Energy-Efficient Information Retrieval for Correlated Source Reconstruction in Sensor Networks

Qing Zhao, *Member, IEEE*, and Lang Tong, *Fellow, IEEE*

***Abstract*— We consider information retrieval in a wireless sensor network deployed for the reconstruction of a spatially correlated signal field. Referred to as QUality-of-service specific Information REtrieval (QUIRE), the proposed protocol optimizes the network performance under the metric of information rate per Joule while ensuring a given QoS. Based on the density of sensor deployment and the QoS specified by the maximum distortion for reconstructing the signal field, QUIRE partitions the sensor network into disjoint and equal-sized cells. The cell size is chosen to minimize the number of transmissions required for a given QoS by exploiting the spatial correlation of the signal field. Adopting the cross-layer design methodology that integrates opportunistic carrier sensing and optimal cell activation, QUIRE eliminates redundant transmissions and fully utilizes the channel reception capability in a fading environment.**

***Index Terms*— Wireless Sensor Networks. Information Retrieval. Energy Efficiency. Opportunistic Transmission.**

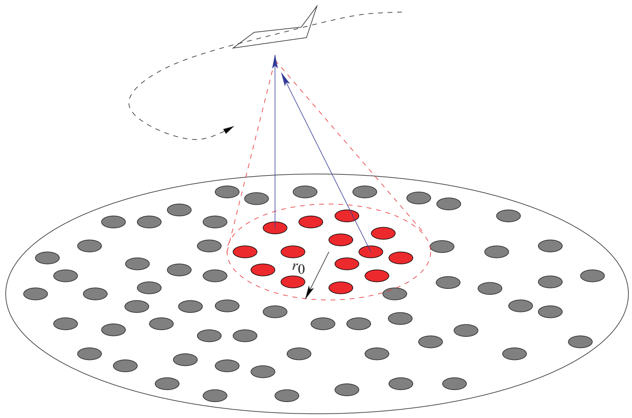

Fig. 1. Sensor network with mobile access point.

I. Introduction

A. Information Retrieval in Sensor Networks

ONE of the major applications envisioned for large-scale wireless sensor networks is the monitoring of certain physical phenomenon which can be modelled as a spatially correlated signal field. In this type of applications, sensors can be preprogrammed to take measurements at specific time instants and transmit their data periodically to the end-user. We consider the scenario where the information collected by sensors is retrieved by a mobile access point (AP). As illustrated in Figure 1, during the information retrieval (IR) operation, the AP broadcasts a beacon to activate sensors in its coverage area. Activated sensors then transmit, according to an IR protocol, their data to the AP through a common wireless channel.

An IR protocol determines which set of sensors data should be collected from and how the wireless channel is shared among activated sensors. The first component of an IR protocol distinguishes its design from that of a medium access control (MAC) protocol in communication networks. It is necessitated by two unique characteristics of sensor networks. First, differing from a communication network whose purpose is to provide service to each individual user, a sensor network is deployed for a single end-user; quality-of-service (QoS) is defined at the network level instead of for each individual node. Second, data from proximate sensors can be highly correlated due to node redundancy and the spacial correlation of the underlying signal field. These two characteristics of sensor networks indicate that it is unnecessary and inefficient to retrieve data from every sensor node.

B. Performance Measure for IR Protocols

A common performance measure for MAC protocols in communication networks is throughput in the unit of packets per unit time. For IR protocols in sensor networks, however, throughput does not necessarily represent protocol efficiency due to the correlation among sensor measurements. Two packets collected from proximate sensors do not provide as much information as two packets from sensors far apart. Merely counting the number of packets can be misleading. Furthermore, throughput does not take into account energy consumption which is of paramount importance in sensor networks.

A good IR protocol should be the one that accomplishes, within the minimum amount of time using the minimum amount of energy, the task the sensor network is deployed for. For the type of applications considered in this paper, the task is accomplished if the information collected by the AP is

Manuscript received December 22, 2004; revised November 30, 2006 and February 22, 2006; accepted May 20, 2006. The associate editor coordinating the review of this paper and approving it for publication was L.-C. Wang. This work was supported in part by Army Research Laboratory CTA on Communication and Networks under Grant DAAD19-01-2-0011 and National Science Foundation under Grants ECS-0622200 and CCR-0311055. Part of this result was presented at MILCOM 2003 and SPSC 2003.

Q. Zhao is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 (e-mail: qzhao@ece.ucdavis.edu).

L. Tong is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 (e-mail: ltong@ece.cornell.edu).

Digital Object Identifier 10.1109/TWC.2007.04885.

sufficient for reconstructing the signal field for a given QoS that can be specified by, for example, the maximum distortion of the reconstruction.

Let I denote the minimum number of packets that need to be collected from sensors in order to reconstruct the signal field for a given QoS. Clearly, this quantity I is a function of the QoS and the statistics of the signal field. It is, however, independent of the IR protocol. Let L and E denote, respectively, the amount of time and energy required for an IR protocol to accomplish the task[1]. We measure, in the unit of packets per unit time per Joule, the performance of the protocol by

$$\eta \stackrel{\Delta}{=} \frac{I}{\mathbb{E}[LE]}, \quad (1)$$

where $\mathbb{E}[\cdot]$ denotes expectation. Besides bringing energy expenditure into the picture, this metric also correctly characterizes the amount of information provided by the collected packets. If a protocol collects more packets than necessary (more than I packets), those redundant packets do not contribute to the performance but rather increase the energy expenditure E and possibly system latency L. We point out that in the design of optimal IR protocols, the quantity I is of little interests. Without loss of generality, we set I to unity and call this performance measure η the IR efficiency.

C. Contributions and Limitations

In this paper, we propose a QUality-of-service specific Information REtrieval (QUIRE) protocol that maximizes the IR efficiency for a given QoS. QUIRE consists of three components. First, based on the density of sensor nodes and the QoS specified by the maximum distortion of the reconstruction, QUIRE partitions the sensor field into disjoint and equal-sized cells. The cell size is chosen as such that the signal field within a cell can be reconstructed, within the maximum distortion, from one sample collected from the center area of the cell. Second, an opportunistic carrier sensing scheme is used to coordinate the transmissions of the sensor nodes located in the center area of a cell. Since the measurement from one sensor located in the center area is sufficient for the reconstruction of the whole cell, only one sensor should transmit to minimize energy consumption and avoid unnecessary interference. With the opportunistic carrier sensing scheme which utilizes the channel state information, QUIRE not only eliminates redundant transmissions, but also ensures that the sensor with the best channel transmits. This opportunistic characteristic of QUIRE further reduces the energy consumption. Third, Based on a Markov representation of the data collection process, QUIRE utilizes an optimal sensor activation scheme to maximize the IR efficiency η.

