World Engineering & Applied Sciences Journal 6 (3): 136-146, 2015
ISSN 2079-2204

DOI: 10.5829/idosi.weasj.2015.6.3.22229

Double Cluster Head Based Reliable Data Aggregation for Wsn

[1]D. Suresh and [2]K. Selvakumar

[1]Department of Computer Science and Engineering,
University of Annamalai,Annamalai Nagar, Chidambaram, India
[2]Department of Computer Science and Engineering, Annamalai University,
Annamalai Nagar, Tamilnadu, India

Abstract: In Wireless Sensor Network (WSN), sensors are deployed to collect the data and send them to the base station. During data aggregation, the collected data is merged before sending it to the base station. Since a node is typically deployed in locations that are accessible to malicious attackers, information guarantee of the data fusion process is important. To keep the nodes safe from the attacks, Double Cluster Head based Reliable Data Aggregation for WSN approach has been proposed in this paper. In this approach, the two cluster heads namely main and sub-ordinate cluster heads are selected based on the parameters such as residual energy, minimum average distance from the member, nodes timer and node degree using particle swarm optimization technique. Then a direct voting approach is implemented, where the nodes act as witness nodes which help the base station to verify the data and find if any node is affected in the network. Results show that in the proposed aggregation technique, the attackers could compromise 1% of all packets, but because of DCRDA 50% of corrupt packets are identified.

Key words: Data aggregation · Clustering · CRR scheme · Master Cluster Head · Node Density

INTRODUCTION

WSN: WSN consists of inexpensive sensor nodes, each node having continuous sensing capability with limited communication power. WSN is applied in commercial, civil and military applications including vehicle tracking, climate monitoring and intelligence, medical and agriculture, etc. Sensor nodes have inbuilt chips and software for processing specific function. The security application of a wireless sensor network gives one the ability to collect data for analysis remotely and detects attack. In military applications, the information passed over the network must be secure.

Many of the features of the wireless sensor networks give rise to challenging problems.

The most important three characteristics are:

- Sensor nodes are the ones, which are prone to maximum failures.
- Sensor nodes make use of the broadcast communication pattern and have severe bandwidth restraint.
- Sensor nodes have limited amount of resources.

Data Aggregation: Data aggregation is considered as one of the fundamental distributed data processing procedures for saving the energy and minimizing the medium access layer contention in wireless sensor networks. Data aggregation is presented as an important pattern for routing in the wireless sensor networks. The basic idea is to merge the data from various sources, reroute it with the elimination of the redundancy and thus reducing the number of transmissions and saving the energy. The inbuilt redundancy in the raw data is gathered from various sensors. This can be prevented by the in-network data aggregation. Additionally, these operations use raw materials for obtaining application specific information. To preserve the energy in the system for maintaining longer lifetime in the network, it is important for the network to maintain high incidence of the in-network data aggregation.

Cluster Based Aggregation: Clustering is an important step in the process of data analysis with applications to numerous fields. Informally, clustering is defined as the

Corresponding Author: D. Suresh, Department of Computer Science and Engineering,
University of Annamalai, Annamalai Nagar, Chidambaram, India.

problem of partitioning data objects into groups (clusters), such that objects in the same group are similar, while objects in different groups are dissimilar. The process of grouping the sensor nodes in a densely deployed large-scale sensor network is known as clustering. The intelligent way to combine and compress the data belonging to a single cluster is known as data aggregation in cluster based environment.

Issues: There are some issues involved with the process of clustering in a wireless sensor network. First issue is, how many clusters should be formed that could optimize some performance parameter. Second could be how many nodes should be taken in to a single cluster. Third important issue is the selection procedure of cluster-head in a cluster. Another issue is that user can put some more powerful nodes, in terms of energy, in the network, which can act as a cluster-head and other simple node work as cluster-member only.

The cluster head means data aggregator nodes send fuse these data to the base station. This cluster head or aggregator node may be attacked by malicious attacker. If a cluster head is compromised, then the base station (sink) cannot be ensure the correctness of the aggregate data that has been send to it. Several copies of the aggregate result may be sent to the base station (sink) by uncompromised nodes. It increase the power consumed at these nodes.

Problem Identification: In our previous work, two cluster heads namely main and sub-ordinate cluster heads are selected based on the parameters such as residual energy, minimum average distance from the member, timer of the node and node degree using particle swarm optimization technique. When timer of node expires, node declares itself as cluster head. When a source node wants to transmit a data to sink, energy efficient routing protocol is used based on the parameters such as the expected number of retransmissions and link failure probability. In this technique, each cluster member node sends the data to main cluster head. The aggregated data from the main cluster head is transmitted to sink through sub-ordinate cluster head.

Literature Review: Woo-Sung Jung *et al*. [1] have proposed a hybrid clustering based data aggregation scheme. The proposed scheme can adaptively choose a suitable clustering technique depending on the status of the network, increasing the data aggregation efficiency as well as energy consumption and successful data transmission ratio Their simulation results show that the effectiveness of the proposed scheme. However, the overhead analyses, delay metric and data aggregation are not handled.

