Adaptive Localized QoS-Constrained Data Aggregation and Processing in Distributed Sensor Networks

Jin Zhu, *Student Member*, *IEEE*, Symeon Papavassiliou, *Member*, *IEEE*, and Jie Yang, *Student Member*, *IEEE*

Abstract—In this paper, an efficient Quality of Service (QoS)-constrained data aggregation and processing approach for distributed wireless sensor networks is investigated and analyzed. One of the key features of the proposed approach is that the task QoS requirements are taken into account to determine when and where to perform the aggregation in a distributed fashion, based on the availability of local only information. Data aggregation is performed on the fly at intermediate sensor nodes, while at the same time the end-to-end latency constraints are satisfied. Furthermore, a localized adaptive data collection algorithm performed at the source nodes is developed that balances the design tradeoffs of delay, measurement accuracy, and buffer overflow, for given QoS requirements. The performance of the proposed approach is analyzed and evaluated, through modeling and simulation, under different data aggregation scenarios and traffic loads. The impact of several design parameters and tradeoffs on various critical network and application related performance metrics, such as energy efficiency, network lifetime, end-to-end latency, and data loss are also evaluated and discussed.

Index Terms—Sensor networks, distributed networks, data aggregation, quality of service.

1 Introduction

A distributed sensor network is usually a self-organized system composed of a large number of sensor nodes that collaborate with each other in order to measure different parameters that may vary with time and space, and send the corresponding data to a collector center for further processing. The collaboration among different sensor nodes is mostly realized through multihop network architectures due to their energy-efficiency and scalability features [1], [2]. Furthermore, since in sensor networks the data in the neighboring nodes are considered highly correlated [3], [4], [5], localized data processing [6], [7], [8] and data aggregation [9], [10], [11], [12] might dramatically decrease the amount of information to be transmitted.

Although several research works in the literature have discussed the problems of developing efficient routing and data aggregation processes mainly for energy savings/minimization in sensor networks (e.g., [8], [9], [10], [13], [14], [15]), several issues associated with the data aggregation process with the specific objective of meeting the task requirements (i.e., QoS-constrained data aggregation) are not yet well addressed. Given the fact that many sensing tasks present some strict reporting quality requirements (e.g., in a time critical application an obsolete sensor report that may exceed a given time threshold is discarded), development of efficient and feasible strategies that perform data aggregation in a distributed manner and with energy efficiency, in order to meet various quality requirements such as end-to-end latency and measurement accuracy, is of high research and practical importance.

In this paper, we study the data gathering and aggregation process in a distributed, multihop sensor network under specific QoS constraints. For a network, the data collection and dissemination is basically divided into two parts: the original data collection at the end nodes (i.e., source nodes) and the data transmission from the source nodes to the collector center through the intermediate nodes. The end nodes are the ones that are responsible for performing the actual measurements and for the collection of the required samples, while the intermediate nodes receive, process and forward samples originated from other nodes to the collector center.

Since in a distributed multihop sensor network the resulting end-to-end QoS heavily depends on the actual system conditions, traffic load, and the actions taken by each intermediate node, in the following the emphasis is placed on the operations performed at the intermediate sensor nodes. Therefore, we first present and analyze a QoS-constrained Data Aggregation and Processing approach (Q-DAP) that is performed at the intermediate nodes in a totally distributed fashion. Each intermediate sensor node determines independently whether or not to perform data aggregation randomly with some specific probability that is precalculated according to the resource conditions and the specific task requirements. One of the

- *J. Zhu is with the Department of Industrial Technology, University of Northern Iowa, Cedar Falls, IA 50614. E-mail: jin.zhu@ieee.org.*
- *S. Papavassiliou is with the Department of Electrical and Computer Engineering, National Technical University of Athens, Iroon Polytechniou 9, Athens 15773, Greece. E-mail: papavass@mail.ntua.gr.*
- *J. Yang is with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102. E-mail: jxy9918@njit.edu.*

Manuscript received 21 Feb. 2005; revised 30 May 2005; accepted 10 Aug. 2005; published online 26 July 2006.
Recommended for acceptance by I. Stojmenovic, S. Olariu, and D. Simplot-Ryl.
For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDSSI-0163-0205.

1045-9219/06/$20.00 © 2006 IEEE Published by the IEEE Computer Society

main principles of the proposed scheme is that the network does not need to be formed into clusters to perform the data aggregation, while the task QoS requirements are taken into account to determine when and where to perform the aggregation in a distributed fashion, based on the availability of local only information. Furthermore, a Localized Adaptive Data Collection and Aggregation (LADCA) approach for the end nodes is also proposed, which provides a method of adjusting measurement accuracy related parameters at the source nodes, in order to allow the system to adapt to the changing conditions.

It should be noted that, in the literature, the study of QoS guarantee in sensor networks is usually focused on the routing protocols tailored to meet the requirements (e.g., [9], [14], [16], [13]), while in this paper the emphasis is placed on the study of localized data collection and aggregation strategies that should be implemented at each individual sensor node in a distributed fashion, which is complementary to the applied QoS routing protocols, and can enhance the capability of QoS guarantee in sensor networks.

The remainder of the paper is organized as follows: In Section 2, we first introduce and describe the proposed Q-DAP approach that is performed at the intermediate nodes. Then, a statistical model that represents the data aggregation and report delivery process in sensor networks is described, and applied to the study of the performance of the proposed Q-DAP approach in terms of the end-to-end delay and probability of success in the report delivery process. In Section 3, we describe the proposed LADCA approach which is performed at the end nodes, with the objective of evaluating and balancing the tradeoffs among measurement accuracy, delay, and sensor buffer overflow. In Section 4, the performance of the proposed approach, under different data aggregation scenarios and traffic loads, is analyzed and evaluated through simulation and, finally, Section 5 concludes the paper.

2 QoS-Constrained Data Aggregation and Processing (Q-DAP) at Intermediate Nodes

In the following, we assume that when a sensor node receives a packet or message from its neighbor, it is able to either perform local processing and aggregation or just forward (relay) it, according to the QoS requirements of the corresponding applications. Here, the procedure that a sensor node locally generates and/or processes a measurement packet in which new data may be aggregated is referred to as reporting, while the corresponding new/updated packet is referred to as a report.

Therefore, the operation of the Q-DAP approach can be described as follows: When a sensor node receives a report from its neighbor, it first determines whether or not it would perform data aggregation on the report. The following different cases may occur. 1) If the delay constraint can be satisfied, the sensor node defers the report for a fixed time interval τ with probability γ, during which the node processes and aggregates any reports that arrive, and generates a new report before transmitting it to the next hop. With probability of $(1-\gamma)$ the sensor node will directly try to forward the report without introducing any deferred period. 2) If the delay constraint can be satisfied only if the report is not deferred, the sensor node simply tries to forward this report to the next hop. 3) If the delay constraint cannot be satisfied in any case, the sensor node will discard the report, to avoid further wasting of any additional resources.

In this paper, the considered end-to-end QoS constraint is the end-to-end latency requirement D of a report, that may aggregate other data or reports along its path from the source to the collector center. If the report is delivered to the collector center within the given latency constraint D after its initial generation, it will be considered as a successful delivery. At intermediate nodes, the delay from the origination of a report to the intermediately receiving of the report is checked, and the report will be discarded if the delay is larger than the requirement D. The actual procedure of performing this check and making the appropriate decision is an implementation specific issue. For instance, assuming that the nodes can be synchronized, time stamps can be added in the packets and the intermediate nodes can calculate the delay between the current time and the time when the packets are generated, and then compare this delay with the delay constraint to determine whether to discard or forward the packet. Alternatively, a time to live field, with an initial value equal to each packet's delay requirement can be used, which will be reduced appropriately as the packet is forwarded through other nodes towards the collector center.

It should be also noted that at a sensor node, for a received report, in addition to the possible deferred period τ, there is some additional waiting time caused by the transmission of the previous report at the node. The relation between these delays depends actually on the traffic load and system conditions, and is linked to the performance of the data aggregation process. The energy-efficiency that is achieved via aggregation during the deferred periods along the transmission path, is mainly due to the traffic reduction that is achieved by the data aggregation. In some cases under light load, the end-to-end delay may increase due to the introduction of the deferred period τ, since some packets that otherwise could have been transmitted, may have to wait for the aggregation. However, as the traffic load increases, in a system without data aggregation the network becomes congested and the waiting time at each node becomes the dominant factor. Since in principle the waiting time is significantly affected by the network load, performing data aggregation and thus reducing the network traffic load, will result in reduction of the end-to-end delay in the sensor network. In the proposed algorithm, γ and τ are configurable system parameters, and their actual impact on the achievable system performance is analyzed and studied in detail later in this paper.