The proposed protocol is by no means intended to solve the IR problem in general. It targets at certain applications, applies to certain network architectures, and considers one particular approach to eliminating redundancy among sensor measurements. First, QUIRE is developed mainly for the clock-driven and query-driven applications [1] where the IR process can be initiated by the end-user. For event-driven applications such as target detection and tracking, the applicability of the proposed protocol is limited. Second, QUIRE assumes one-hop transmission from the sensor nodes to the AP. It is applicable to sensor networks with mobile access [2], [3] where an AP traverses the network collecting data directly from the sensors. The proposed protocol may also be applied to sensor networks with a cluster structure such as LEACH [4] and its variation [5]. It is, however, not obvious how QUIRE can be used under the multi-hop ad hoc architecture [6]. Finally, there are, in general, three approaches to reducing data correlation hence improving the IR efficiency: distributed source coding, data aggregation, and sampling. The first approach considered in [7], [8] is information theoretic in nature which inevitably encounters large delay and high complexity. The second approach considered in [9], [10] is commonly used under the multi-hop ad hoc architecture where data are aggregated at intermediate sensor nodes along a multi-hop route to a gateway node. It is thus strongly coupled with the design of routing protocols. In this paper, we use sampling to reduce the correlation among collected data, *i.e.,* only data from a fraction of sensor nodes are retrieved. This approach, also considered in [11], [12], is more compatible with the network architecture considered in this paper.

D. Related Work

MAC design for wireless sensor networks has received considerable attention (see [6] for a survey). These protocols, however, do not directly address QoS specified by the applications. There is also a growing body of literature on the energy efficiency of sensor networks [13]–[15]. The primary focus of these results is the analysis of sensor network lifetime under the multi-hop ad hoc architecture.

Perhaps [11] and [16] are the most relevant work to this paper. In [11], the authors explicitly exploit node redundancy. They develop an adaptive scheme for each sensor to determine independently whether to transmit so that a fixed total number of transmissions occur in each slot. The difference between [11] and our work is that QoS in [11] is defined as the total number of transmissions that should occur in each slot and an independent channel from each sensor to a remote base station is assumed. In [16], a protocol named opportunistic ALOHA is proposed for the IR operation in sensor networks with mobile AP. Different from the protocol proposed in this paper, opportunistic ALOHA uses throughput instead of energy efficiency as the performance measure. It does not address the issue of QoS. Being a sensor-initiated random access protocol, it does not eliminate redundant transmissions.

II. The Problem Statement

A. The Signal Field

Let $\mathcal{D}$ denote the sensor field with area A and $S(\mathcal{D})$ the signal field being monitored. We pose the following assumptions on $S(\mathcal{D})$.

- A1 $S(x, y)$ for all $(x, y) \in \mathcal{D}$ have common mean μ and variance ν^2.
- A2 $S(\mathcal{D})$ is spatially homogeneous, *i.e.,* the correlation of two points in $\mathcal{D}$ is determined by the Euclidean distance between them.

[1]Both L and E can be random variables due to the randomness in the fading channel and the sensor locations.

A3 The correlation function $R(d)$ is continuous and monotonically decreasing in $[0, d_{max}]$ where d_{max} is the maximum distance between two points in $\mathcal{D}$.

We assume the sensor nodes form a two-dimensional Poisson field with mean λ nodes/m^2. If a sensor locates at (x, y), it measures the value (one realization) of $S(x, y)$ and generates a packet containing its measurement to be transmitted to the AP. Note that for this type of application, sensors may need to acquire their locations via GPS or other position estimation schemes [17], [18] after deployment.

B. The Wireless Fading Channel

During the IR operation, sensors are activated by and synchronized to the beacon signal broadcasted by the AP. The channel can then be considered slotted with the slot length equal to the transmission time of one packet. We assume that the channel between an activated sensor and the AP is constant within a slot but fades independently and identically across slots and activated sensors. Let $\gamma_i(t)$ denote the channel fading gain from sensor i to the AP in slot t. The distribution F_γ of $\gamma_i(t)$ is then independent of i and t. For Rayleigh fading, F_γ is an exponential function.

Due to the i.i.d. distribution of the channel gain across slots and activated sensors, the reception capability of the AP can be modelled by $C_{n,k}(F_\gamma)$, the probability of having k successes in a slot with n transmissions. The multiaccess channel between the sensors and the AP is then fully characterized by the following reception matrix $\mathbf{C}$ which is a function of the fading distribution F_γ.

$$\mathbf{C}(F_\gamma) = \begin{pmatrix} C_{1,0}(F_\gamma) & C_{1,1}(F_\gamma) & & \\ C_{2,0}(F_\gamma) & C_{2,1}(F_\gamma) & C_{2,2}(F_\gamma) & \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \tag{2}$$

We point out that this general channel model, first considered in [19], includes the single-user collision channel and channels with capture as special cases. It also models networks using spread spectrum (as considered in the simulation examples in Section V) and/or multiple antennas at the AP where successful reception of simultaneously transmitted packets is possible. Define

$$\mathcal{C}_n \stackrel{\Delta}{=} \sum_{k=1}^{n} k C_{n,k}, \qquad n_0 \stackrel{\Delta}{=} \arg\max_n \mathcal{C}_n. \tag{3}$$

That is, when n_0 packets are transmitted simultaneously, the expected number of successfully received packets is maximized. Simultaneously enabling more than n_0 transmissions results in a reduction of successful receptions and an increase in energy expenditure.

C. QoS Requirement

Let $\mathcal{A}$ denote all the points whose measurements are collected during an IR operation. Let $\bar{\mathcal{A}}$ denote the complement of $\mathcal{A}$ in $\mathcal{D}$. Then the signal field $S(\mathcal{D})$ is reconstructed by approximating a point (x_0, y_0) in $\bar{\mathcal{A}}$ with a point in $\mathcal{A}$ that is closest to (x_0, y_0)[2]. Let $d_{(x_0,y_0),(x,y)}$ denote the distance

[2]More sophisticated reconstruction techniques can be used. Here we choose a simple method for the tractability of the analysis.

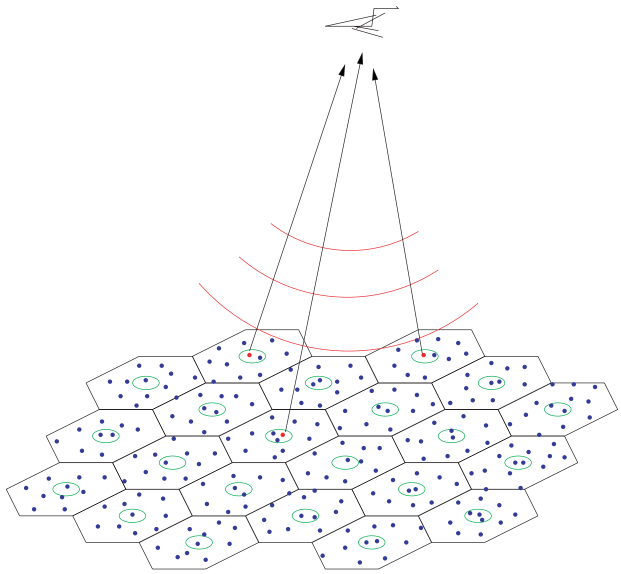

Fig. 2. The cell structure of QUIRE.

between (x_0, y_0) and (x, y). The estimate $\hat{S}(x_0, y_0)$ of the signal field at (x_0, y_0) is given by

$$\begin{aligned} \hat{S}(x_0, y_0) &= S(x_1, y_1), \\ (x_1, y_1) &= \arg\min_{(x,y)\in\mathcal{A}} d_{(x_0,y_0),(x,y)}. \end{aligned} \tag{4}$$

The QoS is characterized by the maximum distortion D in terms of the mean square error (MSE) and the outage probability[3] P_o. In other words, with probability no smaller than $1 - P_o$, every point in $\mathcal{D}$ can be estimated with an MSE no larger than D.

