ENERGY EFFICIENT LOCALIZED CLUSTERING TECHNIQUE FOR ECOMMERCE APPLICATIONS

Abraham Amal Raj B[1], Mahaveer Sain[2]

([1]Department of Computer Science & Informatics, Maharishi Arvind University, Jaipur, Rajasthan)

([2]Associate Professor, Department of Computer Science & Informatics, MAISM, Jaipur, Rajasthan)

ABSTRACT:

Wireless sensor networks (WSNs) cover significant portion of earth's surface, are power-driven through source of finite energy similar to batteries. Thus, WSN's need operates within stringent energy boundaries. Sensing, computing and communication are the major energy consuming activities of a wireless sensor node, of which communication consumes a comparatively higher amount of energy. By reducing the volume of data to be communicated significant level of energy can be conserved. The proposed work contemplates the spatial and temporal correlation between data collected by a WSN. A Localized Energy efficient Clustering Approach (LECA) is a distributed clustering approach where Cluster Heads (CH) are selected depend on their weights using passive clustering technique. Data are sent to base station on a multi hop backbone through a connected dominant set comprises of only CHs. Sensor nodes transmits only a sub set of sensed data to CH through a dual prediction framework. The LECA distributes Cluster heads evenly over the region of interest, thus maintaining uniform and balanced coverage of the region. The proposed method achieves longer life span of WSN, with optimal level of accuracy on collected data. The coverage improvement is also compared with other hierarchical clustering protocols, and found to be best.

Key words:

Wireless sensor networks, energy conservation, uniform coverage, distributed clustering, data reduction

1. Introduction:

Today's industrial, military, and healthcare facilities must include wireless sensor networks because of their lower implementation costs, smaller size, and higher computational capacity. Numerous industrial, commercial, and environmental monitoring applications use WSNs. Owing to their simple installation and reduced maintenance, they proliferated to innumerable disciplines. WSNs are applied for variety of monitoring and surveillance applications. A WSN is made up of independent sensor nodes that are physically spread throughout the targeted area and report on one or more interesting metrics. In most of the cases the nodes collect the raw data from the environment and communicate it to the remote host. A sensor is a node's front end that collects physical input. The electrical output is served on to the processor and initial data processing is done. Then, the preprocessed data is allowed to transmit through an RF transceiver. All the tasks are powered by a battery on the node.

The lifespan of the sensor nodes is restricted since they are independent systems with finite energy resources. The majority of WSN are deployed in hostile situations, making charging challenging. The system life cycle can be extended by energy conservation, ensuring continuing service. Energy is consumed by 3 major components of a sensor node, of which transmission module is the prime consumer.

The very purpose of deploying the WSN is to collect the accurate data over the region during the time span. The data accuracy is one of the major qualities of service factor for WSN. In WSN the system data accuracy is highly dependent on the sampling frequency and the efficiency of the medium. By increasing the amount of samples sent to the host, the host can more accurately reconstruct the measured signal. By increasing the number of samples sent to the sink, the data accuracy can be improved. But, the increased transmission results in decreased residual power on the battery. For a WSN both data accuracy and the node life time are important parameters. So there is a need to optimize the both, with application defined constraints.

Most of the environmental parameters and physical phenomena exhibit slower changes in time domain. The relationship between multiple points of a signal in time domain is termed as temporal correlation. It is possible to significantly reduce the quantity of data transfer by accurately predicting the temporal correlation of a recorded signal. Due to their close proximity, closer sensors measure values that are nearly equivalent in the physical space. By identifying the change in magnitude of the given signal over the space ample redundant data can be avoided.

The proposed work involves grouping of nodes by their relative location. The clustering is allowed to initiate by a highest weight node that asserts itself as CH. The neighbors join the cluster. Then, from the non-clustered nodes, uppermost weight node tends to initiate clustering. This can be continued, till the entire nodes are clustered. The cluster members' start reporting the sensor data to their cluster heads through dual prediction.By clustering of sensor nodes spatial correlation is accomplished and by DPR temporal redundancy is reduced.

2. Related Works

Numerous studies have been conducted on energy-efficient clustering on WSNs. The majority of the studies focus on energy efficiency while maintaining data accuracy. None of the publications address long-term coverage of the region of interest. Only if the distributed sensor network covers a greater area for a longer period of time is it considered to be efficient. The primary focus of this effort is energy efficient coverage.

Different methods are used to study energy conservation in WSN [1]. Some techniques use energy-efficient routing as one of their instruments for wireless node energy balancing [2]. Some techniques make use of MAC-based approaches, which ON/OFF the node circuits at specific time windows to significantly save energy costs [3].

By using clustered aggregation techniques, spatial correlation between the sensor nodes is employed to minimize data transmission. Many clustering techniques are listed [4]. In WSN, clustering is used in various applications to accomplish a variety of objectives.

It is possible for clustering to be centralised [5], in which case the base station gathers node parameters and notifies the cluster chiefs and their members. For the little network, this one works well. Distributed clustering approaches [6], where the grouping decision is performed by individual sensor nodes by obtaining just local information from one hop neighbours, are ideally suited for large scale networks. Clustering may also take into account single-hop [7] or multiple-hop [8] networks. The overhead during CH election is reduced by the use of heterogeneous nodes[9].

One of the principal and top-most popular clustering techniques for energy-efficient WSN is LEACH [7]. Random CH selection in leach causes close CHs and an unequal load distribution, which is an issue. Other works [10] have node degree included as from many of the selection criteria for choosing a CH.

By determining similarities between sensor data, clustering is accomplished. The base station [11] receives the data from the nodes, and based on the similarity of the data, assigns the nodes to the proper cluster.

Depending on the selection criterion, different clustering methodologies utilise different parameters for CH selection. Node degree and node id, among other criteria, are taken into account using deterministic method [12], where these values are fixed during the deployment itself. Adaptive algorithms [13] take mobility, node residual energy, etc. into account. Throughout the network's operational range, these parameters change. For cluster head selection, hybrid algorithms [14] take into account both deterministic and adaptive factors.

A lot of research has been done on how to use data's temporal correlation. The temporal relationship between time series generated by sensors is calculated using linear regression techniques [15]. The schemes struggle with decreased accuracy and a lack of responsiveness to dynamic input signal fluctuations. [16] employs ARIMA-based techniques to forecast sensor data based on prior values. Dual prediction approach employing LMS-based adaptive filters was proposed by Saintini et al. [17]. This method has the advantage of being able to make predictions without using any prior models.

For the various Enterprise IT applications, the vast amount of heterogeneous data can be made open source and linked [18]. It must follow the prescribed format. IOT offers the ability to link URIs, utilise Resource Description Framework, and give users quick access from a variety of sources via the Internet. It is used for IoT and service-oriented WSN integration for e-commerce.

By determining and establishing the cluster head's remaining energy threshold for the entire network, the number of cluster selections in the e-commerce marketing system is decreased [19]. The original cluster head is replaced once the region within the cluster has been partitioned, and the best cluster head is established based on the density of marketing nodes in various locations and the cluster's minimal energy usage. The appropriate number of cluster heads is determined based on the network's minimum energy consumption because the density of marketing nodes varies across different types of e-commerce.

