International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:10, No:3, 2016

Secure and Efficient Transmission of Aggregated Data for Mobile Wireless Sensor Networks

A. Krishna Veni, R.Geetha

***Abstract*—**Wireless Sensor Networks (WSNs) are suitable for many scenarios in the real world. The retrieval of data is made efficient by the data aggregation techniques. Many techniques for the data aggregation are offered and most of the existing schemes are not energy efficient and secure. However, the existing techniques use the traditional clustering approach where there is a delay during the packet transmission since there is no proper scheduling. The presented system uses the Velocity Energy-efficient and Link-aware Cluster-Tree (VELCT) scheme in which there is a Data Collection Tree (DCT) which improves the lifetime of the network. The VELCT scheme and the construction of DCT reduce the delay and traffic. The network lifetime can be increased by avoiding the frequent change in cluster topology. Secure and Efficient Transmission of Aggregated data (SETA) improves the security of the data transmission via the trust value of the nodes prior the aggregation of data. Since SETA considers the data only from the trustworthy nodes for aggregation, it is more secure in transmitting the data thereby improving the accuracy of aggregated data.

***Keywords*—**Aggregation, lifetime, network security, wireless sensor network.

I. Introduction

WIRELESS Sensor Networks (WSNs) are utilized by the individual where the monitoring of physical and environmental conditions becomes impossible. WSNs are used in the applications like wildlife monitoring, forest fire detection, meteorology, health care monitoring, under sea navigation, military surveillance. The network is constructed by building nodes from few hundred to several thousand. Each node is self-configurable where they can be configured automatically when they are deployed [1].

The nodes are deployed in the region where the monitoring is to be done. The deployment of nodes can be of two types, they are random deployment where the nodes can be dropped from an aircraft and regular deployment in which the nodes are well planned and fixed manually.

The main problem in WSNs is the energy of the nodes and maximizing the network's lifetime. The network's lifetime should be increased without any data loss during the transmission.

To increase the energy of the nodes, sleep scheduling can be done; but, in order to transmit the data in the shortest path to the destination, many protocols are found and by imposing those onto the nodes can drain the energy. So the data aggregation and routing protocols can be used. Routing protocols [10] are based on clustering mechanisms are used in order to efficiently collect the data based on the node's location. The topology of the network is found and then the data can be sent to the destination [2].

The topology management plays a crucial role in reducing the constraints such as limited energy, node failure, and long range communication within the network. The data collection schemes are used to collect the data and transmit it to the sink node in the network. A few of the data collection schemes, according to the topologies are multipath, chain, tree, cluster and hybrid [11].

The rest of the paper is organized as follows. Related work is discussed in Chapter II. Proposed Work, SETA is discussed in Chapter III. Chapter IV discusses the simulation results followed by the conclusion in Chapter V.

II. Related Work

In [3], the authors propose an alternative clustering scheme with the Self-configurable clustering mechanism which uses the Backup Cluster Head (BCH) to lessen the failure of the CH in the network and thereby saving the energy of the nodes in the network.

In [4], the authors propose a data collection algorithm such as Tree-Cluster-Based Data-Gathering Algorithm (TCBDGA) which uses the weight based tree construction that efficiently gathers the data. The TCBDGA algorithm reduces the energy consumption of the nodes in the network and can balance the whole network.

In [5], the authors propose an efficient clustering scheme even for the isolated nodes. The clustering scheme is done even with the isolated nodes and using the protocol Regional Energy Aware Clustering with Isolated Nodes (REAC-IN). The REAC-IN protocol is very useful in combining the isolated nodes among the cluster, within the network.

In [6], the authors have explained the variety of data aggregation methods and the security in it. It is designed to reduce the communication overhead of each sensor node. Some of the privacy features such as the encrypted and the unencrypted protocols are followed.

In [7], the authors have explained the optimization of the data in the clustered network. The clustering mechanism is used to alleviate the energy consumption of the nodes. The joint optimization is proposed in order to protect the CHs. It is applicable to the large-scale sensor network where the distributed clustering scheme is possible.

