A Cluster-based Energy Balancing Scheme in Heterogeneous Wireless Sensor Networks

Jing Ai, Damla Turgut, and Ladislau Bölöni

Networking and Mobile Computing Research Laboratory (NetMoC)
Department of Electrical and Computer Engineering
University of Central Florida
Orlando, FL 32816,
{jingai,turgut,lboloni}@cpe.ucf.edu

Abstract. In this paper, we propose a novel, cluster-based energy balancing scheme. We assume the existence of a fraction of "strong" nodes in terms of abundant storage, computing and communication abilities as well as energy. With the transformation of the flat network infrastructure into a hierarchical one, we obtained significant improvements in energy balancing leading to a longer connected time of the network. The improvement is quantified by mathematical analysis and extensive numerical simulations.

1 Introduction

Unbalanced energy consumption is an inherent problem in wireless sensor networks, and it is largely orthogonal to the general energy efficiency problem. For example, in a data gathering application, multi-hop wireless links are utilized to relay information to destination points called *sinks*. Inevitably, the nodes closer to the sink will experience higher traffic and higher energy consumption rate. These nodes will be the first ones which run out of power. Algorithms which allow "routing around" failed nodes will increase the load even more on the remaining active nodes close to the sink.

Our proposed cluster-based energy balancing scheme is intended to ameliorate the above energy unbalancing phenomena. We exploit the observation that in a heterogeneous sensor network there are nodes which are more powerful in terms of energy reserve and wireless communication ability. We transform the flat communication infrastructure into a hierarchical one where "strong" nodes act as clusterheads to gather information within the clusters and then communicate with the sink directly via single-hop link. In such a way, the "hot spot" around the sink is divided into multiple regions around the clusterheads in the hierarchical infrastructure. Furthermore, these distributed regions will assume fewer burdens due to the smaller scale of sensor nodes within the clusters.

The remainder of this paper is organized as follows. Section 2 describes in detail the proposed model of cluster-based energy balancing scheme for heterogeneous wireless sensor networks. Section 3 presents the analysis as well as the performance study on the proposed scheme. Section 4 reviews related work in the literature with reference to energy-efficiency schemes for wireless sensor networks. Finally, section 5 concludes the paper and outlines directions for ongoing research.

2 A cluster-based energy balancing scheme

2.1 Motivation

Wireless sensor networks are composed by a large amount of small, resource constrained devices, called *sensor nodes*, which have limited sensing, computing and wireless communication abilities. These sensor nodes usually collaborate with each other via multi-hop links. The multi-hop organization presents many advantages, from the increase of the network capacity, ability to perform data fusion and a more efficient energy utilization. However, under many scenarios, multi-hop sensor networks are utilizing energy in an unbalanced manner.

To illustrate this phenomena, let us consider a simple, unidirectional example in Figure 1. We assume that all nodes communicate only with their neighbors and all the nodes are sending their observations back to the sink. We assume the nodes to be equi-distant, and thus the dissipated energy being roughly the same for each node. Normally, if all nodes have the same initial energy upon deployment, the node closer to the sink will drain earlier since it has heavier forwarding burden. Moreover, the further nodes which may still have plentiful energy supplies cannot find the routes to the sink. The energy unbalancing problem will aggravate with the increase of the network depth (defined as the largest number of hop from a node to the sink) [9].
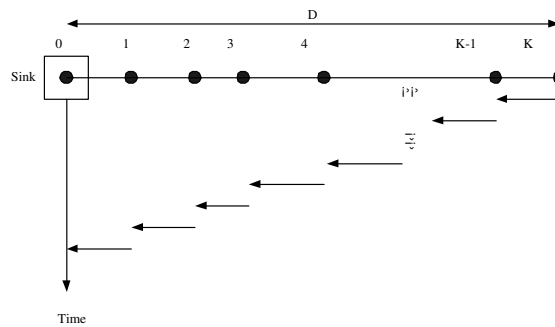

Fig. 1. $(K + 1)$-node line network assumed to be connected. It illustrates a transmission schedule when only node K transmits a data packet to node 0 (sink) via multi-hop links.

