Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.5, No.4, 2014

The Hybrid Algorithm for Data Collection over a Tree Topology in WSN

Shivendra dubey[1], Chetan Agrawal[2] and Anurag jain[3]

1. PG Scholar Department of Computer Science and Engineering, RITS, RGPV University, Bhopal, India
2. Asst. Prof. Department of Computer Science and Engineering, RITS, RGPV University, Bhopal, India
3. Head, Department of Computer Science and Engineering, RITS, RGPV University, Bhopal, India

[1]Shivendra.dubey5@gmail.com
[2]Chetan.agrawal12@gmail.com
[3]Anurag.akjain@gmail.com

Abstract

Wireless sensor networks have wide range of application such as analysis of traffic, monitoring of environmental, industrial process monitoring, technical systems, civilian and military application. Data collection is a basic function of wireless sensor networks (WSN) where sensor nodes determine attributes about a phenomenon of concern and transmits their readings to a common base station(sink node). In this paper, we use contention-free Time Division Multiple Access (TDMA) support scheduling protocols for such data collection applications over tree-based routing topology. We represent a data gathering techniques to get the growing capacity, routing protocol all along with algorithms planned for remote wireless sensor networks. This paper describes about the model of sensor networks which has been made workable by the junction of micro-electro-mechanical systems technologies, digital electronics and wireless communications. Firstly the sensing tasks and the potential sensor network applications are explored, and assessment of factors influencing the design of sensor networks is provided. In our propose work using data compression and packet merging techniques; or taking advantage of the correlation in the sensor readings. Consider continuous monitoring applications where perfect aggregation is achievable, i.e., every node is capable of aggregate the entire packets expected from its children as well as that generate by itself into a particular packet before transmit in the direction of its sink node or base station or parent node.

Keyword: Aggregation, Data Converge-cast, Data fusion, Energy Efficiency, Routing and TDMA.

Introduction

Data collection beginning a set of sensors to an ordinary sink (base station) over a tree-based routing topology is a basic traffic prototype in wireless sensor networks. This many-to-one communication pattern in which data flows from numerous nodes to a single node is known as converge cast. The objective of the data-gathering problem is to send the sensed data from all sensor nodes to a base station. One important thing is defined as the base station (sink node) collect data from every sensor node at only one time. The intention of algorithms which execute data gathering is to enlarge the total number of round of communication earlier than the nodes expire and the network becomes untreatable. This means least amount energy should be consumed and the transmission should take place among smallest delays, which are disputed requirements. Hence, the arrangement of energy with delay metric is used to compare dissimilar algorithms, because this metric measures speedy and energy-efficient data gathering [1]. The aggregate converge cast is appropriate when tough spatial association present in the data, or the principle is to collect summarized information such as the greatest sensor reading. On the further hand, raw data converge-cast is appropriate when every sensor reading is uniformly significant, or the correlation is negligible. As radio communication consumes a large fraction of power supply, it is critical to design energy-efficient routing algorithms for data aggregation in wireless sensor networks. In the existing works, the problem of designing energy-efficient routing algorithms has been extensively considered in both general multi-hop wireless networks, and the particular backdrop of sensor networks.

