# Extending Network Lifetime for Precision-Constrained Data Aggregation in Wireless Sensor Networks

Xueyan Tang
School of Computer Engineering
Nanyang Technological University
Singapore 639798
asxytang@ntu.edu.sg

Jianliang Xu
Department of Computer Science
Hong Kong Baptist University
Kowloon Tong, Hong Kong
xujl@comp.hkbu.edu.hk

***Abstract*— This paper exploits the tradeoff between data quality and energy consumption to extend the lifetime of wireless sensor networks. We consider the applications that require some aggregate form of sensed data with precision guarantees. Our key idea is to differentiate the precisions of data collected from different sensor nodes to balance their energy consumption. This is achieved by partitioning the precision constraint of data aggregation and allocating error bounds to individual sensor nodes in a coordinated fashion. Three factors affecting the lifetime of sensor nodes are identified: (1) the changing pattern of sensor readings; (2) the residual energy of sensor nodes; and (3) the communication cost between the sensor nodes and the base station. We analyze the optimal precision allocation in terms of network lifetime and propose an adaptive precision allocation scheme that dynamically adjusts the error bounds of sensor nodes. Experimental results using real data traces show that the proposed scheme significantly improves network lifetime compared to existing methods.**

## I. Introduction

Rapid advances in sensing and wireless communication technologies have made feasible the deployment of wireless sensor networks for a wide range of applications such as ecosystem monitoring and traffic surveillance [1], [2]. A wireless sensor network typically consists of a base station and a group of sensor nodes (see Figure 1). The sensor nodes are responsible for continuously capturing environmental data such as temperature and wind. They are also capable of communicating with each other and the base station through radios. The base station, on the other hand, serves as a gateway for the sensor network to exchange data with external applications to accomplish certain missions. It collects the sensor readings and converts them into a form requested by the applications (e.g., average temperature reading). This conversion process is called *aggregation*. Primarily designed for monitoring purposes, many sensor applications require continuous aggregation of sensed data [3].

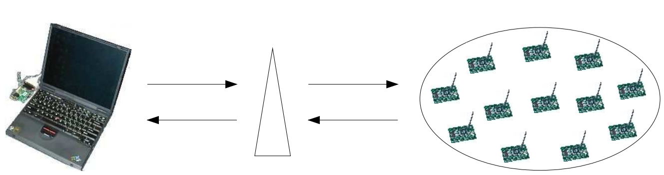



Fig. 1. System Architecture

While the base station can have continuous power supply, the sensor nodes are usually battery-powered. They are inconvenient to replace once deployed in the field. Sometimes, replacement is even impossible (e.g., sensor nodes in a hard-to-reach area). Thus, energy efficiency is a critical design consideration of wireless sensor networks. In these networks, communication is a dominant source of energy consumption [4]. Continuous exact data aggregation requires substantial energy consumption because each sensor node has to report every reading to the base station. Unlike the strict data semantics emphasized in traditional databases, to save energy, many sensor applications allow *approximate data aggregation* with precision guarantees [5], [6], [7], [8]. The precision can, for example, be specified in the form of quantitative error bounds: "average temperature reading of all sensor nodes within an error bound of 1°C." In approximate data aggregation, the sensor nodes do not have to report all readings to the base station. Only the updates necessary to guarantee the desired level of precision need to be sent.

When a sensor node runs out of energy, its coverage is lost. The mission of a sensor application would not be able to continue if the coverage loss is remarkable. Therefore, the practical value of a sensor network is determined by the time duration before it fails to carry out the mission due to insufficient number of "alive" sensor nodes. This duration is referred to as the *network lifetime* [1]. It is both mission-critical and economically desirable to manage sensed data in an energy-efficient way to extend the lifetime of sensor networks. However, this is a challenging task in that the sensor nodes are inherently heterogeneous in energy consumption. First, the data captured by different sensor nodes may change at different rates. This implies the sensor nodes need to report data at different rates. Second, the wireless communication cost heavily depends on the transmission distance [9]. Due to the geographically distributed nature of sensor networks, the sensor nodes are likely to differ significantly in the energy cost of sending a message to the base station. Even if all sensor nodes report data at the same rate, their energy consumption can be highly unbalanced, thereby reducing network lifetime.

In this paper, we investigate the optimization of network lifetime for approximate data aggregation. Our key idea is to differentiate the quality of data collected from different sensor nodes to balance their energy consumption. This is achieved by partitioning the precision constraint of data aggregation and allocating error bounds to individual sensor nodes in a coordinated fashion. Our contributions are as follows:

- We identify three factors affecting the lifetime of sensor nodes in the context of approximate data aggregation: (1) the changing pattern of sensor readings; (2) the residual energy of sensor nodes; and (3) the communication cost between the sensor nodes and the base station.
- We analyze the optimal precision allocation in terms of network lifetime. To the best of our knowledge, this is the first study on optimizing precision allocation to balance energy consumption in wireless sensor networks.
- We develop a sample-based precision allocation method and prove its optimality. Based on this, an adaptive precision allocation scheme is proposed to dynamically adjust the error bounds of sensor nodes. The scheme models in-network aggregation in multi-hop networks.
- We present an experimental evaluation using real data traces. Experimental results show that the proposed scheme significantly improves network lifetime compared to existing methods.

The rest of this paper is organized as follows. Section II summarizes the related work. Section III describes the system model and gives some basic definitions. Section IV analyzes the optimal precision allocation in single-hop networks and then proposes an adaptive precision allocation scheme. Section V extends the adaptive scheme to multi-hop networks. The experimental setup and results are discussed in Section VI. Finally, Section VII concludes the paper.

## II. Related Work

Wireless sensor networks have attracted much research effort in recent years. In the networking community, researchers have focused on optimizing network related operations such as routing and media access [10], [11], but they usually did not take application-level data semantics into consideration. In the database community, researchers have focused on exact query processing over sensed data [12], [13], [14] and paid little attention to trading data quality for energy efficiency.

Han *et al.* [7] investigated the management of sensor states for approximate query processing. In an earlier work, we developed a two-tier storage scheme for approximate query processing in object tracking sensor networks [15]. However, these studies were limited to queries for the readings of individual sensor nodes, in which case the precision can be set independently for different sensor nodes. Different from [7], [15], in this paper, we consider continuous data aggregation where the precision settings of different sensor nodes are inter-related. Olston *et al.* [5] investigated burden-based precision adjustment for continuous queries over distributed data streams. They aimed at minimizing the total communication cost between data sources and the data sink. Sharaf *et al.* [6] implemented a simple uniform precision allocation for in-network data aggregation. Deligiannakis *et al.* [8] further optimized the allocation to reduce the number of messages in the network. However, none of these studies has taken energy and lifetime models into consideration. Thus, their proposed techniques are not effective in handling the energy constraints in wireless sensor networks. Moreover, some algorithm design was based on a given changing pattern of sensor readings (e.g., one-dimensional random walk [5]). Hence, the algorithms may not work well for other changing patterns. Different from existing work, in this paper, we develop a generic framework for extending network lifetime. Our proposed techniques are applicable to any changing pattern of sensor readings.

Besides precision setting, some researchers investigated approximate representation of sensor readings with sophisticated data structures [16], [17]. These studies are complementary to our work.

## III. System Model

We consider a network of $n$ sensor nodes geographically distributed in an operational area. They monitor immediate surroundings and periodically collect local measurements such as temperature and wind. The rate at which the sensor nodes capture data readings is called the sensing rate, denoted by $S$. The period between two successive readings is called an *epoch*. The base station gathers data from the sensor nodes.

### A. Data Aggregation Model

We consider approximate data aggregation with precision guarantees [6], [7], [8]. Data impreciseness is measured by the quantitative difference between an approximate value and the exact value. The sensor application specifies the precision constraint of data aggregation by an upperbound $E$ on data impreciseness (called the *error bound*). That is, on receiving an aggregate data value $A'$ from the sensor network, the application would like to be assured that the exact aggregate value $A$ lies in the interval $[A'-E, A'+E]$.