Additionally, it is necessary to detect malicious nodes in the area using little infrastructure and processing. As a result, it is anticipated that an agent-based method will preserve the node's present status [20]. The ratio of packets delivered to packets received will be used to rate each node. Each node votes for the succeeding node in e-commerce models based on the ratio of packets transmitted to packets received. Additionally, the proposed agent-based framework incorporates a node's reputation from its nearby nodes into the calculation of trust.

These existing models require more computation, storage, and complex security calculations. Still, the previous schemes should be upgraded by using the adaptation of the interior parameters.

The majority of the currently available work either focuses on temporal correlation-based data reduction or spatial correlation-based data aggregation. For energy-efficient data collecting, LECA makes use of spatiotemporal correlation. LECA also upholds the highest standards of data integrity. The approach uses straightforward and computationally lightweight techniques, which simplify implementation and use less energy to run.

3. LECA

Assumptions: Before creating the energy-efficient technique, the proposed work makes the following assumptions.

1. Topology is static in one
2. Every node is aware of its position
3. Each sensor node can reach CH with a single hop.
4. CH can make one or more hops to reach the base station.

In order to lower the network's bandwidth requirements in addition to the energy consumption of the sensor nodes, clusters were specifically designed. A single cluster head (CH), coupled to numerous neighbouring sensor nodes, is responsible for attributing the clusters in the network. The cluster head gathers the data from the sensor nodes, aggregates it, and then sends it to the base station. This procedure lowers the nodes' energy costs and lowers the likelihood of data collisions. The traditional cluster's energy model is provided as

$$\sum_{i\in N}\sum_{j\in N}\{\, l y_{ij}\,(D_{ij} + E + E_{DA}) + l_{ag} F_i x_i \,\} \ldots (1)$$

Where y_{ij}=1, when 'j' node is a CH cluster member of 'i' node.

y_{ij}=0, otherwise.

l, (length) data packet size from cluster member to CH.

E, energy required to receive a data bit.

E_{DA} , energy for data aggregation

l_{ag}, size (length) of aggregated data packet

x_i =1, if node ‘i’ is a CH.

x_i =0, otherwise.

Data transmission power is provided from cluster member to CH as

$$D_{ij} = \begin{cases} E + \varepsilon_{fs} d_{ij}^2 & ;if\ d_{ij} < d_0 \\ E + \varepsilon_{mp} d_{ij}^4 & ;if\ d_{ij} > d_0 \end{cases} \text{..........(2)}$$

Where d_{ij}, the distance from sensor node ‘i’ to sensor node ‘j’.

d_0, the threshold distance.

ε_{fs},

ε_{mp},

Data transmission power is provided from the CH to the base station as

$$F_i = \begin{cases} E + \varepsilon_{fs} f_i^2 & ;if\ f_j < d_0 \\ E + \varepsilon_{mp} f_i^4 & ;if\ f_i > d_0 \end{cases} \text{..............(3)}$$

f_i, the distance measured from ‘i’, CH node to Base station.

Weight based election of CH

From the set of nodes the cluster heads are identified by their weights. The node with highest weight than its neighbors in a region is selected as cluster head. The node weight comprises of two components namely residual power and node priority. In each region, a node with highest residual power and priority is elected as cluster head. The significance of the components is user defined.

Let ‘R’ be the current data gathering round, for any node ‘i’, N_i be the count of rounds being as Cluster Head. The priority P_i of node ‘i’ at round R is given as

$P_i=1-(N_i/R)$(6)

The residual power component helps in distributing the nodal energy equally among all nodes. The node degree component improves the efficiency of communication subsystem, so that energy efficient communication is achieved. The priority component reduces consecutive loading on high degree nodes. The weight of node i,Wi_{total}at any point of time is given as

$$Wi_{total} = W_1 P_i + W_2 E_{ir} \text{ ....(7)}$$

Where E_{ir} isthe residualenergy of node ‘i’, in round r.

Distribution of CH

Uniform distribution of CH can only guarantee complete representation from the region of interest. Most of the previous methods involve a probabilistic approach for CH election, which cannot guarantee uniform distribution of CH. Being a deterministic approach, in order to achieve uniform distribution of CH over the region of interest, MECA insists on certain criteria’s.

Any node, which is a one hop neighbor of existing CH, is not eligible for a CH.

$$w_j = \sum_{i \in N} \sum_{j \in N} (k_{ij} \,||x_i) w_j \text{ .... (8)}$$

Where w_j,is the weight of node ‘j’,

x_i = 0; if node ‘i’ is a cluster head, otherwise x_i =1.

k_{ij}= 0; if node ‘j’ is a one hop neighbor of node ‘i’, otherwise k_{ij}=1.

LECA algorithm defines a connected dominating set (CDS) of CHs through which a CH reaches BS. A CH node reaches the base station in one hop, if there is no intermediate CH node in between the CH and BS. The CH nodes which have intermediate nodes reach the BS through multiple hops on intermediate CH nodes. A CDS comprises only of CHs is shown in figure.1.

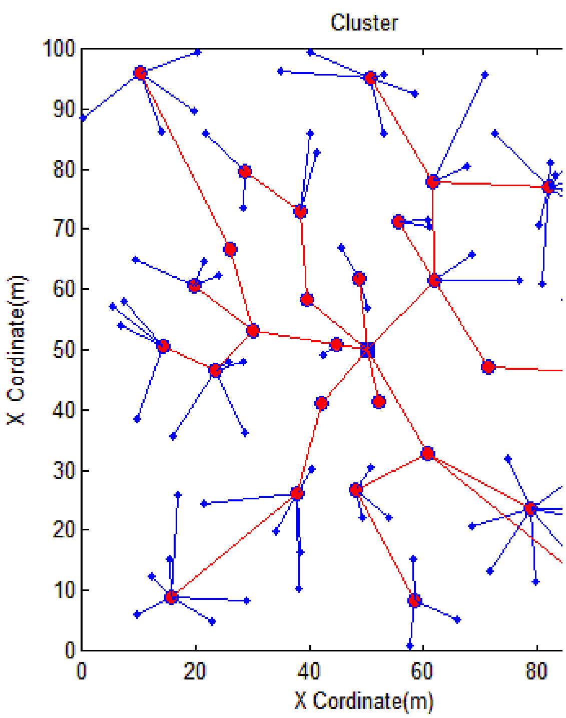

Fig.1.Multi hop routing through CHs

The energy spending of the CH can be rewritten as

$$\sum_{i\in N}\sum_{j\in N}\{\, l y_{ij}(E + E_{DA}) + l_{ag}F_i x_i + l_{jf} x_j m_j(\mathrm{E} + F_i x_i)\,\}$$
………………..(9)

Where l_{jf},is the length of forwarded packet from CH 'j'.

m_j=1, if CH 'j' is in the leaf of CH 'i'.

m_j=0, Otherwise.

As the second constraint says, the CHs near the BS are over loaded, since most of the data are reaching BS through them. in order to distribute load equally among CHs, the CHs nearby the base station are of small size. The CHs in the border region are of large size. The clusters of varying sizes are shown in Fig.2.