In [8], the authors have proposed a scheme for aggregation of data i.e., compressed data aggregation. Compressed data

R.Geetha is Associate Professor with S.A. Engineering College and Research Scholar with Vel Tech Dr. RR & Dr. SR Technical University, Chennai, India (Phone:+919841998698,email:geetha@saec.ac.in)

A. Krishnaveni is PG Scholar with Computer Science and Engineering department of S.A. Engineering College, Chennai, India

aggregation in which the data that is collected from the nodes are compressed in order to achieve the recovery fidelity in the sink. It can be applied in order to recover the data that is lost during the aggregation of the data, and at the receiving end, it must be decompressed.

In [9], the authors propose a scheme for scheduling and target coverage. In data collection, it needs to schedule the coverage area and the aggregates the data and it is sent to the sink node. It makes use of the Coverage and Data Collection Tree (CDCT) in transmitting the sensed information from the sensor nodes to the root node and finally to the sink node. It also concentrates on the coverage area of the sensor nodes in which the sensors can gather the information. It provides the maximum lifetime of the network.

In [12], the authors have proposed a way to send the data in a secure manner to the sink node. Here, the trustworthiness of the node is estimated and then based on that value, the data is sent to the sink node by the neighboring nodes in the network.

In [13], the authors have provided Trust-Aware Routing Framework which eliminates the attacker node from the multi-hop routing. It uses the trust values of the nodes in the network, through which the data can be sent to the sink node in a secure manner.

In [14], the authors have proposed Trust and Energy-aware Routing Protocol (TERP) in order to identify the malicious and trustworthy node in the network. It is based on the trust value of the neighbor node which determines whether to transmit the data to the sink node. If the trustworthiness of the node is greater, than that neighbor node is taken into account for the transmission.

The above schemes are not successful in providing a reliable network in mobility, traffic delay and accuracy of aggregation. VELCT scheme is proposed for the data collection in order to mitigate the problems of coverage, delay, traffic, end-to-end connection, etc., and the SETA is proposed for secure transmission of data.

III. Proposed Work (SETA)

The proposed work consists of the following divisions such as the A. Network Construction, B. Cluster Formation, C. DCT Formation, D. Data Aggregation, E. CH Re-election, F. Secure Aggregation and G. Data Transmission.

The system architecture is given in Fig. 1 with the collection of nodes which form the network. Once the network is constructed, the cluster formation is started with the selection of the CH. All the clusters are formed throughout the entire network and then the tree construction is initiated. The DCT is constructed in order to find the shortest path to the sink node by using the algorithm given in Section III *C*. The Data Collection Node (DCN) in the DCT is identified to transmit the data to the Sink Node. The DCN node just collects the data from the CH and forwards it to the sink. Inter and intra-cluster communication is permitted among the clusters, in order to identify whether the same data is aggregated or not. If there is same data, then it will be ignored.

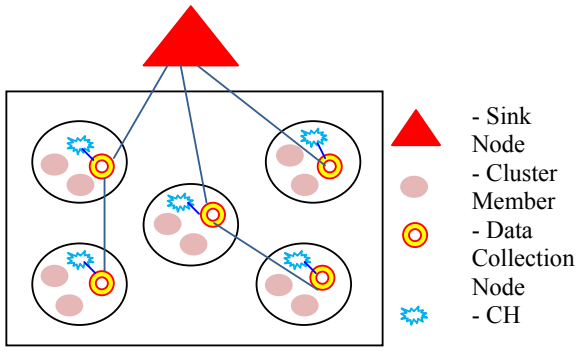

Fig. 1 System Architecture

A. Network Construction

The nodes are distributed densely in the monitoring area in order to form the network. Once the nodes are in a dense manner, the network is formed in that region. A network is constructed by the nodes which are deployed densely.

B. Cluster Formation

The entire sensor node selects the CH based on the threshold value, the node with maximum energy. After the election of the CH, the cluster is formed which gives the intra-cluster communication. The data from the Cluster Members (CM) are transmitted to the CH in the entire network. DCT formation is initiated after the cluster is formed in the network.