In approximate data aggregation, not all sensor readings have to be sent to the base station in every epoch. To reduce communication cost, the designated error bound on aggregate data can be partitioned and allocated to individual sensor nodes (we shall call it *precision allocation*). Each sensor node updates a new reading with the base station only when the new reading significantly deviates from the last update to the base station and violates the allocated error bound. This is called a *precision-driven update*. The designated precision of aggregate data must be guaranteed provided that each sensor node updates the readings with the base station according to its allocated error bound. Therefore, the error bounds allocated to individual sensor nodes have to satisfy certain *feasibility constraints*. Different aggregation functions impose different constraints. In this paper, we consider three commonly used types of aggregations: SUM, COUNT and AVERAGE. Exploring other aggregation functions such as MIN/MAX is an important topic of future work. For SUM and COUNT aggregations, to guarantee a given error bound $E$ on aggregate data, the total error bound allocated to the sensor nodes cannot exceed $E$, i.e.,

$$\sum_{i=1}^{n} e_i \leq E, \tag{1}$$

where $e_i$ is the error bound allocated to sensor node $i$. For AVERAGE aggregation, the total error bound allocated to the sensor nodes cannot exceed $n \cdot E$, i.e.,

$$\sum_{i=1}^{n} e_i \leq n \cdot E, \tag{2}$$

where $n$ is the number of sensor nodes.

Eligible precision allocation under the feasibility constraint is not unique. For example, in a network of 10 temperature sensor nodes, if the error bound on AVERAGE aggregation is 1°C, we can allocate an error bound of 1°C to each sensor node. Alternatively, we can also allocate an error bound of 5.5°C to a selected node and an error bound 0.5°C to each of the remaining nodes. This offers the flexibility to adjust the energy consumption of individual sensor nodes by careful precision allocation. In general, to collect the readings of a sensor node at higher precision (i.e., smaller error bound), the sensor node needs to send precision-driven updates to the base station more frequently, which translates to higher energy consumption. The quantitative relationship depends on the *changing pattern* of sensor readings (e.g., the frequency and magnitude). Without loss of generality, we shall denote the precision-driven update rate of each sensor node $i$ as a function $u_i(e)$ of the allocated error bound $e \geq 0$. $u_i(e)$ is essentially the rate at which the reading captured by sensor node $i$ changes beyond $e$. Intuitively, $u_i(e)$ is a *non-increasing* function with respect to $e$. It is obvious that $u_i(e) \leq S$ and $u_i(\infty) = 0$.

Constraints (1) and (2) share the common characteristic that the total error bound of the sensor nodes is capped by a given value. We shall focus on constraint (1) in our discussion. The analysis and algorithms developed in this paper can be adapted to handle constraint (2) in a straightforward manner. They are also directly applicable to SUM and AVERAGE aggregations over any fixed subset of the sensor nodes.

### B. Energy Model

The base station normally has continuous and sufficient power supply. Thus, we shall assume no energy constraint at the base station. The sensor nodes, on the other hand, are powered by batteries that are inconvenient to replace. Communication has been shown to be the dominant source of energy consumption in wireless sensor networks [1], [4] and is therefore the focus of this paper.

We denote the energy consumed by sensor node $i$ to transmit and receive a data update by $s_i$ and $v_i$ respectively. They can take different forms to cater for a wide range of factors. In the simplest case, if all sensor nodes use a default communication range, $s_i$'s are the same for all nodes. More sophisticatedly, if the sensor nodes know the locations of the receivers [18], [19], they can adapt the power level to the transmission distance. The sensor nodes with longer transmission distances would be associated with higher $s_i$'s. In addition, reliability can also be modeled in the energy cost. Less reliable links are entitled to higher $s_i$'s and $v_i$'s due to possible retransmissions. The exact forms of $s_i$ and $v_i$ are orthogonal to our analysis and beyond the scope of this paper.

Similar to other studies [20], [21], [22], [23], we define the network lifetime as the time duration before the first sensor node runs out of energy. Our analysis is also applicable to redundant sensor deployment where each target location is covered by several sensor nodes. From the viewpoint of network lifetime, the set of sensor nodes monitoring the same location can be converted to a single equivalent node by adding up the energy budgets of these sensor nodes. More generally, if the network lifetime is defined as the time duration before a given portion of sensor nodes run out of energy, our proposed scheme can be applied repeatedly to extend network lifetime after the exhaustion of a sensor node's energy.

The notations we have introduced and will introduce later are summarized in Table I.

TABLE I
SUMMARY OF NOTATIONS

| Notation | Definition |
|---|---|
| $n$ | number of sensor nodes |
| $E$ | designated error bound on aggregate data |
| $S$ | sensing rate |
| $e_i$ | error bound allocated to sensor node $i$ |
| $u_i(\cdot)$ | precision-driven update rate of sensor node $i$ (as a function of its allocated error bound) |
| $s_i$ | energy cost for sensor node $i$ to transmit a data update |
| $v_i$ | energy cost for sensor node $i$ to receive a data update |
| $p_i$ | residual energy of sensor node $i$ |
| $r_i$ | normalized energy consumption rate of sensor node $i$ |

## IV. PRECISION ALLOCATION IN SINGLE-HOP NETWORKS

We start by investigating the precision allocation in a single-hop network where each sensor node sends its local readings to the base station directly. In this case, $s_i$ refers to the energy cost for node $i$ to send a data update to the base station. Single-hop networks are preferred by many applications due to a number of reasons [11]. First, the limitation of sensor designs (e.g., simplex MAC with limited buffer) may make relaying practically infeasible. Second, breaking the transmission into a number of short hops does not necessarily favor energy efficiency compared to a single long hop due to the receiving cost. More importantly, the analysis of precision allocation in a single-hop network provides insights on the allocation in a multi-hop network. The adaptive precision allocation scheme described in this section serves as a building block of the one we shall propose for multi-hop networks in Section V.

### A. Analysis of Optimal Precision Allocation

Consider a snapshot of the network. Let $e_1$, $e_2$, $\cdots$, $e_n$ be the error bounds currently allocated to sensor nodes 1, 2, $\cdots$, $n$ respectively. Since the sensor nodes in a single-hop network are not involved in relaying data from other sensor nodes to the base station, the energy consumption rate of sensor node $i$ is simply

$$u_i(e_i) \cdot s_i.$$

Suppose the residual energy of sensor node $i$ is $p_i$. Then, the expected lifetime of sensor node $i$ is given by

$$\frac{p_i}{u_i(e_i) \cdot s_i}.$$

Therefore, the network lifetime is given by

$$\min_{1 \le i \le n} \frac{p_i}{u_i(e_i) \cdot s_i}.$$

The objective of precision allocation is to find a set of error bounds $e_1$, $e_2$, $\cdots$, $e_n$ that maximizes the network lifetime under the constraint

$$\sum_{i=1}^{n} e_i \le E.$$

We now analyze the optimal precision allocation. For simplicity, we shall assume functions $u_i(\cdot)$'s are continuous and denote the inverse function of $u_i(\cdot)$ by $u_i^{-1}(\cdot)$.

Since $u_i(\cdot)$ is non-increasing, the minimum lifetime of sensor node $i$ is given by

$$l_i = \frac{p_i}{u_i(0) \cdot s_i}.$$

Without loss of generality, suppose

$$l_1 \le l_2 \le \cdots \le l_n.$$

For each pair $(i, j)$ where $i \le j$, consider the error bound $u_i^{-1}(\frac{p_i}{l_j \cdot s_i})$ that makes the lifetime of sensor node $i$ equivalent to the minimum lifetime of sensor node $j$. Since $u_i(\cdot)$ is non-increasing, it follows from $l_j \le l_{j+1}$ that

$$u_i^{-1}(\frac{p_i}{l_j \cdot s_i}) \le u_i^{-1}(\frac{p_i}{l_{j+1} \cdot s_i}).$$

Thus, given any $1 \le j < n$,

$$\sum_{i=1}^{j} u_i^{-1}(\frac{p_i}{l_j \cdot s_i}) \le \sum_{i=1}^{j} u_i^{-1}(\frac{p_i}{l_{j+1} \cdot s_i}) = \sum_{i=1}^{j+1} u_i^{-1}(\frac{p_i}{l_{j+1} \cdot s_i}).$$