Xiangbin Zhu *et al*. [2] have proposed the node clustering algorithm. Optimization of mobile agent in the routing within the cluster strategy for wireless sensor networks to further reduce the amount of data transfer. Through the experiments, using mobile agents in the integration process within the cluster can be reduced the path loss in some extent.

Hiren Thakkar *et al*. [3] have proposed a modification in Low Energy Adaptive Clustering Hierarchy (LEACH) protocol. Their modified protocol is considering residual energy as a criterion for a node to be a cluster head during Cluster head selection and Clusters setup phase. They also proposed multi-level data aggregation among Cluster heads to reduce the packet size which in turn reduces the transmission and receiving energy for a node. They also proposed multi-hop transmission of aggregated data. The data aggregated by Cluster heads will not be transmitted to base station directly but through multi-hop transmission by Cluster heads which are nearer to base station, which in turn reduces the transmission distance and so as energy consumption of nodes. Their main focus to achieve energy efficiency is by reducing packet size by multi-level data-aggregation among Cluster heads and by proper selection of nodes as Cluster heads by considering maximum residual energy of a node as a constraint.

Hung-Ta Pai *et al*. [4] have proposed the base station receives the fusion data and "votes" on the data from a randomly chosen sensor node. The vote comes from other sensor nodes, called "witnesses," to confirm the correctness of the fusion data. Since the base station receives the vote through the chosen node, this node could forge the vote if it is compromised. Accordingly, the witness node must apply cryptographic operations to the vote to prevent this forgery. The cryptographic operation requires more bits than the vote, increasing the transmission burden from the chosen node to the base station. The chosen node consumes too much power. This work improves the witness-based approach using a direct voting mechanism such that the proposed scheme performs better in terms of assurance, overhead and delay. The witness node transmits the vote directly to the base station. Forgery does not pose a problem in this scheme. Moreover, fewer bits are necessary to represent the vote, significantly reducing the power consumption. Performance analysis and simulation results indicate that the proposed approach has a 40-times lower overhead than the witness-based approach. However, data fusion and energy efficiency are not considered.

Weilian Su *et al.* [5] have proposed a data fusion algorithm that combines the cluster-based design of WSNs using fuzzy logic methods. Simulation results show that the algorithm eliminates redundant sensor reports.

Preethi Y.R. *et al*. [6] have proposed a Cluster Based Data Routing for In-Network Aggregation that has some key aspects such as a reduced number of messages for setting up a routing tree, maximized number of overlapping routes, high aggregation rate and reliable data aggregation and transmission & provides the best aggregation quality when compared to other existing algorithms.

Zhonghua Wang [7] has proposed a novel double cluster-head routing policy based on clustering hierarchy routing. According to balancing the number of neighbour nodes, surplus energy and distance weights, it adopts first and second cluster-head mode. It will increase the overall performance of the network, decrease the cluster head energy consumption, balance energy consumption between member nodes and prolong the entire lifetime of network. However, they did not consider the cluster head election time and the node degree in the selection of cluster head. There is no energy consumption model in this method

Mei Yang *et al*. [8] have proposed the Coordinated Robust Routing (CRR) scheme to address the fault tolerance requirements in the layered WSN. The robust routing trees are constructed co-ordinately from the most-outward layer all the way to the sink node. The coordination is done by selecting two dedicated cluster heads for every two clusters in one layer. The problem of selecting dual cluster heads is formulated as a transportation problem, which can be solved using network flow algorithms. By having two cluster heads for every cluster pair, the CRR scheme helps to achieve fault tolerance and energy efficiency with low degree of network redundancy and inter-packet delay.

Weike Chen *et al*. [9] have proposed QoS-based Adaptive Clustering (QAC) algorithm that concerns energy consumption and improves the reliability and the steadiness of wireless sensor networks by establishing a dual cluster-head model. QAC proposes a local-centralized mechanism for electing cluster-head and suggests a parameter to measure the QoS support in hierarchical applications of WSNs. This model can increase the reliability and the steadiness of wireless sensor network by distributing evenly the communication load and the load of data fusion among cluster-heads. The dual cluster-head model can also improve the survival ability of wireless sensor networks and makes the network fitting in with the fierce changes of environment.

Zhang Ruihua *et al*. [10] have proposed a double cluster-heads clustering algorithm using particle swarm optimization (PSO-DH) that generates two cluster heads. The election of the master cluster-head and the vice cluster-head needs consider the state information. Master Cluster Head (MCH) receives and aggregates the data from its member nodes. The aggregation data are sent to the vice one. Vice Cluster Head (VCH) transmits aggregation data to the sink directly. This algorithm can balance the energy consumption, so it can extend the network lifetime effectively. Simulation results show the lifetime of the algorithm is extended for 50% contrast with LEACH.

Stefano Basagni [11] has proposed Distributed Clustering Algorithm and Distributed Mobility-Adaptive Clustering for efficient partitioning of the nodes into clusters with a cluster head and some ordinary nodes. Weight-based criteria are introduced for the cluster formation that allows the choice of the cluster heads based on node mobility-related parameters. It guarantees fast inter- and intra-cluster communication between each pair of nodes. The DCA is easy to implement and its time complexity is proven to be bounded by a network parameter. It combines the easiness of implementation with full adaptation to the mobility of the nodes, even during clustering set up.

