Exact and Heuristic Algorithms for Data-Gathering Cluster-Based Wireless Sensor Network Design Problem

Hui Lin and Halit Üster

***Abstract*—Data-gathering wireless sensor networks (WSNs) are operated unattended over long time horizons to collect data in several applications such as those in climate monitoring and a variety of ecological studies. Typically, sensors have limited energy (e.g., an on-board battery) and are subject to the elements in the terrain. In-network operations, which largely involve periodically changing network flow decisions to prolong the network lifetime, are managed remotely, and the collected data are retrieved by a user via internet. In this paper, we study an integrated topology control and routing problem in cluster-based WSNs. To prolong network lifetime via efficient use of the limited energy at the sensors, we adopt a hierarchical network structure with multiple sinks at which the data collected by the sensors are gathered through the clusterheads (CHs). We consider a mixed-integer linear programming (MILP) model to optimally determine the sink and CH locations as well as the data flow in the network. Our model effectively utilizes both the position and the energy-level aspects of the sensors while selecting the CHs and avoids the highest-energy sensors or the sensors that are well-positioned sensors with respect to sinks being selected as CHs repeatedly in successive periods. For the solution of the MILP model, we develop an effective Benders decomposition (BD) approach that incorporates an upper bound heuristic algorithm, strengthened cuts, and an ε-optimal framework for accelerated convergence. Computational evidence demonstrates the efficiency of the BD approach and the heuristic in terms of solution quality and time.**

***Index Terms*—Benders decomposition (BD), network design, wireless sensor networks (WSNs).**

I. INTRODUCTION

RECENT advances in wireless networking, embedded microprocessors, integration of microelectromechanical systems (MEMS), and nanotechnology have enabled rapid development of low-cost, low-power, and multifunctional sensors [1], [2]. Very small in size, sensors are capable of sensing, data processing, and communicating with each other or with the data sinks. A group of sensors communicating in a wireless medium for the purpose of gathering data and transmitting it to a user (sinks) form a *wireless sensor network* (WSN).

In the WSN applications, the main purpose is to monitor and collect data by the sensors and then transmit this data to the sinks. Tilak *et al.* [3] categorize WSNs as time-driven, event-driven, sink-initiated, and hybrid in terms of data delivery scheme. In the time-driven model, the sensors sense their data continuously at a prespecified rate and send it to the sink periodically. For event-driven and sink-initiated models, sensors report to the sinks only when a certain event occurs or when a request is initiated by the sink. They are well suited to time-critical applications. A hybrid model is defined as a combination of the above methods. In this paper, we consider time-driven sensor network applications such as environmental sensor networks for monitoring ecological habitats.

WSNs present a great opportunity for ecological monitoring that was not possible before due to the remoteness of areas of interest and/or infeasibility of in-person attendance in data collection. Use of WSNs does not only make eco-monitoring possible, but also facilitates more frequent data collection [4]. In addition, as habitat monitoring can be very sensitive to human presence, an unattended WSN provides a noninvasive approach to obtain real-time environmental data. For example, researchers from the University of California, Berkeley, CA, USA, and the College of the Atlantic, Bar Harbor, ME, USA, deploy sensors in Great Duck Island, ME, USA, to monitor the nesting burrows of Leach's Storm Petrels [5], [6]. The PODS project at the University of Hawaii at Manoa, Honolulu, HI, USA, deploys sensors in Volcanoes National Park on the Big Island of Hawaii to monitor the ecological environment and the events around the rare and endangered species of plants [7]. Other applications include studies in the Chihuahuan desert [8], monitoring icecaps and glaciers [9], forest monitoring [10], [11], and others as mentioned in [12]–[14].

The general framework of WSN network operations of our interest can be outlined as follows. Initially, a set of sensors, which are equipped with limited energy resource (e.g., battery) as well as sensing, processing, and communication capabilities, is deployed in a geographical region. Data collected by the sensors are forwarded to specially designated sensors, called clusterheads (CHs), which conduct some processing to aggregate their received data. CHs then forward the data to specific locations, called sinks, either directly or through other CHs. In this underlying setting, which is also depicted in Fig. 1, design of the network refers to the determination of CH and sink locations, while the operation decisions refer to the routing of data from the sensors to the sinks in that network.

Manuscript received November 27, 2011; revised September 16, 2012 and March 19, 2013; accepted May 05, 2013; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor M. Liu. Date of publication May 16, 2013; date of current version June 12, 2014. This work was supported by the National Science Foundation under Grant CMMI-0428831.

H. Lin was with Texas A&M University, College Station, TX 77843 USA. She is now with Bank of America, Dallas, TX 75202 USA.

H. Üster is with the Department of Industrial and System Engineering, Texas A&M University, College Station, TX, 77843 USA (e-mail: uster@tamu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2013.2262153

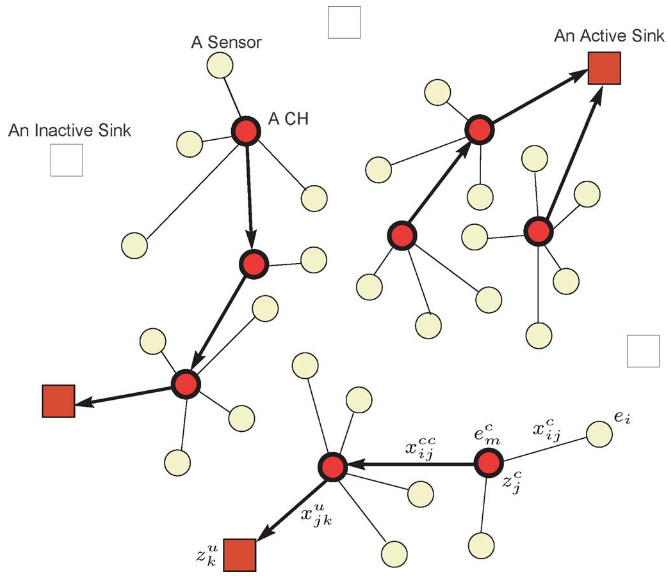

Fig. 1. Sample network, data flow, and notation.

Given limited energy levels at the sensors, as is the case in many applications of WSNs, one of the main concerns in network design and operation is the network lifetime, which we consider in this study to be the time between two sensor deployments. Sensor redeployments may be needed due to several reasons, e.g., having less than a critical number of operational sensors with enough remaining energy in the network. Typically, the lifetime is assumed to be divided into periods of uniform length, and for each period, network design and operations decisions are made in such a way that the number of periods in a deployment cycle is maximized. Consequently, the network lifetime is defined as the number of periods that can be achieved with a deployment.

Topology control and routing are two fundamental problems in effective design and operation of WSNs. The close relationship between these decisions and their relation to network lifetime are especially underlined by the WSN-specific design/operation attributes that include *energy efficiency* and *computation–communication tradeoff*. As mentioned, energy efficiency is a major concern since each sensor has finite and nonrenewable energy resource. Communication–computation tradeoff refers to the fact that communication consumes more energy than performing computations on board a sensor [15]. This is critical as it relates to the energy efficiency. Although the direct communication of a sensor with a sink is preferable for the overall network, this is impractical or, otherwise, leads to excessive energy use, thus shortening the network lifetime [16]. Therefore, routing schemes where the data size is decreased via in-network data aggregation (i.e., using energy for computation rather than communication) along the paths from sensors to a sink (user) are usually preferred.

Having defined the problems, operational attributes, and metrics discussed above, our purpose in this paper is to address the optimum design and operation of a WSN for a period within a deployment cycle. To this end, we consider an optimization approach to effectively integrate topology control and routing decisions in a cluster-based hierarchical network structure in which in-network data aggregation is facilitated for better energy efficiency.

This paper builds on the work described in [17]. In that study, the authors develop and examine three mathematical models whose objectives dictate alternative policies to be employed in each period of a deployment cycle for the purpose of maximizing the network lifetime. The first two models, the minimization of total energy usage in the network and the minimization of the maximum energy usage at a sensor, may face the issue of quick energy drainage, which occurs at certain sensors in the former and in the whole network in the latter. The third proposed model, the minimization of the weighted sum of the total energy usage and the range of remaining energy distribution in the network, improves energy efficiency and provides network lifetime that is significantly longer than the ones by the first two models as shown in experimental studies. However, in [17], the authors do not provide an exact solution method that can be utilized to improve and/or benchmark the solution quality of their heuristic algorithm.

In this paper, we study the third model in [17] further. Specifically, we consider two important extensions of this model.

- First, in the modeling context, we incorporate a total fixed cost term associated with CH selection into the objective function. By setting a higher fixed cost of usage (as a CH) for a sensor with low energy, the model attempts to avoid some well-positioned sensors from being selected as CHs repeatedly in successive periods and to protect these sensors from quick energy depletion. This approach also facilitates a uniform energy consumption profile at the sensors across the network. This is important because in a hierarchical setting, where data flow from sensors to the sinks occurs via CHs, a CH not only functions to capture information in its vicinity, but also as an aggregator/relay node to process and transfer the data generated by other sensors to the sinks. Thus, CHs consume more energy than regular sensors, while the whole network operation enjoys taking advantage of the computation–communication tradeoff.
- Second, observing that the model is amenable to exact solution via Benders decomposition (BD), in the methodological context, we focus our efforts on devising an efficient BD Algorithm as a solution method. In particular, we develop a solution approach that incorporates an effective heuristic algorithm and strengthened Benders cuts in an ε-optimal BD framework. Computational evidence demonstrates the efficient performance of the approach in terms of solution quality and time. In particular, our heuristic algorithm provides a good initial upper bound and facilitates the generation of initial Benders cuts, while the strengthened Benders cuts and ε-optimal framework accelerate the convergence of the BD algorithm.

The remainder of the paper is organized as follows. In Section II, we provide a review of the most related research in WSN design. In Section III, we present the details of the system model along with our assumptions, and in Section IV, we introduce a detailed problem definition and present the optimization model. The solution methodology is developed in Section V, and numerical results from a computational study are presented in Section VI. Finally, conclusions and future research directions are summarized in Section VII.

II. Related Literature

Clustering of sensors has been shown effective in prolonging sensor network lifetime in the literature. The basic idea is to organize WSNs into a set of clusters, and within each cluster, sensors transmit the collected data to their CHs. Each CH aggregates its received data and forwards it to the sink either directly or via relaying through other CHs. This is beneficial in terms of energy efficiency in three ways: 1) Hierarchical structure facilitates a multihop sensor-to-sink data transfer scheme which eliminates the quick energy drainage at the sensors that are away from the sink; 2) data aggregation is performed at the CHs to reduce data redundancy so that energy savings in communication are realized; and 3) periodic reclustering can balance the energy consumption by reassigning the CHs and the sinks and adjusting the routing in the network. Noting that the reviews on routing in WSNs include [31]–[33] and general reviews are given in [34], [35], and more recently in [36], in what follows, we specifically discuss the works that are more closely related to this research in the context of network topology and data routing.