C. DCT Formation

The CM sends the aggregated data to the CH. The CH sends the data to the DCN from which it is sent to the sink node. DCT algorithm is given in Algorithm. 1.

The DCT Algorithm consists of the following process, (i.e.,) A, B, C. The DCN is generated by selecting the Neighbor Node (NN) from the Sink Node, which is the process A. The Process B states that from the previously generated DCN, the next DCN is formed. The process C is continuing either process A or B over the entire network. If the Hop Count (HC) =1 and the NN! = CH then that node's identity is stored in the TIN array. If the array contains the Best value of CNI then that node is selected as DCN or else it is just a CM. The Process is repeated for the entire network.

Algorithm. 1. DCT Algorithm

1. Start.
2. Deploy sensor nodes.
3. Elect CH over the entire network.
4. From sink elect 1 hop distance NN to generate DCN (A).
5. From previous DCN generate other DCN (B).
6. Continue DCT for the entire network (C).
7. If HC=1 && NN! =CH->Assign CNI to TIN Array else Go Back.
8. Tin [] > best->select that nodes as DCN else CM.
9. If FD->Max, Construct using B or else with A.
10. Repeat C until all DCT is Constructed.
11. Stop.

D.Data Aggregation

Once the DCT is constructed, the data aggregation process is initiated. Data aggregation is the method of summing up of the data from the members of the cluster by the CH. All the CM sends the sensed data to the CH. The CH collects the aggregated data from its members. The CH then sends the data to the sink node by the DCT through the DCN node. The DCT Communication is initiated which uses the Direct Sequence Spread Spectrum (DSSS) to transfer the data from CH to DCN and then to Sink. DCN collects the aggregated data from the CH or other DCN. The DCN holds the aggregated data and it sends to the sink node. The data is sent through the DCT tree in which the best path is chosen for transmission of data. The above data aggregation process is repeated until all the clusters send the data to the sink node through the DCN node in the DCT tree. The aggregation of data by the construction of the tree identifies the optimal path in the network between the CH and the sink node. The Data aggregation is given for one cycle. After the completion of the first cycle, the re-election of CH is done in Section III *E*.

E. CH Re-Election

Once the data transmission is completed, the CH's energy is checked. If the energy of the CH is in a sustainable manner, then the next set of transmission takes place, otherwise re-election of CH is to be done. The CH is re-elected after the transmission of data through the tree that is constructed. The re-election is based on the energy of the CH after the transmission of data. Once the energy of the CH is drained, it cannot send its cluster information to the sink node. So, the re-election of the CH is used.

F. Secure Aggregation

After the re-election of CH, the CMs are ready to transmit the data to the sink node. The data that is being sent by the CM must be secure that the attacker should not receive the data. So, the proposed method SETA uses the trust manager which manages the trust values for all the nodes in the network. It enables the node to identify the trust value of the neighbor node; if it has high trust, then the node sends the data through that node. Else, the node selects another path for the transmission of data using the trust values. Then the algorithm for the transmission of the data using the trust is given in Algorithm 2.

Algorithm. 2. Trust Algorithm for Data Transmission

1. Start.
2. Node N identifies NN.
3. Finds the trust value of NN.
4. If (Trust Value > Max) then
5. Aggregation of data.
6. Else
7. Choose another NN with Max Trust Value.
8. Stop.

The node N identifies its Neighbor Node (NN) and it finds the trust value of the NN. If it has maximum trust value, then the data is sent through that NN or else the node N chooses another NN with Maximum trust value.

G.Secure Data Transmission

After the secure aggregation of data, the data are to be transmitted to the sink node. The transmission of the data is followed by the DCT algorithm which chooses the DCN. The node collects the trust values of Neighbor Node, using which the data is securely transmitted to the sink node. The SETA method uses the trust value approach and by identifying the trust value of the nodes, transmits the data using the DCT algorithm with the DCN thereby transmitting data successfully to the sink node.

