs idle listening. This kind of cycle is similar to the $(\gamma; \eta)$ algorithm [13], in which the duration of a sensor in a sleep mode is distributed exponentially with a mean of $1/\eta$ and the duration that, in an active status, is distributed exponentially with a mean of $1/\gamma$. For a queue server, the data admission control is quite important to obtain optimal reward during its service time. From the existing related works, it is evident that few of these consider the data admission control problem for

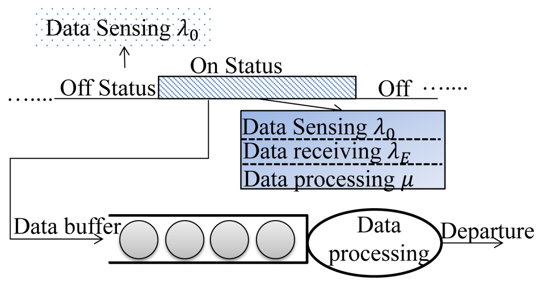

FIGURE 1. The sleep/active model of a sensor node.

an optimal reward for data processing during the sleep/active cycles.

In this paper, we investigate a model of admission control in sensor nodes with the MDP method. We consider a stochastic model of sensor nodes in which each node randomly and alternatively stays in the active or sleep mode. When a sensor node is in the active status, it can sense data packets, transmit packets, and receive packets. When the sensor is in the sleep status, it does not interact with the external world and only senses data packets. A reward is obtained when data from a sensor are received. At the same time, a storage cost also exists for data storing. With the stochastic model, we consider that the sensor node makes decision strategies on whether to access new data in the buffer for further processing, or reject the data to achieve optimal performance. Furthermore, for our proposed optimal admission control policy, we also investigate performance on energy consumption.

III. MODEL DEVELOPMENT

We consider a WSN in which static sensor nodes are randomly located in a given region. A sensor node's main purpose is to perform data gathering and then transmit the data to the next sensor node. There exists a storage cost for data storage, a switching on cost when a sensor is switched on, and a switching off cost when a sensor is switched off. In the meantime, a reward also exists for the admitted data. To study the optimal admission control problem with the sleep/active scheme, in this section, we present how to model this problem with the continuous time Markov decision process (CTMDP).

A. ASSUMPTIONS

An illustration of the working model of a sensor node is provided in Fig.1. In order to better describe the proposed model, the following assumptions and notations are introduced for the sensor node being investigated.

(1) The duration of a sensor in a sleep mode is distributed exponentially with a mean of $1/\eta$. In the sleep mode, the sensor node disconnects from the external world and there is no communication with other sensors. But the sensor can sense data with its sensing subsystem according to a Poisson process at a rate of λ_0. After the sleep duration, the sensor ends its sleep status and returns to the active status.

(2) The duration that a sensor spends in the active status is distributed exponentially with a mean of $1/\gamma$. During the active status, the sensor node may
   a) generate packets according to a Poisson process at a rate of λ_0,
   b) receive packets originating from other sensors in accordance to a Poisson process at a rate of λ_E, or
   c) process (transmit or relay) data packets with negative exponential distribution with a mean rate of μ.

   In the active status, the packets generate and relay packets originating from other sensors to fit Poisson distribution, respectively. Hence, we define these data in active status following a Poisson process with a rate of λ_1. It is clear that $\lambda_1 = \lambda_0 + \lambda_E$. When the sensor node ends its active status, it will return to the sleep status.

(3) For a sensor node, accepting a data packet would obtain R units of reward. In the meantime, once new data is generated in its buffer, the sojourn times in the node will result in storage costs. Let $f(i)$ be the storage cost rate per unit time with i data packet. $f(i)$ can be assumed to be positive, increasing, unbounded function. Meanwhile, transforming the sensor from the sleep status to the active status would generate E_1 units of *on* cost to the sensor and transforming the sensor from the active status to the sleep status would generate E_2 units of *off* cost to the sensor.

(4) The information sensed by a sensor node is organized into data units of fixed size that can be stored at the sensor in a buffer of infinite capacity; the buffer is modeled as a centralized FIFO queue. Sensor nodes cannot simultaneously transmit and receive. The wireless channel is assumed to be error-free, which is to say, if a data packet is transmitted, it will successfully arrive at its destination node.

B. OBJECTIVE FUNCTION

We now model the sensor nodes' behavior to accept or reject data packets as a stochastic dynamic program problem, in which a decision rule or a policy prescribing a procedure for action selection in each state at a specified decision epoch should be optimally determined. In general, a policy π is a sequence $\pi_1, \pi_2, \ldots$ of decision rules where π_i is the decision selected when the sensor is in the i state, which determines how to select an action after completion of the $(i-1)$th transition. A policy π is said to be *stationary*, if for each decision epoch, the decision is only related to its status. We denote $\pi(s)$ by the action to take when the system occupies state s. Given a policy π, denote the total expected infinite-horizon discounted reward when starting from state s by $v_\alpha^\pi(s)$, here $\alpha > 0$ is the discount factor so that a reward r received has present value $re^{-\alpha t}$ after some time t. With these descriptions in mind, our objective in this paper is to find an optimal policy π that can bring the maximum total expected discounted reward from a sensor for every initial state s. Mathematically speaking, our objective is to find a

policy π so as to maximize the following function:

$$v_\alpha^\pi(s) = E_s^\pi\left\{\int_0^\infty e^{-\alpha t} r(s_t, a_t)dt\right\}, \quad (1)$$

where s_t stands for the state at time t, a_t is the action to take at state s_t, and $r(s_t, a_t)$ is the total reward obtained when action a_t is selected at state s_t.

C. CTMDP FRAMEWORK FOR SENSOR'S BEHAVIOR

To find the optimal policy, we introduce a CTMDP which can be uniquely identified by the following five components: state space, action space, decision epochs, the reward function, and the transition probabilities.

- **State space:** At each decision epoch, the system occupies a state. We denote the set of possible states of a sensor by S. Let the state variable consist of the status of the sensor (sleep or active), the amount of data in the buffer and the most recent packet event on the sensor. So, state space

$$S = \{s : s = \langle\delta, i, b\rangle,\ i \geq 0\}.$$

Here, $\delta \in \{0, 1\}$ and $b \in \{A, D, C\}$; $\delta = 0$ denotes the sensor node is in the sleep state, $\delta = 1$ means the sensor is in the active status; i refers to the number of the data packets in the buffer (including the one currently in transmission); $b = A$ stands for the recent packet event is an arrival of a data, $b = D$ means the recent packet event is a departure of a data packet, and $b = C$ means the recent packet event is a switching of the sensor node.

- **Action space:** Action refers to things that the decision maker can do on any particular state s. In this model, no matter which status the sensor is, when the most recent event is an arrival of a new data packet, the node may take an action, denoted by a_A, to accept the arrival data, or take an action, denoted by a_R, to reject the new arrival data. If the most recent event is a process of a data packet or a switching of the sensor node, then the sensor takes an action, denoted by a_C, to continue. Thus, we have the action space for different states as follows:

$$A_{\langle 0,i,A\rangle} = A_{\langle 1,i,A\rangle} = \{a_A, a_R\},$$

and

$$A_{\langle 1,i,D\rangle} = A_{\langle 0,i,C\rangle} = A_{\langle 1,i,C\rangle} = \{a_C\}.$$

- **Decision epochs:** The decision epochs are those time points in which a new data packet has arrived, a data packet is processed, or a new status occurs (switching on or switching off). At each decision epoch, let $\tau(s, a)$ be the sojourn time starting from state s with action a. Therefore, based on the superposition property of exponential distributions, $\tau(s, a)$ will be an exponential random variable with a rate, say $\beta(s, a)$, and the probability that the next decision epoch occurs within t time units is give by

$$P(\tau(s, a) \leq t) = 1 - e^{-\beta(s,a)t}, \quad t \geq 0.$$

In this model, $\beta(s, a)$ can be written as,

$$\beta(s,a) = \begin{cases} \lambda_0 + \eta, & s = \langle 0, i, A\rangle,\ a{=}a_A \text{ or } a_R,\ i{\geq}0, \\ \lambda_1 + \mu + \gamma, & s = \langle 0, i, C\rangle,\ a = a_C,\ i > 0, \\ \lambda_1 + \gamma, & s = \langle 0, i, C\rangle,\ a = a_C,\ i = 0, \\ \lambda_1 + \mu + \gamma, & s = \langle 1, i, A\rangle,\ a = a_A,\ i \geq 0, \\ \lambda_1 + \gamma, & s = \langle 1, i, A\rangle,\ a = a_R,\ i = 0, \\ \lambda_1 + \mu + \gamma, & s = \langle 1, i, A\rangle,\ a = a_R,\ i > 0, \\ \lambda_1 + \mu + \gamma, & s = \langle 1, i, D\rangle,\ a = a_C,\ i > 1, \\ \lambda_1 + \gamma, & s = \langle 1, i, D\rangle,\ a = a_C,\ i = 1, \\ \lambda_0 + \eta, & s = \langle 1, i, C\rangle,\ a = a_C,\ i \geq 0. \end{cases}$$

- **Transition probability**: Let $q(m|s, a)$ denote the probability that the system occupies state m in the next epoch, if at the current epoch the system is at state s and the decision maker takes action $a \in A_s$. The function $q(m|s, a)$ is called a transition probability function, which should be specific for each problem and satisfies that $\sum_{m\in S} q(m|s, a) = 1$. In this model, we assume that when an active sensor node is switching off before a data packet is processed, the data packet will be put in the waiting space and be reprocessed whenever the sensor is on active status again. For any i, we define $\bar{i} = \max\{i - 1, 0\}$. The transition probability is as follows.

