Optimal Management of Rechargeable Biosensors in Temperature-Sensitive Environments

Yahya Osais, F. Richard Yu, and Marc St-Hilaire
Department of Systems and Computer Engineering
Carleton University, Ottawa, ON, K1S 5B6 Canada
yosais@sce.carleton.ca, {richard_yu, marc_st_hilaire}@carleton.ca

***Abstract*—Biosensors are tiny wireless medical devices which are attached or implanted into the body of a human being or animal to monitor and control biological processes. They are distinguished from conventional sensors by their biologically derived sensing elements. Biosensors generate heat when they transmit their measurements and their temperature rises when recharged by electromagnetic energy. These phenomena translate to a temperature increase in the tissues surrounding the biosensors. If the temperature increase exceeds a certain threshold, the tissues might be damaged. In this paper, we discuss the problem of finding an optimal operating policy for a rechargable biosensor under a strict maximum temperature increase constraint. This problem can be formulated as a Markov decision process with an average reward criterion. The solution is an optimal policy that maximizes the average number of samples which can be generated by the biosensor while observing the constraint on the maximum safe temperature level. Due to the exponential nature of the problem, a heuristic policy is proposed. The performance of the policies is studied through simulation. A greedy policy is used as a baseline for comparison.**

I. Introduction

Biosensors are tiny wireless devices which are attached or implanted into the body of a human or animal to monitor and control biological processes. Unlike conventional wireless sensors, biosensors are energy- as well as temperature-constrained. Also, their sensing elements are biological materials such as enzymes and antibodies which are integrated into transducers for producing electrical signals in response to biological reactions and changes. They are powered by either rechargable built-in batteries or by continuosly sending electric energy in the form of electromagnetic waves. Two famous applications of biosensors are the geodesic sensor network developed by EGI corporation [1] and the artificial retina [2].

The use of batteries necessitates periodic recharging which can be performed using energy resulting from vibration, motion, light and heat. However, a more mature approach is to wirelessly collect energy from a radio frequency (RF) source and then convert it into usable power. This approach is widely used in industry to transfer data and power to biosensors. It is also more practical since many sensors can be recharged simultaneously.

Biosensors are temperature-constrained due to the heat generated as a result of their operation in temperature-sensitive environments like the human body. Radiation due to wireless communication and recharging are the major sources of heat. The generated heat manifests itself as a temperature increase inside the tissues which might be damaged if the blood flow is less than optimal.

In this paper, we study a stochastic control problem which arises when a rechargable biosensor operates in a temperature-constrained environment like the human body. In this problem, the state of the biosensor is characterized by its current temperature and energy levels and uncertainty exists due to the random behavior of the wireless channel between the biosensor and base station. The objective is to operate the biosensor in such a way that the average number of samples generated by the biosensor is maximized while the maximum safe temperature level is not exceeded. To that end, the control problem is formulated as a Markov Decision Process (MDP) and solved to obtain an optimal operating policy.

The remainder of the paper is organized as follows. First, related work is presented. Second, the system model and its MDP formulation are described. Besides, an example is given to illustrate the viability of the proposed MDP model. After that, a greedy policy and heuristic policy are discussed. Then, they are compared with the optimal policy using simulation. Finally, conclusions and further directions for research are given.

II. Related Work

Tang et al. [3] were the first to propose rotating the cluster leadership in a cluster-based biosensor network to minimize the heating effects on human tissues. They proposed a genetic algorithm for computing a minimal temperature increase rotation sequence. Since computing the temperature increase due to a sequence is computationally expensive, they proposed a scheme for estimating the possible temperature increase due to a sequence.

In another work, Tang et al. [4] addressed the issue of routing in implanted biosensor networks. They proposed a thermal-aware routing protocol that routes the data away from high temperature areas referred to as hot spots. The location of a biosensor becomes a hot spot if the temperature of the biosensor exceeds a predefined threshold. The proposed protocol achieves a better balance of temperature increase and shows the capability of load balance.

