Optimal Lifetime Aware Data Aggregation for WSNs using Optimization Techniques

Youssef Emhemmad Mohammad Youssef, Klinsega and Wilson Jeberson

Department of Computer Science and I.T. Sam Higginbottom University of Agriculture, Technology and Sciences (SHUATS), Allahabad – 211007, Uttar Pradesh, India; 13PHCOE103@shiatsmail.edu.in, klinsega.jeberson@shiats.edu.in, jeberson_w@shiats.edu.in

Abstract:

Objectives: The goal of this work to identify the bloated state problem or increase of the agent size in the network. **Methods/Findings:** In this case, one of the mobile agent's issues is Bloated State Problem (BSP) which resolved by the SLMADA protocol, but by using this solution will have an effect on the aggregation time and the lifetime of the MAs. In order to resolve this issue, we designed in this paper a proposal called (OLDA-WSN protocol). Three different algorithms are the components of the proposed protocol. The grouping optimization algorithm to form the clustering, in this way, reducing the energy consumption of the nodes. The aggregated node evaluated between all nodes of the network for that we have used the node selection algorithm. **Application:** Also, we have used the modified task scheduling algorithm to calculate the waiting time and aggregated delay. This will be reducing the time-bloated state problem. We have used NS2 software to simulate the proposed protocol. Through analysis, it turned out to us that the proposed protocol has improved a network lifetime, the performance of throughput and success ratio.

Keyword: Cluster Formation, Data Aggregation, Group Optimal Algorithm, Modified Task Scheduling Algorithm, Node Selection Algorithm

1. Introduction

When we want a mesh design of a multi-hop (WSN), our adoption upon to the requirements of sensing application. The main goal of sensor networks is to monitoring physical phenomena in a particular area, and then the collected data sent to the sink¹. Wireless Sensor Network (WSN) consists of many numbers of energies, computational constraint nodes, and communication. In order, to overcome energy constraints recharging or replacing the batteries of the nodes in (WSN) is almost impossible, once the nodes deployed in hostile regions². We must balance communication between the nodes in (WSN) to keep the (WSN) alive for a long time. Applications of time importance, such as forest fire detection, battlefield monitoring and data receipt via the sink, on requiring specific delays to avoid disasters^{3.4}. Thus, a protocol design needed that enhances the life of the network and provides the information to the sink with limited delay.

The direct transmission provides a slight delay, but energy consumption in the WSN nodes increased. On the other side, multi-hop communication is energy-efficient as nodes have to send over a shorter distance, and the energy consumption is directly proportional to the distance^{5.6} but it increases the delay. Also, we should select direct transmission or multi-hop transmission between CH and member nodes, and between CH and other CHs or BS to balance between facing by the data and the energy consumption of a node. To reducing WSNs energy consumption, the (In-network) data fusion and clustering have proven as effective. The bottlenecks to networks, it can happen by clustering, which is causing an extra delay in the process of data aggregation.

Neighboring sensor nodes in WSNs generated data, which is often highly correlated and redundant. The relaying or sending of many redundant data, the sensor nodes will be consuming large energy for that. Furthermore, a large number of data transfer cause congestion and data collisions. These are the reasons that led to the turning up of data aggregation technology. We can define the process of data aggregation in the network as a process of aggregation data from nodes, which deployed in the network and eliminating excess transmission and providing fused information to the sink node. There are three main types of data aggregation ways, clustering data aggregation is the first one, in this type, data are collected and aggregated at Cluster Node (CN), after that, this data transmitted to a sink node^{7.8} hop-by-hop aggregation is the second one, in this type, data aggregated will be at each intermediate node⁹, partial aggregation is the third one, in this type, data aggregation satisfies energy threshold or a time¹⁰. Nevertheless, each type of data aggregation above has drawbacks. In the first type, the WSNs, always used an algorithm to find the CH of each cluster in WSN, which in turn causes waste energy and time⁸. In the second type (hop-by-hop aggregation), this type suffers from energy consumption is unbalanced as well as long delay in transmission, also, suffers from nodes on the transmission direction¹¹ Aggregating function. Here, always, the period of time is decided by a given request data accuracy or time threshold, but the energy consumption of the nodes in the network always not taken into consideration¹²⁻¹⁴. As we all know, the lifetime of a network is always based on the lifetime of nodes in a network. A network has to must perform reorganization when a node's energy runs out. And therefore, as we know that the transmission delay can shorten in the partial aggregation, but it still suffers from the short lifetime of the network because of unbalancing in the lifetime of the nodes. To find the fault nodes the Clone-based Dynamic and Distributed Agent Migration (CDDAM)¹⁵ is used. For multiple agents in¹⁶ an itinerary based an exploratory approach is introduced, and it uses along with that, a method called Iterated Local Search (ILS).

The other sections of the presented paper are divided as follows: Section two presents the related work, and the data aggregation major problem is discussed in the third section, the detailed study of the proposed protocols given in the fourth section, in the fifth section the simulation results are given, in section six the conclusion of the paper is given.

2. Literature Survey

In¹⁷ the proposed protocol called Hamilton Energy-Efficient Routing (HEER) it's an improved energy-efficient clustered and delay-aware protocol. During the network initialization phase, the clusters will be forms by HEER, the Hamilton Path was created by a greedy algorithm, and every node in each cluster will link by this path for the purpose of data transmission. In HEER proposal, it is not required to the reformation, the cluster; each node on the HEER path will take its turn to become CH. By comparing the traditional cluster-based protocols with HEER, we will notice that the design allows HEER to balance the load and save on network administration energy. If we compared HEER with different types of classic chain-based protocols such as PEGASIS, we would find that in HEER it is not necessary that every node must have the information of the global location, and that it does not suffer from long delay as well. For HEER protocol simulators have been used MATLAB, and the simulation results compared with some chain-based and cluster-based protocols, network lifetime was improved by HEER over existing protocols with maintaining the data transfer delay rate. From the simulation, HEER proposed, respectively managed to achieved 21.2 times and 16.7 times more rounds than (Intra-grid-PEGSIS and PEGASIS) which are chain-based protocols, and compared with cluster-based protocols (LEACH-EE and LEACH), HEER was achieved 66.5% and 40.6% more rounds than both of them. Plus all of this, compared to cluster-based protocols, the HEER can get rid of all transmission delay, and 99% compared to chain-based protocols into the network.

