Sensor Networks Routing

Dr M R S Mitra
Professor & Dean of CSE, IT & MCA Dept., India
E-mail: dr.mrsmitra@rediffmail.com

M N Santhi
Assistant Professor, Department of CSE, India
E-mail: g3mango@rediffmail.com

M V P Kumar
Assistant Professor, Department of CSE, India
E-mail: mvphani@gmail.com

P R Sudha
Assistant Professor, Department of CSE, India
E-mail: raghasudha.rit@gmail.com

Abstract

In order to provide load balancing in clustered sensor deployment, the upstream clusters (near the BS) are kept smaller in size as compared to downstream ones (away from BS). Moreover, geographic awareness is also desirable in order to further enhance energy efficiency. But, this must be cost effective, since most of current location awareness strategies are either cost and weight inefficient (GPS) or are complex, inaccurate and unreliable in operation.

This paper presents design and implementation of a Load Balanced (LBA) Clustering Protocol for Wireless Sensor Networks. A mathematical formulation is provided for determining the number of sensor nodes in each cluster. This enables uniform energy consumption after the multi-hop data transmission towards BS. Either the sensors can be manually deployed or the clusters be so formed that the sensor are efficiently distributed as per formulation. The latter strategy is elaborated in this contribution. Methods to provide static clustering and custom cluster sizes with location awareness are also provided in the given work. Finally, low mobility node applications can also be implement the proposed work.

Keywords: LBA, Network, Routing, Organization

1. Introduction

Wireless Sensor Network (WSN) applications are characterized by dense Sensor Node (SN) deployment owing to disasters or battery depletions. Most of the applications like volcano reporting, battlefield surveillance, forest fire detection etc suffer from non replace ability or recharge ability of depleted batteries. The redundant nodes ensure that a minimal degree of fidelity is achieved even in the case of non functionality of few of the nodes. But, this density inherently brings about scalability and communicational efficiency issues. Clustering or hierarchical organization of network nodes is considered as a proficient rescue to these problems.

Majority of clustering protocols suffer from the limitations of frequent re-clustering, complex Cluster Head (CH) selection process and location unawareness. Re-clustering is required to prevent both the inter-cluster and intra-cluster 'hot-spot' problems. But, frequent re-clustering in order to rotate the roles of SNs leads to energy expenditure and sometimes needs lot of maintenance of current network topology. The CH selection too is complex in many protocols, as sometimes being based on residual energy of each SN requiring flooding of control packets and thus, again leading to energy expenditure. In others it is either random (Taewook Kang, Jangkyu Yun, Hoseung Lee, Icksoo Lee, Hyunsook Kim, Byunghwa Lee, Byeongjik Lee, Kijun Han. 2007) or based on certain probability function (D. Culler, D. Estrin and M. Srivastava.). Moreover, the clusters near the Base Station (BS) not only send aggregated data of their own members but also require sending data for other clusters at the locations farther away from BS than itself. Thus, in order to provide load balancing the clusters near the BS must be of smaller size and the size may gradually increase as we move away from BS. Recently,

only few of the approaches address to this problem and almost none of them provide a mathematical approach for optimal number of node to be placed in the clusters.

In this paper we propose a mathematical model for optimal placement of nodes in each cluster for a uniform load balancing in a location aware multi hop clustering protocol. The paper not only proposes a non uniform clustering scheme but also provides a way for multihop data routing. The protocol uses the location of a node as its ID for the purpose of routing. This prevents an overhead of maintaining a lot of IDs in case of extremely dense deployments. Moreover, the protocol doesn't require any GPS device or complex uncertain logics to determine this location; rather it uses a grid based, BS controlled approach to determine the location of a region. Moreover the protocol saves all the overheads incurred in dynamic cluster formation and CH selection, by providing a static cluster formation and a CH selection in a round robin manner. Finally, each round consists of more than one steady phases unlike LEACH, which requires re-clustering and CH selection for each round. This also saves a lot of energy and time.