The above two works have motivated us to explore further the bioeffects of implanted biosensor networks. As a result, we noticed a lack of information on how to optimally operate an implanted biosensor network when bounds such as the

978-1-4244-3574-6/10/$25.00 ©2010 IEEE

maximum temperature increase exist. Most of the existing works assume that energy is the only limiting factor in the operation of wireless sensor networks. However, this is not the case in biosensor networks where the increase in temperature is a serious limiting factor.

We have approached the problem of how to optimally operate a biosensor network from the perspective of sensor scheduling and activation in traditional wireless sensor networks. Sensor scheduling is concerned with the problem of how to dynamically choose a sensor for communication with the base station. On the other hand, sensor activation is concerned with the problem of when a sensor should be activated. Many interesting works have been done in this regard. Next, these works are breifly reviewed.

In [5], the sensor scheduling problem is formulated as an MDP. The objective is to find an operating policy that maximizes the network lifetime. The state of a sensor is characterized by its current energy level only. Considering only the energy level at each sensor gives rise to an acyclic (i.e., loop-free) transition graph which enables the MDP model to converge in one iteration. On the other hand, if the temperature of each sensor is included in the model, the transition graph of the underlying MDP becomes cyclic. This is because when the sensor cools down (i.e., its temperature decreases), it transitions back to a less hot state. An MDP model whose transition graph is cyclic needs more time to converge.

Dynamic sensor activation in networks of rechargable sensors is considered in [6]. The objective is to find an activation policy that maximizes the event detection probability under the constraint of slow rate of recharge of the sensor. The state of the system is characterized by the energy level of the sensor and whether or not an event would occur in the next time slot. The recharge event is random and recharges the sensor with a constant charge. The model does not include the state of the wireless channel which is very crucial when temperature is considered.

Body sensor networks [7] with energy harvesting capabilities are another kind of WSNs in which each sensor has an energy harvesting device that collects energy from ambient sources such as vibration, light and heat. In this way, the more costly recharging method which uses radiation is avoided. The interaction between the battery recharge process and transmission with different energy levels is studied in [8]. The proposed policies utilize the sensor's knowledge of its current energy level and the state of the processes governing the generation of data and battery recharge to select the appropriate transmission mode for a given state of the network.

III. System Model

Figure 1 shows the system under study where a mobile subject, in this case an animal, has a biosensor implanted into its body. Both the biosensor and RF power source are under the control of the base station which initiates the measurement process. The base station generates three control signals: *Sleep* and *Send* targeted at the biosensor and *Recharge* targeted at the RF power source. The system state information is assumed to be available to the base station before it generates a control signal.

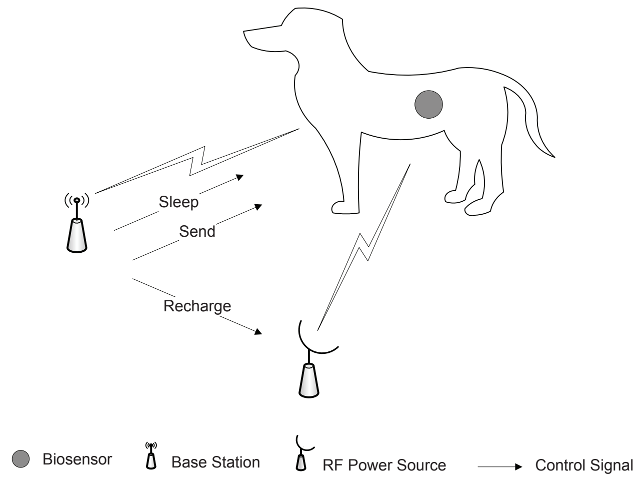

Fig. 1. Setup of the system under study.

The biosensor has a built-in battery which is recharged by an RF power source. The role of the biosensor is to monitor and report interesting physiological events such as heart rate and blood pressure. The location of the biosensor represents a critical point since it experiences the maximum temperature increase. This is because the tissues surrounding the biosensor might be heated continuously due to the local radiation generated by the biosensor itself and the radiation generated by the base station while recharging the biosensor. The biosensor becomes incapable of detecting and reporting events if it does not have enough energy for transmission under any channel condition or the increase in its temperature reaches a prespecified threshold. In this case, the biosensor is either recharged or put into sleep in order for its temperature to decrease.

