This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2997984, IEEE Sensors Journal

Maximum Throughput under Admission Control with Unknown Queue-length in Wireless Sensor Networks

Xiaolu Zhang, Demin Li, Yihong Zhang

***Abstract*—In a Wireless Sensor Network (WSN) node, as the input traffic increases, the throughput can be assumed to first increase and then start to decrease, indicating congestion in the buffer. This suggests the need for an admission control mechanism to maintain high throughput as the arrival traffic increases. Considering the stochastic nature of WSNs, the information of the queue-length of arrival or newly sensed data packets can be unknown to a sensor node. This paper proposes a probabilistic admission control model with the maximum throughput for the node. In the proposed model, a reward when a data packet arriving to a sensor is accepted (not rejected) for transmission is considered, but a holding cost per unit time for the delay of accepted data packets in the sensor is also incurred. For the sensor node, by constructing a suitable Markov decision process (MDP), a probabilistic admission control algorithm on how to accept data packets on sleep and active phases to achieve a maximum throughput is proposed. Furthermore, for the identified $(p;q)$ model, the energy consumption of the node in active and sleep phases, as well as the energy consumption switching from active to sleep per unit time and vice versa is investigated. An extensive simulation is implemented. The numerical results show that the problem is effectively solved by an optimal scheme with high energy efficiency. The results of this paper can be applied in designing optimal sensor nodes in WSNs**

***Index Terms*—Maximum throughput, probabilistic admission control, queue-length, discounted reward, energy performance**

I. INTRODUCTION

WIRELESS Sensor Networks are widely used in many areas, such as environmental monitoring, disaster warning and industrial intelligence, due to the simple deployment and low cost. In WSNs, nodes sense variables from the environment and exchange collected data with other nodes through data packets. Throughput is a metric in WSNs. As the throughput increases, the load of the node increases.

This work is supported by the NSF of China under Grant No.18K10454, No.71171045, No.61772130 and No.61301118; the Special Project Funding for the Shanghai Municipal Project "Big Data Management System of UAVs" under No.JMRH-2018-1042; the Innovation Program of Shanghai Municipal Education Commission under Grant No.14YZ130; the international S&T Cooperation Program of Shanghai Science and Technology Commission under Grant No.15220710600; the Fundamental Research Funding for the Central Universities of Ministry of Education of China under Grant No.18D110408 and No.CUSF-DH-D-2015058.

Xiaolu Zhang, Demin Li, and Yihong Zhang are with the College of Information Science and Technology and Engineering Research Center of Digitized Textile and Fashion Technology of Ministry of Education, Donghua University, Shanghai 201620, China (e-mail:xiaoludhu@mail.dhu.edu.cn, deminli@dhu.edu.cn, zhangyh@dhu.edu.cn)

Corresponding author: Yihong Zhang.

Admission control policies for maintaining high throughput for a network have been an ongoing research topic for a long period of time [1]. In admission control models, there exists a holding cost for the delay of data packets in a sensor. The holding cost increases as the increase of data packets in the sensor's buffer [2], [3], [4]. Developing an optimal admission control policy in data gathering process is meaningful in WSNs.

The traffic characteristics and communication patterns in WSNs are partial and vague [5], [6]. Information has a decisive effect during the admission control process. However, many authors consider this process to be ideal and they suppose systems can acquire all the information to make the correct decision [7], [8], [9]. For the complexity of WSNs, a node may work with partial information. For example, in [10], considering the information of the queue-length at the transmitter is not known at the receiver, authors have proposed a sleep/wake scheduling at the wireless node. In the model, packets are assumed to be received with probability p. In [11], an active queue management (AQM) mechanism is proposed to estimate the congestion and avoid the congestion by means of dropping packets with different probabilities for different arrival packets. It is meaningful to study the admission control process of sensor nodes in the case of partial information, in which a sensor node may not know the information of the queue-length of arrival data packets.

In this paper, with the sleep/active scheme, we consider to obtain a maximum throughput during the data admission process with the Markov Decision Process (MDP) model. We mainly consider a two-phase model. The first stage is used to manage the admission of arrivals and the second one is used to manage the service. In the fist stage, the receiver or sensing subsystem of a sensor node does not know the information of the queue-length of arrival or newly sensed data packets. However, the common information which includes the data packets' arrival rate, service rate, the sleep and active duration features of the sensor node is known.

The contributions of this paper are as follows.

- First, we propose a novel $(p;q)$ policy for data admission process. If a sensor node does not know arrival data packets' queue-length information, it is not wise to accept all arrival data packets. The holding cost for the delay of data packets in the sensor will be increased as the increase of data packets. We propose a $(p;q)$ admission model for the sensor node. In the sleep status, the sensor accepts data packets with probability p and on active status, it accepts data packets with probability q. The proposed model can reflect the different behavior of a sensor node on sleep and active status.

- Second, we present a MDP-based method to the problem of maximum throughput under data admission process with the $(p;q)$ policy. For a sensor node, an immediate reward is obtained once a data packet is accepted into the sensor. However, a holding cost is incurred during the delay of data packets in the sensor. Based on the MDP, a decision rule during data admission control is obtained. Results of the optimal value of $(p;q)$ for the maximum throughput under data admission control process are verified. The solution of the $(p;q)$ 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. It also can be implemented by the MDP-Based dynamic optimization methodology in [12].
- Finally, we derive the sensor node performance from energy consumption (energy consumption switching from the sleep status to the active status and switching from the active status to the sleep status, average energy consumption in the sleep status and in the active status.) as the sensor in sleep/active mode changes with the optimal $(p;q)$ policy, which can be used to redesign the working process of sensor nodes for a better performance.

The rest of this paper is organized as follows. Related work is reviewed in Section II. We describe the system model and formulate our problem in Section III. We present the maximum throughput with admission control for the proposed model which is also the most important part of this paper in Section IV. The performance evaluation from energy consumption is given out in Section V. The experimental results are presented in Section VI. Finally, We conclude our work in Section VII.

II. Related work

Throughput is a metric in the device design and implementation of WSNs. A lot of work is done to analyze the throughput of WSNs. The queue model is an important tool to analyze the performance of sensor nodes. For example, in [13], the average network throughput is analyzed with a proposed finite queuing model of a sensor node. In [14], the authors propose a Markov model to describe the behavior of SMAC with a finite queue capacity. With the model, the expected throughput of SMAC under variable number of nodes, queue capacities, contention window sizes, and data arrival rates is studied. In [15], a stochastic model of WSNs in which each sensor node randomly and alternatively stays in an active mode or a sleep mode is investigated. In the model, authors analyze the throughput with the Markov model.These research work ([13], [14], [15]) indicate that as the arrival rate or service rate of data packets increases, the value of the throughput can be assumed to increase. Meanwhile, the cost for maintaining the sensor increases, too.

Despite that many proposed mechanisms that can be used to analyze and optimize the throughput in a WSN, in some scenarios the required constraints for the throughput can not be available in the network. It is meaningful to control arrival data packets' cost in a sensor node's operating when we maintain the required throughput. Thus, an admission control mechanism can be used to monitor arrival data packets.

Admission control can be addressed with regard to energy consumption, resource utilization, or feasibility. It can also be addressed at different levels, like the packet/MAC level, flow/connection level, node level, or service level [16]. In [17], during data admission process, a threshold N has been derived to obtain minimum power consumption for a sensor node while considering each different data packets' arrival rate. In [18], to obtain the maximum discounted reward during data admission process, a $(M;N)$ policy about when to admit arrival data packets and when to reject arrival data packets is verified. In [19], the authors propose real-time measurements of the energy consumption by individual applications, then they propose an optimal admission control policy and a post-admission policing mechanism at the node-level. The approach trades between energy consumption and user rewards. Few of these existing work ([16], [17], [18], [19]) have addressed admission control for throughput in WSNs. And most of them assume all the information are known when the node makes decision.