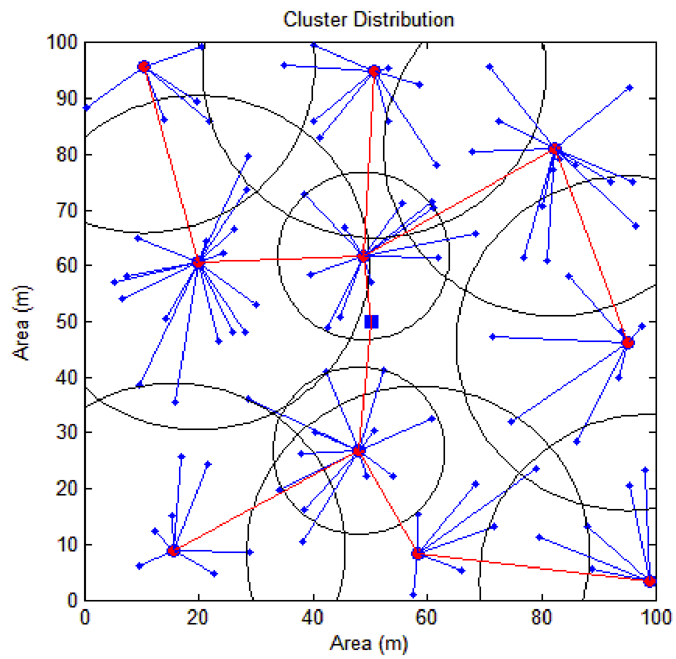

Fig.2.Clusters near and far

By this method the CH nodes are elected based on multiple parameters, which leads to efficient load balancing among sensor nodes. Since the method selects only one CH from one hop vicinity the CHs are distributed well over the sensing region. The weight parameter forces out the border nodes and remote nodes from becoming CH, thus their energy is preserved for continuous coverage of low density and border regions.

Passive Clustering:

It is necessary to alternate the role of CH among nodes so as not to load a fewer nodes with more task than other nodes. Clustering is repeated periodically to heal disconnected regions due to node death and to distribute energy consumption through all nodes. Since the clustering is also based on adaptive parameters that vary dynamically, periodic re clustering is necessary. Since the network is time synchronized after m rounds of data gathering, the cluster head give up its role. LECA follows distributed clustering, where the decisions made by information's gathered from local neighbors. This method consumes a significant amount of nodal energy towards nodal information sharing. LECA reduces this clustering overhead by supporting passive clustering methodology.

During the clustering phase, every node calculates its weight. Every node in the vicinity can proclaim itself a CH, after a certain delay. The proclamation delay of node as CH is inversely proportional to its weight. So that the node that first proclaims itself as CH becomes the CH. Once a node hears the proclamation, it stays away from contention to become CH. If it doesn't receive any CH advertisement till its order, the node now announces itself as CH. If a node receives multiple advertisements from higher weight nodes, the node selects the nearest higher weight node as its CH. The non CH nodes now start replying to their nearest CH to join as cluster members. When compared to other clustering approaches, this approach has very low clustering overhead. The amount of data transmitted to construct clusters is reduced and the clustering process consumes much lesser energy.

Dual prediction based reporting

Dual prediction method involves a source node and cluster head having equivalent dimension of data history and a prediction engine to estimate future data. As indicated in fig.5, the process of prediction allowed to take place in both source node and cluster head simultaneously. Both the predictors output similar values. The source node then allowed comparing the projected output with the desired data from the sensor. If the predicted and desired data are almost equal, then there is no data transmission from the source node to cluster head. The cluster head adds up its predicted data on the data base. The data is delivered if the difference between the predicted data and the actual data exceeds a certain value to the cluster head.

In this dual prediction approach, the cluster head exploits a Normalized LMS (NLMS) filter to predict the sensor reading of the source nodes with certain accuracy instead of direct communication. As a result, there are fewer communications between nodes and the cluster head, and it is possible to skip the periodic radio broadcast, which is energy-intensive. The transmission of the subset of all samples is the major goal of this strategy. Each sensor node in this model has a prediction model that was trained using the data from previous sensor measurements. For predicting the data y, the previous data series x[k] is multiplied with its corresponding weight factor w[k], where k is the history length,

$$y[k] = w^t[k]\,x[k] \qquad \text{........................(10)}$$

The prediction is again verified with the actual data d, the deviation is calculated as e.

$e[k] = d[k] - y[k]$.....................(11)

Based on the error value, the weight values are updated for next iteration, the new weight factor is

$w[k+1] = w[k] + \mu x[k] e[k]$...(12)

On every new iteration, a new weight value is built from the previous weight values. The change in weight is made in multiples of step. Here μ is the step size. The value of μ can be between 0 to $1/E_x$. Where E_x is the mean input power. The major issue in LMS based filter is choosing the correct value of μ. Bigger the μ values tend to immediate convergence, but high amplitude oscillations around zero error state leads to unstable filter predictions. Smaller μ values result in stable filter prediction, but with slower convergence. The value μ should be optimized for faster convergence and higher stability.

$0 \leq \mu \leq 1/E_x$……………………..(4)

$E_x = \frac{1}{M}\sum_{k=1}^{M} x[k]^2$……………(5)

The correct choice of learning rate μ guarantees stability of the LMS adaptive filter. The proposed work uses a variant of LMS, Normalized LMS (NLMS), so that the sensitivity of filter with respect to input x(n) is reduced. This method doesn't need to initialize the initial step size, hence makes the method more generalized for various applications.

In LMS
$\underline{w}[k+1]=\underline{w}[k]+\mu\underline{x}[k]e[k]$…………(6)

In normalized LMS
$\underline{w}[k+1]=\underline{w}[k]+\mu\underline{x}[k]e[k]/\underline{x}^t[k]\underline{x}[k]$..(7)

Modified dual prediction framework

The proposed work deals with three main aspects of LMS based dual prediction. The proposed work concentrates on improving the rate of convergence during normal mode and reducing the deviation during standalone mode, which in turn increases the efficiency of dual prediction system. The work also improve the accuracy of prediction through which reduces the mean error of the overall system. The system develops a reliability enhancement mechanism by employing a counter and an inverted acknowledgement scheme.

Although the NLMS normalises the step size after each iteration, the rate of convergence is the same for all dual prediction system modes. By using a variable step size algorithm that manages the convergence process, the suggested method increases the effectiveness of the NLMS filter even further based on the current mode of operation. The step size, μ can be adjusted from 0 to $1/E_x$. The pace of convergence depends on the value of. Faster convergence results from larger step sizes, and vice versa. Faster deviations may also be caused by larger step sizes. The value of μ decides the speed of convergence. Faster convergence results from larger step sizes, and vice versa. Faster deviations may also be caused by larger step sizes. The step size is set as $(1/E_x)/D$. Here D is a non-zero integer value. For the duration of normal mode, the D tends to a minimum value, so μ has a larger value, which enables faster convergence. Once the prediction converges below the error threshold, during standalone mode the value of D is increased to higher values so that μ can be set at lower value, which results in fine-tuned convergence and lesser deviations.