The best resource utilization is achieved when every sensor node has the same rate of energy dissipation (or as close as possible), such that the network remains functional for the maximum possible time. Such a forwarding schedule is theoretically obtainable, by the algorithms proposed by Bhardwaj and Chandrakasan [3].

Although the proposed algorithm executes on polynomial time only, it also requires the global knowledge of the traffic, and thus is not feasible except for centrally managed networks and very large data packets. For a typical sensor network, where the individual measurements are small, the collection of global traffic information would be as expensive as the actual data communication itself.

Our algorithm proposes a cluster-based organization of the traffic, which does not require global information, and proposes to ameliorate the energy unbalancing problem by decreasing or confining the network depth within each cluster.

2.2 Scheme description

In a heterogeneous sensor network, we identify a subset of nodes as "strong" nodes with more powerful communication capabilities and energy resources. Instead of the flat organization of nodes, we assume a hierarchical structure where the strong nodes act as clusterheads. The clusterheads should be able to form a connected backbone between themselves such that they can communicate without relying on regular nodes. We assume two types of communication: one between the the regular nodes and the clusterheads with low transmission power, and the communication between clusterheads with higher transmission range spawning larger distances. In a practical deployment, these two types of traffic may be carried on different frequency bands or encoding techniques.

During the initialization phase, strong nodes broadcast their willingness to act as clusterheads. The sensor nodes decide to which cluster they wish to belong based on the strength of signal from the broadcast: the stronger the signal, the closer the clusterhead is and therefore the clusterhead with the strongest signal is chosen. At this point multiple clustering algorithms can be used, provided that they can be adapted to the specific condition of having a pre-determined clusterhead. On the other hand, algorithms which rely on dynamic leader election [7] are not appropriate for this purpose.

After the clusters are formed, the sensor nodes can use various algorithms for energy-efficient schedule for transmission such as in [9]. The clusterhead gathers the information from the sensor node within its cluster via multi-hop link and then forwards the aggregated information to the sink through the backbone of clusterheads.

This approach has all the desirable properties of similar schemes [5], such as localized traffic and scalability. The clusterheads are the natural points to implement data fusion and data compression algorithms. First, there is a potential correlation in the data from neighboring sensor nodes (given their physical proximity), and second, the higher energy resources of the strong nodes allows them to execute more complex computations. The proposed clustering scheme reduces the depth of the average multi-hop path to the clusterhead and transforms the single heavy "hot spot" around the sink to various distributed lighter "hot spots" around corresponding clusterheads. The ratio of the strong nodes to regular nodes determines the average depth of the multi-hop path inside the cluster. The essence of proposed the scheme explores the tradeoff between the multi-hop communication within the clusters and single-hop communication among clusters to achieve a better utilization of the energy resources.

3 Performance evaluation

3.1 Preliminaries

To facilitate the performance analysis, we make the following assumptions:

i There is only one sink node with abundant energy resources.

ii There are N identical regular sensor nodes uniformly distributed in a planar disk whose radius is R.

iii There are S identical "strong" nodes with pre-determined locations in the same area such that they form clusters of roughly equal size.

iv The regular sensor nodes consume their energy much faster than the strong nodes such that the bottleneck is the energy of the regular nodes.

v The maximum transmission range, r, of regular sensor nodes ensures the connectivity of the network while the transmission range of "strong" nodes is large enough for strong nodes and the sink to form a connected backbone.

vi There is no interference between the communication in the backbone and the intra-cluster communication.

vii The nodes may fail only when they deplete their energy resource.

viii All nodes deploy an ideal MAC protocol and there is no collision among packets.

ix All nodes have an ideal sleep scheduling and consume energy only during transmission and reception.

The energy consumption of a sensor node is divided between the three components of a wireless sensor: sensing, computation and communication components [3].

1. Sensing: We assume that every sensor node captures b bits/sec data from its environment. The energy needed to sense a bit of data is (α_3). Thus, the energy consumed for sensing is $p_{sense} = \alpha_3 b$.