A sensor node works like a queue server [1], [15], [20], [21]. In a queue, a server can only know partial information and not know the information of the queue-length of arrival or newly sensed data packets. In this paper, we investigate a model for a maximum throughput on admission control with sleep/active scheme in sensor nodes using MDP method. In our model, a sensor node has no knowledge about the number of arrival or newly sensed data packets. Thus, we assume that on sleep status, data packets are accepted by the sensor node with probability p. On active status, data packets will be accepted by the sensor node with probability q. With this, we model the admission process with MDPs, in which, a reward will be obtained when a data packet from a sensor is accepted and at the same time, a holding cost also exists for data packets' delay in the sensor. We focus on the problem that a sensor node how to make decision strategies to make sure the long term discounted reward for accepting an arrival data packet is not less than that for rejecting it in the sleep and active phase, respectively. With the decision strategies, we consider how to get a maximum throughput. Finally, we study the influence of the proposed $(p;q)$ policy on energy consumption.

III. Maximum Throughput Base Model

We consider a WSN in which static sensor nodes are randomly located in a given region. All sensor nodes are mainly used to sense variables from the environment and exchange the collected data with other nodes through data packets. To reduce energy consumption, all sensor nodes are configured with the sleep/active scheme. In real-application environments, due to the existence of signal interference and noise, the packet cannot be received successfully [22]. To be more practical and realistic, instead of taking the deterministic network model, we define a probabilistic network model, where a sensor node cannot know the information of the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2997984, IEEE Sensors Journal

queue-length of arrival or newly sensed data packets. So, in the network mode, sensor nodes receive arrival data packets with probability. There exists a holding cost for data packets' delay in the sensor, a switching on cost when a sensor is switched on and a switching off cost when a sensor is switched off. In the meanwhile, a reward also exists for an admitted data packet. To study the optimal admission control problem with the sleep/active scheme for the maximum throughput, in this section, we provide the detailed description of our model.The descriptions of the symbols and notations used in this paper can be summarized in the Table I.

TABLE I: Summary of Notations

Symbols	Definition
η	Switching-on rate.
γ	Switching-off rate.
λ_0	Data packets' sensing rate.
λ_E	Data packets' arriving rate from other sensors.
μ	Data packets' processing rate.
$f(i)$	Holding cost rate per unit time with i data packets.
E_1	Switching-on cost per time.
E_2	Switching-off cost per time.
p	Data packets received probability on sleep status.
q	Data packets received probability on active status.

A. Assumptions

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

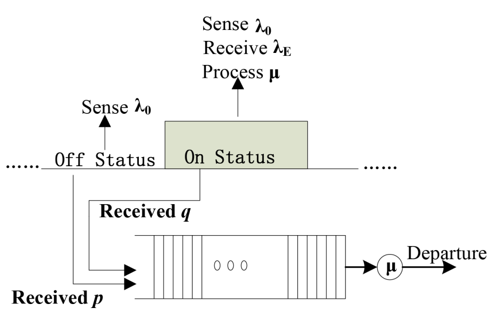

Fig. 1: The working model of a sensor node.

(1) The duration of a sensor in a sleep mode is distributed exponentially with a mean of $1/\eta$. On sleep status, the sensor can generate data packets with its sensing subsystem according to a Poisson process at a rate of λ_0.

(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
   - generate packets according to a Poisson process at a rate of λ_0;
   - receive packets coming from other sensors in accordance to a Poisson process at a rate of λ_E; and
   - process (transmit or relay) data packets with negative exponential distribution with a mean rate of μ.

   From the definition above, we can define these data packets in active status follows a Poisson process with a rate of λ_1. It is clearly that $\lambda_1 = \lambda_0 + \lambda_E$. When a sensor node ends its active status, it will switch to the sleep status.

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

(4) On sleep status, data packets are received by a sensor node with probability p, and on active status, data packets are received by a sensor node with probability q.

(5) 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. The maximum throughput problem under data admission control

With the above assumptions, the maximum throughput problem under data admission control process in a sensor node can be formulated as follows:

Given:

- The sensor node works on a sleep/active scheme at the η/γ rate.
- On sleep status, data packets arrive at a rate of $\lambda_0 p$. On active status, data packets arrive at a rate of $\lambda_1 q$ and is processed at a rate of μ.
- A sensor is in the sleep status with i data packets with the probability of $\pi_{S_i} = P(S_i)$ and in the active status with i data packets with the probability of $\pi_{R_i} = P(R_i)$, where $i = 0, 1, 2, \cdots$.

Objective: Find the probability (p, q) to maximize the throughput Q of the sensor node which is defined as the average number of the data packets transmitted per unit time as follows.

$$Q(p,q) = \sum_{i=1}^{\infty} P(R_i)\mu. \tag{1}$$

Subjective To:

During data admission process,

1) on sleep status, the total expected discounted reward for an arrival data packet's acceptance of a sensor is not less than that for its rejective; and
2) on active status, the total expected discounted reward for an arrival data packet's acceptance of a sensor is not less than that for its rejective.

From Equation (1), we can see that to obtain the maximum throughput during data admission process in the sensor node, it is important to solve the data admission control problem. In the following sections, we will discuss our proposed solutions for obtaining the maximum throughput during data admission process in detail.

IV. Proposed Model for Maximum Throughput

In this section, for the proposed objective function, we firstly need to seek the throughput during data admission process in the sensor node and then, we need to solve the problem of the constraints of our proposed objective function. With these work, we can finally to obtain the optimal probability $(p;q)$ for maximum throughput during data admission process.

A. Throughput during data admission process

During the data admission process, we can get that with the sleep/active scheme, data packets arrive at a rate of $\lambda_0 p$ on sleep status, and at a rate of $\lambda_1 q$ on active status. Meanwhile, the sensor node process data packets at a rate of μ on active status. Thus, the transition rate diagram can be illustrated as Fig.2.

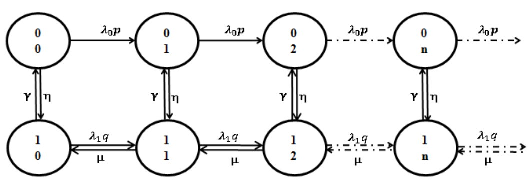

Fig. 2: Transition rate diagram of a sensor node.

Theorem 1. *The throughput of a sensor node in the investigated WSNs is*

$$Q(p,q) = \frac{\lambda_1 \eta q + \lambda_0 \gamma p}{\gamma + \eta}.$$

Proof: The throughput Q of a sensor node which is defined as the average number of the data packets transmitted per unit time. Thus,

$$Q(p,q) = \sum_{i=1}^{\infty} P(R_i)\mu.$$

The corresponding transition rate matrix T of the constructed multi-dimensional Markov process in Fig.2 can be given by

$$T = \begin{bmatrix} B_0 & A_0 & 0 & 0 & 0 & \cdots \\ A_2 & A_1 & A_0 & 0 & 0 & \cdots \\ 0 & A_2 & A_1 & A_0 & 0 & \cdots \\ 0 & 0 & A_2 & A_1 & A_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

where

$$B_0 = \begin{bmatrix} -(\eta + \lambda_0 p) & \eta \\ \gamma & -(\gamma + \lambda_1 q) \end{bmatrix}, A_0 = \begin{bmatrix} \lambda_0 p & 0 \\ 0 & \lambda_1 q \end{bmatrix},$$

and

$$A_1 = \begin{bmatrix} -(\eta + \lambda_0 p) & \eta \\ \gamma & -(\mu + \gamma + \lambda_1 q) \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 \\ 0 & \mu \end{bmatrix}.$$