The state of the biosensor is characterized by two variables which are the current temperature T_t and energy level E_t. There are $\tau + 1$ safe temperature levels; i.e., $T_t \in \{0, 1, ..., \tau\}$ where the zero temperature level represents the normal body temperature and τ is an upper limit which must not be exceeded. Initially, the biosensor has a total energy of $\mathcal{E}_0$ which is also the capacity of its battery. The energy required for the biosensor to successfully transmit its measurement to the base station is determined by the state of the channel at the time of transmission. This transmission energy is denoted by $\mathcal{E}_{w_i}$, where w_i is the i^{th} state of the wireless channel. The temperature increase due to a transmission energy of $\mathcal{E}_{w_i}$ units is denoted by $\mathcal{T}_{w_i}$.

At any instant of time, the base station may decide to recharge the biosensor, let it transmit its measurement or put it into sleep. The time required for a full recharge is random. During this time, interesting events may occur but they will not be reported by the biosensor since it is being recharged. Also, the biosensor may be put into sleep for a random amount of time during which no measurements can be produced.

The system can mathematically be modeled as a discrete-state system which evolves in discrete time. Therefore, the time axis is divided into slots of equal duration Δt. At the begining of each time slot, the state of the system is observed and a control signal is generated by the base station accordingly. Each time slot is long enough to transmit a complete packet carrying a measurement. The state of the biosensor is updated as follows. First, the energy level at the biosensor is given by the following equation:

$$E_{t+1} = \begin{cases} E_t & if \quad a_t = Sleep \\ E_t - \mathcal{E}_{w_i} & if \quad a_t = Send \\ E_t + \mathcal{E}_r & if \quad a_t = Recharge \end{cases} \tag{1}$$

where a_t is the action taken by the base station at time t and $\mathcal{E}_r$ is the amount of recharge energy gained by the biosensor. Similarly, the temperature of the biosensor at $t+1$ is given by the following equation:

$$T_{t+1} = \begin{cases} \max\{T_t - \mathcal{T}_s, 0\} & if \quad a_t = Sleep \\ T_t + \mathcal{T}_r & if \quad a_t = Recharge \\ T_t + \mathcal{T}_{w_i} & if \quad a_t = Send \end{cases} \tag{2}$$

where $\mathcal{T}_s$ is the amount by which the temperature of the biosensor decreases when it is put to sleep. In the same way, $\mathcal{T}_r$ and $\mathcal{T}_w$ are the amounts by which the temperature of the biosensor increases when it is recharged and when it is allowed to transmit its measurement, respectively.

$\mathcal{T}_{w_i}$ is not constant since the Specific Absorption Rate (SAR) due to the biosensor changes with the change in transmission energy. SAR is a measure of the level of radiation absorbed by the human body when exposed to RF radiation and is expressed in units of W/Kg. SAR records the rate at which radiation energy is absorbed per unit mass of tissue [9]. It is a function of the current provided to the antenna of the biosensor in accordance with the transmission energy requirement for each channel state. Therefore, before $\mathcal{T}_{w_i}$ due to a transmission can be calculated, the induced SAR must be computed. Then, $\mathcal{T}_{w_i}$ can be calculated using the Pennes's bioheat equation [10] after it is transformed into a discrete form using the finite-difference time-domain method [3], [11].

IV. MDP Formulation

The purpose of the MDP formulation of the system described in the previous section is to find a policy π that prescribes the best action to take in each possible state of the system so as to maximize the long-term expected number of samples generated by the biosensor. The policy π will turn out to be a stationary one. That is, it is independent of time and depends only on the state of the system. The elements of the MDP model are given in this section.

A. State Set

The state of the system at time t is described by the following 3-dimensional vector:

$$s_t = (T_t, E_t, W_t), \tag{3}$$

where T_t, E_t and W_t are the current temperature of the biosensor, its energy level and transmission power required for successful transmission at time t, respectively. The total number of system states is $|S| = |T| \times |E| \times |W|$, where $|T|$, $|E|$ and $|W|$ are the numbers of possible temperatures, residual energies and transmission energy levels, respectively.