2.1 Data Aggregation Modeling

By using proper routing mechanisms, we assume that each report goes through each node only once, and nodes always forward the report to other nodes that are closer to the collector center. Therefore, assuming that l nodes are visited from the source node to the collector center, we denote the set of these sensor nodes as $G_l = \{s_1, s_2, \ldots, s_l\}$. Without loss of generality, we assume that the distances between the sensor nodes and the collector site are arranged in decreasing order, i.e., $d_1 > d_2 > \ldots > d_l$, where d_i is the distance between node s_i and the collector center.

Let us also denote by $t_i^{(R)}$ the reporting time at node s_i which includes the time period for data aggregation, by $t_i^{(F)}$

the forwarding time at node s_i which accounts for the transmission time including the potential retransmission time due to channel contention (this time is related to the report length, the bandwidth and the communication success probability), and by $t_i^{(P)}$ the propagation time from node s_i to next node s_{i+1} which depends on the distance between the two nodes. Time periods $t_i^{(R)}$, $t_i^{(F)}$ and $t_i^{(P)}$ are random variables and in the following their corresponding probability density functions (pdf) are denoted by $f_i^{(R)}(t)$, $f_i^{(F)}(t)$, and $f_i^{(P)}(t)$, respectively. Let us denote by t_i the time interval between the point that node s_i receives a report to the point that this report is delivered to node s_{i+1}. If node s_i does not perform data aggregation the corresponding time interval is $t_i^{(F)} + t_i^{(P)}$; otherwise, the time interval is $t_i^{(R)} + t_i^{(F)} + t_i^{(P)}, i \geq 1$. Therefore:

$$t_i = \begin{cases} t_i^{(R)} + t_i^{(F)} + t_i^{(P)}, & \text{with reporting} \\ t_i^{(F)} + t_i^{(P)}, & \text{without reporting,} \end{cases}$$

and its pdf is denoted by $f_i(t)$. Under the assumption that a sensor node performs reporting with probability γ, we have

$$f_i(t) = f_i(t|\text{with reporting})\gamma + f_i(t|\text{without reporting})(1-\gamma).$$

Let us also assume that the time periods are independent of each other, and denote by $F_i^{(R)}(s)$, $F_i^{(F)}(s)$, and $F_i^{(P)}(s)$ the Laplace transforms of $f_i^{(R)}(t)$, $f_i^{(F)}(t)$, and $f_i^{(P)}(t)$, respectively. Applying the Laplace transform to $f_i(t)$, we have

$$\begin{aligned} F_i(s) &= E[e^{-st_i}] = \int_0^\infty f_i(t)e^{-st}dt \\ &= F_i^{(F)}(s) \cdot F_i^{(P)}(s) \cdot \left[\gamma F_i^{(R)}(s) + (1-\gamma)\right]. \end{aligned} \tag{1}$$

In the following, we first assume that no reports will be discarded due to the delay constraint, and obtain the end-to-end delay distribution, which can be used to obtain the probability P_{succ} that the report is delivered to the collector center within the delay constraint D. Then, the probability that the report is discarded due to unsatisfactory end-to-end delay performance can be obtained as $(1 - P_{succ})$. Let us denote the end-to-end delay of a report by $T_L = \sum_{i=1}^L t_i$ and its corresponding pdf by $f_{T_L}(t)$, where the random variable L is the number of hops that are involved in the transmission of a report from the source node to the collector center (including the source node). Thus, the Laplace transform of $f_{T_L}(t)$, denoted by $F_{T_L}(s)$, is given by

$$\begin{aligned} F_{T_L}(s) &= E[e^{-s(t_1+t_2+\ldots+t_L)}] = \sum_{l=1}^\infty E[e^{-sT_L}|L=l]\Pr[L=l] \\ &= \sum_{l=1}^\infty p_L(l)\prod_{i=1}^l F_i(s), \end{aligned} \tag{2}$$

where $p_L(l)$ is the probability mass function of L, where the random variable L represents the number of hops that are involved in the transmission of the report from the source node to the collector center (including the source node). The pdf of T_L can be obtained by using the inverse Laplace transform of $F_{T_L}(s)$, i.e.,

$$f_{T_L}(t) = L^{-1}\{F_{T_L}(s)\} = \sum_{l=1}^\infty p_L(l)L^{-1}\left\{\prod_{i=1}^l F_i(s)\right\}. \tag{3}$$

When $f_{T_L}(t)$ is obtained, the successful probability P_{succ} that a report can reach the collector center within the delay constraint D is given by

$$P_{succ} = P[T_L \leq D] = \int_0^D f_{T_L}(t)dt. \tag{4}$$

When the distribution of $t_i^{(R)}$, $t_i^{(F)}$, and $t_i^{(P)}$ are given, then the $f_{T_L}(t)$ and corresponding P_{succ} can be obtained. However, in practice, it is difficult to obtain an analytical expression for P_{succ}, since the distribution of $t_i^{(F)}$ is generally unknown. In the next section, we use a probabilistic model to develop a lower-bound on the probability P_{succ}.

2.2 Lower Bound on P_{succ}

The end-to-end delay of an independent report that meets the delay constraint and passes through l hops can be represented by

$$T_l = \sum_{i=1}^l t_i^{(R)} + \sum_{i=1}^l t_i^{(F)} + \sum_{i=1}^l t_i^{(P)} \leq D, \tag{5}$$

where, in general, $t_i^{(F)}$ can be upper-bounded based on the largest report length and the corresponding data rate of the sensor network, and $t_i^{(P)}$ can be upper-bounded by the range of the sensor network and the longest distance between two sensor nodes. Thus, we can assume that D is decomposed as $D = D_r(l) + D_f(l) + D_p(l)$, where $D_r(l)$, $D_f(l)$, and $D_p(l)$ are the upper bounds on the end-to-end reporting time, forwarding time, and propagation time, respectively, when the report needs to be delivered to the collector center using l hops. As a result, in our study the constraint that needs to be satisfied regarding the reporting time can be represented as

$$\sum_{i=1}^l t_i^{(R)} \leq D_r(l). \tag{6}$$

Noted that $t_i^{(R)}$ is a function of τ and γ, and (6) provides a worst-case bound on the reporting time under the constraint (5). Therefore, the lower bound on the probability $p_{succ}(l)$ that a specific independent report is delivered to the collector center within the end-to-end constraint, when the distance between the source node and the collector center is l hops, is lower-bounded by

$$p_{succ}(l) = P[S_L(t) < D|L=l] \geq P\left(\sum_{i=1}^l t_i^{(R)} \leq D_r(l)\right). \tag{7}$$

Thus, the probability of success P_{succ} under delay constraint (5) is lower-bounded by

$$P_{succ} = \sum_{l=1}^{L_{\max}} p_L(l)p_{succ}(l) \geq \sum_{l=1}^{L_{\max}} p_L(l)P\left(\sum_{i=1}^l t_i^{(R)} \leq D_r(l)\right), \tag{8}$$

where $L_{\max}$ is the maximum possible number of hops that a packet may go through from the source to the collector. Note that (8) provides a lower bound to the probability of a successful report delivery within the QoS constraint for sensor networks with and without data aggregation schemes. When $\gamma = 0$, P_{succ} is reduced to the probability of a successful delivery in a sensor network without any data aggregation scheme, in which each received report will be forwarded as is (without any deferred period).

Since the QoS routing algorithm deployed in the sensor network is independent of the proposed Q-DAP approach, we assume that if there is no data aggregation scheme deployed in the sensor network, the report can be delivered to the collector center within its end-to-end delay constraint D, through the use of the deployed routing algorithm. Otherwise, the sensor node will not participate in the specific measurement task. That is, we assume $P_{succ,nonaggregation} = 1$. Furthermore, we assume that under the Q-DAP approach, the generated reports will follow the same path as in the case without data aggregation.

