IACSIT International Journal of Engineering and Technology Vol. 2, No.1, February, 2010
ISSN: 1793-8236

Distributed Clustering with Restricted Number of Clusterheads for Energy Efficient Data Gathering in Wireless Sensor Networks

Bhaskar P. Deosarkar, *Member, IACSIT*, Narendra Singh Yadav, *Member, IACSIT* and R.P.Yadav, *FIETE*

***Abstract*—Energy being major constraint and data gathering the core operation, considerable attention has been given to energy efficient data gathering in wireless sensor network research and clustering has emerged as an efficient architecture to achieve it. Low energy adaptive clustering hierarchy (LEACH), proposed in past, guided significant portion of this research and is used as a benchmark in this area. However, use of probabilistic and randomized approach for clusterhead selection results in their variable number, during different rounds of data gathering. Use of spread spectrum technique, to allow interference free parallel communication in different clusters, results in increased data size. The variable number of clusterheads requires the spreading factor to be set corresponding to probable maximum number of clusterheads, to avoid interference. This additional increase in data size increases amount of energy required to transmit it to the base station, directly affecting network lifetime and energy efficiency, both. To reduce the spreading factor and to improve data gathering efficiency a distributed clustering scheme is proposed in this paper, which restricts the number of clusterheads whenever more than the desired numbers of nodes select them to acquire the role. Results of the simulations, carried out in NS2, show that with the proposed scheme the amount of data received at base station per unit of energy consumed is significantly more than the benchmark scheme used. Proposed scheme is observed to provide improvement of 58% in network lifetime and 48.76% in energy efficiency over LEACH.**

Index Terms— Clustering, Data Gathering, Energy Efficiency, Network Lifetime, Wireless Sensor Network

I. INTRODUCTION

Use of clustering for energy efficient data gathering is widely discussed by wireless sensor network researchers, over last decade. These networks consisting of large number of tiny sensor nodes *capable of sensing the environment,* processing the data and communicating with each other or with the centralized administration, are envisioned to have numerous applications (like environment monitoring, traffic monitoring, health monitoring [1][2][3] etc.). Such applications demand unattended dense deployment of tiny sensor nodes. Use of wireless links, tiny nature of sensor nodes and their deployment environment imposes certain restrictions on operation of such networks and major among them is on the battery replacement. Further, communication activities of sensor nodes are considered to be more energy consuming than their computational activities [4]. Therefore, it becomes necessary to make wireless sensor network survive for longer time period with limited available battery power, by reducing communication burden on these nodes to a level as low as possible.

Dense deployment of sensor nodes helps in achieving high degree of fault tolerance during data gathering. Large number of nodes sense the same environment but transmission of similar data generated at these nodes not only result in their unnecessary energy consumption but also it increases redundancy in the data reported at centralized administrator (Most commonly referred as base station). To avoid such redundant data transmission, clustering is said to be more energy efficient and scalable architecture [5][6] in which, with division of the network, sensor nodes organize them into small groups called clusters and then select one sensor node, termed as clusterhead, in each of these clusters to coordinate activities of their member sensor nodes. During data gathering phase, clusterhead collects data from all its cluster members and after aggregation transmits it to base station. With this approach, sensor nodes acting as clusterhead consume more energy compared to their cluster members, due to their long range transmissions to distant base station. When clusterhead energy reduces to a level at which it can not reach base station, the cluster data may not be communicated. This ends with uneven energy consumption of sensor nodes in the network.

A major breakthrough was provided by Low Energy Adaptive Clustering Hierarchy (LEACH) proposed in [7], with suggestion of rotation of clusterhead role among the sensor nodes, for distribution of energy consumption over the network. Motivated from LEACH, so many clustering schemes are reported in the literature, for energy efficient data gathering which may broadly be categorized as

Manuscript received July 14, 2009. Accepted October 22, 2009.

Bhaskar P Deosarkar is Lecturer in Electronics with Government Polytechnic, Nanded, Maharashtra State, India and is presently on deputation for PhD at Malaviya National Institute of Technology, Jaipur, Rajasthan, India.
(phone: +91 9887322955; e-mail: bhaskar44_nanded@yahoo.co.in).
Narendra Singh Yadav is research scholar in the Department of Electronics and Communication Engineering, Malaviya National Institute of Technology, Jaipur, Rajasthan, India (e-mail: narensinghyadav@yahoo.com)
R.P.Yadav is Professor and Head, Department of Electronics and Communication Engineering, Malaviya National Institute of Technology, Jaipur, Rajasthan, India. (e-mail::rp_yadav@yahoo.com).

centralized or distributed. In centralized schemes, e.g. [8] [9] [10], base station is responsible for cluster formation utilizing information like node locations, remaining energy communicated by sensor nodes, individually during each instance of cluster set up. However, collection of all the necessary information at the base station makes them both time and energy consuming. And in distributed schemes, sensor nodes take decision about their clusterhead role autonomously, either stochastically [7] [11] [12] or in deterministic fashion [13] [14].

Centralized schemes are guaranteed to select exactly equal number of nodes as clusterheads, during each round of data gathering, but collection of all the necessary information at the base station is energy consuming [5] and distributed schemes have advantage in this respect.

Among distributed schemes, LEACH is most popular protocol but is reported, in the literature [8] [9] [10] [11] [12] [13] [14] [15] [16], to have many limitations and most of them may be credited to its clusterhead selection strategy used [15].

Apart from disadvantages listed in the literature, another factor affecting energy efficiency of clustering algorithm is the number of nodes selecting them as clusterheads. All clustering schemes use CDMA codes in clusters to avoid interference [16]. To use CDMA code, signal needs to be spread and hence number of bits to be transmitted gets multiplied by the amount of spreading used, increasing energy consumption. Due to randomized approach suggested for clusterhead selection, in LEACH scheme, number of clusterheads during different rounds is not fixed which demands use of enough amount of additional spreading [19], to avoid loss of data due to inter cluster interference.

Secondly, in LEACH scheme, the role of clusterhead is proposed to be rotated among the sensor nodes after specific time, termed as *round*, and a node once become a clusterhead is not supposed to participate in clusterhead selection process for next 1/P rounds, where P is the desired percentage number of nodes to be selected as clusterhead during each round. After 1/P rounds, if selects it to be clusterhead, i.e. second time, a sensor node may not have sufficient energy to complete the round and in such cases early death of clusterhead may cause its cluster member data to get lost.