Heinzelman *et al.* [18] develop a data aggregating cluster-based routing protocol Low Energy Adaptive Clustering Hierarchy (LEACH). In LEACH, they assume a single-hop CH-to-sink connection and adopt the randomized rotation of CHs to ensure a balanced energy consumption. However, such assumptions may not guarantee network connectivity. Younis and Fahmy [19] propose a hybrid energy-efficient distributed clustering routing (HEED) protocol where the CHs are probabilistically selected based on their remaining energy and the sensors join clusters such that the communication cost is minimized. HEED assumes a multihop connection between the CHs and to the sink. Liu *et al.* [20] suggest a distributed energy-efficient protocol EAP for the general setting in [19]. In EAP, each CH is probabilistically selected based on its ratio of the remaining energy to the average remaining energy of all the neighbor sensors within its cluster range. This is in contrast to HEED that only chooses CHs based on a sensors' own remaining energy. To further extend network lifetime, EAP introduces the idea of "intracluster coverage" that allows a partial set of sensors to be active within clusters while maintaining an expected coverage. Ademola *et al.* [37] also aim to promote a uniform energy usage across the network by minimizing the communication distance among sensors and selecting the CHs based on remaining energy at the sensors.

Since sensors generally send data to the sink in a "many-to-one" (convergecast) fashion, Haenggi [38] points out that some critical sensors closer to the sink appear on most forwarding paths in the network. Specifically, in a multihop cluster-based WSN, the CHs closer to the sink may have quick drainage due to their heavy load in forwarding data to the sink. We note that most of the above-mentioned studies do not explicitly take this factor into account. To ensure a balanced energy consumption, there are some studies, e.g., [21] and [39], that consider an unequal cluster-based routing scheme, i.e., CHs closer to the sink have smaller cluster sizes than those farther from the sink.

In contrast to the above studies that adopt a localized and/or protocol-based methods, Al-Karaki *et al.* [22] present a mathematical formulation by jointly considering the cluster-based routing problem with application-specific data aggregation. Al-Turjman *et al.* [23] propose a mixed integer linear program (MILP) with the objective of minimizing the total network energy consumption while including constraints on fault tolerance simultaneously. In that study, sensors are assumed to forward their data to the sink through specific relay nodes that are equipped with higher energy sources. In [24], a routing problem is considered for networks with flat topologies. For this, a linear programming approach is suggested to maximize the data flow per period. In another study, with similar assumptions but without considering data aggregation, a multicommodity flow approach is provided to maximize lifetime with the use of multiple sinks in a WSN [29]. Wang *et al.* [26] also consider a similar setting with mobile CHs thst are special high-energy sensors and examine the network lifetime under fixed sink location assumption. Efforts toward that end also include consideration of placing specific relay nodes with more energy [25]. Kim *et al.* [28] illustrate the benefit of employing multiple sinks and suggest a mixed integer linear program to determine sink locations. Luo and Hubaux [30] address a routing problem with sink mobility to improve network lifetime. Efficient approximation algorithms for generation of multihop routing trees (single sink) and forests (multiple sinks) are provided in [27]. However, in these studies, flat-routing structures without any CHs or aggregation are considered.

Table I summarizes a comparison of the key related works with both cluster-based and flat topology considerations. In our specific context of cluster-based approaches, we have the following observations: 1) Most of the studies in the clustered sensor network adopt a localized method to select and vary CHs over the periods. Such methods may be biased from the long-term network lifetime perspective. 2) In the majority of the literature, topology control and routing problems are handled separately, thus overlooking the interrelationships among them. 3) The majority of studies on cluster-based WSNs does not consider the use of multiple sinks with mobility. Therefore, we are motivated to investigate a generalized and integrated topology control and routing problem using optimization techniques. In doing so, we particularly consider a multiobjective optimization model that combines energy usage and remaining energy characteristics and that simultaneously considers cluster-based topology control and routing decisions in a multiple-sink WSN having a hierarchical network structure to facilitate data aggregation. For our mathematical model, we develop a solution algorithm based on joint use of Benders decomposition and an effective heuristic and analyze both algorithmic and network characteristics in an extensive computational study.

III. System Model

A. Network Model

In our problem of interest, sensors are deployed in a two-dimensional field, and the candidate sinks, which have no energy limitations, are located around the periphery of the sensor field. As in [17], each sensor is assumed to communicate its position, obtained via triangulation [40], [41], to the user in the beginning of a deployment cycle.

As sensors collect data, they form packages to forward to their CH based on a schedule. In our model, we assume a fixed sensor data generation rate that is also the rate of data forwarding from

TABLE I
RELATED FUNDAMENTAL WORK IN CLUSTER-BASED SENSOR NETWORKS

References	Model	Criterion	Sinks	Remark
Heinzelman et al. [18]	PR	EE	S	randomized rotation of CHs
Younis & Fahmy [19]	PR	EE	S	CHs selection based on residual energy
Liu et al. [20]	PR	EE	S	intra-cluster coverage
Ever et al. [21]	PR	EE	S	unequal clustering
Al-Karaki et al. [22]	OR	EE	S	application with specific data aggregation
Al-Turjman et al. [23]	OR	FT	S	specific relay nodes with more energy
Chang & Tassiulas [24]	OR	DF	S	flat topology
Misra et al. [25]	AA	EE	S	survivability requirements, high energy relays, flat topology
Wang et al. [26]	AA	EE	S	mobile high energy relays, flat topology
Wu et al. [27]	AA	EE	M	flat topology, tree and forest construction
Kim et al. [28]	OR	DF	M	sink locations and flat topology
Xue et al. [29]	OR	EE	M	multi-commodity flow network
Luo & Hubaux [30]	OR	EE	M	sink locations and flat topology
This work	OR	EE	M	multi-objective integrated topology and routing, sink and CH locations, data aggregation, hierarchical

Routing Model: PR (Protocol-based), AA (Approximation Algorithm), OR (Optimization-based)
Criterion: EE (Energy Efficiency), DF (Data Flow); FT (Fault Tolerance)
Sinks: M (Multiple), S (Single)

a sensor to its CH. We consider a multilayer hierarchical setting where data flows from sensors to the sinks either directly or through other CHs as depicted in Fig. 1. We assume that sensors are equipped with a dynamic transmit power level control that they utilize to achieve topology control based on the solutions obtained. Required number of CHs and sinks are specified *a priori*, and while we consider varying values in our computational study, we assume that at least one sensor acts as a CH and at least one of the sink locations is active.

Specific CH and sink locations are chosen from their associated candidate sets based on the solution of the mathematical formulation that also specifies routing of collected data from sensors to sinks via CHs. CHs are assumed to perform processing on their received data for eliminating redundancy and data aggregation. However, a CH does not further process the data it receives from another CH for which it serves only as a relay. Similar to [17] and [42], we employ an average aggregation ratio so that an overall view of the sensor field can be maintained after aggregation. This setting gives a dynamic topology where, at the end of each period, the energy information at the sensors is updated and the new CH and sink locations are to be determined for the next period.

B. Communication Model

We assume that a sensor can communicate with any sensor or sink within its vicinity that is assumed to be a disk centered at that sensor. We define the radius of such a disk as the transmission range. Although this model is widely used in the routing and topology control literature, we note that it does not represent actual hardware capabilities in real applications. More importantly, the model makes the assumption that a sensor can reliably communicate with any sensor within its transmission range. In this regard, a few comments are in order.

Since we consider data-gathering WSNs, we make the assumption that data generation and flow occur at a low rate, thus congestion is assumed not to be a source of unreliability. However, link reliability can also be due to the characteristics of the operating environment (e.g., obstacles, interference) or the hardware (e.g., antenna length). To incorporate these aspects, one approach is to estimate link reliability via a log-normal shadowing model [23], [43]. On the other hand, from an application perspective, a viable approach to incorporate link reliability is to conduct a neighborhood discovery in the network [44]–[46]. In this process, link quality indicators (LQIs) or received signal strength indicator (RSSI) can be utilized as measures of link reliability.

Then, to handle link reliability issues within our approach, which relies on mathematical modeling, we argue that a preprocessing can be incorporated to modify the input to optimization. More specifically, in the optimization model (Fig. 2), the $\mathbf{x}$ variables represent the data flow between sensors. If it is determined in the preprocessing that a pair of sensors cannot communicate at the least level of desired reliability, the corresponding flow variable can be set to zero, or it can just be excluded from the model (this also reduces the problem size and, thus, improves algorithmic performance as a by-product benefit). An example of handling such a situation is given in [17].

We further note that an implication of link reliability on network performance is loss of data and/or delayed transmission of data, which cause additional energy consumption due to changes in data transfer schedules and resends. While the link reliability issues in the context of a data-gathering WSN can also be addressed via deployment of dense networks with some level of redundancy built in, this comes at the expense of increased problem sizes as well as more involved operational planning due to concerns such as interference, data aggregation, scheduling, etc.

Model Parameters		*Decision Variables*	
$\mathcal{I}$	set of sensors, $i \in \mathcal{I}$,	x^c_{ij}	fraction of data flow per period from a sensor i to a CH j,
$\mathcal{J}$	set of candidate CHs, $j, m \in \mathcal{J},\ \mathcal{J} \subseteq \mathcal{I}$,	x^{cc}_{ij}	data flow per period from a CH i to a CH j,
$\mathcal{K}$	set of candidate sinks, $k \in \mathcal{K}$,	x^u_{jk}	data flow per period from a CH j to a sink k,
R_i	data generation rate (*bits/unit-time*) at a sensor i,	z^c_j	1 if a sensor j is setup as a CH, 0 o.w.,
D_{pq}	distance (*m*) between any two nodes p and q,	z^u_k	1 if a node k is setup as a sink, 0 o.w.,
w, v	energy dissipation rates for radio and amplifier, resp.,	e_i	energy consumed by a sensor i,
s, c	average aggregation ratio and dissipation rate, resp.,	e^c_m	energy consumed by a CH m,
C, U	number of required CHs and sinks, resp.,	E^R_{max}	maximum remaining energy at a sensor,
E_i	available energy at a sensor i,	E^R_{min}	minimum remaining energy at a sensor.
T	the length of a period.		