Our problem here is to design an IR protocol that optimizes the IR efficiency η for a given QoS (D, P_o).

III. QoS Specific Information Retrieval: Cell Partition

QUIRE consists of two steps. First, based on the density λ the sensor nodes, the autocorrelation $R(d)$ of the signal field, and the QoS (D, P_o), the AP partitions $\mathcal{D}$ into disjoint and equal-sized cells. In the second step, one sensor from each cell transmits according to a transmission control scheme so that the IR efficiency η is maximized. In this section, we present the cell structure of QUIRE.

As shown in Figure 2, QUIRE partitions the network into equal-sized hexagons with radius r. Within each cell, a center area is defined as the radius-r_0 disk at the center of the hexagon (see Figure 3). The sizes (r, r_0) of the cell and the center area are chosen as such that the whole cell can be reconstructed, with MSE no larger than D, from one sample collected from the center area. Let (x, y) denote the point in the center area whose measurement is collected. Let (u, v) be another point in the same cell. The QoS requires that

$$\begin{aligned} \mathbb{E}[(\hat{S}(u,v) - S(u,v))^2] &= \mathbb{E}[(S(x,y) - S(u,v))^2] \\ &= 2\nu^2 - 2R(d_{(x,y),(u,v)}) \leq D. \end{aligned}$$

[3]Due to the random distribution of sensor nodes, the maximum distortion D can not be bounded deterministically by a fixed value for a given finite node density λ.

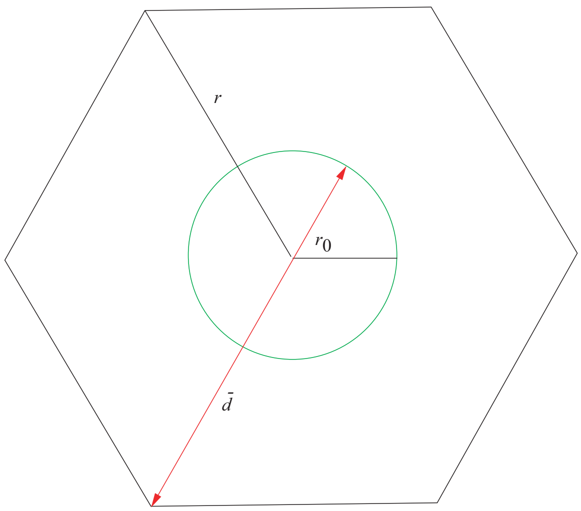

Fig. 3. The size of a cell.

Equivalently, we have $R(d_{(x,y),(u,v)}) \geq (2\nu^2 - D)/2$. It then follows that to reconstruct the cell from one sample collected from the center area, we need

$$\begin{aligned} r + r_0 &\leq \bar{d}, \\ \bar{d} &\triangleq \max(\{d : R(d) \geq (2\nu^2 - D)/2, d \in [0, d_{max}]\}). \end{aligned} \tag{5}$$

With (5) satisfied, the QoS is met if with probability no smaller than $1 - P_o$, at least one sensor is located in the center area of every cell so that one sample can be collected from every cell. This imposes a lower bound on r_0. Specifically, let M denote the total number of resulting cells when the cell size is chosen to be $r = \bar{d} - r_0$. We have, ignoring the boundary effect, $M = \lceil \frac{2A}{3\sqrt{3}(\bar{d}-r_0)^2} \rceil$. To meet (with equality) the requirement on the outage probability, we have,

$$(1 - e^{-\pi r_0^2 \lambda})^M = 1 - P_o \quad \Rightarrow \quad r_0 = \sqrt{-\frac{\log(1 - (1 - P_o)^{\frac{1}{M}})}{\lambda \pi}}, \tag{6}$$

which follows directly from the fact that the numbers of sensors inside disjoint areas are independent for Poisson distribution. We assume that the sensor network is sufficiently dense (λ sufficiently large) so that the above defined r_0 satisfies $r_0 < r < \bar{d}$ for the give QoS. With r_0 chosen, we obtain the cell size r from (5) as $r = \bar{d} - r_0$. Note that we choose the largest r that satisfies (5) to reduce the number of resulted cells hence the number of required transmissions from sensors.

We point out that the cell partition is randomized from one IR operation to the next to ensure uniform energy expenditure among sensors. Furthermore, the cell partition for the upcoming IR operation can be carried out by the access point at the end of the current IR operation. The size r_0 of the center area and the locations of the center points of all cells are broadcasted by the access point so that all sensors outside the center areas do not need to participate even in the sensing task for the next IR operation.

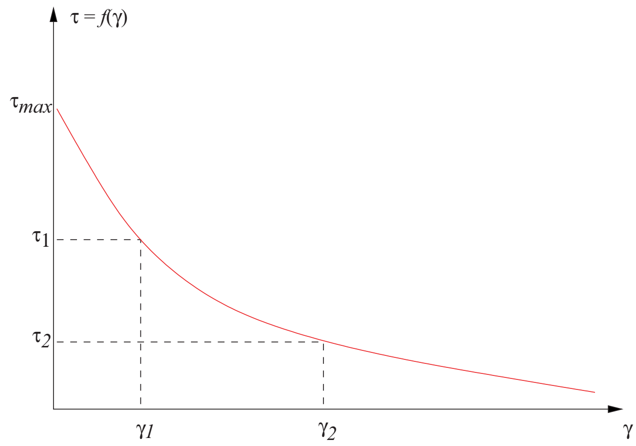

Fig. 4. Opportunistic carrier sensing.

IV. QoS-Specific Information Retrieval: Packet Collection

After the cell partition, the AP activates[4] sensors in the center areas and collects one packet from each cell according to the transmission and activation schemes described below.

A. Distributed In-Cell Transmission Control

When the AP activates the center area of a particular cell, more than one sensor is activated. Since one measurement is sufficient for the reconstruction, only one sensor in the center area should transmit to reduce energy consumption and avoid unnecessary interference. To achieve this, we employ a distributed transmission protocol—opportunistic carrier sensing—proposed in [20], [21]. Specifically, when the AP activates the center area of a cell by broadcasting a beacon, sensors inside the center area measure their channel gain γ using the beacon signal. Based on its own channel gain γ, each sensor chooses a backoff time τ according to a predetermined function $f(\gamma)$ and listens to the channel. A sensor will transmit with its chosen backoff time if and only if no one transmits before it. Considering the relatively small size of the center area, we assume that there are no hidden terminals and the signal propagation delay within the center area of a cell is negligible. Thus carrier sensing ensures that one and only one sensor from the center area transmits. Furthermore, if $f(\gamma)$ is chosen to be a strictly decreasing function as shown in Figure 4, the sensor with the best channel will choose the smallest backoff time thus seizes the channel. This opportunistic carrier sensing not only eliminates redundant transmissions, but also ensures that the sensor with the best channel[5] transmits.

