RESEARCH ARTICLE OPEN ACCESS

Aggregation Mechanism for Reducing Schedule Length in Tree Based Wireless Sensor Networks

M.Karthika[#1], D. Gautham Chakravarthy[*2]
[#]Department of Computer Science & Engineering Coimbatore institute of engineering & Technology, Coimbatore, India.

Abstract
To Explore and evaluate different techniques using real simulation models under multilevel communication paradigm. Data packets are time scheduled on single frequency channel by minimizing the time slots to complete a CONVERGECAST. Scheduling with transmission power control will diminish the effects of interference. The power control helps in reducing the schedule length under single frequency. Scheduling transmissions using multiple frequencies is more efficient than the single frequency. By providing power bounds in the schedule length interference is eliminated. The proposed algorithm can achieve these bounds. The use of multi frequency scheduling is sufficient to eliminate the interference. The data collection is no longer limited by the interference. To provide a proper solution degree-constrained spanning trees and minimal spanning trees are created. This will provide a significant improvement in scheduling performance. Finally, in a schedule length, the collisions in different interference over different channel models have been evaluated.
Keywords-Degree-Constrained Routing Trees,Data Centric Approach, Greedy Aggregation, contention-based vs. contention free protocols, DAC tree construction algorithms, Unit Disk Graphs (UDG)

I. INTRODUCTION

CONVERGECAST, namely the collection of data from a set of sensors toward a common sink over a tree based routing topology, is a fundamental operation in WSN. In many applications, it is crucial to provide a guarantee on the delivery time as well as increase the rate of such data collection. For instance, in safety and mission-critical applications where sensor nodes are deployed to detect oil/gas leak or structural damage, the actuators and controllers need to receive data from all the sensors within a specific deadline, failure of which might lead to unpredictable and catastrophic events. This falls under the category of one-shot data collection. On the other hand, applications such as permafrost monitoring require periodic and fast data delivery over long periods of time, which falls under the category of continuous data collection.

II. LITERATURE SURVEY

A. Degree-Constrained Routing Trees

It explores a hierarchy of techniques using realistic simulation models to enhance the data collection rate. It begin by considering TDMA scheduling on a single channel, reducing the original problem to minimizing the number of time slots needed to schedule each link of the aggregation tree. The second technique is to combine the scheduling with transmission power control to reduce the effects of interference.

Further the data collection rate can be enhanced by the use of degree-constrained routing trees[1]. A degree-constrained minimum-hop tree is constructed using a modified version of Dijkstra's shortest path algorithm. Consider a graph G (V, E) and a given degree constraint max degree. Each node n keeps a value for the number of its children C (n) with an initial value = 0 and hop count to the sink HC (n) with an initial value = ∞. The algorithm starts with a set T that contains the sink node s ($HC(s) = 0$), at each iteration we add a node $m \in T$ to T with the following constraints:

- there is a node $m' \in T$ such that edge $(m,m') \in E$,
- $C(m') <$ max degree - 1,
- The hop count to the sink = $HC(m)$ is minimized.

The updates are made as $HC(m) = HC(m') + 1$ and $C(m') = C(m') + 1$. The algorithm stops when $|T| = |V|$ or when no more nodes can be added since the degree of the all nodes in T have reached the max degree.

If the nodes select their parents according the minimum hop criteria without a degree constraint, all the nodes will select the sink as a parent and this schedule will take n time slots. On the other hand, if we limit the number of connections per node as 2, this will result in 2 sub trees rooted at the sink. If there is enough number of frequencies to eliminate all the interference then the network can be scheduled in 2 time slots and we achieve a factor of $n/2$ reduction in the schedule length.

The gains with the degree constrained trees may be costly in terms of latency due to the increased number of hop distances to the sink node.

B. Data Centric Approach

Sensor networks are distributed event-based systems that differ from traditional communication networks in several ways: sensor networks have severe energy constraints, redundant low-rate data, and many-to-one flows. Data-centric technologies [2], are needed that perform in-network aggregation of data to yield energy-efficient dissemination. In this paper it explores data-centric routing and compare its performance with traditional end-to-end routing schemes. It shows that data-centric routing offers significant performance gains across a wide range of operational scenarios.