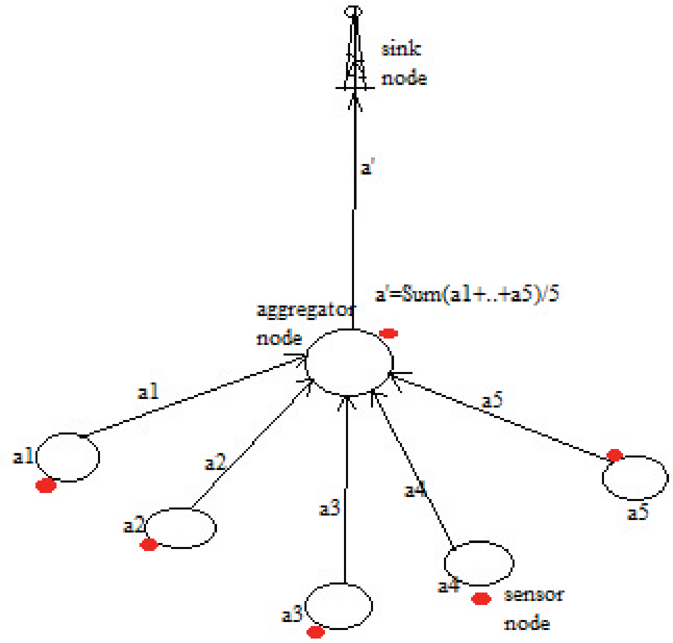

Figure 1: Data Converge-cast

In terms of design constraint and assumptions, the algorithms change mainly in the following dimensions: (a) make use of communication and interference models, (b) implementation method, such as centralized or distributed, (c) topology assumptions, (d) types of data aggregation, for example use of in-network processing vs. raw-data, and (e) capacity of transceivers obtainable on the sensor nodes. Under this setting, minimizing the data collection time for (aggregated/raw-data) converge-cast is equivalent to minimize the number of time slots necessary per frame, called the schedule length, such that every (aggregated/raw) packet starting the source nodes arrive at the sink. As multi-hop TDMA allow spatial use again of time slots, further than one node can transmit simultaneously if their receivers are in non-conflicting parts of the network. There are two types of conflicts that arise: (i) primary conflict, and (ii) secondary conflict. A primary conflict occurs when a node transmits and receives at the one time, or receive further one transmissions destined to it at the same time [4, 6]. A secondary conflict occurs when a node, an intended receiver of only one particular transmission, is also within the coverage of another transmission intended for other nodes. Figure 1 illustrates the data converge-cast technique; here we are representing five sensor nodes a1, a2, a3, a4 and a5, one aggregator node and base station. There are five sensor nodes sending the data to aggregator node for aggregate the data then finally transmit these data to the base station for convey to the another node for further processing[2].

1.1 TDMA – Protocol - Related Sensor Network Properties

In WSN, save the battery power is the most significant issue, by which we can enhance the lifetime of the network. Sensor nodes are assumed to be dead as they are out of battery. One more issue is quick and efficient query response, so that all change in the environment can be detect instantly. On the basis of the above characteristics, the planned TDMA protocol must be energy proficient by minimizing the potential energy waste (presented below) and sends the sense data to the sink without more delay. TDMA protocol decreases the data retransmission since collision does not happen in TDMA protocol. TDMA transmission in multi-hop networks frequently resolve the minimum length conflict free assignment of slots in which all link or node is activate at least once. This is base on theory; there is a lot of independent point-to-point flow in the network. For sporadic traffic, it is very well known that contention-free medium access control (MAC) protocols such as TDMA (Time Division Multiple Access) are better fit for fast data collection, since they can remove collisions and retransmissions and offer assurance on the conclusion time as oppose to conflict based protocol.

The rest of paper is the prepared as follow: In Section 2, we talk about related works. In Section 3, we describe problem definition, state our assumption, methodology. In Section 4, we explain the proposed work and also examine the lower bounds on schedule length for aggregated and raw data converge-cast, In Section 5; we focus present detailed simulation and result. In Section 6, explains the future works. In Section 7, finally we draw our conclusions.

1. Related Work

Large scale dense wireless sensor networks (WSNs) will be increasingly more deploy in special classes of applications for huge monitoring. Because of this reason their very high density of nodes, it is especially expected that information is together spatially and temporally connected can be detected by numerous nodes what can be exploited to save energy, a key feature on these networks. Here, we are able to further and suggest a latest algorithm, called Efficient Data Collection Aware of Spatial Temporal Correlation (EAST), which use straight route for forward the gathered data in the way of the sink node and completely develop both spatial and sequential correlations to perform near real-time data gathering WSNs [2].