We have

$$\pi_i = \pi_0 H^i, \text{for } i = 0, 1, 2, \cdots,$$

where matrix H is the minimal non-negative solution to the matrix-quadratic equations

$$H^2 A_2 + H A_1 + A_0 = 0,$$

and π_0 is the unique position solution of the equations

$$x_0(B_0 + HA_2) = 0 \text{ and } x_0(I - H)^{-1} e = 1,$$

in which e is a two-dimensional column vector with all its components of 1. According to [15], we finally obtain that

$$H = \begin{bmatrix} \frac{(\gamma+\mu)\lambda_0 p}{(\lambda_0 p+\eta)\mu} & \frac{\lambda_0 p}{\mu} \\ \frac{\gamma \lambda_1 q}{(\lambda_0 p+\eta)\mu} & \frac{\lambda_1 q}{\mu} \end{bmatrix},$$

and

$$\pi_0 = [\frac{\eta(\mu - \lambda_1 q) - \gamma\lambda_0 p}{\mu(\eta+\gamma)(\lambda_0 p + \eta)}\gamma, \quad \frac{\eta(\mu - \lambda_1 q) - \gamma\lambda_0 p}{\mu(\eta+\gamma)}].$$

Thus,

$$P(S_0) = \frac{\gamma(\eta\mu - \gamma\lambda_0 p - \eta\lambda_1 q)}{\mu(\gamma+\eta)(\lambda_0 p + \eta)},$$

$$P(R_0) = \frac{(\mu - \lambda_1 q)\eta - \gamma\lambda_0 p}{(\gamma+\eta)\mu}.$$

From the transition diagram in Fig.2, we also can have

$$P(S)\eta = P(R)\gamma,$$

and

$$P(S) + P(R) = 1.$$

Therefore, the throughput of a sensor node in the investigated WSNs is

$$Q(p,q) = \sum_{i=1}^{\infty} P(R_i)\mu = [P(R) - P(R_0)]\mu = \frac{\lambda_1 \eta q + \lambda_0 \gamma p}{\gamma + \eta}.$$

The proof is finished. ■

Remarks. The higher value of p and q are, the higher throughput of a sensor node will be. However, the holding cost of the sensor node will be more, too. Thus, it is very important to control the length of data packets in a sensor node. For the $(p;q)$ scheme, there is no knowledge about the information of the queue-length of arrival or newly sensed data packets. It is useful to investigate the average number of data packets in the sensor node.

Theorem 2. *When data packets' average arriving rate is less than the processing rate, which is to say $\eta\mu > \eta\lambda_1 + \gamma\lambda_0$, then, in the sleep status, the average number of data packets in a sensor node is given by*

$$L_S = \frac{-\lambda_0\lambda_1 pq + \gamma\lambda_0 p + \lambda_0 \mu p + \lambda_1 \eta q}{(\eta\mu - \gamma\lambda_0 p - \lambda_1 \eta q)(\gamma + \eta)}\gamma,$$

and in the active status, the average number of data packets is

$$L_R = \frac{\gamma\lambda_0^2 p^2 + \gamma\lambda_0 \eta p + \lambda_1 \eta^2 q}{(\eta\mu - \gamma\lambda_0 p - \lambda_1 \eta q)(\gamma + \eta)}.$$

Proof: See the Appendix A. ■

For a sensor node, it needs to receive and process (transmit) data packets which is just like a server. Admitting a data packet would bring a reward to the sensor node but holding each data packet in the node would also incur some expenses (or cost). Thus, it is very important to control the admission process in a sensor node.

B. CTMDP framework for sensor nodes' behavior

The behavior of a sensor node about accepting or rejecting arrival data packets can be modeled as a stochastic dynamic programm problem in which a decision rule prescribes a procedure for action selection in each status at a specified decision epoch. In general, a policy π is a sequence π_1, π_2, ... of decision rules where π_i is the decision selected when a sensor is in the ith state, which tells how to select an action after completion of the $(i-1)$th transition. We denote $\pi(s)$ as the action to take when the system occupies state s. Given policy π, denote the total expected infinite-horizon discounted reward when starting from state s by $v_\alpha^\pi(s)$. $\alpha > 0$ is the discount factor so that a reward r received has present value $re^{-\alpha t}$ after some time t. The total expected infinite-horizon discounted reward with state s is

$$v_\alpha^\pi(s) = E_s^\pi\left\{\int_0^\infty e^{-\alpha t} r(s_t, a_t)dt\right\}, \tag{2}$$

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.

With these description, during data admission process, we need to make sure that on sleep status, the total expected discounted reward for an arrival data packet's acceptance of a sensor is not less than that for its rejective and on active status, the total expected discounted reward for an arrival data packet's acceptance of a sensor is not less than that for its rejective. For this purpose, we will introduce the method of continuous-time Markov Decision Process (CTMDP) which can be uniquely identified by the following five components: state space, action space, decision epochs, the transition probabilities, and the reward function.

- **State space:** At each decision epoch, the system occupies a state. We denote the set of possible states of a sensor by S. In our scheme, state space $S = \{s : s = \langle\delta, b\rangle\}$, $\delta \in \{0, 1\}$ and $b \in \{A, D, C\}$. $\delta = 0$ means the sensor is in the sleep status and $\delta = 1$ indicates that the sensor is in the active status. $b = A$ stands for the recent packet event is an arrival of a data packet, $b = D$ means the recent packet event is a process of a data packet and $b = C$ means the recent packet event is a switching of the node.
- **Action space:** In this model, there are three different actions $A_s = \{a_A, a_R, a_C\}$. For a node, when the most recent event is an arrival of a new data packet, it may accept the arrival data packet with action a_A, or reject the new arrival data packet with action a_R. If the most recent event is a process of a data packet or a switching of the sensor node, then the node takes an action, denote by a_C, to continue. Thus, we have $A_{\langle 0,A\rangle} = A_{\langle 1,A\rangle} = \{a_A, a_R\}$ and $A_{\langle 1,D\rangle} = A_{\langle 0,C\rangle} = A_{\langle 1,C\rangle} = \{a_C\}$.
- **Decision epochs:** The decision epochs are those time points when a new data packet is arrived, a data packet is processed or when 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
$$P(\tau(s, a) \le t) = 1 - e^{-\beta(s,a)t}, t \ge 0.$$
With the (p, q) scheme, $\beta(s, a)$ can be written as,
$$\beta\big(\langle\delta, b\rangle, a\big) = \begin{cases} \lambda_0 + \eta, & s = \langle 0, A\rangle, a = a_A \text{ or } a_R, \\ \lambda_0 + \eta, & s = \langle 1, C\rangle, a = a_C, \\ \lambda_1 + \mu + \gamma, & s = \langle 1, A\rangle, a = a_A \text{ or } a_R, \\ \lambda_1 + \mu + \gamma, & s = \langle 1, D\rangle \text{ or } \langle 0, C\rangle, a = a_C. \end{cases}$$
- **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 the (p, q) scheme, a node does not know the exact information about itself. The transition probability is as follows.
For the active status, $s = \langle 0, C\rangle$ and $a = a_C$, the transition probability is,
$$q(m|\langle 0, C\rangle, a_C) = \begin{cases} \frac{\lambda_1}{\lambda_1+\mu+\gamma}, & m = \langle 1, A\rangle, \\ \frac{\mu}{\lambda_1+\mu+\gamma}, & m = \langle 1, D\rangle, \\ \frac{\gamma}{\lambda_1+\mu+\gamma}, & m = \langle 1, C\rangle. \end{cases}$$
It is clear that
$$\begin{aligned} q(m|\langle 1, A\rangle, a_R) &= q(m|\langle 1, A\rangle, a_A) \\ &= q(m|\langle 1, D\rangle, a_C) \\ &= q(m|\langle 0, C\rangle, a_C). \end{aligned}$$
When the sensor is in the sleep status, $s = \langle 1, C\rangle$ and $a = a_C$, the transition probability is,
$$q(m|\langle 1, C\rangle, a_C) = \begin{cases} \frac{\lambda_0}{\lambda_0+\eta} & m = \langle 0, A\rangle, \\ \frac{\eta}{\lambda_0+\eta} & m = \langle 0, C\rangle. \end{cases}$$
For state $s = \langle 0, A\rangle$, and action $a = a_A$ or $a = a_R$, it is obvious that
$$\begin{aligned} q(m|\langle 0, A\rangle, a_R) &= q(m|\langle 0, A\rangle, a_A) \\ &= q(m|\langle 1, C\rangle, a_C). \end{aligned}$$
- **Reward functions:** Because the system state does not change between decision epochs, the expected discounted reward between epochs satisfies
$$\begin{aligned} r(s, a) &= k\big(s, a\big) + c(s, a)E_s^a\left\{\int_0^{\tau(s,a)} e^{-\alpha t}dt\right\} \\ &= k\big(s, a\big) + \frac{c(s, a)}{\alpha + \beta(s, a)}, \end{aligned}$$
where $r(s, a)$ is the reward obtained at status s with action a, $k\big(s, a\big)$ is the lump reward obtained when a new state and action is observed and $c\big(s, a\big)$ is the continuous rewards accumulated between decision epochs.

