DYNAMIC AND EFFICIENT TREE-BASED DATA COLLECTION IN WIRELESS SENSOR NETWORKS

S.Thivyan Ravi Kumar, B.Maheswari, Dr. S.Karthik

Abstract— Periodic collection of data from sensors to sink is the basic operation in Wireless sensor networks. Here the data are aggregated together and they passed to the common sink. This many to one communication paradigm is known as convergecast. Data can be collected by using both single frequency channel and multiple frequency channel. By considering single frequency channel, time slots are reduced but the interference occurs and by using multiple frequency channel the interference is reduced. The TDMA scheduling is used for transmitting the packets. A capacitated minimal spanning tree and a degree constrained spanning tree is used to provide substantial performance over different deployment densities. In capacitated minimal spanning tree, the number of nodes in sub tree is no more than the total number of nodes in remaining sub tree. However capacitated minimal spanning tree does not provide the minimum cost for long suited nodes. A hybrid convergecast tree is proposed and implemented which is used to obtain minimum cost for both long and short suited nodes. To improve the data collection, the capacity is adjusted at each node whenever the packet moves from one sensor node to another sensor node.

Keywords—Convergecast, Data Collection, Degree Constrained spanning tree, Hybrid convergecast tree, Minimal Spanning Tree Time Division Multiple Access Scheduling.

1. INTRODUCTION

Wireless sensor networks consist of distributed autonomous sensor nodes which are used in various applications areas such as health, military, home. Each node receives data from environment and it will forward the data to its base station. The location of each sensor node is not predetermined and each sensor nodes contains its own self organizing capabilities. The unique feature of sensor nodes is the cooperative effort of sensor nodes. Instead of sending the raw data to the nodes responsible for the synthesis, sensor nodes use their processing abilities to locally carry out simple computations and it will transmit only the required and partially processed data. Sensor nodes are used for continuous sensing, detecting events, event ID, sensing the location. The sensor nodes are scattered in the sensor network. Each scattered sensor nodes has the ability to collect and route data to the sink.

In a wireless sensor networks the data collection can be collected by using TDMA scheduling. In TDMA, nodes communicate on different time slots in order to prevent conflicts. By communicating on different time slots, the TDMA based communication provide guarantee delivery. Consider a repeated frame of n timeslots in which each link of the data gathering tree is scheduled once. Once the pipeline is established the sink will receive aggregated information from all nodes in the network once per frame.

In TDMA schedule, time slotted and each slot is long enough to transmit and receive of a single packet. Hence the time slots are grouped as a non – overlapping frames and the schedule for each frame is repeated when the data collection is periodic. The main objective of TDMA design is to minimize schedule length, minimize Latency, minimizing energy consumption, maximizing capacity, maximizing fairness, minimizing communication costs, maximizing the parallel transmission, meeting deadlines, and minimizing interference.

A WSN has small or no infrastructure. It consists of a number of sensor nodes which works together to monitor a region and it is used to obtain data about the environment. The WSN is classified into two types such as structured and unstructured. An unstructured WSN contains the dense collection of sensor nodes. The Sensor nodes are deployed in an ad hoc manner into the field. Once deployed, the network is kept unattended to perform monitoring and reporting functions. Hence network maintenance and failure detection is difficult since there are so many nodes. In structured network, all or some of the nodes are deployed in a pre planned manner. Due to small number of nodes network maintenance is easy and management cost is reduced.

Data collection is the basic operation in wireless sensor networks. The sensor nodes measure attributes about the phenomenon of interest and it transmits the data to a common base station. The

collected data is forwarded to a common sink through sensors over a tree based routing topology. This is a many to one communication paradigm in which data flows from many nodes to a single node and it is called as convergecast. Once the data is collected at the sink, it can be either recorded or stored for future analysis, or it can be processed immediately to take certain actions depending on application requirements.[5]

2. RELATED WORK

Collecting data is the basic operation in wireless sensor networks. Collecting those data in a well-organized manner is more critical when compared to the performance of sensor networks. Each sensor will measure the values at regular time intervals and it will send that to sink node. The total collection of data is called as snapshot.   Siyuan Chen et al.,[4] discussed about the hypothetical limits of data collection capacity. Here the sensor network is a TDMA based network. In the past, the data collection capacity is based on large scale random networks, though most of the sensors are not deployed uniformly and the available sensors will not be as huge as in theory. For this purpose we have to know about the capacity of data collection in a network. The capacity of data collection shows how the sink collects data from sensor nodes quickly by considering its interference conditions. The upper and lower bounds for data collection capacity are constructed based on the protocol interference and disk graph models. A simple BFS tree based method is used to achieve the collection capacity which matches the upper bound [4].

