Multi-Level Reliable Data Aggregation in Wireless Sensor Network

R. Jebakumar [#1], Dr. P. Vivekanandan [*2]

[#] Research Scholar, Department of Computer Science and Engineering, Anna University
Chennai, India

[1] rrjeba@india.com

[*] Professor and Head, Computer Centre, A.C. Tech, Anna University
Chennai, India

[2] vivek@annauniv.edu

***Abstract*—Data aggregation is a key challenging task in wireless senor network to reduce energy consumption and ensures the data reliability and security. Multi-level data aggregation techniques ensure security and reliability of data aggregation in the presence of secure routing. Every movement of data must be followed by secure routing via the dynamically distributed secure node list to ensure the secure path. Three different levels of aggregation are used to add for more security and so it reduces the time consumption of the network. In the first level, the sensors senses its own signal or it may receive from adjacent sensors and checks the duplication then forward it to the nearest storage node in its region, where the storage nodes are special than the ordinary sensor nodes due to additional computation with high resource capacity. In the second level, the storage node receives from sensors and it senses itself and checks the duplication and also it tries to avoid the very closer data to be transmitted. In the Final level, the base station receives the data from the storage nodes and generates the whole data set for its own region based on session ID. The session and region based data sets are used for data aggregation and to store the sensed data. The session based secure keys are used to protect the data and it is controlled by distributed data server. Due to very sensitive data transmission, two different encryptions are applied in high level data aggregation and shield the data for each and every level by using the digital signature to add more security on data.**

Keyword-Wireless sensor network, Multi-level data aggregation, Storage node, Base station, Distributed data server, Session ID

I. INTRODUCTION

In Wireless sensor networks (WSNs), a large number of tiny sensors are collecting the sensed data and rely on multihop small range radio communication to send data to the base station. WSNs can operate in an event-driven model or regular continuous monitoring model for data collection. Here, a regular continuous monitoring model is chosen, where each sensor will monitor its area and periodically sends the collected sensed data to the base station possibly via the relay of other sensors or storage nodes, where the storage nodes are something special than the ordinary sensors due to additional computation such as data aggregation and shield the data for secure communication and its having more storage capacity.

The key objective of data aggregation is to collect and aggregate data in an energy efficient manner so that network lifetime is enhanced, however, it depends on effective energy saving strategies such as sensor scheduling and information processing to reduce the network utilization. One of the important network processing is data aggregation.

The sensors are usually scattered in a sensor field and each of these sensor nodes has the abilities to collect data and send back to the base station through a multihop infrastructure less architecture as shown in Fig. 1. Some amount of energy is consumed during data transmission so energy conservation is the vital factor in sensor network. Data aggregation is the noble technique to save the precious energy of sensor nodes. In most of the sensor networks, the sensors sense the environment based on its desired application and send back to the base station to combine all the information to produce the desired output for user needs. Here, the number of packets are decreased to transmit in the network and can save the energy of sensor nodes due to data aggregation where packets are combined before reaching the base station.

Data confidentiality prefers to data encryption before transmitting at the source node and decrypted at the destination to processing it. However, data aggregation requires any encryption techniques for data communication during aggregation. In WSN, data aggregation is implemented to eliminate data redundancy, reduce data transmission and improve data accuracy [1]. Data aggregation lead to better bandwidth and battery utilization [2], [3], which improves the network lifetime however, the communication constitutes 70% of the

total energy consumption of the network [4]. Data aggregation with secure routing process has to ensure that data are received from the secure node and the data is to be original.

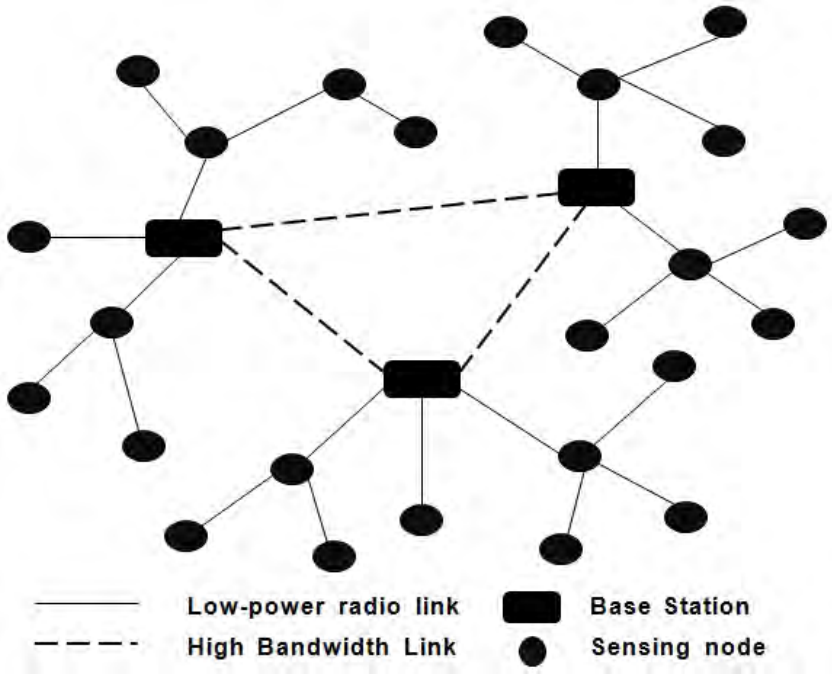

Fig. 1. Sample Architecture of Wireless Sensor Network

II. Structure Based Data Aggregation Techniques

Data aggregation is defined as the process of aggregating the data from multiple sensors to eliminate redundant transmission and provide fused information to the base station [5]. Data aggregation usually involves the fusion of data from several sensors at intermediate nodes and transmission of the aggregated data to the base station. Basically data aggregation can be classified into network topology, network flow, quality of services and many more [6]. In this chapter discussed some structure oriented under network topology based data aggregation techniques.

