QoI-aware Tradeoff Between Communication and Computation in Wireless Ad-hoc Networks

Sepideh Nazemi[†], Kin K. Leung[†] and Ananthram Swami[‡]
[†]Imperial College, London, UK. Email: {s.nazemi10, kin.leung}@imperial.ac.uk
[‡]U.S. Army Research Laboratory, Adelphi, MD, USA. Email: ananthram.swami.civ@mail.mil

***Abstract*—Data aggregation techniques exploit spatial and temporal correlations among data and aggregate data into a smaller volume as a means to optimize usage of limited network resources including energy. There is a trade-off among the Quality of Information (QoI) requirement and energy consumption for computation and communication. We formulate the energy-efficient data aggregation problem as a non-linear optimization problem to optimize the trade-off and control the degree of information reduction at each node subject to given QoI requirement. Using the theory of duality optimization, we prove that under a set of reasonable cost assumptions, the optimal solution can be obtained despite non-convexity of the problem. Moreover, we propose a distributed, iterative algorithm that will converge to the optimal solution. Extensive numerical results are presented to confirm the validity of the proposed solution approach.**

I. Introduction

The rapid growth of smart environments equipped with various types of sensors generates enormous amount of data. Such data must be gathered, transferred and processed to produce useful/meaningful information for end user(s). Considering the practical network constraints such as bandwidth or energy limitations and taking into account the high level of correlation among data generated in these environments, transferring the huge volume of generated data from many sources through the communication infrastructure is very inefficient, if not infeasible. Since data transmission is the main cause of energy consumption in such networks, the idea of conserving energy by reducing the amount of data transmissions has caught the attention of many researchers.

Data aggregation is defined as the process of gathering data from multiple sources (e.g., sensors), routing through multiple hops and processing (i.e., fusing, averaging or compressing) data in order to eliminate redundant transmission and provide fused/aggregated information to the end user(s) [1], [2]. An early work on energy efficient data aggregation developed a data centric-routing scheme called directed diffusion [3]. If the attributes of data generated by the sources match the interest of the sink, a gradient specifying the data rate and the direction of send is set up to identify the data generated by the sensor nodes. Directed diffusion eliminates the number of redundant transmissions by selecting only the useful data for transmission. Therefore, it can conserve a huge amount of energy. Moreover, cluster-based data aggregation protocols such as LEACH [4] and CLUDDA [5] have shown the effectiveness of this idea in prolonging network lifetime. Reviews of data aggregation techniques can be found in [1], [2], [6].

Research was partially sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. Furthermore, the authors would like to thank Dr. Mahsa Derakhshani of Loughborough University for many helpful discussions and comments.

The common assumption among most of the data aggregation work is that the energy required to process data is less than the energy required to send it. Therefore, it is beneficial to perform computation to reduce the data volume for transmission. Very little attention has been paid to computational energy cost while data aggregation is applied in the network [7]. However, Barr and Asanovic [8] investigated energy saving by lossless data compression and showed that with several typical compression tools, there is a net energy increase when compression is applied before transmission. Moreover, they discussed that the choice of how and whether to compress is not obvious and depends on hardware characteristics as well as software factors.

As for computational energy cost, Gallucio et. al. [7] studied the conditions under which aggregation is preferable. Eswaran et. al. [9] applied the network utility maximum (NUM) framework to determine the optimal compression and fusion factors for data aggregation as well as the optimal locations for performing data processing. With the goal of minimizing energy consumption in the network, [10] proposed a heuristic approach for determining the degree of data aggregation at each individual node. Similarly, Sharma et. al. [11] introduced a distributed approximate solution that makes joint compression and transmission decision.

In common with prior work such as [11], [10] and [9], we consider computational cost to obtain optimal aggregation decisions. But we propose a novel distributed solution that efficiently achieves the optimal solution, drawing upon key results in [12]. Moreover, a common assumption in the existing NUM work such as [9] is the concavity of the utility functions, which may not be valid [12]. In contrast, we formulate and solve our problem as a nonconvex optimization problem.

While data aggregation helps optimize the usage of network resources, it is also very important to consider how such aggregation affects the quality of information (QoI) required by end user(s). Data aggregation can cause QoI to deteriorate [13]. The degree to which a system can aggregate information is one of the main determining factors of QoI [14]. Despite the development of several data aggregation protocols that ensure desired Quality of Service, energy efficient data aggregation that guarantees desired QoI has not been well studied. For this reason, we investigate here the trade-off among communication and computation energy costs as well as the QoI requirement in order to determine the optimal degree of data reduction (rate) at each node in the network.