Yafeng Wu et al.,[7] discuss about the improvement of communication performance by using multiple channels. The current multi channel protocols are not suitable for wireless sensor networks due to minimum number of available channels and unavoidable time errors found in networks. Here a novel tree based multi channel scheme for data collection applications is constructed which allocates disjoint trees and exploits parallel transmissions among trees. By using NP complete, the interference is minimized within the trees. A greedy channel allocation algorithm is proposed for use in small number of channels in dense network. A TMCP

protocol is implemented to improve the network throughput and to reduce the packet losses.

In sensor networks the packets generated by each and every node have to reach the sink. This many – to –one communication is known as convergecast. A TDMA schedule is used which minimizes the total time required to complete the convergecast. A simple version of problem is considered where every node generates exactly one packet. Yiin Zhang et al.,[8] constructed a distributed scheduling algorithm for the tree networks that requires at most $\max(3n_k – 1,N)$ time slots for convergecast, where n_k represents the maximum number of nodes in any sub tree and N represents the number of nodes in network. The Distributed scheduling algorithm requires at most 3N time slots in any network. The proposed simulation shows that the number of time slots required is about 1.5 N. Two bounds are required for the packets to be buffered at the node during convergecast. Sleep schedules for nodes are considered for conserve energy. It reduces the energy consumption by at least 50%.Breadth first search tree is considered for convergecast scheduling.

S.Upadhyayula et al.,[7] address the problem of performing the operation of Data Aggregation enhanced Convergecast(DAC) in an energy and latency efficient manner. By assuming as all the nodes in the network have a data item and there is an a priori known application dependent data compression factor, the total data is collected.Two DAC tree Algorithms are constructed. First is the use of minimum spanning tree and the second is the use of single source shortest path spanning tree algorithm. These two algorithms are combined in order to construct an energy optimal DAC tree for any fixed value. The nodes in this tree are scheduled as a collision free communication using a channel allocation algorithm. In order to achieve low latency the ß constraint which puts soft limit on the

minimum number of children a node have in a DAC tree.

Collecting data from sensor nodes is the primary operation in Wireless sensor networks, where sensor nodes determine attributes about the observable fact of interest and transmit their readings to the sink. This is a many-to-one communication, in which data flows from many nodes to a common sink and it is called as convergecast. Once the data is collected it can be either recorded and stored for further analysis or it can be processed immediately to take certain actions depending on the applications requirements.[5]

In WSN the data collection occurs when any external event triggers, such as queries or events or it can be for continuous periodic monitoring without any external triggering occurs. When the number of source node are large or if the data rate is high then carrier sense multiple access protocols, such as CSMA may fail to allocate the medium which cause retransmissions and collisions. For this purpose TDMA is developed. Ozlem Durmaz Incell et al.,[5]describes the contention free Time Division Multiple access scheduling based protocols for collecting data using tree based routing topologies. By using TDMA the nodes can communicate on different slots to prevent interference and conflicts. Consecutive time slots are grouped into non overlapping frames. Hence the schedule for each frame is repeated when data collection is periodic. TDMA also provides guaranteed completion time for data collection. The algorithms are classified according to the design objectives. In order to minimize schedule length, minimize latency, minimize energy consumption, and to maximize fairness the algorithms are analyzed.By using multi-hop TDMA spatial reuse of time slots are allowed and more than one node can transmit simultaneously if the receivers are in non – conflicting parts of the network. In order to minimize the schedule length the raw data convergecast and aggregated convergecast are used.

Wireless sensor network consists of hundreds to thousands of wireless sensor nodes and those nodes are randomly distributed. Low cost sensors contains low signal-to-noise ratio, so it allows higher level of redundancy. Thus the redundancy makes the network more robust to changing environment. The wireless sensor nodes are battery – powered equipment so battery replacement or recharges are impractical. In WSN the energy saving scheme is based on clustering. A network is divided into number of clusters and a cluster head is selected from those clusters. It will collect data from cluster members. By using clustering significant improvement in network lifetime is obtained but it also introduces bottleneck problem.