2. Related Work

Recent years have witnessed many clustering protocols (Taewook Kang, Jangkyu Yun, Hoseung Lee, Icksoo Lee, Hyunsook Kim, Byunghwa Lee, Byeongjik Lee, Kijun Han. 2007)(J. Burrell, T. Brooke and R. Beckwith. 2004)(R. Cardell-Oliver, K. Smettem, M. Kranz and K. Mayer. 2004)(D. Culler, D. Estrin and M. Srivastava.)(A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler and J. Anderson. 2002), LEACH (D. Culler, D. Estrin and M. Srivastava.) being ancestor of most of them. LEACH randomly selects a few SNs as CH, based on certain probability function and rotates their role to balance the energy dissipation of the sensors in the network. This rotation is done after every round. This repetitious set-up processes results in unnecessary energy consumption and delay. Its randomized nature creates clusters with non uniform sizes leading to an uneven load distribution.

PEGASIS (A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler and J. Anderson. 2002) an enhancement over LEACH is a near optimal chain-based protocol. The chain construction is performed in a greedy fashion. In order to find the nearest neighbors for the chain formation the protocol uses RSSI. This ensures smaller inter-nodal distances for providing short communication ranges and thus saves the communication cost.

TEEN (R. Cardell-Oliver, K. Smettem, M. Kranz and K. Mayer. 2004), a LEACH based protocol was developed for time critical applications. It defined two parameters ST and HT to further reduce the overall transmission towards BS. These parameters also provide a reactive behavior to the network in contrast to many others who require periodic transmission of data. HEED incorporates communication range limits and intra-cluster communication cost information for the decision of selection of CH.

3. System Model

In LBA we have assumed that the sensors are distributed in an evenly randomized manner throughout a rectangular field and the network has the following properties:

1) A unique BS with powerful radio, processor and no battery issues.

2) sensor nodes are static.

3) Network is homogeneous in computation and communication capacity.

4) The network is location unaware i.e. physical location of nodes is not known a priori.

5) The transmitter can adjust its amplifier power based on the transmission distance.

The first assumption considers a mobile BS. The sink mobility is desirable (Bin Wang, Dongliang Xie, Canfeng Chen, Jian Ma, Shiduan Cheng. 2007)(A. Parsons. 2003) in many applications in order to provide energy efficiency. Moreover in LBA, the mobility is required only once, i.e. immediately after the deployment of SNs. But, because of any reason, if mobility of BS is a big restriction, the protocol provides an alternative to this. The second assumption of lack of mobility of SNs is typical for WSNs employing some clustering or grouping (Jinyang Li, John Jannotti, Douglas S. J. De Couto, David R. Karger, and Robert Morris. 2000) methodology for network organization. SNs with rapid mobility in network degrade the cluster quality, because they frequently alter the organization of cluster. Assumptions like node homogeneity and advance location unawareness are rather advantageous when hardware costs and resource requirements are key issues. In this paper we use the same communicational energy dissipation model as discussed in(A. Manjeshwar and D. P. Agarwal. 2001).

The clustering algorithm must divide the network into disjoint clusters, i.e. if G (V, E) is network deployment graph, C_i the i^{th} cluster and N_c the number of clusters, then,

$$G(V,E) = \bigcup_{i=1}^{N_c} C_i(V_i, E_i) \quad V_i \subseteq V,\ E_i \subseteq E \qquad \ldots (1)$$

We propose LBA, an energy efficient protocol for location aware grid based clustering in wireless sensor networks. The protocol forms clusters in the form of logical grids which are non-uniform in size. The clusters that are near to the BS need to be smaller in size as compared to the clusters away from it in order to provide an even load distribution over the network. This prevents the clusters near the sink from depleting faster than others. This may lead to either network partition or larger communication ranges for the distant clusters leading to more energy dissipation. The cluster formation is static in nature, i.e. once the clusters are formed no re-clustering is required. This prevents lots of overheads as compared to other clustering protocols (R. Cardell-Oliver, K. Smettem, M. Kranz and K. Mayer. 2004)(D. Culler, D. Estrin and M. Srivastava.) that require frequent re-clustering, requiring high setup energy and time. Moreover, the CH selection is very simple and is done from a locally generated list. The scheduling is done in a round robin manner from this list and hence avoids the complex or the probabilistic nature (D. Culler, D. Estrin and M. Srivastava.) of the CH selection process.

Above all, we provide location awareness in the clusters, which is extremely rare in the present day clustering protocols. Even if some protocols provide location information in the sensor networks it either requires costly hardware (GPS) with all SNs or is very complex and uncertain in nature. LBA only requires one-time mobility of BS at initial setup phase. Sink mobility is not a big demand, as many applications require it. This provides energy efficiency in multi-hop routing by preventing formation of 'hot spots' in network. In time critical applications direct CH-BS communication is preferable and sink mobility provides shorter distances for communicating with CHs, thus further enhance energy efficiency.