When the sensor is in the sleep status, $s = \langle 0, i, C\rangle$, where $i > 0$ and $a = a_C$, the transition probability is,

$$q(m|\langle 0, i, C\rangle, a_C) = \begin{cases} \dfrac{\lambda_1}{\lambda_1 + \mu + \gamma}, & m = \langle 1, i, A\rangle, \\ \dfrac{\mu}{\lambda_1 + \mu + \gamma}, & m = \langle 1, i, D\rangle, \\ \dfrac{\gamma}{\lambda_1 + \mu + \gamma}, & m = \langle 1, i, C\rangle, \end{cases}$$

and if $i = 0$, we have

$$q(m|\langle 0, i, C\rangle, a_C) = \begin{cases} \dfrac{\lambda_1}{\lambda_1 + \gamma}, & m = \langle 1, i, A\rangle, \\ \dfrac{\gamma}{\lambda_1 + \gamma}, & m = \langle 1, i, C\rangle. \end{cases}$$

For any $i > 0$, we have

$$q(m|\langle 1, i, D\rangle, a_C) = q(m|\langle 0, \bar{i}, C\rangle, a_C).$$

Since admitting incoming data migrates the system status immediately, we have that whenever $i \geq 0$ that

$$q(m|\langle 1, i, A\rangle, a_A) = q(m|\langle 0, i + 1, C\rangle, a_C);$$

and

$$q(m|\langle 1, i, A\rangle, a_R) = q(m|\langle 0, i, C\rangle, a_C).$$

Next, for state $s = \langle 1, i, C\rangle$, where $i \geq 0$, with action $a = a_C$, we have

$$q(m|\langle 1, i, C\rangle, a_C) = \begin{cases} \dfrac{\lambda_0}{\lambda_0 + \eta} & m = \langle 0, i, A\rangle, \\ \dfrac{\eta}{\lambda_0 + \eta} & m = \langle 0, i, C\rangle. \end{cases}$$

According to the property that admitting an incoming data migrates the system status immediately again, when $i \geq 0$, we have that

$$q(m|\langle 0, i, A\rangle, a_A) = q(m|\langle 1, i+1, C\rangle, a_C),$$

and

$$q(m|\langle 0, i, A\rangle, a_R) = q(m|\langle 1, i, C\rangle, a_C).$$

- **Reward functions:** Because the system status does not change between decision epochs, the expected discounted reward between epochs will be:

$$\begin{aligned} r(s,a) &= k\big(s,a\big) + c(s,a)E_s^a\left\{\int_0^{\tau(s,a)} e^{-\alpha t}dt\right\} \\ &= k\big(s,a\big) + \frac{c(s,a)}{\alpha+\beta(s,a)}, \end{aligned} \tag{2}$$

where $k\big(s,a\big)$ is the lump reward obtained for a state s when action a is taken, and $c\big(s,a\big)$ is the continuous cost accumulated between decision epochs. In this paper, the reward function consists of four different components: storage cost $f(i)$, reward R for data accepted, switching on cost E_1 per time from sleep status to active status, and switching off cost E_2 per time from active status to sleep status. According to equation (2), we will have

$$k(\langle\delta, i, b\rangle, a) = \begin{cases} R, & b=A,\ a=a_A,\ i\geq 0, \\ -E_1, & \delta=0,\ b{=}C,\ a{=}a_C,\ i\geq 0, \\ -E_2, & \delta=1,\ b{=}C,\ a{=}a_C,\ i\geq 0, \\ 0, & others; \end{cases}$$

and

$$c(\langle\delta, i, b\rangle, a) = \begin{cases} -f(i+1), & b=A,\ a=a_A,\ i\geq 0, \\ -f(\bar{i}), & b=D,\ a=a_C,\ i\geq 1, \\ -f(i), & others. \end{cases}$$

A policy is said to be α-optimal if its expected α-discounted return is maximal for every initial state. And during each decision epoch, the status is stationary. The optimal policy is a stationary deterministic policy. From equation (1), it is easy to realize that the calculation for $v_\alpha^\pi(s)$ could be simplified if $\beta(s,a)$ is a constant for all states s. This is the idea of rate uniformization technique [14]. Based on this idea, we may define the uniformization of our process with corresponding components denoted by notation $\sim$. We will have the process remain in each state for an exponential amount of time with a constant rate $\beta = \lambda_1+\mu+\gamma+\eta$ and the new transition probabilities will be defined as

$$\widetilde{q}(m|s,a) = \begin{cases} \dfrac{q(m|s,a)\beta(s,a)}{\beta}, & m\neq s, \\ 1-\dfrac{[1-q(s|s,a)]\beta(s,a)}{\beta}, & m=s. \end{cases} \tag{3}$$

Based on this, we obtain the following results.

For status $s = \langle 0, i, C\rangle$, when $i > 0$,

$$\widetilde{q}(m|s,a_C) = \begin{cases} \lambda_1/\beta, & m=\langle 1, i, A\rangle, \\ \mu/\beta, & m=\langle 1, i, D\rangle, \\ \gamma/\beta, & m=\langle 1, i, C\rangle, \\ \eta/\beta, & m=s; \end{cases}$$

when $i = 0$,

$$\widetilde{q}(m|s,a_C) = \begin{cases} \lambda_1/\beta, & m=\langle 1, i, A\rangle, \\ \gamma/\beta, & m=\langle 1, i, C\rangle, \\ (\mu+\eta)/\beta, & m=s. \end{cases}$$

For status $s = \langle 1, i, C\rangle$ in which $i \geq 0$ and action a_C, we have

$$\widetilde{q}(m|s,a_C) = \begin{cases} \lambda_0/\beta, & m=\langle 0, i, A\rangle, \\ \eta/\beta, & m=\langle 0, i, C\rangle, \\ (\gamma+\lambda_E+\mu)/\beta, & m=s. \end{cases}$$

From [15], the objective function in equation (1) can now be replaced by the following Bellman equation:

$$v_\alpha(s) = \max_{a\in A}\left\{\widetilde{r}(s,a) + \frac{\beta}{\alpha+\beta}\sum_{m\in S}\widetilde{q}(m|s,a)v(m)\right\}, \tag{4}$$

where

$$\widetilde{r}(s,a) \equiv r(s,a)\frac{\alpha+\beta(s,a)}{\alpha+\beta}.$$

By noting there is only one possible action a_C for states $s = \langle 0, i, C\rangle$ and $s = \langle 1, i, C\rangle$, where $i \geq 0$. Also, there is only one possible action a_C for status $s = \langle 1, i, D\rangle$, with $i > 0$.

From the analysis above, when $i > 0$, for state $s = \langle 0, i, C\rangle$, we have

$$\begin{aligned} &v\big(\langle 0, i, C\rangle\big) \\ &\quad = -E_1 + \frac{1}{\alpha+\lambda_1+\mu+\gamma}\Big\{-f(i)+\lambda_1 v\big(\langle 1, i, A\rangle\big) \\ &\qquad + \mu v\big(\langle 1, i, D\rangle\big) + \gamma v\big(\langle 1, i, C\rangle\big)\Big\}, \end{aligned} \tag{5}$$

and when $i = 0$,

$$\begin{aligned} v\big(\langle 0, i, C\rangle\big) = -E_1 + \frac{1}{\alpha+\lambda_1+\gamma}\Big\{-f(i)+\lambda_1 v\big(\langle 1, i, A\rangle\big) \\ + \gamma v\big(\langle 1, i, C\rangle\big)\Big\}. \end{aligned} \tag{6}$$

For status $v\big(\langle 1, i, D\rangle\big)$ with $i > 0$, we obtain that

$$v\big(\langle 1, i, D\rangle\big) = v\big(\langle 0, \bar{i}, C\rangle\big) + E_1. \tag{7}$$

For status $s = \langle 1, i, C\rangle$, where $i \geq 0$, we get

$$\begin{aligned} v\big(\langle 1, i, C\rangle\big) = -E_2 + \frac{1}{\alpha+\lambda_0+\eta}\Big\{-f(i)+\lambda_0 v\big(\langle 0, i, A\rangle\big) \\ + \eta v\big(\langle 0, i, C\rangle\big)\Big\}. \end{aligned} \tag{8}$$

With each different status, at every decision epoch, the best decision should be made to obtain optimal benefits. Next, we study the optimal admission control policy for this sleep/active model.

IV. OPTIMAL POLICY

Admitting a data packet can reap a reward while holding a data packet in the sensor node will contribute to a storage cost. Therefore, in this section, we study the optimal admission policy of the sleep/active model. Specifically speaking, for each different status, we study the optimal policy of when to admit or reject an arriving data packet.

According to the analysis above, for state $\langle 1, i, A\rangle$, when $i > 0$, we have,

$$\begin{aligned} &v\big(\langle 1, i, A\rangle, a_R\big)\\ &\quad= \frac{1}{\alpha+\beta}\Big\{-f(i)+\lambda_1\, v\big(\langle 1, i, A\rangle\big)+\mu v\big(\langle 1, i, D\rangle\big)\\ &\qquad +\gamma v\big(\langle 1, i, C\rangle\big)+\eta v\big(\langle 1, i, A\rangle\big)\Big\};\end{aligned}$$

and when $i = 0$,

$$\begin{aligned} &v\big(\langle 1, i, A\rangle, a_R\big)\\ &\quad= \frac{1}{\alpha+\beta}\Big\{-f(i)+\lambda_1\, v\big(\langle 1, i, A\rangle\big)+\gamma v\big(\langle 1, i, C\rangle\big)\\ &\qquad +(\mu+\eta)v\big(\langle 1, i, A\rangle\big)\Big\}.\end{aligned}$$

From these results, we can easily know in general that for $i > 0$,

$$\begin{aligned} &v\big(\langle 1, i, A\rangle, a_R\big)\\ &\quad\geq \frac{1}{\alpha+\lambda_1+\mu+\gamma}\Big\{-f(i)+\lambda_1\, v\big(\langle 1, i, A\rangle\big)\\ &\qquad +\mu v\big(\langle 1, i, D\rangle\big)+\gamma v\big(\langle 1, i, C\rangle\big)\Big\}\\ &\quad= v\big(\langle 0, i, C\rangle\big)+E_1;\end{aligned}$$

and for $i = 0$,

$$\begin{aligned} &v\big(\langle 1, i, A\rangle, a_R\big)\\ &\geq \frac{1}{\alpha+\lambda_1+\gamma}\Big\{-f(i)+\lambda_1\, v\big(\langle 1, i, A\rangle\big)+\gamma v\big(\langle 1, i, C\rangle\big)\Big\}\\ &= v\big(\langle 0, i, C\rangle\big)+E_1.\end{aligned}$$