Min Qin and Roger Zimmermanna [12] have proposed a novel Voting-based Clustering Algorithm (VCA) for energy-efficient data dissemination in wireless sensor networks. Here, the sensor nodes vote for their neighbors to elect suitable cluster heads. VCA is completely distributed, location unaware and independent of network size and topology. It combines load balancing, energy and topology information together by using very simple voting mechanisms. VCA can reduce the number of clusters and prolong the lifetime of the network.

Younis, O. *et al*. [13] have proposed a new energy-efficient approach for clustering nodes in ad-hoc sensor networks. Based on this, Hybrid Energy-Efficient Distributed clustering (HEED)is proposed that periodically selects cluster heads. HEED does not make any assumptions about the distribution or density of nodes, or about node capabilities, e.g., location-awareness. HEED incurs low overhead in terms of processing cycles and messages exchanged and achieves fairly uniform cluster head distribution across the network.

Proposed Solution

Overview: As an extension work, a new aggregation technique is proposed by combining the adaptive

clustering and direct voting technique. Initially, clustering process is performed by using Particle Swarm Optimization (PSO). Then, the main cluster head (CH_M) is selected for data transmission to sink. If the volume of traffic is high, the sub ordinate cluster heads (CH_S) can be dynamically chosen for aggregation. The data arrival rate of cluster members will be monitored at CH_M. If the aggregate arrival rate is more than a maximum threshold, CH_M will send an overflow (OF) warning message to the cluster members and the overflow data directly to CH_S. CH_S then receives both the aggregated data and the overflow data from the CH_M. Then CH_S sends the final aggregated data towards the sink. Then, the direct voting technique is applied. When CH_M send its aggregated results to CH_S, then CH_M will serve as witness nodes. The witness nodes begin to vote on the transmitted result, directly to the base station, thereby validating the results of the cluster head. This technique helps to identify the faults and attacks on cluster heads [14-21].

The notations used in the proposed solution are summarized in the following table.

Notations	Definitions
E_{res}	Residual Energy
N_i	Node
E_i	Initial energy of the node
E_{tx} and E_{rx}	Energy utilized at the time of transmission and reception of data
D	Average distance
Tx_r	Transmission range
h	Hop count among the nodes
η	Operation wavelength
P_{tx}	Power transmitted by the sensor
P_{rx}	Sensitivity of the receiver
α and β	Transmitter gain and receiver gain
ε	Reflected power co-efficient of receiving antenna
t	Time at which node sends the cluster head election message
T_{LCH}	Lasting time of electing CH
τ	Randomly generated real value
E_{ar}	Average residual energy
ND	Node Density
ζ	Communication range
K	Number of links among the member nodes and cluster head
P_F	Link failure probability
n	Number of failure count
E_{RTx}	Expected number of re-transmissions
R_c	Cluster radius
Txr_{min}	Minimum transmission range
D_{exp}	Expected distance from cluster member to CH
P_i and V_i	Position and velocity
F_i	Fitness function
α_1, α_2, α_3, α_4	Weight Value
P_{bi}	Local best position
P_{gi}	Global best position
λ	Learning factor
R	Random numbers
ω	Inertia weight
A_{data}	Aggregated data arrival traffic
N	Number of data arrivals
T	Time at which the data arrived.
Th	Threshold
R_i	Reputation value
A_1, A_2	Weight factors
D_{tx}	Current data
P	Historical Data
Th_R_i	Threshold for reputation value
M`	Size of the witness set
P_f	Probability
W_N	Number of agreeing node
A_{nodes}	Number of witness nodes which agreed with the transmitted result
D_{nodes}	Number of unpolled witness nodes

Estimation of Metrics

Residual Energy: After performing one data communication, the residual energy (E_{res}) of each node (N_i) is estimated by:

$$E_{res} = E_i - (E_{tx} + E_{rx}) \quad (1)$$

where, E_i = Initial energy of the node
E_{tx} and E_{ra} = energy utilized at the time of transmission and reception of data.

Distance: The minimum average distance among the sensor nodes is defined as the product of transmission range and corresponding hop counts. Average distance is given below:

$$D = Ta_r * h \quad (2)$$

where,
Ta_r = Transmission range
h = Hop count among the nodes.

Nodes Timer: The nodes timer is used during the selection of the cluster head. It is used to reduce the transmission of competing message of single node twice.

Each node in the network can send the cluster head election message at time t1 or t2, which is expressed using the following equation [17],

$$t_1 = \frac{T_{LCH}}{2} * \frac{E_{ar}}{E_{res}} * \tau , \qquad E_{res} \geq E_{ar} \tag{3}$$

$$t_2 = T_{LCH} - \frac{T_{LCH}}{2} * \frac{E_{res}}{E_{ar}} , \quad E_{res} < E_{ar} \tag{4}$$

where,

T_{LCH} = Lasting time of electing CH and it is pre-defined.
τ = Randomly generated real values distributed in the range [0.9, 1]
E_{ar} = Average residual energy.