The slot structure resulted from this transmission scheme is illustrated in Figure 5. We point out that to ensure the

[4]If the AP can adjust its coverage area (area within which the beacon signal can be heard) via beamforming, it can activate sensors in desired area by broadcasting a beacon signal. Otherwise, the AP can broadcast the location of the center point of the cell to be activated, and each sensor, by calculating its distance to the center point, can determine whether it should be active for transmission.

[5]Here we assume reciprocity. When the propagation delay within the center area is negligible, $f(\gamma)$ can be any decreasing function with range $[0, \tau_{max}]$ where the maximum backoff time τ_{max} can be any positive number. Under significant propagation delay, $f(\gamma)$ needs to be chosen judiciously to ensure the efficiency of the opportunistic carrier sensing as addressed in [21].

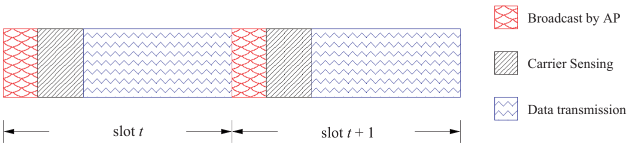

Fig. 5. Slot structure of QUIRE.

efficiency of the opportunistic carrier sensing, simultaneously enabled cells can be spread out (as shown in Figure 2) so that sensors in the center area of one cell will not hear the transmissions from another cell in applications resulting in small cell size.

Opportunistic carrier sensing changes the statistics of the channel gain γ seen by the AP, thus affecting the channel reception capability. Consider, for example, Rayleigh fading where γ is exponentially distributed with mean θ. Let $\check{\gamma}$ denote the channel gain from a transmitting sensor to the AP. We have, from the fact that opportunistic carrier sensing ensures the sensor with the best channel transmits, $\check{\gamma} = \max\{\gamma_1, \gamma_2, \cdots, \gamma_J\}$ where J is the number of sensors inside a nonempty center area[6]. It has the following truncated Poisson distribution

$$P[J=j] = \frac{e^{-\lambda_0}}{1-e^{-\lambda_0}} \frac{(\lambda_0)^j}{j!}, \quad \lambda_0 = \lambda\pi r_0^2, \quad j > 0.$$

We then obtain the distribution of $\check{\gamma}$ as follows.

$$\begin{aligned} F_{\check{\gamma}}(x) &\stackrel{\Delta}{=} P[\check{\gamma} \le x] = \sum_{j=1}^{\infty} P[\check{\gamma} \le x \mid J=j] P[J=j] \\ &= \sum_{j=1}^{\infty} P[J=j] \prod_{i=1}^{j} P[\gamma_i \le x] \\ &= \sum_{j=1}^{\infty} \frac{e^{-\lambda_0}}{1-e^{-\lambda_0}} \frac{(\lambda_0)^j}{j!} (1-e^{-\frac{x}{\theta}})^j \\ &= \frac{1}{1-e^{-\lambda_0}} (e^{-\lambda_0 e^{-\frac{x}{\theta}}} - e^{-\lambda_0}). \end{aligned} \tag{7}$$

Clearly, $\check{\gamma}$ is no longer exponentially distributed. With the opportunistic carrier sensing, the reception matrix $\mathbf{C}$ is determined by $F_{\check{\gamma}}$ instead of F_γ. In Section V, examples are given to demonstrate the construction of $\mathbf{C}(F_{\check{\gamma}})$ and the impact of opportunistic carrier sensing on the channel reception capability.

B. Optimal Cell Activation

We now develop an optimal cell activation scheme to maximize the IR efficiency η. The key idea is to fully exploit the channel reception capability characterized by $\mathbf{C}(F_{\check{\gamma}})$. We assume here $\mathbf{C}(F_{\check{\gamma}})$ has been obtained and known to the AP.

At the beginning of an IR operation, the AP queues up all M cells and activates in each slot N cells starting from the head of the queue. At the end of each slot, the AP detects whether this slot is empty. An empty slot implies that no sensor is located in the center areas of these N cells [7]. These N cells

[6]The channel gain seen by the AP excludes the case of an empty center area where no transmissions occur.

[7]The probability of having one or more empty center areas is bounded below P_o. See (6).

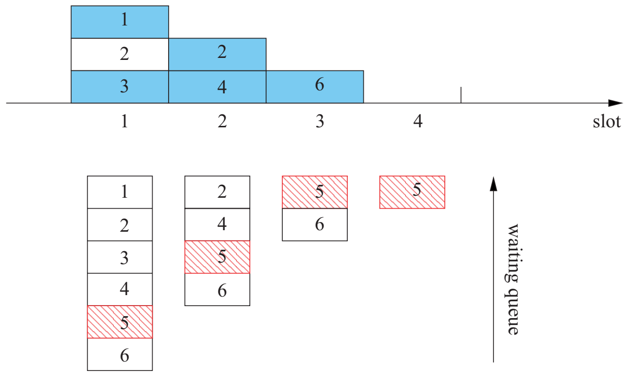

Fig. 6. Optimal cell activation of QUIRE ($M=6$, $N=3$, and the 5th cell (shaded) has an empty center area. The numbered rectangles on the time axis indicate transmissions from the corresponding cells. Successful transmissions are illustrated by shaded rectangles.

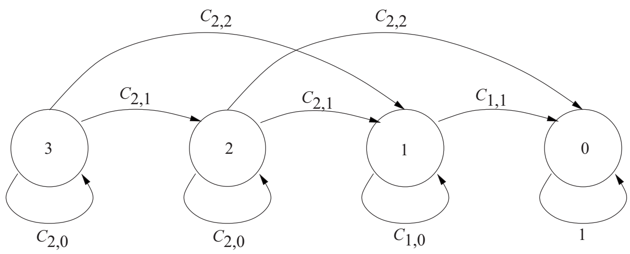

Fig. 7. The Markov chain for Case 1 ($M=3$ and $N=2$).

are thus removed from the queue. On the other hand, if this slot is not empty and k ($k \ge 0$) packets are successfully received, those k cells from which a packet is received are removed from the queue. This procedure continues until the queue becomes empty (all M cells are processed). The basic procedure of this cell activation scheme is illustrated in Figure 6.

With this queue structure, the only parameter to be designed is the number N of simultaneously activated cells. The optimal N^* should be chosen by maximizing the IR efficiency $\eta \stackrel{\Delta}{=} 1/\mathbb{E}[LE]$, or equivalently, minimizing $\mathbb{E}[LE]$,

$$N^* = \arg \min_{N=1,\cdots,N_{max}} \mathbb{E}[LE \mid N], \tag{8}$$

where N_{max} is the maximum number of simultaneously activated cells to be considered. For typical applications, N_{max} can be set to n_0 defined in (3). In order to determine N^*, we calculate $\mathbb{E}[LE \mid N]$ by considering the following two cases.

1) Case 1: Every cell has a nonempty center area: This scenario happens with probability $1-P_o$. In this case, N transmissions occur in all slots except the last few when the number of cells in the queue is smaller than N. In large sensor networks, we typically have $M \gg N \sim n_0$. Thus, $E \approx \omega N L$ where ω is the energy used for transmitting one packet[8]. It then follows that

$$\mathbb{E}[LE \mid N] = \omega N \mathbb{E}[L^2 \mid N]. \tag{9}$$

[8]We assume that the AP is not energy constrained and thus consider only the energy consumed by sensors. Furthermore, we ignore the energy consumed by sensors in listening to the beacon signal and in carrier sensing, which is negligible compared to the energy consumed in transmission. It is, however, straightforward to extend the analysis to include these overheads in energy consumption.