The grid size for clusters can be set by BS according to the granularity of the area to be monitored. More is the granularity more are the number of clusters, Nc. Thus a proper trade off must be there between grid size (Gs), and Nc. A very big value of Nc is not desirable since it affects both energy and communicational efficiency. An optimal approach is to have as many as 5% of SNs as CHs. But in LBA it may depend on resolution of location awareness required. The clustering algorithm incurs a complexity of O |1| at node level as compared to O |n| or even worse in some cases. By O|1| we mean to say that the complexity of the protocol implementation does not depend upon the node density or on the number of nodes in the cluster. There is a negligible load over the individual node for the cluster formation; rather our protocol exploits the resources at the BS for most of the clustering maneuver.

Finally, LBA is not just a cluster formation protocol but also provides a multi hop path for data forwarding and aggregation. It does not require all nodes to use radio in a powerful transmission mode since most of the SN-CH and CH-CH communication is local. This saves enormous energy as compared to (R. Cardell-Oliver, K. Smettem, M. Kranz and K. Mayer. 2004)(D. Culler, D. Estrin and M. Srivastava.)(A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler and J. Anderson. 2002), as they require direct communication with BS.

4. Implementation

This paper uses a simple model for radio hardware energy consumption (A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg. 2006). Thus, for transmitting an *l*-bit message through the distance *d*, the energy that radio spends is:

$$E_{Tx}(l,d) = \begin{cases} lE_{elec} + l\varepsilon_{fs}\, d^2 & ,d < d_0 \quad \ldots \quad (2) \\ lE_{elec} + l\varepsilon_{mp}\, d^4 & ,d \geq d_0 \quad \ldots \quad (3) \end{cases}$$

for receiving this message, the radio expends:

And

$$E_{Rx}(l) = lE_{elec} \qquad \ldots \qquad (4)$$

The electronics energy, $Eelec$, depends on factors such as the digital coding, modulation, whereas the amplifier energy, $\varepsilon_{fs}\, d^2$or, $\varepsilon_{mp}\, d^4$ depends on the transmission distance and the acceptable bit-error rate. Like any other hierarchical routing protocol LBA also has a setup phase and a steady phase.

Once the nodes are deployed, the BS will locate itself at a distance, B_d, from one edge of the field. It then broadcasts a signal with strength, B_s, sufficient to reach all SNs. On receiving this signal the nodes calculate their distance from the BS based on the received signal strength using the free energy dissipation model (MICA2 Series):

$$P_r(d) = P_{tx} \times \frac{\varepsilon}{d1^2} \qquad \ldots \qquad (5)$$

Where, the power of received signal is P_r, $d1$ is distance and ε the attenuation coefficient.

Form (5) $d1$ can be calculated as: $$d1 = \frac{r}{\sqrt{p_r}} \cdots \quad (6)$$

Where, r is constant.

(All energies using (2), (3) and (4))

1) First, according to the granularity of location information needed, let the application decide the number of vertical grids, G_n, to be formed. The numbering of grids is as shown on the Figure1.

2) Now the n[th] grid ($G_i = n$), i.e. the grid nearest to the BS has the following energy equation,

$$E_n = G_n * E_{BS} + (G_n - 1) * e_r + (n_{nc} - 1) * e_r \quad \ldots (7)$$

E_{BS}, is the energy required to transmit G_n, packets to BS. n_{nc} is the number of nodes in n[th] grid (one is CH itself, thus $(n_{nc} - 1)$) and e_r is the energy for receiving the cluster member and inter cluster packets.

3) And for $G_i < n$ the energy equation will be:

$$E_i = G_i * E_{IC} + (G_i - 1) * e_r + (n_{ic} - 1) * e_r \quad \ldots (8)$$

E_{IC} is energy for inter cluster transmission, n_{ic} is the number of nodes in i[th] grid.

4) For load to be balanced energy of any two grids must be same, i.e. (7) = (8). Equating the R.H.S. of two equations and simplifying them we get:

$$(n_{nc} - n_{ic}) = (G_i - G_n) + [G_i E_{IC} - G_n E_{BS}] / e_r \quad \ldots (9)$$

5) Also for some other i' grid:

$$(n_{nc} - n_{i'c}) = (G_{i'} - G_n) + [G_{i'} E_{IC} - G_n E_{BS}] / e_r \quad \ldots (10)$$

$n_{i'c}$ is the number of nodes in i' grid.