In¹⁸ presented a delay-aware network structure for WSNs with (In-network) data fusion technique. The proposed structure organizes sensor nodes into clusters of different sizes such that each cluster can communicate with the fusion center in an interleaved manner. An optimization process was proposed to optimize intracluster communication distance. When compared simulation results with some existing aggregation structures, shown that the proposed network structure can reduce delay in data aggregation processes. In addition, keep the total energy consumption at low levels, provided that data are only partially fusible.

In¹⁹ introducing the Energy-Delay Index for Trade-off (EDIT) to improve the objectives of both delay and energy, by designing a routing algorithm. For CHs selection and

"next hop", the EDIT was used by considering energy and or delay requirements a given application. The proposed system was derived by using two different distances between the node and a sink called Euclidean Distance (ED) and Hop-Counting (HC), and it is also proven using realistic parameters of the radio to obtain the data closest to the testbed implementation. The results aspire to given out enough insights to others before carrying out the testbed. The authors state that their work was the first attempt to get out an energy-delay trade-off by using two different types of distances for delay restricted applications.

In²⁰ a real-time routing protocol was proposed for Mobile Wireless Sensor Nodes (MWSNs) in industrial environments, which provides Quality Of Service (QoS) in terms of delay time. At the start, to get out location information, the information, which in the nodes routing table and the geographic routing will be used for this purpose. The node next to the target node was then calculated to reduce hop counts and quickly achieve packet data transfers fast. It optimizes packets redirecting hops, which in turn reduce delay time effectively. The results of the simulation results of the algorithm compared to the existing ones indicate that the algorithm was useful and suitable for WSNs.

In¹⁰ provided an ideal network lifetime by balancing energy consumption between nodes in the network. Here, they proposed a partial compilation of the waterfalls, controlled by a group of waterfalls that pushing rate vectors. There are two contributions in this work; the first one was to propose partial aggregation of waterfalls and to model it with queuing theory. The second one was that the optimum network lifetime was mathematically achieved and suggested an algorithm close to the optimal level of a given transmission delay. The proposed algorithm results were compared with energy-efficient algorithms which are exist, and from getting out the evaluation results find out that the proposed algorithm was efficiency.

In²¹ proposed a new model of mobile data collection that reduces the data latency significantly. Using a combination of a touring strategy based on clustering and a data collection mechanism based on wireless communication, for mobile data collection proposed a new model greatly reduces data access time. By blending both data collection mechanism based on wireless communication as well as touring strategy based on clustering, they showed that delays could be significantly reduced without compromising the advantages of the Mobile Data Collector (MDC) approach. Using largescale simulation studies, analyzed the performance of the proposed model when compared with other existing models, it was confirmed that the proposed model reduced the package delay to more than half.

In²² the delay performance in WSN was studied with the topology of the cluster-tree. The end-to-end delay in such a network can depend heavily on the relative location between each of resource allocations of the Cluster Heads (CHs), the sink and the sensors. For real-time traffic, packets sent with excessive delay were dropped. Given the schedule assignments for each CH for a local and inter-cluster traffic transmission, an analytical model was also developed to find the end-to-end transmission delay distribution of packets created from different clusters. From the results, derived the packets drop rate, Next, a guideline was proposed to find schedule assignments for all (CHs in WSN). To achieve the balanced and minimum rate of packet drop of traffic originating from different levels of the cluster-tree. Simulation results were displayed to verify the analysis and demonstrate the effectiveness of the proposed schedule allocation scheme for (CH).

In²³ developed the (EDAL) protocol that provides energy delay. Precisely, EDAL was formulated by addressing the energy cost of the packet transmitting in WSNs in a manner similar to the cost of delivering goods in OVR and addressing packet differences similar to delivery times. After that problem that was resolved by EDAL was NP-hard they introduced both the distributed heuristic based on ant colony gossiping and the centralized Meta heuristic based on Tabu Search (TS), to get approximate solutions. Design of their algorithm also took into account the load balancing of individual nodes to maximize the lifetime of the system. When used the algorithm with compressive sensing, its will leads to reduce the traffic generated amount in the network. Both of a small-scale and a large-scale were evaluated using NS3, based on the hardware testbed.

3. Problem Methodology

3.1 Problem Definition

In Data Aggregation (DA) technique, the data always wait for a period of time to be sent to collect more data at a node; all collected data are aggregated into several representative data and then transmitted to an upper neighboring node. In data aggregation, BSP caused an increase in energy consumption of the node. In order to resolve the issue²⁴ have designed a scalable and load-balanced scheme for mobile agents-based data aggregation (SLMADA) protocol. This proposal has been tested through different network scenarios and based on the simulation, this proposal has been evaluated. On SLMADA protocol the agent itinerary will be dynamically, and an agent will decide at each hop, which path it should take. The SLMADA protocol determines a maximum data payload limit for the agent and uses the agent cloning technique. According to its payload size, the agent makes a clone of itself. SLMADA protocol according to the current state of a network will help the MA to change his path dynamically. To provide a scalable solution and a loadbalanced, an agent cloning technique is used. Through mathematical analysis, we notice the emergence of (BSP) as well as its impact on the performance of the network. Data collection is a scheme that saves energy by reducing the amount of data transfer. Non-aggregation consumes a lot of energy, full aggregation causes a long delay in transmission, and partial aggregation can balance energy other WSN offerings.