In this work, we examine the following primary question - how fast can information be gathered from a wireless sensor network organized as a tree? Here we focused and can provide those who similar to deeper approaching for data collection and routing techniques in wireless sensor networks. The theoretical analysis shows that this algorithm can reach O (1) -approximation ratio for the linear networks. Here we also observe the impact of the delay-constraint on the worst-case used for the planar networks [5].

The prior broadcast trees are not appropriate for converge-cast; because converge-cast is repeal transmit process. The tree protocols construct an estimated load-balancing converge-cast tree. Furthermore, the modify algorithm dynamically adjusts the tree structure to avoid breaking tree links. The tree modifies only needs localized information and operations at the sensors [6].We categorize the algorithms according to their common design objectives, categorize the subsequent four as the most basic and most studied with respect to data collection in WSNs: (i) minimizing the schedule length, (ii) minimizing latency, (iii) minimizing energy consumption, and (iv)maximizing fairness [13]. For an application which is concentricity in nature such as an environment controlling system or health controlling system, the tree based data aggregation is the mainly proficient technique, in condition of energy as well as data collection time [14].

We present a competent data reporting control method in a cluster-based hierarchical wireless sensor network, which has two mechanisms: (i) intra-cluster data reporting control (Intra-DRC) process and (ii) inter-cluster control (Inter-DRC) process. The Intra-DRC process controls the amounts of traffic generate in a cluster by select a definite figure of data coverage nodes base on the chosen throughput particular by the end system. In other word, the Inter-DRC structure offers differentiate coverage paths foundation from a cluster to a sink node base on the traffic distinctiveness [16]. A data gathering tree is usually produced as a sub network of a wireless sensor network. Power maintenance is of fundamental significance in such networks, and with periodic sleep–wake cycle for sensor nodes is one of the most proficient methods for power maintenance. The difficulty of scheduling sleep–wake cycle of nodes in a data gathering tree under deadline constraint [17]

2.1 Network Topology and Communication Pattern

Kulkarni defines three type of communication pattern in infrastructure less or wireless sensor networks [6]; (a). Broadcast, (b).Converge-cast, and (c).Local gossips. In broadcasting, any information is transmitted to all sensor nodes. If any update happens in sensor program, query of sensor query-processing architectures, or manage packets which are used by the whole system or transmitted by the base station or sink. In converge-cast communication model, groups of specific sensor nodes they have require of communication transmit data to a sensor node. The destination node might be a cluster head, a data fusion center, or a sink node (a base station). In cluster based networks, cluster heads converse with their own members, these every members may be everyone the neighbors' of cluster head or a separation of neighbors of cluster head. This type of communication pattern is called multicast communication pattern.

2.2 Data Collection Technique

In WSNs', Networks are divided into three categories-

2.2.1 Tree-based Technique

In tree based structure [4], the root of the tree is called a sink node. Figure 2 illustrates the tree based technique in which the sink node, monitored data by the sensor node is composed where continuous query are posed and processed. In WSN tree data run beginning leaf nodes to parent node or sink node (base station) during the intermediate nodes. These intermediate nodes are call as aggregator node. These aggregator nodes obtain data from its child nodes. Aggregator nodes procedure and add their individual monitored data and send to its root node (parent node).

