Local Optimization for Energy Efficiency and QoI in In-Network Processing

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

***Abstract*—In-network processing (INP) is a promising method that allows aggregation of data as a means to optimize the utilization of network resources without violating the quality of information (QoI) requirements. Given the potentially large scale and dynamic network environment, optimization of INP requires a distributed framework that can adapt easily to network changes and user's QoI requirements. We develop the principle for designing such a distributed mechanism to facilitate control of INP. We formulate the INP problem as a non-linear optimization problem in order to minimize the total energy consumption by all associated nodes while satisfying QoI requirements. High inter-dependency among the nodes renders the optimization problem intractable. To cope with this, we propose a fully distributed, but suboptimal, approach which we call Locally Constrained Optimization. We prove that under a set of conditions, the INP process can be fully distributed while performing closely to the optimality. The significance of the proposed distributed framework is that it requires each node to make independent decisions locally for data aggregations; this naturally enhances robustness and resiliency of network and data load dynamics.**

I. Introduction

In-Network Processing (INP) primarily aims to aggregate (e.g., compression, fusion and averaging) data from different sources with the objective of reducing the data volume for further transfer, thus reducing energy consumption and increasing the network lifetime [1]. When INP is applied in an information network, it is crucial to consider how such data processing affects the quality of information (QoI) at the receiving end, which is represented by multi-dimensional metrics [2], e.g., information accuracy, completeness, reliability and timeliness. How INP should be carried out for satisfactory QoI at the user level remains an open research issue. For example, [3], [4] investigate the QoI and introduce various models to maximize a utility function of QoI. However, they do not consider INP or the impact of further INP process on aggregated information. For this reason and in sharp contrast to other related literature, the objective of this paper is to introduce a fully distributed framework to facilitate controlling of INP process at intermediate nodes while satisfying the end user QoI requirements.

In a dynamic environment, such as military field, fully distributed methods are very desirable due to the ability to readily adapt to the network changes and constraints and deal with a huge amount of data generated in this environment. For this purpose, as a starting point, we consider the amount of information that user needs to receive as a QoI parameter. Note that required amount of data can represent the accuracy metric of QoI [3].

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. Stuart Farquhar of DSTL for many helpful discussions and comments.

II. Problem Formulation and Assumptions

Assume a Wireless Sensor Network (WSN) monitoring the level of chemical contamination in an area. A user approaches the area and wants to know the level of chemical contamination of a specific part of the area. The user sends a query to the network specifying his area of interest and the amount of information that user needs to receive (e.g., number of data packets) as a QoI requirement threshold. We assume a data aggregation tree is formed after the user queries the network.

Furthermore, we assume that the total energy consumption of a node consists of the energy spent in receiving p_R , computing p_C and transmitting p_T its data. Among these operations, data transmission typically uses more energy than the others [5]. Data generated in the information networks has a large amount of redundancy due to the spatial and temporal correlation among sensors. Therefore, it is possible to aggregate data as a means of reducing energy consumption for transmission and reception, without sacrificing QoI for the end users. On the other hand, the more the data are aggregated, the higher the computational energy cost is. Therefore, an energy trade-off exists among the energy that each node spends for data reception, transmission and computation. A key for the energy trade-off is the data reduction rate (denoted by δ between 0 and 1), which is the ratio of the amount of aggregated data to that of all original input data at a node. The optimal energy trade-off is determined by choosing the optimal reduction rate at each node.

We formulate the problem of finding the optimal data reduction rates at all nodes as an optimization problem. Our goal is to choose the data reduction rates at all nodes involved in order to minimize the total energy consumption, while ensuring that the amount of aggregated data for the end user exceeds a specified QoI threshold γ.

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

where *n* is the total number of nodes in the network aggregation tree and $p_i(\delta_i, y_i)$ is the total energy consumption which is a function of the volume of input data y_i and the data reduction rate δ_i at node i. Since node *r* is responsible for delivering the required information to the user, the constraint $y_r\delta_r \geq \gamma$ specifies the minimum volume of aggregated data that the user requires from all source nodes in the area of interest.

We refer to (1) as a global optimization (GO) problem. Even though the problem has only a single constraint, data reduction rate must be chosen optimally at every node so that the constraint for the end user can be satisfied.

The GO problem in (1) is a variant of 0/1 knapsack problem [6]. Therefore, intractability and interdependency among nodes motivate the need for a distributed method where no centralized operator is required and complex network structures, such as multi-level data aggregation tree can be handled easily. To this end, we will propose our fully distributed framework, a solution to the GO problem.

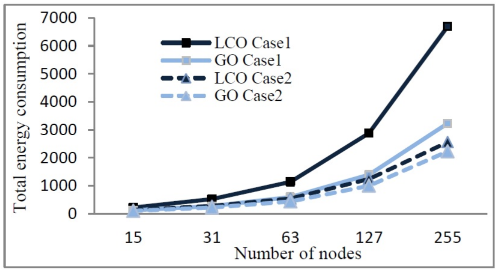

Fig. 1. Impact of computation and transmission cost on total energy consumption.

III. Local Optimization Approach

We introduce our model the Local Constrained Optimization framework (to be referred to as LCO model) as follows:

$$\begin{aligned} &\sum_{i=1}^{n} \min_{\delta_i} p_i(\delta_i, y_i) \\ &\text{s.t. } \delta_i y_i \geq \gamma \\ &\quad y_i = \sum_{j \in C_i} \delta_j y_j, \quad for\ i = 1, ..., n \end{aligned} \quad , \tag{2}$$

where C_i denotes the set of children nodes of node i in the aggregation tree.