A. Data Aggregation in Flat Networks

All Sensor nodes are playing the same role in flat networks. The data aggregation is accomplished by data centric routing where the sink generally transmits a query message to the sensors, such as flooding and sensors which have data matching the query send reply messages back to the sink. The choice of a specific communication protocol depends on the particular application at hand [6].

1) Flooding and Gossiping: All Sensor nodes are playing the same role in flat networks. The data aggregation is accomplished by data centric routing where the sink generally transmits a query message to the sensors, such as flooding and sensors which have data matching the query send reply messages back to the sink. The choice of a specific communication protocol depends on the particular application at hand [6].

2) *Directed diffusion:* Directed diffusion [7] is a popular data aggregation paradigm for wireless sensor networks. It is a data-centric and application-aware paradigm. Such a scheme combines the data coming from different sources enroute to the sink by eliminating redundancy and minimizing the number of transmissions. In this way, it saves the energy consumption and increases the network lifetime of WSNs. In directed diffusion, the base station requests data by broadcasting interests, which describes a required task to be implemented by the network. The interest is defined using a list of attribute-value pairs such as name of objects, interval, duration and geographical area. Each node receiving the interest can cache it for later use. As the interest is broadcasted through the network hop-by-hop, gradients are setup to draw data satisfying the query towards the requesting node. A gradient is a reply link to the neighbor from which the interest was received. It contains the information derived from the received interest's fields, such as the data rate, duration and expiration time. Each sensor that receives the interest sets up a gradient toward the sensor nodes from which it received the interest. This process continues until gradients are setup from the sources all the way back to the base station. In this way, several paths can be established, so that one of them is selected by reinforcement. The sink resends the original interest message through the selected path with a smaller interval, hence reinforcing the source node on that path to send data more frequently.

3) SPIN: SPIN [8] is among a data centric routing mechanism. The SPIN is to name the data using high-level descriptors or meta-data. Before the transmission, metadata are exchanged among sensor nodes via a data advertisement mechanism, which is the main feature of SPIN. When each node receiving new data, advertises it to its adjacent nodes and interested adjacent are retrieve the data by sending a request message. SPIN's meta-data negotiation explains the classic problems of flooding such as redundant information, overlapping of sensing areas and resource blindness and achieving a lot of energy efficiency. Here no standard meta-data formats are available and it is assumed to be application dependent.

4) Rumor routing: Rumor routing [9] is alternative variation of Directed Diffusion and is generally intended for contexts in which geographic routing principles are not applicable. It is between event flooding and query flooding and the idea is to route the queries to the nodes that have detected a particular event rather than flooding the whole network to retrieve information about the going on events. In order to flood the events through the network, this routing algorithm employs long-lived packets called by agents. When an event occurs, it reports such event to its local table and generates an agent. Agents travel the network in order to spread information about that event to distant nodes. Hence, the cost of flooding the whole network is to be avoided. Rumor routing maintains only one path between source and destination as it opposed to Directed Diffusion where data to be sent through multiple paths at low rates.

5) Gradient-Based Routing: Gradient-Based Routing [10] is also another version of directed diffusion, which aims to distribute traffic evenly throughout the network to increase the network lifetime. The main idea is to memorize the number of hops when interest is diffused through the whole network. Each node can calculate a parameter called by height of the node, which is the minimum number of hops are required to reach the base station. The variance between a nodes height and that of its neighbor is considered the gradient on that link. Finally the packet is forwarded on a link with the largest gradient.

B. Data aggregation in Hierarchical Networks

A flat network can result in unnecessary communication and computation loads at the sink node, resulting in a faster depletion of its battery power [5]. The loss of the sink node breaks down the functionality of the network. Hence the view of scalability and energy efficiency, some hierarchical data-aggregation approaches have been proposed. Hierarchical data aggregation [5] involves data fusion at special nodes, which decreases the number of messages communicated to the sink. This increases the energy efficiency of the network.

C. Data Aggregation in Cluster-Based Network

In energy-constrained sensor networks of large size, it is inefficient for sensor nodes to transmit the data straight to the sink. In such situations, sensors can transmit data to a local data aggregator or cluster head which aggregates data from the sensors in its cluster and transmits the aggregated data to the sink. This results in significant energy savings for the sensors. The cluster heads can communicate with the sink directly or via other cluster heads.

1) *LEACH:* LEACH [11] is the first clustering protocol and it provides a conception of round. LEACH protocol runs with many rounds and each round contains two states namely, cluster setup state and steady state. In cluster setup state, it forms cluster in self-adaptive mode; in steady state, it transfers data. The second state taken more time than the time of first state for saving the protocol payload.

2) E-LEACH: E-LEACH [12] protocol also divided into rounds, first, every node has the same probability to turn into cluster head, that mean nodes are randomly selected as cluster heads, in second rounds, the residual energy of each node is not the same after one round and taken into account for the selection of the next cluster heads. That mean which nodes have more energy will become cluster heads.

3) *TL-LEACH:* Two-level Leach [13] protocol; cluster head collects information from cluster member and in spite of sending it to directly base station it sends it to another cluster head that lie between the cluster head and base station as a relay station.

4) *M-LEACH:* In M-LEACH [14], multi-hop communication is selected among cluster heads. These cluster heads transmits data to the corresponding cluster head according to the selected optimal path, which is nearest to base station. Finally, this cluster head sends data to base station. It makes communication mode from single hop to multi-hop between cluster heads and base station.

5) *LEACH-C:* LEACH-C [15] protocol can produce well performance by dispersing the cluster heads through the network. In the set-up phase of LEACH-C, each node sends facts about its current location and residual energy level to the sink. In addition to determining good clusters, the cluster head needs to ensure that the energy load is equally distributed among all the nodes.