So, to restrict number of clusterheads during each data gathering round and to reduce loss of data due to sudden death of clusterhead, a clustering scheme termed as Restricted Clusterhead LEACH (RC-LEACH) is proposed, in this paper, which for increasing energy efficiency does not allow sensor nodes located in farthest network corners, as seen by the base station, to become clusterhead and then whenever more than specific number of nodes declare their candidature for clusterhead role, restricts the number to desired one.

The simulations carried out in NS-2 [17], with incorporation of MIT uAMPS project sensor network framework [18], show that the proposed approach has capability to provide an improvement, over LEACH, in terms of data gathering efficiency, energy efficiency and network lifetime.

Major contributions of this paper include:

- Performance of LEACH scheme is evaluated to observe the effect of spreading factor and consistency of its performance.
- A distributed clustering scheme with restricted number of clusterheads is proposed.
- LEACH and RC-LEACH are simulated in NS-2.
- Analysis of RC-LEACH is carried out in terms of energy efficiency, energy consumption rate, data gathering rate and number of nodes alive through comparison of its performance with LEACH.

Rest of the paper is organized as follows. Network model used, for the study carried out, is described briefly in section II. Overview of LEACH scheme along with its performance in terms of its consistency over different runs of simulations and effect of spreading factor are presented in section III. Proposed clustering scheme is described in section IV. Parameters used for simulation of RC-LEACH are described and simulation results are presented in section V and finally, section VI concludes the paper.

II. Network Model

In this section, network and radio energy models used for the study, carried out in this paper, are described.

A wireless sensor network consisting of *S* nodes (s1, s2, s3,…,…,Sp) is considered to be deployed over a rectangular region of size M x N with coordinates in the range of [0, 0] to [M, N] such that M > 0 and N > 0. Sensor nodes are considered to be deployed in a random fashion across the network and once deployed are assumed to be static. They are considered not to have any location information or location finding hardware installed with them and depends on Received Signal Strength Indicator (RSSI) for estimation of distance to the transmitter. The sink node or base station is assumed to be located outside the network at [Bs(x), Bs(y)] with a minimum distance d_{min} from the network and is considered to be constraint free. In data centric applications like environment monitoring, sensor nodes are densely deployed over a comparatively smaller geographical region and in such scenario, sensor nodes are considered to be within the communication range of each other and are expected to be capable of reporting to base station. Hence, with no loss of generality these nodes are assumed to have sufficient range to reach each other. The nodes further are assumed to estimate their accurate distances to base station with received signal strength of the beacons transmitted by base station. The sensor nodes are assumed to have the capacity to eliminate data redundancy and to reduce the communication load through data aggregation. They are further assumed to be equipped with CDMA facilities.

Radio energy model as in [7] is used for this study which uses a 914 MHz radio. The node radio energy consumed in transmission is

$$E_{Tx}(m,d) = \begin{cases} m \times E(elec) + (m \times E_{fs} \times d^2) & d < d_o \\ m \times E(elec) + (m \times E_{mp} \times d^4) & d \geq d_o \end{cases} \quad (1)$$

where, *m* is the number of bits transmitted, *d* is the distance between transmitter and receiver and d_o is the distance constant referred as crossover distance. And for receiving the *m* bit message the node radio consumes

$$E_{Rx}(m) = m \times E(elec) \quad (2)$$

Contribution of computations to the energy consumption is considered to be negligible in this analysis, as communication cost is much larger than that of computations. The assumed energy required for running the transmitter and receiver electronic circuitry *E(elec)* is 50nJ/bit and for acceptable SNR required energy for transmitter amplifier for free space propagation (E_{fs}) is 100pJ/bit/m^2 and for two ray ground (E_{mp}) is 0.0013pJ/bit/m^4. Crossover distance d_o is assumed to be 87m.

III. Performance Of LEACH Scheme

This section presents a brief overview of LEACH scheme, parameters used during simulations, targeted performance metric and performance of LEACH in terms of its consistency during different runs of simulations and effect of spreading factor.

A. *Overview of LEACH Scheme*

LEACH [7], [19] is a distributed clustering protocol that utilizes randomized rotation of clusterheads for even distribution of energy consumption among sensor nodes in the network. Operation of LEACH is divided into rounds. Each round is of fixed time duration. During each round sensor nodes autonomously decide on their role as clusterhead or a general sensor node by evaluating a threshold function, which is dependent on the desired percentage of nodes to select as clusterhead (*P*), current round number *(r)* and total number of alive nodes (*N*), and its comparison with a random number generated. The sensor nodes once selected to perform the role of clusterhead are not allowed to compete for the role for next 1/P rounds. Each round consists of a set-up phase followed by a steady state phase. During set-up phase, nodes organize them into clusters with each cluster having its own clusterhead, with exchange of short cluster formation messages among them, utilizing CSMA/CA MAC [20] protocol. Clusterheads then prepare the TDMA schedule for their clusters and distribute it to their respective member nodes, to utilize for data forwarding during steady-state phase. The steady state phase consists of transmission of many data frames. Duration of the data frame is dependent on the data size generated at the sensor node and number of data frames during each steady phase is dependent on time required for cluster set up, frame size and round duration. Member nodes transmit their data to their respective clusterhead which then performing aggregation, to eliminate the redundancy, pass it on to base station. To reduce inter-cluster interference, nodes in different cluster communicate using different codes utilizing Direct-Sequence Spread Spectrum (DS-SS).

B. *Simulation Parameters*

Network simulator NS-2 was augmented with MIT uAMPS project (NS-2 extension) sensor network framework and different parameters were set for performance evaluation of LEACH scheme. NS-2, widely used and accepted simulator among network research community, has capabilities to simulate the wired as well as wireless environments. NS-2 is a discrete event simulator in which actions are associated to events. The proposed algorithm was implemented using the node structure as in MIT framework. The channel was assumed to be symmetrical and to have only system losses and not the propagation loss. 100 nodes were considered to be distributed over a network area of size 100 meter × 100 meter. The data packet of 500 bytes and packet header of 25 bytes were considered. Desired number of nodes to be selected as clusterheads during each round of data gathering was set to 5. Propagation model used was two ray ground. The clusterhead change time was set to 20 seconds. All nodes were assumed to have initial energy of 2 Joules and thus, the network to have 200 Joules of energy. Channel bandwidth was set as 1Mbps. Base station was assumed to be located at (50, 175), at a distance of 125 meters from the network center.