6) (9) - (10) and simplifying,

$$n_{i'c} - n_{ic} = (G_i - G_{i'}) + (G_i - G_{i'}) E_{IC} / e_r \quad \ldots (11)$$

But, for adjacent grids $(G_i - G_{i'}) = 1$

$$\therefore n_{i'c} = n_{ic} + \Delta \quad \ldots (12)$$

Where, $\Delta = 1 + E_{IC} / e_r$

7) Let, x be the number of nodes in (n-1)[th] grid, then for any i^{th} grid ($i < (n-1)$)

$$n_{ic} = x + ((n-1) - i)\Delta \quad \ldots (13)$$

8) Also from (9),

$$n_{nc} = n_{ic} - (G_n - G_i) - [G_n E_{BS} - G_i E_{IC}] / e_r$$

… (14)

For $(i = n - 1)$, $(G_n - G_i) = 1$ and $(n_{ic} = x)$, so simplifying (14) we get,

$$n_{nc} = x - G_n (E_{BS} - E_{IC}) / e_r - \Delta \quad \ldots (15)$$

9) From (n-1) grid onwards the number of nodes is in Arithmetic Progression. So, the sum of all nodes from 1to n-1 grid is,

$$S_{n-1} = (G_n - 1)[x + (G_n - 2)\Delta / 2] \quad \ldots (16)$$

Now, if N is the total number of nodes then,

$$N = S_{n-1} + n_{nc} \quad \ldots (17)$$

Solving and simplifying (17) we get,

$$x = N / G_n + (E_{BS} - E_{IC}) / e_r - \Delta(G_n - 3) / 2 \quad \ldots (18)$$

10) Using (18), (13) and (15) we can calculate the optimal number of nodes in each vertical grid for achieving a uniform load distribution. Now, if ρ is the density of nodes in network, then we can easily calculate the area required for each vertical grid, a,

$a = x / \rho$, for the (n-1)[th] Grid.

11) For calculating the exact number of nodes in each cluster, divide this area with the number of horizontal grids, H_G. This gives area of each cluster, A_C, beginning from near to BS till end of field. The BS can now easily estimate the positioning of these logical grid lines over the network for cluster formation.

$A_C = a / H_G$, for the (n-1)th Grid.

12) Now, ether the sensors can be manually deployed in such a way that the nodes are optimally distributed in the field as per the formulation or the following strategy can be used.

13) The BS based on above calculations broadcasts the Cluster Formation Packet (CFP). The header of this packet is the allocated Cluster ID and remaining fields are upper left and bottom right distance coordinates of the respective cluster $((X_1, Y1),(X_2, Y2))$.

14) The receiving SNs now check their own distance coordinates with respect to CFP in order to find the cluster they belong to.

$X_1 \le d1 \le X_2$, $Y_1 \le d2 \le Y_2$

Less frequently mobile nodes can repeat this process to calculate new locations after a predefined time. The cluster grids can remain same. The nodes now locally broadcast their Cluster IDs along with their respective distance coordinates ($d1$, $d2$) in form of a beacon packet. This is done in order to let all cluster members (with same Cluster ID) to know each other. In order to prevent collisions in this initial stage we can use some multiple access approach (TDMA etc.) or nodes may generate a small random number to decide the time after which to broadcast this beacon packet. The TDMA schedule may be set by BS also.

5. Performance Evaluations

The following simulation parameters are considered for the implementation of EEPUSH:

The distance between the BS and the network is taken as 100m.

the size of data packet is *512* bytes

the electronic power is *50* nJ/bit

free space attenuation coefficient is *12* pJ/bit/m^2

multipath attenuation coefficient is *0.0012* pJ/bit/m^4

nodes' initial energy is *6.0* J

A node is treated as dead when its remaining energy is less than *0.002* J

For simplicity, error free communication links are assumed. We assume a square network field.

Selection of an optimal number of clusters to achieve the desired granularity of location awareness is important for an application. First we will see how the selection of number of grid lines along each of the adjacent edges affects the number of clusters formed. The relation is simple; if N_V is number of grid lines along vertical edge and N_H along the horizontal edge then the number of clusters, Nc formed are:

$$Nc = (N_H + 1) * (N_V + 1) \quad \ldots (20)$$

More number of clusters is generally not desirable as they increase average energy consumption per cluster but in LBA it depends on the granularity of the location awareness required.