This implies $\sum_{i=1}^{j} u_i^{-1}(\frac{p_i}{l_j \cdot s_i})$ is non-decreasing with increasing $j$. Note that when $j = 1$,

$$\sum_{i=1}^{j} u_i^{-1}(\frac{p_i}{l_j \cdot s_i}) = u_1^{-1}(u_1(0)) = 0.$$

Therefore, given an error bound $E > 0$ on aggregate data, if $\sum_{i=1}^{n} u_i^{-1}(\frac{p_i}{l_n \cdot s_i}) > E$, there must exist a $j^*$ $(1 \le j^* < n)$ such that

$$\sum_{i=1}^{j^*} u_i^{-1}(\frac{p_i}{l_{j^*} \cdot s_i}) \le E < \sum_{i=1}^{j^*+1} u_i^{-1}(\frac{p_i}{l_{j^*+1} \cdot s_i}).$$

Since

$$\sum_{i=1}^{j^*+1} u_i^{-1}(\frac{p_i}{l_{j^*+1} \cdot s_i}) = \sum_{i=1}^{j^*} u_i^{-1}(\frac{p_i}{l_{j^*+1} \cdot s_i}),$$

we have

$$\sum_{i=1}^{j^*} u_i^{-1}(\frac{p_i}{l_{j^*} \cdot s_i}) \le E < \sum_{i=1}^{j^*} u_i^{-1}(\frac{p_i}{l_{j^*+1} \cdot s_i}).$$

Hence, there also exists an $l^*$ $(l_{j^*} \le l^* < l_{j^*+1})$ such that

$$\sum_{i=1}^{j^*} u_i^{-1}(\frac{p_i}{l^* \cdot s_i}) = E. \quad (3)$$

On the other hand, if $\sum_{i=1}^{n} u_i^{-1}(\frac{p_i}{l_n \cdot s_i}) \le E$, since $u_i(\cdot)$'s are non-increasing and $u_i^{-1}(0) = \infty$, there exists an $l^*$ $(l^* \ge l_n)$ such that

$$\sum_{i=1}^{n} u_i^{-1}(\frac{p_i}{l^* \cdot s_i}) = E. \quad (4)$$

For convenience, we shall denote $j^* = n$ in this case so that (4) is consistent with (3).

*Theorem 1:* An optimal precision allocation is given by

$$e_i^* = \begin{cases} u_i^{-1}(\frac{p_i}{l^* \cdot s_i}) & 1 \le i \le j^*, \\ 0 & j^* < i \le n, \end{cases}$$

which has a lifetime of $l^*$.

**Proof:** See [24] for details. □

Theorem 1 implies that the sensor nodes with high residual energy ($p_i$), slow change in readings (i.e., low $u_i(0)$), and low communication cost ($s_i$) may be assigned zero error bounds. The sensor nodes allocated non-zero error bounds in an optimal precision allocation must be equal in the energy consumption rate normalized by the residual energy:

$$r_i = \frac{u_i(e_i) \cdot s_i}{p_i}.$$

We shall call $r_i$ the *normalized energy consumption rate*. To extend network lifetime, it is important to *balance* the normalized energy consumption rates of the sensor nodes.

### B. Adaptive Precision Allocation

In practice, the exact forms of $u_i(\cdot)$'s (i.e., the changing patterns of sensor readings) may not be known a priori and they may even change dynamically. Thus, we propose a sample-based precision allocation method, which will further be used to design an adaptive precision allocation scheme. The key idea is to let each sensor node report to the base station a number of *sample error bounds* and the associated normalized energy consumption rates based on historical sensor readings. The base station optimizes precision allocation based on these samples to extend network lifetime. Since the general relationships between error bounds and precision-driven update rates are not known, an additional constraint here is that the error bound of each sensor node can only be set to one of its samples. Such allocations are called *sample precision allocations* and the one that maximizes network lifetime is called the *optimal sample precision allocation*.

Assume that each sensor node estimates $m$ samples. For each sensor node $i$, let $e_{i,1} < e_{i,2} < \cdots < e_{i,m}$ be the list of sample error bounds, and $r_{i,1}, r_{i,2}, \cdots, r_{i,m}$ be the associated normalized energy consumption rates. It follows that $r_{i,1} \geq r_{i,2} \geq \cdots \geq r_{i,m}$. Suppose the smallest sample error bounds for the sensor nodes do not add up to the designated bound on aggregate data, i.e., $e_{1,1} + e_{2,1} + \cdots + e_{n,1} \leq E$.[1] Algorithm 1 presents the pseudocode to compute the optimal sample precision allocation.

**Algorithm 1** Optimal Sample Precision Allocation in a Single-Hop Network

**Input:**

$E$: error bound of aggregate data

$e_{i,*}, r_{i,*}$: sample error bounds and normalized energy consumption rates

**Output:**

$e_{i,x_i}$: error bound of each sensor in optimal allocation

1: **for** $i = 1$ to $n$ **do**
2: $\quad x_i = 1$;
3: **end for**
4: **while** $\min_{1 \leq i \leq n} x_i \neq m$ **do**
5: $\quad j = \arg \max_{1 \leq i \leq n, x_i \neq m} r_{i,x_i}$;
6: $\quad$ **if** $e_{j,x_j+1} + \sum_{i \neq j} e_{i,x_i} \geq E$ **then**
7: $\quad\quad$ break;
8: $\quad$ **end if**
9: $\quad x_j = x_j + 1$;
10: **end while**

Initially, the error bound of each sensor node is set to its smallest sample (steps 1 to 3). In each iteration of steps 4 to 10, the error bound of the sensor node having the highest energy consumption rate is replaced with its next smallest sample. The iteration stops if a new replacement would make the total bound of the sensor nodes exceed the designated bound on aggregate data (steps 6 to 7). The worst-case time complexity of Algorithm 1 is $O(mn)$.[2]

*Theorem 2:* The sample precision allocation computed by Algorithm 1 maximizes network lifetime.

**Proof:** See [24] for details. □

Adaptive precision allocation works by adjusting the error bounds of sensor nodes periodically. The interval between two successive adjustments is called an *adjustment period*, which is much longer than an epoch. At the beginning of an adjustment period, each sensor node $i$ selects a list of sample error bounds $e_{i,1}, e_{i,2}, \cdots, e_{i,m}$. The sensor node keeps track of the precision-driven update rates $u_{i,1}, u_{i,2}, \cdots, u_{i,m}$ of these error bounds as it captures new readings. At the end of the adjustment period, the sensor node computes the normalized energy consumption rate $r_{i,j}$ of each $e_{i,j}$ as[3]

$$r_{i,j} = \frac{u_{i,j} \cdot s_i}{p_i}.$$

The sample error bounds $e_{i,j}$'s and normalized energy consumption rates $r_{i,j}$'s are then reported to the base station. On receiving the samples, the base station computes the optimal sample precision allocation using Algorithm 1. If the computed error bounds are different from those currently allocated to the sensor nodes, they are then sent to the sensor nodes for their adjustments.

Algorithm 1 and Theorem 2 are generic in that they are applicable to any list of samples. In this paper, we propose to choose a set of sample error bounds that are exponentially spaced for each sensor node. The closer the samples to the current error bound, the smaller the difference between neighboring samples. The motivation is to adjust the error bounds at coarser granularity when they are significantly far away from the optimum, and adjust them at finer granularity when they are close to the optimum. Let $e_i$ be the current error bound of sensor node $i$. Then, the sample error bounds of $i$ range from $\frac{1}{2}e_i$ to $\frac{3}{2}e_i$. Given the number of samples $m = 2k + 1$, the sample error bounds are selected as

$$\frac{1}{2}e_i, \frac{3}{4}e_i, \cdots, \frac{2^k - 1}{2^k}e_i, e_i, \frac{2^k + 1}{2^k}e_i, \cdots, \frac{5}{4}e_i, \frac{3}{2}e_i.$$

## V. Precision Allocation in Multi-Hop Networks

### A. Modeling In-Network Aggregation

Multi-hop networks are necessary if the base station is beyond the radio coverage of some sensor nodes. In-network aggregation is an important technique to reduce the network traffic of data collection in multi-hop networks [25], [6], [8], [17]. Specifically, the sensor nodes are organized in a tree structure rooted at the base station. On receiving data from its children, each intermediate node aggregates the data before forwarding them upstream, thereby cutting down the volume of data transmitted over the upper-level links in the tree. Under this architecture, $s_i$ refers to the energy cost for node $i$ to send a data update to $i$'s parent, and $v_i$ refers to the energy cost for node $i$ to receive a data update from a child.