C. *Performance Metrics*

Following performance metric was targeted for this study.

1) *Energy Efficiency:* It is defined as the number of data units received at base station per unit of network energy consumed. This is a very important performance metric for data gathering algorithm as its effectiveness depends on the amount of data received under the given energy constraint.

2) *Energy Consumption Rate:* It is defined as the amount of energy consumed per unit time. This also is an important parameter as it gives estimate about duration over which the network may continue to work with available energy and may be seen as a measure of network lifetime.

3) *Data Gathering Rate:* Number of data units received at the base station per unit time is defined as data gathering rate. This gives the quickness of the data gathering algorithm.

4) *Number of Nodes Alive:* This plot of simulation run time versus number of nodes alive gives the time over which network runs. This may be used as a measure of uniformity of energy consumption. It may also be seen as another measure for network lifetime.

D. *Simulation Results*

1) *Performance Consistency*

To observe behavioral consistency of LEACH, scheme was simulated in NS-2.27 extended with MIT uAMPS project (NS-2 Extension) sensor network framework for five different runs and trace files were generated to record the data.

With analysis of the generated trace files, observations about energy efficiency, energy consumption rate, data gathering rate and network lifetime are presented below.

Following Fig.1 shows the amount of data units received with respect to energy consumed in Joules over different runs of simulation. It can be observed from the figure, that the amount of data units received during different runs of simulations is almost constant till the network consumes 80 Joules of energy and after which, a substantial variation is observed during different runs of simulation. This variation may be due to variation in selected number of clusterheads during different rounds and their distribution over the network.

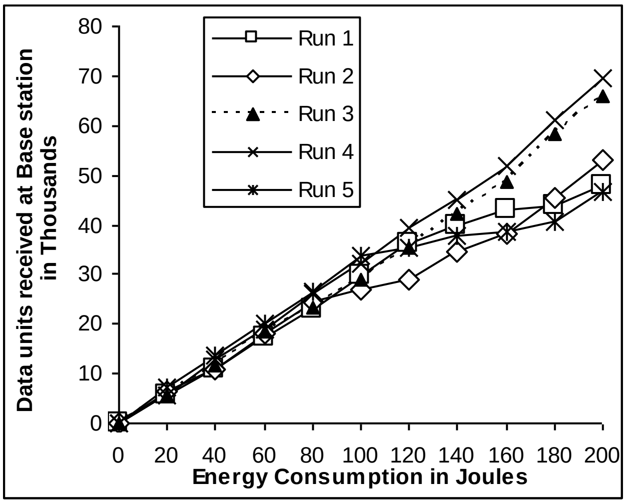

Figure 1: Energy Efficiency of LEACH Scheme

Following Fig.2 is an enlarged version of Fig.1, for results after 80 Joules of energy consumption, to highlight variation in the amount of data units received as a function of energy consumption. It is observed through careful analysis of trace files that the data units received varies from 46811 to 69638, over different runs of simulations; with 18.64% standard deviation around 56635 mean number of data units. Observed 95% confidence interval is 56635±9254. Thus, it does not guarantee to produce similar results during different rounds. It can be concluded from these results that the number of clusterheads and their distribution over network greatly affects the energy efficiency of LEACH scheme.

Energy consumption in Joules is plotted as a function of simulation time for different runs of simulations of LEACH scheme in the Fig. 3, to follow. It is observed that the time over which network runs successfully and gathers the data at the base station varies from 387.6 seconds to 560 seconds, with 17.1% standard deviation around the mean of 461.96 seconds. Variation observed with 95% confidence interval of 461.96±69.08 seconds may be credited to concentration of clusterheads only in a part of the network, selection of no clusterhead during some rounds and selection of variable number of clusterheads.

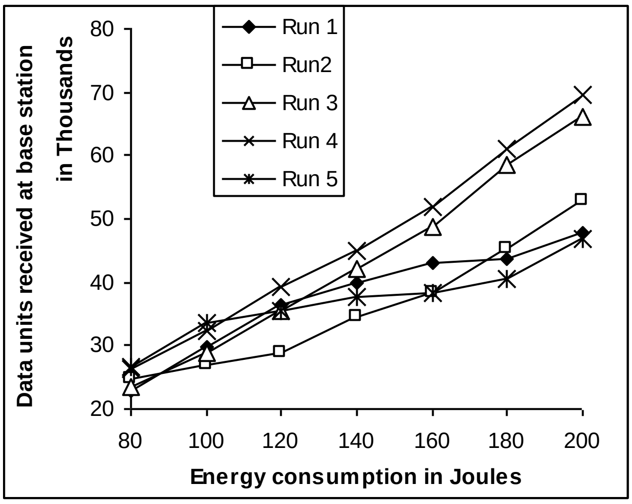

Figure 2: Energy efficiency of LEACH Scheme (Enlarged)

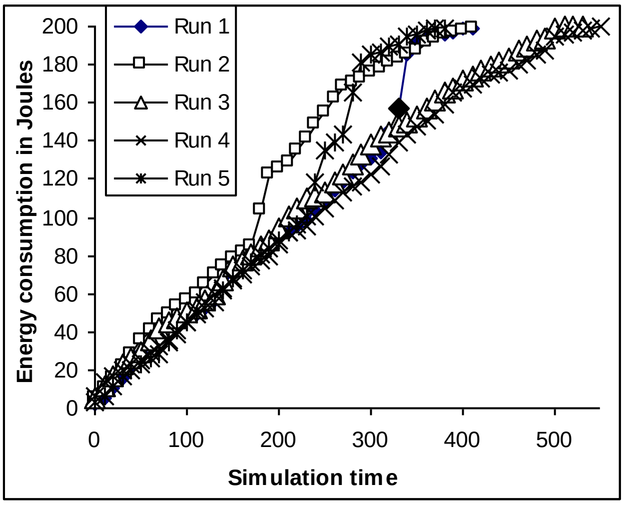

Figure 3: Energy Consumption Rate of LEACH Scheme

In the case of selection of more than desired number of sensor nodes as clusterhead, increased number transmissions to base station increases the network energy consumption.

