Abstract—This paper covers a novel routing algorithm called Multi-Group based LEACH (MG-LEACH) that has been utilized the redundant deployed sensor nodes to improve the network life time. It has been suppressing the correlated data gathered by the sensor nodes by monitoring the similar event. Thus reduces not only the data transmission inside the clusters but also conserve the energy of deployed sensor nodes consequently improve the overall network lifetime. This is a simple idea that has been implemented over LEACH protocol however it is valid for almost all clustering based routing algorithms/protocols specially those variants based upon frame work of LEACH. The proposed routing algorithm has been simulated using MATLAB to verify the efficiency in enhancing network life time. A critical evaluation of routing algorithm is conducted to determine the relevance and applicability in increasing network life time. Simulation results confirmed that it has performed better than LEACH and enhanced network life time up to approximately 90%.

Index Terms—improvement research, LEACH algorithm, transformation strategy, wireless sensor networks

I. INTRODUCTION

A Wireless Sensor Network is composed of a large number of nodes disseminated over a region to accumulate information. The sensor nodes communicate among themselves through the wireless channel to self-organize into a single-hop or multi-hop network and transmit the gathered data towards one or more sink node also known as base stations. Each node has one or more sensors, embedded processors and low-power radios, and by and large it is a battery operated device. Characteristically, these nodes organize to execute a common specified task. Low power capacities of sensor nodes result in very limited coverage and communication range compared to other mobile devices [1-2].

For effectively cover the targeted area, sensor networks are composed of large number of sensor nodes. With the latest issues on climate change, WSNs can be utilized to track changes that influence the atmosphere using a network of sensors to collect environmental variables such as temperature, humidity and pressure. One of the several benefits of these sensors is their capability to operate unattended which is ultimate for unapproachable areas. Conversely, while WSNs are progressively more equipped to handle a number of these composite functions, in-network processing such as data aggregation, information fusion, computation and transmission activities requires these sensors to utilize their energy proficiently in order to extend their effective network life time.

It may also have application specific supplementary optional components such as a location finding system (e.g. GPS), a power generator and mobilizes. Sensing units are typically composed of two subunits: sensors and Analog to Digital Converters (ADCs).

The analog signals generated by the sensors based upon the monitored phenomenon are converted to digital signals by the ADC, and then transferred to the processing unit. The processing unit commonly associated with a small storage unit handled the procedures that make the sensor node collaborate with the other nodes to perform the assigned sensing tasks.

A transceiver unit joins the node in to the network. One of the most imperative components of a sensor node is the power unit. There may also have some additional subunits, which are highly application dependent. Each node has the capability to sense elements of its environment, execute simple computations, and communicate among its adjacent nodes or transfer monitored data directly to Base Station (BS).

Wireless sensor network has many constraints associated with it such as limited computational and battery power along with inadequate storage capacity, so management of resources cautiously is utmost important. A problem of dominant importance for a WSN is to maximize the lifetime of the network which bounds with the amount of energy stored in each node. In most of the applications sensor nodes are equipped with small, irreplaceable batteries with limited power capacity [3-4].

Sensor nodes are exceptionally prone to failures also it may not have global identification (ID) due to the large amount of overheads. Sensor nodes are densely deployed depending upon used applications. Hence, the primary goal of a WSN is to generate information from raw local data acquired (sensed data) by individual sensor node by extending the life time of WSN as much as possible. The resource constrained nature of sensor nodes poses the unique challenges to the design of WSNs for planned applications. The limited power of sensor nodes mandates the design of energy-efficient communication protocol. Most of the proposed communication protocols improve the energy efficiency to a certain extent by exploiting the collaborative nature of WSNs and its con-elation characteristics.

Multiple sensor nodes within the vicinity of the sensed phenomenon might generate similar data. In monitoring systems, various sensor nodes detect a common target of interest simultaneously and the data composed in this way are generally highly correlated and redundant. In contrast to other networks, this fact can be exploited at the network layer to improve energy efficiency and bandwidth requirements [5]. Data aggregation is also an imperative paradigm for condensing data hence energy of the net-
work is spent proficiently without principally losing the accuracy of the final aggregated data.