$$\textbf{(P)} \qquad \text{Min} \qquad t_1(1/|\mathcal{I}|)(\sum_{m\in\mathcal{J}} e^c_m + \sum_{i\in\mathcal{I}} e_i) + (E^R_{max} - E^R_{min}) + t_2(\sum_{m\in\mathcal{J}} z^c_m/E_m) \tag{1}$$

subject to

$$\sum_{k\in\mathcal{K}} (w + v\,D^2_{mk})\,T\,x^u_{mk} + \sum_{j\in\mathcal{J}\setminus\{m\}} (w + v\,D^2_{mj})\,T\,x^{cc}_{mj} + \sum_{j\in\mathcal{J}\setminus\{m\}} w\,T\,x^{cc}_{jm} + \sum_{i\in\mathcal{I}} (w + c\,s)\,R_i\,T\,x^c_{im} = e^c_m \qquad \forall\, m \in \mathcal{J} \tag{2}$$

$$\sum_{j\in\mathcal{J}} (w + v\,D^2_{ij})\,R_i\,T x^c_{ij} = e_i \qquad \forall\, i \in \mathcal{I} \tag{3}$$

$$\sum_{k\in\mathcal{K}} x^u_{mk} + \sum_{j\in\mathcal{J}\setminus\{m\}} x^{cc}_{mj} - \left(\sum_{j\in\mathcal{J}\setminus\{m\}} x^{cc}_{jm} + (1-s)\sum_{i\in\mathcal{I}} R_i\,x^c_{im}\right) = 0 \qquad \forall\, m \in \mathcal{J} \tag{4}$$

$$\sum_{j\in\mathcal{J}} x^c_{ij} = 1 \qquad \forall\, i \in \mathcal{I} \tag{5}$$

$$x^c_{ij} \le z^c_j \qquad \forall\, i \in \mathcal{I}, j \in \mathcal{J} \tag{6}$$

$$x^{cc}_{mj} \le \sum_{i\in\mathcal{I}} R_i z^c_j \qquad \forall\, m, j \in \mathcal{J} \tag{7}$$

$$x^u_{jk} \le \sum_{i\in\mathcal{I}} R_i z^u_k \qquad \forall\, j \in \mathcal{J}, k \in \mathcal{K} \tag{8}$$

$$x^c_{jk} \le \sum_{i\in\mathcal{I}} R_i z^c_j \qquad \forall\, j \in \mathcal{J}, k \in \mathcal{K} \tag{9}$$

$$\sum_{j\in\mathcal{J}} z^c_j = C \tag{10}$$

$$\sum_{k\in\mathcal{K}} z^u_k = U \tag{11}$$

$$e_i \le E_i \qquad \forall\, i \in \mathcal{I} \tag{12}$$

$$e^c_j \le E_j \qquad \forall\, j \in \mathcal{J} \tag{13}$$

$$z^c_j\,E_j - e^c_j \le E^R_{max} \qquad \forall\, j \in \mathcal{J} \tag{14}$$

$$(1 - z^c_i)\,E_i - e_i \le E^R_{max} \qquad \forall\, i \in \mathcal{I} \tag{15}$$

$$E^R_{min} \le E_i - e_i \qquad \forall\, i \in \mathcal{I} \tag{16}$$

$$E^R_{min} \le E_j - e^c_j \qquad \forall\, j \in \mathcal{J} \tag{17}$$

$$z^c_j,\ z^u_k \in \{0,1\} \qquad \forall\, i \in \mathcal{I},\ j \in \mathcal{J},\ k \in \mathcal{K} \tag{18}$$

$$x^c_{ij},\ x^{cc}_{ij},\ x^u_{jk},\ e_i,\ e^c_j,\ E^R_{max},\ E^R_{min} \ge 0 \qquad \forall\, i \in \mathcal{I},\ j \in \mathcal{J},\ k \in \mathcal{K} \tag{19}$$

Fig. 2. Notation and overall formulation $(\mathbf{P})$.

C. Energy Consumption Model

We employ a simple energy dissipation model that reflects the operational network characteristic of interest [18], [47]. To estimate the energy dissipation for transmitting x_{pq} (bits) of data from node p to node q, we use the path-loss model $v\,D^a_{pq}\,x_{pq}$ where v (J/bit/m^a) is a constant, D_{pq} is the distance (in meters) between p and q, and $2 \le a \le 4$. To calculate the power requirement at a receiver node q, we use the model $w\,x_{pq}$ where w (J/bit) is a constant. Therefore, transmitting x_{pq} (bits) of data from node p to node q dissipates $(w + v\,D_{pq}^{a})\,x_{pq}$, and receiving the same amount of data dissipates $w\,x_{pq}$. We assume that the radio dissipates $w = 50$ nJ/bit to run the transmitter or receive circuitry and $v = 100$ pJ/bit/m^2 for the transmit amplifier as in [18] and several others. In addition, we employ a dissipation rate of $c = 50$ pJ/bit for data aggregation/processing efforts at a CH. While we note that a higher a value is more appropriate to account for interference in uneven terrain, for computational purposes, we assume an a value of 2 (in our formulation in Fig. 2), which represents the ideal case in terms of the efficiency of communications. With higher a values, energy consumption levels are expected to increase due to increased path loss and, in turn, a reduced network lifetime is expected. We further note that our modeling and methodology framework presented in the paper is directly applicable without any modification even for different a values or with any other energy consumption model (e.g., [48]) since energy consumption model is an input as a function into our optimization approach [as in constraints (2) and (3) stated in Fig. 2]. We further note that it is possible to employ varying a values for different pairs of sensors based on the reliability estimate of their communication that can be determined as described in Section III-B.

D. Network Lifetime Model

In the literature, network lifetime is commonly defined as the number of periods in a deployment cycle. However, the definition of the end of a deployment cycle is typically application-specific. For example, recently, there are studies in which the end of a deployment cycle is said to be reached when a percentage of alive-and-connected relay nodes fall below a specific threshold level [10], [11], [49]. Since relay nodes are assumed to be special sensors and more expensive than the regular sensors, these studies assume an objective of maximizing connectivity under constraints on network lifetime and a minimum number of relay nodes. In our study, we consider cluster-based sensor networks where CHs are chosen from regular sensors as in [17]–[19]. The end of a deployment cycle is reached when some sensors fail to generate readings due to energy depletion and/or it is not possible to transmit the data generated in the network to the sinks. Thus, the purpose of prolonging network lifetime (via efficient use of network resources) corresponds to obtaining a maximum number of successive periods. In this case, simply minimizing the energy dissipation does not necessarily prolong the network lifetime as it may leave the sensors in the network with a wide disparity in energy levels [17], [19]. This eventually leads to heavy use of some sensors as CHs and their energy expiration results in the end of network lifetime. The goal of prolonging network lifetime needs to be achieved via reducing the energy consumption while ensuring a uniform energy depletion across the network.

IV. Problem Definition and Formulation

Given a set of sensors deployed and available at the beginning of a period, our problem is defined as the determination of the CHs and sinks to be employed and the data flow from sensors to sinks through CHs in that period. In doing so, the average energy usage and the variation in remaining energy distribution in the network is minimized while encouraging the adoption of high-energy sensors as CHs. We assume that any sensor can take the role of a CH while noting that if there are special high-energy sensors deployed for this role, our model can be used as is since its objective is already designed to promote high-energy sensors for selection as CHs. We also assume that a discrete set of points around the sensor field is designated as candidate sink locations. We provide the detailed notation and the overall formulation (P) for our problem in Fig. 2.

In this formulation, the objective function represents the weighted sum of three terms with weights t_1, one (taken as base weight), and t_2 for those terms, respectively. The first term is the average energy consumption that is minimized for better efficiency in energy usage. The second term gives the range of remaining energy levels, which is minimized to spread the energy usage (drainage) more uniformly across the sensors to promote prolonged operational network lifetime. The last term represents the fixed cost associated with locating the CHs. The last term effectively instills in the model the ability to protect low-energy sensors from being selected as CHs even if they are well-placed for this role in the network.

Constraint (2) provides the energy consumed by a CH node m, e_m^c, by adding the transmitter and receiver energy consumption implied by its interactions with other CHs, sinks, and sensors. Constraint (3), on the other hand, assigns the values of the total energy consumed by a sensor i transmitter. Specifically, note that the variable x_{ij}^c is nonzero if sensor i is planned to interact with any other sensor j that is acting as a CH. If j is not a CH, then z_j^c becomes zero and, by (6), x_j^c is assigned a value of zero. Also note that z_j^c would not arbitrarily be assigned a value of one in the solution as this would unnecessarily increase the objective value which is minimized. Constraint (4) states the data flow balance at each CH, and constraint (5) guarantees that each sensor is assigned to one CH. Constraints (6)–(9) assign the values of binary variables related to CH and sink location selections. Constraints (10) and (11) establish the required number of CHs and sinks, respectively. Constraint sets (12) and (13) ensure that the total energy consumed at a sensor cannot exceed the total available energy at the corresponding sensors. Constraint sets (14) and (15) give the maximum remaining energy at a sensor, and constraint sets (16) and (17) give the minimum remaining energy at a sensor. Finally, (18) and (19) include the integrality and nonnegativity of the decision variables.

V. Solution Approach Based on Benders Decomposition

Benders decomposition [50] is a solution approach for mixed integer linear programming problems such as the one in Fig. 2, and it has been successfully employed for solving a wide array of large-scale optimization problems. This technique is based on the idea of exploiting the special structure of the problem at hand; it separates the original formulation into two smaller easier-to-solve problems called a *master problem* and a *subproblem*.

The master problem accounts for all the integer variables and the associated portion of the objective function and the constraints of the original problem. It also embodies the information regarding the subproblem portion of the problem via use of an additional (continuous) auxiliary variable and a set of constraints called *Benders cuts*. On the other hand, the subproblem includes all continuous variables and the associated constraints

$$\text{Max}\quad Z_{\text{DSP}} = \sum_{i\in\mathcal{I}}\beta_i - \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\hat{z}_j^c\gamma_{ij} - \sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}\setminus\{j\}}(\sum_{i\in\mathcal{I}}R_i\hat{z}_j^c)\,\delta_{jm} + \sum_{j\in\mathcal{J}}E_j\{\hat{z}_j^c\,\theta_j - \sigma_j - \tau_j\}$$
$$+ \sum_{i\in\mathcal{I}}E_i\{(1-\hat{z}_i^c)\,\eta_i - \pi_i - \rho_i\} - \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}(\sum_{i\in\mathcal{I}}R_i)\{\hat{z}_k^u\,\lambda_{jk} + \hat{z}_j^c\,\mu_{jk}\} \tag{20}$$

subject to

$$(w+v\,D_{jk}^2)\,T\,A_j + \alpha_j - \lambda_{jk} - \mu_{jk} \le 0 \qquad \forall j\in\mathcal{J},\ k\in\mathcal{K} \tag{21}$$

$$(w+v\,D_{jm}^2)\,T\,A_j + w\,T\,A_m + \alpha_j - \alpha_m - \delta_{jm} \le 0 \qquad \forall j,\, m\in\mathcal{J} \tag{22}$$

$$(w+c\,s)R_i\,T\,A_j + (w+v\,D_{ij}^2)R_i\,T\,B_i - (1-s)\,R_i\,\alpha_j + \beta_i - \gamma_{ij} \le 0 \qquad \forall i\in\mathcal{I},\ j\in\mathcal{J} \tag{23}$$

$$-A_j + \theta_j - \sigma_j - \tau_j \le t_1(1/|\mathcal{I}|) \qquad \forall j\in\mathcal{J} \tag{24}$$

$$-B_i + \eta_i - \pi_i - \rho_i \le t_1(1/|\mathcal{I}|) \qquad \forall i\in\mathcal{I} \tag{25}$$

$$\sum_{j\in\mathcal{J}}\theta_j + \sum_{i\in\mathcal{I}}\eta_i \le 1 \tag{26}$$

$$\sum_{j\in\mathcal{J}}\sigma_j + \sum_{i\in\mathcal{I}}\pi_i \ge 1 \tag{27}$$

$$\gamma_{ij},\,\delta_{jm},\,\lambda_{jk},\,\mu_{jk},\,\theta_j,\,\eta_i,\,\pi_i,\,\sigma_j \ge 0 \qquad \forall i\in\mathcal{I},\ j,m\in\mathcal{J},\ k\in\mathcal{K} \tag{28}$$

$$A_j,\,B_i,\,\alpha_j,\,\beta_i\ \text{unrestricted} \qquad \forall i\in\mathcal{I},\ j\in\mathcal{J} \tag{29}$$

Fig. 3. Benders dual subproblem.

in the original problem. Solving the dual of the subproblem provides information about the subproblem portion of the original objective function, and this information is communicated to the master problem via Benders cuts.