Secondly, with increase in number of clusters the number of codes required to avoid inter-cluster interference increases. If spreading factor is fixed as K, it can allow interference free communication for clusters up to K-1. If number of clusters exceed this limit and selected clusterheads get concentrated in a part of the network, the inter cluster interference in inevitable.

The number of data units received as a function of simulation time is plotted for five different runs of simulation of LEACH scheme, in the following Fig. 4. It can be observed from the figure that, the reception rate of data units at base station is almost constant during different runs of simulation. However, the number of data units received before exhaustion of network energy varies during different runs of simulation. The numerical figures quoted in the legend brackets give the network lifetime which show a large variation in the time over which base station receives the data, during different runs of simulations.

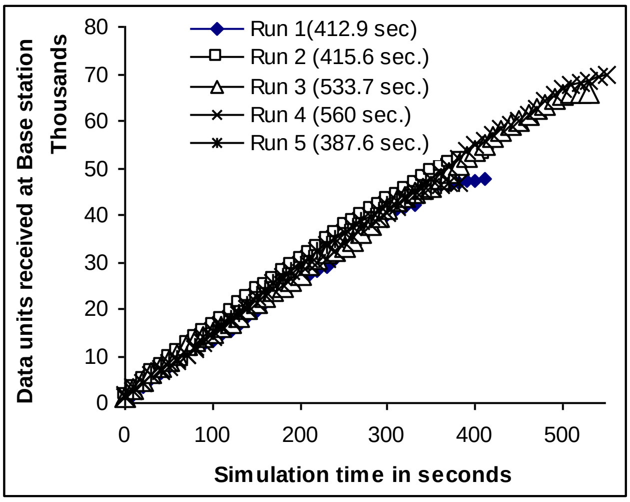

Figure 4: Data Gathering Rate of LEACH Scheme

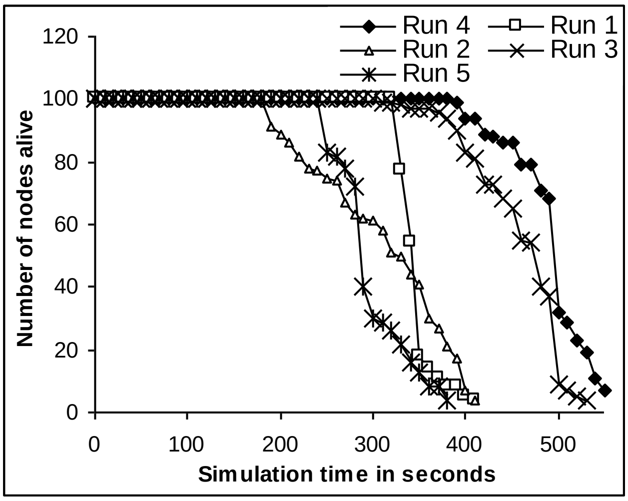

Figure 5: Number of Nodes Alive

Above Fig. 5 gives the number of nodes alive over a simulation run time for different runs of simulation of LEACH scheme. The time at which first node dies varies from 200 seconds to 400 seconds and time at which the number of nodes alive attains a value less that the desired number of clusterheads varies from 387.6 seconds to 560 seconds. Hence, a considerable variation in the rate of node death is observed, over different runs of simulation.

Thus, a general observation that can be made from above results is that there is a significant difference in performance of LEACH during different runs of simulations.

2) Effect of Spreading Factor

LEACH uses DS-SS to make simultaneous communication possible in different clusters. The number of codes required to achieve this is a function of desired number of clusters. One additional code is required for clusterhead to base station data transmission to avoid collisions of these data with the control messages within other clusters, generated during set up. But, LEACH scheme utilizes a randomized approach for clusterhead selection and sensor nodes can make their own decision regarding this role independently, during different rounds of data gathering. Hence, the number of clusterhead varies during different rounds of data gathering. It is observed during different trial runs of simulations that this number varies from 0 to 10, during different rounds of data gathering and during different runs.

To observe the effect of spreading factor on performance of LEACH, simulations were carried out with different spreading factors for sufficient number of times by varying it from 6 to 11 and the best result from each such set of results is presented to observe its capability.

For the set up used, expected number of clusters was set to 5 with 100 nodes forming the network over an area of 100m x 100m. Hence the minimum spreading factor required to be set is 6 and as there can be as much as 10 nodes selecting them as clusterheads, during some round, maximum spreading factor is considered to be 11. Hence, the spreading factor was varied from 6 to 11, to see its

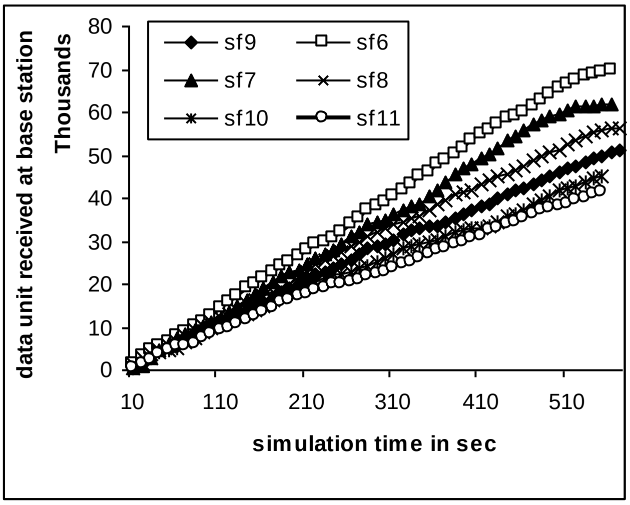

Figure 6: Effect of Spreading factor on Data Gathering Rate

effect. The best observed results, among the results obtained by running the simulations several times, are presented below to highlight the effect of spreading factor on the performance of LEACH and to observe its capability.

Above Fig. 6 gives a plot of number of data units received in thousands as a function of simulation time. It can be observed that the number of data units received linearly increase with time for different spreading factors. Variation in the slope for different spreading factors show that the number of data units received per unit time decreases with increase in the spreading factor. Secondly, the run time over which the network survives is almost constant with different spreading factors. This shows that the LEACH scheme may achieve a network lifetime up to 560 seconds. It is to be noted that the results presented in this section are the best results from the set of observed results, during several runs of simulations, due to which network lifetime appears to be almost same for different spreading factors but there observed is a large variation in it for each spreading factor over different runs of simulation.