II. OVERVIEW

The objective of clustering algorithms differs but mainly depends upon the application requirements. The main objectives for the network clustering are [6-8]:

- **Load Balancing.** Optimal requirement of clustering is equal sized clusters that plays and important role for extending the network lifetime as it avoids the exhaustion of the energy of a subset of Cluster-Heads (CHs) at high rate and impulsively making them dysfunctional. Even the distribution of sensor nodes extremely influence data delay. When CHs execute data aggregation, it is imperative to have comparable number of nodes in the clusters so that the pooled data report becomes ready almost at the same time for further processing at the base-station or at the next tier in the network.

- **Fault Tolerance.** In a number of applications, sensor nodes may need to operate in harsh environments so nodes are typically exposed to augmented use of malfunction and physical damage. Enduring the failure of CHs is usually necessary in such applications to avoid the loss of imperative sensor data. One way to recuperate from a CH failure is to reconfigure or re-cluster the network. Aligning the rule of backup CHs is another scheme persuaded in the literature for recovery from a CH failure. Dividing the clustering operation in rounds and rotating the role of CHs among nodes in the cluster can also be a means for fault-tolerance in addition to their load balancing advantage.

- **Increased Connectivity and Reduced Delay.** Except if Cluster-Heads (CHs) have very long-haul communication capabilities (e.g. satellite link) inter-CH connectivity is an important condition in many applications. This is principally correct when CHs are picked from the deployed sensor nodes population. The objective of connectivity could be limited to guarantee the availability of a path from every CH to the base-station or it could be more restraining by imposing a bound on the length of the path.

- **Minimum Cluster Count.** Specified objective is principally rich nodes. The network designer often common when CHs are specialized resource likes to utilize the minimum number of such nodes as they tend to be more expensive and vulnerable than sensors.

- **Maximum Network Longevity.** As sensor nodes are characterize by their energy constrained nature, so the lifetime is a major concern especially for the applications of WSNs in harsh environments. In Heterogeneous Sensor Networks, when Cluster-heads (CHs) are more affluent in resources than sensors, it is imperative to minimize the energy for intra-cluster communication.

- **On the other side, when Cluster-heads (CHs) are standard sensors, their lifetime could be extended by limiting; their load and by rotating the role of cluster-heads among member node. Adaptive clustering is also a viable choice for achieving network endurance.**

- **Low Energy Adaptive Clustering Hierarchy (LEACH)** was pioneering research work in terms of hierarchical network topologies. LEACH is based on the deployment of homogeneous nodes and presumes that the deployed nodes along with sink node are stationary. It is an adaptive clustering-based protocol using randomized rotation of cluster-heads to evenly distribute the energy load among the sensor nodes in the network so it is an energy efficient routing algorithm used to reduce energy consumption in the network.

As mentioned earlier the clustering task is rotated among the nodes, based on duration. In LEACH, Direct communication is used by each CH to forward the data to the base station (BS). It is an application-specific data dissemination protocol that uses clusters to prolong the life of the wireless sensor network. LEACH used aggregation (or fusion) technique to combines or aggregates the monitored information into a smaller size of data that carry only imperative information by cutting down the redundant data of all individual deployed sensor nodes. In LEACH, network is divided in to several clusters of sensor nodes, which are raised by utilizing localized coordination and control mechanism.

It has been used not only to cut the amount of data that are transmitted to the sink, but also to make routing and data distribution more scalable and robust. Considering that energy dissipation of the sensor nodes based upon the distance and the data size to be transmitted, LEACH attempts to transmit data over short distances and reduce the number of transmission and reception operations. The three important features of LEACH are:

1) Localized co-ordination and control for cluster setup.
2) Randomized cluster head rotation.
3) Local compression to reduce global data communication.

The proposed clustering schemes search out to accomplish some characteristics for the producing clusters. The following are the relevant attributes:

1) **Cluster Count:** The set of cluster heads may be predetermined as a result the number of clusters is fixed in proposed algorithms. If the cluster heads are selected arbitrarily from the deployed sensor nodes usually produced variable number of clusters.