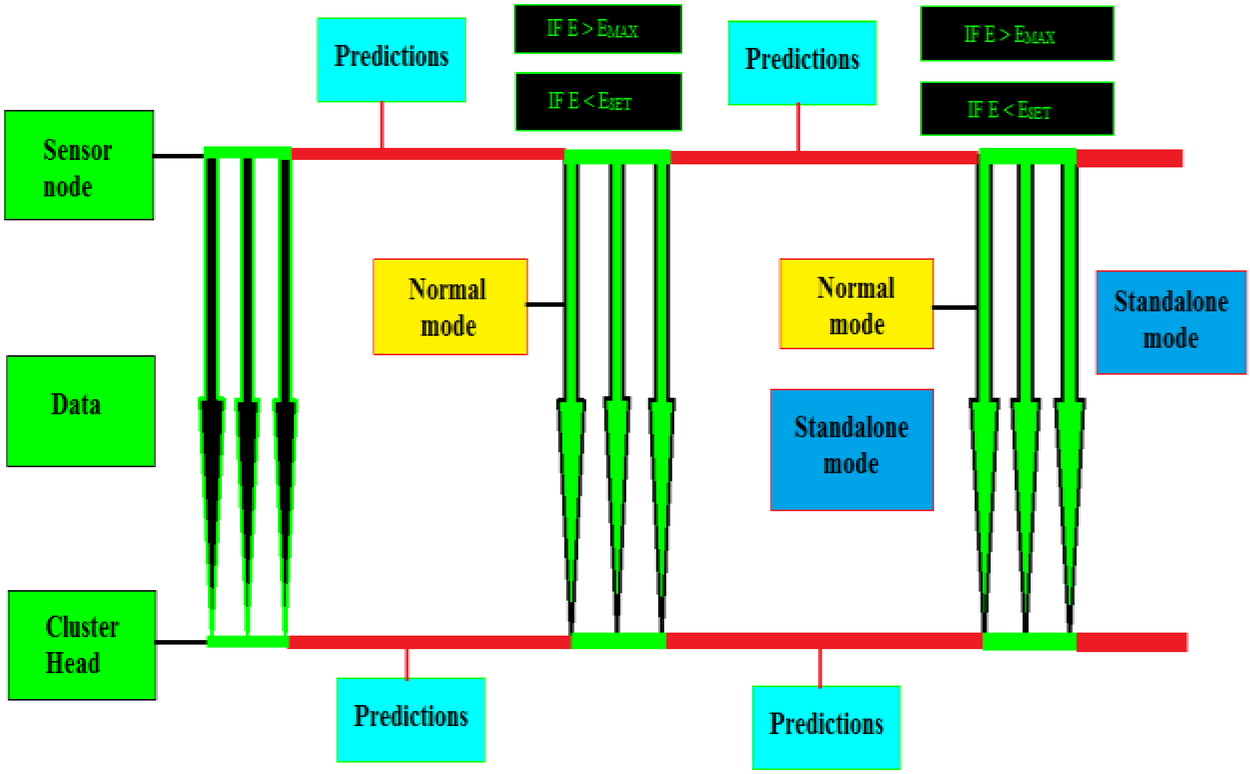

Fig.3. Dual Prediction based reporting

The dual prediction systems switch dynamically between the normal mode and standalone mode based on the error value. The accuracy and timing of switching decides the energy efficiency and prediction error of the dual prediction scheme. The delayed switching from normal to standalone mode causes excess data transmission in turn reduces the energy

Parameter	**Value**
E	5 nJ/bit
ε_{fs}	10 pJ/ bit/m2
ε_{mp}	0.0013 pJ/bit/m4
E_o	0.5 J
E_{DA}	5 nJ/bit/message
d_0	70 m
Message size	4000 bits
P_{opt}	0.1
Er_{max}	0.05°C
Er_{set}	0.005°C

efficiency. A delayed switching from standalone to normal mode results in larger deviations increases error and needs more iteration to get converged in to the threshold boundary. So an accurate switching mechanism is needed to govern the switching process.

Previous works on dual prediction make the switching from normal to standalone mode after certain number of consecutive predictions within the threshold boundary. So on every mode change several transmissions need be made. But with the proposed method, which is based on a new error level, the no. of transmissions after the mode change is reduced to fewer transmissions. Here a new error level Er_{set} is introduced. As shown in Fig.3 during normal mode, if the difference between predicted and desired value is less than Er_{set}, then the mode is switched to standalone mode. During standalone mode, the changeover is made after error value exceeds Er_{max}.

5. Results& Discussion

Since WSN has multidimensional characteristics, the proposed algorithm has been evaluated with multiple metrics. In WSN energy efficiency and data accuracy are the most important issues. By reducing the amount of data transmitted energy can be well conserved. But an accurate data collection needs more quantity of data to be transmitted. So the proposed algorithm balances between energy and data accuracy to achieve optimal performance. The LECA algorithm improves the life span with optimal data accuracy and coverage. Here the LECA clustering approach is verified on MATLAB platform.

Table 1.Simulation Parameters

The framework is simulated on both homogeneous and heterogeneous node environments. In both the environments LECA performs better than its equivalent algorithms. Table 1 indicates the simulation parameters.

Data accuracy

The very purpose of sensor network is to collect accurate data around the region. The data accuracy is not directly proportional to the data reduction. In the proposed work, the data accuracy is measured in terms MSE and MAE. With the help of dual prediction, the amount of data transmitted from the sensor node to cluster head can be reduced significantly. The data reduction comes at the cost of a predefined data error.

The suitability of data to fulfill its function in a particular environment is a measure of data quality. In WSN scenario data accuracy and the cost of data are important parameters. In a WSN the data is collected from spatially distributed points with discrete time span, thus No network collects 100% accurate data. With a small compromise in data accuracy, huge data cost can be reduced along with bandwidth reduction

For long-term data gathering in WSNs with limited bandwidth, approximate data collection is a wise option. The acquired sensor data typically exhibit intrinsic spatial-temporal correlations in many real-world application settings with densely placed sensor nodes. By investigating these correlations, sensor data can be compressedly acquired while staying below application-dependent error limitations.

In LECA, the dual prediction scheme is implemented between sensor node and the cluster head. The selection of LMS filter coefficients minimizes the deviation of predicted data from the actual data. The LMS filter with two level error threshold assures reduced data transmissions. Being an adaptive filter, the system adapts its parameters with respect to the change in data. The filter quality towards prediction is analyzed with the accuracy of data within the error threshold and the amount of data reduction achieved. The figure shows the similarity between desired and predicted data. The data are acquired from intel Berkley lab. The prediction accuracy is evaluated with different node data.