The significance of introducing (2) is explained as follows. It turns out that under a reasonable set of assumptions, the order of the minimization and summation operators in (2) can be switched, while yielding the same optimal solution as follows.

$$\begin{aligned} &\min_{\{\delta_i\}} \sum_{i=1}^{n} p_i(\delta_i, y_i) \\ &\text{s.t. } \delta_i y_i \geq \gamma \\ &\quad y_i = \sum_{j \in C_i} \delta_j y_j, \quad for\ i = 1, ..., n \end{aligned} \quad . \tag{3}$$

We refer to (3) as Constrained Optimization (CO) problem. While (3) and (1) show a global optimization model, the difference between the CO problem (3) and the GO problem (1) is that in the CO problem, each node has its own QoI constraint, while the GO problem has only one constraint for the root node *r*. We shall develop conditions (assumptions) under which the LCO leads to such a global optimal solution as follows.

Let p_i denotes the total energy consumption of node i as

$$p_i = p_{i_R} + p_{i_C} + p_{i_T}, \tag{4}$$

where p_{i_R}, p_{i_C} and p_{i_T} denote the energy spent in receiving, aggregating and transmitting data by node i, respectively. Then, we assume

$$p_{i_R} = f(y_i) = e_R y_i. \tag{5}$$

$$p_{i_T} = g(y_i, \delta_i) = e_T y_i \delta_i, \tag{6}$$

$$p_{i_C} = k(y_i, \delta_i) = e_C y_i q_i(\delta_i), \tag{7}$$

$$y_i = \sum_{j \in C_i} y_j \delta_j, \tag{8}$$

where e_R, e_T and e_C are the energy consumed in receiving, transmitting and processing one unit of data respectively. δ_i and y_i show the reduction rate and the amount of received data respectively. $q_i(\delta_i)$ is a scaling function for energy consumption of computation which is a decreasing differentiable function of the reduction rate δ_i and captures the influence of the reduction rate δ_i on p_{i_C}. Furthermore, (8) shows the amount of data that node i receives.

There may be a concern that the linear model in (5) to (7) is unable to adequately adjust to all the characteristics of communication and computation in the network (e.g., coding and processing functions); however, as a general assumption in [5] we assume a linear model here.

With these assumptions, we present the following theorem which introduces sufficient conditions under which the order of the minimization and summation operators in (3) can be switched while preserving optimality.

Theorem. *Given the energy consumption represented by (4) to (7) and the amount of received data in (8) for each node in a single level data aggregation tree, the LCO model (2) is equivalent to the CO (3) as follows:*

$$\begin{array}{ll} \min_{\{\delta_i\}} \sum_{i=1}^{n} p_i(\delta_i, y_i) = & \sum_{i=1}^{n} \min_{\delta_i} p_i(\delta_i, y_i) \\ s.t.\ \delta_i y_i \geq \gamma & s.t.\ \delta_i y_i \geq \gamma \\ y_i = \sum_{j \in C_i} \delta_j y_j & y_i = \sum_{j \in C_i} \delta_j y_j \\ for\ i = 1, ..., n & for\ i = 1, ..., n \end{array} . \tag{9}$$

Proof. Duo to space limitation we omit the proof here. ■

The importance of introducing the LCO model is that under practical parameter setting it can provide very close approximate solution to the GO problem as illustrated by Fig 1.

We evaluate the LCO modelled in (2) under two different parameters settings; Case 1 ($e_R = e_T = e_C = 0.00024J$) and Case 2 ($e_R = e_T = 0.00024J, e_C = 0.00012J$). QoI threshold γ is assumed to be 5 data packets for both cases. We compare the LCO results to the optimal value of global optimization (GO) in (1). We assume that the full binary aggregation tree rooted at node *r* is formed for the request of information by a user. The function $q_i(\delta_i) = \frac{1}{\delta_i} - 1$, $\delta_i > 0$ is considered to reflect the impact of the data aggregation process on energy consumption.

As Fig. 1 illustrates, imposing individual constraints on each node and forcing them to satisfy the user QoI threshold leads to a gap between energy consumption of the GO and LCO models. However, under a practical network setting presented by Case 2, the LCO model performs very close to the GO value. The reason for this exciting result is, due to less expensive computation cost, each node aggregates more data and as a result, consumes less energy for data transmission.

In conclusion, we have proven that under a set of reasonable assumptions the optimal data reduction rate can be determined by each node based on local information in a fully distributed manner. Computer simulations show that the new framework, LCO, can perform very close to the global optimum for parameter settings where communication costs more than computation.

References

[1] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, “In-network aggregation techniques for wireless sensor networks: a survey,” *IEEE Wireless Communications*, 2007.

[2] C. Bisdikian, L. M. Kaplan, M. B. Srivastava, D. J. Thornley, D. Verma, and R. I. Young, “Building principles for a quality of information specification for sensor information,” in *FUSION*, 2009.

[3] E. Ciftcioglu, A. Yener, R. Govindan, and K. Psounis, “Operational information content sum capacity: Formulation and examples,” in *FUSION*, 2011.

[4] E. N. Ciftcioglu, A. Yener, and M. J. Neely, “Maximizing quality of information from multiple sensor devices: The exploration vs exploitation tradeoff,” 2013.

[5] 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 Transactions on Mobile Computing*, 2012.

[6] M. M. Güntzer and D. Jungnickel, “Approximate minimization algorithms for the 0/1 knapsack and subset-sum problem,” *Operations Research Letters*, 2000.