A. Base Benders Decomposition Approach

In each iteration of Benders algorithm, the master problem is resolved to optimality with the addition of a Benders cut. This gives a lower bound for the original problem (P), and values for the integer variables are then substituted into the subproblem. The dual subproblem is then solved to produce an upper bound for (P) and a set of dual variables values that are used to generate a new Benders cut for the master problem in the next iteration. This process is repeated until a termination condition, usually a small optimality gap between the lower bound and the upper bound, is met. In Benders approach, it is known that if the iterations are allowed to continue long enough, an optimal solution is obtained as the Benders cuts recover the complete feasible polyhedron of the overall problem. Although, theoretically, this is not efficient as the number of Benders cuts is exponential in problem size, in practice, a very good optimality gap can be obtained if the algorithm is designed carefully with problem-specific enhancements as we develop for our problem.

Our formulation employs the binary variables z_j^c and z_k^u ($\mathbf{z}$ for brevity) associated with CH and sink selection, continuous variables x_{ij}^c, x_{ij}^{cc}, and x_{jk}^u ($\mathbf{x}$ for brevity) for routing decisions and energy related variables e_i, e_m^c, $E_{\max}^R$, and $E_{\min}^R$ ($\mathbf{e}$ for brevity). The structure of our problem presents a natural decomposition scheme for the Benders approach: For fixed $\mathbf{z}$ values in (P), we obtain a routing problem that is an efficiently solvable linear program. The master problem is obtained by excluding routing related decisions and constraints in (P), and it is an integer program that involves much smaller numbers of variables and constraints than the problem (P). Therefore, at each iteration, the solution of the master problem gives a tentative network configuration. i.e., the selection of CH and sink locations as specified by the $\mathbf{z}$ variable values and the subproblem provides the optimal data routing $\mathbf{x}$ and energy values $\mathbf{e}$ under that fixed configuration. Also, the master problem and the subproblem provide information to obtain lower and upper bounds on the objective value of the original problem, respectively. This is in contrast to the heuristic methods that only give feasible solutions and cannot guarantee a solution quality. In each iteration of the algorithm, a new Benders cut is added to the master problem by using the dual subproblem solution; the lower bound is therefore nondecreasing.

Next, we describe each component of the BD framework as well as its algorithmic enhancements in detail. Later, we provide the overall algorithm and the specifics of its each iteration completed to solve (P) in Section V-E and Fig. 6.

B. Benders Subproblem and Its Dual

The subproblem, denoted as $\mathrm{SP}(\mathbf{x},\mathbf{e}\,|\,\hat{\mathbf{z}})$, can easily be obtained from the overall formulation (P) as follows: Third term in the objective function is excluded; constraints (10) and (11) are excluded; and the $\mathbf{z}$ values are fixed at $\hat{\mathbf{z}}$ in the rest of the constraints that are in $\mathrm{SP}(\mathbf{x},\mathbf{e}\,|\,\hat{\mathbf{z}})$. Intuitively, for given binary variables $\hat{\mathbf{z}}$ associated with fixed CH and sink locations whose locations are known as dictated by the master problem, the subproblem $\mathrm{SP}(\mathbf{x},\mathbf{e}\,|\,\hat{\mathbf{z}})$ is essentially a linear minimization problem that determines the data routing scheme from sensors to sinks via CHs and energy usage/status in the network.

In the Benders framework, rather than solving the subproblem (primal), we solve its dual, which is denoted as $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$ stated in Fig. 3 for our specific problem. At optimality, the objective values of the dual and the primal subproblems are equivalent due to the duality theorem. To develop the dual of $\mathrm{SP}(\mathbf{x}, \mathbf{e} \,|\, \hat{\mathbf{z}})$ given in Fig. 3, we define the dual variables A_j, B_i, α_j, β_i, γ_{ij}, δ_{jm}, λ_{jk}, μ_{jk}, ρ_i, τ_j, θ_j, η_i, π_i, and σ_j corresponding to the constraints (2)–(9) and (12)–(17), respectively.

After solving the dual subproblem $\mathrm{DSP}(\cdot \,|\, \hat{\mathbf{z}})$, following Benders cuts (BCuts) are generated using $\mathrm{DSP}(\cdot \,|\, \hat{\mathbf{z}})$'s objective function (20), the values of dual variables, and an auxiliary continuous variable B that actually refers to the optimum value of a subproblem's objective function

$$\begin{aligned} B \geq & \sum_{i\in\mathcal{I}} \hat{\beta}_i - \sum_{j\in\mathcal{J}} E_j\left(\hat{\sigma}_j + \hat{\tau}_j\right) + \sum_{i\in\mathcal{I}} E_i\left(\hat{\eta}_i - \hat{\pi}_i - \hat{\rho}_i\right) \\ & - \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \hat{\gamma}_{ij} z_j^c - \sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}\setminus\{j\}} \left(\sum_{i\in\mathcal{I}} R_i\, \hat{\delta}_{jm}\, z_j^c\right) \\ & + \sum_{j\in\mathcal{J}} E_j \hat{\theta}_j\, z_j^c - \sum_{i\in\mathcal{I}} E_i\, \hat{\eta}_i\, z_i^c \\ & - \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} \left(\sum_{i\in\mathcal{I}} R_i\right) \left\{\hat{\lambda}_{jk}\, z_k^u + \hat{\mu}_{jk}\, z_j^c\right\}. \end{aligned} \tag{30}$$

C. Benders Master Problem

The master problem $\mathrm{MP}(\mathbf{z} \,|\, \cdot)$ can be stated as in Fig. 4, which is obtained from the overall formulation (P) by adopting its third term in the objective function and the requirements on the number of CHs and sinks given by constraints (10) and (11). The real-valued variable is contained in the set of Benders cuts (34) as given in (30). The master problem $\mathrm{MP}(\mathbf{z} \,|\, \cdot)$ is essentially a minimization problem that gives a tentative network configuration, selection of CH and sink locations, and a lower bound of the original model.

At each iteration, we obtain a new dual solution of $\mathrm{DSP}(\cdot \,|\, \hat{\mathbf{z}})$, substitute it into constraint (30), add it to the $\mathrm{MP}(\mathbf{z} \,|\, \cdot)$, and then resolve the master problem to obtain a new set of values of the binary variables z_j^c and z_k^u.

D. Approaches for Accelerating the BD Algorithm

We observe that the direct implementation of classical BD approach in our model often converges slowly. This is due to the following reasons.

1) In the absence of a set of dual variables, BD approach starts the iterative procedure by solving the master problem without any Benders cuts (34). However, the initial selection of cuts can have a profound effect upon the performance of Benders algorithm [51].
2) Due to the degeneracy of the subproblem $\mathrm{SP}(\mathbf{x}, \mathbf{e} \,|\, \hat{\mathbf{z}})$, there exists multiple dual optimal solutions for $\mathrm{DSP}(\cdot \,|\, \hat{\mathbf{z}})$. This means that multiple sets of dual values are possible to provide the same optimum solution to the dual subproblem. Thus, it is important to obtain an optimal solution to $\mathrm{DSP}(\cdot \,|\, \hat{\mathbf{z}})$ so that a stronger cut of the form (30) is generated.
3) The master problem $\mathrm{MP}(\mathbf{z} \,|\, \cdot)$ must be solved each iteration a new Benders cut (30) is added in (34). Thus, as the number of iterations increases, the complexity and the size of $\mathrm{MP}(\mathbf{z} \,|\, \cdot)$ increases dramatically, and, consequently, solving $\mathrm{MP}(\mathbf{z} \,|\, \cdot)$ becomes very time-consuming.

$$\text{Min} \quad Z_{\text{MP}} = t_2\Big(\sum_{j\in\mathcal{J}} z_j^c / E_j\Big) + B \tag{31}$$

subject to

$$\sum_{j\in\mathcal{J}} z_j^c = C \tag{32}$$

$$\sum_{k\in\mathcal{K}} z_k^u = U \tag{33}$$

$$\text{Benders cuts in the form of (30)} \tag{34}$$

$$z_j^c,\, z_k^u \in \{0,1\},\ B \geq 0 \qquad \forall\, j \in \mathcal{J}, k \in \mathcal{K} \tag{35}$$

Fig. 4. Benders master problem.

1: **initialize** *Maxiter*, Ψ, g, $Z(\mathcal{S}^b) = \infty$;
2: **while** *Maxiter*> 0 and $G^* > 0$ **do**
3: $\quad\mathcal{I}^R = \{i \in \mathcal{I} : E_i \geq TH_\Psi\}$;
4: $\quad\mathcal{J} = \mathcal{I}^R$, solve model (P) with `TiLim` to obtain $Z(\mathcal{S}^c)$;
5: $\quad G^* = Z(\mathcal{S}^b) - Z(\mathcal{S}^c)$;
6: $\quad$**if** $G^* > 0$ **then**
7: $\qquad\mathcal{S}^b = \mathcal{S}^c$, $Z(\mathcal{S}^b) = Z(\mathcal{S}^c)$;
8: $\quad$**end if**
9: $\quad\Psi = max\{0, \Psi - g\}$;
10: $\quad$*Maxiter* = *Maxiter*-1;
11: **end while**
12: **return** $\mathcal{S}^b$ and $Z(\mathcal{S}^b)$

Fig. 5. Upper bound heuristic (UBH).

In order to circumvent these difficulties, we explore several techniques to accelerate the convergence of the BD algorithm as discussed next.