Furthermore, if action a_R is the best action at the state $\langle 1, i, A\rangle$, where $i \geq 0$, we have

$$v(\langle 1, i, A\rangle, a_R) = v(\langle 0, i, C\rangle) + E_1. \tag{9}$$

Next, for $i \geq 0$, if an action a_A is taken on the state $\langle 1, i, A\rangle$, by conducting a similar analysis as above, we will know in general that

$$\begin{aligned} &v\big(\langle 1, i, A\rangle, a_A\big)\\ &= \frac{\alpha+\lambda_1+\mu+\gamma}{\alpha+\beta}R+\frac{1}{\alpha+\beta}\Big\{-f(i+1)+\lambda_1\, v\big(\langle 1, i+1, A\rangle\big)\\ &\qquad +\gamma v\big(\langle 1, i+1, C\rangle\big)+\mu v\big(\langle 1, i+1, D\rangle\big)+\eta v\big(\langle 1, i, A\rangle\big)\Big\}\\ &\geq R+\frac{1}{\alpha+\lambda_1+\mu+\gamma}\Big\{-f(i+1)+\lambda_1\, v\big(\langle 1, i+1, A\rangle\big)\\ &\qquad +\gamma v\big(\langle 1, i+1, C\rangle\big)+\mu v\big(\langle 1, i+1, D\rangle\big)\Big\}\\ &= v\big(\langle 0, i+1, C\rangle\big)+E_1+R.\end{aligned}$$

And if action a_A is the best action at the state $\langle 1, i, A\rangle$,

$$v(\langle 1, i, A\rangle, a_A) = v(\langle 0, i+1, C\rangle) + E_1 + R. \tag{10}$$

From the above analysis, it is not difficult to verify that for $i \geq 0$,

$$v\big(\langle 1, i, A\rangle\big) = E_1 + \max\{v\big(\langle 0, i+1, C\rangle\big)+R,\, v\big(\langle 0, i, C\rangle\big)\}. \tag{11}$$

Similarly, for any $i \geq 0$ with status $s = \langle 0, i, A\rangle$, we get that in general,

$$v\big(\langle 0, i, A\rangle, a_R\big) \geq v\big(\langle 1, i, C\rangle\big) + E_2, \tag{12}$$

and if action a_R is the best action at the state $\langle 0, i, A\rangle$,

$$v\big(\langle 0, i, A\rangle, a_R\big) = v\big(\langle 1, i, C\rangle\big) + E_2. \tag{13}$$

Furthermore, we have in general,

$$v\big(\langle 0, i, A\rangle, a_A\big) \geq v\big(\langle 1, i+1, C\rangle\big) + R + E_2, \tag{14}$$

and if action a_A is the best action at the state $\langle 0, i, A\rangle$,

$$v\big(\langle 0, i, A\rangle, a_A\big) = v\big(\langle 1, i+1, C\rangle\big) + R + E_2. \tag{15}$$

Finally, we will also have the following result, which is similar as that in equation (11), that

$$v\big(\langle 0, i, A\rangle\big) = E_2 + \max\{v\big(\langle 1, i+1, C\rangle\big)+R,\, v\big(\langle 1, i, C\rangle\big)\}. \tag{16}$$

Definition 1: For any $i \geq 0$, a discrete function $f(i)$ is convex on i if

$$f(i+1) - f(i) \geq f(i) - f(i-1),$$

and is concave on i if

$$f(i+1) - f(i) \leq f(i) - f(i-1).$$

Lemma 1: If both $v\big(\langle 1, i, C\rangle\big)$ and $v\big(\langle 0, i, C\rangle\big)$ are concave and non-increasing on i with $i \geq 0$, the optimal policy for the data admission control scheme is a control limit policy. Specifically, there exist two integers, say M and N, such that decision

$$d\big(\langle 1, i, A\rangle\big) = \begin{cases} a_A, & \text{if } i \leq M,\\ a_R, & \text{if } i > M;\end{cases} \tag{17}$$

and

$$d\big(\langle 0, i, A\rangle\big) = \begin{cases} a_A, & \text{if } i \leq N;\\ a_R, & \text{if } i > N.\end{cases} \tag{18}$$

Proof: From Theorem 11.5.3 in [15], we can determine that the optimal policy of the scheme is a stationary deterministic policy. Thus, our problem can then be reduced to finding a deterministic decision rule d. For any $i \geq 0$, let

$$\Delta v_{C1}(i) = v(\langle 1, i+1, C\rangle) - v(\langle 1, i, C\rangle),$$

and

$$\Delta v_{C0}(i) = v(\langle 0, i+1, C\rangle) - v(\langle 0, i, C\rangle).$$

We know that both $\Delta v_{C0}(i) \leq 0$, $\Delta v_{C1}(i) \leq 0$ and both of them are non-increasing. For states $\langle 1, i, A\rangle$ and $\langle 0, i, A\rangle$, whenever $i \geq 0$, we have the decision rule

$$d(\langle 1, i, A\rangle) = \begin{cases} a_A, & \Delta v_{C0}(i) > -R, \\ a_R, & \Delta v_{C0}(i) \leq -R, \end{cases}$$

and

$$d(\langle 0, i, A\rangle) = \begin{cases} a_A, & \Delta v_{C1}(i) > -R, \\ a_R, & \Delta v_{C1}(i) \leq -R. \end{cases}$$

Therefore, if $d(\langle 1, M, A\rangle) = a_R$, for $i > M$, we have $d(\langle 1, i, A\rangle) = a_R$. Similarly for states $\langle 0, i, A\rangle$, if $d(\langle 0, N, A\rangle) = a_R$, for $i > N$, we have $d(\langle 0, i, A\rangle) = a_R$. Therefore, the optimal policy for the admission control scheme in sensor nodes is a control limit policy.

Lemma 2 [16]: Suppose that $f(i)$ is convex and nondecreasing on i, $i \geq 0$, and let $R \geq 0$ be a constant; then

$$g(i) \equiv max\{-f(i),\ R - f(i+1)\},$$

is concave and non-increasing.

Theorem 1: Suppose that the cost function $f(i)$ is convex and nondecreasing on i ($i \geq 0$), then the data admission control scheme in wireless sensor nodes is a control limit policy.

Proof: We use value iteration method to verify it. We have the following analysis:

(1). Set $j = 0$, $v^j(s) = 0$ and substitute this into equations (5), (6) and (8). We can obtain that for any $i > 0$,

$$v^1(\langle 0, i, C\rangle) = -E_1 - \frac{f(i)}{\alpha + \lambda_1 + \mu + \gamma},$$

and

$$v^1(\langle 0, 0, C\rangle) = -E_1 - \frac{f(i)}{\alpha + \lambda_1 + \gamma}.$$

For $i \geq 0$,

$$v^1(\langle 1, i, C\rangle) = -E_2 - \frac{f(i)}{\alpha + \lambda_0 + \eta}.$$

For $i \geq 0$, $f(i)$ is convex and non-decreasing, thus, we obtain that both $v^1(\langle 1, i, C\rangle)$ and $v^1(\langle 0, i, C\rangle)$ are concave and non-increasing on i where $i \geq 0$.

(2). Set $j = j + 1$ and by using these results in equations (7), (11), and (16), as well as Lemma 2, we know that $v^j(\langle 0, i, C\rangle)$ and $v^j(\langle 1, i, C\rangle)$ are concave and non-increasing on i, where $i \geq 0$.

(3). As the iteration continues, i.e., as j goes to $+\infty$, $v^j(\langle 0, i, C\rangle)$ will converge to $v(\langle 0, i, C\rangle)$ and $v^j(\langle 1, i, C\rangle)$ will converge to $v(\langle 1, i, C\rangle)$, where $i \geq 0$. Therefore, $v(\langle 0, i, C\rangle)$ and $v(\langle 1, i, C\rangle)$ are concave and non-increasing on i, where $i \geq 0$. From Theorem 1, the optimal policy for the proposed model must be a control limit policy. The verification is completed.

From the analysis above, we can determine that, supposing that the cost function $f(i)$ is convex and nondecreasing on i, $i \geq 0$, for the data admission control scheme in wireless sensor nodes, there exists a $(M; N)$ policy. That means that when the sensor is in the sleep status, the arrival data will be rejected if the number of data packets is equal to or more than N; when the sensor is in the active status, the arrival data will be rejected if the number of data packets is equal to or more than M.

The optimal strategy has been verified to be a control limit policy. However, how to find the optimal threshold value is always a challenging problem in the subject. The traditional way of finding it is to use the simulation or numerical analysis based upon the specific values of the inputs. There are no theoretical results on this yet. We reach this value identification theoretically and have obtained the following bound results.

Theorem 2: Under the condition in Theorem 1, if denoted by $\Delta f(n) = f(n) - f(\overline{n})$, we will have

- When the sensor is in the sleep status, the arrival data will be rejected if the amount of data is more than the optimal threshold N, where N is bounded by N^*, that is $N \leq N^*$ and

$$N^* = \max\left\{n : \Delta f(n) \leq \frac{(\alpha + \lambda_1 + \mu + \gamma)(\alpha + \eta)R}{\alpha + \beta}\right\}.$$

- When the sensor is in the active status, the arrival data will be rejected if the amount of data is more than the optimal threshold M, where M is bounded by M^*, that is $M \leq M^*$ and

$$M^* = \max\{m : \Delta f(m) \leq (\alpha + \mu + \gamma)R\}.$$

Proof: From **Theorem 1** above, we can obtain that in the sleep status, when $i \geq N$, the sensor should reject arrival data. So, we have,

$$\begin{cases} v(\langle 1, N+1, C\rangle) - v(\langle 1, N, C\rangle) \leq -R, \\ v(\langle 1, N, C\rangle) - v(\langle 1, \overline{N}, C\rangle) > -R. \end{cases}$$

Thus, by further noting the equation (13) and (15) with the corresponding conditions, we have

$$\begin{aligned} &v(\langle 1, N+1, C\rangle) - v(\langle 1, N, C\rangle) \\ &= \frac{f(N) - f(N+1) + \eta\left[v(\langle 0, N+1, C\rangle) - v(\langle 0, N, C\rangle)\right]}{\alpha + \eta} \\ &\leq -R, \end{aligned}$$

and

$$v(\langle 1,N,C\rangle)-v(\langle 1,\overline{N},C\rangle) = \frac{f(\overline{N})-f(N)-\lambda_0 R+\eta\left[v(\langle 0,N,C\rangle)-v(\langle 0,\overline{N},C\rangle)\right]}{\alpha+\lambda_0+\eta} > -R.$$