6) *V-LEACH:* In V-LEACH [16] protocol, besides having a cluster head in the cluster, there is a vice cluster head that takes the role of the cluster head when it dies. Due to this the cluster nodes data will always reach the base station and no need to elect a new cluster head each time even the cluster head dies. This will extend the overall network life time.

D. Chain-Based Data Aggregation

The chain-based data aggregation is that each sensor transmits to its closest neighbour node only. A chain-based data-aggregation protocol namely Power-Efficient Data Gathering Protocol [17] for Sensor Information Systems (PEGASIS).Here the nodes are organized into a linear chain for data aggregation. The sensors can form a chain by employing a greedy algorithm or the cluster head can determine the chain in a centralized manner. In data-gathering round, when a node receives data from one of its neighbour nodes, fuses the data with its own, and transmits it to its other neighbor node along the chain. In greedy chain formation approach, some nodes having relatively distant neighbour nodes along the chain. So it is alleviated by not permitting such nodes to become leaders.

E. Tree-Based Data Aggregation

The main aspect of tree-based networks is to construction of an energy efficient data aggregation tree. An Energy-Aware Data Aggregation Tree (EADAT) algorithm is proposed in [18]. Here the base station broadcast a control message periodically. When receiving this message for the first time, the node will start a timer. The end time is inversely proportional to the nodes remaining energy. The timer is to be refreshed when a node receives this message during the timer count down. In E-span (energy-aware spanning tree algorithm) [19], the source node which has the highest remaining energy is chosen as the root. If there are multiple neighbour nodes with equal distance, the node which has more remaining energy is selected as parent. It considers distance as main parameter and remaining energy as second one. After the aggregation and data transmission, the remaining energy of these nodes is ended quickly. So the node failure and network cannot coverage region completely.

F. Grid-Based Data Aggregation

Grid-Based Data Aggregation [20] based on dividing the region monitored by a sensor network into several grids. Here, a set of sensors is assigned as data aggregators in fixed regions of the sensor network. The sensors in a specific grid transmit the data straight to the data aggregator of that grid. The sensors within a grid do not communicate with each other. In the network aggregation is similar to grid-based data aggregation with two key differences, such as, each sensor within a grid communicates with its neighboring nodes. Any sensor within a grid can assume the role of a data aggregator until the last node dies.

III. Security Issues in Data Aggregation

In data aggregation, Confidentiality and integrity are the key security issues; data confidentiality is to protect the sensitive transmitted data from passive attacks, it is particularly vital in a hostile environment, where the wireless channels is vulnerable to eavesdropping. The complicated encryption and decryption operations can use up the sensors power quickly [21]. Another security issue is data integrity, which avoids the compromised source nodes or aggregator nodes from significantly changing the final aggregation value [22]. Sensor nodes are easy to be compromised due to lack expensive tampering-resistant hardware and even that hardware might not always be reliable. A compromised node can alter, forge or discard messages.

Two different methods can be used for secure data aggregation in WSN [9], first one is hop-by-hop encrypted data aggregation and second one is end-to-end encrypted data aggregation. Earlier, the sensed data is encrypted by the sensors and decrypted by the aggregator nodes. The aggregator's nodes are aggregate the data and encrypt the aggregation result again to transmit it and finally the sink node gets the final encrypted aggregation result and decrypts it for computation. There are two main practical issues involved in implementing data encryption at the sensors [23] namely, the size of the encrypted message and next execution time for encryption at the sensor nodes. Another main aspect of security in sensor networks is secret keys establishment between the sensor and the base station. A security protocol proposed for sensor networks [24] which address the key establishment problem. Here, all nodes are trust the base station at the network creation time and each node is given a master key which is shared with the base station. A message authentication code is used, and the keys for encrypting the data and computing the code are derived from the master key using a pseudo random function. When a key is compromised, a new key is derived without transmitting confidential information.

IV. Network Model

In this network hierarchy Fig. 2, the sensor nodes are placed in the most bottom of the network. After sensing it does some calculation for verifying the sensed data and it will avoid the duplication of data due to the low level aggregation. Then it transmits the data to the storage node or adjacent sensor which one is more secure for data communication where the storage nodes are also sensors but it elected through high resource availability than the ordinary sensors. Storage and sensor nodes are forwarding nodes; it checks the source node for the authentication when the data is received. It won't forward to the packet directly, does the aggregation based on the aggregation level and forward it to the base station or nearest node which one is very nearest to the BS or DDS.