1) Upper Bound Heuristic Algorithm: We devise an efficient heuristic algorithm, called Upper Bound Heuristic (UBH), that provides a feasible solution to overall problem (P) without much computational effort. The aim of our heuristic algorithm is to find a good upper bound and facilitate the generation of good initial Benders cuts. We use the solution, specifically CH and sink selections given by $\mathbf{z}$, obtained from the heuristic as an input and solve the dual subproblem $\mathrm{DSP}(\cdot \,|\, \hat{\mathbf{z}})$ for generating an initial Benders cut so that it can be added to the master problem in the following iteration. This is in contrast to initially solving the $\mathrm{MP}(\mathbf{z} \,|\, \cdot)$ without any cuts in a typical BD implementation.

We design the heuristic in a way to avoid well-positioned sensors being selected as CHs repeatedly in successive periods and to protect low-energy sensors from being selected as CHs. For this purpose, we consider only a subset of sensors, $\mathcal{I}^R \subset \mathcal{I}$, with high energy as the set of candidate CHs $\mathcal{J}$. In particular, to determine the $\mathcal{I}^R$ set, we use a threshold value TH_Ψ calculated as Ψ% of the average initial energy level at the sensors, i.e., $TH_\Psi = (\Psi/100) * (\sum_{i\in\mathcal{I}} E_i/|\mathcal{I}|)$ and $\mathcal{I}^R = \{i \in \mathcal{I} : E_i \geq TH_\Psi\}$.

In the UBH, given in Fig. 5, we proceed as follows. First, we note that its core algorithm (lines 3–10) works in an iterative fashion. At each iteration, we determine the set $\mathcal{I}^R$ based on a threshold value TH_Ψ (line 3); solve the model (P) assuming $\mathcal{J} = \mathcal{I}^R$ (line 4). Thus, we obtain the current solution $\mathcal{S}^c$ represented by the CH and sink selections $\mathcal{C} = \{j \in \mathcal{I} : z_j^c = 1\}$

```
initialize algorithmic parameters for UBH,
  Iterno = 0;
Apply UBH to S^b and record Z(S^b);
Set UB= Z(S^b) and z̄_best = S^b;
Solve DSP(·|ẑ) to obtain the values for all dual variables;
Generate the initial Benders cut and incorporate it into MP(z|·);
Solve MP(z|·) to obtain the values for ẑ and Z_MP;
while MP(z|·) has a feasible solution do
    Iterno = Iterno + 1;
    Solve DSP(·|ẑ) to obtain the values for
    all dual variables and Z_DSP.
    if Z_MP − B + Z_DSP < UB then
        UB = Z_MP − B + Z_DSP;
        ẑ_best = ẑ;
        Update the incumbent value UB in constraint (37);
    end if
    Generate BCuts and incorporate them into MP(z|·);
    Solve MP(z|·) to obtain the value for ẑ and Z_MP.
end while
Solve SP(x, e|z̄_best) for all continuous variables x̄_best and ē_best;
Return (x̂_best, ê_best, ẑ_best) and the UB.
```

Fig. 6. ε-optimal BD algorithm.

and $\mathcal{D} = \{k \in \mathcal{K} : z_k^u = 1\}$, respectively, along with the objective value $Z(\mathcal{S}^c)$ (line 4). While solving the model $(\mathbf{P})$, we employ a stopping criterion given by a `TiLim` (CPLEX parameter) time limit to alleviate the problem of excessive runtimes. If an improvement G^* (line 5) over the best solution $\mathcal{S}^b$ is obtained, then $\mathcal{S}^c$ becomes the new $\mathcal{S}^b$ (lines 6–8). We decrease Ψ by a parameter value g (line 9), update the set $\mathcal{I}^R$, and then resolve the problem (P). The algorithm terminates when no improving solution is found or it reaches the maximum iteration *Maxiter*.

This procedure is very effective in terms of solution quality and serves the purpose of generating the initial Benders cut with inexpensive computational times. As illustrated later in Section VI, combining the upper bound heuristic and BD framework promotes faster convergence, especially for larger instances.

2) Strengthening the Benders Cuts: Due to the degeneracy of the subproblem $\mathrm{SP}(\mathbf{x}, \mathbf{e} \,|\, \hat{\mathbf{z}})$, there exist multiple dual optimal solutions for $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$, each defining a different Benders cut; some cuts are stronger (reduce the solution space more effectively) than the others. Hence, it is important to identify the optimal dual solution corresponding to a stronger Benders cut of the from (30). Magnanti and Wong [51] define the strongness of a Benders cut for a general optimization problem given by $\mathrm{Min}_{y \in Y, z \in R}\{z : z \geq f(u) + y\, g(u), \forall u \in U\}$ as follows: The cut $z \geq f(u^1) + yg(u^1)$ dominates or is stronger than the cut $z \geq f(u) + y\, g(u)$ if $f(u^1) + yg(u^1) \geq f(u) + y\, g(u), \forall u \in U$ with a strict inequality for at least one point $y \in Y$. The use of the strengthened Benders cuts can facilitate better lower bounds and increase the algorithm efficiency, as shown for various problem settings in [51]–[54].

For our problem, we adopt a two-phase approach to strengthen the Benders cuts [52], [54]. This is based on the observation that, in the DSP objective (20), if $\hat{z}_j = 0$, one can modify its associated dual variable values (e.g., γ_{ij} in the second term) without changing the optimal objective function value Z_{DSP}, provided that the feasibility with respect to (21)–(28) is maintained. Recall that the DSP's objective is directly employed to devise a Benders cut (30).

Specifically, in the first phase, by solving $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$, we only obtain the values of the dual variables for which the associated binary variables $\hat{z}_j^c$ and $\hat{z}_k^u$ have values equal to 1. Hence, the elimination of the remaining dual variables (whose associated $\mathbf{z}$ values are zero) in the first phase cannot affect the objective function value (20), and we obtain a partial optimal solution. In the second phase, we fix the values of the dual variables obtained from the first phase and solve for other dual variables using a modified version of $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$ given in (36). The detailed description of two-phase approach is given as follows.

In Phase I, we only obtain the values of the dual variables for which the associated binary variables $\hat{z}_j^c$ and $\hat{z}_k^u$ have values equal to 1. We denote the reduced $\mathcal{J}$ set $\mathcal{J}_R$ and the reduced $\mathcal{K}$ set $\mathcal{K}_R$. i.e., $\mathcal{J}_R = \{j \in \mathcal{J} : \hat{z}_j^c = 1\}$ and $\mathcal{K}_R = \{k \in \mathcal{K} : \hat{z}_k^u = 1\}$. We also consider a similarly defined set $\mathcal{I}_R$ in (25) and denote $\mathcal{I}_R$ as a set of sensors not being selected as CHs, i.e., $\mathcal{I}_R = \mathcal{I} \backslash \mathcal{J}_R$. We solve the dual subproblem $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$ after this reduction with sets $\mathcal{J}_R$, $\mathcal{K}_R$, and $\mathcal{I}_R$ in (25). By doing so, we can solve a reduced version of $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$ without changing the objective function value.

In Phase II, we focus on computing the dual variables for which the associated binary variables $\hat{z}_j^c$ and $\hat{z}_k^u$ have values equal to 0. To this end, we solve the following linear programming problem:

$$\begin{aligned} \mathrm{Max} \quad & \sum_{j \in \mathcal{J}} E_j \theta_j - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \gamma_{ij} - \sum_{j \in \mathcal{J}} \sum_{m \in \mathcal{J} \backslash \{j\}} \left(\sum_{i \in \mathcal{I}} R_i \right) \delta_{jm} \\ & - \sum_{i \in \mathcal{I}} E_i\, \eta_i - \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} \left(\sum_{i \in \mathcal{I}} R_i \right) \{\lambda_{jk} + \mu_{jk}\} \\ & \text{subject to} \quad (21)\text{–}(29). \end{aligned} \tag{36}$$

Note that, in problem (36), the objective function represents the sum of all the coefficient associated with $\hat{z}_j = 0$ given in the Bender cut (30) and the constraints are the same as $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$. By doing so, we aim to obtain the values of the remaining variables in (21)–(27) so as to generate a stronger cut. Also, in order not to affect the objective function value in $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$, the values of the dual variables associated with $\hat{z}_j = 1$ in Phase II need to remain the same as in Phase I. Specifically, the values of the dual variables found in Phase I are substituted in the problem (36). Then, we solve the problem (36) to obtain the values of the dual variables for which the associated binary variable values $\hat{z}_j^c$ and $\hat{z}_k^u$ are zero.

The complete set of dual variable values obtained after the application of two phases in this way are then utilized to generate a Benders cut (30).

3) ε-Optimal Approach: As mentioned before, in the basic BD algorithm, we add a new Benders cut into the master problem $\mathrm{MP}(\mathbf{z} \,|\, \cdot\,)$ at each iteration. Thus, as the number of iterations increases, the complexity and the size of $\mathrm{MP}(\mathbf{z} \,|\, \cdot\,)$ increase dramatically, and, consequently, $\mathrm{MP}(\mathbf{z} \,|\, \cdot\,)$ becomes more difficult to solve. In order to decrease the solution time of $\mathrm{MP}(\mathbf{z} \,|\, \cdot\,)$, we utilize the ε-optimal approach introduced in [55]. Specifically, we add one additional constraint to the $\mathrm{MP}(\mathbf{z}\,|\,\cdot\,)$ given as

$$t_2\left(\sum_{j\in\mathcal{J}} z_j^c/E_j\right) + B \leq UB(1-\varepsilon) \tag{37}$$

where UB and ε denote the best upper bound and the acceptable optimality gap in solving (P), respectively. Then, while solving the $\mathrm{MP}(\mathbf{z}\,|\,\cdot\,)$ in an iteration of the ε-optimal BD algorithm, we seek for a feasible solution rather than an optimum solution. By doing so, the runtime for the $\mathrm{MP}(\mathbf{z}\,|\,\cdot\,)$ can be substantially reduced at each iteration. Using the values of the $\mathbf{z}$ variables given by this feasible solution, we then solve $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$ and generate a new Benders cut.

Note that the feasible solution obtained from $\mathrm{MP}(\mathbf{z}\,|\,\cdot\,)$ is no longer a valid lower bound for the problem. The ε-optimal Benders algorithm terminates when $\mathrm{MP}(\mathbf{z}\,|\,\cdot\,)$ cannot produce a feasible solution, thus verifying that the best upper bound is within ε from optimality.

E. Overall ε-Optimal BD Framework

In order to improve the computational efficiency, our algorithm brings together the above components including the UBH and strengthening of the Benders cuts in an ε-optimal BD framework presented in Fig. 6. Note that *Iterno*, *UB*, *LB*, and $(\hat{\mathbf{x}}_{\mathbf{best}}, \hat{\mathbf{e}}_{\mathbf{best}}, \hat{\mathbf{z}}_{\mathbf{best}})$ denote the number of iterations, the best upper bound, the best lower bound, and the best feasible solution, respectively.

In particular, we first apply the UBH to obtain a feasible solution (an upper bound) and solve the dual subproblem $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$ for generating an initial strengthened Benders cut so that it can be added to the master problem $\mathrm{MP}(\mathbf{z}\,|\,\cdot\,)$ in the beginning (lines 1–6).