2. Computation: The computational power of a sensing node is used for operations, such as data aggregation. It is difficult to quantify the energy used for data agregation in absolute terms without specific knowledge about the nature of the data. However, in our analysis, we are interested in the *relative* performance of the hierarchical organization against a flat network of sensor nodes. We will assume that any scheme will benefit both organizations approximately equally, thus we will ignore this term in our calculations.

3. Communication: We use the following model for the energy dissipation used for communication [8]:

$$p_{tx}(n_1, n_2) = (\alpha_{11} + \alpha_2 d(n_1, n_2)^n)b \qquad (1)$$

$$p_{rx} = \alpha_{12} b \qquad (2)$$

where $p_{tx}(n_1, n_2)$ is the power dissipated in node n_1 when it is transmitting to node n_2, $d(n_1, n_2)$ is the distance between the two nodes, n is the path loss index, and the α_i are positive constants.

Typical values of the parameters α_i are shown in Table 1.

3.2 Analysis

The energy consumption of the wireless sensor network is determined by the spatial distribution of the sensor nodes. Although in our approach the strong nodes are in pre-determined locations, the distribution of the locations of the regular nodes is essentially

Table 1. Typical values of parameters α_i [6].

α_{11}	45 nJ/bit
α_{12}	135 nJ/bit
α_2	10 pJ/bit/m^2 ($n = 2$)or 0.001 pJ/bit/m^4 ($n = 2$)
α_3	50 nJ/bit

random. Thus, our analysis will be based on establishing lower bounds of the energy consumptions. We will rely on two theorems introduced in [2]:

Theorem 1: Given D and number of intervening relays ($K - 1$) as shown in Figure 1, $P_{link}(D)$ is minimized when all the hop distances are equal to $\frac{D}{K}$.

This theorem gives us a bound of energy dissipation rate in a line network via multi-hop links. It is interesting to note that increasing the number of hops can effectively decrease the transmission power while increase the reception power. There is an optimal number of hops K_{opt} which minimizes the total energy dissipation by trading of the power consumed for transmission and reception.

Theorem 2: The optimal number of hops K_{opt} is always one of

$$K_{opt} = \lfloor \frac{D}{d_{char}} \rfloor or \lceil \frac{D}{d_{char}} \rceil \tag{3}$$

where the distance d_{char}, called the characteristic distance, is independent of D and is given by

$$d_{char} = \sqrt[n]{\frac{\alpha_1}{\alpha_2(n-1)}} \tag{4}$$

We conclude, that for any path loss index n, the energy cost of transmitting a bit can always be made linear with distance. Moreover, for any given distance D, there is an optimal number K_{opt} of intervening nodes. Using more or less than this optimal number leads to energy inefficiencies.

In the following, we will use the above theorems to compare the energy consumption of a wireless sensor network with a flat network architecture to the clustering scheme proposed in this paper.

Case I: flat network architecture. In our environment, there are N identical sensor nodes uniformly distributed in a planar disk of radius R. Using the results from [2], we derive the lower bound of the energy dissipation rate:

$$P_{flat\_network} \geq (\sum_{i=1}^{N} \alpha_1 \frac{n}{n-1} \frac{d_i}{d_{char}} - N\alpha_{12})b \tag{5}$$

where, d_i is the distance of sensor i from the center of the disk.

Thus, the expected value of the lower bound of dissipated energy can be derived as follows:

$$E[min(P_{flat\_network})] = E[(\sum_{i=1}^{N} \alpha_1 \frac{n}{n-1} \frac{d_i}{d_{char}} - N\alpha_{12})b]$$

$$= \alpha_1 \frac{n}{n-1} E[\sum_{i=1}^{N} \frac{d_i}{d_{char}}] - N\alpha_{12}b$$

$$= \alpha_1 \frac{n}{n-1} \frac{1}{d_{char}} E[\sum_{i=1}^{N} d_i] - N\alpha_{12}b \qquad (6)$$

$$= [\alpha_1 \frac{n}{n-1} \frac{RN}{2d_{char}} - N\alpha_{12}]b$$

Case II: hierarchical clustering scheme. According to the clustering scheme described above, when S "strong" nodes are deployed, S clusters will automatically be formed. In each cluster, the expected number of nodes is $\frac{N}{S}$.