Like in a single-hop network, each sensor node $i$ is allocated an error bound $e_i$ for its local readings. We shall call it the *local error bound*. The total error bound allocated to the sensor nodes in the subtree rooted at sensor node $i$ is referred to as its *gross error bound*, denoted by $E_i$. Note that the gross and local error bounds of a leaf sensor node are the same.

In a multi-hop network, each sensor node guarantees to update its parent whenever the partial aggregate result over the subtree rooted at it changes beyond its gross error bound. Specifically, in each epoch, a leaf sensor node sends its reading

[1] Our proposed sample selection method (to be discussed later in this section) satisfies this constraint.

[2] As shall be shown in Section VI, a small $m$ is sufficient to achieve near optimal network lifetime.

[3] The residual energy of a sensor node can be estimated from its battery voltage [2], [3].

to the parent if the reading has changed beyond its local error bound since the last update to the parent. Each intermediate sensor node in the tree maintains the data value reported by each child as well as its local sensor reading at the time of its last update to the parent. In each epoch, the sensor node re-aggregates the data values and sends the updated aggregate value to its parent if (a) it has received data updates from at least one child; or (b) the local sensor reading has changed beyond the local error bound since the last update to its parent. If neither event (a) nor (b) occurs, the intermediate sensor node does not need to update its parent because the aggregate value cannot have changed beyond its gross error bound.

Let $U_i$ be the rate of data updates sent by each sensor node $i$ to its parent. If $i$ is a leaf node, $U_i$ is simply $i$'s precision-driven update rate, i.e., $U_i = u_i(e_i)$. If $i$ is an intermediate node, $U_i$ depends on the rates at which events (a) and (b) occur. We first analyze event (a). In each epoch, the probability that an $i$'s child $c$ sends a data update to node $i$ is $\frac{U_c}{S}$, where $S$ is the sensing rate. For simplicity, we assume that the changes in the readings at different sensor nodes are independent. Then, the probability that at least one child sends a data update to sensor node $i$ in an epoch is given by

$$1 - \prod_{c \in C_i} \left(1 - \frac{U_c}{S}\right),$$

where $C_i$ is the set of $i$'s children. On the other hand, in each epoch, the probability that event (b) occurs can be approximated[4] by $\frac{u_i(e_i)}{S}$. Therefore, the probability that sensor node $i$ sends a data update to its parent in an epoch is

$$1 - \left(1 - \frac{u_i(e_i)}{S}\right) \cdot \prod_{c \in C_i} \left(1 - \frac{U_c}{S}\right).$$

As a result,

$$U_i = S \cdot \left(1 - \left(1 - \frac{u_i(e_i)}{S}\right) \cdot \prod_{c \in C_i} \left(1 - \frac{U_c}{S}\right)\right). \quad (5)$$

Taking into consideration the energy consumed in sending and receiving data updates, the normalized energy consumption rate of sensor node $i$ is given by

$$\frac{U_i \cdot s_i + \sum_{c \in C_i} U_c \cdot v_i}{p_i}. \quad (6)$$

### *B. Adaptive Precision Allocation*

As shown in Section IV, to extend network lifetime, it is important to balance the normalized energy consumption rates of the sensor nodes, i.e., to minimize the maximum rate. Adaptive precision allocation in a multi-hop network also works by adjusting the error bounds of sensor nodes periodically. At the beginning of an adjustment period, each sensor node $i$ selects a list of sample local error bounds $e_{i,1}$, $e_{i,2}$, $\cdots$, $e_{i,m}$ and a list of sample gross error bounds $E_{i,1}$, $E_{i,2}$, $\cdots$, $E_{i,m}$. Like in a single-hop network, each sensor node keeps track of the precision-driven update rates $u_{i,1}$, $u_{i,2}$, $\cdots$, $u_{i,m}$ of the sample local error bounds for its local readings. At the end of the adjustment period, the sample gross error bounds and the associated data update rates and energy consumption rates are computed and propagated in a bottom-up manner from the leaf sensor nodes to the base station. Specifically, each sensor node $i$ sends a list $\langle E_{i,1}, U_{i,1}, R_{i,1} \rangle$, $\langle E_{i,2}, U_{i,2}, R_{i,2} \rangle$, $\ldots$, $\langle E_{i,m}, U_{i,m}, R_{i,m} \rangle$ to its parent, where $U_{i,j}$ is the rate of data updates sent by node $i$ to its parent under the optimal allocation of $E_{i,j}$ in the subtree rooted at $i$, and $R_{i,j}$ is the corresponding maximum normalized energy consumption rate of the sensor nodes in the subtree.

If $i$ is a leaf sensor node, its sample gross error bounds are the same as its sample local error bounds. Thus, $U_{i,j}$'s are simply the precision-driven update rates $u_{i,j}$'s. Like the sensor nodes in a single-hop network, $i$ computes the normalized energy consumption rate of $e_{i,j}$ as

$$r_{i,j} = \frac{u_{i,j} \cdot s_i}{p_i}.$$

Since $i$ is a leaf node, $R_{i,j}$'s are simply $r_{i,j}$'s.

If $i$ is an intermediate sensor node, it collects the lists of sample gross error bounds, data update rates and energy consumption rates from all of its children. Together with the precision-driven update rates $u_{i,1}$, $u_{i,2}$, $\cdots$, $u_{i,m}$ measured locally, sensor node $i$ computes the optimal sample precision allocation for each gross error bound $E_{i,j}$ using Algorithm 2. Given a sample gross error bound $E_{i,j}$, all sample local error bounds $e_{i,h}$ where $e_{i,h} < E_{i,j}$ are considered (step 3). For each $e_{i,h}$, the optimal allocation of $E_{i,j} - e_{i,h}$ among $i$'s children is computed using Algorithm 1 (step 4). Suppose $E_{c,x_c}$ is the gross error bound of each child $c$ in the optimal allocation. Then, $U_{c,x_c}$ is the corresponding data update rate from $c$ to $i$, and $R_{c,x_c}$ is the corresponding maximum normalized energy consumption rate of the sensor nodes in the subtree rooted at $c$. The data update rate from sensor node $i$ to its parent and $i$'s normalized energy consumption rate are then computed based on (5) and (6) respectively (steps 6 and 7). The maximum normalized energy consumption rate of the sensor nodes in the subtree rooted at $i$ can then be computed (step 7). The sample local error bound $e_{i,h}$ and the corresponding optimal allocation among $i$'s children that lead to the lowest maximum energy consumption rate are selected as the optimal sample precision allocation for $E_{i,j}$ (steps 8 to 13). The worst-case time complexity of Algorithm 2 is $O(m^2 \cdot |C_i|)$, where $|C_i|$ is the number of $i$'s children. Sensor node $i$ keeps the optimal allocation for each gross error bound $E_{i,j}$, and sends the list $\langle E_{i,1}, U_{i,1}, R_{i,1} \rangle$, $\langle E_{i,2}, U_{i,2}, R_{i,2} \rangle$, $\ldots$, $\langle E_{i,m}, U_{i,m}, R_{i,m} \rangle$ to its parent.

The base station, on receiving the lists from all of its children, computes the optimal sample precision allocation among the children using Algorithm 1. The computed error bounds are then sent to the sensor nodes for their adjustments in a top-down manner. The base station sends to its children the gross error bounds allocated to them. An intermediate sensor

[4]This is a conservative estimate. Strictly speaking, the rate at which event (b) occurs is lower than the precision-driven update rate $u_i(e_i)$. This is because the up-to-date sensor reading of $i$ is also conveyed in the updates to $i$'s parent when event (a) occurs.