- BSP of the mobile agent was resolved by the SLMADA protocol, but the aggregate time and the lifetime of the MAs can affect, by this solution. As we know, at nodes, the waiting time is always decided by corresponding applications and is adjustable as well. One can adjust the wait time very short when the data are wanted, with short transmission delay and vice versa. Sensor data can be consolidated from multiple nodes before being redirected to neighboring nodes, and thus, energy consumption can be greatly reduced, that happens, when using (In-network) data aggregation approach.
- Furthermore, most of the energy consumed during receiving and transmitting data is because of the power required to turn the circuitry of a sensor node Off and On. Therefore, a reduction in data transmission can result in a significant reduction in energy consumption.

To overcome those issues, an OLDA-WSN is proposed using optimization techniques. The main objective of the proposed (ORDA-WSN) is to overcome the time-bloated state problem to achieve efficient data aggregation.

The main contributions of proposed (OLDA-WSN) as follow:

- 1. In (OLDA-WSN), optimal algorithm set is used to form clusters; this reduces the energy consumption of the nodes.
- 2. We provided a node selection algorithm to computation the Aggregated Nodes (ANs) between several nodes in a network.
- 3. The modified task scheduling algorithm used to compute the waiting time and aggregation delay for the re-forwarding data, thus, reduces the time-bloated state problem.
- 4. We also design the analytical model of the proposed protocol with its assessment. The results were compared with existing approaches and the evaluation results shown the efficiency of our proposed protocol.

3.2 OLDA-WSN Model

Figure 1 illustrates a system model of the (OLDA-WSN) protocol. The figure has shown a simple network, where divided the network into clusters, where every cluster contains the aggregated nodes as well as nodes. From the source, data will be sent to the destination and nodes.

As we know the sensor nodes are homogeneous, so the nodes were deployed in NS2 into the flat are allocated omni-directional antennas with the same transmission range, and in SLMADA protocol the field of observation was circular, spanning in a two-dimensional plane.

The sensor nodes do not change their location once done the deployment and nodes are stationary, assigning location coordinates to them is a onetime task during the same.

In a multi-hop manner, the data generated at the nodes will be transmitted to a sink node.

All neighboring nodes are in the same transmission range, and the distance between two neighboring nodes is the same.



Figure 1. System model.

The node density is varied by 50, 100, 150, 200, 250 and 300, with fixed network area as $100m \times 100m$.

4. Proposed Methodology

In this section, the optimal lifetime aware data aggregation is proposed for WSN (OLDA-WSN) protocol is presented for solving the time-bloated state problem to achieve efficient data aggregation. The operation of the proposed methods, OLDA-WSN, consists of three phases: 1. cluster formation, 2. identification of the aggregated node, and 3. computing waiting time of aggregation delay. This section describes each of these phases in detail.

4.1 Group Optimal Algorithm used to form the Clustering

The grouping algorithm is used to cluster formation. It is typically used to reduce energy consumption at nodes.

4.1.1 Cluster Formation

During node selection, network lifetime is the main challenge. For effective data transmission, the optimal cluster is determined. This will reduce the loss of data. The clustering is thus formed by the group optimal algorithm. In the cluster, the Aggregate Node (AN) is selected to transmit the aggregated data to a sink node for data analysis. We assumed that the WSN environment contained n sensors, which are given as Ni.

Energy: The consists of WSN have distributed sensors²⁵. All sensor nodes are powered by a battery, for this reason, maintaining the sensor node energy is a major challenge, in case the energy consumption is high, and this would short the network lifetime.

$$E_{dis}(N_i) = (E_{elec} \times P_i + E_{pa} \times b_i \times ||N_i - AN||^4);$$

if $||N_i - AN|| \ge d_0$ (1)

$$E_{dis}(N_i) = (E_{elec} \times P_i + E_{fs} \times b_i \times ||N_i - AN||^2);$$

if $||N_i - AN|| < d_0$ (2)

$$L_{d_0} = \sqrt{\frac{E_{fs}}{E_{pa}}} \tag{3}$$

Where E_{elec} indicates electronic energy, $E_{dis}(N_i)$ indicate energy dissipation, bi indicates the number of bytes sent

by the node, the power amplifier energy is indicated by Epa and Efs is the free space energy. Normally the electrical energy depends on the factors like modulation, filtering and, coding.

$$E_{elec} = E_{tx} + E_{agg} \tag{4}$$

Where the transmission energy is indicated by E_{tx} and the E_{agg} indicate the energy of data aggregation. $||N_i - AN_j||$ is the distance between the ith node and the jth Aggregator Node (AN). When a normal node communicates with AN at the jth position, there is a loss in the energy at the AN. This energy loss is due to the electrical energy at the receiver side and the data bytes received by the AN. The energy dissipation at the jth AN is:

$$E_{dis}(AN_i) = E_{elec} \times b_i \tag{5}$$

When the data are transmitting and receiving, the details of ANs and the sensor nodes are updated. Based on the energy dissipation of nodes and energy of nodes the updating will be. As follows the formula will be used for updating:-

$$E_{t+1}(N_i) = E_t(N_i) + E_{dis}(N_i)$$
(6)

$$E_{t+1}(AN_i) = E_t(AN_i) + E_{dis}(AN_i)$$
(7)

The energy continues to update in the nodes until the energy of the nodes becomes (0) or the node becomes dead.