To avoid the bottleneck problem the Chi-Tsun Cheng[1] modifies the network topology from star based to tree based. The network structure is formed based on two approaches namely top down approach and bottom up approach and the network formation may be either centralized or decentralized. By comparing with minimum spanning tree and a single hop network the author shows that this network structure is efficient in terms of collecting data at the time.

Wireless sensor network consists of large number of small sensors with low power transceivers will be an effective tool for gathering data in variety of environment [2]. The sensors will sense the data and it will help to monitor and aggregate the data. The important requirements for gathering data includes network lifetime, scalability and load balancing. In order to utilize the energy efficiently different methods are used such as cluster-based, tree-based and chain-based.

Gurpreet Singh Chhabra and Dipesh Sharma[2] uses both cluster based and tree based protocols to improve the power consumption by improving First node death(FND) and it avoids the communication directly between sink and sensor nodes. The proposed method reduces the power consumption on avoiding communication directly between sink and sensor nodes. It also increases the network lifetime as compared with others. This cluster based protocols works on two phases. In first phase the network lifetime is minimized by balancing the energy consumption nodes and in second phase the communication overhead in reduced by forming the tree structure.

Wireless sensor networks consists of sensor nodes which are connected to each other. The sensor nodes will collect data and it will forward it to the common base station. The sink broadcasts the request and one node tries to transmit data with the help of

other nodes. This WSN is based on three categories: Cluster based, Chain based and Tree based. When one node fails then consequently the chain of data transmission will be lost in chain based structure. In Cluster based structure, the cluster head or aggregator node may be attacked by malicious attacker. Thus if the data collection in not performed efficiently, traffic and energy consumption occurs.

Hanieh Alipour and Alireza Nemaney Pour[3] proposes an efficient approach for data collection in wireless sensor network by using Member Forward list. Here a tree based algorithm is used for forwarding the data and this MFL helps other nodes to find the route for forwarding data when previous forwarder node has failed. By using this MFL shortest path is identified for forwarding data to the sink. This protocol decreases the latency, increases the energy efficiency and prevents the algorithm from repeating when a node has failed or died for some reasons.

In [6] Single frequency channel is used for collecting the data with the aim of minimizing the number of timeslots .But by using it the timeslots are reduced but interference occurs. Here the sink remains as a bottleneck, receiving data over different paths does not reduce the schedule length. Due to this problem this paper proposed and used the multiple frequency channels in order to reduce the interference. Here the capacitated minimal spanning tree is constructed where the number of nodes in a sub tree is no more than half the total number of nodes in remaining sub trees.The main problem in capacitated minimal spanning tree is that it does not provide the minimum cost for long suited nodes.

OUR CONTRIBUTION

In [6] the capacitated minimal spanning tree is proposed and used but it has a drawback for long suited nodes. The cost will be high when it is used for long suited nodes. A hybrid convergecast tree is proposed which is used to obtain minimum cost for long suited nodes. To improve data collection the capacity is also adjusted at each node whenever the packet moves from one sensor node to another sensor node. TDMA scheduling and multiple frequency channels are proposed to obtain the data in a fast and correct manner.

HYBRID CONVERGECAST TREE

For transmitting the packets in short distance the capacitated minimal spanning tree is used to obtain minimum cost. A hybrid convergecast tree is proposed for transmitting the packets with minimum cost for long suited nodes. The following steps are involved for developing the hybrid convergecast tree:

2.1 WIRELESS SENSOR NETWORK SETUP

Wireless Sensor Networks consists of sensor nodes which are connected to one another and these sensors are used for transmitting the packets to the common base station. While constructing the network we consider the network size, sensor radius, sensor cost, sensor period, transmission period, transmission radius, transmission cost. Here the network size may vary from small to large depending on the number of sensors needed. Sensor radius shows the individual range where each sensor can sense within the network of its own.

The sensor cost denotes the rate of each sensor and sensor period shows the battery power of each sensor. Transmission radius denotes the area in which the sensor can sense the incoming packet and it can forward the packet up to the range. Transmission cost shows the amount required for transmitting the packets.