**Algorithm 2** Sample Precision Allocation at Sensor Node $i$ in a Multi-Hop Network

**Input:**

$E_{i,j}$: sample gross error bound of sensor node $i$

$E_{c,*}, U_{c,*}, R_{c,*}$: sample gross error bounds, data update rates and maximum normalized energy consumption rates received from each child $c$ of $i$

**Output:**

$E_{c,y_c}$: gross error bound of each child $c$ of $i$ in optimal allocation

$e_{i,y_i}$: local error bound of sensor node $i$ in optimal allocation

$U_{i,j}$: data update rate from sensor node $i$ to its parent in optimal allocation

$R_{i,j}$: maximum normalized energy consumption rate in the subtree rooted at $i$ in optimal allocation

```
R_{i,j} = +∞;
for h = 1 to m do
    if e_{i,h} < E_{i,j} then
        compute the optimal sample allocation of error bound E_{i,j} − e_{i,h} among i's children using Algorithm 1 based on E_{c,*} and R_{c,*};
        for each child c of i, let E_{c,x_c} be the error bound of c in the optimal allocation, then U_{c,x_c} is the corresponding data update rate from c to i, and R_{c,x_c} is the corresponding maximum normalized energy consumption rate of the sensor nodes in the subtree rooted at c;
        U_i = S · (1 − (1 − u_{i,h}/S) · ∏_{c∈C_i} (1 − U_{c,x_c}/S));
        R_i = max((U_i·s_i + ∑_{c∈C_i} U_{c,x_c}·v_i)/p_i, max_{c∈C_i} R_{c,x_c});
        if R_i < R_{i,j} then
            R_{i,j} = R_i;
            U_{i,j} = U_i;
            y_i = h;
            for each child c of i, y_c = x_c;
        end if
    end if
end for
```

node, on receiving its allocated gross error bound, retrieves the corresponding optimal allocation which contains a local error bound and a set of gross error bounds for its children. The intermediate sensor node applies the local error bound to its local readings and sends the gross error bounds to its children. A leaf sensor node, on receiving its allocated gross error bound, simply takes it as the local error bound.

Similar to adaptive precision allocation in a single-hop network, the sample gross and local error bounds of each sensor node are exponentially spaced around the current gross and local error bounds respectively. Let $E_i$ and $e_i$ be the current gross and local error bounds of sensor node $i$ respectively. Given the number of samples $m = 2k+1$, the sample gross error bounds are selected as

$$\frac{1}{2}E_i, \frac{3}{4}E_i, \cdots, \frac{2^k-1}{2^k}E_i, E_i, \frac{2^k+1}{2^k}E_i, \cdots, \frac{5}{4}E_i, \frac{3}{2}E_i,$$

and the sample local error bounds are selected as

$$\frac{1}{2}e_i, \frac{3}{4}e_i, \cdots, \frac{2^k-1}{2^k}e_i, e_i, \frac{2^k+1}{2^k}e_i, \cdots, \frac{5}{4}e_i, \frac{3}{2}e_i.$$

## VI. Performance Evaluation

### A. Experimental Setup

We have developed a simulator based on ns-2 (version 2.26) [26] and NRL's sensor network extension [27] to evaluate the proposed adaptive precision allocation scheme. Table II summarizes the system parameters and their settings.

TABLE II

System Parameters and Settings

| Parameter | Setting |
|---|---|
| Number of Sensor Nodes ($n$) | 10 (single-hop network)<br>20 (multi-hop network) |
| Energy Consumption for Sending a Message | $s \cdot (\alpha + \beta \cdot d^q)$<br>$\alpha = 50$ nJ/b<br>$\beta = 100$ pJ/b/m$^2$<br>$q = 2$<br>$s$: message size<br>$d$: transmission distance |
| Energy Consumption for Receiving a Message | $s \cdot \gamma$<br>$\gamma = 50$ nJ/b<br>$s$: message size |
| Power Consumption in Sleeping Mode | 0.016 mW |
| Initial Energy Budget at Each Sensor Node | 0.1 J |
| Epoch | 1 time unit |

The simulator includes the detailed models of the MAC and physical layers for wireless networks. The sensor nodes can operate in one of three modes: sending message, receiving message, and sleeping. These modes differ in energy consumption. We used similar energy models to those in other studies [11], [21], [28]. The energy consumed by a sensor node to send a message is $s \cdot (\alpha + \beta \cdot d^q)$, where $s$ is the message size, $\alpha$ is a distance-independent term, $\beta$ is the coefficient for a distance-dependent term, $q$ is the exponent for the distance-dependent term, and $d$ is the transmission distance. The energy consumed by a sensor node to receive a data update is $s \cdot \gamma$, where $\gamma$ is a coefficient independent of transmission distance. The settings of these parameters are shown in Table II. The power consumption in the sleeping mode was set at 0.016 mW. For simplicity, the energy overhead of mode switching is ignored. The initial energy budget at each sensor node was set at 0.1 Joule. The epoch was assumed to be 1 time unit.

We simulated a single-hop network of 10 sensor nodes and a multi-hop network of 20 sensor nodes. Their layouts are shown in Figures 2 and 3.

We made use of the data provided by the Tropical Atmosphere Ocean (TAO) project [29] to simulate the sensor readings. Oceanographic and meteorological data are collected from a wide range of mooring sites in the TAO project for improved detection, understanding and prediction of El Nino

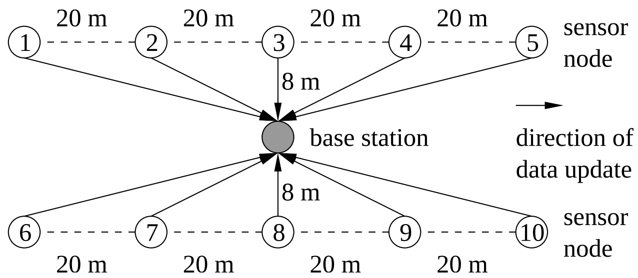



Fig. 2. Single-Hop Network Layout

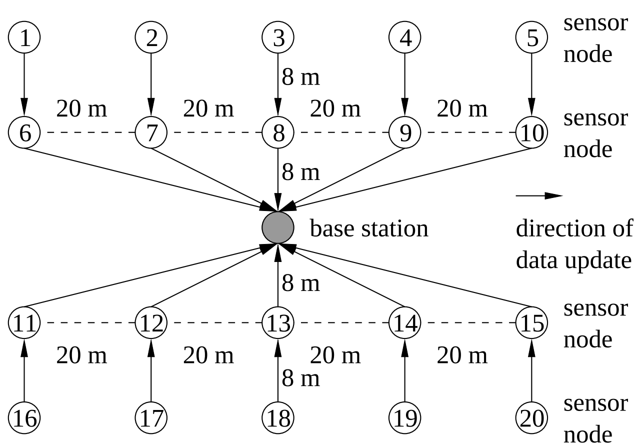



Fig. 3. Multi-Hop Network Layout

and La Nina. We selected a subset of the sites and mapped them to the sensor nodes in our networks. We used the air temperature (AT) and winds (WIND) data from 1999 to 2000 in our experiments. The data of different sites are similar in magnitude. Figure 4 shows some representative segments of the AT and WIND data traces. In general, the WIND data fluctuate more widely than the AT data.

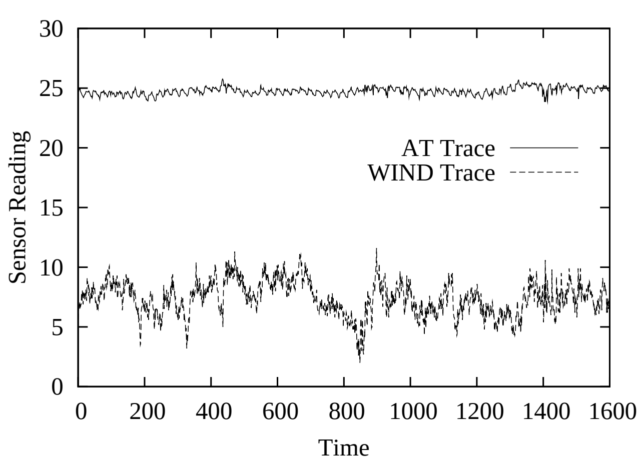



Fig. 4. Sample Data Traces

The base station computes the AVERAGE aggregation of the readings collected from all sensor nodes with a designated error bound $E$. As discussed in Section III-A, in this case, the total error bound allocated to the sensor nodes should be capped by $n \cdot E$, where $n$ is the number of sensor nodes. The experiments started with the error bound uniformly allocated to the sensor nodes, i.e., each sensor node is allocated an error bound of $E$. The following precision allocation schemes were simulated for performance comparison. Table III lists the scheme-specific parameters and their settings. We measured the energy consumption of each sensor node and the network lifetime in the experiments.