During data admission control process, the holding cost is quite important, which usually cannot be neglected in WSNs [15], [23]. So, the reward function consists of four different parts in our scheme: holding cost rate $f(i)$, reward R for a data packet's acceptance, switching 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 the analysis of our proposed model, we have

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

and

$$c(\langle\delta,b\rangle,a)=\begin{cases}-pf(L_S+1)-(1-p)f(L_S),\\ \qquad\qquad \delta=0,b=A,a=a_A,\\ -f(L_S),\delta=0,b=A,a=a_R,\\ -f(L_R),\delta=0,b=C,a=a_C,\\ -f(L_S),\delta=1,b=C,a=a_C,\\ -f(\overline{L_R}),\delta=1,b=D,a=a_C,\\ -qf(L_R+1)-(1-q)f(L_R),\\ \qquad\qquad \delta=1,b=A,a=a_A,\\ -f(L_R),\delta=1,b=A,a=a_R.\end{cases}$$

A policy is said to be α-optimal if its expected α-discounted return is maximal for every initial state. During each decision epoch, the status is stationary. The optimal policy is a stationary deterministic policy. From [24], the constraint function in equation (2) can now be replaced by the following Bellman equation:

$$v_\alpha(s)=\max_{a\in A}\left\{r(s,a)+\frac{\beta}{\alpha+\beta}\sum_{m\in S}q(m|s,a)v(m)\right\}. \quad (3)$$

By noting there is only one possible action a_C for states $s=\langle 0,C\rangle$, $s=\langle 1,C\rangle$, and $s=\langle 1,D\rangle$. Therefore, we have

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

$$\begin{aligned}&v(\langle 1,C\rangle)\\ &=-E_2+\frac{1}{\alpha+\lambda_0+\eta}\{-f(L_S)+\lambda_0 v(\langle 0,A\rangle)+\eta v(\langle 0,C\rangle)\},\end{aligned}$$

and

$$v(\langle 1,D\rangle)=v(\langle 0,C\rangle)+E_1+\frac{f(L_R)-f(\overline{L_R})}{\alpha+\lambda_1+\gamma+\mu}.$$

C. Optimal control policy

With the CTMDP model, here, we study the problem how to make sure the total expected discounted reward for an arrival data packet's acceptance of a sensor is not less than that for an arrival data packet's rejective.

Theorem 3. *For the $(p;q)$ scheme, to satisfy the constraints of admission control, the decision rule is*

$$d(\langle 1,A\rangle)=\begin{cases}a_A, & R\geq\frac{f(L_R+1)-f(L_R)}{\alpha+\lambda_1+\mu+\gamma},\\ a_R, & R<\frac{f(L_R+1)-f(L_R)}{\alpha+\lambda_1+\mu+\gamma},\end{cases}$$

and

$$d(\langle 0,A\rangle)=\begin{cases}a_A, & R\geq\frac{f(L_S+1)-f(L_S)}{\alpha+\eta+\lambda_0},\\ a_R, & R<\frac{f(L_S+1)-f(L_S)}{\alpha+\eta+\lambda_0}.\end{cases}$$

Proof: See the Appendix B. ■

Remarks. From this theorem, it can be seen that the switching on cost E_1 and switching off cost E_2 have no influence on the decision rule during admission process.

D. Optimal $(p;q)$ policy

With the above models of throughput and the admission control, in this subsection, we focus on the problem how to determine the optimal $(p;q)$ policy for the maximum throughput under the admission control constraints.

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

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

Definition 2. Let's define that

$$g_1(p,q)=f(L_R+1)-f(L_R)-R(\alpha+\lambda_1+\mu+\gamma),$$

and

$$g_2(p,q)=f(L_S+1)-f(L_S)-R(\alpha+\lambda_0+\eta).$$

From theorems 1-3, we finally can get that our objective function in Equation (1) can be written as

$$\begin{aligned}&\max_{p,q}\ \frac{\lambda_1\eta q+\lambda_0\gamma p}{\gamma+\eta}, \quad (4)\\ &\text{s.t.}\ \ g_i(p,q)\leq 0,\ i\in\{1,2\},\\ &\qquad 0\leq p\leq 1,\ 0\leq q\leq 1.\end{aligned}$$

To solve Equation (4), we can relax the admission control constraints and solve the problem by introducing the Lagrange multipliers method [25]. Let's define that $g_3(p,q)=p, g_4(p,q)=q, g_5(p,q)=p-1, g_6(p,q)=q-1$, and $L(p,q,\nu)=-Q(p,q)+\sum_j \nu_j g_j(p,q)$. Thus, Equation (4) can be written as

$$\begin{aligned}&\nabla L(p,q,\nu)=0, \quad (5)\\ &g_i(p,q)\leq 0,\ i\in\{1,2\}\\ &0\leq p\leq 1, 0\leq q\leq 1,\\ &\nu_i g_i(p,q)=0, \quad i=1,...,6,\\ &\nu_i\geq 0, \quad i=1,...,6.\end{aligned}$$

Theorem 4. *Assume that $f(i)$ is convex on i, where $i \geq 0$, and data packets' arriving rate is less than data packets' processing rate, which is to say $\eta(\mu - \lambda_1) > \gamma\lambda_0$, then, there exists a vector of multipliers $\nu^* \geq 0$ such that (p_0, q_0, ν^*) is a saddle point of the Lagrangian function*

$$L(p, q, \nu) = -Q(p, q) + \sum_j \nu_j g_j(p, q),$$

satisfying

$$L(p_0, q_0, \nu) \leq L(p_0, q_0, \nu^*) \leq L(p, q, \nu^*),$$

$\nu \geq 0$, and for all (p, q), where $p \in [0, 1]$ and $q \in [0, 1]$. Meanwhile, (p_0, q_0) is the optimal solution of Equation (4).

Proof: It is clear that $Q(p, q)$ is differential on p, q. $\eta(\mu - \lambda_1) > \gamma\lambda_0$, and $f(i)$ is convex on i, where $i \geq 0$, we can get that $g_i(p, q)$ is also differential on p, q. According to theorem 5.1.2 and theorem 5.1.3 in [26], there exists a vector of multipliers $\nu^* \geq 0$ such that (p_0, q_0, ν^*) is a saddle point of the Lagrangian function

$$L(p, q, \nu) = -Q(p, q) + \sum_j \nu_j g_j(p, q),$$

satisfying

$$L(p_0, q_0, \nu) \leq L(p_0, q_0, \nu^*) \leq L(p, q, \nu^*),$$

$\nu \geq 0$, for all (p, q), where $p \in [0, 1]$ and $q \in [0, 1]$. Meanwhile, (p_0, q_0) is the optimal solution of Equation (4).

This completes the proof. ■

Corollary 1. If let the probabilities for data admission in different status are the same ($p = q$), the cost function $f(i)$ is convex and non-decreasing on $i (i \geq 0)$, and $\eta(\mu - \lambda_1) > \gamma\lambda_0$, the optimal $(p; q)$ policy is

$$p = q = \min\{p : g_1(p, q) \geq 0, g_2(p, q) \geq 0\}.$$

- If $p = 0$, the sensor should reject all arrival packets, the throughput $Q = 0$;
- If $p = 1$, the sensor should accept all arrival packets, the throughput $Q = \frac{\lambda_1\eta + \lambda_0\gamma}{\gamma + \eta}$; and
- If $p \in (0, 1)$, the throughput $Q = \frac{\lambda_1\eta + \lambda_0\gamma}{\gamma + \eta} p$.

