International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 12 Issue 2 –JANUARY 2015.

Enhancing Aggregated Data Converge cast in Wireless Sensor Networks

Y. Nalini Priyanka[1] and A. L. Sreenivasulu[2]

[1]*PG Student, Dept. of CSE, Intell Engineering College, Affiliated to JNTUA, Andhra Pradesh, India*

[2]*Assistant Professor in Dept. of CSE, Intell Engineering College, Affiliated to JNTUA, Andhra Pradesh, India*

***Abstract*— In WSNs, the segregation of data from a set of sensors toward a frequent sink over a tree based routing topology is a primary operation. In such networks, Convergecast is a pattern of communication where data flows from many sources to a single sink node. The fundamental question is "how fast can information are collected from a wireless sensor network organized as tree?" Convergecast network requires proper coordination among nodes to avoid collisions. In order to address this, we consider time scheduling on a single frequency channel with the objective of minimizing the number of time slots that are required to complete a converge cast. Scheduling is combined with transmission power control to mitigate the effects of interference, and to show that while power control helps in reducing the schedule length under a single frequency, scheduling transmissions using multiple frequencies is more proficient.**

Keywords: Wireless Sensor Networks, Multi frequency scheduling, Aggregated Convergecast , Raw data convergecast.

I. Introduction

Wireless sensor systems are deployed in unattended and hostile environments such as battlefield, adversary environments. This model of sensor networks includes a central location for data collection, allowing the operators to retrieve aggregated data from any of the nodes in the network in a distributed manner. Convergecast is a many-to-one communication paradigm in which data flows from many nodes to a single node over a tree-based routing topology.

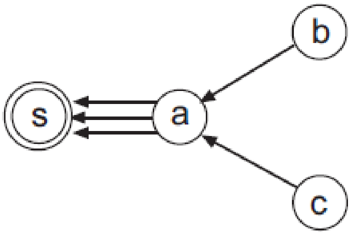


Fig 1. Convergecast - data flows from nodes a, b, and c towards a single nodes

Once data is collected at the sink, it can either be recorded and stored for future analysis, or can be processed immediately. Data collection can either be triggered by external sources, such as queries or events as and when they appear, or can be for continuous periodic monitoring without any external triggering. Different objectives can be associated with data collection based on the environment in which the sensors are deployed. For instance, disaster early warning applications include detection of forest fire, gas/oil leaks, or structural damage identification, bursty traffic generated by events needs to be delivered to the sink as quickly and as reliably as possible to prevent catastrophes.

The fundamental question is: "How fast can data be streamed from a set of sensors to a sink over a tree-based topology?" The primary limiting factors of fast data collections are:

Interference in the wireless medium

Half-duplex transceivers on the sensor nodes

Topology of the network.

We have two types of data collection:

- Aggregated convergecast - packets are aggregated at each hop. This is applicable when a strong spatial correlation exists in the data.
- raw-data convergecast - packets are individually relayed toward the sink. This is applicable when every sensor reading is equally important, or the correlation is minimal.

In the proposed work, we show the improvement due to the routing structure that comes from using capacitated minimal spanning trees for raw-data convergecast, where the number of nodes in a sub tree is no more than half than the total number of nodes in the remaining sub trees.

II. EXISTING SYSTEM

For fast data collection the contention free medium access control protocols such as Time Division Multiple Access are better fit, since they could eliminate collisions and retransmissions and provide guarantee on the completion time when compared to contention-based protocols for periodic traffic. To acheive additional improvement, on the achievable schedule lengths the transmission power control is combined with scheduling, and we use multiple frequency channels to permit more concurrent transmissions. The problem of constructing conflict free Time Division Multiple Access schedules under the simple graph-based interference model has been proved to be NP-complete. Also, Time Division Multiple Access scheduling (TDMA) enables spatial reuse with which if multiple frequencies are combined the data collection rate frequently no longer remains limited by interference but only by the topology of the network.

Like routing algorithms, aggregation algorithms should also be aware of the network topology and queries which are propagated by root, based on these selects aggregation function and aggregate the nodes. Under regular, heavy traffic conditions, contention-free medium access control (MAC) protocols, such as Time Division Multiple Access (TDMA),nodes communicate on different time slots to prevent conflicts thus eliminate collisions, overhearing, and idle listening, the main sources of energy consumption in wireless communications. TDMA-based communications can provide provable guarantee on the completion time of data collection, for instance, in time detection of events. Another key aspect of time-slotted communication is the robustness of the network during peak loads. When the number of source nodes is many or the data rates are high, carrier-sense multiple access protocols, such as CSMA, may fail to successfully allocate the medium, causing retransmissions and collisions. The use of simple graph-coloring based scheduling schemes might be idealistic but may fail in practice, because interference is not a binary phenomenon, and two nearby concurrent transmissions can actually be successful only if the interference level is tolerable.