Summary of contribution: We study a class of multi-hop wireless networks where nodes are logically arranged as a tree and every node processes and aggregates data received from its children nodes. The aggregated data is transferred and further processed toward the root node. We formulate the problem of energy-efficient data aggregation with QoI constraint as a nonconvex optimization problem. We define the optimal data reduction rate as the degree to which data can be efficiently reduced while guaranteeing the required QoI for the end user. The proposed problem is intrinsically a nonconvex problem, which is hard to solve in general. However, for a set of reasonable energy cost structures, we can find the optimal solution by transforming the original problem to an equivalent one. By utilizing and analysing the KKT conditions, we prove that the nonconvex optimization problem can be solved exactly as the associated optimal duality gap is shown to be zero.

As a second contribution of this paper, we devise a distributed algorithm based on gradient descent, and show that the method can achieve the optimal solution efficiently. We evaluate our proposed method under different parameter settings and illustrate the performance of our proposed method through extensive simulations.

The rest of the paper is organized as follows. We formulate the problem and define our network topology and assumptions in section II. In section III, we introduce our solution approach and discuss the optimality of the proposed solution. We present a distributed algorithm that can achieve the optimal solution in section IV and discuss the performance of the proposed framework through extensive simulations in section V .

II. Problem formulation

Data generated in sensor networks often has some degree of redundancy due to spatial and temporal correlations among information observed or collected by various sensors. Therefore, it is possible to aggregate data as a means to optimize utilization of limited network resources. Considering energy as a critical resource in networks such as wireless sensor networks (WSNs), it may be desirable to reduce energy consumption for transmission and reception by aggregating the huge amount of data into a smaller volume. However, the greater the degree of data aggregation, the higher the amount of energy consumed for computation. Therefore, a trade-off exists among the amount of energy that each node spends on data reception, transmission and computation. We define the ratio of the volume of aggregated data to that of all incoming data at each node as the *data reduction rate* denoted by δ, $0 \leq \delta \leq 1$. The reduction rate is the degree by which a node can aggregate its received data effectively, and is a determining factor for QoI.

A. Assumptions

We assume that a data aggregation tree is formed among all involved nodes in the network after the user requests information from the network. The root node, *r*, of the tree is responsible for delivering the required information to the end user. Without loss of generality, it is assumed that only leaf nodes generate data and each of the other nodes in the tree receives data from its children nodes, processes and forwards aggregated data to its parent node.

Let the total energy consumption of node *i* denoted by F_i consist of energy spent in receiving e_{iR}, computing e_{iC} and transmitting e_{iT} as follows:

$$F_i = e_{iR} + e_{iC} + e_{iT}, \tag{1}$$

where

$$e_{iR} = \epsilon_{iR} y_i, \tag{2}$$

$$e_{iT} = \epsilon_{iT} y_i \delta_i, \tag{3}$$

$$e_{iC} = \epsilon_{iC} y_i l_i(\delta_i), \tag{4}$$

and ϵ_{iR}, ϵ_{iC} and ϵ_{iT} are the energy consumed in receiving, processing and transmitting one unit of data at node *i*, respectively. For the leaf nodes, ϵ_{iR} denotes the energy cost for observing or sensing or creating one unit of data. Since the greater the degree of data aggregation (i.e., the smaller amount of aggregated data produced after processing), the higher the energy consumption for computation, $l_i(\delta_i)$ is a scaling function to capture this characteristics of the computation energy consumption e_{iC}. We assume that $l_i(\delta_i)$ is a decreasing differentiable function of the reduction rate. Let $l_i(\delta_i)$ be defined as follow:

$$l_i(\delta_i) = \frac{1}{\delta_i} - 1, \tag{5}$$

where $\delta_i > 0$. Given (1) to (4), we define $f_i(\delta_i)$ as the total energy consumed by node *i* for one unit of receiving data:

$$f_i(\delta_i) = \epsilon_{iR} + \epsilon_{iC} l_i(\delta_i) + \epsilon_{iT} \delta_i. \tag{6}$$

Since it is assumed that only leaf nodes generate data, the total amount of data recieved by node *i* is

$$y_i = \sum_{\mathrm{j} \in \mathrm{C_i}} \delta_j y_j \quad i = 1 \ldots N \;, \tag{7}$$

where C_i denotes the set of children nodes of node i. y_i is assumed to be a constant value if *i* is a leaf node. Moreover, we assumed that the QoI delivered to the end user is a function of data reduction rate and received data associated with the root node which is responsible for delivering information to the end user, as denoted by $q_r(\delta_r, y_r)$.

B. Objective

We aim to minimize the total energy consumed by the network while meeting the QoI constraint for the user. We formulate the problem as follows:

$$\begin{array}{ll} \min\limits_{\boldsymbol{\delta}} & \sum\limits_{\mathrm{i}=1}^{\mathrm{N}} F_i(\delta_i, y_i) \\ \text{s.t.} & q_r(\delta_r, y_r) \geq \gamma \end{array} ,$$

where *N* is the total number of nodes in the data aggregation tree. $F_i(\delta_i, y_i)$ is the cost function of node i, which is a function of the total volume of input data y_i received from its children nodes and its data reduction rate δ_i. $\boldsymbol{\delta}$ is a vector of reduction rates for all nodes. $q_r(\delta_r, y_r)$ specifies the QoI function. Since the root node, *r*, is responsible for delivering the required information to the end user, the QoI constraint is associated only with the root node. γ indicates the QoI requirement threshold specified by the end user. Even though the problem has only a single QoI constraint associated with the root node, the data reduction rate must be chosen optimally at every node so that the total energy consumption is minimised while the QoI constraint for the end user can be satisfied.