Proof: $\eta(\mu - \lambda_1) > \gamma\lambda_0$, thus, both $L_R(p, q)$ and $L_S(p, q)$ are non-decreasing on p and q. The cost function $f(i)$ is convex and non-decreasing on $i (i \geq 0)$. Therefore, both $g_1(p, q)$ and $g_2(p, q)$ are non-decreasing on p and q.

When the sensor is in the active status, the sensor node should accept data packets with probability q, where

$$q = \min\{p : g_1(p, q) \geq 0\}.$$

When the sensor is in the sleep status, the sensor node should accept data packets with probability p, where

$$p = \min\{p : g_2(p, q) \geq 0\}.$$

$p = q$, finally, we have

$$q = p = \min\{p : g_1(p, q) \geq 0, g_2(p, q) \geq 0\}.$$

From the function of the throughput, it is obvious that $Q(p, q)$ is non-decreasing on p and q. Thus, to get the maximum throughput, p and q should be as large as possible. With the constraints of data admission control, it is clear that if $p = 0$, the sensor should reject all arrival packets, the throughput $Q = 0$; if $p = 1$, the sensor should accept all arrival packets, the throughput $Q = \frac{\lambda_1\eta + \lambda_0\gamma}{\gamma + \eta}$; and if $p \in (0, 1)$, the throughput $Q = \frac{\lambda_1\eta + \lambda_0\gamma}{\gamma + \eta} p$.

The proof is completed. ■

V. Energy Consumption Analysis

The analysis of power consumption is very important in wireless sensor networks [27]. According to [15], we consider the energy consumption in terms of the sensor node status, number of packets transmitted, and the switches from one status to another. We have 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_s: the power consumption for sensing per unit time in the sleep status;
- e_{tr} : the transmitter power consumption per data packet in the active status;
- e_{or} : the operation power consumption per unit time in the active status;
- e_{os} : the operation power consumption per unit time in the sleep status.

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

Theorem 5. *For the $(p; q)$ scheme, 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 node consumes e_{sr} milliwatts of power each time it switches from the sleep status to the active status. The expected number of switching times from the sleep mode to the active mode per unit time is $\sum_i P(S_i)\eta$. Thus, we have

$$E_{SR} = \sum_i P(S_i)\eta e_{sr} = \frac{\gamma\eta}{\gamma + \eta} e_{sr}.$$

The node consumes e_{rs} milliwatts of power each time it switches from the active status to the sleep status. The expected number of switching times from the active mode to the sleep mode per unit time is $\sum_i P(R_i)\gamma$. Therefore,

$$E_{RS} = \sum_i P(R_i)\gamma e_{rs} = \frac{\gamma\eta}{\gamma + \eta} e_{rs}.$$

The proof is finished. ■

Remarks. From Theorem 5, with the increase of γ, the probability that the sensor is on the sleep status increases, and the rate of switching-off also increases. With the increase

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2997984, IEEE Sensors Journal

of η, the probability that the sensor is on the active status increases, and the rate of switching-on increases, too. Thus, with the increase of γ and η, the average switching energy consumed per unit time increases.

Theorem 6. *The average energy consumption in the sleep mode E_S is*

$$E_S = \frac{\gamma\lambda_0 p + \lambda_0\mu p + \lambda_1\eta q - \lambda_0\lambda_1 pq}{(\eta\mu - \gamma\lambda_0 p - \lambda_1\eta q)(\gamma+\eta)}\gamma e_s;$$

and the average energy consumption in the active mode E_{TR} is

$$E_{TR} = \frac{\gamma\lambda_0^2 p^2 + \gamma\lambda_0\eta p + \lambda_1\eta^2 q}{(\eta\mu - \gamma\lambda_0 p - \lambda_1\eta q)(\gamma+\eta)}e_{tr}.$$

Proof: When the node is in the sleep status, it would consume e_s milliwatts of power for per packet sensing. Thus, we have

$$E_S = \sum_{i=1}^{\infty} iP(S_i)e_s = \frac{\gamma\lambda_0 p + \lambda_0\mu p + \lambda_1\eta q - \lambda_0\lambda_1 pq}{(\eta\mu - \gamma\lambda_0 p - \lambda_1\eta q)(\gamma+\eta)}\gamma e_s.$$

When the sensor node is in the active status, it would consume e_{tr} milliwatts of power for per packet transmitting. So,

$$E_{TR} = \sum_{i=1}^{\infty} iP(R_i)e_{tr} = \frac{\gamma\lambda_0^2 p^2 + \gamma\lambda_0\eta p + \lambda_1\eta^2 q}{(\eta\mu - \gamma\lambda_0 p - \lambda_1\eta q)(\gamma+\eta)}e_{tr}.$$

The proof is finished. ■

Remarks. In theorem 6, as the increase of data packets' arriving rate λ_0, λ_1, both E_S and E_{TR} increase and with the increase of data packets' processing rate μ, both E_S and E_{TR} decrease.

Theorem 7. *The operation energy consumption in the sleep mode E_{OS} is*

$$E_{OS} = \frac{\gamma}{(\gamma+\eta)\lambda_0 p e_s}e_{os}E_S;$$

and the operation energy consumption in the active mode E_{OT} is

$$E_{OR} = \frac{\eta}{(\gamma+\eta)(\mu-\lambda_1 q)e_{tr}}e_{or}E_{TR}.$$

Proof: When the node is in the sleep status, it can sense variables from the environment and switch to Phase R. If we denote the operation time by $T_{S,i}$, when starting from the time when there are i data packets in phase S, we have the following probability distribution function:

$$P\{S|S_i < t\} = \frac{\gamma}{\gamma+\eta}\sum_{k=i}^{\infty} P_{S_k}(t) = \frac{\gamma}{\gamma+\eta}\sum_{k=i}^{\infty}\frac{(\lambda_0 pt)^k}{k!}e^{-\lambda_0 pt},$$

where $t \geq 0$.

Thus, we have the probability density function as follows:

$$f_{T_{S,i}} = \frac{(\lambda_0 pt)^{i-1}}{(i-1)!(\gamma+\eta)}e^{-\lambda_0 pt}\lambda_0 p\gamma,$$

where $t \geq 0$.

Therefore, the average energy consumption for the duration of the operation with i data packets received in the sleep status, can be given by

$$\begin{aligned} E_{OS,i} &= e_{os}E[T_{S,i}] \\ &= e_{os}\int_{t=0}^{\infty}\frac{(\lambda_0 pt)^i}{(i-1)!(\gamma+\eta)}e^{-\lambda_0 pt}\gamma dt \\ &= \frac{\gamma e_{os}}{(\gamma+\eta)\lambda_0 p}i. \end{aligned}$$

The average energy consumption for the operation of a sensor in the phase S is

$$E_{OS} = \sum_{i=1}^{\infty} P(S_i)E_{OS,i} = \frac{\gamma e_{os}}{(\gamma+\eta)\lambda_0 p e_s}E_S.$$

When the sensor node is in the active status, it can sense, receive, transmit data packets and switch to phase S. If we denote the operation time by $T_{R,i}$, when starting from the time when there are i data packets in phase R, we have the following probability distribution function:

$$P\{R|R_i < t\} = \frac{\eta}{\gamma+\eta}\sum_{k=i}^{\infty} P_{R_k}(t),$$

where $t \geq 0$.

According to [28], for an M/M/l queue, with arrival rate $\lambda_1 q$ and service rate μ, which begins operation at $t = 0$ with n data packets in the sensor, the probability that there are $i(i \geq 0)$ data packets in the sensor at time t is given by

$$\begin{aligned} P_{R_i}(t) =& e^{-(\lambda_1 q+\mu)t}[(\frac{\lambda_1 q}{\mu})^{(i-n)/2}I_{i-n}(2t\sqrt{\lambda_1 q\mu}) \\ &+ (\frac{\lambda_1 q}{\mu})^{(i-n-1)/2}I_{i+n+1}(2t\sqrt{\lambda_1 q\mu}) \\ &+ (1-\frac{\lambda_1}{\mu}q)(\frac{\lambda_1 q}{\mu})^i\sum_{j=i+n+2}^{\infty}(\frac{\lambda_1 q}{\mu})^{-j/2}I_j(2t\sqrt{\lambda_1 q\mu})] \end{aligned}$$

where $t \geq 0$, and $I_i(y)$ is the infinite series for the modified Bessel function of the first kind.