Furthermore, we obtain

$$\begin{cases} v(\langle 0,N+1,C\rangle)-v(\langle 0,N,C\rangle) \\ \qquad \leq \dfrac{-R(\alpha+\eta)+f(N+1)-f(N)}{\eta}, \\ v(\langle 0,N,C\rangle)-v(\langle 0,\overline{N},C\rangle) \\ \qquad > \dfrac{-R(\alpha+\eta)+f(N)-f(\overline{N})}{\eta}. \end{cases}$$

According to equations (5)-(6), through mathematical manipulations, for any $i\geq 0$, we get

$$\begin{aligned} v(\langle 0,i,C\rangle)-v(\langle 0,\bar{i},C\rangle) &= \frac{1}{\alpha+\lambda_1+\mu+\gamma}\Big\{f(\bar{i})-f(i)+\lambda_1(v(\langle 1,i,A\rangle) \\ &\quad -v(\langle 1,\bar{i},A\rangle))+\mu(v(\langle 0,\bar{i},C\rangle)-v(\langle 0,\overline{i-1},C\rangle)) \\ &\quad +\gamma(v(\langle 1,i,C\rangle)-v(\langle 1,\bar{i},C\rangle)\Big\}. \end{aligned} \tag{19}$$

By noting that for all $i\geq 0$, $v(\langle 1,i,A\rangle)$, $v(\langle 0,i,C\rangle)$ and $v(\langle 1,i,C\rangle)$ are concave and non-increasing on i, we have

$$v(\langle 0,i,C\rangle)-v(\langle 0,\bar{i},C\rangle) \leq \frac{f(\bar{i})-f(i)}{\alpha+\lambda_1+\mu+\gamma}. \tag{20}$$

Therefore, we have

$$\begin{cases} v(\langle 0,N+1,C\rangle)-v(\langle 0,N,C\rangle) \leq \dfrac{f(N)-f(N+1)}{\alpha+\lambda_1+\mu+\gamma}, \\ v(\langle 0,N,C\rangle)-v(\langle 0,\overline{N},C\rangle) \leq \dfrac{f(\overline{N})-f(N)}{\alpha+\lambda_1+\mu+\gamma}. \end{cases}$$

Finally, we obtain that

$$\frac{-R(\alpha+\eta)+f(N)-f(\overline{N})}{\eta} < \frac{f(\overline{N})-f(N)}{\alpha+\lambda_1+\mu+\gamma}.$$

That means,

$$f(N)-f(\overline{N}) \leq \frac{(\alpha+\lambda_1+\mu+\gamma)(\alpha+\eta)}{\alpha+\beta}R.$$

By noting the monotonic non-decreasing property of $f(i)-f(i-1)$ because of the convexity property of function $f(i)$, the upper bound result in equation (21) is now verified.

In regards to the bound value of optimal threshold M, similar as above analysis, we have

$$\begin{cases} v(\langle 0,M+1,C\rangle)-v(\langle 0,M,C\rangle) \leq -R, \\ v(\langle 0,M,C\rangle)-v(\langle 0,\overline{M},C\rangle) > -R. \end{cases}$$

With equation (19), we get

$$v(\langle 0,M,C\rangle)-v(\langle 0,\overline{M},C\rangle) \leq \frac{f(\overline{M})-f(M)-\lambda_1 R}{\alpha+\lambda_1+\mu+\gamma}.$$

Thus, we obtain

$$-R < \frac{f(\overline{M})-f(M)-\lambda_1 R}{\alpha+\lambda_1+\mu+\gamma}.$$

That means,

$$f(M)-f(\overline{M}) \leq (\alpha+\mu+\gamma)R.$$

The bound result is equation (23) is finally verified.

Remark 1: From Theorem 2, it is straightforward that

- The optimal threshold value $N=0$ when
$$\frac{(\alpha+\lambda_1+\mu+\gamma)(\alpha+\eta)}{\alpha+\beta}R \leq f(1)-f(0).$$
- The optimal threshold value $M=0$ when
$$(\alpha+\mu+\gamma)R \leq f(1)-f(0).$$

Similar to the analysis of the above Theorem 2, we will have the following more closed up bound for the Optimal threshold M and N.

Corollary 1: Under the condition in Theorem 1,

- When the sensor is in the sleep status, the arrival data will be rejected if the amount of data is more than optimal threshold N, where N is bounded by N^{**}, i.e., $N\leq N^{**}$ and
$$N^{**} = \max\Big\{n : \sum_{k=0}^{\bar{n}} b_k[f(n-k)-f(\bar{n}-k)] < K_0\Big\}, \tag{21}$$
where the constants K_0 and b_k $(k=0,1,2,\ldots,\bar{n})$ are given by
$$\begin{aligned} b_0 &= \frac{\alpha+\beta}{\eta}; \\ b_k &= \frac{\mu^k(\alpha+\eta)^{k-1}(\alpha+\eta+\gamma)}{[(\alpha+\lambda_1+\mu+\gamma)(\alpha+\eta)-\gamma\eta]^k}; \\ K_0 &= \frac{(\alpha+\lambda_1+\mu+\gamma)(\alpha+\lambda_0+\eta)-\gamma\eta}{(\alpha+\lambda_0+\eta)\eta}(\alpha+\eta)R \\ &\quad -\frac{\gamma\lambda_0}{\alpha+\lambda_0+\eta}R. \end{aligned} \tag{22}$$
- When the sensor is in the active status, the arrival data will be rejected if the amount of data is more than the optimal threshold M, where M is bounded by M^{**}, that is $M\leq M^{**}$ and
$$M^{**} = \max\Big\{m : \sum_{k=0}^{\overline{m}} d_k[f(m-k)-f(\overline{m}-k)] < K_1\Big\}, \tag{23}$$
where the constants K_1 and d_k $(k=0,1,2,\ldots,\overline{m})$ are given by
$$\begin{aligned} d_k &= \frac{\mu^k(\alpha+\lambda_0+\eta)^k(\alpha+\lambda_0+\eta+\gamma)}{[(\alpha+\mu+\gamma)(\alpha+\lambda_0+\eta)-\gamma\eta]^k}; \\ K_1 &= [(\alpha+\mu+\gamma)(\alpha+\lambda_0+\eta)-\gamma\eta]R. \end{aligned} \tag{24}$$

Remark 2: It is straightforward that $M^{**} \leq M^*$ and $N^{**} \leq N^*$.

Remark 3: From the **Corollary 1**, it is readily known that

- The optimal threshold value $N = 0$ when
$$\frac{(\alpha + \lambda_1 + \mu + \gamma - \frac{\gamma\eta}{\alpha+\lambda_0+\eta})(\alpha + \eta)}{\alpha + \beta} R \leq f(1) - f(0).$$
- The optimal threshold value $M = 0$ when
$$\frac{(\alpha + \lambda_1 + \mu + \gamma)(\alpha + \lambda_0 + \eta) - \gamma\eta}{\alpha + \lambda_0 + \eta + \gamma} R \leq f(1) - f(0).$$

V. ENERGY CONSUMPTION ANALYSIS

Power consumption analysis is very important in wireless sensor networks. In this section, based on the results above, we study the energy performance. According to [13], we consider the energy consumption in terms of the sensor status, number of packets transmitted, and the switches from one status to another. We provide the following definitions for power consumption.

e_{sr}: the power consumption when the sensor switches from the sleep status to the active status;

e_{rs}: the power consumption when the sensor switches from the active status to the sleep status;

e_{tr} the transmitter power consumption per data packet in the active status;

e_s: the operation power consumption per unit time in the sleep status.

In this section, we derive the formula of the steady-state probability of the node when there are $i (i \geq 0)$ packets (including the one being processed and the others that are waiting) in the sensor node. According to the analysis above, there are three different state diagrams for the proposed optimal admission control scheme, $N = M, N > M, N < M$. Here, we denote

- $P(R_i)$ as the steady-state probability of the node when there are i $(i \geq 0)$ data packets in the referenced sensor node, which is in phase R of the active mode.
- $P(S_i)$ as the steady-state probability of the node when there are i $(i \geq 0)$ data packets in the referenced sensor node, which is in phase S of the sleep mode.
- For the case $N = M$, we denote by
$$\pi_{R_i} = P(R_i),\ \ \pi_{S_i} = P(S_i),\ and\ \pi_i = (\pi_{S_i}, \pi_{R_i})$$
for $i = 0, 1, 2, 3, \cdots, N$.
- For the case $N > M$, we denote by
$$\pi_{R_i}^N = P(R_i),\ \ \pi_{S_i}^N = P(S_i),\ and\ \pi_i^N = (\pi_{S_i}^N, \pi_{R_i}^N)$$
for $i = 0, 1, 2, 3, \cdots, N$.
- For the case $N < M$, we denote by
$$\pi_{R_i}^M = P(R_i),\ \ \pi_{S_i}^M = P(S_i),\ \ and\ \pi_i^M = (\pi_{S_i}^M, \pi_{R_i}^M)$$
for $i = 0, 1, 2, 3, \cdots, M$.

As long as the formula of the steady-state probability is derived, it is not difficult to determine various energy consumption measures of the sensor node. Here, some results are listed to demonstrate how to utilize this formula to obtain the sensor node's performance measures.

Theorem 3: The average energy consumed per unit time switching from the sleep status to the active status is $\frac{\gamma\eta}{\gamma+\eta} e_{sr}$ and the average energy consumed per unit time switching from the active status to the sleep status is $\frac{\gamma\eta}{\gamma+\eta} e_{rs}$.

Proof: The sensor consumes e_{sr} milliwatts of power each time it switches from the sleep status to the active status. The expected number of switching from the sleep mode to the active mode per unit time is $\sum_{i=0}^{\max\{N,M\}} P(S_i)\eta$. Thus, we have
$$E_{SR} = \sum_{i=0}^{\max\{N,M\}} P(S_i)\eta e_{sr} = \frac{\gamma\eta}{\gamma + \eta} e_{sr}.$$

The sensor consumes e_{rs} milliwatts of power each time it switches from the active status to the sleep status. The expected number of switching from the active mode to the sleep mode per unit time is $\sum_{i=0}^{\max\{N,M\}} P(R_i)\gamma$. Therefore,
$$E_{RS} = \sum_{i=0}^{\max\{N,M\}} P(R_i)\gamma e_{rs} = \frac{\gamma\eta}{\gamma + \eta} e_{rs}.$$

The proof is finished.