Mobility: This parameter describes the position of the nodes as well as their movement and acceleration in the network and illustrates the velocity of the node in the network region²⁶. Thus, for analyzing the performance of the routing all this will be used. Consider two nodes n_i and n_k and the position is given by (p_i, l_i) and (p_k, l_k) such that $s^i \,\epsilon \, (p_i, l_i)$ and $s^k \,\epsilon \, (p_k, l_k)$. The variable velocity of n_i and n_k is given by ϕ_1 and ϕ_2 . Initially, the nodes are placed at the h_1 and h_2 . After a certain time, the node will move to a new position, and it is given by (p_i^{new}, l_i^{new}) and (p_k^{new}, l_k^{new}) . The Euclidean distance is given by:

$$h_{(pl,0)} = \sqrt{|p_i - p_k|^2 + |l_i - l_k|^2}$$
(8)

The nodes traverse with a velocity of δ_{ni} and δ_{nk} and the distance is given by:

$$h_1 = \delta_{n_i} \times t \tag{9}$$

$$h_2 = \delta_{n_k} \times t \tag{10}$$

The new location is given by:

$$p_i^{new} = p_i^{old} \,\delta_{n_i} \times t \cos\phi \tag{11}$$

$$l_i^{new} = l_i^{old} \,\delta_{n_i} \times t \cos\phi \tag{12}$$

Then the new distance of the node at k (n_k) is given by:

$$p_k^{new} = p_k^{old} \,\delta_{n_i} \times t \cos\phi \tag{13}$$

$$l_k^{new} = l_k^{old} \,\delta_{n_i} \times t \cos \phi \tag{14}$$

The new distance is given by:

$$h_{(p^{new}l^{new},t)} = \sqrt{|p_i^{new} - p_k^{new}|^2 + |l_i^{new} - l_k^{new}|^2}$$
(15)

4.1.2 Group Optimal Algorithm

Grouping algorithm is used for gathering the nodes that are having the same characters. The clusters are formed for the proper packet transformation. Initially, the first node is taken for group formation. Each parameter above is calculated for each node and compared to get the character of each node. After those nodes having the same character, all will be grouped in the same group:

$$G_i(t+1) = G_i(t) + \eta_{1i}(D_i - x_i(t) + \eta_{2i}(Q - x_i(t)))$$
(16)

$$G_i(t+1) = G_i(t) + \eta_{1i} D_i - \eta_{1i} x_i(t) + \eta_{2i} Q - \eta_{2i} x_i(t)) \quad (17)$$

$$G_i(t+1) = G_i(t) + \eta_{1i} D_i + \eta_{2i} Q - (\eta_{2i} + \eta_{2i}) x_i(t))$$
(18)

Where $G_i(t + 1)$ and $G_i(t)$ are the search space dimension of ith population in two iterations (t + 1)th and tth, η_{1i} and η_{2i} are the random numbers of ith population in the random interval [0,1], the best positions are given by D_i and Q and the global best solution of the population. Calculated the position of the population as follow:

$$x_i(t+1) = x_i(t) + G_i(t+1)$$
(19)

$$x_{i}(t+1) = x_{i}(t) + G_{i}(t) + \eta_{1i}D_{i} + \eta_{2i}Q - (\eta_{2i} - \eta_{1i})x_{i}(t)$$
(20)

Algorithm 1. Grouping optimal algorithm for clustering formation		
Input	$n_{i}, E_{(t+1)}(N_{i}), E_{(t+1)}(AN_{j})$	
Output	Cluster formation	
1	Initialization of the dolphins (Aggregator Nodes) $n_{_{i}}$	
2	Calculate the initial position of the dolphin $G_i(t)$ and $G_i(t + 1)$	
3	Calculate $G_i(t + 1)$ and $x_i(t + 1)$	
4	Predefined probability is calculated $\rho(j) = \rho + (1 - \rho 1) \times \frac{j_i^{pp} - 1}{(j_z)^{pp} - 1}$	
5	Convergence varies with the predefined probability	
6	Calculate $E_{(t+1)}(N_i)$, $E_{(t+1)}(AN_j)$	
7	Calculate the Euclidean distance of the node n_i and AN_j .	
8	Calculate search position $G_i(t + 1)$ and $G_i(t)$	
9	Update the best and the global best positions (optimal cluster formation)	
10	Update the new position and distances.	
11	end	
Return Cluster Formation.		

4.2 Node Selection Algorithm for the AN

4.2.1 Data Aggregation

The increased lifetime for (WSN) by removing redundant data transmission is the major aim of data aggregation protocols. It is considered that in WSN the data transmission occurs in a multi-hop manner. In multi-hop approach, every node transmits data to the node which is near to it. The near node is considered as the sink. The nodes which are placed in a closely packed area may sense the data of the same category. This approach cannot be considered as an energy efficient approach. Therefore, a clustering-based approach is an improved method of data transmission. In this approach, data will be sent by nodes to Cluster Head (CH). After that, the (CH) performed the aggregation. In the row data that are available at the receiver end, the aggregation is performed. The data are sent to the sink, after the data aggregation. When the aggregation is performed over the (CH) will also consume energy. But in homogeneous sensor network clusterhead will soon die out. Thus, re-clustering is performed again to maintain the processing operation. This will also consume a large amount of energy.

Node Terminology: where into two subsets the sensor nodes will be divided, this type is considered as a normal model. The subsets are named as the server nodes and the client nodes. Some request will give by the client nodes for the services. The server nodes will give the required services. But in some advanced models, every node is able to give the services to all the other nodes. Different type of nodes is listed below:

- 1. *Normal Sensor Node:* It is a client node. This node is usually used to send data through the data aggregated node to the BS. Aggregator algorithms are used in order to find a good server.
- 2. *Aggregator Node:* it is considered as a server node. The node function is to aggregate data which come from different sensor nodes.
- 3. *Malicious Node:* any node which performs the malicious activity is classified as a malicious node.
- 4. *Intermediate Node:* the node used to forward data consider as an intermediate node, this type of nodes knows as a relay node.
- 5. *Idle Sensor node:* if the node is in sleep mode, it is considered as an idle sensor node.

Topology: different typologies are for the nodes to operate. In this work, we are considering dynamic topology.

Node behavior: each node will get the details of the neighborhood nodes. Thus, every node is aware of the character of each node.