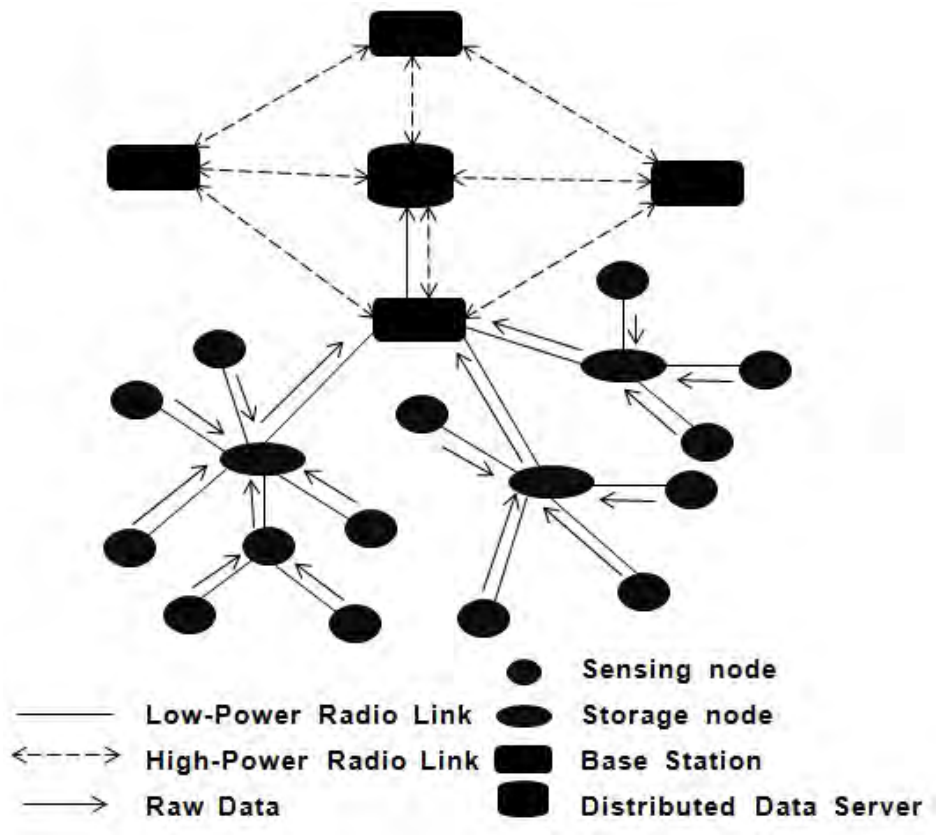

Fig. 2. Schematic of Wireless Sensor Network Architecture

. Base stations contain more storage capacity, computation power and high bandwidth for data transmission. All base stations are directly communicating with the users and the other part of the network through the queries. Here, one or more base stations are connected with a single region and it stores the specific region data itself and forwards to the DDS for storing the data permanently. The DDS is directly connected with all BSs. In this network, the base stations can solve the query without interacting with the DDS if it has the required data. When the base stations receive the sensed data from the sensors or the storage nodes then it does the higher level data aggregation to forward to the DDS. This process always happens for the data collection, aggregation and storing.

V. METHOD

The DDS distributes the keys to all secure nodes when the previous session is expired. In DDS, a database is maintained for key management where the keys are stored and related information such as session ID and key pool and randomly chooses keys based on session ID. The fast RSA algorithm is better to choose for protecting the data due to the computation cost. The digital signature is added after encryption is over on the data for more security on it. A spanning tree is maintained to disseminate the random keys to all secure nodes based on the proposed hierarchy. In the key distribution, all base stations receive the session key from the DDS after generating the random keys and the BSs are flooding the keys to all adjacent secure nodes either storage nodes or sensor nodes in the network. Furthermore, the storage nodes or sensors are flooding the keys to all secure nodes based on the network hierarchy until all nodes are received.

In this proposed model, the aggregation is done by three different levels as it shown in Fig. 3. The sensors collect their own readings and check duplication of reading in the same session. If duplication is occurred it will automatically reject and simultaneously it is forward to the storage node within its same region in level 1. In the next level (level 2), the storage nodes check redundancy of data and occurrences of closer data set based on the similar session under the same region. Finally, in level 3, the base station should avoid data redundancy, closer data set and bind it based on session ID with its region.

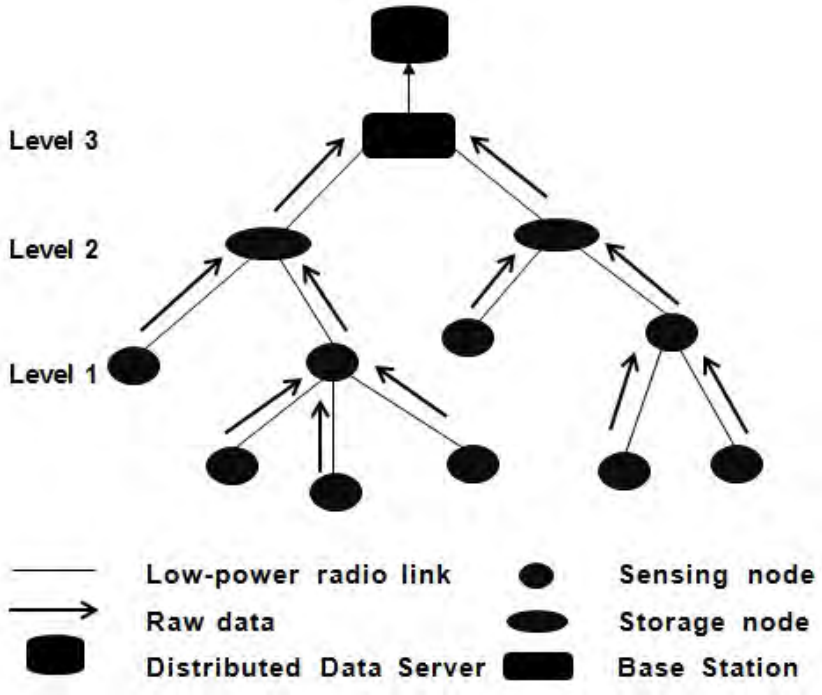

Fig. 3. Network Hierarchy for Data Aggregation

TABLE I
Symbols Used in Algorithms

Symbol	Description
SID	Session ID
RD_{iSID}	Sensor Reading from i^{th} node and a part from SID
PID	Previous Session ID
TDS	Temporary Data Set
OD_{SID}	Output Data a part from SID
EC_{SID}	Encryption Key for SID
SG_j	Signature for j^{th} node
STr	Storage Node for r^{th} region
S	Source Node
SN	Secure Node list
MDP	Misbehaviour Detection Process
DS_{rSID}	Data Set for r^{th} region a part from SID
ODS_{rSID}	Output Data Set for r^{th} region a part from SID
ODS_{rPID}	Output Data Set for r^{th} region a part from PID
EC_{PID}	Encryption Key for PID
DS_{rPID}	Data Set for r^{th} region a part from PID
FD	Fusion Data set
BS_r	Base Station for r^{th} region
DC_{PID}	Decryption Key for PID
DC_{SID}	Decryption Key for SID
RN_r	r^{th} Region Node List
δ	Threshold value for Variation of Sensor reading