2) **Stability:** When the cluster count differs and the node's membership evolves overtime, the clustering scheme is said to be adaptive. On the other side, it is considered fixed if sensors do not switch among clusters and the number of clusters keeps steady right through the network lifespan.

3) **Intra-Cluster topology:** Some clustering schemes are based on direct communication between sensor and its elected cluster head. However, occasionally multi-hop connectivity is mandatory between sensor nodes and cluster head, especially when sensor's communication range is limited or the cluster head count is restricted.

4) **Inter-cluster head connectivity:** When the cluster head does not have long haul communication facilities, cluster head connectivity to the base station has to be provisioned. Another approach is to assume that cluster head would be able to directly reach the base station.

LEACH [9] is first proposed dynamic routing protocol, comprises of two phases named as setup phase and steady state phase. During set up phase, clusters are created followed by selection of cluster heads. Few cluster head are selected randomly which broadcast an advertisement message to the entire network declaring it a new cluster head. Every node after receiving advertisement message decides its cluster based on received signal strength of advertisement message. After selection of cluster head, sensor nodes send message to register with the cluster head of...
their choice. Cluster head send the time schedule to the registered nodes so that can send the data using TDMA approach. In the next phase known as steady state phase sensor nodes can send the observed data to the cluster heads on their time slot using the approach.

Cluster head send compressed aggregated data to the central station usually know as base station. After predetermined time period, new cluster head are elected again randomly. This repetition of electing new cluster heads is known as new round so LEACH is based on large number of rounds. Randomization process is used in LEACH to rotate cluster heads, which achieves a factor of 8 improvements as compare to the direct approach before the first node die.

III. METHOD AND ALGORITHM

LEACH employs the technique of randomize rotation of high-energy Cluster-heads (CH) position rather than selecting in static manner. It provides an opportunity to all sensor nodes to perform a role of CHs and avoid the battery exhaustion of an individual sensor node which results break down of these individual sensor nodes and drastically effects network live time.

The operation of LEACH is divided into rounds; while each round is consist of two phases known as:

1) Setup phase
2) Steady-state phase

Set-up phase is used to create Clusters followed by Advertisement and schedule creation. Steady-state phase is used for data aggregation, compression, and transmission to the (BS) base station or sink node.

The basic equation for the routing algorithm is shown below:

$$y_m(k) = wh_i + w_2h_2 + ... + w_nh_n$$ (7)

Assuming the ideal output is $y(k)$, the performance index function is:

$$E(k) = \frac{1}{2} \left( y(k) - y_m(k) \right)^2$$ (8)

The equation of basic function is as equation (9) as follows:

$$\partial_j(C_{ikl} \partial_{k} h_j + e_{ik} \partial_{k} f_j) - \rho \mu_j = 0$$ (9)

Under the linear relationship, basic equation is shown in equation (2):

$$\partial_j(e_{ikl} \partial_{k} h_j - \eta_{ik} \partial_{k} f_j) = 0$$ (10)

The linear differential equation can be expressed into the following simplified forms:

$$L(\nabla, \omega) f(x, \omega) = 0,$$

$$L(\nabla, \omega) = T(\nabla) + 2 \omega^2 \rho J$$ (11)

In which,

$$T(\nabla) = \begin{bmatrix} T_a(\nabla) & t_1(\nabla) \\ t_i(\nabla) & - \tau(\nabla) \end{bmatrix}, \quad J = \begin{bmatrix} \delta_{ik} & 0 \\ 0 & 0 \end{bmatrix},$$

$$f(x, \omega) = \begin{bmatrix} u_k(x, \omega) \\ q(x, \omega) \end{bmatrix}$$ (12)

$$T_{ik}(\nabla) = \partial_jC_{ikl} \partial_j, \quad t_i(\nabla) = \partial_je_{ikl} \partial_j,$$

$$\tau(\nabla) = \partial_j\eta_{ik} \partial_j$$

Consider an infinite situation, we have the equation (5) in the following:

$$L^0 = \begin{bmatrix} \gamma_{ikj}^0 & e_{ikj}^0 \\ e_{ikj}^0 & -\eta_{ikj}^0 \end{bmatrix}$$ (13)