TABLE III

SCHEME-SPECIFIC PARAMETERS AND SETTINGS

| Parameter | Setting |
|---|---|
| **All Schemes** | |
| Data Update Message (data value + timestamp) | 8 bytes |
| Error Bound Adjustment Message (error bound + timestamp) | 8 bytes (except Uniform-PA) |
| **Adaptive-PA** | |
| Adjustment Period | 500 time units |
| Number of Samples ($m$) | 7 |
| Sample Report Message ($m$ samples + timestamp) | $4m$ + 4 bytes (single-hop network) $8m$ + 4 bytes (multi-hop network) |
| **Burden-PA** | |
| Adjustment Period | 200 time units |
| Shrink Percentage | 5% |
| **PGain-PA** | |
| Adjustment Period | 500 time units |
| Shrink Percentage | 40% |
| Sample Report Message (potential gain + timestamp) | 8 bytes |

- **Proposed Adaptive Precision Allocation (Adaptive-PA):** This is the adaptive precision allocation scheme proposed in Sections IV-B and V-B. By default, each sensor node selected $m$ = 7 sample error bounds and the adjustment period was set at 500 time units. The performance impacts of $m$ and adjustment period are investigated in Section VI-B. In a single-hop network, since the base station is aware of the current error bounds of the sensor nodes, it can infer the sample error bounds tracked by each sensor node which are exponentially spaced around the current bound. Therefore, only the estimated energy consumption rates need to be reported to the base station at the end of the adjustment period. The payload size of a sample report message was thus set at $4m$ + 4 bytes.[5] Similarly, in a multi-hop network, since each intermediate sensor node is aware of the current gross error bounds of its children, it can infer the sample gross error bounds tracked by each child. At the end of the adjustment period, only the estimated update rates and maximum energy consumption rates

[5]A timestamp of 4 bytes was assumed to be included in all messages for ordering and consistency purposes.

are reported by a sensor node to its parent. Thus, the payload size of a sample report message was set at 8$m$ + 4 bytes.

- **Uniform Precision Allocation (Uniform-PA)**: The error bound is evenly partitioned among the sensor nodes, i.e., the precision allocation remains the initial one. This is a simple and static scheme which does not differentiate the sensor nodes by the changing pattern of sensor readings, residual energy, and communication cost with the base station.

- **Burden-based Precision Allocation (Burden-PA)**: Olston *et al.* [5] presented a burden-based precision allocation scheme for aggregate queries over distributed data streams. Their objective was to minimize the total communication cost between data sources and the data sink. In our experiments, the energy consumed by each sensor node to send a data update to the base station was taken as a measure of its communication cost. Burden-PA works by periodically reducing the error bound of each sensor node by a *shrink percentage* and redistributing the leftover portion among the sensor nodes. As suggested by [5], the shrink percentage was set at 5%. The default adjustment period was set at 200 time units which showed the best performance in our experiments (see Section VI-B).

- **Potential-Gain-based Precision Allocation (PGain-PA)**: To reduce the number of messages in the network, Deligiannakis *et al.* [8] presented a precision allocation scheme for data aggregation based on online estimation of potential gains. Similar to Burden-PA, PGain-PA periodically reduces the error bound of each sensor node by a *shrink percentage* and redistributes the leftover portion among the sensor nodes. As suggested by [8], the shrink percentage was set at 40%. The default adjustment period was set at 500 time units which showed the best performance in our experiments (see Section VI-B).

### B. Effect of System-Specific Parameters

First, we investigate the performance impact of the number of sample error bounds $m$ in the proposed Adaptive-PA scheme. Figure 5 shows the network lifetime for different $m$ values when the error bound $E$ was set at 0.4.[6] Note that when $m$ = 1, the current error bound is the only sample. Thus, the optimal sample precision allocation computed by Algorithm 1 is always the same as the current allocation. Since the experiments started with uniformly allocated error bounds, Adaptive-PA degenerates to Uniform-PA.

The flexibility of precision allocation increases with $m$. As seen from Figure 5, an $m$ value of 3 improves the network lifetime by over 40% compared to that of $m$ = 1 for both

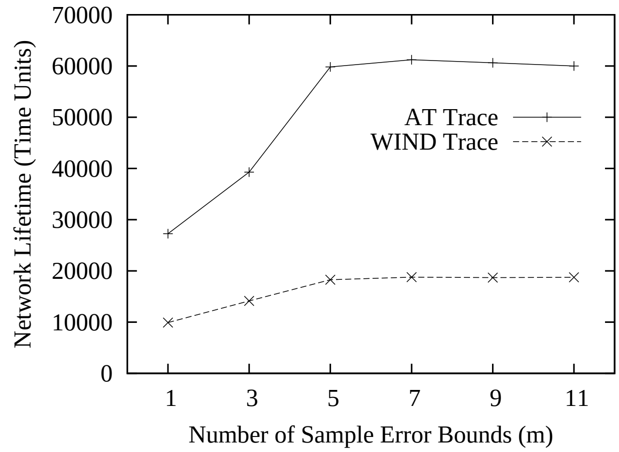



Fig. 5. Network Lifetime vs. Number of Sample Error Bounds in Adaptive-PA (Single-Hop Network, $E$ = 0.4)

traces. The network lifetime increases rapidly with $m$ up to 5. It is generally insensitive to $m$ when $m$ exceeds 5. The default $m$ was set at 7 in the remaining experiments.

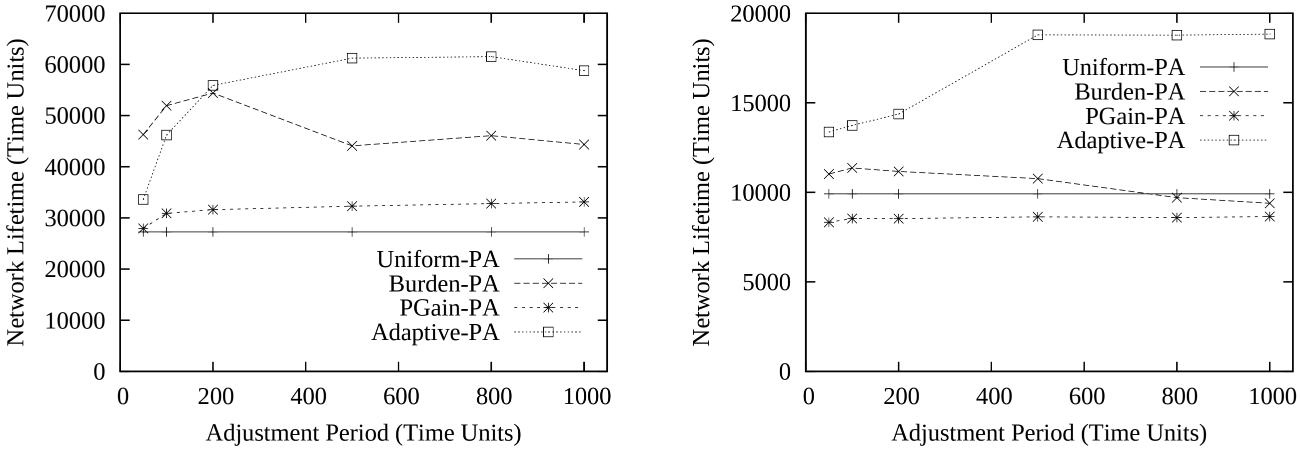
Adaptive-PA, Burden-PA and PGain-PA all adjust the error bounds of sensor nodes periodically. The setting of adjustment period reflects a tradeoff between overhead and adaptivity. In general, the shorter the adjustment period, the higher the overhead. On the other hand, the longer the adjustment period, the less adaptive the precision allocation scheme. Figure 6 shows the network lifetime for different adjustment periods when $E$ was set at 0.4. As expected, the graph of network lifetime is convex for most combinations of precision allocation scheme and data trace. We have selected the default adjustment period for each allocation scheme as the one that showed the best performance (i.e., 500, 200 and 500 time units for Adaptive-PA, Burden-PA and PGain-PA respectively).