Then, the while loop, lines 7–17, implements the ε-optimal approach. Specifically, using the values of the $\mathbf{z}$ variables given in the current feasible MP solution, we solve $\mathrm{DSP}(\,\cdot\,|\,\hat{\mathbf{z}})$ to obtain a feasible (upper bound) solution (lines 8–12). Notice that Z_{MP} includes the auxiliary variable (B) value, thus we calculate an upper bound by first adjusting the master objective solution so that it represents only the last term of the (P)'s objective and then add to it the objective of the dual solution representing the first two terms of the (P)'s objective. The best upper bound UB and the best solution $\hat{\mathbf{z}}_{\mathrm{best}}$ are updated if improved by the current solution $\hat{\mathbf{z}}$ (line 13), and a new iteration is started. Then, we generate a new strengthened Benders cut (line 15), add it to MP, and solve $\mathrm{MP}(\mathbf{z}\,|\,\cdot\,)$ to obtain a feasible solution so that a new iteration can be started in line 8. The algorithm terminates when the $\mathrm{MP}(\mathbf{z}\,|\,\cdot\,)$ cannot find a feasible solution, which verifies the best upper boundUB is within ε from the optimal solution. Once the iterations are completed, we solve the subproblem $\mathrm{SP}(\mathbf{x}, \mathbf{e}\,|\,\hat{\mathbf{z}}_{\mathrm{best}})$ to obtain the values of continuous variables (line 18). Upon the completion of the algorithm, we report the best feasible solution along with the best upper bound (line 19).

VI. Computational Study

In this section, we first conduct a computational study to establish the performance of Benders decomposition algorithm in a single-period setting by generating several test instances of varying problem sizes. The results also illustrate the benefit of utilizing the upper bound heuristic, strengthened Benders cuts, and ε-optimal framework for Benders decomposition. Second, using a sample network generated in a similar fashion, we analyze the effects of the modeling and algorithmic approaches on the network energy profile, configuration, and routing aspects. The computational experiments are performed on a machine with two 2.66-GHz Intel XEON processors and 12.0 GB RAM, and the algorithms are implemented in C++ utilizing Standard Template Library (STL) and Concert Technology when CPLEX 11 was used.

TABLE II
Problem Settings Used in Computational Testing

Setting I – Small instances					Setting II – Large instances				
Cls	$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{K}\rvert$	C	U	Cls	$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{K}\rvert$	C	U
S1	50	8	3	1	L1	150	8	9	1
S2				2	L2				2
S3			6	1	L3			18	1
S4				2	L4				2
S5		16	3	1	L5		16	9	1
S6				2	L6				2
S7			6	1	L7			18	1
S8				2	L8				2
S9	75	8	5	1	L9	200	8	12	1
S10				2	L10				2
S11			9	1	L11			24	1
S12				2	L12				2
S13		16	5	1	L13		16	12	1
S14				2	L14				2
S15			9	1	L15			24	1
S16				2	L16				2
S17	100	8	6	1	L17	250	8	15	1
S18				2	L18				2
S19			12	1	L19			30	1
S20				2	L20				2
S21		16	6	1	L21		16	15	1
S22				2	L22				2
S23			12	1	L23			30	1
S24				2	L24				2

A. Settings for Computational Study

For computational testing, we consider our test instances with a wide range of input data values for the problem parameters. As shown in Table II, we generate a total of 48 problem classes in two settings, Setting I—Small instances and Setting II—Large instances, by varying the number of sensors $|\mathcal{I}|$, the number of candidate sinks $|\mathcal{K}|$, the number of required CHs C, and the number of required sinks U. For Setting I, we consider $|\mathcal{I}|$ values of 50, 75, and 100. For Setting II, we consider $|\mathcal{I}|$ values of 150, 200, and 250. Similar to the settings in [17], for each value of $|\mathcal{I}|$, we consider two levels for $|\mathcal{K}|$ as 8 and 16; two levels for C as 6% and 12% of $|\mathcal{I}|$; and two levels for U as 1 and 2.

For each of the 48 problem classes, we generate 10 random instances. Sensor locations are generated randomly based on uniform distribution in a square of size N (meters) sensor field. Coordinates of $|\mathcal{K}|$ potential sink location are also randomly generated based on uniform distribution along the periphery of the sensor field. We consider N values of 100 for Setting I and 200 for Setting II. Note that all of the input and algorithmic parameter values are set as mentioned previously unless stated otherwise.

Similar to the settings in [17] and [19], we assume that the initial energy levels E_i at the sensors are uniformly distributed in the range [0.1, 0.5] J in all of the instances. We note that the energy levels at the sensors change differently at each period within a deployment cycle. Thus, although sensor energy levels are expected to be very similar at the first period of a deployment cycle, typically in the following periods within the same cycle, initial energy levels are expected to be varied. It must also be noted that measuring or estimating E_i at the beginning of a period is very difficult in practice, thus we assume that it is simply calculated based on initial energy and energy consumption at the sensor in that period.

Other parameters are set as in [17], specifically, we set the period length T as 4000 time-units; the average aggregation ratio s as 0.3; data generation rate R_i as 10 bits/time-unit; and the weight t_1 as 5. In order to determine a reasonable value of CH fixed cost coefficient (weight) t_2, we solve a set of small size instances to optimality using varying values of t_2. After observing the effects of changing t_2 values on the algorithmic performance and the network topology—specifically, by seeking faster runtimes to reach a relatively robust set of CH and sink selections—we set t_2 as 0.01.

Finally, for algorithmic parameters Ψ, g, and *Maxiter*, introduced in Section V-D, we note the following. Since the upper bound heuristic works in an iterative fashion, we set the maximum iteration to 30 so as to save the computational time. Also, we consider only a subset of sensors with higher energy as the set of candidate CHs. At the beginning, we assume initial $\Psi = 140$, i.e., only the sensors with the remaining energy greater or equal to 140% of the average remaining energy of all the sensors can be selected as CHs. For next iteration, we decrease Ψ by $g = 10\%$ for enlarging the search space with the attempt to find a better solution in a relatively short time. We conduct an empirical study by varying values of *Maxiter*, initial Ψ, and g and observe that the above values work fine in terms of faster runtime and good CH selections.

B. Results on Algorithmic Performance

We evaluate the performance of Benders decomposition algorithm (Fig. 6) on the basis of solution quality and time under an optimality gap (ε) of 2% in single-period settings. In addition, we evaluate the performance of our upper bound heuristic (Fig. 5) via utilizing two different benchmarks.

1) For Setting I, we utilize the exact solutions for benchmarking. We obtain the optimal solution for model (P) by using the exact branch-and-cut implementation in CPLEX.
2) For Setting II, we obtain the lower bound from the BD approach as another benchmark solution to evaluate the effectiveness of the heuristic algorithm.

Main results for Setting I (small instances, S1–S24) are summarized in Fig. 7, which shows the average solution times (over 10 instances, in seconds) for three alternative approaches to solve the problem (P). For each instance, the lower part of the corresponding bar gives the solution time when branch-and-cut (via CPLEX) is used to solve (P) to 2% optimality gap. The mid part reports the same for the heuristic approach (Fig. 5), and the top part is for the ε-optimal Benders algorithm (Fig. 6). We note that the optimal solution times for solving (P) using CPLEX are not reported as they are significantly longer than the runtimes for obtaining solutions with CPLEX with 2% optimality gap.

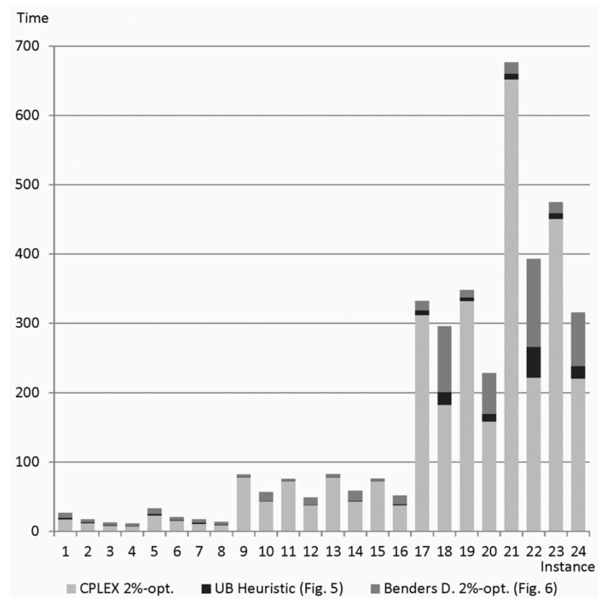

Fig. 7. Solution times for Setting I: small instances S1 to S24.

For fair comparison purposes between CPLEX and the BD approach, we report each approach's runtimes for a termination criterion of 2% optimality gap.

It is clear that the runtimes with the heuristic approach are significantly lower than runtimes with the other approaches (for classes S1–S16, the corresponding band in the bars are very narrow). In addition, the BD approach provides significantly lower runtimes when compared to CPLEX runtimes. In terms of solution quality of the heuristic approach, we calculate the optimality gaps as $\Delta^{\mathrm{H}} = 100 * (Z^{\mathrm{H}} - Z^{\mathrm{O}})/Z^{\mathrm{O}}$ for each instance, where Z^{O} and Z^{H} represent the optimal (more specifically, 10^{-4}%-optimal) objective function value and heuristic value for an instance, respectively. The mean and the range of the average optimality gaps over 24 small instances are extremely low, given as 0.19% and 0.46 (with a maximum of 0.51% and a minimum of 0.04%), respectively. This illustrates that the upper bound heuristic approach provides near-optimal solutions.

In Fig. 8, we summarize the computational results for Setting II (large instances, L1–L24). As the problem size increases, CPLEX runtime increases drastically. Even for 2% optimality gap, large instances cannot be solved with CPLEX within reasonable times. Thus, in Fig. 8, we only report results for the ε-optimal BD approach and the heuristic algorithm. While the solution times with the heuristic approach are significantly lower, the BD approach provides solutions with guaranteed 2% optimality in very reasonable times, even for these large instances, indicating good scalability. In order to assess the solution quality of the heuristic approach, we calculate the optimality gap as $100 * (Z^{\mathrm{H}} - Z^{\mathrm{LB}})/Z^{\mathrm{LB}}$, where Z^{LB} represents the lower bound from the base Benders approach (executed with a termination criterion of 2% optimality gap) and Z^{H} represents the heuristic solution value. The mean of the average values of gaps thus obtained is 2.08% over 24 classes, with a maximum of 2.27% and a minimum of 1.98%. Noting that the lower bound solution already corresponds to a known 2% optimality gap, these results demonstrate that the upper bound heuristic provides consistently good-quality solutions within short runtimes.