2.2 TDMA SCHEDULING

In order to prevent the inconsistency which occurs during transmission, the TDMA scheduling is used. In TDMA the time is slotted and each slot is used for transmitting and receiving the packets. For example consider a time as 9sec and if there are 3 nodes the time will be allocated as 3 sec for each channel and within the time the packets will be sent. Sometimes the allocated seconds will be wasted due to completion within the allocated time. For this purpose, in hybrid convergecast tree as given in the above example assume that the same time is initially set. If the channel completes the job within the given time ie within 2sec then the   3sec will be changed to 2sec for remaining channels. Likewise the time is adjusted whenever the packets move from one sensor node to another sensor node. Thus the time is not wasted.

2.3 POWER CONTROL AND MULTI CHANNEL SCHEDULING

By using single frequency channel the time is reduced but interference occurs. Here all packets are transmitted at the same frequency. When multiple frequency channels are used the interference is reduced because each channel uses different frequency. Same frequency can be allocated for the channel which is far away from the current channel for reducing the overlapping. Thus this power control helps us to minimize the interference.

2.4 IMPLEMENTATION OF ROUTING TREES

In the routing tree implementation the routing structure comes from using capacitated minimal spanning tree and hybrid convergecast tree for raw data convergecast where number of nodes in a sub tree is no more than half the total number of nodes in the remaining sub tree. By considering this the overload at single side is reduced. Hence whenever a new node is inserted, the load at each node is considered by comparing the node below to it and also with the nodes which are nearer to it and based on the load a new node will be inserted. Thus while constructing the tree the capacity is adjusted at each node.

3. SIMULATION RESUTS

Here we evaluated the performance of the proposed scheme in comparison to Capacitated Minimal Spanning tree. The following table shows the parameters for both RCCT and CMST. Here the network life time analysis and Delay analysis are evaluated and it is proved that RCCT performance is best when compared to CMST.

TABLE I. Network Life Time Analysis

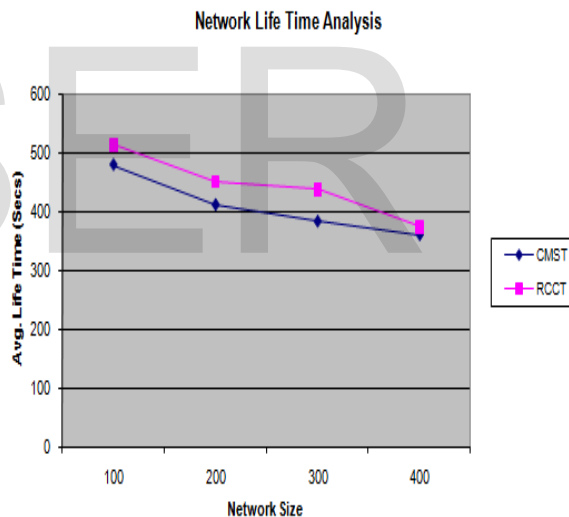| Network Size	CMST	RCCT
100	478	514
200	410	451
300	382	438
400	359	373

As shown in the table I the network lifetime of RCCT is high when compared with CMST. It considers both long and short suited nodes. In table II the delay analysis is evaluated and compared. By considering long suited nodes the delay is high in CMST. Because it is well suited for short suited only.

TABLE II. Delay Analysis

Network Size	CMST	RCCT
100	$9.83\ e^{-09}$	$5.79\ e^{-09}$
200	$9.83\ e^{-09}$	$4.47\ e^{-09}$
300	$7.15\ e^{-09}$	$1.40\ e^{-09}$
400	$4.88\ e^{-09}$	$3.79\ e^{-09}$

4. PERFORMANCE EVALUATION

The performance of both RCCT and CMST is evaluated and it is shown in the figure 1 and figure 2.The performance is evaluated by considering up to 400 nodes. Network life time analysis is high when compared to CMST. Delay is low in RCCT
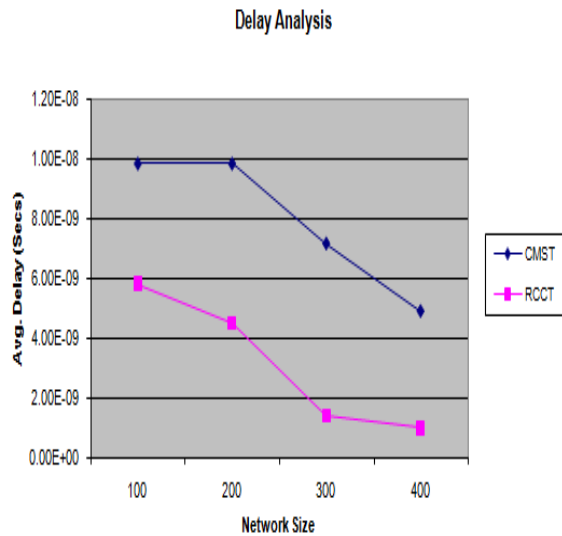
Figure 1