Node Degree: Node Density (ND) is closely linked with the node degree, which reveals the number of neighbours. It is used to find the average distance of neighbour nodes that are connected to the nodes [17].

$$ND = \frac{N_{Deg}}{\pi * \zeta^2} \text{ (for nodes per unit area)} \tag{5}$$

where ζ = communication range

Expected Number of Re-Transmissions: Let K is the number of links among the member nodes and cluster head. This reveals that K number of transmissions is required to deliver the packets successfully to CH.

Let P_F be the link failure probability
Probability of K successful transmission (i.e. end-to-end data delivery to CH) = $(1- P_F)^K$.
Probability of at least one unsuccessful data delivery = $1-(1- P_F)^K$.
Hence, n failure count after one successful data delivery is given by

$$P[n] = [1-(1- P_F)^K]^n * (1- P_F)^K \tag{6}$$

The expected number of re-transmissions leading to one successful data delivery is estimated using the following equation.

$$E_{RTx} = \frac{1}{(1 - P_F)^h} \tag{7}$$

where h = hop count

The hop count is given by,

$$h = \frac{2R_c}{3Txr_{\min}\sqrt{K}} \tag{8}$$

where,

R_c = Cluster radius
Txr_{min} = Minimum transmission range
$2R_c/3\sqrt{K} = D_{exp}$ = Expected distance from cluster member to cluster head.

Clustering: The sensor nodes are uniformly distributed within the network. A Particle Swarm Optimization (PSO) based clustering technique is then applied. The steps involved in the clustering process are as follows:

- Swarm particles (SP_i) are initialized in the deployed network such that the particle's position is randomly dispersed in space. Each SP_i represents a search window equivalent to the nodes position and velocity (P_i, V_i).
- Each SP_i monitors the parameters of each node that includes the residual energy, minimum average distance from the member, nodes timer and node degree.
- Based on the monitored parameters, fitness function (F_i) of each particle is estimated based on below Eq. (10).

$$F_i = \alpha_1 f_1(i) + \alpha_2 f_2(i) + \alpha_3 f_3(i) + \alpha_4 f_4(i) \tag{9}$$

where,

$f_1(i) = E_{resi}$
$f_2(i) = D_i$
$f_3(i) = NT_i$
$f_4(i) = ND_i$
i.e. $F_i = (\alpha_1 * E_{res_i}) + (\alpha_2 * D_i) + (\alpha_3 * NT_i) + (\alpha_4 * ND_i)$ (10)

α_1, α_2, α_3 and α_4 are weight values that are tuned between [0, 1] according to the monitored parameters.

- The local best (P_{bi}) and global best (P_{gi}) value of fitness and position of each particles is estimated.
- The position of P_{bi} and P_{gi} based on following condition.