### C. Comparing Adaptive-PA with Existing Schemes

Figure 7 shows the network lifetime as a function of the designated error bound $E$ on data aggregation for different precision allocation schemes in the single-hop network of Figure 2. It can be seen that the lifetime increases with error bound. The proposed Adaptive-PA scheme significantly outperforms the other schemes for both traces tested.

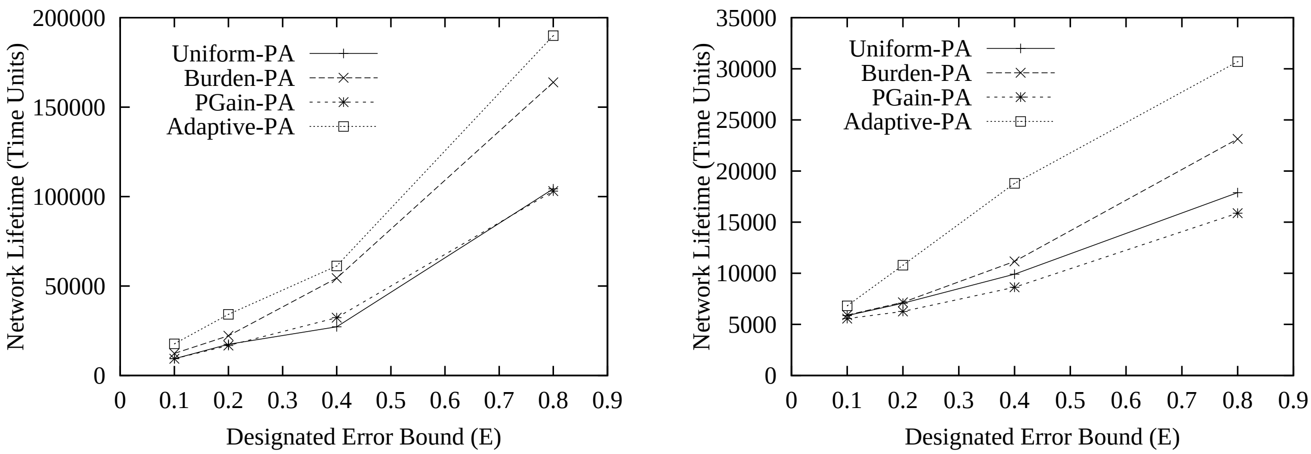
Even when the readings at all sensor nodes follow similar changing patterns, it is not desirable to allocate the same error bound to all sensor nodes due to their geographically distributed nature. In a single-hop network, the sensor node farther away from the base station consumes more energy in sending a data update than the sensor node closer to the base station. Among the four precision allocation schemes examined, Uniform-PA and PGain-PA do not take this heterogeneity into consideration. Thus, as shown in Figure 7, Adaptive-PA improves the network lifetime by a factor up to 2.2 compared to Uniform-PA and PGain-PA. To show the importance of balancing energy consumption in extending network lifetime, we plot in Figure 8 the total energy consumed by each sensor

[6] Only the experimental results of the single-hop network (Figure 2) are reported in this section to show the effect of system-specific parameters. The results of the multi-hop network (Figure 3) have similar trends.



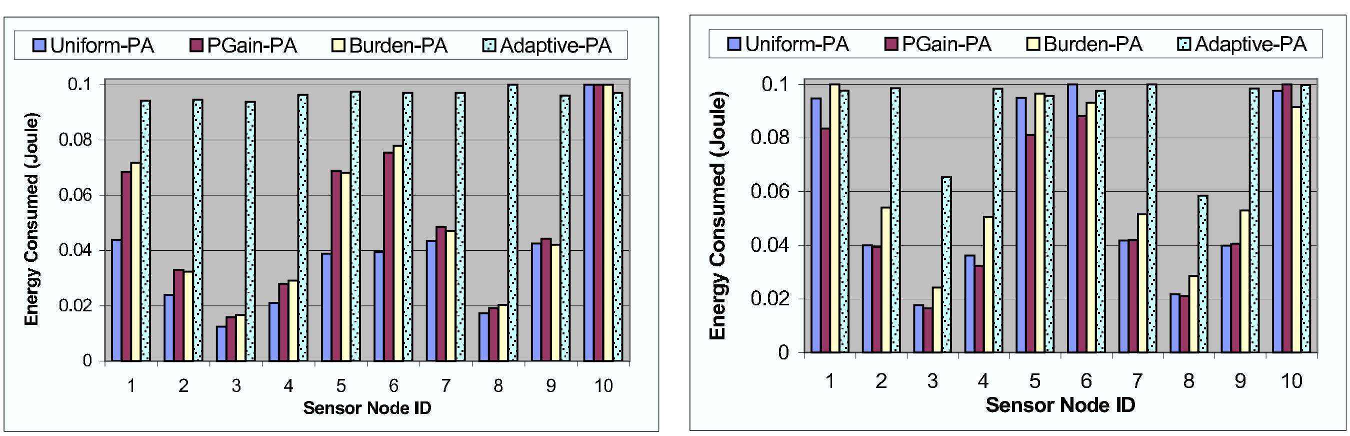

(a) AT Trace (b) WIND Trace

Fig. 6. Network Lifetime vs. Adjustment Period (Single-Hop Network, $E$ = 0.4)




(a) AT Trace (b) WIND Trace

Fig. 7. Network Lifetime vs. Designated Error Bound (Single-Hop Network)




(a) AT Trace (b) WIND Trace

Fig. 8. Energy Consumed at Different Sensor Nodes (Single-Hop Network, $E$ = 0.4)

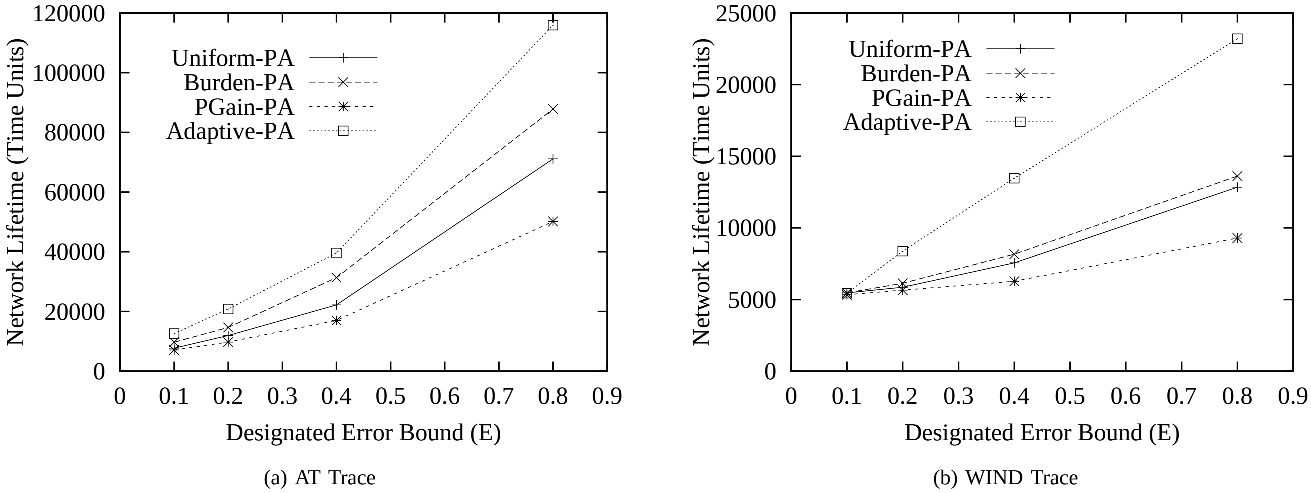



Fig. 9. Network Lifetime vs. Designated Error Bound (Multi-Hop Network)

node by the time when the first sensor node runs out of energy (i.e., the network lifetime elapsed). As can be seen, under Adaptive-PA, most sensor nodes were close to exhausting their energy when the network lifetime elapsed. However, under Uniform-PA and PGain-PA, the sensor nodes close to the base station (i.e., nodes 3 and 8 in Figure 2) consumed as low as 15% of the energy only.