B. Action Set

In each time slot, the base station chooses an action based on the current state of the system. There are three possible actions: *Sleep*, *Recharge* and *Sample*. The *Sleep* action puts the biosensor into sleep so that its temperature can decrease, the *Recharge* action recharges the biosensor and the *Sample* action lets the biosensor generate a measurement and transmit it to the base station. The *Sleep* action can be performed at every system state. The *Recharge* and *Sample* actions, however, can only be performed at system states where the next temperature of the biosensor is within the safe temperature range. In addition, the *Sample* action can only be performed at system states where the remaining energy is sufficient to make a successful transmission.

C. Reward Function

Since the objective is to maximize the expected number of samples that can be generated by the biosensor, the reward function is defined as

$$R(s, Sample) = 1.$$

This means that one unit of reward is earned every time the *Sample* action is performed. The long-run expected sum of rewards represents the average number of samples that can be generated by the biosensor with an initial energy of $\mathcal{E}_0$ units and maximum temperature increase of τ units.

D. Transition Probability Function

After the action taken by the base station is performed, the system transits to a new state according to the transition probabilities of the present state of the wireless channel. Thus, the behavior of the system is described by $|A|$ transition probability matrices, each of size $|S| \times |S|$. Each matrix is denoted by $P_{s_t,s_{t+1}}(a)$ which is the probability that choosing an action a when in state s_t will lead to state s_{t+1}. More formally, $P_{s_t,s_{t+1}}(a)$ can be written as the following:

$$P[s_{t+1}|s_t, a_t] = P[W_{t+1}|W_t] \tag{4}$$

E. Value Function

The problem of finding an optimal policy for maximizing the average number of samples is formulated as an infinite-horizon MDP using the average reward criterion [12]. So, let $V_\pi(s_0)$ be the expected number of samples given that the policy π is used with an initial state s_0. Then, the maximum expected number of samples $V_{\pi^*}(s_0)$ starting from state s_0 is given by

$$V_{\pi^*}(s_0) = \max_\pi V_\pi(s_0) \tag{5}$$

where π^* is the optimal policy that achieves the maximum expected number of samples at all system states.

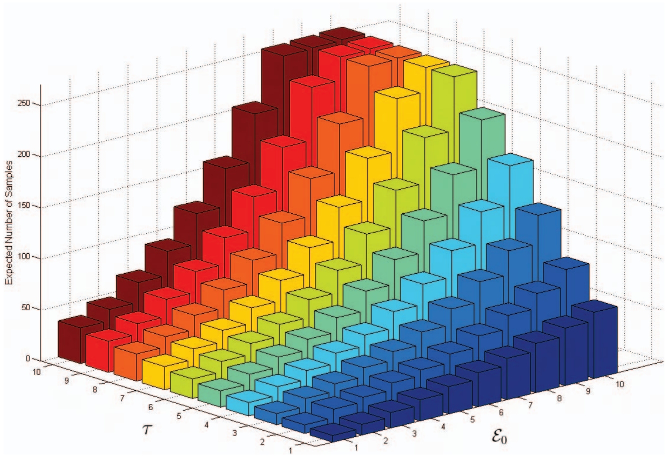

Fig. 2. Expected number of samples as a function of τ and $\mathcal{E}_0$.

Algorithm 1 Greedy Policy

Require: S: Set of possible system states
A: Set of possible actions at each system state

for $i = 1$ to $|S|$ **do**
  if Action(i, Sample) is True **then**
    $Policy(i) = Sample$
  else if Action(i, Recharge) is True **then**
    $Policy(i) = Recharge$
  else
    $Policy(i) = Sleep$
  end if
end for

The famous value iteration algorithm [13] is used to numerically solve the following recursive equation for $n > 0$

$$V_n(s) = \max_{a \in A(s)} \left[R(s,a) + \sum_{s_{t+1} \in S} \mathbb{P}(s_t, s_{t+1}, a) V_{n-1}(s_{t+1}) \right] \tag{6}$$

In (6), the subscript n denotes the iteration index. As $n \to \infty$, $V_n \to V_{\pi^*}$.