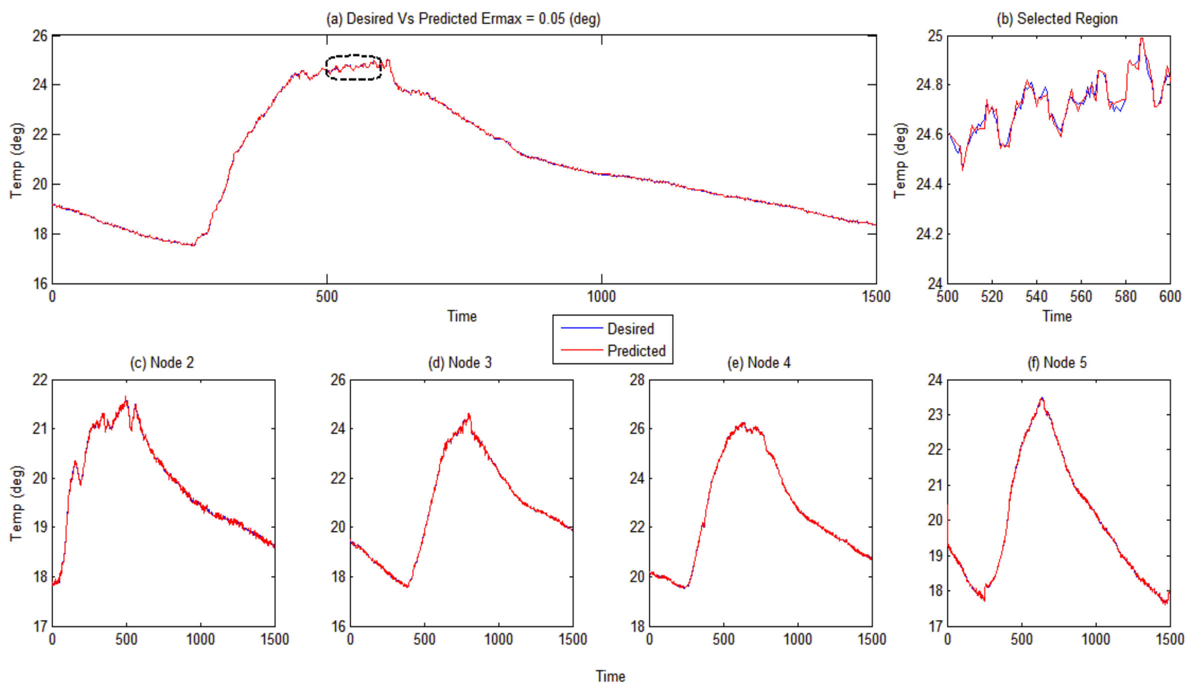

Fig.4. Comparison of Predicted and desired signals (Er_{max}:0.05)

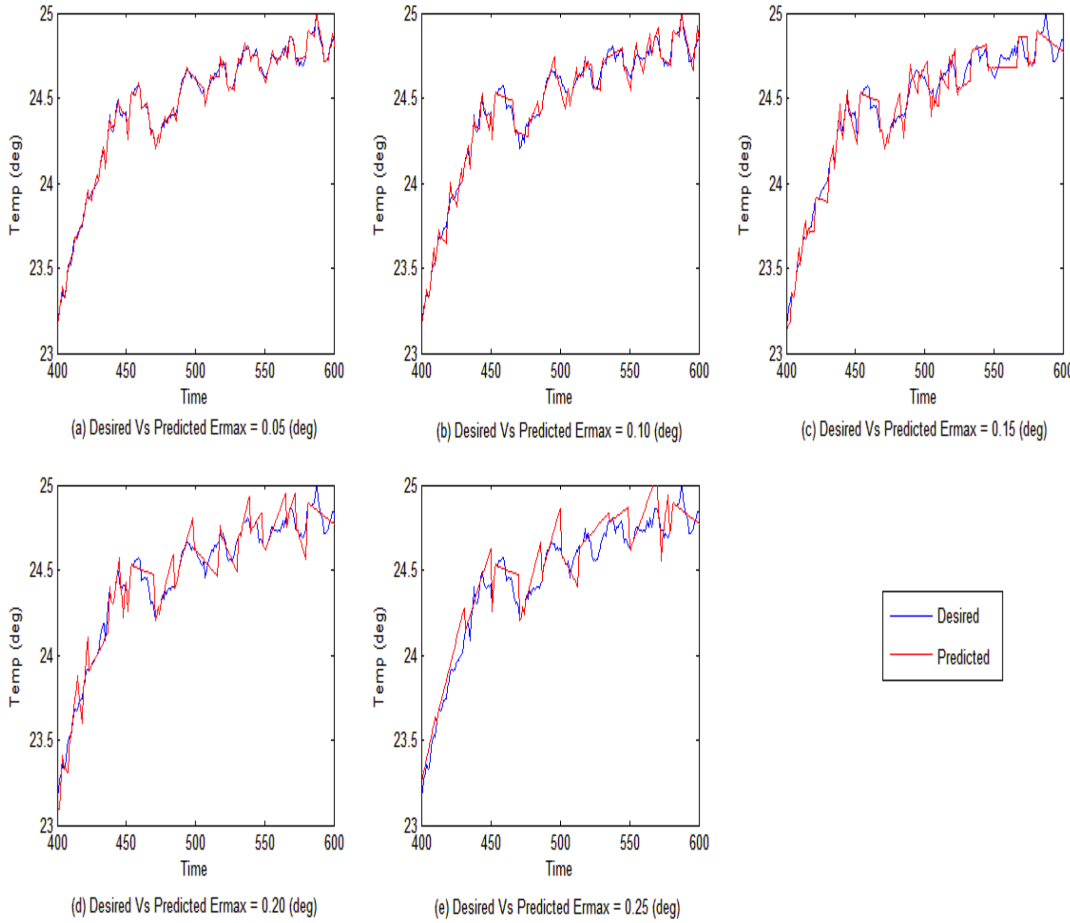

Fig 5.Comparison of predicted and desired signals with varyingEr_{max}

In this work the nodal temperature data form node 1,2,3,4 and 5 are taken .Figure 4a. shows the desired and predicted signal from node 1. Figure 4b.is the zoomed view of node1's data from time epoch 500 to 600. Figures 4c,4d,4e and 4f depicts the desired and predicted signals of node 2,3,4 and 5.

The DPR mechanism is evaluated with different error thresholds. The data from node 1is evaluated with different error threshold values. The comparison of desired and predicted signals for various error thresholds is depicted in Fig.5. Higher the error threshold higher the deviation of predicted signal for the desired one. In Fig 5, the dynamic region of data change is considered and performance is analyzed. The error threshold is varied from 0.05°C to 0.25°C. Being adaptive in nature the NLMS filter converges well with the desired signal even in the highly dynamic environment.

The energy efficiency of the proposed DPR frame work is analyzed. The filters efficiency is the measure of percentage of data reduction for a given error threshold.Fig.6, shows the percentage of data transmitted for different data sets with different error thresholds. Error threshold is varied from 0.05°C to 0.25°C. The amount of data transmission is also reduced from 36% to 5%. This data reduction has huge impact in nodal energy preservation and band width requirement. User can set the application specified error threshold get the maximum data cost reduction.

The data reduction depends on the data's dynamic nature and error threshold. In a dynamic context, the data reduction lowers in order to maintain the necessary level of projected signal accuracy. The proposed DPR framework conserves more data in node 2 and lesser data in node 3. The difference in data reduction between static and dynamic environments is significant in low error thresholds. With the increased error threshold, the difference is reduced.