In the Q-DAP approach, if data aggregation is not performed at node i, the reporting time $t_i^{(R)} = 0$, while if data aggregation is performed with probability γ, $t_i^{(R)} = \tau$. It is clear that the longest delay that a report may experience due to data aggregation is $l\tau$, when the number of hops between the source to the collector center is l. If $l\tau \leq D_r(l)$, the end-to-end delay can be guaranteed even if at each node data aggregation is performed, i.e., $\gamma = 1$. Thus, we can have $p_{succ}(l) = 1$ when $l\tau \leq D_r(l)$. When $l\tau > D_r(l)$, if all the intermediate nodes perform data aggregation and reporting with a deferred period τ, the end-to-end delay of a report may exceed the delay constraint. The maximum number of data aggregation and reporting that can be performed to guarantee the delay constraint, determined by the upper bound on the reporting time $D_r(l)$, is given by

$$C(l) = \left\lfloor \frac{D_r(l)}{\tau} \right\rfloor. \tag{9}$$

That is, the lower bound of P_{succ} is equal to the probability that a report experiences at most $C(l)$ times of data aggregations and reporting along its path. Therefore, the probability $p_{succ}(l)$ is lower-bounded by

$$p_{succ}(l) \geq p_{succ}^{(LB)}(l) \triangleq \begin{cases} \sum_{k=0}^{C(l)} \binom{l}{k} \gamma^k (1-\gamma)^{l-k}, & C(l) < l \\ 1, & C(l) \geq l. \end{cases} \tag{10}$$

Consequently, the probability of success P_{succ} under delay constraint is lower-bounded by

$$P_{succ} \geq P_{succ}^{(LB)} \triangleq \sum_{l=1}^{L_{max}} p_L(l) p_{succ}^{(LB)}(l). \tag{11}$$

In the following, we discuss and evaluate the relation between this lower bound approximation (11) and the actual value of P_{succ}. For this purpose, we consider a sensor network where the sensor nodes are uniformly distributed in a disk area with radius R, and each node has a fixed limited transmission range r. We assume that each node always transmits a report as far as possible within its transmitting range and, therefore, the maximum number of hops is $M = \lceil \frac{R}{r} \rceil$. Assuming that the packet generation rate is the same for all nodes, the number of packets arriving at

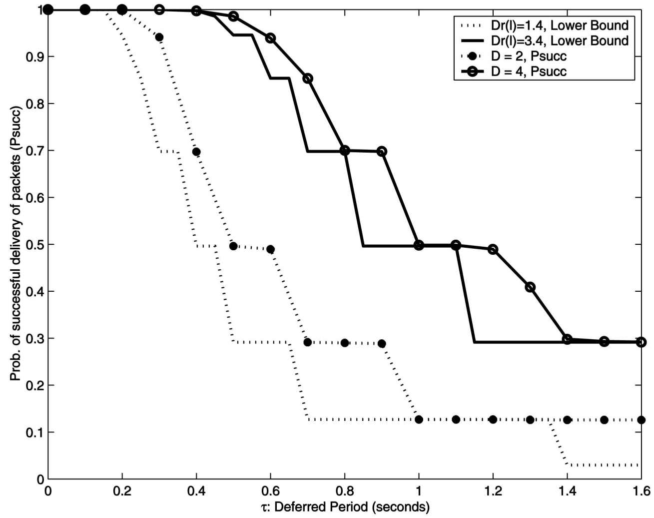


Fig. 1. Lower bound approximation of P_{succ} ($P_{succ}^{(LB)}$) and actual P_{succ} for $\gamma = 0.5$ and different delay constraints D, as a function of deferred period τ.

the collector center through l hops is proportional to the number of sensor nodes within the ring area l hops away from the collector center. So, if the total number of sensor nodes is large, the probability $p_L(l)$ can be approximated by

$$p_L(l) = \begin{cases} \frac{2l-1}{M^2}, & 1 \leq l \leq M \\ 0, & l > M. \end{cases} \tag{12}$$

Let us also assume that the report arrival process at each sensor node follows Poisson distribution with rate λ. In this case, in Fig. 1, the curves of the corresponding probabilities are plotted as functions of the deferred period τ, for a sensor network with $M = 10$, $\lambda = 20$, and $\gamma = 0.5$. In this figure, the lower bound approximation $P_{succ}^{(LB)}$ is obtained by using (10) and (11), and the actual P_{succ} is obtained analytically using (4) under the assumption that the report arrival process is Poisson, while $t_i^{(F)}$ and $t_i^{(P)}$ are exponentially distributed. In this figure, we plot four different curves, which represent the corresponding probabilities for delay constraints $D = 4$ and $D = 2$ seconds, respectively. Correspondingly, in the lower bound calculation, we assume that $D_r(l) = 3.4$ and $D_r(l) = 1.4$.[1] The results in Fig. 1 demonstrate that, in general, a smaller D will result in a lower P_{succ}, while the lower bound approximation demonstrates similar trend with the actual performance of P_{succ}. Based on these results, we conclude that $P_{succ}^{(LB)}$ provides an accurate lower bound approximation of the probability of successful report delivery for all values of τ.

3 QoS-Constrained Data Collection and Aggregation at the End Nodes

In this section, taking into account that each sensor operates autonomously and without any central control, we introduce an adaptive localized algorithm for the data collection

1. These values are obtained based on the fact that in the analytical calculation of P_{succ}, almost all packets spend less than 0.6 time units for the transmission and propagation through their delivery from the source to the collector center. Thus, in the corresponding lower bound approximation, we assume that $D_r(l) = 3.4$ and $D_r(l) = 1.4$, respectively.

and aggregation at the end nodes, with the objective of balancing the tradeoffs among energy-efficiency, delay requirement, accuracy, and buffer overflow probability.

The procedure adopted here in order to collect the appropriate samples at the end nodes is described as follows: In a sensor node s_i, one sample is collected every $T_{\triangle,i}$ unit times. However, if at some point the change of the sensed signal is beyond a predefined threshold $\theta_{\triangle,i}$, a sample is also collected independent of the time. We also assume that a sensor node will collect and save a total number of $N_{s,i}$ samples before it originates a packet transmission to disseminate this information to the appropriate destination. Based on this data collection procedure, the measurement quality or accuracy is determined by parameters $\theta_{\triangle,i}$ and $T_{\triangle,i}$. It should be noted that for energy efficiency purposes multiple samples are collected and aggregated together in a single packet.

Intuitively, we can argue that there is a tradeoff among the various parameters and performance metrics involved in this scenario, such as: $N_{s,i}$, $T_{\triangle,i}$, $\theta_{\triangle,i}$, delay, energy efficiency, and node buffer size. Specifically, as $N_{s,i}$ or $T_{\triangle,i}$ increase, the energy efficiency of the data collection and transmission increases, while the corresponding delay and buffer size requirements at each sensor increase as well. On the other hand, the accuracy of the collected data increases as $\theta_{\triangle,i}$ and $T_{\triangle,i}$ decrease. Therefore, an adaptive algorithm that realizes the data collection process by taking into account the system conditions and the task related quality of service requirements (in terms of accuracy, delay, etc.) can be summarized as follows: The initial values of parameters $N_{s,i}$, $T_{\triangle,i}$, and $\theta_{\triangle,i}$ are first determined according to the delay requirements and the desired accuracy. Then, based on the communication conditions, as they are expressed and represented by some local measurements (e.g., observed data departure rate $\mu_{e,i}$), these parameters may be adjusted according to some desired objectives. One such objective is to adjust the parameters so that the expected probability of buffer overflow is lower than some prespecified overflow threshold $P_{of,th}$.

3.1 Localized Adaptive Data Collection and Aggregation Approach (LADCA) Process

Let us denote by $t_{s,i}$ the time interval from the point that a sample is collected till the point that this sample is successfully transmitted out of the sensor node s_i (i.e., transmitted to the intermediate node). Since the distance between nodes in a sensor network is usually limited, the propagation delay is considered negligible. If $D_{q,i}$ denotes the corresponding average delay requirement, then we need: $E(t_{s,i}) \le D_{q,i}$. Furthermore, let us denote by $\theta_{q,i}$ and $T_{q,i}$ the desired accuracy. That is: $\theta_{\triangle,i} \le \theta_{q,i}$ and $T_{\triangle,i} \le T_{q,i}$. Let the average overhead of data packets be L_{ov} bits and the size of a sample be b bits. Then, the energy-efficiency coefficient $\beta_{en,i}$ can be defined as the following ratio: $\beta_{en} = \frac{bN_{s,i}}{bN_{s,i}+L_{ov}}$. The buffer size for storing the collected data is assumed to be B_{sz} and $bN_{s,i} \le B_{sz}/2$. Based on these notations and definitions, the cost of a data collection policy $C_{\pi(N_{s,i},T_{\triangle,i},\theta_{\triangle,i})}$ is defined as:

$$C_{\pi(N_{s,i},T_{\triangle,i},\theta_{\triangle,i})} = w_1 \frac{E(t_{s,i})}{D_{q,i}} + w_2 \frac{\theta_{\triangle,i}}{\theta_{q,i}} + w_3 \frac{T_{\triangle,i}}{T_{q,i}} + w_4 \frac{1-\beta_{en}}{\beta_{en}}, \tag{13}$$