All sensor nodes having its own sensors data where the data source(S) is NULL and it may receive the data from some of its adjacent sensors where the S is not NULL. The received data (RD_{iSID}) will be rejected if the data is received from not a secure node or the data is already available in the temporary data set (TDS) for the same session ID (SID) with its region (r). If the above case is failed then the sensor sends the data to the nearest storage node or sometimes to the base station too, simultaneously it will add the data (RD_{iSID}) to the temporary data set (TDS) ($TDS \cup RD_{iSID}$) to avoid duplication in same node at the specified session only. If RD_{iSID} is not a duplicate one (not $RD_{iSID} \in TDS$), then it sends RD_{iSID} to the storage node or via its adjacent sensors. The sensors encrypt the data$(EC_{SID}(RD_{iSID}))$, add session ID $(EC_{SID}(RD_{iSID}) + SID)$ and apply signature on it $((EC_{SID}(RD_{iSID}) + SID) + SG_j)$ for data reliability before sending it. The algorithm was discussed in the Aggregation_Level1 is shown below.

Algorithm Aggregation_Level1 (S, RD_{iSID})

1. If (IsSecure (S))
2. If (Not($SID = PID$))
3. $TDS = NULL$
4. End If
5. If (Not ($S = NULL$))
6. $RD_{iSID} = DC_{SID}((RD_{iSID} - SG_i) - SID)$
7. End if
8. If ((RD_{iSID} is not valid) or ($RD_{iSID} \in TDS$)) then
9. Drop RD_{iSID}
10. Else
11. $OD_{SID} = (EC_{SID}(RD_{iSID}) + SID) + SG_j$
12. $TDS = TDS \cup OD_{SID}$
13. $PID = SID$
14. Send (OD_{SID}, ST_r)
15. End If

16.End If

17.End Algorithm

Whenever the receiver node receives its children readings or any request from the source node S, it computes to identify whether it processes the request or not by IsSecure algorithm. At first, it checks the data is receiving from its own($S = NULL$). When the above condition is failed (S is not NULL), it checks the request that is coming from secure node(S) or not via the distributed secure node list (SN). If S is not a member of SN (not $S \in SN$) then the receiver node will send a request to misbehaviour detection process (MDP) to verify the node for secure communication or not. If S is a member of SN ($S \in SN$) then it returns true for process the request and returns false for discard the given request when it fails.

Algorithm IsSecure (S)

1. If (Not($S = NULL$))
2. If (Not($S \in SN$))
3. Send Request to MDP
4. Return (False)
5. End If
6. End If
7. Return (True)
8. End Algorithm

During the aggregation process for the middle and above (level 2 and level 3), it checks (in algorithm Get_FD) not only the data duplication alone but also considered the closer data set. Here a threshold value (δ) is used for measuring the very closer data set which belongs to the received the data (RD_{iSID}) under the same session (SID) with its own region(r). To get the very closer data set (FD), it checks the existing data available in the same data set (DS_{rSID}) with the data range between ($RD_{iSID} + \delta$) and ($RD_{iSID} - \delta$). If FD is empty, it means no data falls under the above range and returns NULL to add the input data as a member in DS_{rSID}, if it is not, the FD contains one or two data and simply returned it (FD) to update the DS_{rSID}.

Algorithm Get_FD (RD_{iSID}, DS_{rSID})

1. $\exists FD \subseteq DS_{rSID}$: Between ($RD_{iSID} + \delta, RD_{iSID} - \delta$)
2. If (IsEmpty (FD))
3. Return (NULL)
4. Else
5. Return (FD)
6. End If
7. End Algorithm

In level 2, the complete process defined in algorithm Aggregation_Level2. Whenever the storage node receives its children readings from S or its own($S = NULL$), first, it checks the request that is coming from secure node(S) or not using the IsSecure algorithm. If S is a member of SN ($S \in SN$) then its checks the data set for the present session ID (DS_{rSID}), if DS_{rSID} is not presently available, then it creates a new data set (DS_{rSID})

to store the new incoming data under the new session ID and simultaneously it sends the previous session data set (DS_{rPID}) to its base station. Before sending the previous data set, it encrypts ($EC_{PID}(DS_{rPID})$), adds the session ID ($EC_{PID}(DS_{rPID}) + PID$) and puts the digital signature on it ($(EC_{PID}(DS_{rPID}) + PID) + SG_j$). During the data aggregation in storage node, first it checks and removes the signature for the received data ($RD_{iSID} - SG_i$) and session ID ($(RD_{iSID} - SG_i) - SID$). Finally, it decrypts the data ($DC_{SID}((RD_{iSID} - SG_i) - SID)$) and gets the original data set (RD_{iSID}) if it is received from other secure nodes (Not($S = NULL$)). The remaining part of the algorithm (Aggregation_Level2) defines the aggregation process of the received data or its own data. During the data aggregation, it checks the data redundancy ($RD_{iSID} \in DS_{rSID}$) and drops RD_{iSID} if the duplication is occurred. Else, using the Get_FD function, it gets the closer data set FD. If FD is NULL (closer data set is not available) then it simply adds the RD_{iSID} to the DS_{rSID} ($DS_{rSID} \cup RD_{iSID}$). When the closer data set is available (FD is Not NULL), first it computes the average of RD_{iSID} and FD, then adds the result data to the DS_{rSID} ($DS_{rSID} \cup FD$) and removes the closer data set from the DS_{rSID}($DS_{rSID} - FD$).