Remark 4: From Theorem 3, we can determine that with the increase of γ, η, the average switching energy consumed per unit time also increases.

Theorem 4: The average energy consumption in the active mode E_{TR} is provided by
$$E_{TR} = \begin{cases} \sum_{i=0}^{N} i\pi_{R_i} e_{tr}, & \text{if } N = M; \\ \sum_{i=0}^{M} i\pi_{R_i}^M e_{tr}, & \text{if } N < M; \\ \sum_{i=0}^{N} i\pi_{R_i}^N e_{tr}, & \text{if } N > M; \end{cases}$$
and the average energy consumption in the sleep mode E_S is provided as follows.
$$E_S = \begin{cases} \sum_{i=0}^{N} i\pi_{S_i} e_s, & \text{if } N = M; \\ \sum_{i=0}^{M} i\pi_{S_i}^M e_s, & \text{if } N < M; \\ \sum_{i=0}^{N} i\pi_{S_i}^N e_s, & \text{if } N > M. \end{cases}$$

Proof: When the sensor node is in the active status, it would consume e_{tr} milliwatts of power for per packet transmitted. The average energy consumption in the active mode E_{TR} should be $\sum_{i=0}^{\max\{N,M\}} iP(R_i)e_{tr}$. When the sensor node is in the sleep status, its only operation is data sensing. It would consume e_s milliwatts of power for per packet sensing. The expected number of data packets in the sleep status is $\sum_{i=0}^{\max\{N,M\}} iP(S_i)$. Therefore, firstly, we need to know the probability of status R_i and S_i.

Fig.2 shows the transition rate diagram with $N = M$. We denote by
$$A_0 = \begin{bmatrix} \lambda_0 & 0 \\ 0 & \lambda_1 \end{bmatrix},\quad A_1 = \begin{bmatrix} -(\eta + \lambda_0) & \eta \\ \gamma & -(\mu + \gamma + \lambda_1) \end{bmatrix},$$

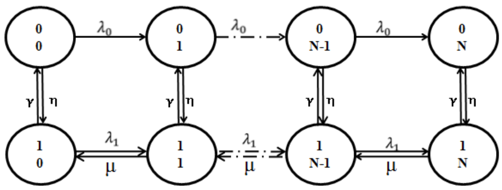

FIGURE 2. Transition rate diagram with $N = M$.

$$A_2 = \begin{bmatrix} 0 & 0 \\ 0 & \mu \end{bmatrix}, \quad A_{N_0} = \begin{bmatrix} -\eta & \eta \\ \gamma & -(\mu+\gamma) \end{bmatrix},$$
$$B_0 = \begin{bmatrix} -(\eta+\lambda_0) & \eta \\ \gamma & -(\gamma+\lambda_1) \end{bmatrix}.$$

For Fig.2, the corresponding transition rate matrix Q of the constructed multi-dimensional Markov process can be given by

$$Q = \begin{bmatrix} B_0 & A_0 & 0 & \cdots & 0 & 0 & 0 \\ A_2 & A_1 & A_0 & \cdots & 0 & 0 & 0 \\ 0 & A_2 & A_1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_2 & A_1 & A_0 \\ 0 & 0 & 0 & \cdots & 0 & A_2 & A_{N_0} \end{bmatrix}.$$

According to Theorem 1 on page 720 in [17], we have,

$$\begin{cases} \pi_N C_N = 0; \\ \pi_i = \pi_{i+1} A_2(-C_i^{-1}) \quad \text{for } 0 \le i \le N-1; \end{cases}$$

with the normalization condition $\sum_{i=0}^{N} \pi_i e = 1$, where

$$C_i = \begin{cases} B_0, & \textit{if } i = 0; \\ A_1 + A_2(-C_{i-1}^{-1})A_0, & \textit{if } 1 \le i \le N-1; \\ A_{N_0} + A_2(-C_{i-1}^{-1})A_0, & \textit{if } i = N. \end{cases}$$

And e is a two-dimensional column vector with all its component of 1, i.e., $e = (1, 1)^T$. Finally, we have

$$E_{TR} = \sum_{i=0}^{N} i\pi_{R_i} e_{tr},$$

and

$$E_S = \sum_{i=1}^{N} i\pi_{S_i} e_s.$$

When $N > M$ and $N < M$, the proof of the results on the average energy consumption E_{TR} in the active mode and E_S in the sleep mode adheres to the same method as that of $N = M$, which can be seen in the APPENDIX.

Remark 5: In the special case when $M = N$ and $\lambda_0 = \lambda_1$, our above probability matrix form can be deducted to a scale closed form similar as that in paper [18]. However, our current uniform matrix representations in both results of probability formula and energy consumption formula would be more easy in computational implementation for all three cases for $M = N$, $M > N$, and $M < N$, as well as the general condition that $\lambda_0 \neq \lambda_1$.

TABLE 1. Simulation parameters of common information.

parameters	λ_1	λ_0	μ	γ	η	R	E_1	E_2
values	0.3	0.1	0.5	0.05	0.1	200	4	1

VI. NUMERICAL ANALYSIS

In this section, we investigate the optimal policy of the proposed model with the traditional duty cycling scheme (the $(\gamma; \eta)$ model) which is similar to [13]. In the $(\gamma; \eta)$ model, on sleep status, the sensor node can only sense data, but on active status, it can sense data, receive data from other nodes and process data. We focus on the values of total expected discounted reward $v(s)$ and show the effects on energy consumption under the optimal $(M; N)$ policy and $(\gamma; \eta)$ policy. Let the cost function be $f(i) = i^2$, $i \ge 0$, and discount factor $\alpha = 0.1$, which fits the theorems' requirements.

A. DISCOUNTED REWARDS PERFORMANCE

To evaluate the average discounted reward with these two different policies, we set the parameters for analysis as in TABLE 1. As shown in TABLE 1, we assume the arrival rates on sleep status and active status are smaller than the service rate, which represents a light traffic load. Algorithm 1 illustrates the steps to perform $(M^{**}; N^{**})$ policy.

Algorithm 1 The Admission Control Algorithm for an Optimal Discounted Reward

set i as the number of data in the sensor node (containing the processing one)
set $R, E_1, E_2, f(i)$ /*set the reward value and cost value in sensor node*/
set $\delta = 0$ as sleep, $\delta = 1$ as active
initialize $\lambda_0, \lambda_1, \mu, \alpha, \beta, \eta, \gamma$ /*the sensor node obtains the real-time features of itself */
calculate $(M^{**}; N^{**})$ according to Corollary 1
if $\delta = 0, i \ge M^{**}$ **then**
    the sensor node refuses arrival data
else
    if $\delta = 1, i \ge N^{**}$ **then**
        the sensor node refuses arrival data
    else
        continue
    end if
end if

With the set parameters, according to the theorems above, we can obtain the following numerical results. Fig.3 illustrates the discounted reward of different status with $(M; N)$ policy. From Fig.3, we can determine that the discounted reward decreases as the number of data increases which fits our theoretical conclusion of Theorem 1. We can determine when the number of data i is less than 7, the discounted reward of $v(\langle 1, i, A\rangle)$ is the largest among all the discounted rewards. And when i is larger than 7, the discounted reward of $v(\langle 1, i, D\rangle)$ is the largest. In Fig.4, $v_0(i) = v(\langle 1, i, C\rangle) - v(\langle 0, i, A\rangle)$, and $v_1(i) = v(\langle 0, i, C\rangle) - v(\langle 1, i, A\rangle)$. From Fig.4, it is clear to see that in the sleep status, when the number of

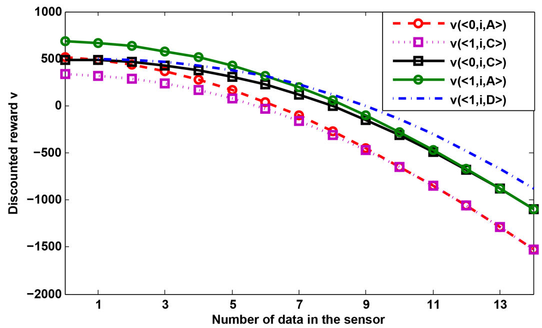

FIGURE 3. Discounted reward on different status.

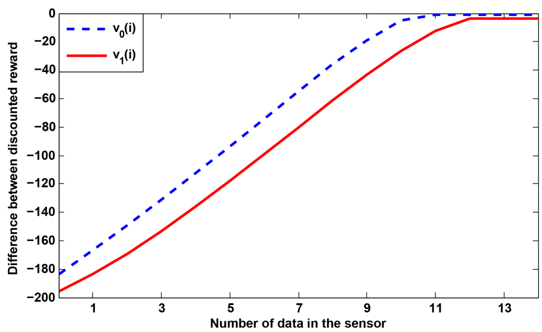

FIGURE 4. Difference between discounted reward.

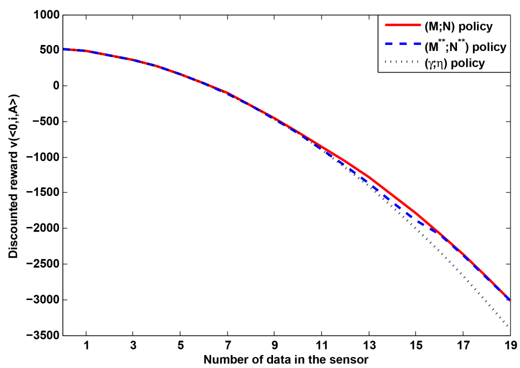

FIGURE 5. Discounted reward with $\langle 0, i, A\rangle$.

data is more than 10, the sensor should reject the arrival data. And when the number of data is more than 11, the sensor should reject the arrival data in the active status. Therefore, the optimal policy in this case is $(M; N) = (12; 11)$. According to Corollary 1 above, we can determine that $(M^{**}; N^{**}) = (19; 17)$. Fig.5-Fig.7 illustrate the discounted reward of status $\langle 0, i, A\rangle$, $\langle 1, i, A\rangle$ and $\langle 1, i, D\rangle$ with three different policies, $(M; N)$ policy, $(M^{**}; N^{**})$ policy and $(\gamma; \eta)$ policy. From these figures, it is clear to see that the discounted reward of $(M; N)$ policy is the largest among these three policies. The discounted reward of $(M; N)$ policy is not excessively larger than that of $(M^{**}; N^{**})$ policy. Specifically, when the number of data is larger than M^{**} the discounted reward of these two different policies are the same. However, the discounted reward of $(\gamma; \eta)$ policy is the smallest among these three policies.