The number of data units received in thousands at base station with respect to different spreading factor is plotted in following Fig. 7.

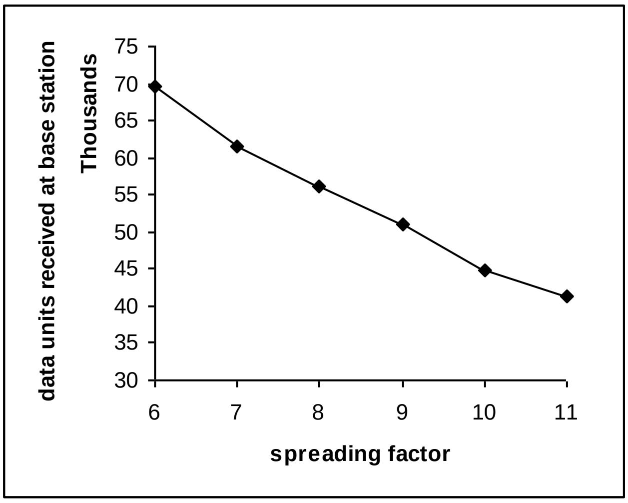

Figure 7: Effect of Spreading Factor on Energy Efficiency

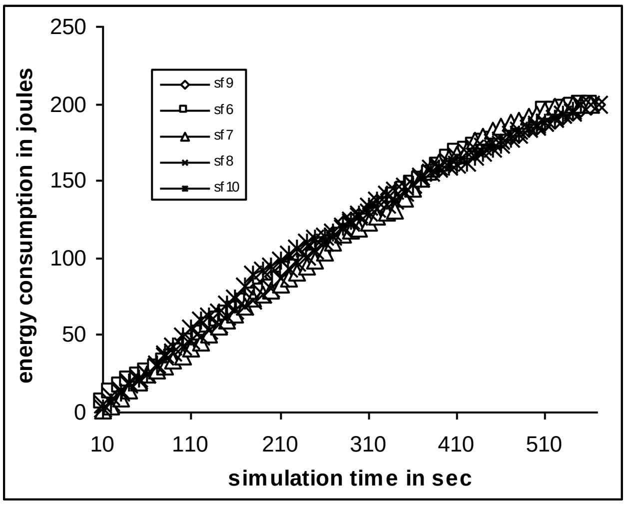

Figure 8: Effect of Spreading Factor on Energy Consumption Rate

A significant reduction in data units received at base station is observed with increase in the spreading factor. This conclusion matches with the conclusion drawn from Fig. 6. This is obvious due to the fact that, with increase in spreading factor number of bits required to be transmitted get increased and the amount of energy consumed is a linear function of number of bits transmitted. Thus there is a direct relation of spreading factor to the energy efficiency of clustering algorithm.

The energy consumption of the network as a function of simulation time is presented in above Fig. 8. It can be observed that, over simulation run time energy consumption varies linearly and is almost same for different spreading factors. Due to this similar energy consumption and increase in number of bits transmitted with increase in the spreading factor, the number of data units received per unit time decreases as shown earlier in Fig. 6. Thus it can be inferred from Fig. 6 and Fig. 8 that the amount of energy consumed per data unit increases with increase in the spreading factor thereby decreasing the energy efficiency of wireless sensor network.

Thus, a common conclusion that can be drawn is, that the spreading factor greatly affects energy efficiency and data gathering rate, both.

IV. Proposed Clustering Scheme

As discussed in earlier section, performance of LEACH, the benchmark clustering scheme, proposed for energy efficient data gathering is not consistent over different runs of simulations, which may be due to improper distribution of clusterheads over the network area and variable number of nodes selecting themselves as clusterheads during different rounds. Secondly, it is also clear that there is a great impact of spreading factor on the performance of LEACH which suggests selection of limited number of nodes, as clusterheads during different data gathering rounds, for it to be more energy efficient. Another observation about LEACH is that when number of rounds exceeds the ratio of number of nodes and expected number of clusterheads, threshold value is set to maximum by all the nodes which forces them all to select as clusterheads and transmit their data to base station and may cause many of them to die suddenly. So, to restrict the number of clusterheads to a desired number, a clustering hierarchy with restricted number of clusterheads is proposed as RC-LEACH in which whenever more than desired number of clusterheads is selected, the excess numbers of nodes voluntarily gets away from the competition and restrict the number. To avoid sudden death of nodes RC-LEACH reduces the clusterhead change time after specific number of rounds.

In proposed RC-LEACH, nodes estimate their distances to base station and based on this distance and the minimum and maximum distances of base station from the network, divides the network into concentric strips. Operation of RC-LEACH also is divided into rounds and each round consists of a set up phase and steady state phase. During set up phase sensor nodes take a probabilistic decision about their role during a particular round and decide to be a clusterhead or a general node. Sensor nodes in the farthest strips, from the base station, are restricted from selecting them as clusterhead.

The detailed operation of proposed RC-LEACH is as follows:

A. *Base Station Distance Estimation*

The base station advertises a *Hello message.* Upon receipt of this message each node in the network estimates its distance to base station (D_{tobs}) using RSSI. Due to work in [21] [22] and [23] it is fairly assumed that the nodes can estimate their distances to transmitter, using RSSI. Again it is assumed that base station knows its location with respect to the network and its distance from the network. Thus, base station has the knowledge of its minimum (D_{bsmin}) and maximum distances (D_{bsmax}) from the network.

B. Clusterhead Selection:

After estimation of base station distance, threshold parameters termed as *netwidth* and *nodewidth* are computed by each sensor node using estimated distance and knowledge about the minimum and maximum distance of the base station from the network, as follows:

$$netwidth = D_{bs\max} - D_{bs\min} \tag{3}$$

$$nodewidth = D_{tobs} - D_{bs\min} \tag{4}$$

Based on these parameters a term is introduced as *round cost* and is defined as

$$round\ \text{cost} = ceil(10 \times \frac{nodewidth}{netwidth}) \tag{5}$$

is evaluated. Less is the *round cost*; less is the amount of energy consumed by the sensor node, in transmission of its data to base station, when selected as clusterhead.