V. Example

Figure 2 shows the expected number of samples which is expressed as a function of the maximum safe temperature level (τ) and initial energy ($\mathcal{E}_0$). Clearly, τ plays a critical role. This is due to the temperature increase caused by the recharge action. For example, for the same initial energy, the expected number of samples increases as τ is varied. Increasing τ enables the *Recharge* action to be performed more often. On the other hand, as one would expect, if τ is fixed and ($\mathcal{E}_0$) is varied, the expected number of samples slightly increases when τ is small. However, when τ is large (≥ 6), the maximum possible expected number of samples can be achieved when $\mathcal{E}_0$ is at its maximum value. Therefore, for this particular example, if $\mathcal{E}_0 = 10$, the optimal value for τ is 6.

VI. Heuristic Policy

Algorithm 2 Heuristic Policy

Require: S: Set of possible system states
A: Set of possible actions at each system state

$\alpha = \frac{T}{\tau}$
$\beta = \frac{E}{E_0}$
for $i = 1$ to $|S|$ **do**
  if Action(i, Sample) is True & $\alpha \leq \beta$ **then**
    $Policy(i) = Sample$
  else if Action(i, Recharge) is True & $\alpha \geq \beta$ **then**
    $Policy(i) = Recharge$
  else
    $Policy(i) = Sleep$
  end if
end for

Since it is difficult to describe the structure of the optimal policy, a heuristic policy is proposed in this section. The goal is to design a policy which mimics the behavior of the optimal policy as close as possible. However, before presenting such a policy, a greedy one is given to provide insight into the design of any heuristic policy.

The greedy policy is computed using Algorithm 1. The inputs to this algorithm are the set of possible system states and set of feasible actions for each system state. The computed policy is greedy in the sense that for each system state, the feasibility of actions is checked in the following order: *Sample*, *Recharge* and *Sleep*. The first feasible action is associated with the corresponding system state.

As will be shown by simulations in the next section, the greedy policy is poor since it is based on a fixed order of actions. Therefore, Algorithm 1 needs to be extended to allow for a dynamic selection of actions. This objective is accomplished by introducing two control parameters: α and β. With these two control parameters, the *Sample* and *Recharge* actions are not selected in a specific order or whenever they are feasible. Algorithm 2 shows how the new heuristic policy is computed.

The essence of Algorithm 2 is as follows. If the current temperature (denoted by T) of the biosensor is low and its current energy level (denoted by E) is high, then the condition $\alpha \leq \beta$ would more likely be true and thus the *Sample* action could be executed. However, this would not be the case when the available energy is very close to zero. In this case, the opposite condition (i.e., $\alpha \geq \beta$) would more likely be true and thus a *Recharge* could be performed. If neither of the two conditions is true, the biosensor is put to sleep and thus its temperature decreases.

VII. Simulation Results

In this section, the performance of the optimal, greedy and heuristic policies is compared using simulation. The impact of various system parameters on the performance of the system is also evaluated. The simulation was performed using a simulator written in Matlab [14]. Each simulation was run

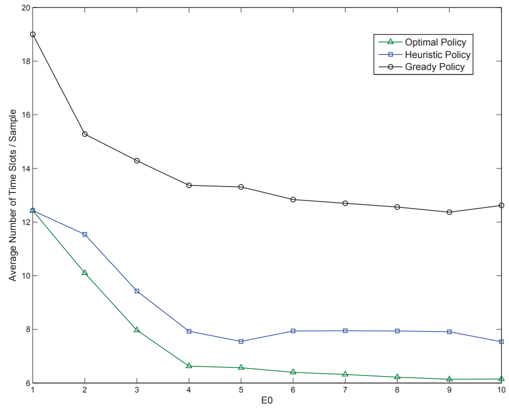

Fig. 3. Average number of time slots needed to generate a sample when τ is fixed at 5 and E_0 is varied.

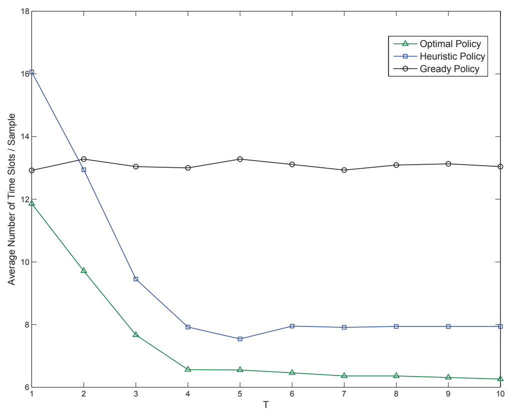

Fig. 4. Average number of time slots needed to generate a sample when E_0 is fixed at 5 and τ is varied.

for a duration of 100000 time slots and each data point is the average of 10 simulation runs. The number of channel states (W) is four and the channel state boundaries are randomly generated.