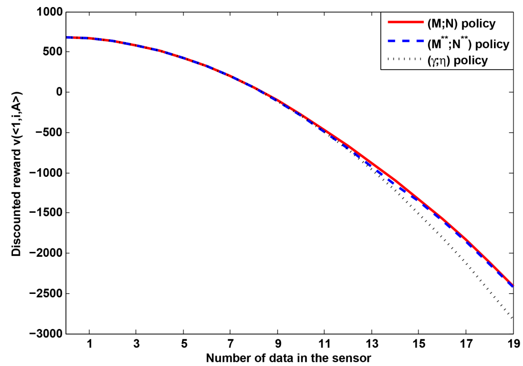

FIGURE 6. Discounted reward with $\langle 1, i, A\rangle$.

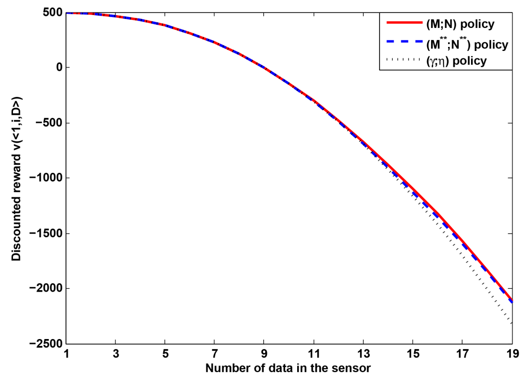

FIGURE 7. Discounted reward with $\langle 1, i, D\rangle$.

TABLE 2. difference value between $(M; N)$ and $(M^{**}; N^{**})$.

i	$\Delta v(\langle 0, i, A\rangle)$	$\Delta v(\langle 1, i, A\rangle)$	$\Delta v(\langle 1, i, D\rangle)$
0	0.0002	0.0001	\
1	0.0005	0.0001	0.0000
2	0.0013	0.0003	0.0001
3	0.0035	0.0007	0.0001
4	0.0100	0.0019	0.0003
5	0.0280	0.0053	0.0007
6	0.0787	0.0147	0.0019
7	0.2215	0.0411	0.0053
8	0.6233	0.1139	0.0147
9	1.7559	0.3111	0.0411
10	4.9567	0.8256	0.1139
11	14.0445	2.0569	0.3111
12	40.0765	4.3574	0.8256
13	66.2332	10.0738	2.0569
14	86.0162	32.3362	4.3574
15	83.6131	57.2263	10.0738
16	13.2552	82.0925	18.1378
17	17.1293	93.8007	26.5104
18	14.1760	28.3520	34.2587
19	11.7319	23.4637	28.3520
20	9.7091	19.4182	23.4637

TABLE 2 shows the difference value of discounted reward between $(M; N)$ policy and $(M^{**}; N^{**})$ policy with status $\langle 0, i, A\rangle$, $\langle 1, i, A\rangle$ and $\langle 1, i, D\rangle$. From TABLE 2, we can see that as i increases, the difference value between $(M; N)$ and $(M^{**}; N^{**})$ policy on status $\langle 0, i, A\rangle$, $\langle 1, i, A\rangle$ and $\langle 1, i, D\rangle$

TABLE 3. Simulation parameters of energy consumption.

parameters	e_{tr}	e_{rs}	e_{sr}	e_s
values	$40\mu w$	$0.5\mu w$	$15\mu W$	$10\mu W$

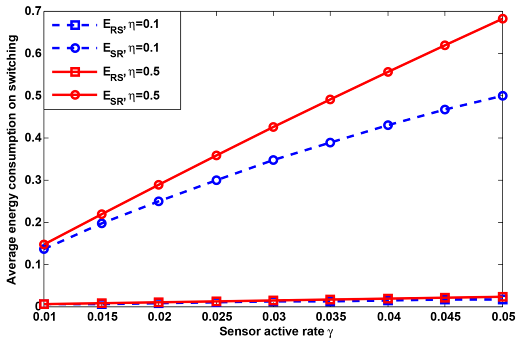

FIGURE 8. Energy consumption switching from sleep to active.

first increases and then decreases. Compared with the value of a discounted reward, the difference value is not too large. Therefore, with $(M^{**}; N^{**})$ policy, the sensor node system can also obtain good performance on discounted reward.

From the simulations above, with a given $(\gamma; \eta)$ algorithm, our work is to optimize its data admission process. It is clear our work increases the complexity because of our continuous MDP model with infinite state space initially. The complexity analysis of the MDP is indeed a hot topic in the area and it is not that easy to get the specific results, even in the polynomial running time. Some new results on the computational complexity of the infinite-horizon discounted-reward Markov Decision Problem (MDP) with a finite state space and a finite action space have been included the papers [19], [20]. However, based on our newly achieved bound results, compared with complexity of $(M; N)$ policy, the computational complexity of our new bound $(M^{**}; N^{**})$ has been significantly reduced because we can directly calculate without using any iterative methods. This can be testified also by the Algorithm 1 and Algorithm 2 as introduced in this section for actual computation.

B. ENERGY CONSUMPTION PERFORMANCE

In this section, we analyze the energy performance of the optimal $(M; N)$ policy and $(\gamma; \eta)$ policy with storage cost $f(i) = i^2, i \geq 0$. In order to evaluate the node energy consumption, we set the parameters for analysis, as in TABLE 3. With the optimal discounted reward, Algorithm 2 concerns performance on energy consumption with theorems in our manuscript. Here, we give an example in our algorithm 2, which concerns the energy consumption on active status with different λ_0. We can see as λ_0 changes, the $(M; N)$ policy will change and it is clear that the complexity of Algorithm 2 is much more than that of Algorithm 1.

It is obvious that the energy consumption from sleep status to active status of the $(M; N)$ policy is the same as that of the $(\gamma; \eta)$ policy; the same goes for the energy consumption

Algorithm 2 The Admission Control Algorithm With an Optimal Discounted Reward Concerning Energy Consumption

set i as the number of data in the sensor node (containing the processing one)
set $e_{rs}, e_{sr}, e_{tr}, e_s$ /*set the index value of energy consumption in sensor node*/
set $R, E_1, E_2, f(i)$ /*set the reward value and cost value in sensor node*/
set $\delta = 0$ as sleep, $\delta = 1$ as active
initialize $\lambda_1, \mu, \alpha, \beta, \eta, \gamma$ /*the sensor node obtains the real-time features of itself */
for $\lambda_0 = 0.15 : 0.05 : 0.40$ **do**
    calculate $(M; N)$
    calculate E_{TR} according to Theorem 4
end for

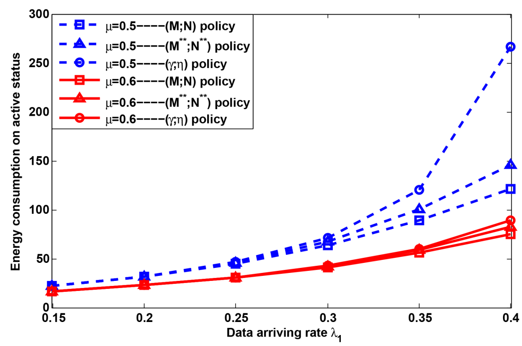

FIGURE 9. Energy consumption on active status with $\lambda_0 = 0.1$.

from active status to sleep status. In Fig.8, E_{RS} stands for the energy consumption from active status to sleep status, and E_{SR} represents the energy consumption from sleep status to active status. The energy consumption functions E_{RS} and E_{SR} are increasing functions on η and γ, and it is clear that the energy consumption from sleep to active is much more than that from active to sleep.

Fig.9 shows the average energy consumption on active status with different policies. For the three policies, it is shown that as λ_1 increases, the energy consumption on active status increases and with the increase of μ, the energy consumption decreases. The energy consumption on active status for the $(\gamma; \eta)$ policy is the largest among the three policies. And the energy consumption on active status of the $(M^{**}; N^{**})$ policy is a little larger than that of the $(M; N)$ policy.

Fig.10 illustrates the energy consumption in the sleep status with different policies. For these policies, we can obtain that with the increase of λ_1, the energy consumption in the sleep status increases. And with the increases of μ, the energy consumption on sleep status decreases. The energy consumption on sleep status for the $(\gamma; \eta)$ policy is the largest among the three policies. And the energy consumption on sleep status of the $(M^{**}; N^{**})$ policy is a little larger than that of the $(M; N)$ policy.

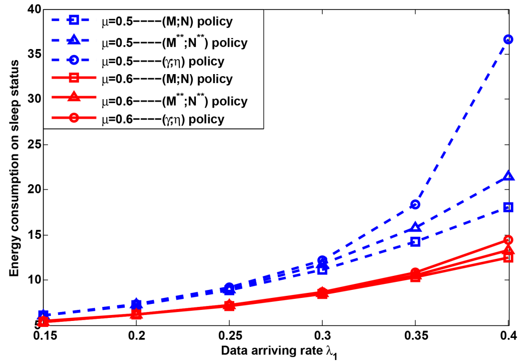

FIGURE 10. Energy consumption on sleep status with $\lambda_0 = 0.1$.

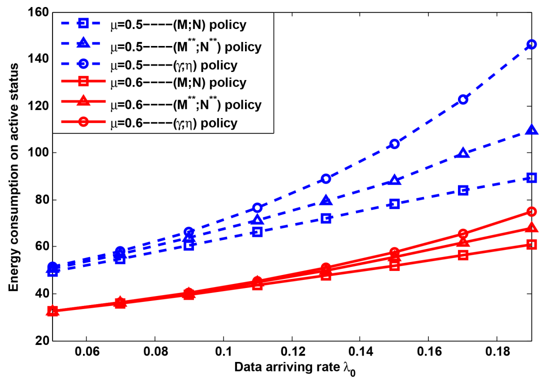

FIGURE 11. Energy consumption on active status with $\lambda_1 = 0.3$.