Sensor networks are typically event-based systems. A sensor network consists of one or more “sinks” which subscribe to specific data streams by expressing interests or queries. The sensors in the network act as “sources” which detect environmental events and push relevant data to the appropriate subscriber sinks. For example, there may be a sink that is interested in a particular spatial-temporal phenomenon (“does the temperature ever exceed 70 degrees in area A between 10am and 11am?”). During the given time interval all sensors in the corresponding spatial portion of the network act as event based publishers. They publish information toward the Subscribing sink if and when they detect the indicated phenomenon.

Because of the requirement of unattended operation in remote or even potentially hostile locations, sensor networks are extremely energy-limited. However since various sensor nodes often detect common phenomena, there is likely to be some redundancy in the data the various sources communicate to a particular sink. In-network filtering and processing techniques can help conserve the scarce energy resources.

Data aggregation has been put forward as an essential paradigm for wireless routing in sensor networks. The idea is to combine the data coming from different sources enroot – eliminating redundancy, minimizing the number of transmissions and thus saving energy. This paradigm shifts the focus from the traditional address centric approaches for networking (finding short routes between pairs of addressable end-nodes) to a more data centric approach (finding routes from multiple sources to a single destination that allows in-network consolidation of redundant data).

It has focused on the case where there is a single sink. Although this is a reasonable scenario for many applications, it is reasonable to ask what would happen if there were additional sink.

C. Greedy Aggregation

In-network data aggregation is essential for wireless sensor networks where energy resources are limited. In this paper, it explores greedy aggregation that adjusts aggregation points to increase the amount of path sharing, reducing energy consumption.

Greedy aggregation differs from opportunistic aggregation in path establishment and maintenance. To construct a greedy incremental tree, a shortest path is established for only the first source to the sink whereas each of the other sources is incrementally connected at the closest point on the existing tree.

In greedy approach [3], each exploratory sample also contains an energy cost for delivering this sample from the source to the current node. In addition, each source on the existing tree (*i.e.,* a source on an established path) also generates an on-tree incremental cost message which corresponds to each new exploratory sample received. The incremental cost message contains the incremental energy cost required for delivering the corresponding exploratory sample to the existing tree. This incremental cost message is only sent and updated along the aggregation tree toward the sink. To find the closest point on the tree, the incremental energy-cost field can be updated only by closer nodes (*i.e.,* nodes which have received the corresponding exploratory sample at lower cost). In this greedy approach, the most preferred neighbor to reinforce is a neighbor which has delivered the exploratory sample or its corresponding incremental cost message at the lowest energy cost.

Greedy approach constructs an energy-efficient aggregation tree using data-centric reinforcement mechanisms and prunes inefficient paths using a greedy heuristic for weighted set-covering problems.

D. Contention-based vs. Contention free protocols

In this paper [4], it explores the potential for using 802.15.4 based radios for wireless sensing in low-latency hard real-time discrete event control applications. Given fixed sensing and actuation delays, such deadlines can usually be translated into communication delay deadlines. A message that does not reach its destination before this deadline may cause the machine to go into an error condition that requires its temporary halting or resetting. In machines today, sensors and actuators communicate to the controller via cables and cater to hard real-time communication latencies ranging from 5-50ms depending on the specifics of the machine. Inherent to most discrete event control systems is the unpredictable nature of the traffic i.e., it is hard to predict when, how many, or which sensors will be triggered to communicate at the same time to the controller. This is because the communication is primarily event-driven and it is often impossible to predict the times and nature of occurrence of external events.

In general, traffic bursts are common in most discrete event systems because, i) a single event may lead to several sensors triggering at the same time and ii) more than one event may occur at the same time. In the rest of the paper we shall refer to such event driven bursts as sensor bursts. In the event of a sensor burst, then, messages from all the sensors must reach the controller within the specified deadline since the controller can take appropriate action only upon receiving all the inputs. Failure of receipt of a message from even one sensor may lead to unpredictable failures in the system forcing it into an error recovery state.