To calculate $\mathbb{E}[L^2 \mid N]$, we represent the IR process by a finite-state discrete Markov chain. It can be shown that the number of unprocessed cells (cells left in the queue) at the beginning of each slot forms a Markov chain. Specifically, at the beginning of a slot, the network is in state j if there are j $(j = 0, \cdots, M)$ unprocessed cells. An example of the state diagram of this Markov chain is illustrated in Figure 7. In general, the transition probability from state j to state k is given by

$$p_{j,k} = \begin{cases} C_{\min(N,j),j-k} & \text{if } 1 \leq j \leq M \\ & \text{and } j - \min(N,j) \leq k \leq j \\ 1 & \text{if } j = k = 0 \\ 0 & \text{otherwise} \end{cases} . \tag{10}$$

With state 0 defined as the absorbing state, $\mathbb{E}[L^2 \mid N]$ is the second moment of the absorbing time of this Markov chain when the initial state is M. Define e_j and v_j as, respectively, the first and the second moments of the remaining time until absorption given that the current state is j. Let $\mathbf{e} \triangleq [e_1, \cdots, e_M]^t$ and $\mathbf{v} \triangleq [v_1, \cdots, v_M]^t$. We then have [22],

$$\mathbf{e} = (\mathbf{I} - \mathbf{P})^{-1}\mathbf{1}, \qquad \mathbf{v} = (\mathbf{I} - \mathbf{P})^{-1}(2\mathbf{e} - \mathbf{1}), \tag{11}$$

where $\mathbf{P}$ is the transition probability matrix (after removing the absorbing state 0) with entries specified by (10), $\mathbf{I}$ and $\mathbf{1}$ denote, respectively, an identity matrix and a vector with all entries equal to 1. After solving v_M from (11), we obtain, from (9),

$$\mathbb{E}[LE \mid N] = \omega N v_M.$$

2) Case 2: At least one cell has an empty center area: This scenario happens with probability P_o. Since the numbers of sensors inside the center areas are i.i.d for different cells, the probability q that a particular cell has an empty center area is given by $q = 1 - (1 - P_o)^{1/M}$. Typically, we have P_o close to 0 and M a large number, which leads to $q \approx 0$. It thus suffices to consider the case where only one cell has an empty center area.

We first consider $N = 1$. In this case, one empty slot occurs during the data collection. In the rest $L - 1$ slots, exactly one transmission takes place, *i.e.,* $E = (L - 1)\omega$. Thus,

$$\mathbb{E}[LE \mid N = 1] = \omega\mathbb{E}[L^2 \mid N = 1] - \omega\mathbb{E}[L \mid N = 1]. \tag{12}$$

Since the number of slots needed for successfully transmitting one packet is geometrically distributed with success probability $C_{1,1}$, we have, considering we have $M - 1$ packets and one empty slot,

$$\mathbb{E}[L \mid N = 1] = \frac{M - C_{1,0}}{C_{1,1}},$$

$$\mathbb{E}[L^2 \mid N = 1] = \frac{(M - C_{1,0})^2 + (M - 1)C_{1,0}}{C_{1,1}^2}.$$

We now calculate $\mathbb{E}[LE \mid N]$ for $N > 1$. Suppose that the mth cell in the waiting queue at the beginning of an IR operation has an empty center area. We partition, by the beginning of the slot in which the mth cell will be activated for the first time, the IR process into two phases with lengths L_1 and L_2, respectively (in the example given in Figure 6, we have $m = 5$, $L_1 = 1$, and $L_2 = 3$). In Phase 1, N transmissions occur in each slot. For Phase 2, since the AP can only recognize an empty center area through an empty slot for a general channel reception matrix $\mathbf{C}$, the mth cell will be activated in each slot in this segment until, in the end, an empty slot occurs. Hence, $N - 1$ transmissions take place in each slot of this phase[9]. This leads to

$$\begin{aligned} &\mathbb{E}_m[LE \mid N] = \mathbb{E}_m[\omega(L_1 + L_2)(NL_1 + (N-1)L_2) \mid N] \\ &= \omega(N\mathbb{E}_m[L_1^2 \mid N] + (N-1)E_m[L_2^2 \mid N] \\ &\quad + (2N-1)\mathbb{E}_m[L_1 L_2 \mid N]), \end{aligned} \tag{13}$$

where $\mathbb{E}_m[\cdot]$ denotes the expectation conditioned on the mth cell having an empty center area. Similar to Case 1, a discrete-time Markov chain is constructed for each phase to compute $\mathbb{E}_m[L_1^2 \mid N]$, $\mathbb{E}_m[L_2^2 \mid N]$, and $\mathbb{E}_m[L_1 L_2 \mid N]$ as detailed in Appendix. We can then obtain $\mathbb{E}_m[LE \mid N]$ from (13). Averaging over m which is uniformly distributed over $\{1, \cdots, M\}$, we reach the solution for Case 2. The final result is obtained by combining Case 1 and Case 2. With $\mathbb{E}[LE \mid N]$ calculated for $N = 1, \cdots, N_{max}$, the optimal N^* can be obtained from (8).

We point out that the calculation of N^* can be carried out off-line. Little computation is required at the AP during the IR operation. For applications with high QoS requirement (P_o approaches 0), N can be chosen by considering only Case 1.

TABLE I
PARAMETERS OF CELL PARTITION FOR QUIRE

λ	0.2	0.4	0.6	0.8	1
r_0	3.61	2.51	2.03	1.748	1.558
r	6.39	7.49	7.97	8.252	8.442
M	378	275	243	227	217

V. SIMULATION EXAMPLES

A. Simulation Setup

The signal field is a 200m × 200m square. The correlation function $R(d)$ and the maximum distortion D are such that $\bar{d}$ as defined in (5) is 10m. The outage probability is $P_o = 0.1$. In Table 1, we give the size r_0 of the center area and the resulting total number M of cells for various node density λ. Clearly, a larger λ leads to a smaller M.

We consider Rayleigh fading where the channel gain γ is exponentially distributed with mean 1. The sensor transmission power P_t (normalized by the pass loss) and the noise variance σ^2 are such that the average SNR at the AP is given by $\frac{P_t}{\sigma^2} = 5dB$. We consider a DS-CDMA system where each sensor uses a random spreading sequence with spreading gain $G = 16$. The AP uses the linear MMSE receiver. We assume that a packet is successfully demodulated if the signal to interference ratio (SINR) at the linear receiver output is greater than a threshold $\beta = 3dB$, which is a function of the modulation, the error control code, and the target BER. According to Tse and Hanly [23], the SIR at the receiver can be approximated by a simple function of the received power for random spreading codes, large network size, and high

[9]Toward the end of data collection when the number of unprocessed cells is smaller than N, fewer than $N - 1$ transmissions occur. However, $N - 1$ is a good approximation for large M and specifies the worst scenario.