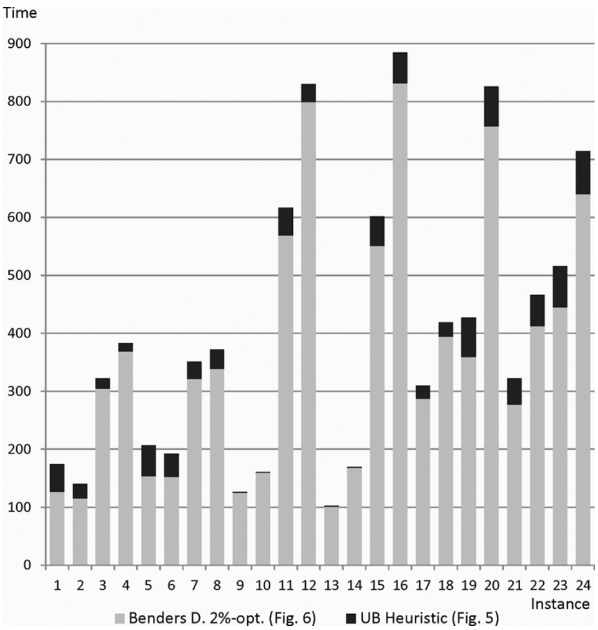

Fig. 8. Solution times for Setting II: large instances L1 to L24.

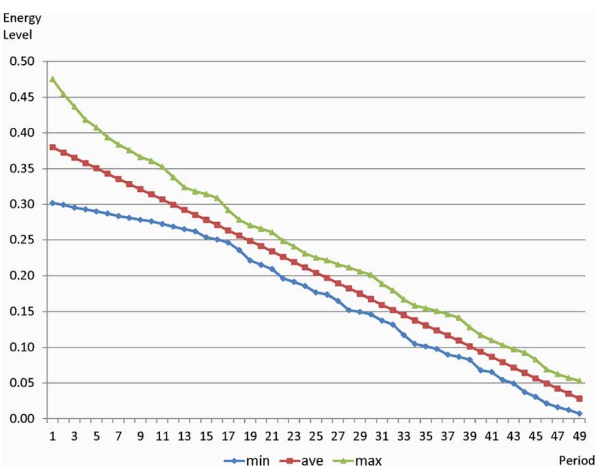

Fig. 9. Network remaining energy statistics in progression.

In general, we can conclude that: 1) our upper bound heuristic approach provides high-quality solutions with much less runtime than CPLEX; 2) ε-optimal Benders decomposition method, which amalgamates various problem and solution characteristics, is very effective in addressing a rather complex problem; and 3) combining the heuristics and ε-optimal BD approach promotes faster convergence, especially for larger instances, and provides a lower bound for which the quality of any solution can be precisely quantified via calculation of an optimality gap.

C. Results on a Sample WSN

We next study the effects of the modeling and algorithmic approaches on the energy profile and network configuration using a sample network generated as described above. In order to observe energy usage, as suggested in [17], we construct the graph in Fig. 9, which shows the minimum, average, and maximum energy levels in the whole network for each period during its lifetime. Clearly, the variation in sensor energy levels is controlled firmly by the suggested approach; We do not observe a

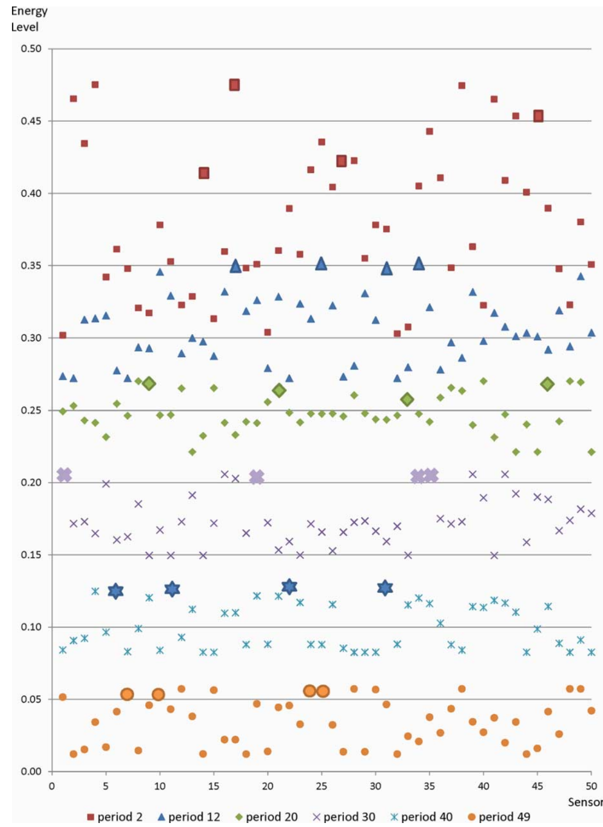

Fig. 10. Sensor energy levels and CHs in selected periods.

situation where some sensors are exploited significantly as CHs to cause them substantial energy drainage while others being lightly used due to not being strategically well positioned to participate in data transfer to sinks.

In this particular sample network, the network lifetime corresponds to a total of 49 periods. In Fig. 10, we plot the individual sensor energy levels for six selected periods (for easy discernibility in the scatter plot). We observe that mostly the sensors with high energy levels are selected as CHs. However, a closer inspection in Fig. 10 reveals that, in a given period, although there are several sensors with high energy levels, the optimization approach selects as CHs the ones that are the most energy-efficient from the network performance perspective.

To gain further insights into our suggested approach, we provide network configurations for four sample periods in Fig. 11. Given the sensor energy distribution at each period, we characterize five break points as follows:

a	the minimum energy value;
b	the average energy value;
c	the maximum energy value;
a_1	the energy value is equal to $a * 1.02$;
c_1	the energy value is equal to $c * 0.98$;

and define six different energy levels via six colors as follows:

gray	the node with minimum energy value a;
pink	nodes whose energy value is in $(a, a1]$;
green	nodes whose energy is in $(a1, b]$;
purple	nodes whose energy is in $(b, c1)$;

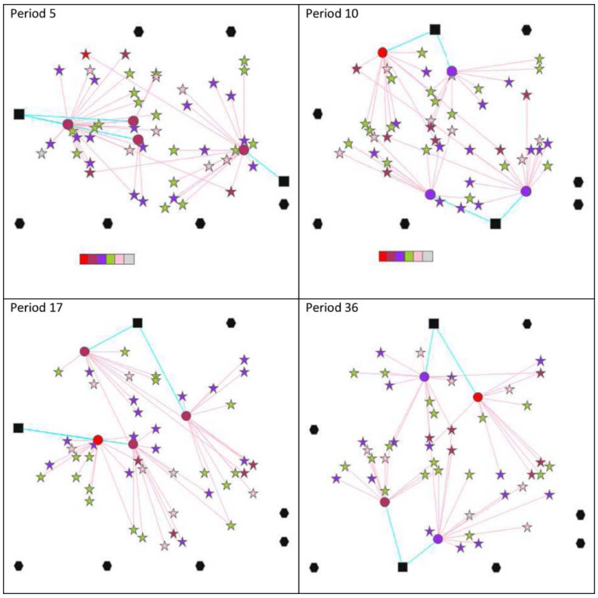

Fig. 11. Sample network configurations.

maroon nodes whose energy is in $[c1, c)$;

red the node with maximum energy value c.

In addition, we employ different shapes for node types. Specifically, *star* represents a regular sensor, *circle* represents a node selected as a CH, *square* represents a selected sink, and *hexagon* represents a candidate sink. In Fig. 11, we observe that an optimum network is not always obtained by selecting higher-energy sensors as CHs. For example, in periods 10 and 36, sensors with a mix of energy levels are selected as CHs. Even the highest energy sensor may not be picked as a CH in an optimal configuration, e.g., as in period 5. Furthermore, the locations of good CHs are not necessarily close to sinks, whose locations also change, but can be more central as in periods 5 and 17. In summary, both the position and the energy level of sensors need to be taken into account in an integrated fashion in designing the networks, and this is effectively facilitated by an optimization approach.

VII. Concluding Remarks

In this paper, we study an integrated topology control and routing problem in cluster-based WSNs. We develop an MILP model to determine the multiple sink and CH locations as well as the data routing scheme over a time horizon. We propose a new objective as the minimization of a weighted sum of the average energy usage, the range of remaining energy distribution, and the energy-based fixed cost for selecting the CHs. By doing so, our model avoids some well-positioned sensors being selected as CHs repeatedly in successive periods to protect low-energy sensors from quick energy depletion while facilitating a uniform energy consumption profile in the network. Furthermore, the integrated approach effectively combines the location and energy-level aspects of the sensors while determining the CHs in conjunction with the locations of the sinks to design the underlying network.

On the methodology side, we develop an effective ε-optimal BD approach that incorporates an upper bound heuristic algorithm and strengthened cuts. Specifically, we devise a heuristic algorithm that provides a good feasible solution so as to facilitate the generation of an initial Benders cut. We adopt a two-phase approach to strengthen the Benders cuts and utilize the ε-optimal approach to decrease the master problem solution times.

Computational evidence demonstrates the performance of the approach in terms of solution quality and time. In particular, the optimal solutions obtained by CPLEX and the lower bounds obtained by the BD verify the high quality of the heuristic solutions for small and large instances, respectively. The availability of good lower bounds is facilitated by the good initial Benders cut and the strengthened Benders cuts.

This study can be extended in several directions. One extension of our work is to incorporate the coverage problem into the integrated topology control and routing problems, i.e., we exploit the high spatial redundancy of sensors by only allowing a subset of sensors active for a given period of time, whereas all other sensors save energy being in inactive state. Since we currently focus on time-driven sensor networks applications pertaining to continuously monitoring ecological habitats (animals, plants, micro-organisms), another interesting extension in the future is to reformulate the models to suit for the time-critical applications.

Acknowledgment

The authors thank the associate editor and three anonymous referees for their helpful comments on the earlier versions of this paper.

References

[1] C. Y. Chong and S. P. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proc. IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.

[2] K. Sohraby, D. Minoli, and T. F. Znati, *Wireless Sensor Networks: Technology, Protocols, and Applications*. Hoboken, NJ, USA: Wiley, 2007.

[3] S. Tilak, N. B. Abu-Ghazaleh, and W. R. Heinzelman, "A taxonomy of wireless micro-sensor network models," *Mobile Comput. Commun. Rev.*, vol. 6, no. 2, pp. 28–36, 2002.

[4] J. Porter, P. Arzberger, H. Braun, P. Bryant, S. Gage, T. Hansen, P. Hanson, C. Lin, F. Lin, T. Kratz, W. Michener, S. Shapiro, and T. Williams, "Wireless sensor networks for ecology," *BioScience*, vol. 55, pp. 561–572, 2005.

[5] A. Cerpa, J. Elson, D. Estrin, L. Girod, M. Hamilton, and J. Zhao, "Habitat monitoring: Application driver for wireless communications technology," in *Proc. ACM SIGCOMM Workshop Data Commun. Latin Amer. Carib.*, 2001, pp. 20–41.