With evaluation of above equation (5), virtually the network is divided in ten concentric strips with nodes in each strip having similar *round cost.* The nodes with larger *round cost* will be at a larger distance from base station and when selected as clusterhead are expected to consume more energy, as compared to the nodes with comparatively smaller *round cost.* Further sensor nodes located in farthest network corner (seen from base station), if select them as clusterheads, may get more penalized due to their distances to base station. Such farthest corner nodes are not better candidates for clusterhead role due to their own energy consumption also. Hence, sensor nodes with *round cost* more than or equal to seven are not allowed to take part in clusterhead selection process.

During each setup phase, sensor nodes evaluate a threshold function as in LEACH and compare it with a random number between 0 and 1 generated at each node, considering the parameters evaluated by equations (3), (4) and (5). If the generated random number is smaller than the calculated threshold, the node decides to become a clusterhead otherwise as a general node. In the case of LEACH, threshold function increases with increase in round number which increases the probability of nodes selecting them as clusterhead creating more number of clusters. The increased number of clusters results in increased number of clusterheads transmitting aggregated cluster data to a distant base station and results in increased energy consumption of the network. This may cause early death of these nodes and reduce the overall network lifetime. So to reduce the probability of selection, the threshold function is scaled down by a factor introduced as *scale factor* which is defined as:

$$scale\ factor = \frac{netwidth - nodewidth}{netwidth} \tag{6}$$

Thus, the threshold function *Th(n)* for any node *n* to be evaluated is defined as follows:

$$Th(n) = \begin{cases} \frac{P}{N - rP} & r < 0.3 \times total\ rounds \\ \frac{P}{N - rP} \times scale\ factor & r \geq 0.3 \times total\ rounds \\ 0 & round\ \text{cost} \geq 7 \end{cases} \tag{7}$$

where, P is the expected number of clusters, N is the total nodes in the network, r is the current round number and

$$totalrounds = \frac{number of\ alive\ nodes}{desired\ number of\ clusters} \tag{8}$$

For distribution of energy consumption over the network, sensor nodes selected as clusterhead during last $1/P$ rounds are not allowed to become clusterhead during current round, as in LEACH.

In LEACH, the role of clusterhead is rotated after a specific interval termed as *clusterhead change time* and when none of the sensor nodes select itself as a clusterhead during some data gathering round, all nodes are expected to transmit their sensed data directly to the base station. This forcefully consumes their energy and reduces the network lifetime. To avoid such energy consumption, whenever nodes do not receive any clusterhead advertise over a specific period, after entering in a setup phase, the setup phase is reinitiated as suggested in [24].

When current round number equals *totalrounds*, current round number is reset to zero. With current round number equal to *totalrounds*, it is expected that all eligible nodes have been in the clusterhead role at least once and consumed their energy in transmitting their cluster data to base station. After resetting the current round number, if a node is selected as clusterhead it may not have sufficient energy to complete the round i.e. over *clusterhead change time* which may result in its earlier death and cluster data may not be reported to base station. To avoid such non- reporting of data and node death, *clusterhead change time* is then reduced or in other words, the round is shortened.

C. Clusterhead advertise

The selected clusterheads advertise their selection over the network area using CSMA/CA MAC protocol and all other nodes despite of their decision about clusterhead role listen to it. Use of CSMA/CA MAC protocol avoids the *hidden terminal* problem [25] and also ensures the ordered reception of clusterhead advertises at each node. To have exactly the same order at each node, each node creates a clusterhead list and if it decides to be clusterhead it advertises its selection over the network and simultaneously appends its clusterhead list. Thus clusterheads are listened by all other nodes in the order they advertise their selection and hence there expected are identical clusterhead lists at each node.

D. Restricting the number of clusterheads

If the number of clusterheads in the clusterhead list is more than the desired number of clusterheads required then the clusterhead list is truncated by keeping only first required number of clusterheads, by each node including the clusterheads. If the selected clusterhead is one from the truncated clusterhead, it behaves as a general node and leaves the clusterhead role for particular round. Thus, with maintenance of clusterhead list at each node and truncating it, if it exceeds the desired number, number of clusterheads during any round of data gathering are restricted to a desired number.

E. Join request

The nodes decided to be general nodes, send joining request to the nearest of clusterhead from clusterhead list, using CSMA/CA MAC protocol, so as to avoid collisions due to hidden terminal problem.

F. Schedule distribution

Each clusterhead prepares a TDMA schedule for the nodes in its cluster and advertises it over the network, again using CSMA/CA MAC protocol.

G. Data transmission (steady state phase)

Data transmission from sensor nodes to base station takes place in two phases: from sensor node to clusterhead and from clusterhead to base station, as in LEACH. During their designated time slot, sensor nodes transmit their sensed data to their respective clusterhead. Upon receiving data from all its member nodes, clusterhead aggregates the data and transmits thus aggregated data to base station.

V. Simulation Results

The results of simulations carried out in NS2 are presented in this section. The proposed RC-LEACH scheme was implemented and simulated in NS-2 and its performance is compared with LEACH. As the number of clusterheads during any round were to be restricted to the desired number of clusterheads i.e. 5, the spreading factor was set to 6. Rest of the parameters, for RC-LEACH, was set similar to LEACH. Performance metric same as defined above for evaluation of LEACH scheme was used. For LEACH, the simulation set up as used, earlier in Section-III, to observe its performance consistency was used.

Following Fig. 9 shows the number of data units received as a function of energy consumption in Joules. It can be observed that, with increase in energy consumption number of data units received at the base station increases almost linearly but the number is always higher in case of LEACH-RC as compared to LEACH. This implies that, the proposed LEACH-RC outperforms LEACH scheme in terms of energy efficiency.

Fig. 10, to follow, shows the variation in the remaining network energy, in Joules, as a function of simulation time in seconds. It can be observed that, the remaining energy decreases with increase in simulation time in case of LEACH and LEACH-RC, both. However, the amount of energy expended per unit time is less for LEACH-RC as compared to LEACH due to which the time taken for network energy to exhaust is larger in the case of LEACH-RC. Thus, proposed LEACH-RC gives an extended network lifetime.