Burden-PA, on the other hand, considers the heterogeneity in communication cost due to transmission distance. However, the objective of Burden-PA is to minimize the total communication cost. Figure 7 shows that Burden-PA results in a shorter network lifetime by up to 41% than Adaptive-PA. This implies minimizing network-wide total energy consumption does not necessarily balance the energy consumption of the sensor nodes. As seen from Figure 8, under Burden-PA, sensor nodes 3 and 8 consumed less than 30% of the energy when the network lifetime elapsed.

It is also interesting to note that even under Adaptive-PA, the sensor nodes are sometimes not well balanced in energy consumption. For example, sensor nodes 3 and 8 consumed around 60% of the energy for the WIND trace. This is because these sensor nodes have low energy costs to communicate with the base station. Even if they are assigned error bounds close to 0, their energy consumption is not as high as the other nodes for the WIND trace. This is consistent with the analysis of optimal precision allocation in Section IV-A.

We have implemented in-network aggregation in the experiments. Figure 9 shows the results for the multi-hop network in Figure 3. The relative performance of the precision allocation schemes remains similar to that in the single-hop network. The network lifetime increases rapidly with error bound. For example, under Adaptive-PA, increasing $E$ from 0.1 to 0.2 prolongs the network lifetime by 65% and 54% for the AT and WIND traces respectively. This demonstrates the effectiveness of approximate data aggregation in improving energy efficiency.

Comparing the performance of different precision allocation schemes, Adaptive-PA significantly outperforms the other schemes for both traces tested. In general, the improvement increases with increasing error bound. This is because a large error bound gives more flexibility in optimizing precision allocation to balance the energy consumption of the sensor nodes. As seen from Figure 9, Adaptive-PA outperforms Uniform-PA, Burden-PA and PGain-PA by factors up to 1.8, 1.7 and 2.5 respectively. Comparing Figures 7 and 9, it can also be observed that the relative improvement of Adaptive-PA over Burden-PA is greater in the multi-hop network than in the single-hop network. This is because Burden-PA considers data sources independently and no in-network aggregation is modeled. Our proposed Adaptive-PA scheme accounts for the impact of in-network aggregation in multi-hop networks and models energy consumption more accurately.

## VII. Conclusion

We have exploited the tradeoff between data quality and energy consumption to extend the lifetime for precision-constrained data aggregation in wireless sensor networks. We partition the precision constraint and allocate error bounds to individual sensor nodes in a coordinated fashion. The purpose of precision allocation is to differentiate the quality of data collected from different sensor nodes, thereby balancing their energy consumption. We have analyzed the optimal precision allocation in terms of network lifetime and have proposed an adaptive precision allocation scheme that dynamically adjusts the error bounds of sensor nodes. Experimental results using real data traces show that: (1) due to geographically distributed nature of sensor networks, uniform precision allocation does not perform well even if the readings at all sensor nodes follow similar changing patterns; (2) to extend network lifetime, it is more important to balance the energy consumption of the sensor nodes than to minimize network-wide total energy consumption of the sensor nodes; (3) the proposed adaptive precision allocation scheme significantly outperforms existing methods over a wide range of system configurations.

## Acknowledgment

Jianliang Xu's work was partially supported by a grant from the Research Grants Council of the Hong Kong SAR, China (Project No. HKBU 2115/05E).

## References


[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[2] R. Szewczyk, E. Osterweil, J. Polastre, M. Hamilton, A. Mainwaring, and D. Estrin, "Habitat monitoring with sensor networks," *Communications of the ACM*, vol. 47, no. 6, pp. 34–40, June 2004.

[3] J. Gehrke and S. Madden, "Query processing in sensor networks," *IEEE Pervasive Computing*, vol. 3, no. 1, pp. 45–55, January–March 2004.

[4] G. J. Pottie and W. J. Kaiser, "Wireless intergrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, May 2000.

[5] C. Olston, J. Jiang, and J. Widom, "Adaptive filters for continuous queries over distributed data streams," in *Proc. ACM SIGMOD'03*, June 2003, pp. 563–574.

[6] M. A. Sharaf, J. Beaver, A. Labrinidis, and P. K. Chrysanthis, "TiNA: A scheme for temporal coherency-aware in-network aggregation," in *Proc. ACM MobiDE'03*, Sept. 2003, pp. 69–76.

[7] Q. Han, S. Mehrotra, and N. Venkatasubramanian, "Energy efficient data collection in distributed sensor environments," in *Proc. IEEE ICDCS'04*, Mar. 2004, pp. 590–597.

[8] A. Deligiannakis, Y. Kotidis, and N. Roussopoulos, "Hierarchical in-network data aggregation with quality guarantees," in *Proc. EDBT'04*, Mar. 2004, pp. 658–675.

[9] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall, 1996.

[10] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks," in *Proc. ACM MobiCom'00*, Aug. 2000, pp. 56–67.

[11] J. Pan, Y. T. Hou, L. Cai, Y. Shi, and S. X. Shen, "Topology control for wireless sensor networks," in *Proc. ACM MobiCom'03*, Sept. 2003, pp. 286–299.

[12] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "The design of an acquisitional query processor for sensor networks," in *Proc. ACM SIGMOD'03*, June 2003, pp. 491–502.

[13] X. Li, Y. J. Kim, R. Govindan, and W. Hong, "Multi-dimensional range queries in sensor networks," in *Proc. ACM SenSys'03*, Nov. 2003, pp. 63–75.

[14] Y. Yao, X. Tang, and E.-P. Lim, "In-network processing of nearest neighbor queries for wireless sensor networks," in *Proc. DASFAA'06*, Apr. 2006.

[15] J. Xu, X. Tang, and W.-C. Lee, "EASE: Energy-conserving approximate storage for querying object tracking sensor networks," in *Proc. IEEE SECON'05*, Sept. 2005.

[16] J. Considine, F. Li, G. Kollios, and J. Byers, "Approximate aggregation techniques for sensor databases," in *Proc. IEEE ICDE'04*, Mar. 2004, pp. 449–460.

[17] N. Shrivastava, C. Buragohain, D. Agrawal, and S. Suri, "Medians and beyond: New aggregation techniques for sensor networks," in *Proc. ACM SenSys'04*, Nov. 2004, pp. 188–200.

[18] A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proc. ACM MobiCom'01*, July 2001, pp. 166–179.

[19] D. Niculescu and B. Nath, "Ad Hoc Positioning (APS) using AoA," in *Proc. IEEE Infocom'03*, Apr. 2003, pp. 1734–1743.

[20] O. Younis and S. Fahmy, "Distributed clustering for ad-hoc sensor networks: A hybrid, energy-efficient approach," in *Proc. IEEE Infocom'04*, Mar. 2004, pp. 629–640.

[21] Y. T. Hou, Y. Shi, and H. D. Sherali, "Rate allocation in wireless sensor networks with network lifetime requirement," in *Proc. ACM MobiHoc'04*, May 2004, pp. 67–77.

[22] C. Buragohain, D. Agrawal, and S. Suri, "Power aware routing for sensor databases," in *Proc. IEEE Infocom'05*, Mar. 2005, pp. 1747–1757.

[23] I. Kang and R. Poovendran, "Maximizing network lifetime of broadcast over wireless stationary ad hoc networks," *ACM/Kluwer Mobile Networks and Applications*, vol. 11, no. 2, Apr. 2006.

[24] X. Tang and J. Xu, "Extending network lifetime for precision-constrained data aggregation in wireless sensor networks" (extended version), to be available from http://www.ntu.edu.sg/home/asxytang/.

[25] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TAG: A tiny aggregation service for ad-hoc sensor networks," in *Proc. USENIX OSDI'02*, Dec. 2002, pp. 131–146.

[26] "The network simulator - ns-2," http://www.isi.edu/nsnam/ns/.

[27] "NRL's sensor network extension to ns-2," http://nrlsensorsim.pf.itd.nrl.navy.mil/.

[28] W. Heinzelman, *Application-Specific Protocol Architectures for Wireless Networks*. Ph.D. Thesis, MIT, 2000.

[29] "Tropical Atmosphere Ocean (TAO) Project," http://www.pmel.noaa.gov/tao/data_deliv.