Let the amount of data required by the end user be the QoI requirement. i.e., $q_r(\delta_r, y_r) = y_r \delta_r$. Therefore, the optimization problem is given by:

$$\begin{array}{ll} \min\limits_{\boldsymbol{\delta}} & \sum\limits_{\mathrm{i}=1}^{\mathrm{N}} F_i(\delta_i, y_i) \\ \text{s.t.} & y_r \delta_r \geq \gamma \end{array} . \tag{8}$$

III. Solution Approach

From equations (1)-(8), we see that the optimization problem in (8) is intrinsically a nonconvex problem. In general, there are no known simple necessary and sufficient conditions for determining global/local optima for nonconvex optimization problems. However, by analysing the conditions of the network and parameter settings, it is possible to transform (8) into an equivalent problem such that the global optimal solution to the problem can be obtained effectively.

Let *K* be the set of all leaf nodes and $h(k)$ denote the depth of node *k* in the aggregation tree. Assume that the root node *r* is located at level 0 and its depth is 0. Let

$\tau_k = (n_{k,h(k)}, n_{k,h(k)-1}, \ldots, n_{k,1}, n_{k,0})$ denote the unique path from node $n_{k,h(k)}$ to $n_{k,0}$ where by definition $n_{k,0} \triangleq 0$ (i.e., the root) and $n_{k,h(k)} \triangleq k$ (i.e., the node itself). Then, $n_{k,i}$ is the node at the i^{th} hop from the root in the unique path τ_k. Using this notation, we introduce the following Theorem which demonstrates how the problem of the total energy consumption over all the nodes can be converted into an equivalent problem by considering the total energy consumption over each unique path from each leaf node to the root.

Theorem 1. *The optimization problem in (8) is equivalent to the problem (9)*

$$\begin{aligned} \underset{\boldsymbol{\delta}}{min} \quad & \sum_{\mathrm{k}\in\mathrm{K}} y_k\Big(f_k(\delta_k) + \sum_{\mathrm{i}=0}^{\mathrm{h(k)}-1} f_{k,i}(\delta_{k,i}) \prod_{m=i+1}^{h(k)} \delta_{k,m}\Big) \\ s.t. \quad & \sum_{\mathrm{k}\in\mathrm{K}} y_k\Big(\prod_{i=0}^{h(k)} \delta_{k,i}\Big) \geq \gamma \end{aligned} \quad , \qquad (9)$$

where y_k denotes the amount of data generated by each leaf node k. $\delta_{k,i}$ and $f_{k,i}$ are the reduction rate and the total energy consumption for a unit of data corresponding to i^{th} node in the unique path from the root to leaf k, respectively.

Proof. According to the assumptions for energy consumption in (1)-(4), the total energy consumption of each node *i* is directly proportional to the total volume of data y_i received from all its children where the proportionality constant is given by $f_i(\delta_i)$ in (6). Since the data received at each node is the sum of all output data from its children (7), the linear relationship reveals that the total energy consumption at each node *i* is equal to the sum of energy spent on each of the data streams received from different children nodes. As only leaf nodes are assumed to generate data, we can reformulate the problem in (8) over all nodes in the network as one based on data generated at each leaf node. ∎

Despite the different representation, (9) still represents a nonconvex optimization problem. The Karush-Kuhn-Tucker (KKT) conditions are necessary conditions for the global optimal solution to the nonlinear primal problem in (9) (e.g., see [15]). By analysing the KKT conditions for the dual problem associated with (9), we prove in the following that the primal and dual problem have zero duality gap.

Let $d(\lambda)$ be the Lagrangian dual function of (9) where λ is the Lagrangian multiplier (price) associated with the QoI constraint in (9). Then, the dual optimization problem is given by:

$$\max_{\lambda} d(\lambda) = L(\boldsymbol{\delta}^*, \lambda) \quad , \qquad (10)$$

where $L(\boldsymbol{\delta}, \lambda)$ is the Lagrange function given by

$$\begin{aligned} L(\boldsymbol{\delta}, \lambda) \;=\; & \sum_{\mathrm{k}\in\mathrm{K}} y_k\Big(f_k(\delta_k) + \sum_{\mathrm{i}=0}^{\mathrm{h(k)}-1} f_{k,i}(\delta_{k,i}) \prod_{m=i+1}^{h(k)} \delta_{k,m}\Big) \\ & -\lambda\Big(\sum_{\mathrm{k}\in\mathrm{K}} y_k(\prod_{i=0}^{h(k)} \delta_{k,i}) + \gamma\Big). \end{aligned} \qquad (11)$$

and $\boldsymbol{\delta}^*$ is obtained from

$$\boldsymbol{\delta}^* = \mathrm{argmin}\; L(\boldsymbol{\delta}, \lambda)\;. \qquad (12)$$