Consider the propagation, instead the equation (13) with the following form:

$$C(x) = C^0 + C^1(x) \cdot \eta(x) = \eta^0 + \eta^1(x) \cdot \rho(x) = \rho_0 + \rho_1(x)$$ (14)

Then we have equation (15) to (18):

$$C^1 = C - C^0, \quad e^1 = e - e^0, \quad \eta^1 = \eta - \eta^0, \quad \rho_1 = \rho - \rho_0$$ (15)

The containing inclusions can be simplified into the following integral equation set:

$$f(x, \omega) = f^0(x, \omega) + \int_{-\infty}^{\infty} S(x-y')L^?F(y') dy'$$ (16)

$$+ \rho_\omega \int_{-\infty}^{\infty} g(R)T_f(y')S(y') dy'$$

In view of the following relationship

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iy_3} dy_3 = \delta(k_3)$$ (17)

Equation (8) can be converted into the following form:

$$f(y, \omega) = f^0(y, \omega) + \int_{-\infty}^{\infty} S(y-y', \omega) LF(y', \omega) dy'$$ (18)

$$+ \rho_\omega \int_{-\infty}^{\infty} g(y-y', \omega) Lf(y', \omega) dy'$$

http://www.i-joe.org
In which, $S$ is cylinder cross section, $y = (x_1, x_2)$, and

$$g(y - y', \omega) = \frac{1}{(2\pi)^2} \int_0^{2\pi} k \, dk \int_0^{\infty} \exp(-ik(y - y')) \, dk$$

Suppose $k = 0$, $g(\kappa, \omega)$ can be obtained from Equation (17).

For such kind of material, general form of equation (10) is expressed as following equation (20-22):

$$g_{ik}(\kappa, \omega) = \frac{1}{\rho \omega^2} \left\{ \frac{\beta^2}{k^2 - \beta^2} \theta_{ik} \right\}$$

$$+ k_{ik} \left( \frac{1}{k^2 - \alpha^2} + \frac{1}{\kappa^2 - \beta^2} \right) + m_{ik}$$

$$g_{ik}(\kappa, \omega) = \frac{1}{\eta \omega^2} \left\{ \frac{\beta^2}{k^2 - \beta^2} \right\}$$

$$\gamma_{ik}(\kappa, \omega) = \frac{1}{\rho \omega^2} \left\{ \frac{\beta^2}{k^2 - \beta^2} \right\}$$

Cluster-heads directly communicate with sink node in LEACH protocol without considering the distance factor. Single-hop communication is used to transmit aggregated data, no matter the distance between Cluster-head and sink is far or near. It will consume lot of energy if the distance is far, so multi hop routing is the standard solution also proposed in various variants based upon LEACH.

Number of Cluster is predefined factor in LEACH protocol, also cluster size varies in different proportion. While electing Cluster-heads node density is not considered which result of creating unequal size clusters with different number of member nodes. The effect of predefined and unequal cluster size is consumption of energy unevenly or disproportionately. In this way, some nodes drain out their energy earlier. Network does not remain fully connected so efficiently affect network life time. For synchronization, the clusters having less sensor nodes must wait for the clusters having more sensor nodes significantly affects network efficiency. Considering density function and calculating optimal number of clusters is the standard solution for the mentioned problem also proposed in number of variants based upon LEACH.

CHs are selected deterministically using threshold equation that based only on probability factor and does not include or consider residual energy of deployed sensor nodes. Residual energy of sensor nodes is an imperative factor that should also consider in threshold equation while electing Cluster-heads. If through this stochastic mechanism a node having less residual energy is selected as a Cluster-head, it may carry out of energy instantly or after operating for a short time. When this Cluster-head (CH) node dies all nodes belonging to the cluster will lose communication ability. It directly effects the network life time.

IV. EXPERIMENT RESULT

It is quite expensive to deploy hundreds of sensor nodes to a particular field to conduct experiments to evaluate proposed clustering algorithms. Simulations are the best form to formulate proposed algorithm and testify the results. In order to evaluate our proposed algorithm MG-LEACH, we have compared it with LEACH for different values of parameter. We have both algorithms on an identical wireless sensor network. We have used the same random seed to create the identical wireless sensor network topology.