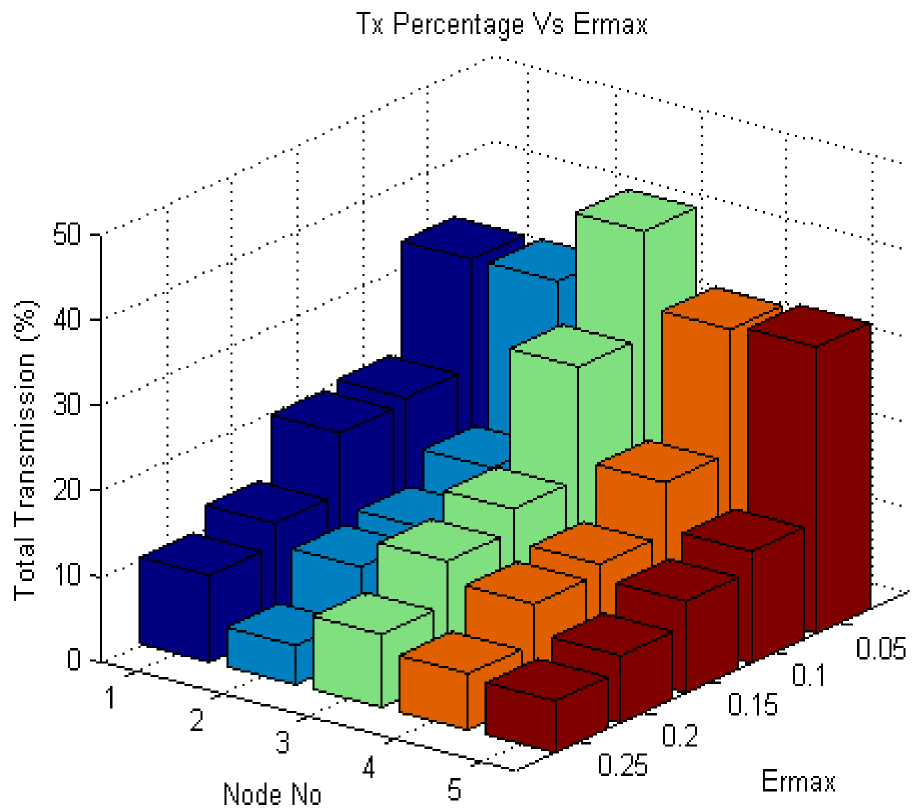

Fig.6.Transmission percentage for various Er_{max}

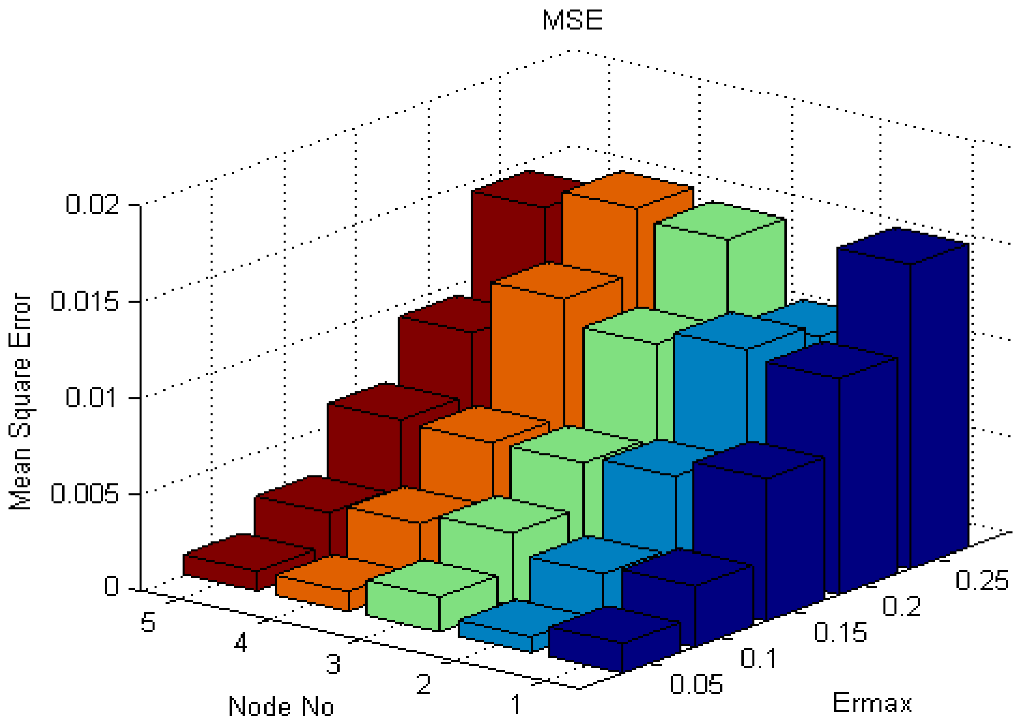

Fig.7. MSE for Various Er_{max}

Data accuracy measured in terms of Mean Squared Error (MSE), Symmetric Mean Absolute Percentage Error (sMAPE), and Mean Absolute Error (MAE) [16]. Mean Squared error can be given as, from eq (11).

$$MSE = \frac{1}{k}\sum_{k=1}^{L} e_k^2 \dots\dots\dots \quad (11)$$

Fig.7. shows the MSE of 5 different data sets with varying error threshold from 0.05°C to 0.25°C. Data accuracy of a prediction filter can be measured using multiple metrics. Here the Mean Absolute Error can be derived as from eq(11) is written as.

$$MAE = \frac{1}{k}\sum_{k=1}^{L} |\, e_k| \dots\dots\dots\dots\dots\dots \quad (12)$$

sMAPE for the prediction frame work is evolved using eq(13).

$$sMAPE = \frac{1}{k}\sum_{k=1}^{L} \frac{e_k}{(|d_k| + |y_k|)/2} \dots\dots (13)$$

Fig.8. and Fig 9 shows the MAE and MAPE of 5 different data sets with varying error threshold from 0.05°C to 0.25°