Let $\boldsymbol{\delta}^*$ and λ^* be any primal and dual optimal solution. The KKT conditions stem from the fact that the gradient of the Lagrangian function must be zero [16]. That is: $\nabla L(\boldsymbol{\delta}^*, \lambda^*) = 0$. We consider the stationary condition for two different types of nodes namely, the root (node 0) and any other node *j* (i.e., *j* can be a leaf or an intermediate node) in the aggregation tree. As shown in the Appendix, the optimal data reduction rates δ_0 and δ_j for these node types are obtained as follows:

$$\delta_0(\lambda) = \sqrt{\frac{\epsilon_{0C}}{\epsilon_{0T} - \lambda}}\;. \qquad (13)$$

$$\delta_j(\lambda) = \sqrt{\frac{\epsilon_{jC}}{\epsilon_{jT} + \sum_{\mathrm{i}=0}^{\mathrm{h(j)}-1}(f_{j,i}(\delta_{j,i})\Delta_{j,i}) - \lambda(\Delta_j)}}\;, \qquad (14)$$

where $\Delta_{j,i} = \prod_{m=i+1}^{h(j)-1} \delta_{j,m}$ and $\Delta_j = \prod_{m=0}^{h(j)-1} \delta_{j,m}$. Note that feasibility and satisfaction of constraints on δ are discussed in Section IV-C.

Equations (13) and (14) are called the price-based solution functions for the problem (9), because they are expressed as the function of the Lagrangian multiplier (price) λ. Equations (13) and (14) are for the specific cost function (6). For a general cost function which is decreasing and differentiable, δ can be uniquely determined from $f'(\delta)$. By applying results in [12], we have the following Theorem.

Theorem 2. *The strong duality property holds for the primal and the dual problem in (9)-(10). Furthermore, an iterative algorithm exists to obtain the optimal solution for both problems.*

Proof. Observe from (13) and (14) that the optimal data reduction rates for all nodes in the aggregation tree are continuous functions of the price variable λ in the feasible range, including the optimal value of λ^*. Based on Theorem 1 in [12], this continuity property guarantees that the duality gap for (9) and (10) is zero and that the optimal solutions for the primal-dual problems can be obtained by an iterative method. ∎

IV. DISTRIBUTED SOLUTION FRAMEWORK

Given that the dual problem in (10) is a linear function of λ (see eq.11), it is possible to utilize the gradient methods to solve the dual problem.

The gradient descent method is a popular technique to find local optima. At each step of the iteration, the search continues in the negative direction of the gradient of the function. The gradient descent recursion for solving (10) is given by:

$$\lambda^{(t+1)} = \lambda^{(t)} - \alpha\Big(\frac{\partial L(\delta^*(\lambda), \lambda)}{\partial \lambda}\Big)\;, \qquad (15)$$

where *t* is the iteration index and $\alpha \geq 0$ is the step size.

Since all price-based solution functions, namely (13) and (14), associated with problem (9) are continuous over λ's domain, one can devise an iterative algorithm based on (15) that will converge to the global optimal solution, as suggested by Theorem 2.

A. Distributed Algorithm

Due to the complex relationships among nodes, imposed by the tree structure and the result in (7), the problem (9) and its corresponding Lagrangian function (11) cannot be easily separated to develop a distributed solution. However, a careful observation of the price-based solution functions in (13) and (14) reveals that the optimal data reduction at a node (say node *i*) only depends on the optimal reduction rates of all ancestors of node *i* in the aggregation tree and the optimal price value λ^*. That is, if a node knows just the solutions associated with its ancestors and the optimal price value λ^*, it can calculate its δ. We exploit this critical observation to devise our distributed solution. Algorithm1 presents the pseudocode of the proposed technique. Steps for the Phase1 and Phase2 operations of the

Algorithm 1 are presented as Algorithm 2 and Algorithm 3 respectively. Note that *t* denotes the iteration index. An initial value λ_0 is considered as an input to the Algorithm 1.

Algorithm 1 Iterative Distributed Algorithm

Input: Initial value for price value λ.
Output: The optimal reduction rates associated with all nodes and the optimal price value.