where $w_i,\ i = 1,..,4$ are the weights of delay, accuracy (measured by $\theta_\triangle$ and $T_\triangle$) and energy-efficiency. The appropriate weights can be selected according to the different application requirements and the corresponding system considerations, resources, and priorities. The initial values of parameters $N_s, T_\triangle$, and $\theta_\triangle$ are selected to minimize the above weighted cost, as follows:

$$(N_{s_{0,i}}, T_{\triangle_{0,i}}, \theta_{\triangle_{0,i}}) = \arg\min C_{\pi(N_{s,i},T_{\triangle,i},\theta_{\triangle,i})}, \tag{14}$$

$$E(T_{s,i}) \le D_{q,i}, N_{s,i} \le B_{sz}/2b, \tag{15}$$

$$\theta_{\triangle,i} \le \theta_{q,i}, T_\triangle \le T_{q,i}, \tag{16}$$

where $\arg\min C_{\pi(N_{s,i},T_{\triangle,i},\theta_{\triangle,i})}$ refers to the arguments $N_{s,i}, T_{\triangle,i}, \theta_{\triangle,i}$ that minimize function $C_{\pi(N_{s,i},T_{\triangle,i},\theta_{\triangle,i})}$.

For simplicity in the representation and notation, in the following, we ignore the index i of the corresponding parameters. Let t_{sc} be the time interval required to collect N_s samples in order to generate a packet, and t_{st} be the time interval from the point that a packet is ready for transmission till the point that is actually transmitted successfully. t_{st} includes the possible queueing delay as well as the transmission time. Then, we have:

$$E(t_s) = E(t_{sc}) + E(t_{st}). \tag{17}$$

The sample arrival process consists of two components: a periodic arrival process with rate $\frac{1}{T_\triangle}$, and a nondeterministic arrival process which depends on threshold $\theta_\triangle$. In the following, we assume that the latter follows a Poisson distribution with rate $\lambda_s(\theta_\triangle)$. Let us also denote by Y the random variable of the interarrival time between two samples. Then, it can be shown that:

$$E(t_{sc}) = \frac{N_s - 1}{2} E(Y) = \frac{(N_s - 1)T_\triangle}{2(1 + T_\triangle \lambda_s)}. \tag{18}$$

In order to obtain $E[t_{st}]$, we need to calculate the corresponding queueing delay. In the following, for demonstration purposes, we first assume that the probability of buffer overflow is very small and therefore the system can be treated as a system with infinite buffer. Later on, the buffer size is taken into consideration as well. Since there are two patterns of sample arrivals, one periodic with rate $\frac{1}{T_\triangle}$ and one Poisson with rate λ_s, the system can be viewed as a combination of an $D/G/1$ and an $M/G/1$ system. Thus, the average queueing delay $E(W)$ is given by: $E(W) = p_s E(W_D) + (1 - p_s)E(W_M)$, where $p_s = \frac{1}{1+\lambda_s T_\triangle}$, and W_D and W_M are the corresponding queueing delays of the $D/G/1$ and $M/G/1$ systems, respectively. The service time depends on the transmission rate and the probability of collisions, while the queueing delay can be obtained when the average service time and the second moment of service time are given. Under the assumption that at the beginning no collision occurs and the service time depends only on the transmission rate, denoted by R (bits per unit time), the average data departure rate μ is constant for given N_s, L_{ov}, and b:

$$\mu = \frac{N_s R}{bN_s + L_{ov}} \text{ samples/unit time.} \tag{19}$$

Then, the average queueing time is easily obtained as: $E(W) = \frac{(1-p_s)\lambda_s}{2\mu(\mu-\lambda_s)}$. Then, we obtain:

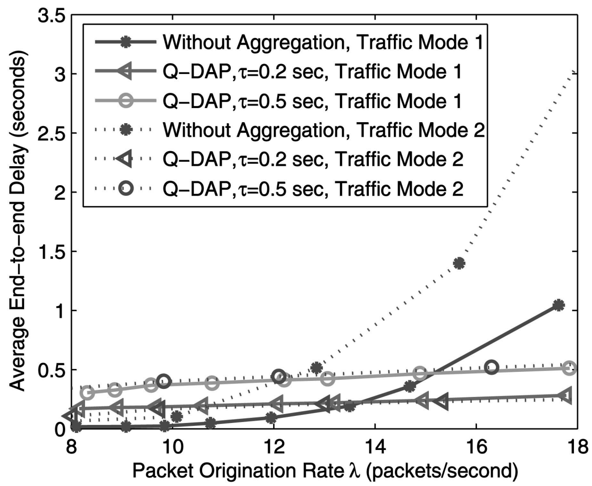


Fig. 2. Cost as a function of N_s for different $T_\triangle$ and $\theta_\triangle$ values.

$$E(t_{st}) = E(W) + \frac{1}{\mu} = \frac{2\mu - \lambda_s(1+p_s)}{2\mu(\mu - \lambda_s)}. \quad (20)$$

Furthermore, by substituting (18) and (20) into (17), we can obtain:

$$E(t_s) = \frac{T_\triangle}{2(1+T_\triangle\lambda_s)}\left[N_s - 1 + \frac{\lambda_s^2}{\mu(\mu-\lambda_s)}\right] + \frac{1}{\mu}. \quad (21)$$

From (19), we have that $\lim_{N\longrightarrow\infty}\mu = \frac{R}{b}$ and $\mu < \frac{R}{b}$ and, thus,

$$\frac{T_\triangle}{2(1+T_\triangle\lambda_s)}\left[N_s - 1 + \frac{\lambda_s^2}{\frac{R}{b}\left(\frac{R}{b}-\lambda_s\right)}\right] + \frac{1}{\frac{R}{b}} < E(T) \le D_q.$$

Therefore, an upper bound of N_s is obtained as follows:

$$N_s < 2\left(\frac{1}{T_\triangle} + \lambda_s\right)\left(D_q - \frac{b}{R}\right) - \frac{\lambda_s^2 b^2}{R(R-\lambda_s b)}. \quad (22)$$

It can be shown that $\lambda_s \propto \frac{1}{\theta_\triangle}$. Therefore, we assume that $\lambda_s = \frac{\alpha}{\theta_\triangle}, \alpha > 0$. For a stable system, $\mu > \lambda_s + \frac{1}{T_\triangle}$ is required and, therefore, a lower bound of N_s can be obtained as:

$$N_s > \frac{L_{ov}}{R/\left(\frac{\alpha}{\theta_\triangle} + \frac{1}{T_\triangle}\right) - b}, \quad (23)$$

where $\left(\frac{\alpha}{\theta_\triangle} + \frac{1}{T_\triangle}\right) < \frac{R}{b}$. Thus, the selection of N_s has to satisfy (15), (22), and (23). In this case, (13) can be expressed as

$$C_{\pi(N_s,T_\triangle,\theta_\triangle)} = \frac{w_1}{D_q}\left[\frac{T_\triangle}{2(1+T_\triangle\lambda_s)}\left(N_s - 1 + \frac{\lambda_s^2}{\mu(\mu-\lambda_s)}\right) + \frac{1}{\mu}\right] + w_2\frac{\theta_\triangle}{\theta_q} + w_3\frac{T_\triangle}{T_q} + w_4\frac{L_{ov}}{bN_s}. \quad (24)$$

In Fig. 2, the data collection process cost is depicted as a function of N_s for different values of $T_\triangle$ and $\theta_\triangle$. The corresponding tradeoffs among the various parameters involved in the overall process can also be seen and evaluated by this figure. For a given set of $T_\triangle$ and $\theta_\triangle$, the optimal value of N_s can be identified.

When the initial values are determined as explained above, the sensing nodes will collect data using these parameters. The actual data departure rate equals to the data arrival rate for a system with infinite buffer size. For a system with finite buffer size, the actual data departure rate is lower than the data arrival rate since buffer overflow may occur. When the data departure rate decreases during a period due to retransmission or collisions (when the network load increases or the channel conditions deteriorate), the probability of buffer overflow will increase. The tradeoff that arises here is that we can lower the buffer overflow probability by decreasing the provided accuracy of the sensed variables. By letting μ_e be the actual departure rate, the probability of overflow is:

$$P_{of} = 1 - \frac{\mu_e}{\frac{\alpha}{\theta_\triangle} + \frac{1}{T_\triangle}}.$$

If we denote by $P_{of,th}$ the buffer overflow threshold, i.e., $P_{of} \le P_{of,th}$, then an adaptive data collection algorithm at the end nodes, based on local only information, can be described as follows:

- A sensor node periodically checks its current P_{of}.
- If $P_{of} > P_{of,th}$, increase $\theta_\triangle$ and $T_\triangle$ so that $P_{of} = P_{of,th}$ and $\min C_{\pi(N_s,T_\triangle,\theta_\triangle)}$.
- If $P_{of} < \eta_1 P_{of,th}$ and the available buffer size is larger than a given threshold, decrease $\theta_\triangle$ and $T_\triangle$ so that $P_{of} = \eta_2 P_{of,th}$, $0 < \eta_1 < \eta_2 \le 1$ and $\min C_{\pi(N_s,T_\triangle,\theta_\triangle)}$.