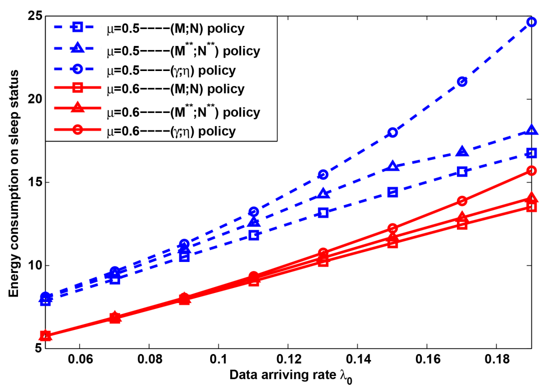

FIGURE 12. Energy consumption on sleep status with $\lambda_1 = 0.3$.

Fig.11 shows the influence of data arriving rate λ_0 on the average energy consumption on active status with different policies. It is shown that as λ_0 increases, the energy consumption on active status increases and with the increase of μ, the energy consumption decreases for the three polices. However, the energy consumption on active status for the $(M; N)$ policy and the $(M^{**}; N^{**})$ policy are much smaller than that of the $(\gamma; \eta)$ policy as λ_0 increases. And the energy consumption on active status of the $(M^{**}; N^{**})$ policy is a little larger than that of the $(M; N)$ policy.

Fig.12 illustrates the influence of the data arrival rate λ_0 on the average energy consumption on sleep status with different policies. For these three policies, it is shown that

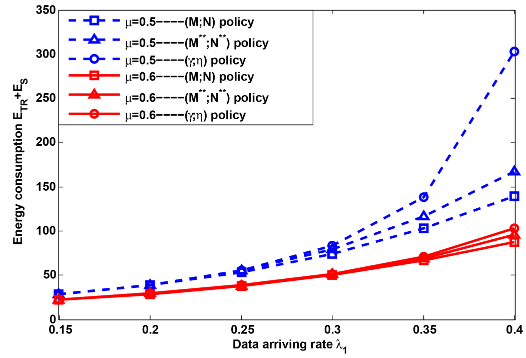

FIGURE 13. Energy consumption $E_{TR} + E_S$ with $\lambda_0 = 0.1$.

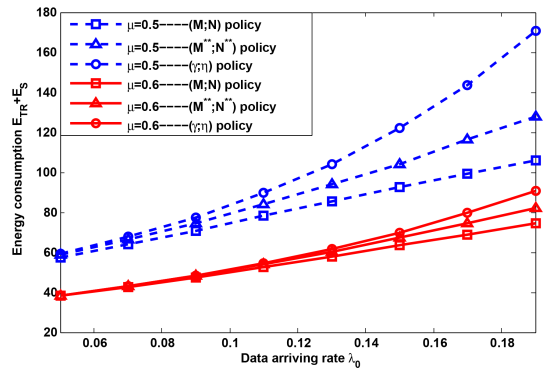

FIGURE 14. Energy consumption $E_{TR} + E_S$ with $\lambda_1 = 0.3$.

as λ_0 increases, the energy consumption on active status increases and with the increase of μ, the energy consumption decreases. The energy consumption on sleep status for the $(\gamma; \eta)$ policy is the largest among the three policies. And the energy consumption on sleep status of the $(M^{**}; N^{**})$ policy is a little larger than that of the $(M; N)$ policy.

From Figs.8-12 we can determine that the average switching energy consumption of different policies are the same, the energy consumption on active (sleep) status for the $(\gamma; \eta)$ policy is the largest among the three policies. And the energy consumption on active (sleep) status of the $(M^{**}; N^{**})$ policy is a little larger than that of the $(M; N)$ policy. Figs.13 and 14 show the energy consumption $E_{TR}+E_S$. In Fig.13 and Fig.14, the energy consumption $E_{TR} + E_S$ of the $(\gamma; \eta)$ policy is the largest among the three polices. However, the energy consumption $E_{TR} + E_S$ of the $(M^{**}; N^{**})$ policy is a little larger than that of the $(M; N)$ policy. From the analysis above, we can determine that with the increase of λ_1 and λ_0, the energy consumption of the sensor node increases dramatically. And the $(M^{**}; N^{**})$ policies can be used as the $(M; N)$ policy.

VII. CONCLUSION

In this article, we have studied the admission control problem in sensor nodes. Different from most existing works, we develop stochastic models of the sensor node of WSNs

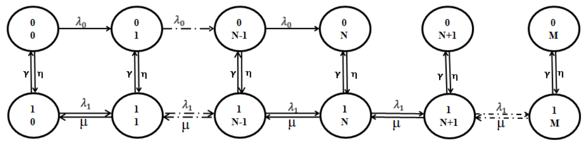

FIGURE 15. Transition rate diagram with $N < M$.

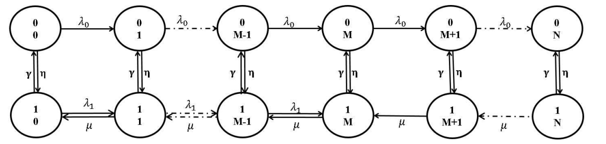

FIGURE 16. Transition rate diagram with $N > M$.

to explore admission control with the sleep/active scheme. In this paper, sensor nodes know common information (e.g., data generating rate, service rate, sensors' active rate, sensors' sleep rate) and the data in the sensors that need to be processed. Based on this, we consider the situation when a reward exists for accepting data packets and a storage cost per unit time is incurred for the data packets in the sensor. We observe that, under certain assumptions, the optimal admission policy for the admission control is a control limit policy. Also, a $(M;N)$ policy exists. When the volume of data in the sleep status is greater than N, and in the active status greater than M, the sensor nodes should reject arriving data to achieve the optimal discounted reward. Finally, we provide an upper bound for the $(M;N)$ policy. We also investigate the performance on energy consumption of the proposed model and provide related results. In addition, we study the expected discounted reward and energy performance of the model through numerical analysis. Based on our results, to obtain the optimal expected discounted reward and reduce energy consumption, when the data arrival rate is large, it is better to choose the $(M;N)$ policy. Our solution of the proposed MDP model can be implemented based on a reference table stored in sensor node's memory for online operations with minimal complexity. And in this paper, we only consider the design of a sensor node; for our future endeavors, we will consider the routing among large scale WSNs.

APEENDIX

PROOF OF THEOREM 4

(1) If $N < M$, the transition rate diagram is illustrated in Fig.15.

In addition to the notations of matrix A_0, A_1, A_2, A_{N_0} and B_0 introduced after Fig. 2 in the proof of Theorem 4, we further need the following notations

$$A_3 = \begin{bmatrix} 0 & 0 \\ 0 & \lambda_1 \end{bmatrix}, \quad A_{N_1} = \begin{bmatrix} -\eta & \eta \\ \gamma & -(\mu+\gamma+\lambda_1) \end{bmatrix},$$

and $A_M = A_{N_0}$. The corresponding transition rate matrix Q_1 with size $M \times M$ of Fig.15 is given by,

$$\begin{bmatrix} D_0^{(1)} & G_0^{(1)} & 0 & \cdots & 0 & 0 & 0 \\ A_2 & D_1^{(1)} & G_1^{(1)} & \cdots & 0 & 0 & 0 \\ 0 & A_2 & D_2^{(1)} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_2 & D_{M-1}^{(1)} & G_{M-1}^{(1)} \\ 0 & 0 & 0 & \cdots & 0 & A_2 & D_M^{(1)} \end{bmatrix},$$

where

$$D_i^{(1)} = \begin{cases} B_0, & \text{if } i = 0; \\ A_1, & \text{if } i = 1, 2, \ldots, N-1; \\ A_{N_1}, & \text{if } i = N,\ N+1, \ldots, M-1; \\ A_{N_0}, & \text{if } i = M. \end{cases}$$

and

$$G_i^{(1)} = \begin{cases} A_0, & \text{if } i = 0,\ 1, \ldots, N-1; \\ A_{N_3}, & \text{if } i = N,\ N+1, \ldots, M-1. \end{cases}$$

According to Theorem 1 on page 720 in [17], we have

$$\begin{cases} \pi_M^M E_M = 0, \\ \pi_i^M = \pi_{i+1}^M A_2(-E_i^{-1}) \quad \text{for } 0 \le i \le M-1, \end{cases}$$

with the normalization condition $\sum_{i=0}^{M} \pi_i^M e = 1$, where

$$E_i = \begin{cases} B_0, & \text{if } i = 0; \\ A_1 + A_2(-E_{i-1}^{-1})A_0, & \text{if } i = 1, 2, \ldots, N-1; \\ A_{N_1} + A_2(-E_{i-1}^{-1})A_0, & \text{if } i = N; \\ A_{N_1} + A_2(-E_{i-1}^{-1})A_3, & \text{if } i = N+1, \ldots, M-1; \\ A_M + A_2(-E_{M-1}^{-1})A_3, & \text{if } i = M. \end{cases}$$

Hence, when $N < M$, we will have

$$E_{TR} = \sum_{i=0}^{M} i\pi_{R_i}^M e_{tr}, \quad \text{and } E_S = \sum_{i=0}^{M} i\pi_{S_i}^M e_s.$$

(2) By using the similar idea in the above item, we can easily get the results also for the case when $N > M$. For the completeness of the results, we also include it as below. In fact, in this case, the transition rate diagram of $(M;N)$ model is as Fig. 16. Similarly, in addition to the notations of matrix A_0, A_1, A_2, A_{N_0} and B_0 introduced after Fig. 2 in the proof of Theorem 4, we further need the following notations in this case:

$$A_4 = \begin{bmatrix} \lambda_0 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_{M_0} = \begin{bmatrix} -(\eta+\lambda_0) & \eta \\ \gamma & -(\gamma+\mu) \end{bmatrix}.$$