Algorithm Aggregation_Level2 (S, RD_{iSID})
1. If (IsSecure (S))
2. If (IsNotExist (DS_{rSID}))
3. Make New DS_{rSID}
4. $ODS_{rPID} = ((EC_{PID}(DS_{rPID}) + PID) + SG_j)$
5. Send (ODS_{rPID}, BS_r)
6. Remove ODS_{rPID}
7. End If
8. If (Not ($S = NULL$))
9. $RD_{iSID} = (DC_{SID}((RD_{iSID} - SG_i) - SID))$
10. End if
11. If ($RD_{iSID} \in DS_{rSID}$)
12. Drop RD_{iSID}
13. Else
14. $FD = Get_{FD}(RD_{iSID}, DS_{rSID})$
15. If (FD is Not NULL)
16. $DS_{rSID} = DS_{rSID} - FD$
17. $FD = Average(RD_{iSID}\ and\ FD)$
18. $DS_{rSID} = DS_{rSID} \cup FD$
19. Else
20. $DS_{rSID} = DS_{rSID} \cup RD_{iSID}$
21. End If
22. End If
23.End If
24.End Algorithm

In higher level data aggregation (Algorithm Agrgation_Level3), the base stations combine all sensors data received from storage nodes or from sensor nodes from its own region ($i \in RN_r$) when the source node (S) is secure. If not ($i \in RN_r$) it redirects the data (RD_{iSID}) to its own base station where the node i belongs to. In base station, the aggregation data stored in a data set (DS_{rSID}) based on session ID with region wise. The base station creates a new data set (DS_{rSID}) for r^{th} region when a new session (SID) is started and simultaneously it forwards the previous data set (ODS_{rPID}) to the DDS. Before forwarding the data (DS_{rPID}), it encrypts the data ($EC_{PID}(DS_{rPID})$), adds the session ID of the previous output ($EC_{PID}(DS_{rPID}) + PID$), again goes for encryption with present session ID ($EC_{SID}(EC_{PID}(DS_{rPID}) + PID)$) and adds the digital signature for its high level data aggregation ($EC_{SID}(EC_{PID}(DS_{rPID}) + PID) + SG_j$). The high level data aggregation is only done by the base station with DDS. Here, all the data set is identified by its region based session IDs. The data set (DS_{rPID}) is forwarded to the DDS and the same data set is stored itself to process the user request. More security is added on data by applying encryption two times using two different keys by comparing with the previous level.

Algorithm Aggregation_Level3 (S, RD_{iSID})
1. If (IsSecure (S))
2. If (Not($i \in RN_r$))
3. ReDirect RD_{iSID} to BS where $i \in RN_r$
4. Exit
5. End If
6. If (IsNotExist (DS_{rSID}))

7. Make New DS_{rSID}
8. $ODS_{rPID} = EC_{SID}(EC_{PID}(DS_{rPID}) + PID) + SG_j$
9 Send (ODS_{rPID}, DDS)
10. End If
11. $RD_{iSID} = DC_{SID}((RD_{iSID} - SG_j) - SID)$
12. For all RD_{iSID} do
13. If ($RD_{iSID} \in DS_{rSID}$) then
14. Continue For
15. Else
16. $FD = Get\_FD\ (RD_{iSID}, DS_{rSID})$
17. If (FD is Not NULL) then
18. $DS_{rSID} = DS_{rSID} - FD$
19. $FD = Average\ (RD_{iSID} + FD)$
20. $DS_{rSID} = DS_{rSID} \cup FD$
21. Else
22. $DS_{rSID} = DS_{rSID} \cup RD_{iSID}$
23. End If
24. End If
25. End For
26.End If
27.End Algorithm

The DDS is the final destination of aggregated data; the authorized users can retrieve any data at any time from different region and various sessions too. When the DDS receives the aggregated data (along with the specified region and session ID) from any one of the base station, it checks the source node (base station) to get the original data that to be stored permanently in the DDS for future purpose too. No aggregation is required here on data but the DDS cares the storage structure based on the indexes for the easiest way to access the data. Before storing the received data, it removes the signature ($RD_{iSID} - SG_j$), decrypts it by DC_{SID} ($DC_{SID}(RD_{iSID} - SG_i)$), removes previous session ID ($DC_{SID}(RD_{iSID} - SG_i) - PID$) and finally decrypts with the help of DC_{PID} ($DC_{PID}(DC_{SID}(RD_{iSID} - SG_i) - PID)$).

VI. Counter Measurement

In the proposed system, the aggregation process is mainly based on the classification of nodes such as sensors, storage node and base station. Three different levels of aggregations are introduced for reliable data aggregation based on its classification. In each level does the different process to achieve the reliable transaction with the minimum network utilization and security too. All the three levels of data aggregation avoid the duplication of data based on session ID with its region. In the middle and above level, it avoids the duplication and does the aggregation with the closer data set for the given received data, where the data set is suitable for the new data. In aggregation, it finds the average of closer data set with the new data. Finally the data set is updated by the above average value and removed the closer data set. Data reassembling process is done as the part of the second and third levels based on its session ID and its region. In addition, except the first two levels (first and second), all the data must be gathered from storage nodes or sensors and verified by the base station and stored it after the higher level aggregation based on the above criteria in the final level.