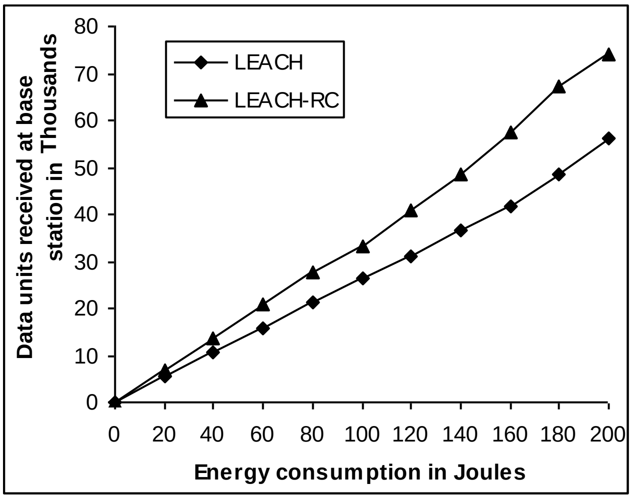

Figure 9: Energy Efficiency

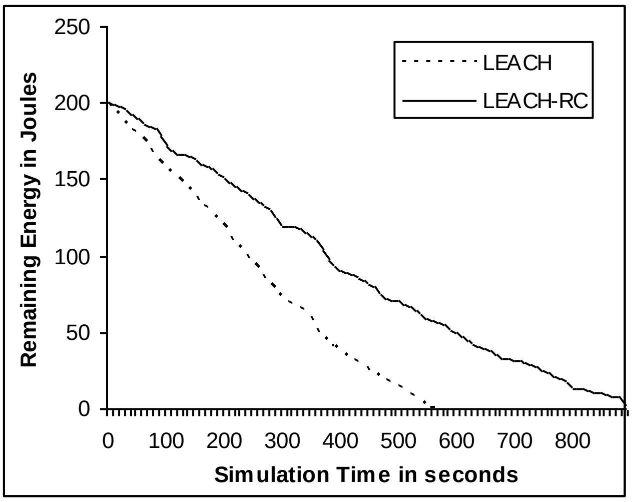

Figure 10: Energy Consumption Rate in terms of Residual Energy

Fig. 11, given below, shows the number of data units received at base station as a function of simulation time in seconds. It can be observed that the amount of data units received increase, with increase in the simulation time. The line corresponding to LEACH is terminated at 560 second indicating that, at this point of time the number of nodes alive is reduced below the desired number of clusterheads. The number of data units received over network lifetime is larger in case of proposed LEACH-RC as compared to LEACH. However, the number of data units received per unit time is comparatively more in the case of LEACH. This implies that the data gathering rate of LEACH is better, but over the network lifetime LEACH-RC outperforms LEACH in terms of energy efficiency.

Number of nodes alive as a function of simulation run time is plotted in Fig. 12, to follow. It can be observed that, in the case of LEACH after 400 seconds number of nodes alive reduces to 4 very quickly whereas in case of LEACH-RC it happens gradually. The sudden death of nodes in LEACH happens due to the fact that, in LEACH when all sensor nodes become clusterhead once round number is set to zero and threshold evaluated for clusterhead

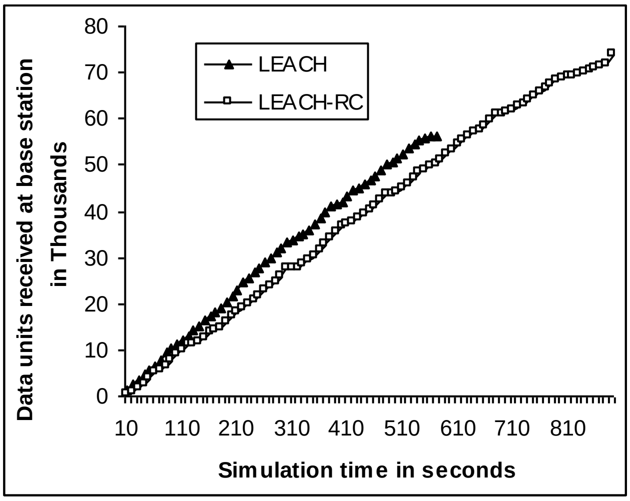

Figure 11: Data Gathering Rate

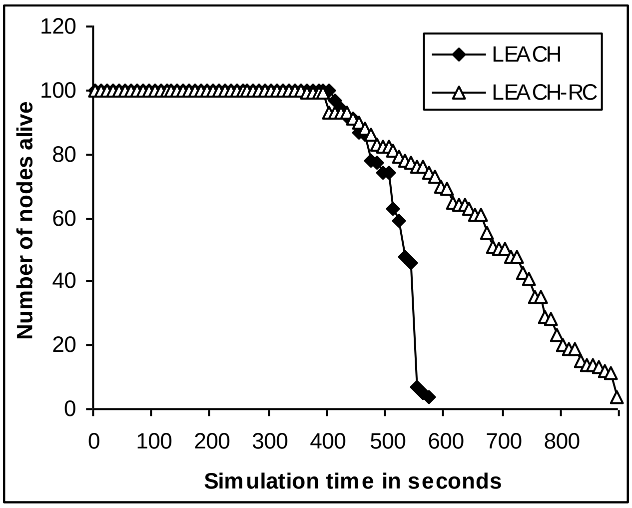

Figure 12: Number of Nodes Alive

selection is maximum (i.e. 1), by all alive nodes, and hence all of them become eligible for clusterhead role. Thus, declaring as clusterhead they transmit their data to base station causing drain out of their energy. Secondly, if selected as clusterhead during successive rounds, the nodes may not have sufficient energy to transmit their cluster data to base station for entire round and may die earlier, as the clusterhead change time remains unchanged. However, in proposed LEACH-RC scheme possibility of sudden death of large number of nodes is reduced by decreasing the clusterhead change time, in such scenario.

Network lifetimes of 560 seconds and 890 seconds are reported for LEACH and LEACH-RC respectively. Thus there recorded is 58% improvement in network lifetime, over LEACH.

VI. Conclusion and Future Scope

A distributed and energy efficient clustering scheme is proposed in this paper as an improvement over LEACH. RC-LEACH restricts the number of clusterheads during different data gathering round to the expected number of clusters to be formed and gives good energy as well as data gathering efficiency. RC-LEACH is shown to outperform LEACH in terms of network lifetime, data gathering efficiency and energy efficiency.