$t \leftarrow 0$;
$\lambda^{(t)} \leftarrow \lambda_0$;
Repeat:
Phase1(*Node_id*): Nodes calculate $\delta(\lambda^{(t)})$ based on (13) and (14);
Phase2(*Node_id*):Nodes calculate their QoI contributions based on (7);
New price is updated based on (15) by the root and sent to all children;
Until *convergence*;

B. Information Exchange

Although each node needs all price-based solutions of its ancestors and the price value to calculate its reduction rate, each node can receive the required information just from its parent. Moreover, it is also possible to reduce the number of messages exchanged among nodes as follows:

The denominator of $\delta_j(\lambda)$ in (14) is :

$$u_j = \epsilon_{jT} + \sum_{\mathrm{i}=0}^{\mathrm{h(j)}-1} (f_{j,i}(\delta_{j,i}) \prod_{m=i+1}^{h(j)-1} \delta_{j,m}) - \lambda(\prod_{m=0}^{h(j)-1} \delta_{j,m}) \; , \quad (16)$$

ϵ_{jT} is a node parameter and it is known to node *j*. Let $w_j := \lambda \prod_{m=0}^{h(j)-1} \delta_{j,m}$ and $s_j := \sum_{i=0}^{h(j)-1} (f_{j,i}(\delta_{j,i}) \prod_{m=i+1}^{h(j)-1} \delta_{j,m})$ which only involve the compression parameters higher up in the tree; hence these can be computed by the parent of node *j*; and that parent node only needs to send these two parameters to all its children. Note that $\lambda^{(t)}$ is the price value calculated at t^{th} round of the algorithm. After calculating $\delta_j(\lambda)$ for all $j \in N$, it is necessary to update λ and check if the algorithm has converged. In order to update λ, $\frac{\partial L(\delta,\lambda)}{\partial \lambda}$ which is in fact the constraint of the problem (9) must be determined first. The algorithm converges to the optimal when $\frac{\partial L(\delta,\lambda)}{\partial \lambda} = 0$. That is, the constraint becomes active.

Algorithm 2 Distributed computation of priced-based function

Input: $n_{k,i}$ ▷ Node Id.
Output: The reduction rate associated with node $n_{k,i}$.

if $n_{k,i} = 0$ **then** ▷ n is the root.
  $s_{k,i} \leftarrow 0$;
  $w_{k,i} \leftarrow \lambda^{(t)}$
else
  $\delta_{k,i} \leftarrow \sqrt{\frac{\epsilon_{iC}}{\epsilon_{iT} + s_{k,i} - w_{k,i}}}$;
  if $n_{k,i} \notin K$ **then** ▷ n is an intermediate node.
    $s_{k,i+1} \leftarrow s_{k,i}\delta_{k,i} + f_{k,i}(\delta_{k,i})$;
    $p_{k,i+1} \leftarrow p_{k,i}\delta_{k,i}$;
    PHASE1($n_{k,i+1}$)
  else
    PHASE2($n_{k,i}$);

Algorithm 3 presents the steps for calculating $\frac{\partial L(\delta,\lambda)}{\partial \lambda}$ and updating the price value λ. As shown in Algorithm 3, each node calculates its contribution to the QoI constraint. By the time the root receives information from its children, it can calculate the new λ based on (15) and check whether the algorithm has converged. If so, the root node will not send an updated message and after a certain amount of time, all nodes will finalize their values of δ as optimal. Otherwise, the root will send a message containing the updated value of λ (i.e., s_0, w_0) to its children.

Algorithm 3 Distributed computation of QoI and the price

Input: $n_{k,i}$ ▷ Node Id.
Output: The updated price value.

if $n_{k,i} \in K$ **then**
  $q_i \leftarrow y_k\delta_k$;
  send q_i to *parent-of-k*;
  PHASE2(*parent-of-k*);
if $n_{k,i} \notin K$ **then**
  if *receive-all-*q_j*-from-children* **then**
    $y_i \leftarrow \sum_{\mathrm{j} \in \mathrm{C_i}} q_j$
  else
    Wait-until-receive-q_j*-from-all-children*;
  $q_i \leftarrow y_i\delta_i$;
  send q_i to *parent-of-i*;
  PHASE2(*parent-of-i*);
if $n_{k,i} = 0$ **then**
  $\lambda^{(t+1)} \leftarrow \lambda^{(t)} - \alpha(\frac{\partial L(\delta,\lambda)}{\partial \lambda})$;
  $t \leftarrow t + 1$;
  if !*algorithm-converge* **then**
    PHASE1($n_{k,i}$)
  else
    return ;

C. Solution Feasibility

Consider the constraint in the problem (9). That is:

$$\sum_{\mathrm{k} \in \mathrm{K}} y_k \Big(\prod_{i=0}^{h(k)} \delta_{k,i} \Big) \geq \gamma \; .$$

By assumption, $0 < \prod_{i=0}^{h(k)} \delta_{k,i} \leq 1, \; \forall k \in K$. Therefore, if $\sum_{k \in K} y_k \leq \gamma$, it does not exist a feasible solution to the problem (9). That means that a feasible solution can exist only if the total amount of data generated in the network is greater than or equal to the QoI requirement.