TDMA-based scheduling algorithms are exploited for fast and timely delivery of data with the objective of minimizing the time to complete convergecast. In these schemes consecutive time slots are grouped into non-overlapping frames, and the schedule for each frame is repeated when data collection is periodic. Under this scheme, minimizing the data collection time for (aggregated/raw-data) convergecast is equivalent to minimizing the number of time slots required per frame, called the schedule length,i.e., aggregated/raw packets from the source nodes reach the sink.

The use of orthogonal codes to eliminate interference where nodes are assigned time slots from the bottom of the tree to the top such that a parent node does not transmit before it receives all the packets from its children. Degree constrained routing topologies are constructed to enhance the data collection rate, though it may not always lead to schedules that have low latency, as because the number of hops in a tree goes up as its degree goes down.

Another protocol TreeMAC, considers the differences in load at different levels of a routing tree and assigns time slots according to the depth, i.e., the hop count, of the nodes on the routing tree, i.e., nodes closer to the sink are assigned more slots than their children in order to mitigate congestion. The problem of minimizing the schedule length for rawdata convergecast on single channel is shown to be NP-complete on general graphs. Maximizing the throughput of convergecast can be done by finding a shortest-length, conflict-free schedule where a greedy graph coloring strategy assigns time slots to the senders and prevents interference. In another routing scheme called disjoint strips transmits data over different shortest paths in which the sink remains as the bottleneck, hence sending data over different paths does not reduce the schedule length.

III. PROPOSED SYSTEM

In this proposed model, we use two types of interference models for evaluation: the graph-based protocol model and the SINR based physical model. In the protocol model, the interference range of a node is equal to its transmission range. In the physical model, the successful reception of a packet from i to j depends on the ratio between the received signal strength at j and the cumulative interference given as

$$SINR_{ij} = \frac{P_i \cdot g_{ij}}{\sum_{k \neq i} P_k \cdot g_{kj} + \mathcal{N}},$$

For which a packet has been received successfully at j if the signal-to-interference-plus-noise ratio, SINRij, is greater than a certain threshold β.

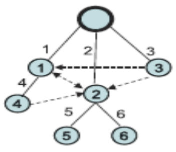
In this work, the aggregated convergecast has been studied in the context of periodic data collection where each source node generates a packet at the beginning of every frame, and raw-data convergecast for one-shot data collection where each node has only one packet to send.

Periodic Aggregated Convergecast

Data aggregation is a commonly used technique in WSN that can able to eliminate redundancy and minimizes the number of transmissions, thus saving energy and improves network lifetime. Data Aggregation can be performed in many ways such as

- suppressing duplicate messages.
- using data compression and packet merging techniques
- taking advantage of the correlation in the sensor readings.

The size of aggregated data transmitted by each node is constant and does not depend on the size of the raw sensor readings. Typical examples of such aggregation functions are MIN, MAX, MEDIAN, COUNT, AVERAGE, etc. The following fig. Illustrates the notion of pipelining with in aggregated converge cast and that of a schedule length on a network of six source nodes.



(a)

	Frame 1						Frame 2					
	S1	S2	S3	S4	S5	S6	S1	S2	S3	S4	S5	S6
s	1	2	3	-	-	-	{1,4}	{2,5,6}	3	-	-	-
1	-	-	-	4	-	-	-	-	-	4	-	-
2	-	-	-	-	5	6	-	-	-	-	5	6

(b)

Fig 2. Aggregated convergecast a) Schedule length of 6 in presence of interfering links b) Node ids in each slot over different frames

The solid lines represent tree edges, and the dotted lines can be represented interfering links. The numbers beside the links represent the time slots at which the links are scheduled to transmit, and the numbers inside the circles denote node ids. The entries in the table list the nodes from which packets are received by their corresponding receivers in each time slot.

In the aggregated convergecast, we first give a lower bound on the schedule length, and then propose a time slot assignment scheme that achieves the bound. Given the lower bound Δ(T) on the schedule length, the algorithm for time slot assignment scheme in the absence of interfering links is given as

Algorithm 1. BFS-TIMESLOTASSIGNMENT

1. Input: $T = (V,\ E_T)$
2. **while** $E_T \neq \phi$ **do**
3. $e \leftarrow$ next edge from E_T in BFS order
4. Assign minimum time slot t to edge e respecting adjacency and interfering constraints
5. $E_T \leftarrow E_T \setminus \{e\}$
6. **end while**

One-Shot Raw-Data Convergecast

One-shot data collection in which every sensor reading is equally important, aggregation might not be desirable or even possible. Each of the packets has to be individually scheduled at each hop en route to the sink. In raw data convergecast, the edges can be scheduled multiple times and there is no pipelining. If all the interfering links are eliminated using multiple frequencies, the only limiting factor in minimizing the schedule length is the half-duplex transceivers.