If $F_i > F(P_{bi})$

Then

Update the position of P_{bi} with the fitness value F_i

End if

If $F_i > F(P_{gi})$

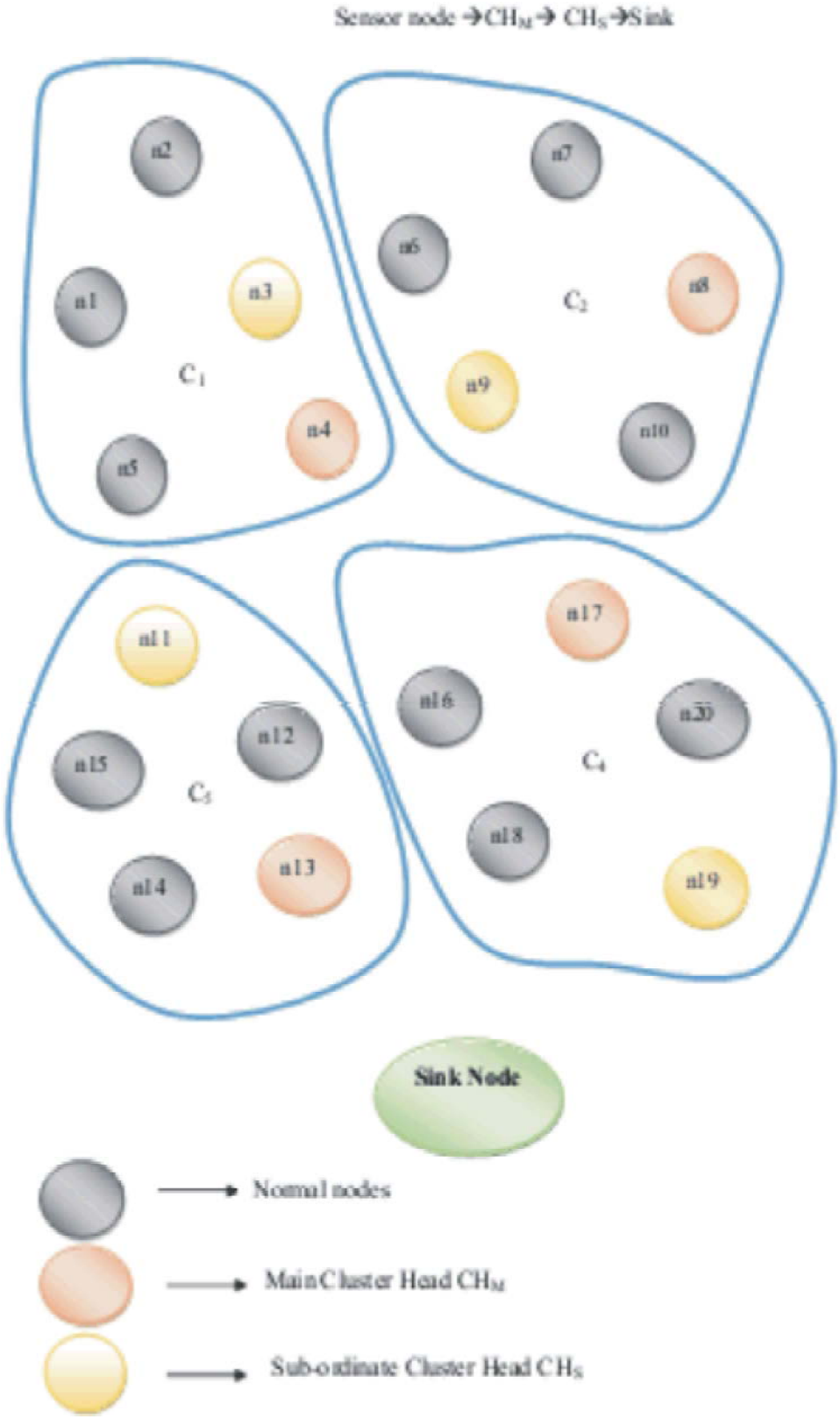

Fig. 1: Cleasering Technique

Then
Update the position of P_{gi} with fitness value F_i
End if

- The velocity and position of each particle is updated using Eq (11) and (12), respectively.

$$V_{bi}(t+1) = (\alpha_1 * E_{res_i}) + (\alpha_2 * D_i) + (\alpha_3 * NT_i) + (\alpha_4 * ND_i) \quad (11)$$

$$Q_{bi}(t+1) = Q_{bi}(t) + V_{bi}(t+1) \quad (12)$$

Here, V is the velocity of the particle, Q is the position of the particle, t is time, λ_1 and λ_2 are learning factors, R_1 and R_2 are random numbers among 0 and 1, P_{bi} and P_{gi} are best and global position of particles, ω is inertia weight.

- The value updated in the global best particle is considered as the best value and the respective node is chosen as the main cluster head (CH_M).
- The next best global value corresponds to the sub-ordinate cluster head (CH_S)
- CH_M transmits the message that contains details about the CH_M and CH_S to all the nodes.
- The data flow from nodes to sink occurs in the following manner

Figure 1 Shows the clustering process where the nodes are grouped into four clusters C_1, C_2, C_3 and C_4. In each cluster, a cluster head is selected (i.e. n4, n8, n13, n17). The cluster head then selects the sub-ordinate cluster head (that is described in section 3.3.1). Node n3, n9, n11 and n19 are the sub-ordinate cluster head.

Selection of Sub-Ordinate Cluster Head: When the cluster traffic is high or the traffic level is higher than the threshold value, then CH_M selects sub ordinate cluster head CH_S. This CH_S is responsible for the aggregation of the data in the cluster and study of data flow between the cluster nodes. If the aggregated data arrival is higher than the threshold, then CH_S sends an overflow (OF) message to all the nodes in the cluster. CH_M sends the overflowed data and aggregated data to CH_S. This received data will be sent to the sink node from CH_S.

Step 1: Initially based on the traffic level in the cluster, CH_M selects a sub ordinate cluster head CH_S. CH_S is responsible for the data aggregation.

Step 2: Then estimate the aggregated data arrival traffic.

$$A_{data} = N / T$$

where N is the number of data arrivals

T is the time at which the data arrived.

Step 3: If the estimated aggregated data arrival is greater than the threshold value, then an overflow (OF) message will be sent to all the cluster nodes. CH_M sends the overflow data to the CH_S. CH_S then transmitted the data to sink. Otherwise, CH_M directly sends the data to sink.

If $A_{data} > Th$,

Then
CH_M send OF message to cluster nodes and OF data to CH_S
CH_S send data to sink
Else
CH_M sends the aggregated data to the sink node
End if

Figure 2 Show the transmission of data from cluster members to the sink. Here, node n3 is a main cluster head. Node n3 selects the node n4 as a sub-ordinate cluster head. Cluster members n5, n1 and n2 send the data to n4. Sub-ordinate cluster head n4 send the aggregated data along with the overflow data to n3. N3 then forwards the data to the sink.

Direct-Voting: The direct voting approach in the witness-based method is designed according to the Message Authentication Code (MAC) of the fusion result at every witness node. The witness nodes are selected according to their reputation values.

This method is sensible when the witness node does not know the fusion result at the selected node. But, in practice, the base station can transmit the fusion result of the selected node to the witness node. Thus, the witness node can gain the transmitted fusion result from the selected node through the base station. The witness node can then relate the transmitted fusion result with its own fusion result. Finally, the witness node can send its vote (agreement or disagreement) on the transmitted result straight to the base station, rather than through the selected node.