FTDMA MACs that use 4 or more transceivers at the controller may be a suitable option for small systems (50-100 sensors), however, FTDMA does not scale well for larger systems. The contention based MAC, T-MALOHA that promises to perform as well or better than FTDMA systems in terms of error probability for sensor burst sizes under 20. The drawback of T-MALOHA however, lies in the fact that its longevity may be much lower than that of FTDMA depending on the frequency of occurrence of events.

The choice between using FTDMA and T-MALOHA is based on the tradeoff between size of the system in terms of number of sensors,The burst sizes that may occur and Frequency of occurrence of sensory events since this may impact T-MALOHA very adversely. When the sensory event frequency is “low” (one event per 10 seconds or less) and burst sizes are below 20, T-MALOHA provides a clear benefit over FTDMA.

E.DAC tree construction algorithms

In this paper [8], WSNs employ battery-powered sensor nodes. Communication in such networks is very taxing on its scarce energy resources. Converge cast – process of routing data from many sources to a sink – is commonly performed operation in WSNs. Data aggregation is a frequently used energy-conversing technique in WSNs. The rationale is to reduce volume of communicated data by using in-network processing capability at sensor nodes. In this paper, they addressed the problem of performing the operation of Data Aggregation enhanced Converge cast (DAC) in an energy and latency efficient manner. They assumed that all the nodes in the network have a data item and there is an a priori known application dependent data compression factor (or compression factor), c, that approximates the useful fraction of the total data collected.

The paper first presents two DAC tree construction algorithms. One is a variant of the Minimum Spanning Tree (MST) algorithm and the other is a variant of the Single Source Shortest Path Spanning Tree (SPT) algorithm. These two algorithms serve as a motivation for our combined algorithm (COM) which generalized the SPT and MST based algorithm. The COM algorithm tries to construct an energy optimal DAC tree for any fixed value of α (= 1 - γ), the data growth factor.

The nodes of these trees are scheduled for collision-free communication using a channel allocation algorithm. To achieve low latency, these algorithms use the β-constraint, which puts a soft limit on the maximum number of children a node can have in a DAC tree. The DAC tree obtained from energy minimizing phase of tree construction algorithms is restructured using the β-constraint (in the latency minimizing phase) to reduce latency.

F. The scaling laws of multi-modal WSNs

In this paper [5], dense wireless sensor networks deployed to observe multiple random processes. The Requirement is to reconstruct an estimate of each random process at the corresponding collector node. This leads to multiple many to-one data gathering wireless channels that interfere with one another. It derive the transport capacity that the network can provide to each process and characterize an achievable rate region for the dense multi-modal network.

The main idea of this scheme is to allow closely located nodes to cooperate with each other in transmitting information to the collector nodes, which comes at a very small cost for densely deployed networks. Each node distributes its observed information about a particular process to its neighbors. Since all the sensor nodes do not need to compete for the same collector node, the concept of spatial frequency reuse comes into play. Thereby other nodes, which are far from this node, will be allowed to simultaneously transmit their observed information about other processes to their neighbors. In the next time slot, all the nodes which correctly decode the transmissions intended to them will cooperate to send the information to the corresponding collector node through beam forming.

It established the gains possible with cooperation between the sensor nodes. It characterized an achievable rate region for dense multi-modal wireless sensor networks using a scheme that exploits the proximity of sensors to allow for efficient cooperation. It have shown that it is possible to observe $O(N^{\beta})$ processes simultaneously, and still achieve a transport capacity of $\Theta(\log(N))$ for each of the processes, with a large number of sensors N and a fixed total average power.

G. Approximation algorithm

One important function of many wireless sensor networks (WSNs) is to gather data from hostile or remote environments. It is expected of such networks to work untended for a long duration. Due to limited energy resources, the above requirements put constraints on the energy usage. Examining various functionalities of sensor networks, communication can be singled out as one function that devours big share of the energy resources.

Data aggregation is an energy conservation technique which tries to reduce the volume of data communicated by collecting local data at intermediate nodes and forwarding only the result of an aggregation operation, such as min and max, towards the sink node. Since a converge cast operation usually follows a broadcast operation, the path taken by a broadcast packet is also used for aggregating data in the converge cast.