4 Performance Evaluation

In this section, the overall performance evaluation of the proposed quality-driven data aggregation approach in multihop sensor networks is accomplished via modeling and simulation using the Optimized Network Engineering Tool (OPNET). Simulation is used in order to evaluate the actual performance of the proposed approach in more realistic environments, where different data generation processes are considered, and a realistic media access control (MAC) protocol based on IEEE 802.11 can be taken into account, including the effect of collisions and retransmissions. First, the achievable performance in terms of the end-to-end delay and overall network energy savings is evaluated, under different data aggregation scenarios and traffic loads. Then, the impact of several design parameters and tradeoffs on various critical network and application related performance metrics, such as the energy efficiency, network lifetime, end-to-end latency, are also evaluated and discussed. Finally, the impact of LADCA algorithm on the data loss due to the buffer overflow at the end nodes is evaluated as well.

It should be noted that the data aggregation modeling approach provided in Section 2.1 is quite general, however, in practice, it is difficult to obtain an analytical expression for P_{succ}, unless the corresponding distributions are known. This is also the reason for obtaining a lower bound approximation for the successful probability in Section 2.2. The accuracy of the lower bound approximation was evaluated in Section 2.2. These results are used in order to estimate P_{succ} and accordingly obtain appropriate values for parameters τ and γ, for the experiments performed in Section 4.2. Similarly, the analytical results from Section 3 (e.g., (24)) are used in order to obtain appropriate initial values of $\theta_\triangle$ and $T_\triangle$, for the experiments performed in Section 4.6.

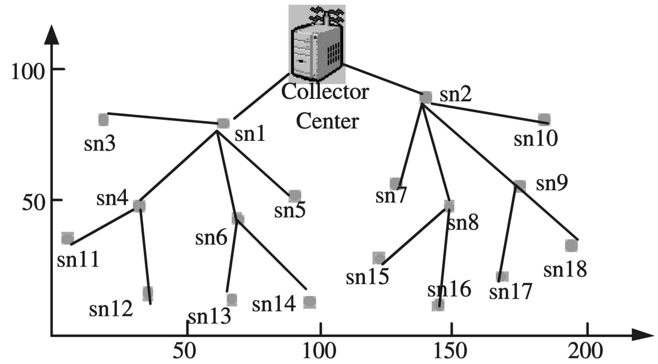


Fig. 3. A reference multihop sensor network.

4.1 Assumptions and Network Reference Topology

Throughout this study a sensor network consisting of 18 nodes and one collector center, distributed in an $100m \times 200m$ area as shown in Fig. 3, is considered. In order to better focus on the study of the impact of the Q-DAP approach on the end-to-end delay and the network energy consumption, we assume that the routing paths are predetermined during the whole network operation. The corresponding routes from the individual sensor nodes toward the collector center are identified by the edges between the various nodes as shown in Fig. 3. The transmission range of each node is assumed to be 50 meters, while the simulation for each scenario lasts for 1,800 seconds. Each simulation scenario is repeated 10 independent times (i.e., each run starts with a different random number seed) and statistical averages are calculated.

When a node begins to transmit, all the neighbors within its transmission range will receive the signal, which is considered as interference for a node if the packet is not destined for it. The MAC protocol adapted here is CSMA/CA and its implementation is based on IEEE 802.11 standard. Rts/Cts messages are exchanged before a data packet is transmitted if the length of the data packets is more than 64 bytes, otherwise the data packet is transmitted without Rts/Cts exchange. The corresponding power consumption of a node under idle/listen, receiving, and transmitting modes is assumed to be 1mW, 10mW, and 36mW, respectively [17]. Furthermore, we assume that the data transmission rate is 1Mbps. Let us also denote by β the aggregation coefficient, which represents the ratio of the new report length after aggregation and reporting, to the total length of all the received packets/reports before aggregation. For instance, $\beta = \frac{\text{length of packet after aggregation}}{\text{total length of original packets}}$ and $0 < \beta \leq 1$. The degree of data aggregation and its relation to the data aggregation coefficient is closely related to the corresponding savings that can be achieved due to both the MAC overhead reduction and the corresponding payload savings of packets with correlated data. It should be noted here that a frame aggregation scheme is one of the possible components in the future 802.11n MAC, where a transmitting station may concatenate multiple packets into a single frame thus reducing the corresponding overheads.

4.2 End-to-End Delay

In this section, we compare the end-to-end delay of the sensor network under the Q-DAP approach with the corresponding results obtained by a system that does not perform any data aggregation. In the following, for demonstration purposes, we assume that the packet length is exponentially distributed with mean 100 bytes, and the aggregation coefficient is considered to be 0.9. In order to compare the achievable delays under different scenarios, we first set the delay constraint D to a very large number, so that there are no packets discarded at the intermediate nodes due to the delay constraint. The corresponding average end-to-end delays, for two different data generation processes at each node, are shown in Fig. 4. Specifically, in Fig. 4a the data generation at each node follows a Poisson process with rate λ, while in Fig. 4b, the data generation follows an ON-OFF bursty process where packets are only

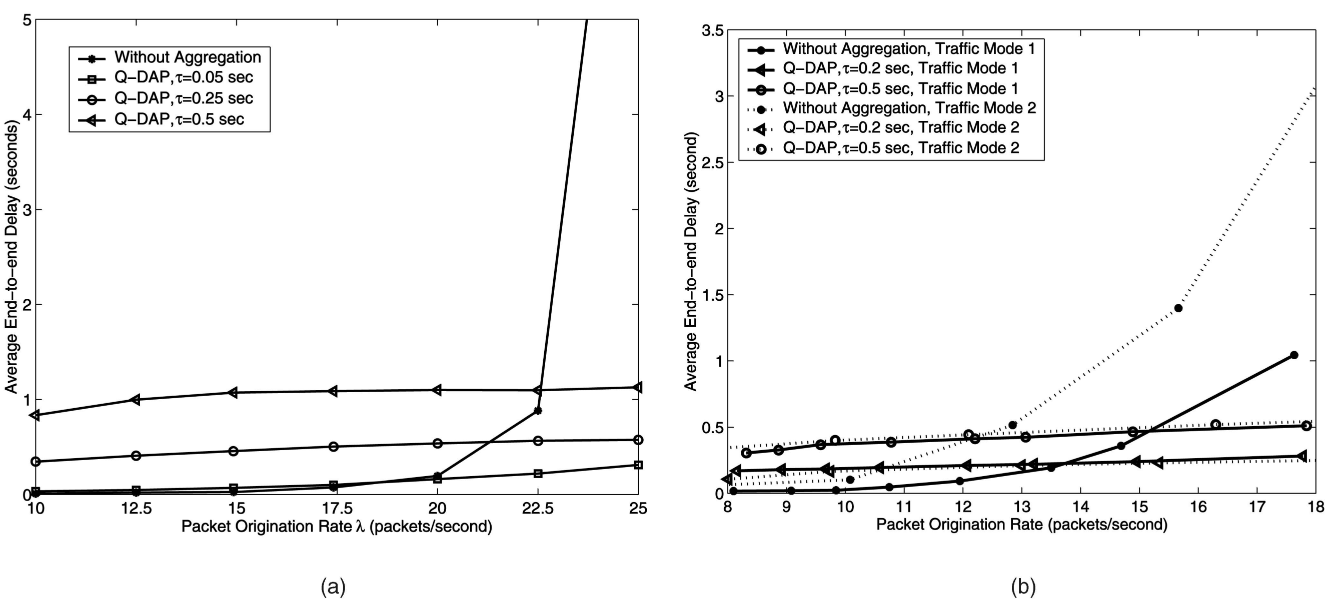
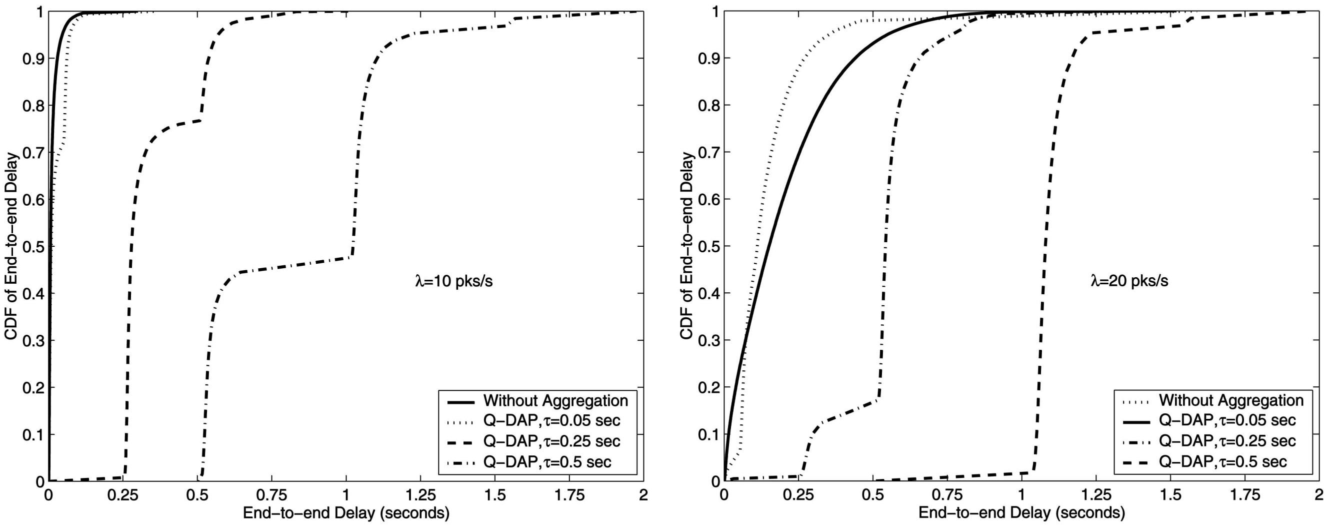