Though RC-LEACH outperforms LEACH scheme, it allows more than desired number of nodes to contend for clusterhead role during current round. Communication of these excess nodes during cluster formation needs to be addressed. Hence, RC-LEACH needs to be improvised to a clustering scheme in which exactly equal number of nodes, that are distributed uniformly over the network area, select them as clusterheads, during each round.

References

[1] I. Akyildiz , E. Stuntebeck, "Wireless underground sensor networks: Research challenges" Elsevier Science Journal of Ad Hoc Networks, Vol.4, 2006, pp. 669–686

[2] I. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirc, "Wireless sensor networks: a survey", Elsevier Science Journal of Computer Networks, Vol.38, 2002, pp. 393–422

[3] M.Nekovee, "Sensor networks on the road: the promises and challenges of vehicular ad hoc networks and grids", In Proc. *Workshop on Ubiquit. Computing and e-Research*, '05, 2005.

[4] L. Ferrigno, S. Marano, V.Paciello and A. Pietrsanto, "Balancing computational and transmission power consumption in wireless image sensor networks", In Proc. IEEE Int. *conf. on Virtual Environments, Human-Computer Interfaces, and MasurementSystems(VECIMS 2005),* July 18-20, 2005, Giardini Naxos, Italy, pp. 61-66.

[5] O. Younis, M. Krunz and S.Ramasubramanian, "Node Clustering in Wireless Sensor Networks: Recent Developments and Deployment Challenges", IEEE Network, May/June-2006, pp. 20-25.

[6] J. Yu and P. Chong, "A survey of clustering schemes for Mobile Ad Hoc Networks", IEEE Communications Surveys and Tutorials, First Quarter 2005, Vol 7, No. 1, pp. 32-47.

[7] W. .Heinzelman, A.Chandrakasan and H.Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks", In Proc.*33rd HICS*, Jan. 4-7, 2000, Vol.2, pp.10.

[8] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks", *IEEE Trans. Wireless Communications,* Vol. 1, No.4, October 2002, pp. 660-670.

[9] P.Tillaport, S. Thammarojsakul, T. Thumthawatworn and P. Santiprabhob, "An Approach to Hybrid Clustering and Routing in Wireless Sensor Networks", In Proc. IEEE Int. Conf. Aerospace, 2005, pp.1-8.

[10] S. Muriganathan, D. Ma, R. Bhasin and A. Fapojuwo, "A Centralized Energy-Efficient Routing Protocol for Wireless Sensor Networks", *IEEE Radio Communication,* March, 2005, pp. S8-S13.

[11] Y. Qian, J. Zhou, L.Qian and K. Chen, "Highly Scalable Multihop Clustering Algorithm for Wireless Sensor Networks", in Proc. *IEEE ICCCAS 2006*, Vol.3, June, 2006, pp.1527-1531.

[12] H.Chen, C. Wu, Y. Chu, C. Cheng and L. Tsai, "Energy Residue Aware (ERA) Clustering Algorithm for Leach-based Wireless Sensor Networks", In Proc. *2nd Int. Conf. Systems and Networks Communications (ICSNC 2007)*,25-31 Aug, 2007, pp.40.

[13] R. Chang and C. Kuo, "An energy Efficient Routing Mechanism for Wireless Sensor Networks" In Proc. *Int. Conf. Advanced Information Networking and Applications (AINA'06),* Apr. 18-20, 2006, Vol.2, pp.5.

[14] M. Rajullah and S. Shimamoto, "An Energy Efficiennt Periodical Data Gathering Protocol Using Deterministic Clustering in Wireless Sensor Networks" In Proc. *Int. Conf. WCNC,* 2007, pp. 3016-3020.

[15] B. Deosarkar, N. Yadav and R.P. Yadav, "Clusterhead Selection in Clustering Algorithms for Wireless Sensor Networks: A Survey" , In *Proc. Int. Conf. Computing, Communication and Networking (ICCCN 2008),* Dec. 18-20, 2008, Karur, Tamilnadu, India.

[16] A. Abbasi, M. Younis, "A survey on clustering algorithms for wireless sensor networks", Elsevier Science Journal of Computer Communications, Vol.30, 2007, pp. 2826-2841.

[17] K. Fall and K. Vardhan, The Network Simulator (ns-2). Available online: http://www.internetworkflow.com/downloads/ns2leach/ns-allinone-2.27.tar.gz

IACSIT
International Association of Computer Science and Information Technology
WWW.IACSIT.ORG

[18] MIT µAMPS project ns2 code extensions. Available online :http://www-mtl.mit.edu/research/icsystems/uamps /research /leach/leach_code.shtml
[19] W. Heinzelman, *Application-Specific Protocol Architectures for Wireless Networks*, Ph.D Thesis, Massachusetts Institute of Technology, June 2000.
[20] Andrew S. Tanenbaum, *Comuter Networks,* Prentice Hall, pp. 892.
[21] J. Rencheng, W. Hongbin, P. Bo and G. Ning, “Research on RSSI-Based Localization in Wireless Sensor Networks”, In Proc *Int. Conf. Wireless communication networking and mobile computing 2008* (Wicom 08), 12-14 Oct.,2008 pp 1-4
[22] Z. Deng, W. Ren and L. Xu, “Localization algorithm based on difference estimation for wireless sensor networks in Pervasive Computing Environment” , In Proc. *Third Int. Conf. Pervasive Computing and Applications, ICPCA 2008, 6-8 Oct, 2008*
[23] W. Wang and Q. Xu, “ RSS-based Monte Carlo localization for mobile sensor networks” , J. IET Communications, Vol..2, No.5, 2008, pp. 673-681.
[24] C. Liu and C. Lee, “Power Efficient Communication Protocols for Data Gathering on Mobile Sensor Networks”, In Proc. *IEEE Int. Conf. Vehicular Technology (VTC’04)*, 2004, pp. 4635-4639.
[25] Hongqiang Zhai and Yugang Fang, “A Solution to Hidden Terminal Problem Over a Single Channel in Wireless AD HOC Networks”, In Proc. *Int. Conf. MiitaryCommunication (MILCOM 2006),* 2006, Oct. 23-25,Washington, DC, pp.1-7.