[6] A. Mainwaring, J. Polastre, R. Szewczyk, and D. Culler, "Wireless sensor networks for habitat monitoring," in *Proc. 1st ACM WSNA*, 2002, pp. 88–97.

[7] E. Biagioni and K. Bridges, "The application of remote sensor technology to assist the recovery of rare and endangered species," *Int. J. High Perform. Comput. Appl.*, vol. 16, no. 3, pp. 315–324, 2002.

[8] S. L. Collins, L. M. A. Bettencourt, A. Hagberg, R. F. Brown, D. I. Moore, G. Bonito, K. A. Delin, S. P. Jackson, D. W. Johnson, S. C. Burleigh, R. R. Woodrow, and J. M. McAuley, "New opportunities in ecological sensing using wireless sensor networks," *Frontiers Ecol. Environ.*, vol. 4, no. 8, pp. 402–407, 2006.

[9] K. Martinez, J. K. Hart, and R. Ong, "Environmental sensor networks," *Computer*, vol. 37, pp. 50–56, 2004.

[10] F. Al-Turjman, H. Hassanein, and M. Ibnkahla, "Connectivity optimization with realistic lifetime constraints for node placement in environmental monitoring," in *Proc. 34th IEEE LCN*, Oct. 2009, pp. 617–624.

[11] F. Al-Turjman, H. Hassanein, and M. Ibnkahla, "Connectivity optimization for wireless sensor networks applied to forest monitoring," in *Proc. IEEE ICC*, Jun. 2009, pp. 1–6.

[12] J. K. Hart and K. Martinez, "Environmental sensor networks: A revolution in the earth system science?," *Earth-Sci. Rev.*, vol. 78, no. 3–4, pp. 177–191, 2006.

[13] P. W. Rundel, E. A. Graham, M. F. Allen, J. C. Fisher, and T. C. Harmon, "Environmental sensor networks in ecological research," *New Phytologist*, vol. 182, no. 3, pp. 589–607, 2009.

[14] L. Selavo, A. Wood, Q. Cao, T. Sookoor, H. Liu, A. Srinivasan, Y. Wu, W. Kang, J. Stankovic, D. Young, and J. Porter, "Luster: Wireless sensor network for environmental research," in *Proc. 5th SenSys*, 2007, pp. 103–116.

[15] Q. Wang and H. Hassanein, "A comparative study of energy-efficient (E^2) protocols for wireless sensor networks," in *Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems*, M. Ilyas and I. Mahgoub, Eds. Boca Raton, FL, USA: CRC Press, 2004, ch. 18.

[16] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Commun. ACM*, vol. 43, no. 5, pp. 51–58, 2000.

[17] H. Üster and H. Lin, "Integrated topology control and routing in wireless sensor network design for prolonged network lifetime," *Ad Hoc Netw.*, vol. 9, no. 5, pp. 835–851, 2011.

[18] W. B. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy efficient communication protocol for wireless micro-sensor networks," in *Proc. IEEE Hawaii Int. Conf. Syst. Sci.*, 2000, pp. 174–185.

[19] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks," *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 366–379, Oct.–Dec. 2004.

[20] M. Liu, J. Cao, G. Chen, and X. Wang, "An energy-aware routing protocol in wireless sensor networks," *Sensors*, vol. 9, no. 1, pp. 445–462, 2009.

[21] E. Ever, R. Luchmun, L. Mostarda, A. Navarra, and P. Shah, "UHEED—An unequal clustering algorithm for wireless sensor networks," in *Proc. 1st SENSORNETS*, Feb. 2012, pp. 185–193.

[22] J. N. Al-Karaki, R. Ul-Mustafa, and A. E. Kamal, "Data agregation and routing in wireless sensor networks: Optimal and heuristic algorithms," *Computer Netw.*, vol. 53, no. 7, pp. 945–960, 2009.

[23] F. Al-Turjman, H. Hassanein, and M. Ibnkahla, "Optimized relay repositioning for wireless sensor networks applied in environmental applications," in *Proc. 7th IWCMC*, Jul. 2011, pp. 1860–1864.

[24] J. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 12, no. 4, pp. 609–619, Aug. 2004.

[25] S. Misra, S. D. Hong, G. Xue, and J. Tang, "Constrained relay node placement in wireless sensor networks: Formulation and approximations," *IEEE/ACM Trans. Netw.*, vol. 18, no. 2, pp. 434–447, Apr. 2010.

[26] W. Wang, V. Srinivasan, and K.-C. Chua, "Extending the lifetime of wireless sensor networks through mobile relays," *IEEE/ACM Trans. Netw.*, vol. 16, no. 5, pp. 1108–1120, Oct. 2008.

[27] Y. Wu, Z. Mao, S. Fahmy, and N. B. Shroff, "Constructing maximum-lifetime data-gathering forests in sensor networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 5, pp. 1571–1584, Oct. 2010.

[28] H. Kim, Y. Seok, N. Choi, Y. Choi, and T. Kwon, "Optimal multi-sink positioning and energy-efficient routing in wireless sensor networks," *Inf. Netw. Inf. Netw.*, vol. 3391, pp. 264–274, 2005.

[29] Y. Xue, Y. Cui, and K. Nahrstedt, "Maximizing lifetime for data aggregation in wireless sensor networks," *Mobile Netw. Appl.*, vol. 10, pp. 853–864, 2005.

[30] J. Luo and J.-P. Hubaux, "Joint sink mobility and routing to maximize the lifetime of wireless sensor networks: The case of constrained mobility," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 871–884, Jun. 2010.

[31] A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Comput. Commun.*, vol. 30, no. 14–15, pp. 2826–2841, 2007.

[32] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Netw.*, vol. 3, no. 3, pp. 325–349, 2005.

[33] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: A survey," *IEEE Wireless Commun.*, vol. 11, no. 6, pp. 6–28, Dec. 2004.

[34] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw.*, pp. 393–422, Mar. 2002.

[35] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, 2008.

[36] I. Akyildiz and M. C. Vuran, *Wireless Sensor Networks*. New York, NY, USA: Wiley, 2010.

[37] P. A. Ademola, N. A. A., and A. O. A. K. K. A., "ANCAEE: A novel clustering algorithm for energy efficiency in wireless sensor networks," *Wireless Sensor Netw.*, vol. 3, no. 9, pp. 307–312, 2011.

[38] M. Haenggi, "Energy-balancing strategies for wireless sensor networks," in *Proc. IEEE ISCAS*, Bangkok, Thailand, 2003, pp. IV-828–IV-831.

[39] G. Chen, C. Li, M. Ye, and J. Wu, "An unequal cluster-based routing protocol in wireless sensor networks," *Wireless Netw.*, vol. 15, no. 2, pp. 193–207, 2009.

[40] A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proc. 7th Annu. Int. Conf. Mobile Comput. Netw.*, 2001, pp. 166–179.

[41] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low cost outdoor localization for very small devices," Department of Computer Science, University of Southern California, Los Angeles, CA, USA, Tech. Rep. 00-729, 2000.

[42] S. Soro and W. B. Heinzelman, "Cluster head election techniques for coverage preservation in wireless sensor networks," *Ad Hoc Netw.*, vol. 7, no. 5, pp. 955–972, 2009.

[43] J. Rodrigues, S. Fraiha, H. Gomes, G. Cavalcante, A. de Freitas, and G. de Carvalho, "Channel propagation model for mobile network project in densely arboreous environments," *J. Microw. Optoelectron.*, vol. 6, no. 1, pp. 236–248, 2007.

[44] M. Kohvakka, J. Suhonen, M. Kuorilehto, V. Kaseva, M. Hannikainen, and T. D. Hamalainen, "Energy-efficient neighbor discovery protocol for mobile wireless sensor networks," *Ad Hoc Netw.*, vol. 7, no. 1, pp. 24–41, 2009.

[45] R. Cohen and B. Kapchits, "Continuous neighbor discovery in asynchronous sensor networks," *IEEE/ACM Trans. Netw.*, vol. 19, no. 1, pp. 69–79, Feb. 2011.

[46] S. Vasudevan, M. Adler, D. Goeckel, and D. Towsley, "Efficient algorithms for neighbor discovery in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 1, pp. 69–83, Feb. 2013.

[47] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

[48] Q. Wang and W. Yang, "Energy consumption model for power management in wireless sensor networks," in *Proc. 4th Annu. IEEE SECON*, Jun. 2007, pp. 142–151.

[49] F. M. Al-Turjman, H. S. Hassanein, W. M. Alsalih, and M. Ibnkahla, "Optimized relay placement for wireless sensor networks federation in environmental applications," *Wireless Commun. Mobile Comput.*, vol. 11, no. 12, pp. 1677–1688, 2011.

[50] J. F. Benders, "Partitioning procedures for solving mixed-variables programming problems," *Numer. Math.*, vol. 4, pp. 238–252, 1962.

[51] T. Magnanti and R. Wong, "Accelerating Benders decomposition: Algorithmic enhancement and model selection," *Oper. Res.*, vol. 29, no. 3, pp. 464–484, 1981.

[52] T. J. V. Roy, "A cross decomposition algorithm for capacitated facility location," *Oper. Res.*, vol. 34, no. 1, pp. 145–163, 1986.

[53] P. Wentges, "Accelerating Benders' decomposition for the capacitated facility location problem," *Math. Methods Oper. Res.*, vol. 44, no. 2, pp. 267–290, 1996.

[54] H. Üster, G. Easwaran, E. Akcali, and S. Cetinkaya, "Benders decomposition with alternative multiple cuts for a multi-product closed-loop supply chain network design mode," *Naval Res. Logistics*, vol. 54, no. 8, pp. 890–907, 2007.

[55] A. Geoffrion and G. Graves, "Multicommodity distribution system design by Benders decomposition," *Manage. Sci.*, vol. 20, no. 5, pp. 822–844, 1974.

Hui Lin received the B.Sc. degrees in automation and engineering economics and M.Sc. degree in electrical engineering from Tianjin University, Tianjin, China, in 2001 and 2004, respectively, and the Ph.D. degree in industrial engineering from Texas A&M University, College Station, TX, USA, in 2010.

She visited the San Diego Supercomputer Center, La Jolla, CA, USA, during July to September of 2007 and 2008 as a participant in the Cyberinfrastructure Experiences for Graduates program. She is currently with Bank of America, Dallas, TX, USA, as an Operations Research Analyst. Her research interests include WSN design, modeling and optimization, and high-performance computing.

Halit Üster received the Ph.D. degree in management science/systems from McMaster University, Hamilton, ON, Canada, in 1999.

He is an Associate Professor with the Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA. His publications appeared in *Ad Hoc Networks, Computers and Operations Research, IIE Transactions, Naval Research Logistics*, and *Transportation Science*, among others. His research interests are in the design of networked systems in logistics and communications and applied optimization.