Fig. 4. Average end-to-end delay as a function of λ. (a) Poisson packet arrival ($\gamma = 1$). (b) Burst packet arrival ($\gamma = 0.9$).



(a) (b)

Fig. 5. Cumulative Distribution Function (CDF) of the end-to-end delay. (a) $\lambda = 10$ packets/second. (b) $\lambda = 20$ packets/second.

generated while the process is in the ON state. For the ON-OFF case we consider two different traffic modes. For traffic mode 1, the duration for which the process stays in the ON and OFF states follows exponential distribution with means 2 and 8, respectively, while for traffic mode 2, the duration for ON state is uniformly distributed between 1 and 50 and the duration for OFF state is exponentially distributed with mean 50 seconds. It can be seen from these figures that without data aggregation, the delay increases exponentially with the increase of the network load (indicated by λ), while under the Q-DAP approach, the delay increases at a much slower rate, since performing data aggregation reduces the network traffic load significantly. When the network load is light, the delay introduced by the Q-DAP strategy is the dominant factor, due to the fact that the sensor node introduces a deferred period of τ to perform the data aggregation, while the corresponding waiting time at each node is negligible. Therefore, in this case, the delay in the sensor network under the Q-DAP approach is larger than the one that can be achieved by a system without aggregation. However, when the network load increases, the waiting time at each node becomes the dominant factor (as compared to τ). Therefore, since the waiting time is significantly affected by the network load, performing data aggregation can reduce the network traffic load and, therefore, result in the reduction of the end-to-end delay in the sensor network. Therefore, as we can observe from Fig. 4, for heavy traffic loads the average end-to-end delay under the Q-DAP is significantly lower than the corresponding delay of a system without any data aggregation.

In Fig. 5, the corresponding cumulative distribution functions (CDF) of the end-to-end delay are shown for $\lambda = 10$ pks/s and $\lambda = 20$ pks/s. This can be used to estimate the successful report delivery for a system with the delay constraint comparable to the end-to-end delay. For instance, based on this, we can choose a delay constraint of $D = 0.6$ seconds for the system with deferred period $\tau = 0.25$ seconds, and a delay constraint of $D = 1.1$ seconds for $\tau = 0.5$ seconds, and then perform experiments in order to obtain the probability of successful packet delivery and actual packet dropping probability P_{drop}, due to the imposed delay constraint.

The corresponding results are shown in Fig. 6. For comparison purposes only, we also present the probabilities that the packets arrive at the collector center within a certain end-to-end delay equivalent to the corresponding delay constraints imposed by the Q-DAP, under a strategy that performs data aggregation (similar to Q-DAP) without, however, discarding packets at the intermediate nodes due to any delay constraints (in the following graph, we refer to these cases as no-packet-drop). As we expected, the successful packet delivery probability of the system with delay constraint $D = 0.6$ and $D = 1.1$ seconds is better than the estimated probability under the strategy that does not discard any packets due to the delay constraints. This happens because the packets that cannot satisfy the imposed delay constraint have been discarded at the intermediate nodes and, therefore, the overall traffic has

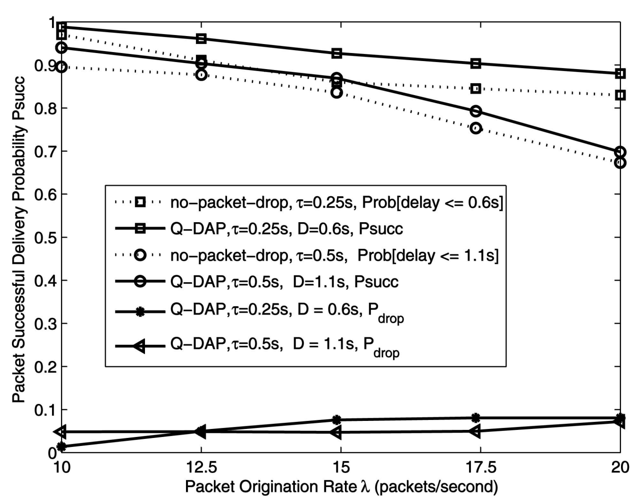


Fig. 6. Probability of successful packet delivery for different delay constraints.

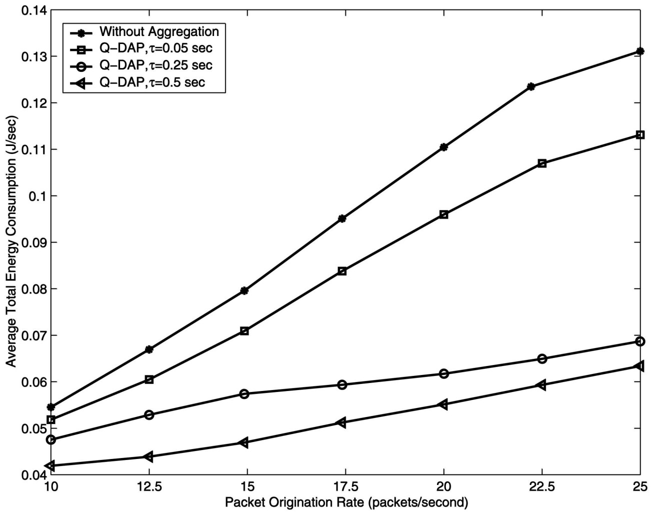


Fig. 7. Average total energy consumption under various traffic conditions ($\gamma = 1$).

been reduced. Furthermore, as can be observed by this figure, the successful packet delivery probability increases as τ decreases, however, as we see later this happens at the cost of higher energy consumption.

4.3 Energy Efficiency

Since the energy consumption for communications is usually considered as the dominant factor compared to that for data processing [18], the proposed Q-DAP approach will result in lower energy consumption and thus extend the lifetime of the sensor network, due to the resulting reduced communication traffic that is achieved by the data aggregation. Throughout this experiment, the energy consumption for the local data processing and aggregation is set to 0.1 nJ/bit. Fig. 7 depicts the total energy consumption in the sensor network under four different scenarios. The first one corresponds to the system where no aggregation is performed, while the other three scenarios correspond to implementations of the Q-DAP approach with different deferred period τ. As can be seen from this figure, the Q-DAP approach outperforms the system without data aggregation, and achieves significant energy savings in the sensor network. Furthermore, it can be also observed from this figure that the energy consumption decreases as τ increases, since when τ increases the average number of packets that can be used for data aggregation increases as well (even for the same traffic load λ).