Computation the level of satisfaction of aggregation: The data sensor node sends to the selected AN that will redirect the data to other ANs or BS based on its quality. Then the received results evaluate by (BS) and it computes its satisfaction with the performed aggregation.

Reward and Punishment: if the received aggregated results were satisfying by (BS), the boost will be made in terms of the pheromone additions to the path leading to the final service provider. If the aggregator has is cheated, a punishment is executed in terms of evaporating the pheromone. One amendment we added to the standards of reward and punishment. Means that the node which has cheated will be given a punishment or a reward if just one node is found to misbehave, why we give punishment to the whole path? It is unfair.

Aggregation of nodes is performed by levels for selecting a particular node. In data aggregation sector, to each node level, time slots are given. The level nodes can send a data packet to their parent during the assigned time slots. To avoid the interference and the latency that occur in this phase the level nodes are subdivided into k concurrent sets. It is considered that data can be transmitted from each node to the others without any and this is expressed as $s_1, s_2,..., s_k$. From this, it is clear that the data from s_i can be assigned to send data from ith time slot.

The aim is to separate different level nodes from the concurrent sets. It is assumed that there is no concurrent sets and k = 0. Out of different levels, L_R takes one node n. This n is inserted into some other concurrent set. In the next step, a new set is formed with k = 1 and the node n is inserted in the s_1 . After that from the next sector, another node m is taken. Then, check whether it conflicts is there with any of the other nodes in the same cluster s_i (i = 1, 2....k) or not. Suppose s_1 found that m is into the sector, otherwise the new concurrent set is found to place the new node m. This process will continue up to all the level nodes are placed in the sectors.

4.2.2 Aggregation of Connected Dominating Nodes

When all level nodes are assigned to a certain time slot to transmitting the data packet to their parents the scheduling the time slots of non-level nodes is performed for aggregation sector. The processing of the algorithm starts from the bottom part of the sector and also the algorithm will assign the time slots for the entire connected node in the sector. The dominant nodes are supposed to be assigned at n the even levels, and the connected nodes are assigned to the odd levels. Concurrent sets in every sector are formed by using the algorithm. The dominating nodes in the level d are divided into n_d if the d is even. Similarly, if the connected node is odd then the level c is divided into n. It is considered that the latency of the aggregation of level nodes is assigned by h, schedule the connected dominating nodes from the lower part of the level to the line. The node $x \in LR$ is usually assigned to the j^{th} set in the layer *i*, the time slot scheduling time of x is calculated by:

$$t = h + j + \sum_{d=H}^{i=2} n_d + \sum_{c=H}^{i=1} n_c$$
(21)

Where i is even, where D Indicates the depth of the data aggregation sector.

$$t = h + j + \sum_{d=H}^{i=2} n_c + \sum_{c=H}^{i=1} n_d$$
(22)

The equation (22) above is valid If i is odd.

Data aggregation scheduling algorithm will be the next step. In this step, the send time of all nodes in the network is assigned. The latency of the nodes is assigned here.

Algorithm 2. Node Selection algorithms for AN			
Input	Level – L_{R} , Sector		
Output	Node selection		
1	Initialization the level (L_R), sector and line values.		
2	Search the nodes to get the details of the node x.		
3	Divide the sector into L_1, L_2, \dots, L_R .		
4	For each node x εL_1 , P(x) = s, e' = e' U {(x, s)}s		
5	For $i = 2$ to R		
6	Calculate upper(x), x ε L _i .		
7	If upper(x) = 0 and x ε L _i , then L _i /x, L _{i+1} = L _{i+1} U {u}.		
8	if <i>i</i> is even, construct the layer independent set D_{i} , and sequence the lower (<i>y</i>) with decreasing order of the set size as $y_1, y_2, \dots, y \in L_{i-1}$, and the set size as $y_1, y_2, \dots, y \in L_{i-1}$;		
9	$j=1$, while $L_i !=$ empty do,		
10	Each node $x \in \text{lower } L_i e^i = e^i U \{(x, L_i)\}, L_i = L_i/U.$		
11	Calculate the concurrent set $s_1, s_2,, s_k$.		
12	Select from line k the level of the node x.		
13	Selection of the node: if the d is even divided into n_d otherwise n_c .		
14	Calculate the time slot for the by using the equation $t = h + j + \sum_{d=H}^{i=1} n_d + \sum_{c=H}^{i=1} n_c$ or $t = h + j + \sum_{d=H}^{i=2} n_c + \sum_{c=H}^{i=1} n_d$		
15	Select the sector, level, and line in which the Dominating node is located.		
Return	Return Node Selection.		

4.3 Modified Task Scheduling Algorithm for Compute the Waiting Time and Aggregation Delay for the Re-Forwarding Data

Node and Network assumptions: normally the network consists of mobile nodes along with error free channels

to make communication of the aggregated data. It is considered that from each node and AN will send a single packet in the assigned time slot. Through multiple links, data are transmitted from nodes located at two or more hops from AN. Thus, it is considered that only one channel is used for multiple time slots.

Network Model: Here, also, we are studying the sector and levels of system modeling. It is assumed that in the whole sector there is number of nodes x. Each sector consists of a group of nodes which is represented by G and in each group consists of wireless links which are represented by K. The sector is divided into a sub-group called levels L₁, L₂,.....L_i. Every level consists of clusters and cluster consists of AN. This AN is used for the making the link, each cluster. The BS is the line which collects the aggregated packets from the AN From each sensor nodes present in the network, the AN are used to collect all information from them. Inside the network, It is considered all nodes are mobile, except for both AN and BS are static. For the aggregation process, trust degree of AN is used. Thus, by using the trust degree of AN, the overall energy consumption will be reduced by taking only the AN. The trust degree is calculated by using the following parameters:

Energy: The equation for the energy that is consumed by the WSN-IoT network is described²⁷.

$$f^{energy} = \frac{f^{energy}(m)}{f^{energy}(n)}$$
(23)

$$f^{energy}(m) = \sum_{j=1}^{M} n E(j)$$
(24)

$$nE(j) = \sum_{i=1,i \in j}^{N} (1 - E(S_i) * E(Y_j)); 1 \le j < M$$
 (25)

$$f^{energy}(n) = M * M_{i=1}^{N} (E(S_i)) * E(Y_j))$$
(26)

 $E(S_i)$ and $E(Y_j)$ indicates the energy of i^{th} normal node and energy of j^{th} node.