We define a metric called average clustering degree, avClstDeg, as the average number of SNs per cluster in a WSN. This is calculated as:

$$avClstrDeg = \left(\sum_{i=1}^{N_c} ClstrDeg_i\right) / N_c \quad \ldots (21)$$

Where, $clstrDeg_i$ is the number of nodes in a particular cluster. Now we will see how Nc affects the avClstrDeg. This reason for this consideration is the fact that generally intra-cluster communication takes place in TDMA manner.

Thus if time slot for one node is t_s, then turn of one node comes after approximately,

$$T(turn) = t_s * (avClstrDeg - 1) \quad \text{Time.} \quad \ldots (23)$$

The bigger the value of $avClstrDeg$, bigger is $T(turn)$. This gives more time to a node for rest between two alternate turns and saves much energy. Figure 3 clearly shows that more is the Nc lesser is the avClstrDeg and thus lesser is the energy efficiency.

Figure 4 shows the relation between the number of vertical grids (N_G), n_{nc} and x. This will help an implementer to set the number of deployed nodes in clusters nearest to the BS and those farther away from it.

Re-clustering are major issues in any clustering protocol, but our protocol saves the entire headache by providing One-Time cluster formation with uniform load distribution. Figure5 gives a comparison of Nc for first six rounds of LEACH and LBA. LEACH constantly changes the logical network configuration whereas LBA shows a consistent behavior, thus providing energy efficiency.

Next we compare initial setup cost and steady phase cost for the LBA with that of LEACH. We can easily visualize from figure 6 and 7 that the difference is mainly because of the fact that the setup process in our approach mainly includes only reception of control packets and the steady phase requires multi hop communication. Because of these two facts our protocol provides energy efficiency and load balancing. Any protocol requiring lots of initial transmissions of control packets for setup process (which is generally the case) will certainly consume more energy as compares to ours.

Now we consider the total energies incurred for different number of rounds of both LBA and LEACH.

The figure clearly depicts that LBA provides more energy efficiency as compared to LEACH protocol. This can be attributed to static and Load balanced non uniform clustering in LBA and least overheads in the next CH selection process that too not after each round but n_s rounds.

$$one\ round = setup\ phase + n_s * steady\ phase \quad \ldots (19)$$

This saves energy and prevents the delay induced.

References

A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg. (2006). Near-optimal sensor placements: Maximizing information while minimizing communication cost, in Proceedings of Information Processing in Sensor Networks.

A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler and J. Anderson. (2002). ”Wireless Sensor Network for Habitat Monitoring. Proceedings of ACM International Workshop on Wireless Sensor and Application, 2002.

A. Manjeshwar and D. P. Agarwal. (2001). "TEEN: a routing protocol for enhanced effciency in wireless sensor networks," In 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, April 2001.

A. Parsons. (2003). Software Engineering a Wireless Sensor Network for Environmental Monitoring. Honours Project, December 2003.

Bin Wang, Dongliang Xie, Canfeng Chen, Jian Ma, Shiduan Cheng. (2007). Deploying Multiple Mobile Sinks in Event-Driven WSN, IEEE International Conference 1, 27-29

Chih-YuWen and William A. Sethares. (2005). Automatic Decentralized Clustering for Wireless Sensor Networks, EURASIP Journalon Wireless Communications and Networking, 5, 686–697©

D. Culler, D. Estrin and M. Srivastava. Overview of Sensor Networks. webpage, http://www.motelab.org/papers/sn-culler.pdf

J. Burrell, T. Brooke and R. Beckwith. (2004). Vineyard Computing: Sensor Networks in Agricultural Production. Pervasive Computing, 2004.

Jeffrey Hightower and Gaetano Borriello, Location systems for ubiquitous computing. (2001). *IEEE Computer*, 34,. 57-66.

Jinyang Li, John Jannotti, Douglas S. J. De Couto, David R. Karger, and Robert Morris. (2000). A scalable location service for geographic ad hoc routing, mobicomm, 1, 56-62.

MICA2 Environmental Data Acquisition Module (MDA300). webpage, http://www.xbow.com/Products/productsdetails.aspx?sid=77

MICA2 Series(MPR4x0). webpage, http://www.xbow.com/Products/productsdetails.aspx?sid=72

R. Cardell-Oliver, K. Smettem, M. Kranz and K. Mayer. (2004). Field Testing a Wireless Sensor Network for Reactive Monitoring. Technical Report UWA-CSSE-04-003, August 2004.