The key idea behind one shot raw-data convergecast is :

1) Schedule transmissions in parallel along multiple branches of the tree, and

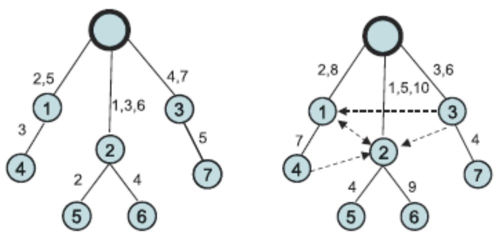
2) Keep the sink busy in receiving packets for as many time slots as possible.

The sink receives from the root of at most one top-subtree in any time slot, and also, the sink is aware of the number of nodes in each top-subtree. Each source node maintains a buffer and its related state, which can be either full or empty depending on whether it contains a packet or not.

Algorithm 2. LOCAL-TIMESLOTASSIGNMENT

1. *node*.buffer = *full*
2. **if** {*node* is sink} **then**
3. Among the eligible top-subtrees, choose the one with the largest number of total (remaining) packets, say top-subtree i
4. Schedule link $(root(i), s)$ respecting interfering constraint
5. **else**
6. **if** {*node*.buffer == *empty*} **then**
7. Choose a random child c of *node* whose buffer is *full*
8. Schedule link $(c, node)$ respecting interfering constraint
9. c.buffer = *empty*
10. *node*.buffer = *full*
11. **end if**
12. **end if**

We define a top-subtree to be eligible if its root has at least one packet to transmit. For a given time slot, we schedule the root of an eligible top-subtree which has the largest number of total (remaining) packets. If none of the top subtrees are eligible, the sink does not receive any packet during that time slot.



(a) (b)

Fig 3. Raw-data convergecast using algorithm LOCAL-TIMESLOTASSIGNMENT a) when interfering links are removed b) when interfering links are present

In wireless networks, excessive interference can be eliminated by using transmission power control i.e., by transmitting signals with just enough power instead of maximum power. TDMA schedule can support as many transmissions as possible in every time slot and has two phases: 1) scheduling and 2) power control. First, the scheduling phase searches for a valid transmission schedule, i.e., largest subset of nodes, where no node is to transmit and receive simultaneously, or to receive from multiple nodes simultaneously. Then, in the given valid schedule, the power control phase iteratively searches for an admissible schedule with power levels chosen to satisfy all the interfering constraints. On the other hand, if the maximum number of iterations has been reached and there are nodes which do not meet the interfering constraint, the algorithm excludes the link with minimum SINR from the schedule and restarts the iterations with the new subset of nodes. The power control phase is repeated until an admissible transmission scenario has been found.

IV. CONCLUSION

In this implemented work, fast convergecast is studied, where nodes communicate using a TDMA protocol to minimize the schedule length. The limitations due to interference, and half duplex transceivers on the nodes are been addressed . Also the studies show that the node-based (RBCA) and link-based (JFTSS) channel assignment schemes are more efficient in terms of eliminating interference.

Using half-duplex radios, after eliminating the achievable schedule length is lower bounded by the maximum degree in the routing tree for aggregated convergecast, and by max(2nk -1,N) for raw-data convergecast.

REFERENCES

[1] S. Upadhyayula and S.K.S. Gupta, "Spanning Tree Based Algorithms for Low Latency and Energy Efficient Data Aggregation nhanced Convergecast (DAC) in Wireless Sensor Networks," Ad Hoc Networks, vol. 5, no. 5, pp.626-648, 2007.

[2] T.Moscibroda,"The Worst-Case Capacity of Wireless Senso Networks," Proc. Int'l Conf. Information Processing in Sensor Networks (IPSN '07), pp. 1-10, 2007.

[3] O ¨. Durmaz Incel and B. Krishnamachari, "Enhancing the Data Collection Rate of Tree-Based Aggregation in Wireless Sensor Networks," Proc. Ann. IEEE Comm. Soc. Conf. Sensor, Mesh and Ad Hoc Comm. and Networks (SECON '08), pp. 569-577, 2008.

[4] Y. Wu, J.A. Stankovic, T. He, and S. Lin, "Realistic and Efficient Multi-Channel Communications in Wireless Sensor Networks," Proc. IEEE INFOCOM, pp. 1193-1201, 2008.

[5] V. Annamalai, S.K.S. Gupta, and L. Schwiebert, "On Tree-Based Convergecasting in Wireless Sensor Networks," Proc. IEEE Wireless Comm. and Networking Conf. (WCNC '03), vol. 3,pp. 1942-1947, 2003.

[6] X. Chen, X. Hu, and J. Zhu, "Minimum Data Aggregation Time Problem in Wireless Sensor Networks," Proc. Int'l Conf. Mobile Ad-Hoc and Sensor Networks (MSN '05), pp. 133-142, 2005.