Considering the original $M/M/1$ queue with an absorbing barrier imposed at zero system size and an initial size of 1, then the expected length of the operation period should be $\frac{1}{\mu-\lambda_1 q}$ [28]. Therefore, the average energy consumption for the duration of the operation with i data packets received in the active status, can be given by

$$E_{OR,i} = e_{or}E[T_{R,i}] = \frac{\eta e_{or}}{(\gamma+\eta)(\mu-\lambda_1 q)}i.$$

The average energy consumption for the operation of a sensor in the phase R is

$$E_{OR} = \sum_{i=1}^{\infty} P(R_i)E_{OR,i} = \frac{\eta e_{or}}{(\gamma+\eta)(\mu-\lambda_1 q)e_{tr}}E_{TR}.$$

The proof is finished. ■

VI. Simulation Results and Analysis

To show the effectiveness of the approach above, we present the numerical results for a set of specific parameters in this section. In our simulation, all data simulations have been carried out with MATLAB 7.14.0 (R2012a) on Intel Core i5-4200M CPU (2.50 GHz, 4GB RAM). We have written MATLAB scripts to evaluate the discussed $(p; q)$ scheme. We perform the simulation for a WSN using the parameters as in [15]. Here we have simulated our model for a single node of a WSN. These evaluations are compared with an ordinary duty cycling scheme (the $(\gamma; \eta)$ model) which is similar to the model in [15]. The $(\gamma; \eta)$ model is a classical model, in which, on sleep status, the sensor node can only generate data packets, but on active status, it can generate data packets, receive data packets from other nodes and process data packets. For the $(\gamma; \eta)$ model, the sensor node receives all arrival data packets.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2997984, IEEE Sensors Journal

A. Discounted rewards performance

To evaluate the average discounted reward with these two different policies, we set the parameters for analysis as in TABLE II. Let the cost function be $f(i) = i^2$, $i \geq 0$, and discount factor $\alpha = 0.1$, which fits the theorems' requirements.

TABLE II: Simulation parameters

parameters	α	λ_1	λ_0	μ	γ	η	E_1	E_2
values	0.1	0.6	0.3	0.9	1	1.1	20	5

As shown in TABLE II, 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 $(p;q)$ policy. With the set parameters, according to the theorems above, we can obtain the following results.

Algorithm 1 Maximum throughput algorithm under the admission control policy

```
set R, E_1, E_2, f(i) /*set the reward value and cost value in sensor node*/
set δ = 0 as sleep, δ = 1 as active
initialize λ_0, λ_1, μ, α, β, η, γ /*the sensor node obtains the real-time features of itself */
calculate (p;q) according to equation (5)
if δ = 0 then
    the sensor node accepts arrival data with probability p
else
    the sensor node accepts arrival data with probability q
end if
```

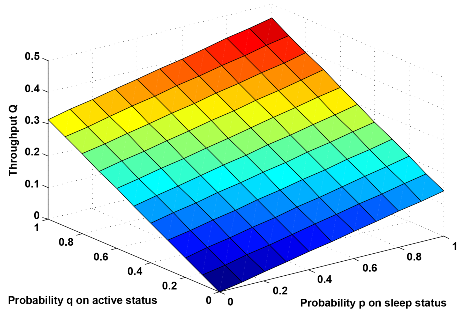

Fig. 3: Throughput of a sensor node.

Fig.3 shows the throughput of a sensor node. In Fig.3, with the increase of p and q, the value of Q increases. It is clear with $(\gamma;\eta)$ policy, the sensor node can achieve the highest throughput.

With the set parameters, our objective function in this simulation is

$$\begin{aligned} \max_{p,q} \quad & \frac{0.66q + 0.3p}{2.1}, \\ \text{s.t.} \quad & \frac{1.089 + 0.09p^2}{1.089 - 0.33p - 0.726q} - 13 \leq 0, \\ & \frac{1.089 + 0.24p - 0.066q - 0.18pq}{1.089 - 0.33p - 0.726q} - 7.5 \leq 0, \\ & 0 \leq p \leq 1,\ 0 \leq q \leq 1. \end{aligned}$$

According to the Lagrange multipliers method, we have $p_0 = 0.67$ and $q_0 = 1.0$. With our Corollary, if $p = q$, we have $p_0 = q_0 = 0.89$.

Fig. 4-6 illustrate the discounted reward of status $\langle 0, A\rangle$, $\langle 1, C\rangle$, $\langle 0, C\rangle$, $\langle 1, A\rangle$ and $\langle 1, D\rangle$ with policy $(\gamma;\eta)$, and different $(p;q)$ policies after 1000 times iterations. Since the value of the discounted reward is very large, here we let the value be $10log(v(s))$. In the simulation, $(p_0;q_0) = (0.67;0.89)$, $(p_1;q_1) = (0.67;1.0)$, $(p_2;q_2) = (0.89;0.89)$. From there figures, it is clear to see that the discounted reward of $(p;q)$ policy is larger than that of $(\gamma;\eta)$ policy.

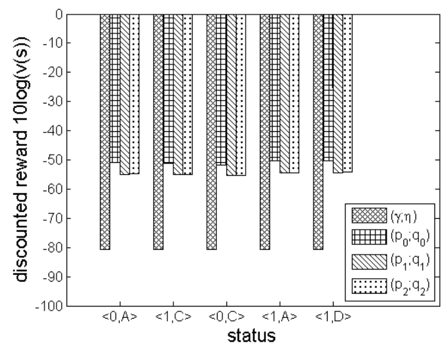

Fig. 4: Discounted reward $R = 5$.

In Fig. 4, the discounted reward of $(p_0;q_0)$ policy is the largest. This is because that the reward R for per admitted data packet is much smaller than its holding cost. It is better to reject all data packets to get maximum discounted reward. However, if $p = q = 0$, there will be no throughput.

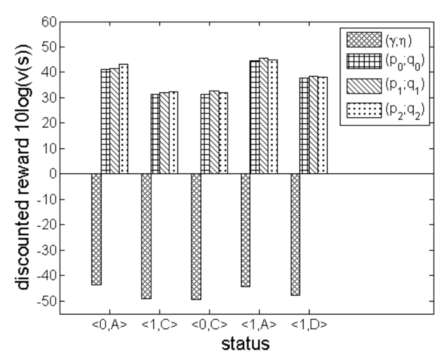

Fig. 5: Discounted reward with $R = 50$.

Fig. 5 shows the discounted reward of different policies when $R = 50$. Comparing with the results in Fig. 4, in Fig. 5, for all policies, as the increase of R, the discounted reward increases. In Fig. 5, the value of the discounted reward of the three $(p;q)$ policies are almost the same. However, the discounted reward of the $(p_1;q_1)$ policy is a little larger than the other two $(p;q)$ policies.

Fig. 6 shows the discounted reward of different policies when $R = 100$. From Fig. 4-6, we can get that as the increase of R, the value of the discounted reward of all policies increase. With different R, the gap of the discounted reward of the three $(p;q)$ policies are very small. From Fig. 5 and 6,

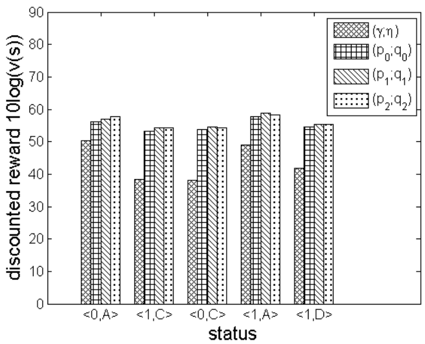

Fig. 6: Discounted reward with $R = 100$.

when the discounted reward is larger than zero, the discounted reward of $(p_1; q_1)$ policy is the largest.