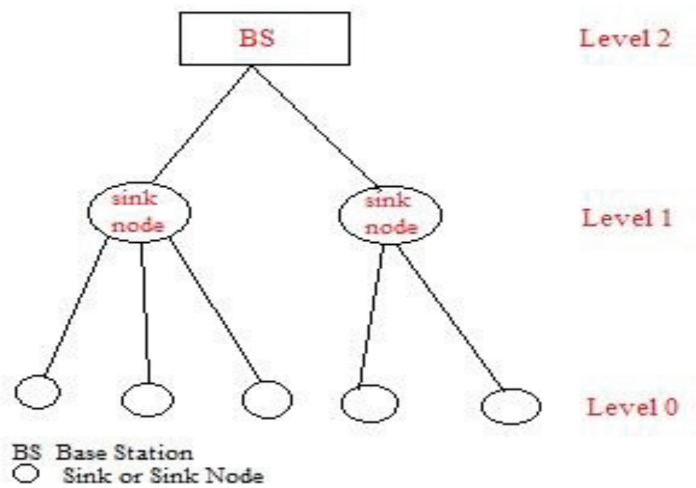

Figure 2: Tree-based techniques

2.2.2 Cluster-based Technique

Figure 3 demonstrate the cluster-based technique [5] in which all clusters have a cluster head and the entire node in this cluster transfer data to the cluster head (CH). This CH transfer or transmit data to CH of one more cluster which is adjacent to sink node or base station or parent node.

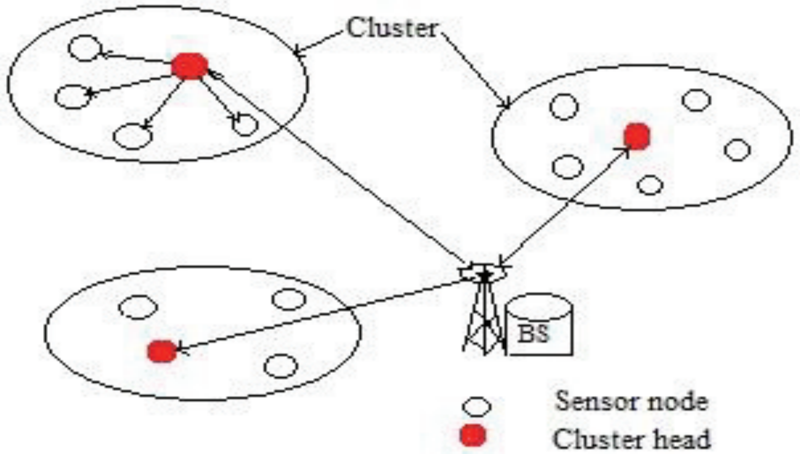

Figure 3: Cluster-based techniques

2.2.2.1. Low-Energy Adaptive Clustering Hierarchy (LEACH)

LEACH is clustering-based protocols that minimize energy dissipation in wireless sensor networks. LEACH randomly decide nodes as cluster-heads and perform sporadic reelection, that's why the rich dissipation adequate by the cluster-heads in communicating by the BS is increase across every nodes of network. All iteration; choice of cluster-heads are called a round [1, 6]. The purpose of LEACH is separated into two phase: set-up and steady. In the set-up phase, every sensor nodes choose an arbitrary number between 0 and 1. If this is fewer than the threshold for nodes n, T (n), the sensor node occur a cluster-head. The threshold T (n) is defined as

$$T(n) = \begin{cases} \frac{P}{1-P\{r*mod(\frac{1}{P})\}} & ; \text{if } n \varepsilon M \\ 0 & ; \text{if not} \end{cases}$$

Where P is the preferred grow of nodes which are cluster-heads, r is the existing around, and M is the group of nodes that has not been cluster-heads in the long-ago 1/P rounds. This insures that all sensors nodes finally spend equivalent energy. After choice, the cluster-heads announce their collection to each and every one node. The entire nodes select their adjacent cluster-heads after that they collect advertisement based on the acknowledged signal potency. The steady segment is of longer time in order to reduce the overhead of cluster pattern. In the steady segment, data communication takes place depended on the TDMA schedule and the cluster-heads execute data aggregation/fusion throughout the local computation [10]. The base station collects only aggregated data from cluster-heads, essential to energy conservation. Once a few period of time in the steady segment, cluster-heads are selected another time during the set-up segment [7].