First order radio model is used as energy dissipation model. This model simulates the energy consumption of each sensor node for transmitting and receiving k bits of data. In every round clusters-heads are elected and clusters are formed. Afterwards, each ordinary node forwards a certain bits of data to its cluster-head. Each cluster-head aggregates the received data and forwards it directly to the base station.

We simulate MG-LEACH along with LEACH algorithm in MATLAB to set up a comparative analysis both for LEACH and MG-LEACH. For the experiment, the random network of 300 Nodes is used. The Base-Station was placed in centre location with dimension (x=100, y=100). The bandwidth of the channel was set to 1 Mbps. Each data message was 4000 bytes long with header packet which is 25 bytes long. The radio electronics energy was set to 50 nJ/bit and the radio transmitter energy to 25 bytes long. The radio electronics energy was set to 50 nJ/bit and the radio transmitter energy was set to 50 nJ/bit for long distances less than 87 and 0.0013 nJ/bit for distances greater than 87m. The energy for performing computations to aggregate data was set to 5 nJ/bit/signal. In order to get improved and quite accurate comments of the algorithm, we establish the same simulation scene 30 times and the taken result is the average of the contain outputs.

For energy model, we assume that each node begins with equal energy and an unlimited amount of data to send to the base station. Once a node runs out of energy, considered as dead node and can no longer transmit or receive data. The valve of x used in set building phase is taken as 20. Initial energy for each node used in simulation is set to 0.5 Joule, while the experiment is repeated with 2 Joule both for LEACH and proposed MG-LEACH.

The main objective of MG-LEACH is to prolong the network life time by utilizing redundant nodes deployed in WSN. As the proposed algorithm is based upon the frame work of LEACH so the team round is used for each of the consecutive periods in which the sensor nodes perform: a predefined constant work. For example, in each round, every sensor node forwards 4000 bits of data to its cluster-head. We provide a summary result table which represents the values of FND(First Node Dies) HND (Half Node Dies) and LND (Last Node Dies) metrics for both of the algorithms simulated. After that, we provide a summary chart which illustrates the values of FND, HND and LND metrics visually. We also generate charts for the distribution of the number of a live sensor nodes and the distribution of the number of clusters per each round. By using these simulation results, we comment on the performance of the simulated algorithms.

Fig.1 illustrates the simulation result that demonstrate relative behavior of both discussed algorithms with parameters values n =300, p= 0.1 and $E_0=0.5$ J. It demon-
strategies Alive Nodes that is taken at y-axis for different time stamps (Rounds) that is taken on x-axis.

Fig. 2 and Fig. 3 illustrates the Pictographic representation of Comparative Analysis of LEACH and MG-LEACH for distribution of the alive sensor nodes with respect to the number of rounds for each algorithm when the define parameters set to n =300, p= 0.05, E₀=0.5 J and n=300, p= 0.05, E₀=2 J respectively.

V. DISCUSSION

Correlated data that has been sensed and collected in sensor network cancelled out before transmission to Base station however the monitored data is transmitted to Cluster heads from member nodes inside the cluster so consume a considerable amount of energy. If the correlated data has been suppressed from sensor nodes to Cluster-heads in clusters then the amount of energy consumed in aggregation process will also be decreased. The key idea behind Multi-group base LEACH (MG-LEACH) is to exploit this redundant information for enhancing network life time.

It is a blend both of centralized and distributed form of routing algorithm. Sub-Groups have been created by the Base-Station centrally using the location information provided by every sensor node at the time of deployment. Other two steps have been highly distributed in nature similar as used in LEACH. Duty cycle for execution of set-up and Steady state phases has been set centrally by Base station.

As revealed in figures above, MG-LEACH outperforms LEACH, considering FND and HND and LND metrics. In the scenario, when the set parameters values were p =0.1, E₀=0.5 J and n=300, MG-LEACH is 83% better than LEACH considering FND metrics while 86% better considering HND metrics. Total life span for proposed algorithm in the discussed scenario has improved 79% better been LEACH. In the scenario, when the set parameters values were p =0.1, E₀=2 J and n=300, MG-LEACH is 101% better than LEACH considering FND metrics while 92% better considering HND metrics. In the same scenario, total life span for proposed algorithm has improved 81% better been LEACH. Similar results has been ascertained when simulation work has been repeated for p=0.05.