The corresponding transition rate matrix Q_2 with size $N \times N$ is given by

$$\begin{bmatrix} D_0^{(2)} & G_0^{(2)} & 0 & \cdots & 0 & 0 & 0 \\ A_2 & D_1^{(2)} & G_1^{(2)} & \cdots & 0 & 0 & 0 \\ 0 & A_2 & D_2^{(2)} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_2 & D_{N-1}^{(2)} & G_{N-1}^{(2)} \\ 0 & 0 & 0 & \cdots & 0 & A_2 & D_N^{(2)} \end{bmatrix},$$

where

$$D_i^{(2)} = \begin{cases} B_0, & \text{if } i = 0; \\ A_1, & \text{if } i = 1, 2, \ldots, M-1; \\ A_{M_0}, & \text{if } i = M,\ M+1, \ldots, N-1; \\ A_{N_0}, & \text{if } i = N. \end{cases}$$

and

$$G_i^{(2)} = \begin{cases} A_0, & \text{if } i = 0, 1, \ldots, M-1; \\ A_{N_4}, & \text{if } i = M,\ M+1, \ldots, N-1. \end{cases}$$

According to Theorem 1 on page 720 in [17], we have,

$$\begin{cases} \pi_N^N L_N = 0, \\ \pi_i^N = \pi_{i+1}^N A_2(-L_i^{-1}) \quad \text{for} \quad 0 \le i \le N-1, \end{cases}$$

with the normalization condition $\sum_{i=0}^{N} \pi_i^N e = 1$, where

$$L_i = \begin{cases} B_0, \quad \text{if } i = 0; \\ A_1 + A_2(-L_{i-1}^{-1})A_0, & \text{if } i = 1, 2, \ldots, M-1; \\ A_{M_0} + A_2(-L_{i-1}^{-1})A_0, & \text{if } i = M; \\ A_{M_0} + A_2(-L_{i-1}^{-1})A_4, & \text{if } i = M+1, \ldots, N-1; \\ A_{N_0} + A_2(-L_{N-1}^{-1})A_4, & \text{if } i = N. \end{cases}$$

Thus, when $N > M$,

$$E_{TR} = \sum_{i=1}^{N} i\pi_{R_i}^N e_{tr}, \quad \textit{and } E_S = \sum_{i=1}^{N} i\pi_{S_i}^N e_s.$$

This completes the proof.

REFERENCES

[1] Q. Wang *et al.*, "A high data rate, multi-nodes wireless personal-area sensor network for real-time data acquisition and control," in *Proc. 1st Int. Conf. Electron. Instrum. Inf. Syst.*, Jul. 2017, pp. 1–5.
[2] K. Ren, K. Zeng, W. Lou, and P. J. Moran, "On broadcast authentication in wireless sensor networks," in *Proc. Int. Conf. Wireless Algorithms, Syst., Appl.*, May 2006, pp. 502–514.
[3] A. Bachir, M. Dohler, T. Watteyne, and K. K. Leung, "MAC essentials for wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 2, pp. 222–248, 2010.
[4] O. Alagoz, M. U. Ayvaci, and J. T. Linderoth, "Optimally solving Markov decision processes with total expected discounted reward function: Linear programming revisited," *Comput. Ind. Eng.*, vol. 87, pp. 311–316, Sep. 2015.
[5] M. A. Alsheikh, D. T. Hoang, D. Niyato, H.-P. Tan, and S. Lin. (2015). "Markov decision processes with applications in wireless sensor networks: A survey." [Online]. Available : https://arxiv.org/abs/1501.00644
[6] M. Buettner, G. V. Yee, E. Anderson, and R. Han, "X-MAC: A short preamble MAC protocol for duty-cycled wireless sensor networks," in *Proc. 4th Int. Conf. Embedded Networked Sensor Syst.*, Oct. 2006, pp. 307–320.
[7] W.-L. Yeow, C.-K. Tham, and W.-C. Wong, "Optimizing application performance through learning and cooperation in a wireless sensor network," in *Proc. IEEE Mil. Commun. Conf.*, Oct. 2007, pp. 1–7.
[8] Z. Ye, A. A. Abouzeid, and J. Ai, "Optimal stochastic policies for distributed data aggregation in wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 17, no. 5, pp. 1494–1507, Oct. 2009.
[9] J. Hao, Z. Yao, K. Huang, B. Zhang, and C. Li, "An energy-efficient routing protocol with controllable expected delay in duty-cycled wireless sensor networks," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2013, pp. 6215–6219.
[10] X. Fei, A. Boukerche, and R. Yu, "An efficient Markov decision process based mobile data gathering protocol for wireless sensor networks," in *Proc. EEE Wireless Commun. Netw. Conf.*, Mar. 2011, pp. 1032–1037.
[11] Z. Yuan, L. Wang, L. Shu, T. Hara, and Z. Qin, "A balanced energy consumption sleep scheduling algorithm in wireless sensor networks," in *Proc. 7th Int. Wireless Commun. Mobile Comput. Conf.*, Jul. 2011, pp. 831–835.
[12] S. Misra and S. Singh, "Localized policy-based target tracking using wireless sensor networks," *ACM Trans. Sensor Netw.*, vol. 8, no. 3, p. 27, Jul. 2012.
[13] Y. Zhang and W. Li, "Modeling and energy consumption evaluation of a stochastic wireless sensor network," *J. Wireless Commun. Netw.*, vol. 1, p. 282, May 2012.
[14] O. Alagoz and M. U. Ayvaci, *Uniformization in Markov Decision Processes*. Hoboken, NJ, USA: Wiley, 2010.
[15] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 2014.
[16] W. Ni, W. Li, and M. Alam, "Determination of optimal call admission control policy in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 1038–1044, Feb. 2009.
[17] D. Gaver, P. Jacobs, and G. Latouche, "Finite birth-and-death models in randomly changing environments," *Adv. Appl. Probab.*, vol. 16, no. 4, pp. 715–731, 1984.
[18] L. Huang and T. T. Lee, "Generalized pollaczek-khinchin formula for Markov channels," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3530–3540, 2013.
[19] Y. Chen and M. Wang. (2017). "Lower bound on the computational complexity of discounted Markov decision problems." [Online]. Available: https://arxiv.org/abs/1705.07312
[20] A. Sidford, M. Wang, X. Wu, L. Yang, and Y. Ye, "Near-optimal time and sample complexities for solving Markov decision processes with a generative model," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 5192–5202.

XIAOLU ZHANG received the bachelor's degree in communication engineering from Donghua University, Shanghai, China, in 2012, where she is currently pursuing the Ph.D. degree in control theory and control engineering with the College of Information Sciences and Technology. From 2016 to 2018, she was studying as a visiting student with the Department of Computer Science, Texas Southern University, Houston, TX, USA. Her current research interests include intelligent systems, intelligent control, network control systems, and the application of wireless sensor networks.

DEMIN LI received the Ph.D. degree in electronic and computer engineering from the Nanjing University of Science and Technology, China, in 1998. From 2002 to 2005, he was a Research Scientist with Nanjing University, China. He is currently a Professor with the Department of Telecommunication and Electronic Engineering, College of Information Science and Technology, Donghua University, Shanghai, China. He has authored more than 60 research papers in journals and conferences. His recent research interests include telecommunication system engineering, wireless mobile networking, mobile decision theory, and mobile decision support systems. He is currently an Associate Chairman of the Circuits and Systems Committee, Shanghai.

WEI W. LI (M'99–SM'06) received the B.Sc. degree from Shaanxi Normal University, Xi'an, China, in 1982, the M.Sc. degree from the Hebei University of Technology, Tianjin, China, in 1987, and the Ph.D. degree from the Chinese Academy of Sciences, Beijing, China, in 1994. He was once an Associate Professor (with tenure mth) with the Department of Electrical Engineering and Computer Science, University of Toledo, Toledo, OH, USA, and a tenure track Assistant Professor with the Department of Electrical and Computer Engineering, University of Louisiana at Lafayette, Lafayette, LA, USA. He is currently a Professor with the Department of Computer Science and the Director/PI of the NSF-CREST Center for Research on Complex Networks, Texas Southern University, Houston, TX, USA. He is the author/coauthor of six books and more than 160 peer-reviewed papers in professional journals and the proceedings of international conferences. His research interests include dynamic control and optimization of energy-efficient wireless sensor networks; evaluation, complexity, power connectivity, and coverage for wireless sensor networks; adaptation, design, and implementation of dynamic models in wireless multimedia systems; theoretic foundations and advanced analysis in real-time, hybrid, and embedded systems; and performance evaluation of biomolecular networks, queueing networks, and reliability systems. He is currently serving as an Editor for several professional journals and is serving or has served as a Steering Committee Member, the General Co-Chair, the Technical Program Committee Co-Chair, the Track Chair, and a Technical Program Committee Member for a number of professional conferences, including the INFOCOM, ICDCS, Globecom, ICC, and WCNC.

WEI ZHAO (F'01) received the bachelor's degree in physics from Shaanxi Normal University, Xi'an, China, and the M.S. and Ph.D. degrees in computer and information sciences from the University of Massachusetts, Amherst, MA, USA, in 1983 and 1986, respectively. He is currently serving the American University of Sharjah as its Chief Research Officer (VPR). From 2008 to 2018, he has served as the eighth Rector (i.e., President) of the University of Macau. He also served as the Dean of the School of Science, Rensselaer Polytechnic Institute, the Director of the Division of Computer and Network Systems with the U.S. National Science Foundation, and the Senior Associate Vice President of Research with Texas A&M University. He has made significant contributions to cyber-physical systems, distributed computing, real-time systems, computer networks, and cyberspace security. He led the effort to define the research agenda of and to create the very first research funding program for cyber-physical systems, when he served as the NSF CNS Division Director, in 2006. His research group has received numerous awards; moreover, their research results have been adopted in the standard of DOD SAFENET (Survivable Adaptable Fiber Optic Embedded Network). In 2011, he was named by the Chinese Ministry of Science and Technology as the Chief Scientist of the National 973 Internet of Things Project. He was a recipient of the Lifelong Achievement Award by the Oversea Chinese Association of Science and Technology, in 2005. In 2007, he was honored with the Overseas Achievement Award by the Chinese Computer Federation. He has been conferred honorable doctorates by 12 universities in the world and academician of the International Eurasian Academy of Sciences.