The individual clusters have a similar structure like the flat network, but we also need to consider both the reception energy consumption of the strong nodes and the energy consumption related to the communication between the strong node and the sink, which follows the equation (1):

$$P_{clustered\_network} \geq S \sum_{i=1}^{\frac{N}{S}} \alpha_1 \frac{n}{n-1} \frac{d_i}{d_{char}} b + \sum_{i=1}^{S} (\alpha_{11} + \alpha_2 d_i''^n)b \qquad (7)$$

where, d_i is a random variable following the uniform distribution over the interval $[0, \frac{R}{\sqrt{S}}]$ and $d_i''^n$ is a random variable following the uniform distribution over the interval $[0, R]$.

Thus, the expected value of the minimum $min(P_{clustered\_network})$ can be derived as follows:

$$E[min(P_{clustered\_network})] = E[S \sum_{i=1}^{\frac{N}{S}} \alpha_1 \frac{n}{n-1} \frac{d_i}{d_{char}} b +$$

$$\sum_{i=1}^{S} (\alpha_{11} + \alpha_2 d_i''^n)b]$$

$$= S\alpha_1 \frac{n}{n-1} E[\sum_{i=1}^{\frac{N}{S}} \frac{d_i}{d_{char}}] + \qquad (8)$$

$$S\alpha_{11}b + \alpha_2 E[\sum_{i=1}^{S} d_i''^n]b$$

$$= [\alpha_1 \frac{n}{n-1} \frac{RN}{2\sqrt{S}d_{char}} + S\alpha_{11} +$$

$$\alpha_2 S \frac{R^n}{n+1}]b$$

An important consequence is that the communication cost of multi-hop links increases with the number of clusters while the communication cost of messaging on the backbone increases with the number of clusters. Thus, there exists an optimal number of clusters which trades off the power consumption between multi-hop and single-hop links to minimize the energy dissipation rate. Applying the techniques of previous two theorems, we can deduct that optimal number clusters is always one of

$$S_{opt} = \lfloor (\frac{\alpha_1 \frac{n}{n-1} \frac{RN}{d_{char}}}{\alpha_{11} + \alpha_2 \frac{R^n}{n+1}})^{\frac{2}{3}} \rfloor \, or \, \lceil (\frac{\alpha_1 \frac{n}{n-1} \frac{RN}{d_{char}}}{\alpha_{11} + \alpha_2 \frac{R^n}{n+1}})^{\frac{2}{3}} \rceil \qquad (9)$$

This result is important from the practical deployment point of view of a sensor network. We need to limit the number of clusters to the one shown in the Equation 9 even if we have a larger number of nodes which, based on their hardware characteristics, would qualify as "strong" nodes.

3.3 Numerical simulation

We will numerically analyze the energy dissipation rate of our scheme compared with the flat network architecture. In addition, we examine the impact of the various parameters, n, N, R, S. We assume a sensor network is composed of $N = 10000$ sensor nodes distributed on a radius of $R = 1000$ meters with communication path loss index $n = 2$ and data bit rate $b = 1bits/sec$.

Figure 2, left, shows the energy dissipation in function of the number of clusters S ranging from $1\% \sim 10\%$ of N. Another property of interest is the optimal *percentage* of clusterheads or strong nodes. Thus, in Figure 2 right, we plot the calculated optimum percentage in function of the total number of nodes, N. We found that the optimal percentage of strong nodes decreases with the number of total sensor nodes and it is between 9% to 2% in a typical field of 10,000 to 100,000 nodes of deployment. Thus, the remarkable gain in energy dissipation rate can be obtained with relatively small percentage of strong nodes.
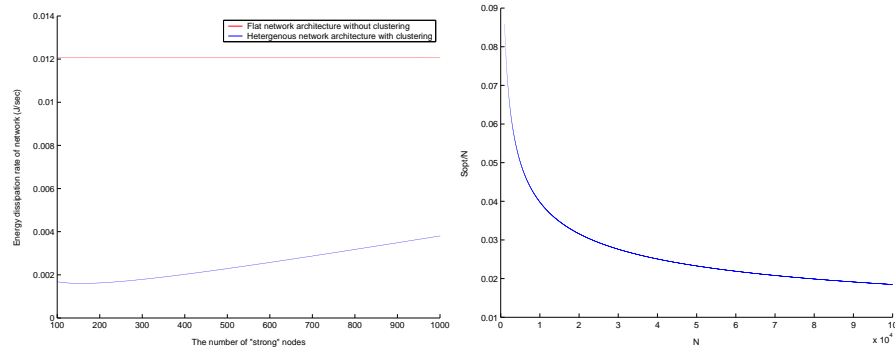