2.2.3 Multipath or Hybrid Technique

The drawback of tree based technique is the damaged robustness of the system. To overcome this drawback, a latest technique was planned by a lot of researcher's in which sending partly gather data to only parent node in aggregation tree, a node might send data more several paths. Fig. 4 illustrates multipath technique, in which every node is able to send data packets to it's possibly numerous neighbors [8]. Therefore data packet flow starting source node to the root node all along several paths, a lot of middle node among leaves node to root node so gathering done in each middle node. The illustration of this technique similar to ring topology, wherever network is separated in to concentric circle with defining level levels according to hop distance from root [1].

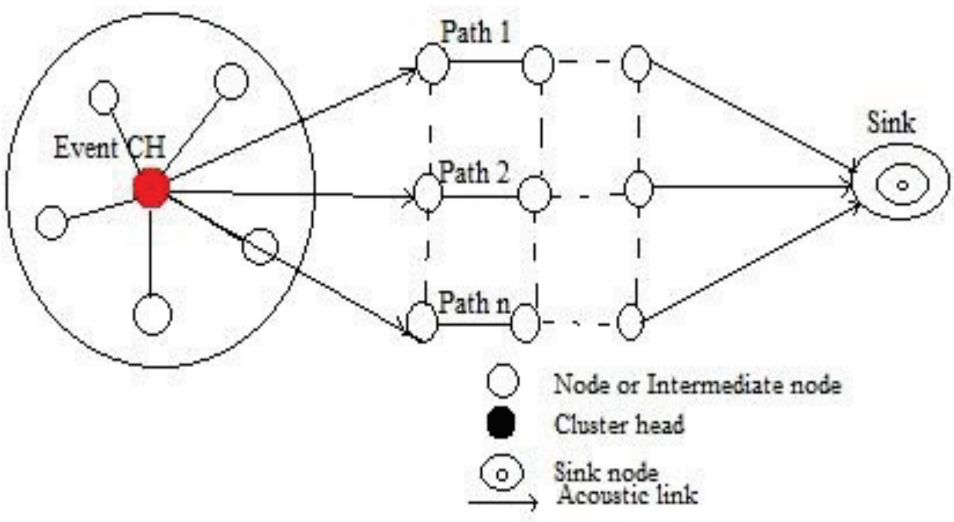

Figure 4: Multipath-based techniques

3. Problem Definition

3.1 Causes of Energy Waste

In WSNs', if the network has limited amount of non-overlapping channel than a dissimilar channel is not assign to every node on the data aggregation tree. Therefore data aggregations introduce data retransmission which is cause by co-channel interference from neighboring sensor nodes and extra battery power is mandatory for retransmission of the data [3]. When a node receive more than one packet at the similar time then these packets collide with each other and retransmissions of these packets are mandatory, which increase the energy utilization. One of the main source of energy waste be idle listening. An idle listening occur when a node listen an idle channel in organize to get probable traffic. The third reason for energy waste is overhearing [2]. Overheating occur when a node receives packets which are going to another destination. The fourth reason of energy wastage is transmit after redundant data packet or else redundant data to the destination.

3.2 Classification Approach and Methodology

Our work has the common purpose of "data collection using TDMA-based communication schedule in WSNs". Our categorization method is primarily based on the design objectives and design issues.

3.2.1. Design Objectives

- *Minimizing Schedule Length*: Minimize the schedule length, or equivalently, minimize the time to whole converge-cast, is the most-studied design objective of data collection in sensor networks. In numerous WSN applications, it is of importance to maximize the rate at which the sink can collect data from the network [5]. For example, it is well-known that in networked structural health monitoring, other than 500 samples per second are mandatory to efficiently detect and localize damages [9]. Minimal schedule length can be achieved by maximizing the reuse of the time slots. Therefore, most of the existing algorithms aim to maximize the amount of parallel transmissions and enable spatial reprocess by devising strategies to eliminate interference.