In order to be able to appreciate the merit of any heuristic policy, a more meaningful criterion is needed. Therefore, in this work, *the average number of time slots needed to generate a sample* is used as a criterion to distinguish between the different policies available to run the system. For example, consider Figure 3. In this Figure, τ is fixed at five while E_0 is varied from one to ten. The greedy policy is very costly since it requires the largest amount of time before a sample can be generated. The difference in the amount of time required by the heuristic policy and that required by the optimal policy stays around two time slots. This is a 75% reduction in time when compared to the greedy policy.

Figure 4 shows the amount of time required to generate a sample when E_0 is fixed at 5 and τ is varied. In this figure, when $\tau = 1$, the greedy policy outperforms the heurisitc policy. A difference of three time slots is observed. This can be explained as follows. In the heuristic policy, the *Recharge* action can be performed in one state only (i.e., when $T = E = 0$. On the other hand, with the greedy policy, the *Recharge* action can be performed in more than one state (i.e., whenever $T = 0$). This, of course, leads to a reduction in the average amount of time needed to generate a sample. Other than that, for $\tau \geq 2$, the heuristic policy is always better than than the greedy policy and its performance is close to that of the optimal policy.

VIII. Conclusions

The increase in temperature due to the heat generated by biosensors is a limiting factor in the operation of biosensor networks. This problem can be modeled as a stochastic control problem using the framework of Markov decision processes. The solution is an optimal policy which ensures that the maximum safe temperature level cannot be exceeded. In order to mitigate the problem of state explosion, we have proposed a heurisitc policy whose performance is comparable to the optimal policy. We have also proposed a performance criterion which more faithfully measures the performance of an operating policy for a biosensor network with a recharge capability.

References

[1] Geodesic Sensor Networks, EGI Corporation. [Online]. Available: http://www.egi.com

[2] L. Schwiebert, S. K. Gupta, and J. Weinmann, "Research challenges in wireless networks of biomedical sensors," in *Proc. Conference on Mobile Computing and Networking*. ACM, 2001, pp. 151–165.

[3] Q. Tang, N. Tummala, S. Gupta, and L. Schwiebert, "Communication scheduling to minimize thermal effects of implanted biosensor networks in homogeneous tissue," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 7, pp. 1285–1294, July 2005.

[4] ——, "Tara: Thermal-aware routing algorithm for implanted sensor networks," in *Distributed Computing in Sensor Systems*, vol. 3560. Springer, 2005, pp. 206–217.

[5] Y. Chen, Q. Zhao, V. Krishnamurthy, and D. Djonin, "Transmission scheduling for optimizing sensor network lifetime: A stochastic shortest path approach," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2294–2309, May 2007.

[6] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable sensor activation under temporally correlated events," *Wireless Netwoeks*, vol. 15, no. 5, pp. 619–635, July 2009.

[7] G. Z. Yang, *Body Sensor Networks*. Springer, 2006.

[8] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," in *Proc. Conference on Information Sciences and Systems*. IEEE, 2008, pp. 704–709.

[9] National Council on Radiation Protection and Measurements (NCRP), "A practical guide to the determination of human exposure to radiofrequency fields," Report No. 119, 1993.

[10] H. H. Pennes, "Analysis of tissue and arterial blood temperature in the resting human forearm," *Journal of Applied Physiology*, vol. 1, no. 1, pp. 93–122, July 1948.

[11] D. M. Sullivan, *Electromagnetic Simulation Using the FDTD Method*. IEEE Press, 2000.

[12] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 2005.

[13] D. P. Bertsekas, *Dynamic Programming and Optimal Control, Volume I*. Athena Scientific, 2000.

[14] The MathWorks, Inc. [Online]. Available: http://www.mathworks.com