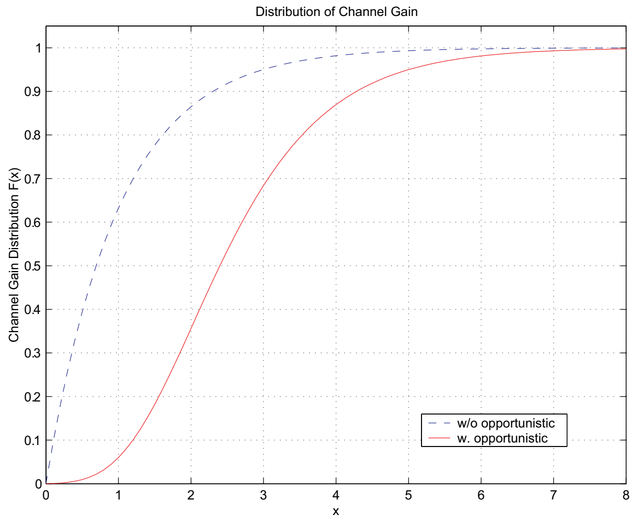

Fig. 8. Improved channel reception capability by opportunistic carrier sensing: channel gain distribution.

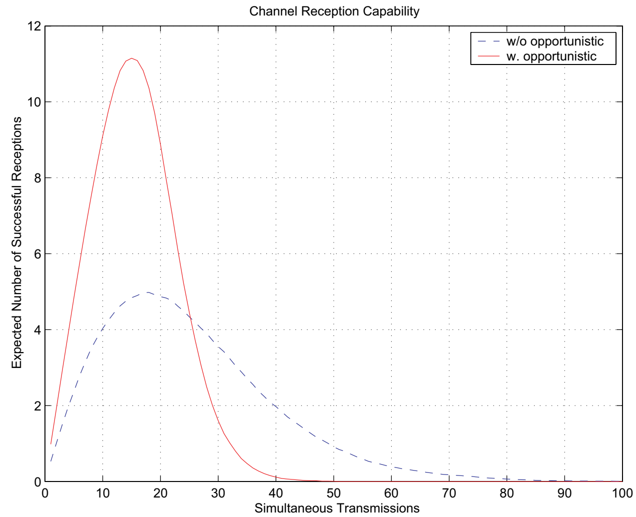

Fig. 9. Improved channel reception capability by opportunistic carrier sensing: expected number of successful receptions.

spreading gain. Specifically, given that K sensors transmit simultaneously, the transmission from sensor i is successful if

$$\frac{\gamma_i}{\frac{\sigma^2}{P_t} + \frac{1}{G}\sum_{k=1,k\neq i}^{K} \frac{\gamma_i \gamma_k}{\beta\gamma_k + \gamma_i}} > \beta. \tag{14}$$

This condition, together with the distribution of the channel gain, completely specifies the channel reception capability. For QUIRE, the channel gain of a transmitting sensor is distributed according to $F_{\check{\gamma}}$ given in (7). By generating realizations of $\check{\gamma}$ according to (7) and applying the receiver model given in (14), we obtain the channel reception matrix $\mathbf{C}(F_{\check{\gamma}})$ numerically.

B. Improved Channel Reception Capability by Opportunistic Carrier Sensing

We first study the impact of opportunistic carrier sensing on the channel reception capability. The node density λ is set to 1.

In Figure 8 we compare the distribution of γ and $\check{\gamma}$. The channel gain γ of the original fading channel is exponentially distributed with mean 1. With opportunistic carrier sensing, the channel gain $\check{\gamma}$ is distributed according to (7). Figure 8 shows clearly that opportunistic carrier sensing changes the channel fading characteristics seen by the AP; $\check{\gamma}$ is more likely to take a large value.

In Figure 9 we compare the channel reception capability $\mathcal{C}_n$ as defined in (3). For the original fading channel, the maximum number of successful receptions is 4.84 which is achieved by simultaneously transmitting 17 packets. With opportunistic carrier sensing, the maximum number of successful receptions is 11.15 achieved by 15 simultaneous transmissions. The channel reception capability is more than doubled with fewer transmissions hence less energy expenditure. Note that when the total number of transmissions is greater than 25, opportunistic carrier sensing reduces the number of successfully received packets. The reason for this is that opportunistic carrier sensing increases the channel gain of each transmission, hence the interference power when the number of simultaneous transmissions is large. In QUIRE, however, by judiciously choosing the number of simultaneously activated cells, this improved channel reception capability can be optimally utilized.

C. Comparison of IR Efficiency η

In Figure 10, we compare the performance of QUIRE, modified TDMA, and ALOHA with optimal transmission probability. The modified TDMA is built upon the cell structure and the opportunistic carrier sensing scheme of QUIRE. Specifically, cells are enabled one by one in each slot and only the sensor with the best channel in the center area transmits. For ALOHA, it is implemented as follows. At the beginning of an IR operation, the AP, based on the total number of sensors in the field, chooses the optimal transmission probability p_t by maximizing the expected number of successful receptions. It than broadcast p_t and all sensors flip a coin with bias p_t to determine whether to transmit in this slot. At the end of this slot, the AP broadcasts the locations of all sensors whose packets are successfully received in this slot. All sensors within $\bar{d}$ distance of these successful sensors will go to sleep; they will not transmit in the future slots of this IR operation. At the beginning of the second slot, the AP, assuming the knowledge of the total number of active sensors (all sensors except those within $\bar{d}$ distance of successful sensors), chooses and broadcasts the optimal transmission probability for this slot. This procedure continues until the whole network is covered (the number of active sensors becomes 0). Note that this implementation of ALOHA does not guarantee the QoS.

From Figure 10 we see that QUIRE achieves 10-fold improvement in IR efficiency at $\lambda = 1$ as compared to modified TDMA and slotted ALOHA with optimal transmission probability. The improvement over modified TDMA, which is built upon the cell structure and the opportunistic carrier sensing scheme of QUIRE, is due to QUIRE's optimal cell activation scheme that fully exploits the channel reception capability. The improvement over ALOHA results from a combination of QUIRE's capability of eliminating redundant

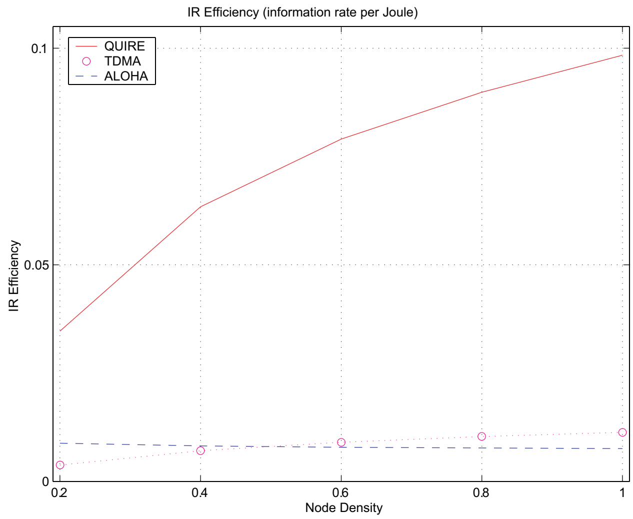

Fig. 10. IR efficiency comparison($\omega = 1.8mJ$).