Delay: Eq. (12) denotes to the delay of the transmission node. The delay values will be between [0,1]. If the number of nodes in the cluster was less, then the delay value will be low. In Eq. (12) the numerator denoted to the highest value node in the WSN, and the total number of nodes denoted by the denominator.

$$f^{delay} = \frac{Max(Y_j)}{N}$$
(27)

Received signal strength: RSS it is measured by taking the distance and transmission energy. The equation of RSS is given by:

$$RSS = \frac{FFE(n,d)}{4\pi D_{i}^{2}} + T_{a,a_{1}/a_{2}}$$
(28)

In WSN the nodes frequently are moving. The velocity of the nodes is the mobility of the nodes.

$$T_{d} = Avg (f^{energy}, RSS, f^{delay})$$
(29)

It is denoted as t is the time slot and the information sensed and collected is denoted by m. At AN all the messages are aggregated. After the aggregation, AN will forward the message to the other AN or BS, which is present in the next time slot. After the first time slot, multiple messages from different nodes will be a collision and this will result in waste energy and the delay. The process of sending the information will be repeated for each time slot until BS receives the finally aggregated message. The common problems in scheduling the aggregated data in the cluster based the network are:

- Scheduling decision is based on the present state for intracluster communication called myopic scheduling.
- Scheduling decision which takes the current and future state of the moving node for inter-cluster communication i.e. for AN and AN to BS is non-myopic.
- Commonly used to increase the throughput, energy consumption by reducing the number of conflicts.
- Delay will be there when communicating the aggregated message to BS.

Parameters for calculating the delay and the waiting time:

Time to the deadline $(D_{\rm T})$: this parameter is used to indicate the time it remains for the packet to before its deadline. This parameter is used for making the routing decisions.

Slack Time (S_T): TTD-EED, proportionately to the remaining hop-count to the sink node along the forwarding

path to judiciously hold packets in the intermediate nodes while surmounting real-time constraints.

End-to-End Delay (D_{EE}) : is the time that takes to deliver the packet from the last forwarding node to the sink node.

Waiting Timeout (W_T) : for a packet at an intermediate node I_M hops away from the sink is calculated as follows²⁸:

$$W_{T} = \frac{D_{T} - D_{EE}}{1 + \left(\frac{I_{M} - 1}{I_{M}}\right)} \alpha$$
(30)

Where α is a constant factor which indicates more time which is used to remain in the safe margin to ensure that the deadline would be met. The prioritization is based on real-time policy and aggregation efficiency. It potentially causes the receivers to randomly delay the data packet transmission to avoid data packet collision.

Selection of a small amount of random delay before sending the data packet indicates the higher prioritization process. It causes the receivers to randomly delay the Clear To Send (CTS) transmission to eliminate CTS collision. By selecting shorter random delay before sending the CTS will indicate the high priority of data transmission. When prioritizing if there is an overhead of data transmission then, CTS will be canceled. Because of the interference between the neighborhood senders, the transmission of multiple CTS data can be removed.

For calculating the priority amount of the nodes, the commonly used parameter is the fitness value. The expression for the fitness is given by:

$$F_{i} = (w) \times \left(1 - \frac{V_{R}}{V_{RL}(C, i)}\right) + (1 - w) \left(\frac{E_{RM}(i)}{E_{IT}(i)}\right) \quad (31)$$

where,

$$V_R = \frac{d(C,S)}{D_T} \tag{32}$$

$$V_{RL} = \frac{d(C,S) - d(N,S)}{E_{HD}}$$
(33)

Where (d) is the distance, S is the sink; E_{HD} is the time to forward a packet from a current node to the next hop

Algorithm 3	• Waiting and Delay time for re-forwarding			
the data				
Input	Aggregator Node - AN, Group - G, Link - K			
Output	Delay and waiting time			
1	Initialization the node x, group the sector as G.			
2	The sector is divided into L_1, L_2, \dots, L_i .			
3	Calculate the value of trust degree T_d and consider it used for collecting the data.			
4	Calculate the V_{R} and V_{RL}			
5	Calculate the fitness value F _i .			
6	$if(F_{i_{1}} > F_{i_{2}})$			
7	Highest priority is the node having fitness value F _{i 1} .			
8	Update F _i as F _{i 1} . Compared to all nodes.			
9	else			
10	Highest priority is the node having fitness value F_{i_2} .			
11	Update F_i as F_{i_2} . Compared to all nodes.			
12	End			
Return Waiting and delay time calculation.				

relay node; E_{RM} and E_{TT} is the residual and initial energy; w is the weight.

5. Result and Discussion

The simulations are performed using the NS2 simulation environment. The simulations are carried out with metrics for both the proposed protocol and existing. From the comparison, it is clear that existing SLMADA has low performance as compared to the proposed system OLDA-WSN. In our proposed the OLDA-WSN consisting of three different approaches are there: group the optimal algorithm, by this algorithm. The energy consumption of the nodes will be reducing. After that, we designed the node selection algorithm to computation the ANs between several nodes in a network. And here, we have done the modified task scheduler algorithm which used to calculate the waiting time and aggregation delay for the re-forwarding data, which reduces the time-bloated state problem.