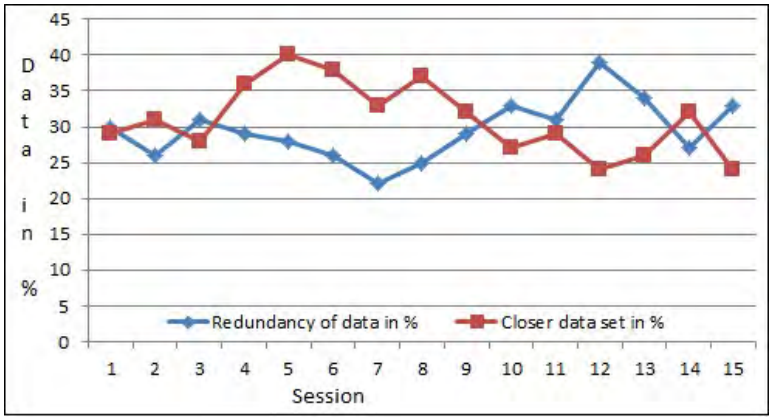

Fig. 4. Aggregation output for 15 continues sessions

Here, we had taken fifteen continues sessions to shown the experimental output. The ratio of closer data set is higher than the redundancy of data most probably, it shown in Fig 4. The value of δ is 0.005 to consider for the process of closer data set. The process of aggregation is basically two parts and the total ratio of each part is shown in Fig 4. The closer data set value is fully depending on the value of δ. The δ value can be changed if the application is ready to accept without affect the originality.

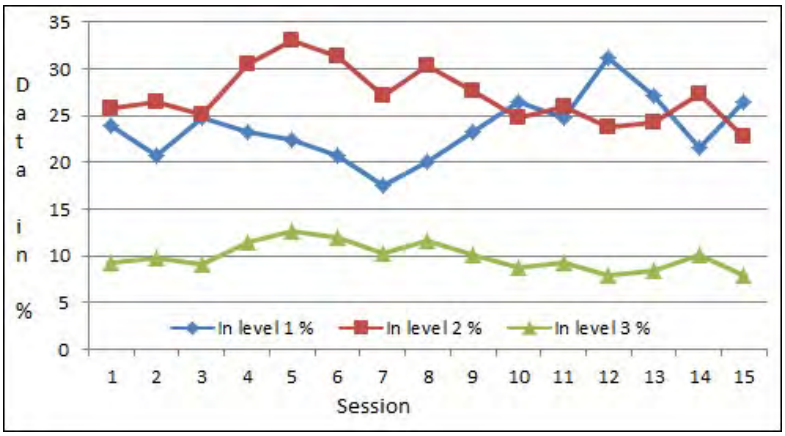

Fig. 5. Aggregation output of each level for 15 continues sessions

The aggregation process of all three different levels is shown in Fig 5. The redundancy ratio of data is smaller than the closer data set and it vary for each level. The aggregation process in level 3 is minimized due to place the storage nodes and it is partially taken care by level 2. If more storage nodes are placed, it will decrease the performance and the aggregation process of level 3 is automatically increased. In some special case, if only one storage node is placed among the sensors in each specific region, totally it avoids the aggregation of closer data set in base station. The Fig 6 had shown how the aggregation process done in level 1 and 2 but not in level 3 because of only one storage node is placed between the base station and the sensors in all regions. The storage nodes required more computation but no more aggregation process at base station because the storage node is a single intermediate node between the sensors and the base station in the above case. If sensors are increased the storage node should be in heavy load, due to this reason add some storage nodes based on the number of sensors and the number of request received at a particular session from sensors in the specified region.

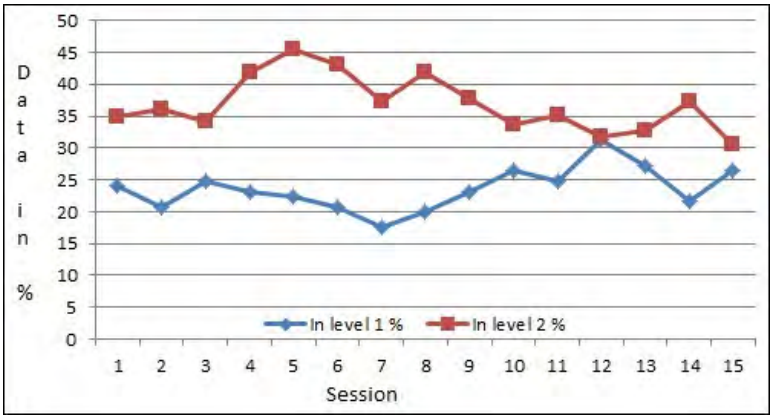

Fig. 6. Aggregation output of each level for 15 continues sessions with single storage node for each region

The proposed system decreases the number of data (avoids the unnecessary data to be send) to be transmitted as much as possible and minimize the computation time too. In addition to this, the reliability of data will be guaranteed with the support of secure routing and secure the data by using the encryption with digital signature. The adversary nodes can't take the data easily because of all data sharing done by the secure nodes only. Due to this, the participation of adversary node will be banned by the help of MDP process that is not discussed in this paper.

Each data appended with session ID and digital signature for secure routing to be reliable data communication. In addition, all receiver nodes check the source node to accept their request and the request is rejected when it receives data from unrecognized node before processing their request. The computation is increased level by level due to their node capacity and avoided more process in the smaller capacity nodes, so the process stability can be maintained by the proposed method.

The single encryption applied on level 1 and level 2 and final level do the double encryption for data security so the adversary node can't access the message easily even it receives the data. The digital signature also used for identity of source node and it protects the data. The signature is changed periodically to shield the data based on the application if it is required. The reliability of data is increased by the proposed method and time taken of each level is varying level by level. The higher level takes more time than the lower one, so the nodes are categorized into different levels and maintains to achieve the good performance of the network.

VII. CONCLUSION

Data aggregation is an energy consuming operation in WSN. In the proposed system, the aggregation is done through the hierarchical based network structure. In the network hierarchy, the storage nodes are placed between the sensors and the base station to increase the data aggregation performance. The data aggregation process is not same for all the nodes which are connected in the network, it vary based on their performance so the resource availability of sensors was considered in this situation. Single encryption applied when the data are transmitted through the low-power radio signal, when it is traveled through the high-power radio link added one more encryption with two different keys for more security. Here the reliability is more due to secure routing via the distributed secure node list. The computation was increased on sensed data for the aggregation based on its levels one by one. The session based secure keys are used to data encryption to protect the data to more secure, and also the digital signature is added for reliability and secure data communication. The session interval, key size and threshold value to find the closer data set chose based on the requirements of application.