Figure 2

5. CONCLUSION

Wireless sensor networks can be used in various applications areas. A sensor networks consists of large number of nodes and the location of each sensor nodes is predetermined. By using TDMA, the nodes communicate on different time slots in order to prevent conflicts. In order to improve the data collection the capacity at each node is adjusted whenever the packet moves from one sensor node to another sensor node. . A hybrid convergecast tree is proposed for transmitting the packets with minimum cost for long suited nodes.

REFERENCES

[1]Chi-Tsun Cheng, Chi K Tse and Francis C.M.Lau, "A Tree Based Data Collection Network structure for Wireless Sensor Networks", Journal of Electronic Science and Technology of China, Vol.6, No.3, September 2008.

[2]Gurpreet Singh Chhabra, Dipesh Sharma, "Cluster Tree Based Data Gathering in Wireless Sensor Networks", International Journal of Soft Computing", ISSN:2231 - 2307,Vol.1, Issue-1, March 2011.

[3]Hanieh Alipour and Alireza Nemaney Pour, "An Efficient Data Collection Approach for Wireless Sensor Networks", World Academy of Science, Engineering and Technology, 2011.

[4]Siyuan Chen, Minsu Huang, Shaojie Tang, Yu Wang, "Capacity of Data Collection in Arbitrary Wireless Sensor Networks", IEEE Transaction on Parallel and Distributed Systems, Vol.23, No.1, January 2012.

[5]Ozlem Durmaz Incell, Amitabha Ghosh and Bhaskar Krishnamachari, "Scheduling Algorithms for Tree-Based Data Collection in Wireless Sensor Networks".

[6]Ozlem Durmaz Incel, Amitabha Ghosh, Bhaskar Krishnamachari, and Krishnakant Chintalapudi, "Fast Data Collection in Tree-Based Wireless Sensor Networks", IEEE Transactions On Mobile Computing, Vol. 11, No. 1, January 2012.

[7]S.Upadhyayula, S.K.S.Gupta, "Spanning Tree based algorithms for low latency and energy efficient data aggregation enhanced convergecast,(DAC) in Wireless sensor Networks", Arizona state University, 22 April 2006.

[8]Yafeng Wu, John A. Stankovic, Tian He and Shan Lin, "Realistic and Efficient Multi – Channel Communications in Wireless Sensor Networks", University of Minnesota.

[9]Yiin Zhang b, Qingfeng Huang , Shashidhar Gandham, "Distributed time-optimal scheduling for convergecast in Wireless Sensor Networks", 7 November 2007.

BIOGRAPHY

Thivyan Ravi Kumar.S received B.Tech degree in Information Technology from Anna University, Chennai,Tamilnadu, INDIA in 2008 and pursuing M.E Software Engineering in SNS College of Technology affliated to Anna University Chennai. His Research includes Wireless Networks and Network Security

B.Maheswari is presently Assistant professor, Department of Computer science and Engineering, SNS college of Technology, affiliated to Anna University- Coimbatore, Tamil Nadu, India. She received her Bachelor of Engineering degree in Computer Science and Engineering in 2007 and received her Master of Computer Science and Engineering degree in SNS College of Technology, Coimbatore, India. Her research interests include network security, Mobile Computing and wireless systems. She has published more number of articles in International/ National Journals and Conferences.

Professor Dr.S.Karthik is presently Professor & Dean in the Department of Computer Science & Engineering, SNS College of Technology, affiliated to Anna University- Coimbatore, Tamilnadu, India. He

received the M.E degree from the Anna University Chennai and Ph.D degree from Ann University of Technology, Coimbatore. His research interests include network security, web services and wireless systems. In particular, he is currently working in a research group developing new Internet security architectures and active defense systems against DDoS attacks. Dr.S.Karthik published more than 35 papers in refereed international journals and 25 papers in conferences and has been involved many international conferences as Technical Chair and tutorial presenter. He is an active member of IEEE, ISTE, IAENG, IACSIT and Indian Computer Society.

IJSER