The two scenarios that are used for the analysis, the payload and the node density using two scenarios the parameters like, delay, lifetime of the network, success ratio, the throughput and average energy is calculated. The total energy that is present in the node in each round of data collection, Is the average energy. The unit of the through is bits per slot, or it can also be defined as the bit per second. Thus, the sum of the data rates that are delivered to all terminals in a network is the aggregate throughput. The time taken for data aggregation is defined as the delay. The amount of time the nodes can be alive is the lifetime of the network. The percentage of successful certificate renewals overall requests during the simulation time is defined as the successive ratio.

For the simulation purpose, the commonly used parameters are taken, and it is listed in Table 1. Number of Nodes which we have taken was from 50 nodes to 300 nodes, the initial energy for each node is taken as the 18720J. The transmission range was 50m.

5.1 Evaluation with Node Density

In this section, parameters such as network lifetime, energy consumption, delay, and throughput and success rate are evaluated based on the number of nodes or node density. The simulation result of energy consumption based on the number of nodes we can see that in the Figure 2, from the figure we can see very clearly that the proposed OLDA-WSN has a reduction in energy consumption compared to other the existing approach such as SLMADA. The OLDA-WSN compared to other the existing approach such as the SLMADA approach consumes 44% lower power. Based on the number of the nodes, aggregation delay simulation result is shown in the Figure 3, we can see that OLDA-WSN is having low aggregation delay as compared to other the existing approach such as SLMADA, the delay is 33% low in the OLDA-WSN protocol as compared to other existing

Table 1.Simulation parameters

Parameter	Values
Number of nodes	50, 100, 150, 200, 250 and
	300
Area	100m × 100m
Transmission range	50 m
Initial battery power	18720J
Maximum data payload size	1024 bytes
Execution time	65s
Simulation time	1000s
BW	2.4 GHz
Data rate	250 kbps
Packet size	97 bytes

approaches, in the Figure 4 illustrates the simulation result of the lifetime of the network in relation to the number of nodes. Compared to the three existing approaches, the OLDA-WSN network has a high network lifetime; there is a 49% improvement in network lifetime when using our proposed OLDA-WSN protocol. The success rate simulation result is illustrated in the Figure 5, it is shown that there is a 59% improvement in the success ration by using our proposed OLDA-WSN protocol, in the Figure 6, the results of the throughput were illustrated, as we can note that the throughput is increased by 44 % when using the OLDA-WSN protocol finally, from all the simulation results, we noted by using the proposed OLDA-WSN protocol, the parameters like the success ratio, lifetime of the network and throughput got increased, and for the other parameters like energy consumption and delay got decreases. Thus, our proposed approach is efficient.

5.2 Evaluation with Maximum Payload

This section gives the simulation analysis of the parameters like energy consumption, aggregation delay, network lifetime, success ratio and the throughput with respect to maximum payload. Figure 7 illustrates the maximum payload of the node with energy consumption.



Figure 2. Energy with number of nodes.



Figure 3. Delay with number of nodes.



Figure 4. Network life with number of nodes.



Figure 5. Success ration with number of nodes.



Figure 6. Throughput with number of nodes



Figure 7. Energy consumption with maximum payload.



Figure 8. Delay with maximum payload.



Figure 9. Network life time.



Figure 10. Success ratio with maximum payload.



Figure 11. Throughput with maximum payload.

We can see very clearly that our proposed OLDA-WSN consumes low power compared to other the existing three protocols. We found that 45% of low energy consumption when proposed OLDA-WSN is used. Figure 8 shows the simulation result of aggregation delay with maximum payload. By using the proposed OLDA-WSN protocol, the aggregation delay will be reduced by 47%. Figure 9 shows the simulation result of network lifetime variation with respect to the maximum payload. From the figure, we can see clearly that there is a 50% improvement in the lifetime of the network. Thus, by using the proposed OLDA-WSN protocol, the lifetime of network is improved as compared with other three the existing protocols. Figure 10 shows the simulation result of the success ratio. By using the proposed OLDA-WSN protocol, it is clear that the success ratio is improved by 38%. Figure 11 shows the simulation of throughput with respect to the maximum payload. It is found that the throughput is improved by 60% when the proposed OLDA-WSN protocol is used. Thus, by using the proposed protocol in the second scenario maximum payload, it is proving that the technique is more effective than the existing three techniques.

6. Conclusion

The goal of this work to identify the bloated state problem or increase of the agent size in the network thus, by proposing OLDA-WSN protocol, it is possible to solve all the above- mentioned problem. The proposed OLDA-WSN protocol is simulated using NS2. For the analysis process, the two scenarios used are the nodes density or number of nodes and the maximum payload. The parameters used for the analysis is energy consumption, aggregation delay, network lifetime, success ratio and throughput. From the analysis, we found that by using our proposed protocol the parameters like network lifetime, success ratio, and throughput is improved. Similarly, delay and energy consumption will be reduced while using the OLDA-WSN protocol. Thus, it is clear that the proposed OLDA-WSN protocol is an efficient approach.

7. Reference

 Liu B, Wu Y. A secure and energy-balanced routing scheme for mobile wireless sensor network, Wireless Sensor Network. 2015; 7:137. https://doi.org/10.4236/ wsn.2015.711012.