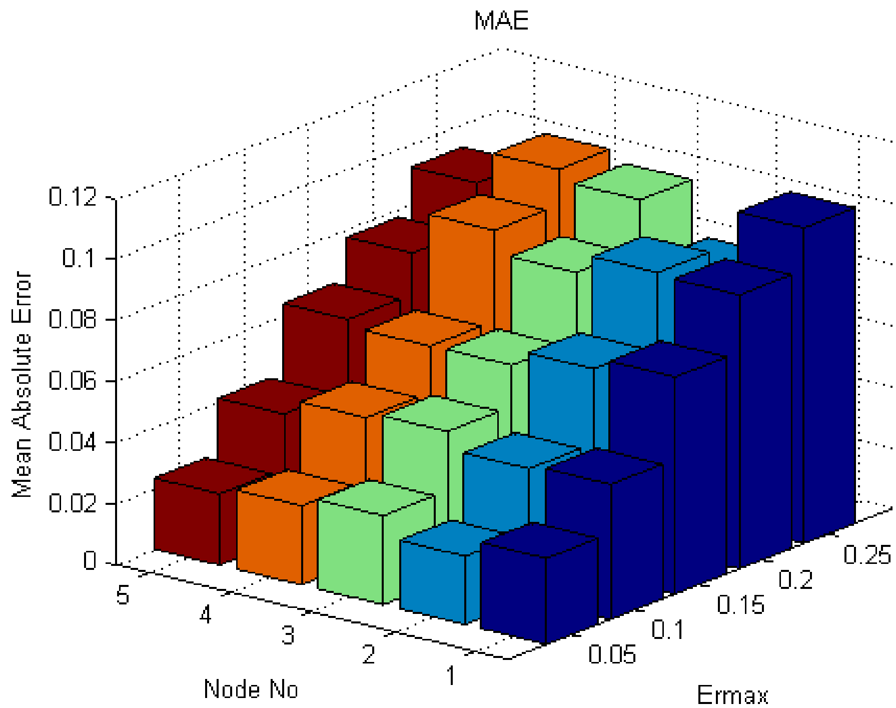

Fig 8. MAE for Various Er_{max}

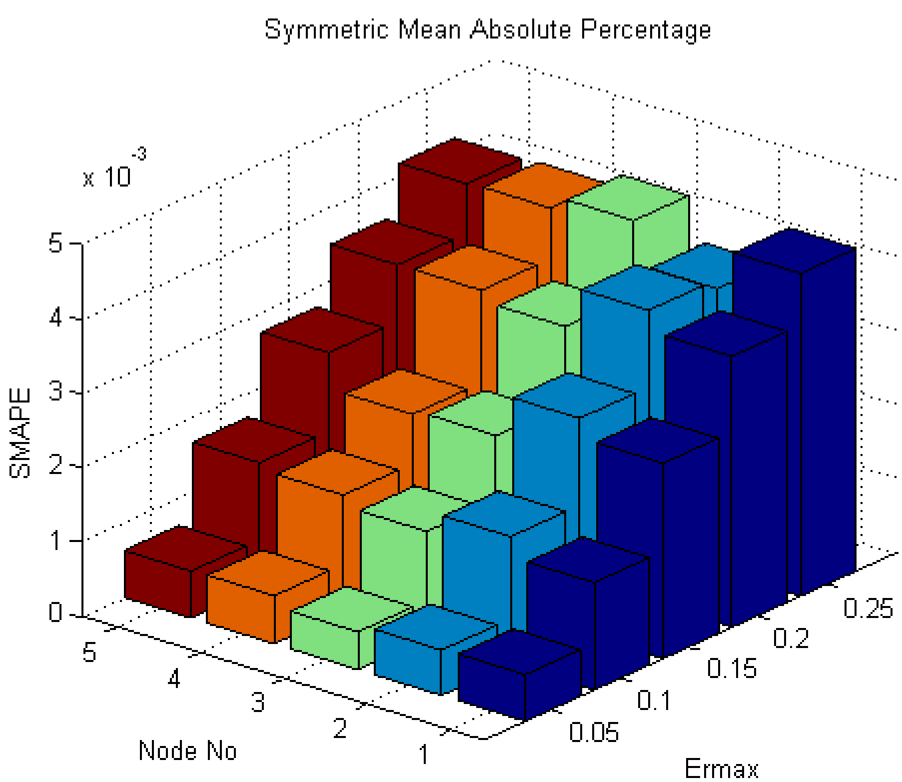

Fig.9.sMAPE for Various Er_{max}

Life span

Life time of a node is an important parameter in evaluating WSN performance. The nodal energy consumption is primarily attributed to data transmissions. LECA is compared with conventional data collection systems and other clustered data aggregation protocols. The results are encouraging, since the proposed system excels all other competing protocols in energy conservation

Homogeneous Network

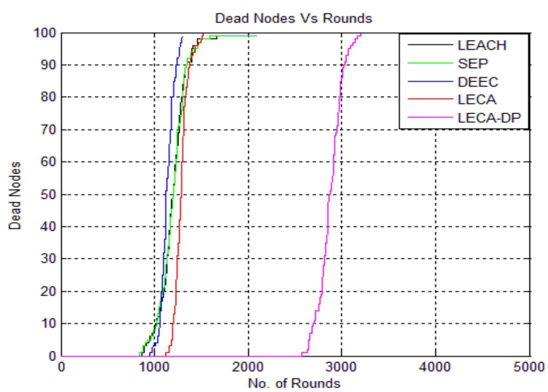

Fig.10.Life span of Nodes in Homogeneous Network

In a homogeneous network all nodes are having equal energy du initialization. In a homogeneous network weight coefficients for node power and n priority are set as 1. LECA improves the time of nodes better than DEEC, LEACH SEP protocols. LECA increases the stabi time of the network up to 20% rounds t other approaches. The close competitor DEEC, but lends itself to more node loss a 30% node dead. LECA algorithm maintains superiority all through the life time of the network as shown in figure. Thus maintains more alive nodes till last node dead. When combining DPR with LECA, the performance has been improved to multiple folds. The LECA framework with an error threshold of 0.05°C in DPR is also evaluated. The life span has been elongated to more than 2500 rounds, which is more than twice of any clustering approach. The network life time is shown in Fig 10.

Multilevel Heterogeneous Network:

In a multilevel heterogeneous network LECA enhances the stability period up to 1.5 times of the LEACH, SEP and DEEC protocols. LECA is highly efficient even in a heterogeneous network with an average of 500 rounds more life span with other protocols as shown in figure. The heterogeneous network is also evaluated with DPR based reporting between sensor node and cluster head. In LECA-DP the stability period is increased to 3 folds. Thus the compromise on 0.05°C in data accuracy resulted in massive energy conservation. The comparison is depicted in Fig.11.

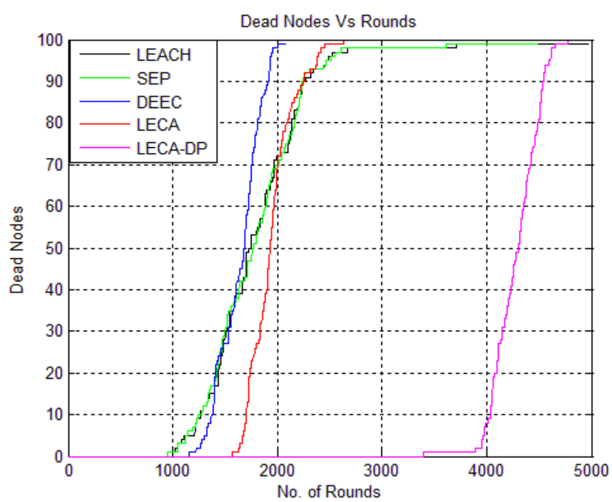

Fig.11. Life span of Nodes in Multi-level Heterogeneous Network

2-Level Heterogeneous network

LECA is tested on networks with different level of heterogeneity. The life time comparison clustering protocols on 2-level heterogeneous network with advanced nodes is shown in Fig.12 and Fig 13.A network with 10% of advanced nodes with 3 time's higher initial energy is defined. Both LECA and DEEC has higher stability time than other conventional LEACH and SEP based protocols. Still LECA slightly improves stability period better than DEEC. LECA also keeps more alive nodes than other protocols up to 90% node dead. After 90% node dead, LEACH maintains more alive nodes. This is a result of LEACH underutilization of advanced nodes.