4.4 Critical Nodes and Network Lifetime

In most of the cases, the operation of the sensor network is completely disrupted, if and only if all the nodes that can directly communicate with the collector center (e.g., one-hop communication from the collector center) "expire," and as a result the lifetime of these nodes is more critical to the network lifetime [2]. In the following, we refer to these nodes as critical nodes. Here, we define the network lifetime as the time interval from the point that the sensor network starts its operation until the point where loss of communication to the collector site by all sensor nodes occurs. With reference to the network topology of Fig. 3, only nodes $sn1$ and $sn2$ can communicate directly with the collector center and, therefore, they are the critical nodes.

TABLE 1
Network Lifetime for Different Values of τ

τ (sec.)	Lifetime		average delay(sec.)	
	$\lambda = 10$	$\lambda = 20$	$\lambda = 10$	$\lambda = 20$
0	1	0.470	0.009	0.195
0.05	1.097	0.473	0.033	0.196
0.25	1.260	0.922	0.346	0.538
0.5	1.499	1.157	0.834	1.099
1	1.880	1.221	2.093	2.201
2	1.975	1.247	4.296	4.325

In order to study the impact of the deferred time τ on the network lifetime, we performed several experiments which correspond to different values of parameter τ, as shown in Table 1. Case $\tau = 0$ represents the system without any data aggregation. Specifically, Table 1 presents the network lifetime (normalized by the lifetime of a system without any data aggregation for $\lambda = 10$ packets/second) and the corresponding average end-to-end delays, under two different traffic loads ($\lambda = 10$ packets/second and $\lambda = 20$ packets/second), for different configurations of parameter τ. From the results presented in this table, we observe that the network lifetime increases as the deferred period τ increases. This happens because the average number of packets that can be used to perform data aggregation increases with τ as well, therefore resulting in reduced communication traffic. We also notice that there exists some value of τ above which the network lifetime increases very slowly as τ increases. For the cases under consideration here, this value is about $\tau = 1$ second for $\lambda = 10$ packets/second and $\tau = 0.5$ second for $\lambda = 20$ packets/second. Furthermore, it can be seen that the average end-to-end delay increases significantly with the increase of τ, and as can be concluded from the results that were presented in Fig. 6, the larger the parameter τ, the higher the probability that a packet may not be delivered within the delay constraint. Therefore, large values of τ will mainly benefit those tasks with loose delay constraints, while the proper value of τ should be identified so that the lifetime of a network can be extended and most of the packets will be delivered to the collector center within the imposed delay constraint.

The network performance may be even improved by allowing different nodes to have different deferred periods. For example, for $\lambda = 20$ packets/second, if we let τ of nodes $sn1$, $sn2$, $sn4$, $sn6$, $sn8$, and $sn9$ be 0.3 seconds and τ of the rest of the nodes be 0.15 seconds, the resulting network lifetime is 0.939, which is longer than the lifetime (0.922) of a system with $\tau = 0.25$ for all the nodes, while at the same time it achieves better successful packet delivery probability ($P_{succ} = 0.94$) than the corresponding $P_{succ} = 0.87$ of the system with the same $\tau = 0.25$.

4.5 The Impact of γ

In Fig. 8, the energy consumption and average end-to-end delay as a function of parameter γ are shown, for $\lambda = 10$ packets/second, $\beta = 0.9$ and $\tau = 0.25$, 0.5 and 1 second, respectively. The results for $\gamma = 0$ correspond to the case that no data aggregation is performed. It can be seen from this figure that as γ increases, the system consumes less energy during the same operation period, while at the same time the average delay increases. So,

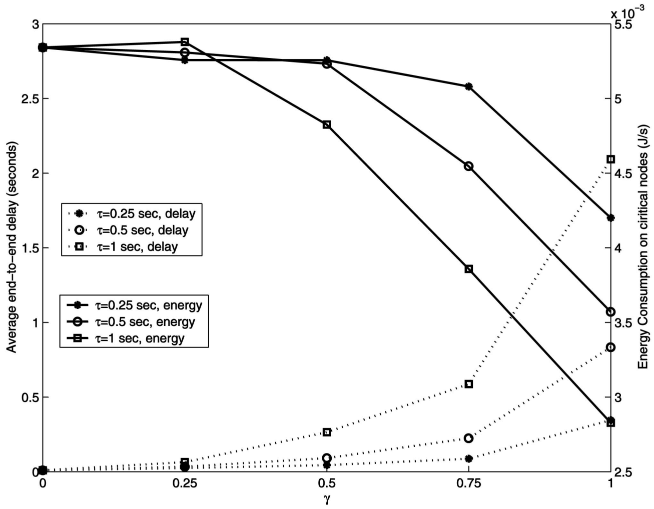


Fig. 8. Average end-to-end delay and energy consumption for different values of τ.

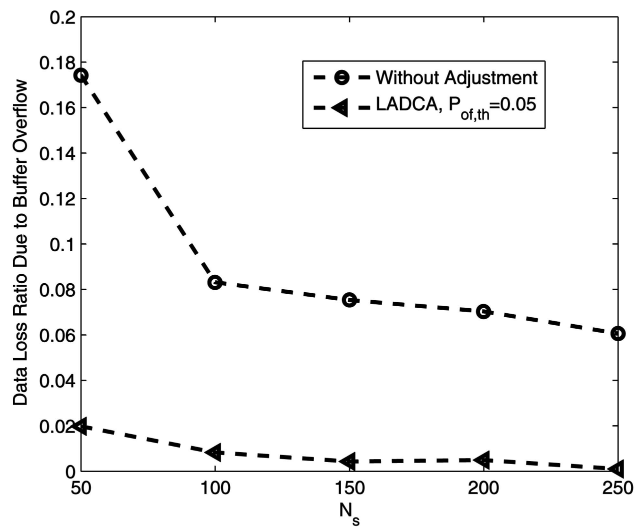


Fig. 9. Data loss ratio at the end nodes as a function of the number N_s of the aggregated samples.

there is also a tradeoff between energy consumption and end-to-end delay. Therefore, the γ provides another adjustable factor for the appropriate design according to the system design requirements and available resources.

4.6 Buffer Overflow and Data Loss at End Nodes

In this section, the impact of LADCA on the data loss due to the buffer overflow at the end nodes is evaluated and discussed. For demonstration purposes, in the following experiment, the sample collection accuracy requirements are set to: $T_q = 0.005$, $\theta_q = 0.01$, 16 bits per sample, $\alpha = 10$, $D_q = 0.5$ second, and $P_{of,th} = 0.05$. The initial accuracy related system parameters are selected as $T_\triangle = 0.0005$ and $\theta_\triangle = 0.001$. Fig. 9 presents the average data loss ratio due to buffer overflow, as a function of the total number of samples N_s that are collected before a packet is generated at the source (end-node), under the proposed LADCA algorithm.

It should be noted that, as described in Section 3.1, LADCA algorithm provides an adaptive data collection method at the end nodes, based on the available local information only, which aims at adjusting the measurement accuracy related system parameters (i.e., $\theta_\triangle$ and $T_\triangle$) in order to allow the system to adjust to the changing conditions and minimize the cost of the associated data collection policy. For comparison purposes, the corresponding data loss ratio for a strategy without such adaptation capabilities is also depicted (we refer to this strategy as "without adjustment" strategy). It can be seen by this figure that through the adaptive method introduced by the LADCA approach, the buffer overflow is well controlled and the corresponding data loss due to buffer overflow decreases significantly. Therefore, when the sensor network traffic and conditions change (i.e., the network load increases or the channel conditions deteriorate), each end node via the proposed localized adaptive collection approach will attempt to readjust the corresponding measurement related parameters based on the interpretation of local information (e.g., P_{of} and μ_e), in order to balance the tradeoffs between delay and accuracy and, thus, decrease/mimimize the actual data loss.

Furthermore, as expected, we observe from this figure that the data loss decreases when the value of N_s increases both under the LADCA strategy and the "without adjustment" strategy. This happens because as the number of collected samples that are aggregated in a single packet for transmission increases, the traffic load injected in the network by each end node as well as the corresponding communication overhead decrease. Therefore, when the network load is heavy the congestion can be reduced and the achievable throughput may improve. The actual rate and values of the corresponding savings depend on the specific data aggregation and compression techniques at the source nodes.