- Norouzi A, Zaim AH. An integrative comparison of energy efficient routing protocols in wireless sensor network, Wireless Sensor Network. 2012; 4:65. https://doi. org/10.4236/wsn.2012.43010.
- Shahraki A, Rafsanjani MK, Saeid AB. A new approach for energy and delay trade-off intra-clustering routing in WSNs, Computers and Mathematics with Applications. 2011; 62:1670–76. https://doi.org/10.1016/j.camwa.2011.06.003.
- Ammari HM. On the energy-delay trade-off in geographic forwarding in always-on wireless sensor networks: A multiobjective optimization problem, Computer Networks. 2013; 57:1913–35. https://doi.org/10.1016/j.comnet.2013.03.009.
- Nguyen NT, Liu BH, Pham VT, Luo YS. On maximizing the lifetime for data aggregation in wireless sensor networks using virtual data aggregation trees, Computer Networks. 2016; 105:99–110. https://doi.org/10.1016/j. comnet.2016.05.022.
- Haseeb K, Bakar KA, Abdullah AH, Darwish T. Adaptive energy aware cluster-based routing protocol for wireless sensor networks, Wireless Networks. 2017; 23:1953–66. https://doi.org/10.1007/s11276-016-1269-0.
- Mantri D, Prasad NR, Prasad R. Two tier cluster based data aggregation (ttcda) for efficient bandwidth utilization in wireless sensor network, Wireless Personal Communications. 2014; 75:2589–606. https://doi. org/10.1007/s11277-013-1489-x.
- Krishnan AM, Kumar PG. An effective clustering approach with data aggregation using multiple mobile sinks for heterogeneous WSN, Wireless Personal Communications. 2016; 90:423–34. https://doi.org/10.1007/s11277-015-2998-6
- Li G, Chen H, Peng S, Li X, Wang C, Yu S, Yin P. A collaborative data collection scheme based on optimal clustering for wireless sensor networks, Sensors. 2018; 18:2487. https://doi.org/10.3390/s18082487. PMid: 30071592, PMCid: PMC6111701.
- Li W, Jia B, Saruwatari S, Watanabe T. Waterfalls partial aggregation in wireless sensor networks, International Journal of Distributed Sensor Networks. 2016; 12:2392149. https://doi.org/10.1155/2016/2392149.
- 11. Joo C, Shroff NB. On the delay performance of in-network aggregation in lossy wireless sensor networks, IEEE/ACM Transactions on Networking. 2014; 22:662–73. https://doi.org/10.1109/TNET.2013.2256795.
- 12. Liu BH, Jhang JY. Efficient distributed data scheduling algorithm for data aggregation in wireless sensor networks,

Computer Networks. 2014; 65:73–83. https://doi. org/10.1016/j.comnet.2014.03.003.

- Parmar K, Jinwala DC. Concealed data aggregation in wireless sensor networks: A comprehensive survey, Computer Networks. 2016; 103:207–27, 2016. https://doi. org/10.1016/j.comnet.2016.04.013.
- 14. Villas LA, Boukerche A, De Oliveira HA, De Araujo RB, Loureiro AA. A spatial correlation aware algorithm to perform efficient data collection in wireless sensor networks, Ad Hoc Networks. 2014; 12:69–85. https://doi. org/10.1016/j.adhoc.2011.08.005.
- 15. Gupta GP, Misra M, Garg K. Energy efficient distributed approach-based agent migration scheme for data aggregation in wireless sensor networks, The Journal of Indian Prosthodontic Society. 2015; 11:148.
- 16. Gavalas D, Venetis IE, Konstantopoulos C, Pantziou G. Mobile agent itinerary planning for WSN data fusion: considering multiple sinks and heterogeneous networks, International Journal of Communication Systems. 2017; 30:e3184. https://doi.org/10.1002/dac.3184.
- Yi D, Yang H. HEER-A delay-aware and energy-efficient routing protocol for wireless sensor networks, Computer Networks. 2016; 104:155–73. https://doi.org/10.1016/j. comnet.2016.04.022.
- Cheng CT, Leung H, Maupin P. A delay-aware network structure for wireless sensor networks with in-network data fusion, IEEE Sensors Journal. 2013; 13:1622–31. https:// doi.org/10.1109/JSEN.2013.2240617.
- Thakkar A, Kotecha K. Cluster head election for energy and delay constraint applications of wireless sensor network, IEEE Sensors Journal. 2014. 14:2658–64. https://doi. org/10.1109/JSEN.2014.2312549.
- Cai H, Zhang Y, Yan H, Shen F, Zhou K, Zhang C. A delay-aware wireless sensor network routing protocol for industrial applications, Mobile Networks and Applications. 2016; 21:879–89. https://doi.org/10.1007/s11036-016-0707-7.
- 21. Kumar AK, Sivalingam KM, Kumar A. On reducing delay in mobile data collection based wireless sensor networks, Wireless Networks. 2013; 19:285–99. https://doi.org/10.1007/s11276-012-0466-8.
- 22. Liu W, Zhao D, Zhu G. End-to-end delay and packet drop rate performance for a wireless sensor network with a cluster-tree topology, Wireless Communications and Mobile Computing. 2014; 14:729–44. https://doi. org/10.1002/wcm.2230.

- 23. Yao Y, Cao Q, Vasilakos AV. EDAL: An energy-efficient, delay-aware, and lifetime-balancing data collection protocol for heterogeneous wireless sensor networks, IEEE/ACM Transactions on Networking. 2015; 23:810–23. https://doi.org/10.1109/TNET.2014.2306592.
- Gupta GP, Misra M, Garg K. Towards scalable and loadbalanced mobile agents-based data aggregation for wireless sensor networks, Computers and electrical engineering. 2017; 64:262–76. https://doi.org/10.1016/j. compeleceng.2017.10.020.
- Kumar R, Kumar D. Multi-objective fractional artificial bee colony algorithm to energy aware routing protocol in wireless sensor network, Wireless Networks. 2016; 22:1461–74. https://doi.org/10.1007/s11276-015-1039-4.
- Yadav AK, Tripathi S. QMRPRNS: Design of QoS multicast routing protocol using reliable node selection scheme for MANETs, Peer-to-Peer Networking and Applications. 2017; 10:897–909. https://doi.org/10.1007/s12083-016-0441-8.
- 27. Reddy MPK, Babu MR. Implementing self adaptiveness in whale optimization for cluster head section in Internet of Things, Cluster Computing. 2018; 1–12. https://doi. org/10.1007/s10586-017-1628-3.
- 28. Harris AF, Kravets R, Gupta. Building trees based on aggregation efficiency in sensor networks, Ad Hoc Networks. 2007; 5:1317–28. https://doi.org/10.1016/j. adhoc.2007.02.021.