IV. Simulation Results

The simulation result is given using the Network Simulator (NS-2) by the scenario taken and then the performance is given by the Throughput, Accuracy and Delay among the nodes in the network. The simulation of the SETA identifies the malicious node in the network and avoids the packet transmission through that particular malicious node. The simulation is taken into account by the workspace containing 900 x 600 m^2. The speed is given by 0 m/s to 140 m/s. The nodes are deployed within the workspace and the network is created. Then the cluster is formed with the CH and then the tree is constructed. The data transmitted by the nodes is 1Mb. The analysis is made between the Cluster Independent Data collection Tree (CIDT) and SETA. The analysis is carried out with the parameters Accuracy, Throughput, Delay. The analysis is carried out in the presence and absence of malicious nodes, which are given in Figs. 2-7. The tables for Accuracy, Delay and Throughput parameters are given in Tables I-III.

Table I consists of the Accuracy of the data that is aggregated during the transmission. The SETA's accuracy of the data with the malicious node is reduced when compared to the SETA's accuracy of data without the malicious node. In the absence of malicious node, the accuracy of data is higher.

TABLE I
COMPARISON OF SETA AND CIDT WITH ACCURACY

No. Of Nodes	Accuracy with Malicious Node		Accuracy without Malicious Node	
	SETA	CIDT	SETA	CIDT
20	0.886	0.876	0.903	0.88
40	0.865	0.852	0.89	0.86
60	0.843	0.831	0.879	0.843
80	0.829	0.816	0.869	0.826
100	0.809	0.796	0.855	0.796

The accuracy of the data is given in Fig. 2 where the SETA works well than the CIDT even in the presence of malicious nodes. The SETA method is providing security to the nodes by the trust value. It is providing a better accuracy to the nodes than the CIDT method.

The accuracy of the data of the SETA method without the malicious nodes is better than the CIDT method as there are

no malicious nodes in the network; the accuracy of data is more which is specified in Fig. 3.

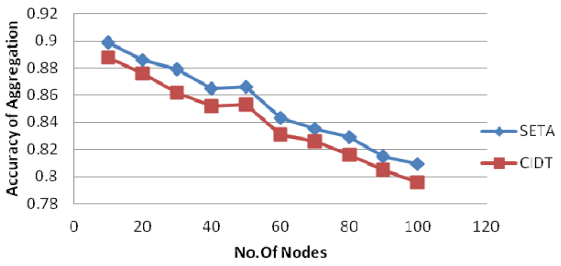

Fig. 2 Accuracy of Aggregation with Malicious Node

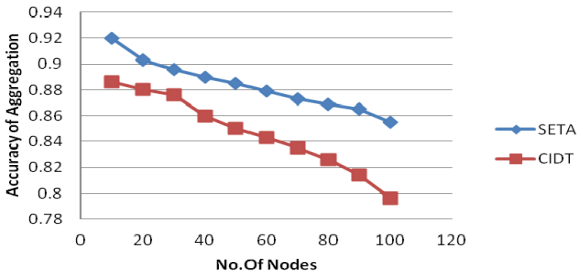

Fig. 3 Accuracy of Aggregation without Malicious Node

Table II consists of the delay in the transmission of data in the presence and absence of malicious nodes. In the presence of malicious node, there is a delay when compared to the delay without malicious node since the trust mechanism is followed. In the absence of malicious node, the delay is comparatively less.

TABLE II
COMPARISON OF SETA AND CIDT WITH DELAY

No. Of Nodes	Delay with Malicious Node		Delay without Malicious Node	
	SETA	CIDT	SETA	CIDT
20	0.25	0.36	0.16	0.25
40	0.45	0.54	0.35	0.46
60	0.64	0.84	0.58	0.69
80	0.84	1.0	0.78	0.86
100	1.0	1.29	0.95	1.09

The delay in the transmission of the data by using the SETA method is less when compared to the CIDT method in the presence of Malicious Nodes. The SETA method works better than the CIDT even if there is a malicious node in the network and moreover by the secure trust, the data is not transmitted to the sink if the node is malevolent which is displayed in Fig. 4.