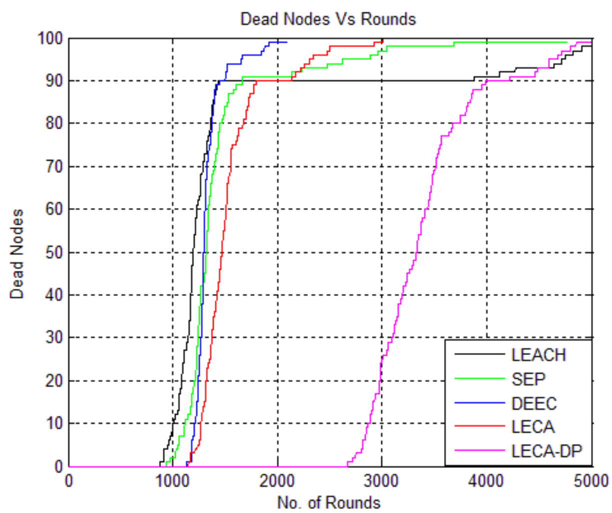

Fig.12.Life span of Nodes in 2-level Heterogeneous Network (m:0.1,a:2)

In a network with 20% advanced nodes with 3 times higher initial energy, LECA maintains higher stability than LEACH and SEP. DEEC has slightly longer stability period than LECA, which lags behind LECA after the death of 10% of its live nodes. After 80% node dead, LEACH maintains more alive nodes with the rounds, again the case of underutilization of advanced nodes.

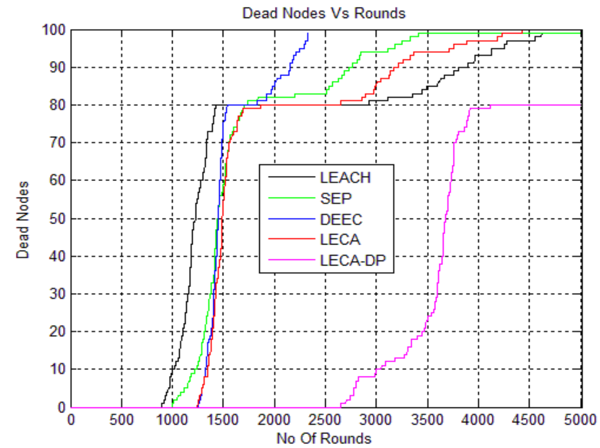

Fig.13.Life span of Nodes in 2-level Heterogeneous Network (m:0.2,a:2)

6. Conclusion

LECA clustering approach is introduced in this article. The LECA approach is evaluated on different types of wireless networks along with previously developed protocols. The LECA approach outperforms the previous approaches in energy efficiency. Then NLMS based dual prediction system is introduced. The combination of LECA and NLMS based DPR completely outperformed all the previous approaches by improving life time of network to multiple folds. The performance of LECA-DP in different types of network resulted in improved performance with meager compromise on data accuracy. LECA-DP due to its general nature can be applied to variety of WSN applications. Being lighter in computation, LECA-DP can be implemented on the low cost processors. The future work involves auto calibration of weight coefficients along with coverage preservation.

References:

1. Anastasi G, Conti M, Francesco MD and Passarella A, “Energy conservation in wireless sensor networks a survey,”AD HOC NETW vol.7, pp.537-568, 2007.
2. JunyoungHeo, Jiman Hong and Yookun Cho, “EARQ: Energy Aware Routing for Real-Time and Reliable Communication in Wireless Industrial Sensor Networks,” IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 5, NO. 1, FEBRUARY 2009.
3. D. Ganesan, A. Cerpa, W. Ye, Y. Yu, J. Zhao and D. Estrin, “Networking Issues in Wireless Sensor Networks,” Journal of Parallel and Distributed Computing, Vol. 64, pp. 799-814,2004.
4. OssamaYounis, Marwan Krunz and SrinivasanRamasubramanian, “Node clustering in wireless sensor networks: recent developments and deployment challenges,” IEEE Network, vol. 20, no.3, pp. 20-25,2006.
5. 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, pp. 1-8,2005..
6. Chuan-Ming Liu, Chuan-Hsiu Lee and Li-Chun Wang, “Distributed clustering algorithms for data-gathering in wireless mobile sensor networks,” J. Parallel Distrib. Comput, vol.67 , pp. 1187 – 1200,2007.
7. Heinzelman WR, Chandrakasan A and Balakrishnan H, “Energy efficient communication protocol for wireless micro-sensor networks,” In: International conference on system science, pp.1-10,4-7 Jan 2000,Maui, Hawaii.
8. P. Ding, J. Holliday and A. Celik,“ Distributed energy efficient hierarchical clustering for wireless sensor networks,” Proceedings of the IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS’05), Marina Del Rey, CA, June 2005.
9. M. Gerla and J.T.C. Tsai, “Multi-cluster, mobile, multimedia radio network,”WirelessNetworks,vol.1,no.3,pp.255–265,1995.
10. G. Gupta and M. Younis, “Load-balanced clustering in wireless sensornetworks,”Communications ICC'03 IEEE International Conference, pp.1848-1852, Anchorage, Alaska, May 2003.
11. Chong Liu, Kui Wu and Jian Pei, “An Energy-Efficient Data Collection Framework for Wireless Sensor Networks by Exploiting Spatiotemporal Correlation,”IEEE Transactions On Parallel And Distributed Systems, Vol. 18, No. 7, July,2007.
12. D. J. Baker and A. Ephremides, “The Architectural Organization of a MobileRadio Network via a Distributed Algorithm,” IEEE Trans. Commun., vol. 29,no. 11, 1981, pp. 1694–701)
13. Fuad Bajaber and Irfan Awan, “Adaptive decentralized re-clustering protocol for wireless sensor networks,” Journal of Computer and System Sciences ,pp.282–292,2011.
14. “HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks,”vol. 3, no. 4,pp. 366-379,October 2004.

15. Carlos.Carvalho,Danielo.G.Gomes,NazimAgoulmine and José Neuman de Souza, “Improving Prediction Accuracy for WSN Data Reduction by Applying Multivariate Spatio-Temporal Correlation,”Sensors,vol.11,no.11,pp. 10010–10037,2011.
16. Guorui Li and YingWang, “Automatic ARIMA modeling-based data aggregation scheme in wireless sensor networks,” EURASIP Journal on Wireless Communications and Networking,vol.85,2013.
17. Santini S and Römer K, “An adaptive strategy for quality based data reduction in wireless sensor networks,” In: 3rd international conference on networked sensing systems, pp.29-36,Chicago, USA, 31 May-2 Jun 2006.
18. N. Kushwaha, R. Mahule, A. P. Singh, O. P. Vyas and B. Singh, "Integration of service oriented WSN and IoT for E-Commerce," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015, pp. 1731-1736.
19. Zhang, Chi. "Research on the Construction of an e-Commerce Marketing System Based on the Wireless Sensor Network." Journal of Sensors 2021 (2021).
20. Reddy, Y. B. (2012). Trust-based approach in wireless sensor networks using an agent to each cluster. International Journal of Security, Privacy and Trust Management (IJSPTM), 1(1), 19-26.