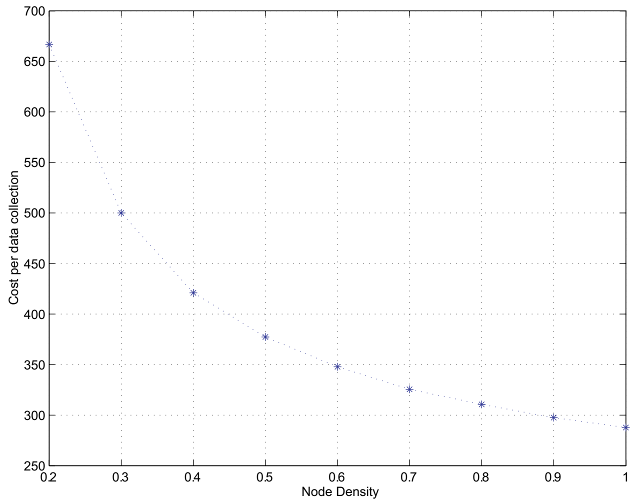

Fig. 11. Network deployment cost per IR operation ($\omega = 1.8mJ$).

transmissions, improving channel reception capability via opportunistic carrier sensing, and optimally activating sensors.

D. Impact of Network Density on Cost Per IR Operation

Since the total number M of cells is a monotonically decreasing function of λ (see Table 1), QUIRE achieves higher IR efficiency in densely deployed networks as shown in Figure 10. This implies that increasing the network density can reduce the cost of each data collection. It is, however, more costly to deploy a dense network. The question we seek to answer with this example is whether the extra cost[10] in deploying a dense network eventually pays off.

We use the same setup as specified in Section V-A and Table 1. Assume that the deployment of one sensor has a unit cost and a sensor becomes nonfunctional (out of power) after transmitting one packet to the AP. The network lifetime ends when the density drops below the level for the existence of r_0, *i.e.,* the network can no longer accomplish the task for the given QoS. The average cost per IR operation is then given by $\frac{\lambda A}{K_d}$ where λA is the expected cost of deploying a sensor network with density λ over an area of Am^2 and K_d is the expected number of IR operations accomplished by the network until it becomes nonfunctional.

Shown in Figure 11 is the cost per IR operation for QUIRE. This result shows that for sensor networks employing QUIRE, it is more efficient to deploy a dense network. Dense network has a longer lifetime and lower cost per network use. This can be used as a rule of thumb for network deployment.

VI. Conclusion

In this paper, we consider IR protocol design for large-scale sensor networks. We explicitly incorporate QoS specified by the network application into the IR protocol design. A new performance measure — IR efficiency η in the metric of information rate per Joule — is introduced that effectively characterizes the amount of information provided by the collected packets and takes energy consumption into consideration. Under this performance measure, we propose QUIRE, an IR protocol that assures a given QoS while maximizing the IR efficiency. It eliminates redundant transmissions and fully exploits the channel reception capability via opportunistic carrier sensing and optimal cell activation.

[10]Here the cost may include the cost of sensors and the cost of deployment.

Appendix: Markov Chain Representation of Case 2

As in Case 1, construct a Markov chain whose state is the number of unprocessed cells and state transitions occur at the beginning of each slot. Since the mth cell will be activated if the number of unprocessed cells belongs to $\{M-m+1,\cdots,M-m+N\}$, we can obtain L_1 as the absorbing time of a Markov chain with state space $\{M-m+1, M-m+2,\cdots,M\}$ and absorbing states $\{M-m+1,\cdots,M-m+N\}$. The transition probability is given by

$$p_{j,k} = \begin{cases} C_{N,j-k} & \text{if } M-m+N+1 \le j \le M, j-N \le k \le j \\ 1 & \text{if } M-m+1 \le j \le M-m+N,\ k=j \\ 0 & \text{otherwise} \end{cases}.$$

The transition probability matrix can be written in the form of $\begin{pmatrix} \mathbf{O} & \mathbf{0} \\ \mathbf{Q} & \mathbf{P} \end{pmatrix}$ where $\mathbf{O}$ and $\mathbf{P}$ correspond to, respectively, the absorbing states $\{M-m+1,\cdots,M-m+N\}$ and the transient states $\{M-m+N+1,\cdots,M\}$. Similar to Case 1, define e_j and v_j as, respectively, the 1st and 2nd moments of the remaining time until absorption given that the current state is j (a transient state). Let u_i be the probability that the absorption occurs at the absorbing state i. We have [22]

$$\begin{aligned} \mathbf{e} &\triangleq [e_{M-m+N+1},\cdots,e_M]^t = (\mathbf{I}-\mathbf{P})^{-1}\mathbf{1}, \\ \mathbf{v} &\triangleq [v_{M-m+N+1},\cdots,v_M]^t = (\mathbf{I}-\mathbf{P})^{-1}(2\mathbf{e}-\mathbf{1}), \\ \mathbf{u} &\triangleq [u_{M-m+1},\cdots,u_{M-m+N}]^t = (\mathbf{I}-\mathbf{P})^{-1}\mathbf{Q}, \end{aligned} \quad (15)$$

from which the 1st and 2nd moment of L_1 can be obtained as e_M and v_M, respectively. As detailed below, $\mathbf{u}$ will serve as the initial condition of the Markov chain constructed for Phase 2.

For Phase 2, the Markov chain has a state space $\{0,1,\cdots,M-m+N\}$ with 0 as the absorbing state. The transition probability is given by

$$p_{j,k} = \begin{cases} C_{\min(N-1,j-1),j-k} & \text{if } 2 \le j \le M-m+N, \\ & \text{and } j-\min(N-1,j-1) \le k \le j \\ 1 & \text{if } 0 \le j \le 1, k=0 \\ 0 & \text{otherwise} \end{cases}.$$

With the similar definition of e_j and v_j ($1 \leq j \leq M-m+N$), we have

$$\begin{aligned} \mathbf{e} &\triangleq [e_1, \cdots, e_{M-m+N}]^t = (\mathbf{I}-\mathbf{P})^{-1}\mathbf{1}, \\ \mathbf{v} &\triangleq [v_1, \cdots, v_{M-m+N}]^t = (\mathbf{I}-\mathbf{P})^{-1}(2\mathbf{e}-\mathbf{1}), \end{aligned}$$

where $\mathbf{P}$ is the transition probability matrix of this Markov chain after removing the absorbing state 0. The 1st and 2nd moments of L_2 can then be obtained as

$$\begin{aligned} \mathbb{E}_m[L_2 \mid N] &= [e_{M-m+1}, \cdots, e_{M-m+N}]\mathbf{u}, \\ \mathbb{E}_m[L_2^2 \mid N] &= [v_{M-m+1}, \cdots, v_{M-m+N}]\mathbf{u}, \end{aligned} \tag{16}$$

where $\mathbf{u}$ is the initial condition of this Markov chain given in (15).