- *Minimizing Latency*: Minimize the data collection latency is significant for applications that are necessary to take certain (precautionary) actions base on deadlines, such as mission-critical and event-based application. Although most algorithms do not consider the standard latency experienced by individual packets at all hop. For example, a line topology might allow higher spatial reprocess, but due to the larger sum of hops from the sources to the sink (base station), it may cause high latency.

- *Minimizing Energy Consumption:* As a major source of energy consumption, radio activity should be managed efficiently. Transmission power control is another well-known technique to decrease energy utilization, contention, and packet losses.

  - *Maximizing Capacity:* Though maximizing the throughput capacity is not measured to be one of the primary objectives of low-rate data collection application over under sized networks, it is significant for large, dense sensor networks and for difficult applications that want efficient delivery of large amounts of Data [15].

  - *Maximizing Fairness:* Fairness is the main objective of WSN application in order to maintain a balanced vision of the sensor environment. In application where all of the sensor readings are significant, fairness becomes a main issue, especially below high data rates. For instance, fair data gathering maybe come necessary for reducing the estimation error in an application involving field reconstruction.

  - *Other Objectives:* Minimize communication costs maximize parallel transmissions, meeting deadlines, minimizing interference and self stabilization are some of the other objectives considered for data collection in WSNs.

  - *Joint Objectives:* In mainly applications, there is not forever a particular objective, but often multiple, and at times conflicting, objectives involved. Some examples include minimize communication cost and delay, minimize energy consumption and completion time of data collection, maximize capacity and falling energy, etc.

3.2.2 Design issues

The issues that make sensor networks a distinct category of ad hoc wireless networks are the following [12]:

- *Mobility of nodes:* Mobility of nodes is not a fixed constraint in sensor networks. For instance, the nodes deploy for sporadic monitoring of earth properties are not necessary to be mobile. However, the sensor nodes that are fixed on the bodies of patients in a post surgery ward of a hospital might be designed to maintain partial mobility.

- *Size of the network:* The total number of nodes within the wireless sensor network be able to much greater than that in a distinctive ad hoc wireless network.

- *Density of deployment:* The densities of nodes in a wireless sensors network vary with the area of application. For instance, military applications need high accessibility of the network, making redundancy a high priority.

- *Power constraints:* The power constraints in wireless sensor networks are a lot of inflexible than those within ad hoc wireless networks. This is primarily because the sensors nodes are predictable to activate in insensitive environmental or geographical situation, through minimum or no human direction and maintenance. The power source used in wireless sensor networks can be classify the subsequent three categories:
  - *Replenish-able power source:* In certain application of wireless sensor networks, the power resource can be replaced when the presented resource is completely worn out.

  - *Non-replenish able power source:* In the some particular applications of wireless sensor networks, the power resource cannot be replenished one time the network has been deployed.

  - *Regenerative power source:* Power resources employed in wireless sensor network that belong to this class have the capacity of regenerate power from the physical limitation under dimension. For instance, the sensor employed for sensing temperature at a power plant can use power source that can produce power by using suitable transducers.

  - *Data/information fusion:* The restricted bandwidth and power constraints require aggregation of bits and information at the middle relay nodes that are responsible for relaying. Data fusion refers to the aggregation of more than a few packets into single earlier than relaying it. These primarily aim at reducing the bandwidth consumed by outmoded headers of the packets and minimizing or reducing the media access delay concerned in transmits multiple packets.

  - *Traffic distribution:* The communication traffic models vary by the domain of application in wireless sensor networks. For example, the environmental sensing applications generate small sporadic packets representing the position of the environmental parameter under study to a fundamental monitoring station. This class of traffic demands low bandwidth. The sensor network employed in detecting border instructions in a military application generates traffic on detection of certain events; in most cases these events might have time constraints for delivery. In contrast, ad hoc wireless networks or infrastructure less network normally take user passage such as digitized and packetized voice stream or data traffic, which difficulty in advanced bandwidth.