REFERENCES

[1] Suat Ozdemir and Hasan Çam, "Integration of False Data Detection With Data Aggregation and Confidential Transmission in Wireless Sensor Networks," *IEEE/ACM Transactions on Networking*, Vol. 18, NO. 3, pp. 736-749, JUNE 2010.

[2] C. Intanagonwiwat,D. Estrin, R.Govindan, and J.Heidemann, "Impact of network density on data aggregation in wireless sensor networks," in *Proc. 22nd Int. Conf. Distrib. Comput. Syst*., Jul. 2002, pp. 575–578.

[3] R. Rajagopalan and P. K. Varshney, "Data aggregation techniques in sensor networks: A survey," *IEEE Commun. Surveys Tutorials*, vol. 8, no. 4, 4th Quarter 2006.

[4] A. Perrig, R. Szewczyk, D. Tygar, V. Wen, and D. Culler, "SPINS: Security protocols for sensor networks," *Wireless Netw. J*., vol. 8, pp. 521–534, Sep. 2002.

[5] Ramesh Rajagopalan and Pramod K. Varsney, "Connectivity analysis of wireless sensor networks with regular topologies in the presence of channel fading ," *IEEE Transactions on Wireless Communications*, Volume 8, Issue 7, pp. 3475-3483, july 2009.

[6] Vaibhav Pandey, Amarjeet Kaur and Narottam Chand, "A review on data aggregation techniques in wireless sensor network," *Journal of Electronic and Electrical Engineering*, Vol. 1, Issue 2, pp-01-08, 2010.

[7] Chalermek Intanagonwiwat, Ramesh Govindan, Deborah Estrin, John Heidemann, and Fabio Silva, "Directed Diffusion for Wireless Sensor Networking," *IEEE/ACM Transactions on Networking*, Volume 11 Issue 1, pp. 2-16, February 2003.

[8] Heinzelman W, Kulik J, Balakrishnan H, "Adaptive protocols for information dissemination for WSNs," *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking* , Seattle, WA, pp. 174 – 185, August 1999.

[9] Braginsky D, Estrin D, "Rumor routing algorthim for sensor networks," *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, Atlanta, GA, pp. 22-31, 2002.

[10] Schurgers C and Srivastava M.B, "Energy Efficient Routing in Wireless Sensor Networks," *In MILCOM Proceedings on Communications for Network-Centric Operations*, Creating the Information Force, McLean, VA, 2001.

[11] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan, "Energy-Efficient Communication Protocol forWireless Microsensor Networks," *Proceedings of 33rd Hawaii International Conference on System Sciences*, 2000.

[12] Fan Xiangning, Song Yulin, "Improvement on LEACH Protocol of Wireless Sensor Network," *Proceedings of the 2007 International Conference on Sensor Technologies and Applications*, pp. 260-264, 2007.

[13] Loscri V, Morabito G, Marano S, "A two-levels hierarchy for low-energy adaptive clustering hierarchy (TL-LEACH)," *in Proceedings of the 62nd IEEE Vehicular Technology Conference (VTC '05)*, pp. 1809-1813, September 2005.

[14] Yuhua Liu, Yongfeng Zhao, Jingju Gao, "A New Clustering Mechanism Based on LEACH Protocol," *Proceedings of International Joint Conference on Artificial Intelligence*, pp. 715-718, 2009.

[15] Wendi B. Heinzelman, Anantha P. Chandrakasan and Hari Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor networks", *IEEE Transactions on Wireless Communications*, Vol. 1, No. 4, 2002.

[16] Bani Yassein M, Al-zou'bi A, Khamayseh Y, Mardini W, "Improvement on LEACH Protocol of Wireless Sensor Network (VLEACH)," *International Journal of Digital Content Technology and its Applications*, Vol. 3, No. 2, pp. 132-136, 2009.

[17] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data Gathering Algorithms in Sensor Networks Using Energy Metrics," *IEEE Trans. Parallel and Distributed Systems*, vol. 13, no. 9, pp. 924-935, Sep. 2002.

[18] Min Ding, Xiuzhen Cheng, Guoliang Xue, "In-network processing for wireless sensor networks with multiple sinks and sources," *Proceeding of 3rd international conference on Mobile technology, applications & systems*, Article No. 53, 2006.

[19] Marc Lee, Vincent W.S. Wong," E-Span and LPT for data aggregation in wireless sensor networks," *Computer Communications*, Volume 29, Issue 13, pp. 2506-2520, August, 2006.

[20] K. Vaidhyanathan et al., "Data Aggregation Techniques in Sensor Networks," *Technical Report, OSU-CISRC-11/04- TR60*, Ohio State University, 2004.

[21] A. Perrig, R. Szewczyk, V. Wen, D. Cullar, and J. D. Tygar, "Spins: Security protocols for sensor networks," *In Proceedings of the 7th Annual ACM/IEEE Internation Conference on Mobile Computing and Networking* , Rome, Italy, pp. 189–199, July 2001.

[22] Hu, L., and D. Evans, "Secure Aggregation for Wireless Networks," *Workshop on Security and Assurance in Ad hoc Networks*, January 2003.

[23] E. J. Duarte-Melo and M. Liu, "Data-gathering wireless sensor networks: organization and capacity," *Computer Networks (COMNET), Special Issue on Wireless Sensor Networks*, Vol 43, Issue 4, November 2003.

[24] J. Girao, D. Westhoff, and M. Schneider: CDA: Concealed Data Aggregation for Reverse Multicast Traffic in Wireless Sensor Networks, *40th International Conference on Communications, IEEE ICC 2005*, Seoul, Korea, May 2005.