Fig. 2. Energy dissipation vs. number of clusters (left) and the optimal percentage of clusterheads to the total number of nodes, N (right)

Next, we determine the optimum number of "strong" nodes with the increase of the R while keeping other parameters unchanged as can be seen in Figure 3. Thus, we visualize the impact of the density of the network on the optimal energy dissipation rate when the optimal number of "strong" nodes is deployed.
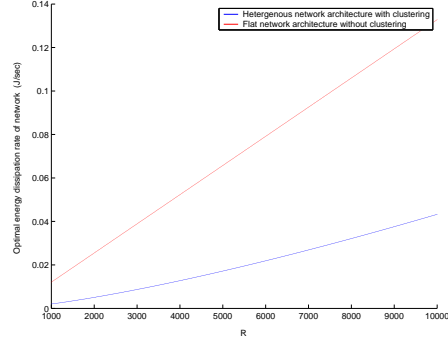

Fig. 3. The impact of network density on the optimal performance of two paradigms of networks where $n = 2$, $N = 10000$.

By repeating the experiments with $n = 4$, we obtained the results in Figure 4 and 5. Contrary to our expectations, our clustering scheme does not show any benefits for this experimental setup. This is explained by the fact that in an environment with large path loss index, the single-hop operations are much more expensive than multi-hop communications. We conclude that the benefits of our scheme is highly dependent on the environment and the it is better adapted for low path loss index values.
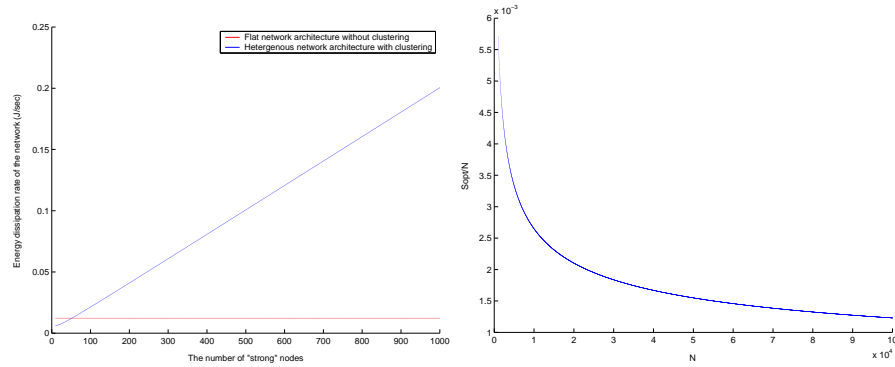

Fig. 4. Performance comparison of two paradigms of networks where $n = 4$, $N = 10000$, $R = 1000$ (left) and the relationship between $\frac{S_{o}pt}{N}$ and N where $n = 4$, $R = 1000$ (right)
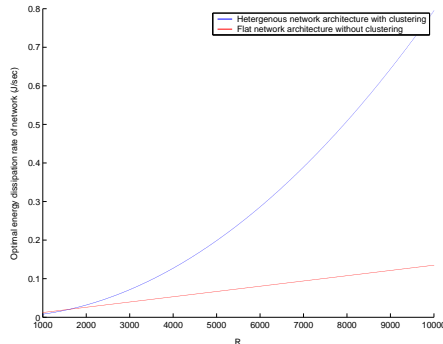
Fig. 5. The impact of network density on the optimal performance of two paradigms of networks where $n = 4$, $N = 10000$.

4   Related Work

The literature of the ad hoc and sensor networks contains several mechanisms for energy-efficient and energy-balanced communication.