Moreover, by definition we have $0 < \delta_i \leq 1$. Therefore, if δ_i, $\forall i \in N$ determined during iterations of Algorithm 2 is outside of this box constraint (i.e., $\delta_i^*(\lambda^*) < 0$ or $\delta_i^*(\lambda^*) > 1$), we map the solution to the upper bound of data reduction rate value (i.e., $\delta_i^*(\lambda^*) \mapsto 1$). The intuition behind this is explained as follow.

As an example, consider (19) and (22). Therefore, we have:

$$\epsilon_T - \frac{\epsilon_C}{\delta_0^2} = \lambda \; .$$

Note that $f_0'(\delta_0)$ is an increasing function of δ_0. During iteration of Algorithm 3, we update the value of λ. The maximum value of λ in this case occurs when δ_0 reaches the upper bound (i.e., 1). Therefore if the optimal value of λ causes δ_0 to attain an imaginary value, we map the solution to the upper bound value. In addition, if $\lambda^* = 0$ and $\epsilon_{0C} \gg \epsilon_{0T}$ then $\delta_0(\lambda) > 1$. In this case, we map the solution to the upper bound as well (i.e., $\delta_0^*(\lambda^*) \mapsto 1$) since, sending

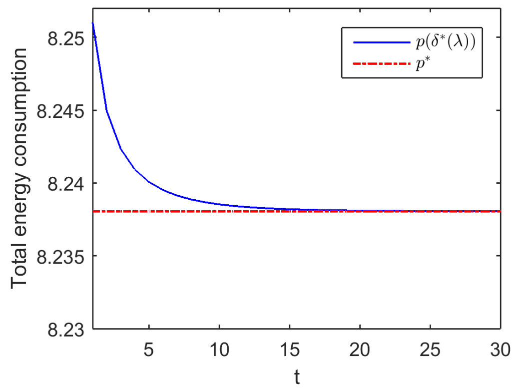

Fig. 1: Convergence of the proposed iterative method vs. iteration index t.

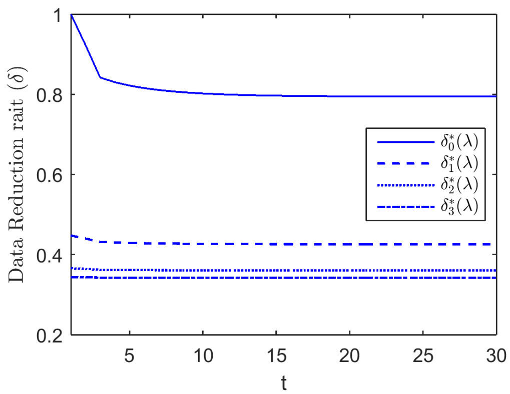

Fig. 2: Convergence of estimated compression rate vs. iteration index t for nodes at different levels of the tree..

all information (choosing $\delta_0 = 1$) is more energy efficient that processing and reducing the data. Note that the same argument is applicable at the intermediate or leaf nodes.

V. Numerical Evaluation

In this section, we present numerical results for evaluating the proposed distributed method. The function $l_i(\delta_i) = \frac{1}{\delta_i} - 1$, for $\delta_i > 0$ is the scaling function. The energy consumption parameters ϵ_R, ϵ_T and ϵ_C are set at $\epsilon_R = \epsilon_T = 0.02, \epsilon_C = 0.01$. Each leaf node generates 15 packets. The QoI threshold γ is assumed to be 5 data packets. In this experiment the feasibility and existence of optimal solution is guaranteed as we assume the amount of QoI requirement by the end use is less than the total amount of data generated in the network. We consider a symmetrical binary aggregation tree with 15 nodes and identical parameters for all nodes. This way, we can compare results of the proposed distributed method with the optimal solution generated by exhaustive search.

Fig.1 presents the convergence of the proposed distributed optimization algorithm versus the number of iterations. The solid line presents the value of the objective function $p(\delta(\lambda))$ at each iteration and the dashed line depicts the optimal value of the objective function denoted by p^* and obtained by an exhaustive search algorithm. As the graph shows, after 20 iterations or so, the distributed algorithm converges to the optimal solution identical to that obtained by exhaustive search p^*.

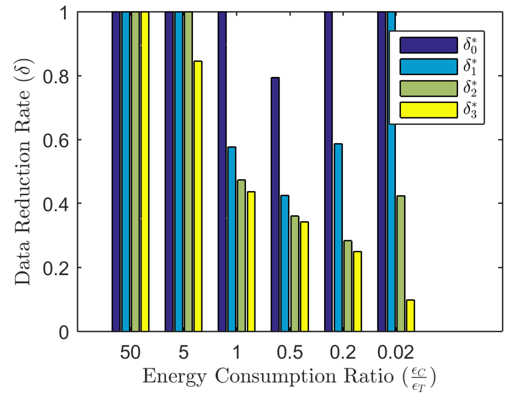

Fig. 3: Data reduction rates variations at different levels of the data aggregation tree under different parameters' settings.