However, research [6] shows that performing data aggregation along this routing path is not energy efficient. The general data aggregation problem is – given m sources and one sink in an n node network (m < n), find a minimum weight sub graph that includes all sources. This is a well known NP-complete problem, known as the Steiner Tree Problem (STP).

There are many approximation algorithms for solving STP. From sensor networks perspective, in the literature there are several heuristic approaches to solve this problem. The problem addressed in this paper is an important special case of the general data aggregation problem in which all the nodes in the network are source nodes.

H. Unit Disk Graphs (UDG)

Fast and periodic collection of aggregated data is of considerable interest for mission-critical and continuous Monitoring applications in sensor networks. In the many-to-one communication paradigm, consider the scenarios where data packets are aggregated at each hop en route to a sink node along a tree-based routing topology and focus on maximizing the data collection rate at the sink by employing TDMA scheduling and multiple frequency channels.

This paper [1] lies in proving that minimizing the schedule length for an arbitrary network in the presence of multiple frequencies is NP-hard, and in designing approximation algorithms with worst-case provable performance guarantees for geometric networks. In particular, a constant factor approximation for networks is modeled as unit disk graphs (UDG) where every node has a uniform transmission range, and a O (Δ (T) log n) approximation for general disk graphs where nodes have different transmission ranges; n is the number of nodes in the network and Δ (T) is the maximum node degree on a given routing tree T.It focuses on the link scheduling problem of maximizing the aggregated data collection rate at the sink node under the setting of TDMA protocols and multiple frequency channels.

III. CONCLUSIONS

In this work, fast converge cast in WSN where nodes communicate using a TDMA protocol to minimize the schedule length is considered. The system addressed the fundamental limitations due to interference and half-duplex transceivers on the nodes and explored techniques to overcome the same. It is found that while transmission power control helps in reducing the schedule length, multiple channels are more effective. Once interference is completely eliminated, the project proves that with half-duplex radios the achievable schedule length s lower-bounded by the maximum degree in the routing tree for aggregated converge cast, and by max (2nk − 1, N) for raw-data converge cast. Using converge cast scheduling algorithms, the time slot length is reduced much.

REFERENCES

[1] A. Ghosh, O. D. Incel, V. A. Kumar, and B. Krishnamachari, "*Multi-channel scheduling algorithms for fast aggregated convergecast in sensor networks*," in MASS '09, Macau, China.

[2] Bhaskar Krishnamachari, Deborah Estrin, Stephen Wicker, "*Impact of data aggregation in wireless sensor networks*," in: International Workshop on Distributed Event-Based Systems (DEBS'02) Vienna, Austria, July 2007.

[3] Chalermek Intanagonwiwat, Deborah Estrin, Ramesh Govindan, John Heidemann, "*Impact of network density on data aggregation in wireless sensor networks*," in: Proceedings of International Conference on Distributed Computing Systems (ICDCS), Vienna, Austria, July 2005.

[4] Chintalapudi and L. Venkatraman, "*On the design of mac protocols for low-latency hard real-time discrete control applications over 802.15.4 hardware*," in IPSN '08, pp. 356–367.

[5] Gopala and H. E. Gamal, "*On the Scaling Laws of Multi-Modal Wireless Sensor Networks*," In Proceedings of the 23 rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), 2004.

[6] Hans J. Promel, Angelika Steger, "*RNC-approximation algorithms for the Steiner tree problems*," in: Proceedings of 14th Annual Symposium on Theoretical Aspects of Computer Science, 2007, pp. 559–570.

[7] O. D. Incel and B. Krishnamachari, "*Enhancing the data collection rate of tree-based aggregation in wireless sensor networks*," in SECON '08, San Francisco, CA, USA, pp. 569–577.

[8] S. Upadhyayula and S. Gupta, "*Spanning tree based algorithms for low latency and energy efficient data aggregation enhanced convergecast (dac) in wireless sensor networks*," Ad Hoc Networks, vol. 5, no. 5, pp. 626–648, 2007.