5 Concluding Remarks

This paper introduced and analyzed an efficient QoS-constrained data aggregation and processing approach for distributed wireless sensor networks. The proposed approach consists of: 1) a QoS-driven data aggregation algorithm (Q-DAP) that aggregates data on the fly at the intermediate nodes in a distributed fashion, therefore reducing the traffic load and the consumed communication energy while at the same time satisfying the latency and measurement quality constraints; and 2) an adaptive localized algorithm (LADCA) for the data collection and aggregation at the end nodes, that balances the design tradeoffs of delay, measurement accuracy, and buffer overflow, and provides a method of adjusting the measurement accuracy related parameters at the source nodes.

An in-depth evaluation of the proposed approach, under different data aggregation scenarios and traffic loads, was performed via modeling and simulation, and the corresponding numerical results demonstrated the significant performance improvements that can be achieved, in terms of several critical operational metrics, such as energy efficiency, improved network lifetime, reduced traffic load, end-to-end delay, etc. In conclusion, given some specific QoS requirements imposed by the task/application under consideration, we can use the proposed approach and accordingly adjust the design parameters τ and γ at the intermediate nodes, as well as the measurement accuracy related parameters at the end nodes, in order to fulfill the required QoS, while at the same time achieve significant energy savings and extend the sensor network operational lifetime.

It should be noted that the proposed approach is evaluated here for a fixed sensor network. However, since this is an adaptive QoS-oriented data aggregation method, it is expected that combined with the appropriate routing mechanism, it would be ideal for deployment in sensor networks with dynamic configuration. Extending it to support dynamic and mobile environments, by allowing the dynamic adjustment of several operational parameters such as the deferred period and the aggregation probability, is part of our current research. Furthermore, additional energy efficiency may also be achieved by considering multiple node energy states based on the relaxation phenomena of the batteries and other possible battery renewal modes (e.g., solar batteries).

Finally, the relationship among the aggregation coefficient, data correlation, and packet concatenation, as well as its corresponding impact on the performance of the proposed approach needs to be further investigated and evaluated. The optimization of the MAC overhead, that depends on several related timers and thresholds, based on the aggregation coefficient is a very interesting issue of high research and practical importance.

REFERENCES

[1] C.-C. Shen, C. Srisathapornphat, and C. Jaikaeo, "Sensor Information Networking Architecture and Applications," *IEEE Personal Comm.*, vol. 8, no. 4, pp. 52-59, 2001.

[2] J. Zhu and S. Papavassiliou, "On the Connectivity Modeling and the Tradeoffs between Reliability and Energy Efficiency in Large Scale Wireless Sensor Networks," *Proc. IEEE Wireless Comm. and Networking Conf.*, vol. 2, pp. 1260-1265, Mar. 2003.

[3] J. Chou, D. Petrovic, and K. Ramchandran, "A Distributed and Adaptive Signal Processing Approach to Reducing Energy Consumption in Sensor Networks," *Proc. IEEE INFOCOM 2003*, vol. 2, pp. 1054-1062, 2003.

[4] S. Pattem, B. Krishnamachari, and R. Govindan, "The Impact of Spatial Correlation on Routing with Compression in Wireless Sensor Networks," *Proc. Third Int'l Symp. Information Processing in Sensor Networks (IPSN)*, pp. 28-35, Apr. 2004.

[5] S.J. Baek, G. de Veciana, and X. Su, "Minimizing Energy Consumption in Large-Scale Sensor Networks through Distributed Data Compression and Hierarchical Aggregation," *IEEE J. Selected Areas in Comm.*, vol. 22, no. 6, pp. 1130-1140, Aug. 2004.

[6] S. Papavassiliou and J. Zhu, "Architecture and Modeling of Dynamic Wireless Sensor Networks," *Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems*, pp. 1-16, July 2004.

[7] V.P. Mhatre, C. Rosenberg, D. Kofman, R. Mazumdar, and N. Shroff, "A Minimum Cost Heterogeneous Sensor Network with a Lifetime Constraint," *IEEE Trans. Mobile Computing*, vol. 4, no. 1, pp. 4-15, Jan.-Feb. 2005.

[8] J. Carle and D. Simplot-Ryl, "Energy-Efficient Area Monitoring for Sensor Networks," *Computer*, vol. 37, no. 2, pp. 40-46, Feb. 2004.

[9] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed Diffusion for Wireless Sensor Networking," *IEEE/ACM Trans. Networking*, vol. 11, no. 1, pp. 2-16, 2003.

[10] S. Olariu, A. Wada, L. Wilson, and M. Eltoweissy, "Wireless Sensor Networks: Leveraging the Virtual Infrastructure," *IEEE Network*, vol. 18, no. 4, pp. 51-56, July/Aug. 2004.

[11] A. Boulis, S. Ganeriwal, and M.B. Srivastava, "Aggregation in Sensor Networks: An Energy-Accuracy Trade-Off," *Proc. 2003 IEEE Int'l Workshop Sensor Network Protocols and Applications*, pp. 128-138, May 2003.

[12] B. Krishnamachari, D. Estrin, and S. Wicker, "The Impact of Data Aggregation in Wireless Sensor Networks," *Proc. IEEE 22nd Int'l Conf. Distributed Computing Systems Workshop*, pp. 575-578, July 2002.

[13] I. Stojmenovic and X. Lin, "Power-Aware Localized Routing in Wireless Networks," *IEEE Trans. Parallel and Distributed Systems*, vol. 12, no. 11, pp. 1122-1133, Nov. 2001.

[14] K. Akkaya and M. Younis, "An Energy-Aware QoS Routing Protocol for Wireless Sensor Networks," *Proc. 23rd Int'l Conf. Distributed Computing Systems Workshops*, pp. 710-715, 2003.

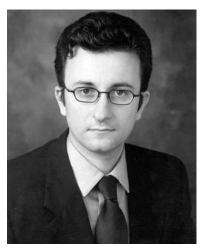
[15] C.L. Barrett, S.J. Eidenbenz, L. Kroc, M. Marathe, and J.P. Smith, "Routing, Coverage, and Topology Control: Parametric Probabilistic Sensor Network Routing," *Proc. Second ACM Int'l Conf. Wireless Sensor Networks and Applications*, pp. 122-131, Sept. 2003.

[16] I. Stojmenovic, "Geocasting with Guaranteed Delivery in Sensor Network," *IEEE Wireless Comm.*, vol. 11, no. 6, pp. 29-37, Dec. 2004.

[17] ASH Transceiver, Designer's Guide, http://www.rfm.com/products/data/tr1000.pdf, 2006.

[18] F. Zhao, J. Liu, J. Liu, L. Guibas, and J. Reich, "Collaborative Signal and Information Processing: An Information-Directed Approach," *Proc. IEEE*, vol. 91, no. 8, pp. 1199-1208, Aug. 2003.

Jin Zhu (S'01) received the BS degree in computer science and the MS degree in communications and information systems from Southwest Jiaotong University, Chengdu, China, in 1997 and 2000, and the PhD degree in electrical engineering from the New Jersey Institute of Technology in 2005. In August 2005, she joined the Department of Industrial Technology, University of Northern Iowa, where she is currently an assistant professor. Her main research interests include the design and modeling of sensor networks, QoS provisioning, resource allocation in ad-hoc and sensor networks, and performance evaluation of stochastic systems. She is a student member of the IEEE.



Symeon Papavassiliou (S'92-M'96) received the diploma in electrical engineering from the National Technical University of Athens, Greece, in 1990 and the MSc and PhD degrees in electrical engineering from Polytechnic University, Brooklyn, New York, in 1992 and 1995, respectively. Currently, he is with the Faculty of the Electrical and Computer Engineering Department, National Technical University of Athens. From 1995 to 1999, Dr. Papavassiliou was a senior technical staff member at AT&T Laboratories in Middletown, New Jersey, and in August 1999 he joined the Electrical and Computer Engineering Department at the New Jersey Institute of Technology (NJIT), where he was an associate professor. Dr. Papavassiliou was awarded the Best Paper Award in INFOCOM '94, the AT&T Division Recognition and Achievement Award in 1997, and the US National Science Foundation (NSF) Career Award in 2003. Dr. Papavassiliou has an established record of publications in his field of expertise, with more than 100 technical journal and conference published papers. His main research interests lie in the areas of computer and communication networks with emphasis on wireless communications and high-speed networks. He is a member of the IEEE.

Jie Yang (S'99) received the BS degree in information engineering and the MS degree in communication and information systems from Xidian University, Xidian, China, in 1996 and 1999, respectively. He also received the PhD degree in electrical engineering from the New Jersey Institute of Technology (NJIT) in 2004. In September 2004, he joined the Spirent Communications, New Jersey, where he is currently a software engineer. His main research interests include end-to-end QoS provisioning, admission control, resource allocation and traffic engineering, and Internet security. He is a student member of the IEEE.