The distributed algorithm converges when the residual inconsistency value equals to zero. i.e., $\frac{\partial L(\delta,\lambda)}{\partial \lambda} = 0$, where the price value gains its optimal value 0.004141. In this experiment, the initial value of the Lagrangian multiplier λ and the step size α were set at 0 and 0.001 respectively. Since we consider a symmetrical aggregation tree with homogeneous nodes, nodes located at the same level of the tree have identical optimal data reduction rates. Fig.2 illustrates the variation of the reduction rates for nodes at level 0, 1, 2 and 3 after each iteration.

We evaluate the performance and correctness of the distributed algorithm under different parameter settings. In particular, we consider the computation energy to transmission energy ratio and test the algorithm under various values from extreme to moderate cases as presented in Table I. The optimal reduction rate at each level associated with these parameters settings is illustrated by Fig.3. It can be seen that when ϵ_C is much greater than ϵ_T (an extreme case) all nodes at each level will send all received data. That means they will not compress data due to extremely high cost of computation. In contrast, if $\epsilon_C \ll \epsilon_T$ (the last) it is beneficial to compress the data at lower levels of the aggregation tree in order to spend less energy for transmitting data. The increasing trend of data reduction rate from leaf nodes to the root node continues among other moderate parameter settings as illustrated by Fig.3. However, this trend is valid only for a symmetrical data aggregation tree. A different pattern can be observed when we consider an irregular data aggregation tree with heterogeneous nodes as illustrated by Fig. 4. In this experiment, ϵ_C is randomly chosen from the interval $[0.01, 0.02]$. Furthermore, it is assumed that $\epsilon_T = \epsilon_R = 0.02$. The input data at each leaf node is 15 packets. The algorithm converged after 560 iterations and the delivered QoI and optimal price are 8 and 0 respectively.

Table I presents optimal price value and delivered QoI corresponding to each parameters setting. The optimal price attains positive values when the delivered QoI is exactly equal to QoI threshold γ. That means the constraint in (9) is active, $\sum_{\mathrm{k}\in\mathrm{K}} y_k(\prod_{i=0}^{h_k} \delta_{k_i}^*) = \gamma$. On the other hand, when the delivered QoI is greater than γ, the Lagrangian multipliers (price) equals zero. This result is compatible with the fact that for any primal and dual optimization problem with zero duality gap, the complementary slackness condition of KKT conditions must be satisfied by both the optimal

$\frac{\epsilon_C}{\epsilon_T}$	λ^*	Delivered QoI	# Iterations	λ_0	α
50	0	120	17	0.0002	0.0000001
5	0	101	20	0.0002	0.0000001
1	0	14	212	0.0002	0.0000001
0.5	0.0041	5	324	0.01	0.001
0.2	0.121	5	112	0.02	0.001
0.02	2.444	5	72	0.1	0.01

TABLE I: Performance of the distributed algorithm under different settings

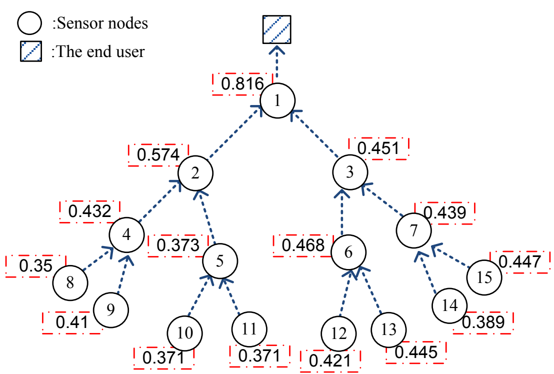

Fig. 4: Data reduction rates for a tree with heterogeneous nodes.

solution and the optimal Lagrangian multiplier. That means $-\lambda^*(\sum_{k\in K} y_k(\prod_{i=0}^{h_k}\delta^*_{k_i})+\gamma)=0$.

The Table also presents the number of iterations required to reach convergence. Note that initial price value of λ can affect the speed of convergence (compare this result to that in the previous experiment).

In this work, we identified practical network conditions and assumptions under which we could show that the duality gap between primal and dual problem is zero. Our future work will involve further investigation on the proposed problem in order to characterize more generic network conditions and assumptions that could lead to the optimal solution.

Appendix

Price-based solution function calculation

Taking the partial derivative of (11) at δ_0 (i.e., root node) , δ_k where $k \in K$ and δ_j where $j \notin K$ we have

$$\frac{\partial L}{\partial \delta_0} = \sum_{k\in K} y_k\left(f_0'(\delta_0)\prod_{m=1}^{h(k)}\delta_{k,m}\right) -\lambda\sum_{k\in K} y_k\left(\prod_{m=1}^{h(k)}\delta_{k,m}\right)=0. \quad (17)$$

$$\frac{\partial L}{\partial \delta_k} = y_k f_k'(\delta_k) + y_k\sum_{i=0}^{h(k)-1}\left(f_{k,i}(\delta_{k,i})\prod_{m=i+1}^{h(k)-1}\delta_{k,m}\right) -\lambda y_k\left(\prod_{m=0}^{h(k)-1}\delta_{k,m}\right)=0 \quad for \;\; k\in K. \quad (18)$$