The delay of SETA is considerably reduced without the malicious nodes than the CIDT method. Without the malicious nodes, the data are transmitted with less delay shown in Fig. 5. There is no delay in the transmission of data and it is transmitted quickly in the network.

Table III displays the comparison of SETA and CIDT with throughput with and without the malicious node. In the presence of malicious node, the throughput of the SETA is comparatively less to SETA without the malicious node.

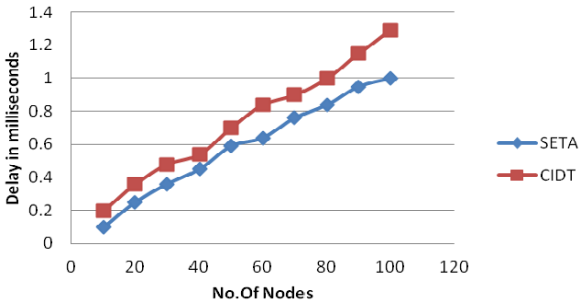

Fig. 4 Delay with Malicious Node

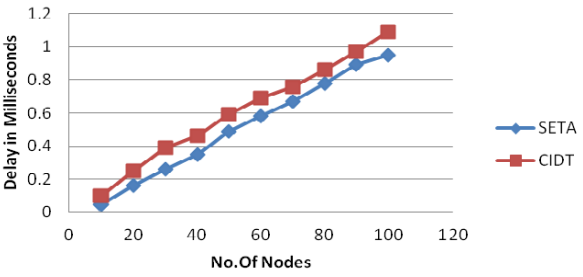

Fig. 5 Delay without Malicious Node

TABLE III
COMPARISON OF SETA AND CIDT WITH THROUGHPUT

No. Of Nodes	Throughput with Malicious Node		Throughput without Malicious Node	
	SETA	CIDT	SETA	CIDT
20	0.886	0.863	0.903	0.885
40	0.865	0.848	0.879	0.87
60	0.85	0.836	0.865	0.85
80	0.836	0.816	0.846	0.83
100	0.809	0.792	0.826	0.803

The throughput of the SETA in the presence of the malicious nodes is better than the CIDT method in which the data transmission per second is higher than the CIDT as shown in Fig. 6. It works well even in the presence of the malicious nodes. The data is more secure by the SETA approach.

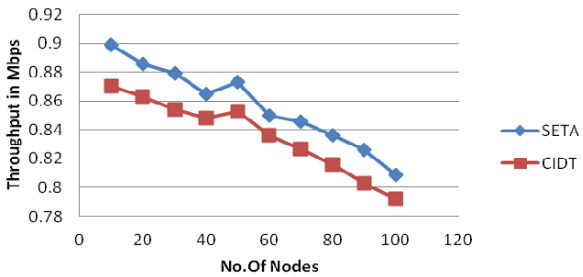

Fig. 6 Throughput with Malicious Node

The throughput of SETA without the malicious node is greater than the CIDT in which there is no malicious node and the transmission of data is faster. The throughput of SETA without the presence of the malicious node in comparison with

CIDT is given in Fig. 7. Since there is no malicious node the throughput of the SETA is higher.