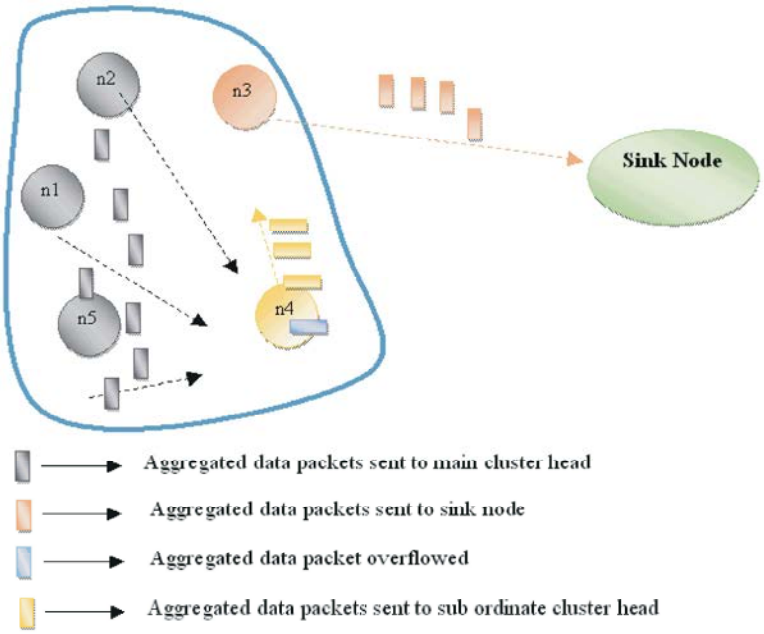

Fig 2: Data Transmission to Sink Node

Estimation of Reputation Values:

Let Ch_i be the cluster member node
Let D_{ot} be the current data transmitted by Ch_i.
Let P be the Ch_i's previous history data for the time period of T seconds.
Let Dr_y be the data set received from the CM_i's neighbor nodes.
The reputation value of MN_i (T_i) can be computed using the following formula

$$R_i = f(A_1 D_{tx}, A_2 P, (1 - A_1 - A_2) DS) \quad (13)$$

where A_1 and A_2 are constant weight factors between 0 and 1.

In the above reputation estimation, the value P plays a major role as it helps in accurate estimation of malicious node activity and data validation. Eq. (13) is utilized in malicious node detection among the cluster member and cluster head.
For detecting the malicious node among the cluster head and base station, the following reputation value is estimated.

$$R_{ri} = f (A_1 D_{tx} , A_2 P) \quad (14)$$

The above reputation value is based on the current data (D_{tx}) and previous historical data of the CH (P).

Malicious Cluster Member Detection:

- Initially, each CH sets a threshold for reputation value (Th_R_i)) with respect to its sensing zone.
- CH monitors the reputation value of data transmitted by its cluster members (CM_i) and detects whether CM_i is malicious or not.

If $R_i < Th\_R_i$

Then

Cm_i is marked as malicious node (MXN_i)

End if

When reputation value of data transmitted by the cluster member is less than the threshold, the member node is considered to be malicious. Otherwise, it is considered as witness node.

Algorithm of Direct Voting Approach:

Step 1: Initially the base station selects a Main Cluster Head (CH_M). Other cluster nodes serve as witness nodes. Define a sequence of witness nodes that includes all witness nodes and let the nodes in the set be randomly ordered. M` = M – 1 represents the size of the witness set in the current round.

Step 2: If the volume of the traffic is high, CH_M selects a Sub-ordinate Cluster Head (CH_S). This CH_S monitors the data flow in the cluster.

$A_{data} > Th$ = send OF message to the nodes

CH_S aggregates the data in the cluster, if the aggregated data (A_{data}) arrival rate is higher than the threshold (Th). The CH_S sends an overflow (OF) messages to all the nodes in the cluster.

Step 3: $P_f = 1$

If in case any compromised or malicious node functions as a witness node, then the node disagrees with the correct result and agrees with the fake result with a probability P_f. If the compromised node tries to make the base station accept the fake result, then it always agrees with the result that is transmitted by other compromised nodes in the cluster. Then at most two rounds of polling have to be implemented.

Step 4: $P_f = 0$

If the compromised node wants to make the polling process run for as long as possible, then it continuously disagrees with the transmitted result.

Step 5: The chosen secondary node transmits its final fusion result to the base station.

Step 6: The base station polls and sends the above fusion result to all the nodes in the witness set by following the order of the witness nodes. The polling process does not break until.

- W_N witness nodes should agree with the fusion result (agreeing nodes), where $1 = W_N = M`$,
- $M` - W_N + 1$ witness nodes should disagree with the fusion result (disagreeing nodes), or
- All witness nodes in the cluster have been polled.

Step 7: A_{nodes} represents the number of witness nodes, which agreed with the transmitted result. D_{nodes} denotes the total number of witness nodes, which disagreed with the transmitted results and the number of unpolled witness nodes.

$A_{nodes} + D_{nodes} = M`$

Step 8: If $A_{nodes} = W_N$,

Then the transmitted result is verified. Break the polling.