We now consider the last term $\mathbb{E}_m[L_1L_2 \mid N]$ in (13). Conditioned on that the absorbtion in Phase 1 occurs at state $M-m+i$ where $i = 1, \cdots, N$, L_1 and L_2 are independent. We thus have

$$\begin{aligned} \mathbb{E}_m[L_1L_2 \mid N] &= \sum_{i=1}^{N} \mathbb{E}_m[L_1L_2 \mid N, i]u_{M-m+i} \\ &= \sum_{i=1}^{N} \mathbb{E}_m[L_1 \mid N, i]\mathbb{E}_m[L_2 \mid N, i]u_{M-m+i} \\ &= \sum_{i=1}^{N} \mathbb{E}_m[L_1 \mid N]\mathbb{E}_m[L_2 \mid N, i]u_{M-m+i} \quad (17) \\ &= \mathbb{E}_m[L_1 \mid N]\mathbb{E}_m[L_2 \mid N], \end{aligned}$$

where $\mathbb{E}_m[L_1 \mid N]$ and $\mathbb{E}_m[L_2 \mid N]$ are given in (15,16). In (17) we have used the independence between L_1 and the state $M-m+i$ that absorbs the chain in Phase 1. This independence follows directly from the definition of L_1—the number of slots until the chain of Phase 1 hits any of the N absorbing states $\{M-m+1, \cdots, M-m+N\}$.

REFERENCES

[1] E. Duarte-Melo and M. Liu, "Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks," in *Proc. IEEE Globecom*, (Taipei, Taiwan), Nov. 2002.

[2] L. Tong, Q. Zhao, and S. Adireddy, "Sensor Networks with Mobile Agents," in *Proc. 2003 Int'l. Symp. Military Communications*, (Boston, MA), Oct. 2003.

[3] G.Mergen, Q. Zhao, and L.Tong, "Sensor networks with mobile access: Energy and capacity considerations," to appear *IEEE Trans. Commun.*, 2007.

[4] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocols for Wireless Microsensor Networks," in *Proc. Hawaaian Int'l. Conf. Systems Science*, (Hawaai,US), Jan. 2000.

[5] S. Lindsey, C. Raghavendra, and K. Sivalingam, "Data Gathering in Sensor Networks Using the Energy*Delay Metric," in *Proc. Int'l. Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobil Computing*, (CA), Apr. 2001.

[6] J. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," *IEEE Communications Magzine*, vol. 40, pp. 102–114, Aug. 2002.

[7] S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed Compression in a Dense Microsensor Networks," *IEEE Signal Processing Mag.*, pp. 51–60, Mar. 2002.

[8] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Mag.*, vol. 21, pp. 80–94, Sept. 2004.

[9] C. Intanagonwiwat, R. Govindan, and D. Esterin, "Directed Diffusion : A Scalable and Robust Paradigm for Sensor Networks," in *Proc. 6th ACM/Mobicom Conference*, (Boston,MA), pp. pp 56–67, 2000.

[10] C. Intanagonwiwat, D. Esterin, R. Govindan, and J. Heidemann, "Impact of network density on data aggregation in wireless sensor networks," in *Proc. Int'l. Conference on Distributed Computing Systems (ICDCS)*, July 2002.

[11] R. Iyer and L. Kleinrock, "QoS Control for Sensor Networks," in *Proc. of 2003 ICC*, May 2003.

[12] R. Willett, A. Martin, and R. Nowak, "Backcasting: adaptive sampling for sensor networks," in *Proc. IPSN*, (Berkeley, CA), 2004.

[13] M. Bhardwaj, T. Garnett, and A. Chandrakasan, "Upper bounds on the lifetime of sensor networks," in *Proc. 2001 IEEE Int'l. Conference on Communications*, pp. 785–790, 2001.

[14] M. Bhardwaj and A. Chandrakasan, "Bounding the lifetime of sensor networks via optimal role assignments," in *Proc. INFOCOM 2002*, (New York), pp. 1587–1596, June 2002.

[15] H. Zhang and J. Hou, "On deriving the upper bound of lifetime for large sensor networks," in *Proc. MobiHoc*, 2004.

[16] P. Venkitasubramaniam, S. Adireddy, and L. Tong, "Sensor Networks with Mobile Agents: Optimal Random Access and Coding," *IEEE Journal Select. Areas Commun.: Special Issue on Sensor Networks*, Sep. 2004.

[17] L. Doherty, K. Pister, and L. Ghaoui, "Convex Position Estimation in Wireless Sensor Networks," in *Proc. 2001 IEEE INFOCOM*, 2001.

[18] A. Savvides, C. Han, and M. Strivastava, "Dynamic Fine-Grained Localization in Ad-Hoc Networks of Sensors," in *Proc. 7th Annual Int'l. Conference on Mobile Computing and Networking*, (Rome, Italy), pp. 16–21, July 2001.

[19] S. Ghez, S. Verdú, and S. Schwartz, "Stability Properties of slotted Aloha with multipacket reception capability," *IEEE Trans. Automatic Control*, vol. 33, pp. 640–649, July 1988.

[20] Q. Zhao and L. Tong, "Quality-of-Service Specific Information Retrieval for Densely Deployed Sensor Network," in *Proc. IEEE Military Communications Intl Symp.*, (Boston, MA), Oct. 2003.

[21] Q. Zhao and L. Tong, "Opportunistic Carrier Sensing for Energy Efficient Information Retrieval in Sensor Networks," *EURASIP Journal on Wireless Commun. and Networking*, no. 2, pp. 231–241, 2005.

[22] D. Isaacson, *Markov Chains, Theory and Applications*. New York: Wiley, 1976.

[23] D. Tse and S. Hanly, "Linear Multiuser Receivers : Effective Interference, Effective Bandwidth and User Capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 641–657, March 1999.

Qing Zhao received the Ph.D. degree in Electrical Engineering in 2001 from Cornell University, Ithaca, NY. From 2001 to 2003, she was a communication system engineer with Aware, Inc., Bedford, MA. She returned to academe in 2003 as a postdoctoral research associate with the School of Electrical and Computer Engineering at Cornell University. In 2004, she joined the Department of Electrical and Computer Engineering at UC Davis where she is currently an assistant professor. Her research interests are in the general area of signal processing, communication systems, wireless networking, and information theory.

Qing Zhao received the 2000 IEEE Signal Processing Society Young Author Best Paper Award. She is an associate editor of the IEEE Transactions on Signal Processing and an elected member of the Signal Processing for Communications technical committee of the IEEE Signal Processing Society.

Lang Tong joined Cornell University in 1998 where he is the Irwin and Joan Jacobs Professor in Engineering. Prior to joining Cornell University, he was on faculty at the West Virginia University and the University of Connecticut. He was also the 2001 Cor Wit Visiting Professor at the Delft University of Technology. He received the B.E. degree from Tsinghua University, Beijing, P.R. China in 1985, and PhD degree in EE from the University of Notre Dame, Notre Dame, Indiana in 1991. He was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University in 1991.

Lang Tong is a Fellow of IEEE. He received the Outstanding Young Author Award from the IEEE Circuits and Systems Society, the 2004 best paper award (with Min Dong) from IEEE Signal Processing Society, and the 2004 Leonard G. Abraham Prize Paper Award from the IEEE Communications Society (with Parvathinathan Venkitasubramaniam and Srihari Adireddy). He is also a coauthor of five student paper awards. He received Young Investigator Award from the Office of Naval Research.

Lang Tong's research is in the general area of statistical signal processing, communication systems, and networks. Using theories and tools from statistical inferences, stochastic processes, and information theory, his research group is investigating various practical problems that arise from broadband wireless systems, mobile ad hoc networks, and large scale sensor networks.