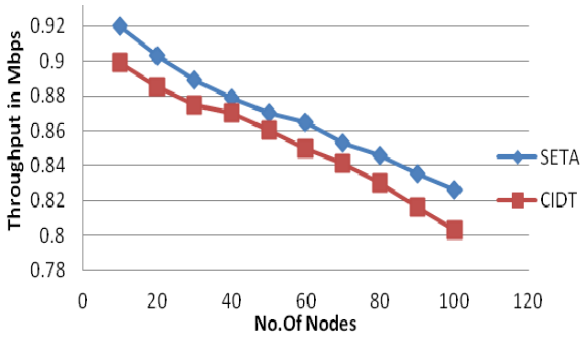

Fig. 7 Throughput without Malicious Node

V. Conclusion

The SETA approach helps to transmit the aggregated data in a secure manner through an efficient path by constructing the DCT between the CHs and the sink node. This scheme helps to find the optimal path in the network so that the aggregated data can be sent through that optimal path by which the data transmission of the data can be effective. The packet delay ratio is analyzed where there is less delay when the node transmits through the shortest path identified by the DCT than CIDT. Moreover, the SETA scheme is secure in finding the malicious node while transmission of the data and not transmitting through that particular malicious node in the network based on the trust value of the node. It is more secure in the transmission of data to the sink node. We have analyzed the security parameters like the accuracy of aggregation, delay in the transmission of data and throughput in the presence and absence of Malicious Nodes.

References

[1] H. Karl and A. Willig, "Protocols Architectures for Wireless Sensor Networks", New York, NY, USA: Wiley, 2000.
[2] R. Velmani and B. Kaarthick, "An Efficient Cluster-Tree Based Data Collection Scheme for Large Mobile Wireless Sensor Networks", IEEE Sensors Journal, Vol. 15, No. 4, April 2015.
[3] Davood Izadi, Jemal Abawajy and Sara Ghanavati, "An Alternative Clustering Scheme in WSN", IEEE Sensors Journal, Vol. 15, No. 7, July 2015.
[4] Chuan Zhu, Guangjie Han, "A Tree-Cluster-Based Data-Gathering Algorithm for Industrial WSNs with a Mobile Sink", Access, IEEE (Volume: 3), May 2015.
[5] Jenq-Shiou Leu, Tung-Hung Chiang, Min-Chieh Yu, and Kuan-Wu Su, "Energy Efficient Clustering Scheme for Prolonging the Lifetime of Wireless Sensor Network with Isolated Nodes", IEEE communications letters, vol. 19, no. 2, February 2015.
[6] Xu Jian, Yang Geng, Chen Zhengyu, Wang Qianqian, "A Survey on the Privacy-Preserving Data Aggregation in Wireless Sensor Networks", Communications, China (Volume: 12, Issue: 5), May 2015.
[7] Zhixin Liu, Yazhou Yuan, Xinping Guan, and Xinbin Li, "An Approach of Distributed Joint Optimization for Cluster-based Wireless Sensor Networks", IEEE/CAA Journal of automaticasinica, vol. 2, no. 3, July 2015.
[8] Liu Xiang, Jun Luo, and Catherine Rosenberg, "Compressed Data Aggregation: Energy-Efficient and High-Fidelity Data Collection", IEEE/ACM Transactions on Networking, Vol.21, No.6, and December 2013.
[9] Zaixin Lu, Wei Wayne Li and Miao Pan, "Maximum Lifetime Scheduling for Target Coverage and Data Collection in Wireless Sensor Networks", IEEE Transactions On Vehicular Technology, Vol. 64, No. 2, February 2015.
[10] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks", Ad Hoc Netw., vol. 3, no. 3, pp. 325–349, 2005.
[11] Q. Mamun, "A qualitative comparison of different logical topologies for wireless sensor networks", Sensors, vol. 12, no. 11, pp. 14887–14913, 2012.
[12] Mohsen Rezvani and Sanjay Jha, "Secure Data Aggregation Technique for Wireless Sensor Networks in the Presence of Collusion Attacks", IEEE Transactions On Dependable And Secure Computing, Vol. 12, No. 1, January/February 2015.
[13] Guoxing Zhan, Weisong Shi and Julia Deng, "Design and Implementation of TARF: A Trust-Aware Routing Framework for WSNs", IEEE Transactions On Dependable And Secure Computing, Vol. 9, No. 2, March/April 2012.
[14] Adnan Ahmed, Kamalrulnizam Abu Bakar, Muhammad Ibrahim Channa, Khalid Haseeb and Abdul Waheed Khan, "TERP: A Trust and Energy Aware Routing Protocol for Wireless Sensor Network", IEEE Sensors Journal, Vol. 15, No. 12, December 2015.