Else If $A_{nodes} < W_N$ and $D_{nodes} < W_N + 1$,

Then no reliable result is valid. Stop the polling.

Else If $D_{nodes} = W_N + 1$,

Then exclude the A_{nodes} witness nodes from the witness set.

Let the first node that disagrees with the transmitted result be the selected node to send the results.

Therefore, the updated size of the witness set, M`, is $D_{nodes} - 1$.

Go to Step 2 for the next round of the polling.

Simulation Results

Simulation Model and Parameters: The Network Simulator (NS2) [13], is used to simulate the proposed architecture. In the simulation, 50 mobile nodes move in a 1000 meter x 1000 meter region for 50 seconds of simulation time. All nodes have the same transmission range of 250 meters. The simulated traffic is Constant Bit Rate (CBR). In our simulation, 4 source nodes send their sensor data to the sink. We have two jamming attack nodes along the same channel.

The simulation settings and parameters are summarized in Table 1.

Performance Metrics: The proposed Double Cluster head based Reliable Data Aggregation (DCRDA) is compared with the Double Cluster-Heads algorithm (DCA) [12]. The performance is evaluated mainly, according to the following metrics.

- Packet Delivery Ratio: It is the ratio between the number of packets received and the number of packets sent.
- Packet Drop: It refers the average number of packets dropped during the transmission
- Delay: It is the time taken by the nodes to transmit the data packets to the receiver.
- Energy Consumption: It is the amount of energy consumed by the nodes to transmit the data packets to the receiver.

Table 1: Simulation Parameters

No. of Nodes	50
Area Size	1000 X 1000
MAC	IEEE 802.11
Transmission Range	250m
Simulation Time	50 sec
Traffic Source	CBR
Packet Size	512
Rate	100,200,300,400 and 500 kbps
Routing Protocol	AODV
Initial Energy	20.1J
Transmission Power	0.660 W
Receiving Power	0.395 W
Antenna	OmniAntenna

RESULTS

Based on Rate: In our experiment, we vary the transmission rate as 100,200,300,400 and 500Kb.

Figure 2 shows the end-to-end delay of DCRDA and DCA techniques for different transmission rate scenario. We can conclude that the end-to-end delay of our proposed DCRDA approach has 18% of less than DCA approach.

Figure 3 shows the delivery ratio of DCRDA and DCA techniques for different transmission rate scenario. We can conclude that the delivery ratio of our proposed DCRDA approach has 42% of higher than DCA approach.

Figure 4 shows the packet drop of DCRDA and DCA techniques for different transmission rate scenario. We can conclude that the packet drop of our proposed DCRDA approach has 25% of less than DCA approach.

Figure 5 shows the energy consumption of DCRDA and DCA techniques for different transmission rate scenario. We can conclude that the energy consumption of our proposed DCRDA approach has 6% of less than DCA approach.

Figure 6 shows the throughput of DCRDA and DCA techniques for different transmission rate scenario. We can conclude that the throughput of our proposed DCRDA approach has 50% of higher than DCA approach.

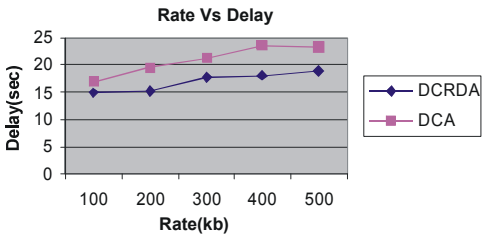

Fig. 2: Rate Vs Delay

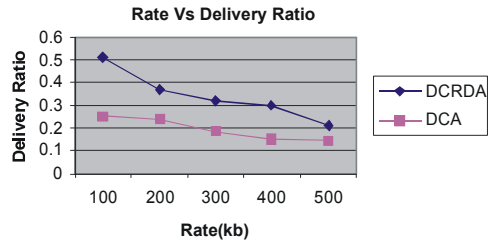

Fig. 3: Rate Vs Delivery Ratio

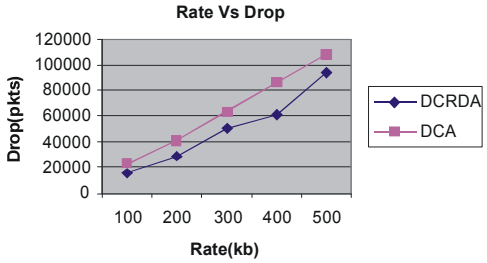

Fig. 4: Rate Vs Drop

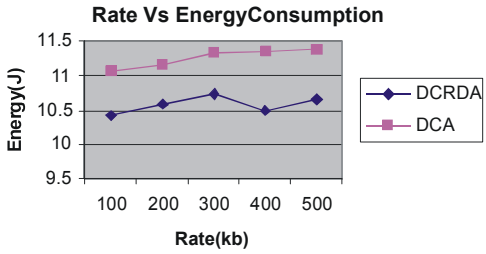

Fig. 5: Rate Vs Energy Consumption

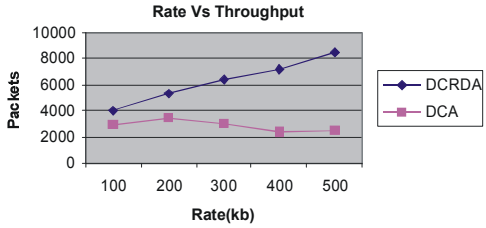

Fig. 6: Rate Vs Throughput

CONCLUSION

In this paper a Double ClusterHead based Reliable Data Aggregation for WSN approach has been proposed. In order to provide a secured data flow between the base station and the nodes. This approach selects two cluster heads namely main and sub-ordinate cluster heads are selected based on the parameters such as residual energy, minimum average distance from the member, nodes timer and node degree using particle swarm optimization technique. A direct voting method has been implemented where the nodes acts as witness nodes. Through these witness nodes, it is possible to verify the data flow between the nodes and the base station. Through this