Redwoods go high tech: Researchers use wireless sensors to study California’s state tree.webpage, http://www.motelab.org/papers/sn-culler.pdf

Taewook Kang, Jangkyu Yun, Hoseung Lee, Icksoo Lee, Hyunsook Kim, Byunghwa Lee, Byeongjik Lee, Kijun Han. (2007). "A Clustering Method for Energy Efficient Routing in Wireless Sensor Networks", Proceedings of the 6th WSEAS Int. Conf. on Electronics, Hardware, Wireless and Optical Communications, Corfu Island, Greece, February 16-19, 2007.

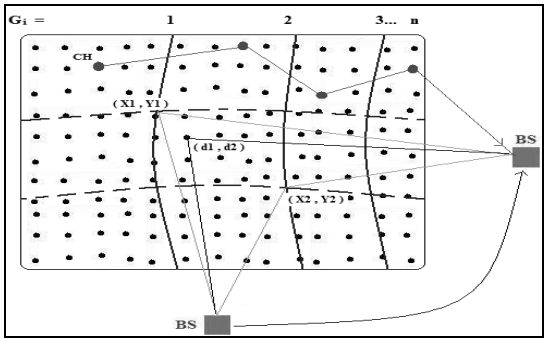

Figure 1. Grid Formation and Distance Coordinate Formation

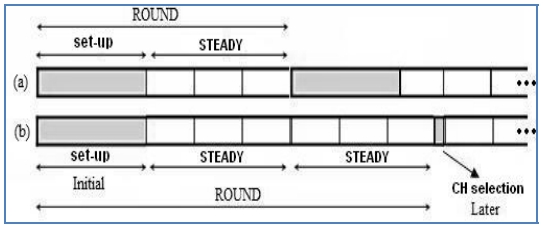

Figure 2. Time-Line a) LEACH b) LBA

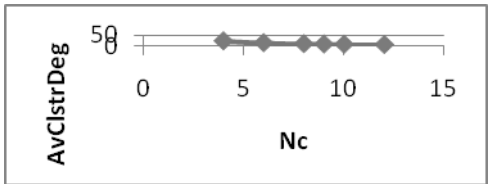

Figure 3. AvClstrDeg Vs N_c

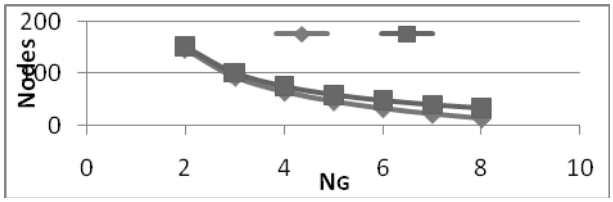

Figure 4. Variation of Nodes in Each Grid w.r.t. N_G

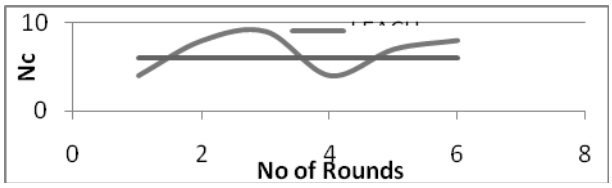

Figure 5. N_c Vs Number of Rounds

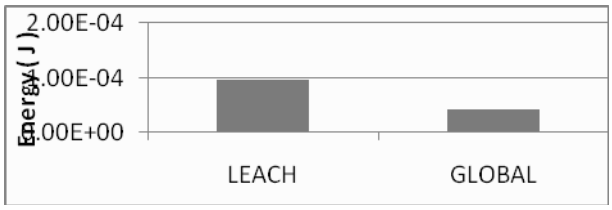

Figure 6. InitialSetup Cost for one CH

Figure 7. Steady phase cost for all CHs

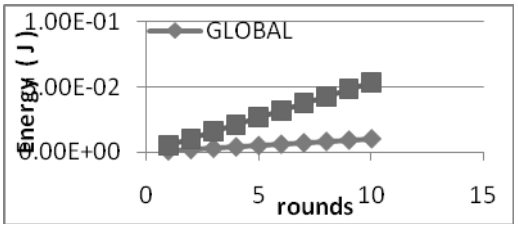

Figure 8. Energy comparison (taking n_s=5)