MG-LEACH has significantly improved the LEACH performance by cutting the correlated information through utilizing redundant nodes present in the network. This is an excellent idea that might be implemented in clustering based routing protocol for enhancing network lifetime.

VI. CONCLUSION

In this paper we have proposed a novel clustering based energy efficient routing algorithm MG-LEACH for wireless sensor networks. The key objective of proposed algorithm is to prolong the lifetime of the wireless sensor network by utilizing the correlated nature of data inside the clusters. MG-LEACH is based upon the framework of LEACH protocol so we also illustrated in detail the shortcoming associated with it; in addition the performance analysis of LEACH has been performed in the thesis. It also covers comparative analysis of energy efficient MAC and Routing protocols used in Wireless Sensor Network.

The number of wireless sensor nodes deployed in the monitor network area is very large from hundreds to thousands but mainly depends upon the used application. As a result, generally data flow in the sensor network is substantial and this huge flow of sensed data will incur significant energy dissipation for sensor nodes. Additionally, the densely deployed nodes incur extremely correlated data. As the deployed sensor nodes are energy constrained so the routing algorithm utilized is required to be highly energy-efficient. On the top, as the energy expenditure is
diverse in the whole area and mostly varies from node to node due to assorted functions and positions in the network, the protocol should be able to balance the energy dissipation of nodes. The distances from nodes to the base station are typically long in a wireless sensor network. For longer distant data transmission will sustain substantial energy dissipation. As a result, the routing protocol should also be capable to minimize the energy consumption of data transmission from nodes to the base station.

REFERENCES

[1] H. Jing, “Node deployment algorithm based on perception model of wireless sensor network,” International Journal of Automation Technology, vol. 9, no. 3, pp. 210-215, April 2015. https://doi.org/10.20965/jat.2015.p0210

[2] H. Jing, “Routing optimization algorithm based on nodes density and energy consumption of wireless sensor network,” Journal of Computational Information Systems, vol. 11, no. 14, pp. 5047-5054, July 2015.

[3] Y. Hu, and X. Zhang, "Aggregation Tree Based Data Aggregation Algorithm in Wireless Sensor Networks," International Journal of Online Engineering, vol. 12, no. 06, pp. 10-15, June 2016. https://doi.org/10.3991/ijoe.v12i06.5408

[4] D. Jiang, Z. Xu, and Z. Lv, "A multicast delivery approach with minimum energy consumption for wireless multi-hop networks," Telecommunication Systems, vol. 62, no. 4, pp. 1-12, August 2015.

[5] Pisano, Andrea, F. Bignami, and R. Santoleri, "Oil Spill Detection in Glint-Contaminated Near-Infrared MODIS Imagery," Remote Sensing, vol. 7, no. 1, pp. 1112-1134, 2015. https://doi.org/10.3390/rs7010112

[6] Y. Li, et al., "Improved Compact Polariometric SAR Quad-Pol Reconstruction Algorithm for Oil Spill Detection," IEEE Geoscience & Remote Sensing Letters, vol. 7, no. 1, pp. 1139-1142, 2014. https://doi.org/10.1109/LGRS.2013.2288336

[7] Reddy, B. Brahma, and Rao, K. Kishan, "A Modified clustering for LEACH algorithm in WSN," International Journal of Advanced Computer Science & Applications, vol. 4, no. 5, 2013.

[8] W. Wang, and Y. Peng, "LEACH Algorithm Based on Load Balancing," Telkomnika Indonesian Journal of Electrical Engineering, vol. 11, no. 9, 2013.

[9] Hui, L. I., Z. Peng, and H. Dong, "An Improved LEACH Algorithm," Electronic Science & Technology, 2014.

AUTHORS

Dong Shujuan and Li Cong are with the Yellow River Conservancy Technical Institute, Henan, China (dongshujuan@chinauniedu.com).

Submitted 09 September 2016. Published as resubmitted by the authors 13 October 2016.