approach, it is possible to identify if any node in the cluster node has got affected by any compromised node or malicious node in the network. By simulation results, we have shown that the proposed aggregation technique enhances the network lifetime by reducing the energy consumption and packet drops.

REFERENCES

1. Jung Woo-Sung, Keun-Woo Lim, Young-Bae Ko and Sang-Joon Park, 2009. A Hybrid Approach for clustering-based Data aggregation in wireless sensor network, Digital society, ICDS Third International conference at Cancun.
2. Zhu, Xiangbin and Wenjuan Zhang, 2010. A Mobile agent-based clustering Data Fusion Algorithm in WSN, International journal of electrical and computer Engineering.
3. Hiren Thakkar, Sushruta Mishra and Alok Chakrabathy, 2012. A Power Efficient Cluster-based Data Aggregation Protocol for WSN (MHML), International Journal of engineering and Innovative Technology (IJEIT), 1(4).
4. Pai, Hung-Ta and Yunghsiang S. Han, 2008. Power-Efficient Direct-Voting Assurance for Data fusion in wireless Sensor networks, IEEE Transactions on Computers, 57(2).
5. Weilian, Su and Theodoros C. Bougioukis, 2007 Data Fusion Algorithms in cluster-based Wireless Sensor networks using Fuzzy Logic Theory, Proceedings of the 11th WSEAS International conference on communications.
6. Preethi, Y.R, C.R. Manjunath and M. Manohar, 2013 Data Routing in In-Network Aggregation in WSN: a Cluster Based approach, International Journal of Modern Engineering Research (IJMER), 3(3).
7. Zhong-hua Wang, Cheng-wu Zou and Min Yu, 2011. Double Cluster-Heads algorithm for wireless sensor networks using PSO, International Conference on Computational and Information Sciences.
8. Mei Yang, Wang Jianping, Z. Gao and Jiang Yingtao, 2005. Coordinated robust routing by dual cluster heads in layered wireless sensor networks, Parallel Architectures, Algorithms and Networks, ISPAN 2005. Proceedings. 8th International Symposium on 2005.
9. Chen Weike, Li Wenfeng, Shou Heng and Bing Yuan Chen, 2006. A QoS-based adaptive clustering algorithm for wireless sensor networks, Mechatronics and Automation, Proceedings of the 2006, International Conference on. IEEE.
10. Zhang Ruihua, Jia Zhiping, Li Xin and Han Dongxue Ruihua, 2011. Double cluster-heads clustering algorithm for wireless sensor networks using PSO, Industrial Electronics and Applications (ICIEA), 6th IEEE Conference on. IEEE, pp: 763-766.
11. Basagni, S., 1999. Distributed Clustering for Ad Hoc Networks, In ISPAN, pp: 310-315.
12. Qin, M. and R. Zimmermann, 2007. VCA: An Energy-Efficient Voting-Based Clustering Algorithm for Sensor Networks, Journal of Universal Computer Science, 13(1): 87-109.
13. Younis, O. and S. Fahmy, 2004. Distributed Clustering in Ad-hoc Sensor Networks: A Hybrid, Energy-Efficient Approach, In INFOCOM.
14. Umashankar, M. and C. Chandrasekar, 2010. Witness based and voting based Data fusion Assurance Mechanism in sensor networks, IJRRAS, 4(2).
15. Bhoopathy, V. and R.M.S. Parvathi, 2011. Energy Efficient Secure Data Aggregation Protocol for Wireless sensor networks, European Journal of Scientific Research, 50(1).
16. Aristides gionis, Heikki manila and Panayiotis Tsaparas, 2007. Clustering aggregation, Journal ACM transactions on knowledge Discovery from Data (TKDD), 1(1).
17. Patil, Nandini S. and P.R. Patil, 2010. Data aggregation in wireless sensor network, IEEE international conference on computational Intelligence and computing Research.
18. Maraiya Kiran, Kamal Kant and Nitin Gupta, 2011. Wireless Sensor Network: A Review on Data Aggregaion, International Journal of Scientific and Engineering Research, 2(4).
19. Network Simulator: http:///www.isi.edu/nsnam/ns.
20. Suresh, Energy Efficient Double Cluster Head Selection Algorithm For WSN, Journal of Theoretical and Applied Information Technology (JATIT).
21. http://en.wikipedia.org/wiki/Low_Energy_Adaptive_Clustering_Hierarchy.