LEACH [5, 6] (Low-Energy Adaptive Clustering Hierarchy) is a cluster-based routing protocol that minimizes a global energy usage by randomized rotation of local clusterheads to evenly distribute the energy load among the sensors in the network. In each round of LEACH operation, the clusters are organized in a set-up phase, followed by a steady-state phase. When clusters are selected, each node decides whether or not to become a clusterhead for the current round. Each node chooses a random number between 0 and 1. The decision is made by the comparison of this number with $T(n)$, which is based on the suggested percentage of clusterheads for the network and the number of times the node has been a clusterhead so far. If $T(n)$ is larger, the node is more likely to be a clusterhead for the current round. Based on the number of nodes in the cluster, the clusterhead node creates a TDMA schedule telling each node when it can transmit. Our scheme is different from LEACH through the fact that it operates in and exploits the properties of a heterogeneous network. Instead of rotating the clusterheads between the homogeneous nodes, the algorithm uses the strong nodes as clusterheads. In addition, we allow multi-hop links for inner-cluster communication. Thus, our algorithm does not require the overhead of the dynamic reconfiguration. On the other hand, it assumes the existence the suitably distributed strong nodes, and its benefits depend on the low path loss indexes.

M. Bhardwaj and A.P. Chandrakasan [3] introduce a role assignment framework to derive bounds of network lifetime for a variety of data gathering scenarios. Althogh their study deals with the lifetime of the networks with limited energy resources, they also implicitly define a transmission schedule for energy-efficient and energy balanced communication. Unfortunately, this transmission schedule can be computed only in the presence of global information, which is not normally available in a deployed network.

E. J. Duarte-Melo and M. Liu [4] propose and analyze a hierarchical scheme which is closer to ours by utilizing two types of nodes to organize the infrastructure of the

network. However, it still needs to dynamically choose clusterheads among "strong" nodes and it deploys single-hop links within a cluster as well.

5 Conclusions

Through introducing a series of "strong" nodes as clusterheads, we change the communication structure of the original data fusion in wireless sensor networks from a flat to an hierarchical one which has better energy-balancing properties. Compared to other energy-balancing schemes, our scheme is rather simple and effective. Future work includes adapting the protocol that does not depend on neither the environment nor the path loss index as well as an extensive simulation work to validate the analytical results.

References

1. I. Akyildiz, W. Su, Y. Sanakarasubramaniam, E. Cayirci. "Wireless Sensor Networks: A Survey". *Computer Networks Journal*, Vol. 38, No. 4, pp. 393 - 422, 2002.
2. M. Bhardwaj, T. Garnett, A.P. Chandrakasan. "Upper Bounds on the Lifetime of Sensor Networks". *In Proceedings of International Conference on Communications (ICC)*, 2001.
3. M. Bhardwaj, A.P. Chandrakasan. "Bounding the Lifetime of Sensor Networks via Optimal Role Assignments". *In Proceedings of INFOCOM 2002*, pp. 1587-1596, New York, June 2002.
4. E. J. Duarte-Melo, M. Liu. "Analysis of Energy Consumption and Lifetime of Heterogeneous Wireless Sensor Networks". *In Proceedings IEEE GLOBECOM*, November 2002, Taipei, Taiwan.
5. W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. "Energy-Efficient Communication Protocol for Wireless Microsensor Networks". *Proceedings of the 33rd International Conference on System Sciences (HICSS '00)*, January 2000.
6. W. R. Heinzelman. "Application-Specific Protocol Architectures for Wireless Sensor Networks". *Ph.D. Thesis, Massachusetts Institution of Technology*, 2000.
7. H. Garcia-Molina. "Election in a distributed computer system". *IEEE Transactions on Computers*, C-31(2), pp. 48-59, 1982.
8. T. Rappaport. *Wireless Communications: Principles & Practice*. New Jersey: Prentice-Hall, Inc., 1996.
9. M. L. Sichitiu. "Cross-Layer Scheduling for Power Efficiency in Wireless Sensor Networks". *In Proceedings of INFOCOM 2004*, Hong Kong, P. R. China, March 2004.