Given that $\prod_{i\in N}\delta_i \neq 0$, by factorising and rearranging the terms in (17) and (18) we have:

$$f_0'(\delta_0)=\lambda\,. \quad (19)$$

$$f_k'(\delta_k) = -\sum_{i=0}^{h(k)-1}\left(f_{k,i}(\delta_{k,i})\prod_{m=i+1}^{h(k)-1}\delta_{k,m}\right)+\lambda\left(\prod_{m=0}^{h(k)-1}\delta_{k,m}\right).$$

To compute $\frac{\partial L}{\partial \delta_j}$, notice that δ_j only affects the energy cost of node j and its ancestors. Let $\tau_j=(n_{j,h(j)},\ldots,n_{j,1},n_{j,0})$ (as defined in Section III). Let y_j denotes the total volume of incoming data at node j. Then

$$\frac{\partial L}{\partial \delta_j} = y_j f_j'(\delta_j) + y_j\sum_{i=0}^{h(j)-1} f_{j,i}(\delta_{j,i})\prod_{m=i+1}^{h(j)-1}\delta_{j,m} -\lambda y_j\prod_{m=0}^{h(j)-1}\delta_{j,m}=0. \quad (20)$$

We treat (20) in the same manner as (17) and (18) to obtain

$$f_j'(\delta_j) = -\sum_{i=0}^{h(j)-1} f_{j,i}(\delta_{j,i})\prod_{m=i+1}^{h(j)-1}\delta_{j,m}+\lambda\prod_{m=0}^{h(j)-1}\delta_{j,m}\,. \quad (21)$$

Notice that we do not need to deference between leaf nodes and intermediate nodes. For the cost model in (6), we have:

$$f_i'(\delta_i)=\epsilon_T-\frac{\epsilon_C}{\delta_i^2}\,. \quad (22)$$

Substituting (22) into (19) and (21) lead to (13) and (14).

References

[1] R. Rajagopalan and P. Varshney, "Data-aggregation techniques in sensor networks: a survey," *IEEE Communications Surveys & Tutorials*, vol. 8, no. 4, pp. 48–63, 2006.
[2] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-network aggregation techniques for wireless sensor networks: a survey," *IEEE Trans on Wireless Communications*, vol. 14, no. 2, pp. 70–87, 2007.
[3] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," in *MobiCom'00*, 2000.
[4] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans on wireless communications*, vol. 1, no. 4, pp. 660–670, 2002.
[5] S. Chatterjea and P. Havinga, "A dynamic data aggregation scheme for wireless sensor networks," in *Program for Research on Integrated Systems and Circuits,Veldhoven, The Netherlands*, 2003.
[6] P. Jesus, C. Baquero, and P. S. Almeida, "A survey of distributed data aggregation algorithms," *IEEE Communications Surveys & Tutorials*, 2015.
[7] L. Galluccio, S. Palazzo, and A. T. Campbell, "Efficient data aggregation in wireless sensor networks: an entropy-driven analysis," in *PIMRC'08*, 2008.
[8] K. C. Barr and K. Asanović, "Energy-aware lossless data compression," *ACM Trans on Computer Systems (TOCS)*, vol. 24, no. 3, pp. 250–291, 2006.
[9] S. Eswaran, J. Edwards, A. Misra, and T. F. LaPorta, "Adaptive in-network processing for bandwidth and energy constrained mission-oriented multihop wireless networks," *IEEE Trans on Mobile Computing*, vol. 11, no. 9, pp. 1484–1498, 2012.
[10] S. Nazemi, K. K. Leung, and A. Swami, "A distributed, energy-efficient and qoi-aware framework for in-network processing," *PIMRC'14*, 2014.
[11] A. B. Sharma, L. Golubchik, R. Govindan, and M. J. Neely, "Dynamic data compression in multi-hop wireless networks," in *ACM SIGMETRICS Performance Evaluation Review*, 2009.
[12] G. Tychogiorgos, A. Gkelias, and K. K. Leung, "A non-convex distributed optimization framework and its application to wireless ad-hoc networks," *IEEE Trans on Wireless Communications,*, vol. 12, no. 9, pp. 4286–4296, 2013.
[13] A. Bar-Noy, G. Cirincione, R. Govindan, S. Krishnamurthy, T. LaPorta, P. Mohapatra, M. Neely, and A. Yener, "Quality-of-information aware networking for tactical military networks," in *IEEE PERCOM Workshops*, 2011.
[14] S. A. Ehikioya, "A characterization of information quality using fuzzy logic," in *IEEE International Conf. on Fuzzy Information Processing Society*, 1999.
[15] G. Tychogiorgos and K. Leung, "Optimization-based resource allocation in communication networks," *Computer Networks*, 2014.
[16] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.