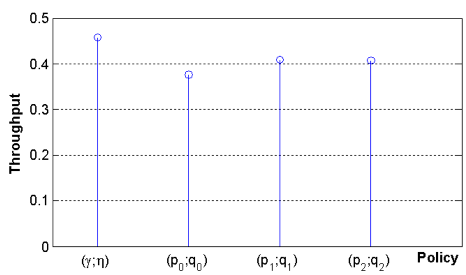

Fig. 7: Throughput with different policies.

Fig. 7 illustrates the throughput of different policies. It is clear that the throughput of $(\gamma; \eta)$ is the highest. For the three $(p; q)$ policies, the throughput of $(p_1; q_1)$ policy is the highest. From the simulation of throughput and discounted reward of different status for different policies, we can get that with $(\gamma; \eta)$ policy, the sensor node can achieve the highest throughput. However, the discounted reward of $(\gamma; \eta)$ policy is the smallest. For the $(p; q)$, if the reward of per admitted data packet is much large, there exists a $(p; q)$ policy to get a maximum throughput along with a maximum discounted reward.

B. Energy consumption performance

In this section, we analyze the energy performance of the optimal $(p; q)$ policy and $(\gamma; \eta)$ policy with holding 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 III. 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 E_{TR} is less than φ.

TABLE III: Simulation parameters of energy consumption

parameters	e_{tr}	e_{rs}	e_{sr}	e_s	e_{os}	e_{or}
values	$40\mu w$	$0.5\mu w$	$15\mu W$	$10\mu W$	$15\mu W$	$25\mu W$

Fig.8 shows the average energy consumption with different policies. The energy consumption on active status (E_{TR}, E_{OR}) is larger than that on sleep status (E_S, E_{OS}). It is clear to see that the average energy consumption of $(\gamma; \eta)$ policy is always larger than that of $(p; q)$ policy. For the different optimal $(p; q)$ policies, the energy consumption on sleep status for data sensing (or the energy consumption on active status for data transmitting) is almost the same. From Fig. 4-8, the optimal $(p; q)$ policy perform better than the $(\gamma; \eta)$ policy on the throughput, discounted reward and 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 $\lambda_0, \lambda_1, \mu, \alpha, \beta, \eta, \gamma$ /*the sensor node obtains the real-time features of itself */
calculate $(p; q)$with equation (4)
calculate E_{TR} according to Theorem 7
while $E_{TR} > \varphi$ **do**
    adjust λ_0
    calculate $(p; q)$with equation (4)
    calculate E_{TR} according to Theorem 7
end while

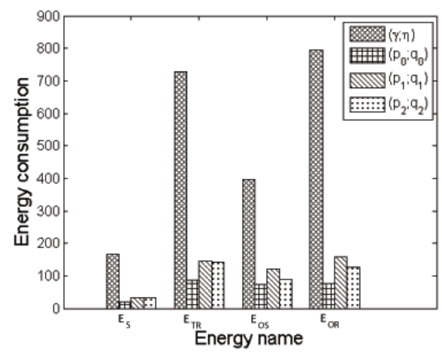

Fig. 8: Energy consumption with different policies.

VII. Conclusion

In this article, we have proposed a $(p; q)$ policy for the problem of the maximum throughput under data admission process in sensor nodes. A sensor node's working process can be modeled as a multi-phase queuing system in which the fist stage is used to manage the admission of arrivals and the second stage is used to manage the service. In the fist stage, the node can not know the information of the queue-length of arrival or newly sensed data packets. Thus, a $(p; q)$ model has been proposed in which a sensor generates data packets with probability p on sleep status and with probability q on active status. In the model, during data admission process, some reward for accepting data packets and a holding cost per unit time for data packets' delay in the sensor have been considered. We have observed that, under certain assumptions, there exists an optimal $(p;q)$ policy to obtain the maximum throughput during data admission process. The optimal value of p and q for the $(p;q)$ policy has been investigated. The performance on energy consumption of the proposed model has been studied. In addition, the throughput, expected discounted reward and energy performance of the model through numerical analysis have been performed. The results of this paper can be applied in designing optimal sensor nodes in wireless sensor networks.

Appendix A
Proof of Theorem 2

Proof:
Here, we define that

$$\rho=\gamma/\eta, \rho_0=\lambda_0/\mu, \rho_1=\lambda_1/\mu,$$
$$s_1=(A-\sqrt{A^2-4B})/2,$$
$$r_1=(A+\sqrt{A^2-4B})/2,$$

where

$$A=\frac{\lambda_0 p(\lambda_1 q+\gamma+\mu)+\lambda_1 q\eta}{\mu(\lambda_0 p+\eta)},$$
$$B=\frac{\lambda_0 pq\lambda_1}{\mu(\lambda_0 p+\eta)}.$$

From Theorem 1, for $i=0,1,2,\cdots$, we have

$$\pi_i=\pi_0 H^i.$$

With the assumption that

$$\eta\mu>\eta\lambda_1+\gamma\lambda_0,$$
$$p+q\neq 0,$$

So, we have $s_1\neq r_1$ and $r_1<1, s_1<1$. Thus, for $i=0,1,2,\cdots$,

$$\begin{cases} P(S_i)=\frac{s_1^{i+1}-r_1^{i+1}}{s_1-r_1}P(S_0), \\ P(R_i)=\frac{s_1^i-r_1^i}{s_1-r_1}\rho_0 pP(S_0)+[\frac{s_1^i-r_1^i}{s_1-r_1}\rho_1 q-\frac{r_1 s_1^i-s_1 r_1^i}{s_1-r_1}]P(R_0). \end{cases}$$

Therefore, in the sleep status, the average number of data packets in the sensor node is given by

$$\begin{aligned} L_S &= \sum_{i=0}^{\infty} iP(S_i) \\ &= \frac{-\lambda_0\lambda_1 pq+\gamma\lambda_0 p+\lambda_0\mu p+\lambda_1\eta q}{(\eta\mu-\gamma\lambda_0 p-\lambda_1\eta q)(\gamma+\eta)}\gamma, \end{aligned}$$

and in the active status, the average number of data packets is

$$L_R=\sum_{i=1}^{\infty} iP_{1,i}=\frac{\gamma\lambda_0^2 p^2+\gamma\lambda_0\eta p+\lambda_1\eta^2 q}{(\eta\mu-\gamma\lambda_0 p-\lambda_1\eta q)(\gamma+\eta)}.$$

If $p+q=0$, it is clear that $L_R=L_S=0$.
The proof is completed. ■

Appendix B
Proof of Theorem 3

Proof: According to the analysis above, when the sensor node is in the active status with $s=\langle 1,A\rangle$, we have,

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

From the result, we can get,

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

If an action a_A is taken on the state $\langle 1,A\rangle$, by doing a similar analysis as above, we will know in general that

$$\begin{aligned} &v(\langle 1,A\rangle,a_A) \\ =&\frac{\alpha+\lambda_1+\mu+\gamma}{\alpha+\beta}Rq+\frac{1}{\alpha+\beta}\{-f(L_R+1)q-f(L_R)(1-q) \\ &+\lambda_1 v(\langle 1,A\rangle)+\gamma v(\langle 1,C\rangle)+\mu v(\langle 1,D\rangle)+\eta v(\langle 1,A\rangle)\}. \end{aligned}$$

If action a_A is the best action at the state $\langle 1,A\rangle$,

$$v(\langle 1,A\rangle,a_A)=\frac{f(L_R)-f(L_R+1)}{\alpha+\lambda_1+\mu+\gamma}q+v(\langle 0,C\rangle)+E_1+Rq.$$

From the above analysis, it is not too hard to verify that,

$$\begin{aligned} &v(\langle 1,A\rangle) \\ =&E_1+v(\langle 0,C\rangle)+\max\{\frac{f(L_R)-f(L_R+1)}{\alpha+\lambda_1+\mu+\gamma}q+Rq,0\}. \end{aligned}$$

