ABSTRACT

Data gathering plays a pivotal role in wireless sensor networks as it affects the lifetime of the sensor network, and has higher energy consumption than that of communication. The possible ways of initiating the data transfer from the sensors to the sink are event driven, query based, time driven and hybrid. For any data gathering scheme, sensors send their readings to the sink, which is classified into two namely, static sink and mobile sink based on the sink mobility. In the static sink approach, sink is placed permanently in the sensing environment, and the data are collected via multi-hop communication. Due to this, the nodes near the sink deplete its energy at the earliest, and result in energy hole problem. In order to address this issue, cluster based data gathering approaches are proposed. In the cluster based data gathering approach, the nodes that are close to each other are grouped together, and a node in the group is randomly selected as cluster head, which collects the information from the remaining nodes in the group called cluster members. This approach reduces the number of nodes that participate in the forwarding of data to the sink. It also reduced the energy hole problem, whereas the lifetime of the CH near the sink remained an issue, which was addressed by the ME data gathering approach.

In ME data gathering, ME visits every sensor node in its transmission range and collects the information from it. Though it reduce the energy hole problem. The data gathering latency is an open issue which is addressed by introducing rendezvous points.

In the rendezvous points, ME starts collecting the information from a particular data collection point instead of visiting every sensor node in its transmission range. However, it improves the speed of the mobile element data gathering by introducing the data collection points, the mobile element should wait until all its sensor nodes upload its data. Hence, data gathering latency remains an issue. Furthermore, sensor nodes transmit its data to the mobile element via multi-hop communication, which results in energy hole problem.

By considering the above mentioned points, the following data gathering approaches are proposed: Threshold based Toward Energy Efficient Big Data Gathering (T-TEEBDG) approach, Swift and Energy Efficient Big Data Gathering (SEEBDG) approach, Energy Efficient Mobile element Data Gathering (EEMDG) approach, Query based Dynamic Clustering (QDC) approach and Genetic Algorithm based Energy Efficient Data Gathering (GAEEDG) approach.

The first three proposed data gathering approaches concentrates on mobile element data gathering issues and their solutions. On the contrary, the last two approaches concentrate on the cluster based static sink data gathering issues and its solutions.

Firstly, the proposed T-TEEBDG approach is a cluster based energy conservation scheme. The main objective is to reduce the data gathering latency and improve the network lifetime by controlling the nodes that are communicating with the ME based on the threshold value. The sensing value of the node that reaches the threshold value is only allowed to communicate with the ME. However, it improves the speed of the ME data gathering by controlling the nodes that are communicating with the ME, data are transmitted to the ME via multi-hop communication because of its limited communication range and hence results in energy hole problem again.

Further to improve the performance of T-TEEBDG, the SEEBDG approaches are proposed. It is a cluster free energy conservation scheme for heterogeneous wireless sensor networks. In order to reduce the multi-hop communication, region partitioning is carried out based on the communication range of the resource rich node named local data collector (LDC). LDC transmits its request using free space model because of its high transceiver capacity. But data is collected from the sensor nodes in the local sensing region via multi-hop communication. Hence, energy hole problem still remains. In order to reduce the multi-hop communication, the nodes that are communicating with the ME is controlled based on the threshold value. Therefore, only minimal amount of nodes goes for multi-hop communication. For further improvement, an Unmanned Aerial vehicle (UAV) is used not only to collect the information but also to improve the speed of the data gathering. However, it reduces the data gathering latency by introducing LDC; ME should visit every LDC and collects the information from it. Hence, the data gathering latency still remains unsolved.

In order to improve this, EEMDG approach is proposed. In EEBDG approach, ME collects the information from only the special nodes named buffering nodes and visits its entire Buffering Node (BN). But, it collects the information only from the BNs that are in critical situation and thus increases the speed of the data gathering and network lifetime.

QDC approach is proposed to improve the performance of the cluster based static sink data gathering approach. In traditional clustering, all the sensor nodes in the sensing environment is considered while forming the cluster and it increases the cluster count and its density. Furthermore, cluster head transmits its collected information to the static sink via other cluster heads, which makes the CH near the sink deplete its energy quickly. The main objective of the QDC approach is to reduce the energy consumption by

controlling cluster count and its density, and increase the network lifetime by using relay element based routing. In which, the nodes satisfying the query are only allowed to participate in the clustering process. Others immediately go to the sleep mode. It reduces the count of clusters formed and increases the network lifetime. The cluster head that proximate to BS is taken as the first level cluster head and it uses a free space model to transmit the data to BS. The cluster heads, which are out of BS proximity is taken as the second level cluster head, which transmits data to BS with the help of relay node. The relay nodes are resource rich nodes that transmit the data to BS via other relay node. Hence, the load is balanced.

To further improve the performance of the QDC approach, a GAEEDG approach has been proposed. In which, the Genetic Algorithm (GA) is applied to minimize the count of the cluster formed and maximize the network lifetime in terms of rounds.

From the simulation results, it can be concluded that these proposed approaches outperforms the benchmarking system in various aspects viz network lifetime, data gathering latency and packet delivery ratio.