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

$$v(\langle 0,A\rangle,a_R)\geq v(\langle 1,C\rangle)+E_2,$$

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

$$v(\langle 0,A\rangle,a_R)=v(\langle 1,C\rangle)+E_2.$$

Furthermore, we have in general,

$$v(\langle 0,A\rangle,a_A)\geq E_2+pR+v(\langle 1,C\rangle)-\frac{f(L_S+1)-f(L_S)}{\alpha+\eta+\lambda_0}p,$$

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

$$v(\langle 0,A\rangle,a_A)=E_2+Rp+v(\langle 1,C\rangle)-\frac{f(L_S+1)-f(L_S)}{\alpha+\eta+\lambda_0}p.$$

Finally, we will also have the following result,

$$\begin{aligned} &v(\langle 0,A\rangle) \\ =&\max\{Rp-\frac{f(L_S+1)-f(L_S)}{\alpha+\eta+\lambda_0}p,0\}+v(\langle 1,C\rangle)+E_2. \end{aligned}$$

For the data admission control model, since both the state space S and the action space A are finite, the reward function $r(s,a)$ is also finite. From Theorem 11.3.2 of [7], the optimal policy is a stationary deterministic policy. Thus, our problem can be reduced to as finding a deterministic decision rule. From the equations above, for states $\langle 1,A\rangle$ and $\langle 0,A\rangle$, we have the decision rule

$$d(\langle 1,A\rangle)=\begin{cases} a_A, & R\geq\frac{f(L_R+1)-f(L_R)}{\alpha+\lambda_1+\mu+\gamma}, \\ a_R, & R<\frac{f(L_R+1)-f(L_R)}{\alpha+\lambda_1+\mu+\gamma}, \end{cases}$$

and

$$d\big(\langle 0, A\rangle\big) = \begin{cases} a_A, & R \geq \frac{f(L_S+1)-f(L_S)}{\alpha+\eta+\lambda_0}, \\ a_R, & R < \frac{f(L_S+1)-f(L_S)}{\alpha+\eta+\lambda_0}. \end{cases}$$

This completes the proof. ■

References

[1] M. Ergen and P. Varaiya, "Throughput analysis and admission control for ieee 802.11 a," *Mobile networks and Applications*, vol. 10, no. 5, pp. 705–716, 2005.

[2] X. Zhang, D. Li, W. W. Li, and W. Zhao, "An optimal dynamic admission control policy and upper bound analysis in wireless sensor networks," *IEEE Access*, vol. 7, pp. 53314–53329, 2019.

[3] K. Ren, K. Zeng, W. Lou, and P. J. Moran, "On broadcast authentication in wireless sensor networks," in *International Conference on Wireless Algorithms, Systems, and Applications*, pp. 502–514, 2006.

[4] A. Bachir, M. Dohler, T. Watteyne, and K. K. Leung, "Mac essentials for wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 2, pp. 222–248, 2010.

[5] G. Sunitha, S. D. Kumar, and B. V. Kumar, "A pre-emptive multiple queue based congestion control for different traffic classes in wsn," in *International Conference on Circuits, Communication, Control and Computing*, pp. 212–218, IEEE, 2014.

[6] X. Yin, X. Zhou, P. Miao, and S. Li, "Admission control with multi-constrained qos providing in wireless sensor networks.," 2010.

[7] J. Zhao, T. Q. Quek, and Z. Lei, "Heterogeneous cellular networks using wireless backhaul: Fast admission control and large system analysis," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2128–2143, 2015.

[8] A. Ksentini, T. Taleb, and K. B. Letaif, "Qoe-based flow admission control in small cell networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2474–2483, 2015.

[9] S. Li, N. Zhang, S. Lin, L. Kong, A. Katangur, M. K. Khan, M. Ni, and G. Zhu, "Joint admission control and resource allocation in edge computing for internet of things," *IEEE Network*, vol. 32, no. 1, pp. 72–79, 2018.

[10] S. T. V. Pasca, V. Srividya, and K. Premkumar, "Energy efficient sleep/wake scheduling of stations in wireless networks," in *2013 International Conference on Communication and Signal Processing*, pp. 382–386, IEEE, 2013.

[11] S. P. Li and Z. P. He, "A aqm model with multi-priority and multi-receiver for mobile ad hoc networks," in *2007 International Conference on Wireless Communications, Networking and Mobile Computing*, 2007.

[12] . G.-R. A. Munir, A., "An mdp-based dynamic optimization methodology for wireless sensor networks," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 23(4), pp. 616–625, 2010.

[13] J. L. J. Luo, L. J. L. Jiang, and H. C. H. Chen, "Finite queuing model analysis for energy and qos tradeoff in contention-based wireless sensor networks," 2007.

[14] Y. O. Y. Ou and W. Heinzelman, "Modeling and throughput analysis for smac with a finite queue capacity," 2010.

[15] Y. Zhang and W. Li, "Modeling and energy consumption evaluation of a stochastic wireless sensor network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, p. 282, 2012.

[16] Y. Iraqi and N. Jabeur, "Service admission control in multi-service sensor networks," in *Proceedings of the 14th IEEE Symposium on Computers and Communications (ISCC 2009), July 5-8, Sousse, Tunisia*, 2009.

[17] D. C. Huang and J.-H. Lee, "A dynamic n threshold prolong lifetime method for wireless sensor nodes," *Mathematical & Computer Modelling*, vol. 57, no. 11-12, pp. 2731–2741.

[18] L. D.-L. W. W. . Z. W. Zhang, X., "An optimal dynamic admission control policy and upper bound analysis in wireless sensor networks," vol. 7, pp. 53314–53329, 2019.

[19] A. Boulis and M. Srivastava, "Node-level energy management for sensor networks in the presence of multiple applications," *Wireless Networks*, vol. 10, no. 6, pp. 737–746.

[20] S. Jabbar, A. Minhas, M. Imran, S. Khalid, and K. Saleem, "Energy efficient strategy for throughput improvement in wireless sensor networks," *Sensors*, vol. 15, no. 2, pp. 2473–2495, 2015.

[21] X. Liu, D. He, and H. Ding, "Throughput maximization for uav-enabled full-duplex relay system in 5g communications," *Physical Communication*, vol. 32, pp. 104–111, 2019.

[22] S. Ji, R. Beyah, and Z. Cai, "Snapshot and continuous data collection in probabilistic wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 3, pp. 626–637, 2013.

[23] W. Ni, W. Li, and M. Alam, "Determination of optimal call admission control policy in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 1038–1044, 2009.

[24] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[25] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[26] W. Forst and D. Hoffmann, *Optimizationłtheory and practice*. Springer Science & Business Media, 2010.

[27] A. S. Toor and A. Jain, "Energy aware cluster based multi-hop energy efficient routing protocol using multiple mobile nodes (meacbm) in wireless sensor networks," *AEU-International Journal of Electronics and Communications*, vol. 102, pp. 41–53, 2019.

[28] J. F. Shortle, J. M. Thompson, D. Gross, and C. M. Harris, *Fundamentals of queueing theory*, vol. 399. John Wiley & Sons, 2018.

Xiaolu Zhang received her Bachelors degree in Communication Engineering from Donghua University, Shanghai, China, in 2012, where she is currently pursuing a Ph.D. degree in control theory and control engineering with the College of Information Sciences and Technology. From August of 2016 to 2018, she was studying as a visiting student at the Department of Computer Science, Texas Southern University, Houston, TX 77004, USA. Her current research interests include intelligent systems, intelligent control, network control systems and the application of wireless sensor networks.

Demin Li received his 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 over 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 in Shanghai.

Yihong Zhang received his Ph.d. degree from HongKong polytechnic university. He is currently an Assistant Dean of College of Information Science & Technology, DongHua university, Shanghai, China. He is the supervisor of postgraduates, an Associate professor and an IEEE member. His mainly research area focus on intelligent system and control, Pattern recognition and robot control, intelligent sensor technology, intelligent digital textile and garment technology, intelligent wearable technology, and etc.