4. Proposed Work

4.1 Periodic Aggregated Converge-cast

Data aggregation is a typically used method in WSN that can eliminate redundancy and reduce the number of transmission, hence save energy and improving network lifetime [11]. Aggregation can execute in numerous ways, such as by suppressing duplicate messages; using data compression and packet merging techniques; or taking advantage of the correlation in the sensor readings. Consider continuous monitoring applications where perfect aggregation is achievable, i.e., every node is accomplished of aggregating the entire packets established from its children as well as that generate via itself into a only packet before transmitting to its parent.

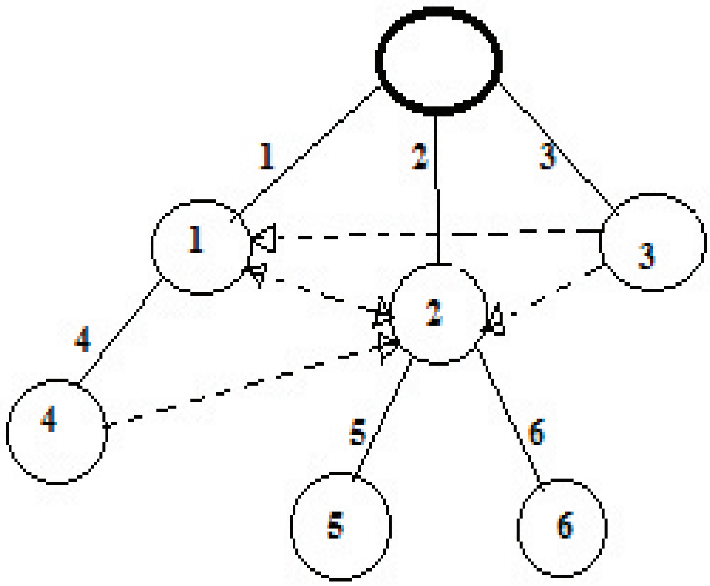

Figure 5: Schedule length of 6 in the presence of interfering links

In Fig.5 illustrates the notion of pipelining in aggregated converge-cast and that of a scheduled length on a network of 6 source nodes. The solid lines stand for tree edges, and the dotted lines stand for interfering links. The numbers beside the links stand for the time slots at which the links are scheduled to transmit, and the numbers within the circles denote node ids.

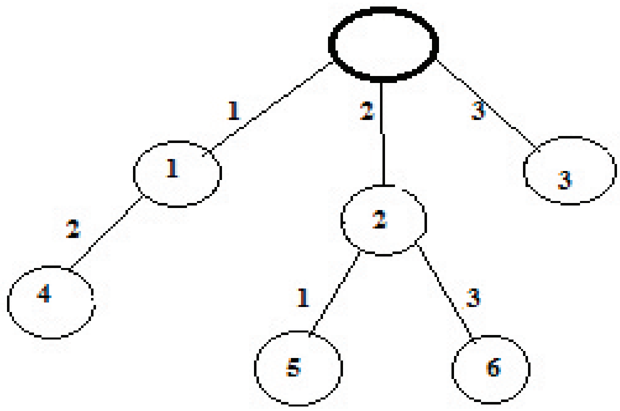

Figure 6: Schedule length of 3 using BFS TIME SLOT ASSIGNMENT when all the interfering links are eliminated for the aggregated converge-cast

The entries in the table the nodes from which packets are received by their subsequent receivers in each time slot. Note that at the ending point of frame 1, the sink (base station) does not enclose packets from nodes 5 and 6; though, as the schedule is repetitive, it receive aggregated packets from 2, 5, and 6 in slot 2 of the subsequent frame. Similarly, the sinks (base station) also receive aggregated packets from nodes 1 and 4 initial from slot 1 of frame 2. The entry {1, 4} and {2, 5, 6} in the table represent single packets comprising aggregated data from nodes 1 and 4, and opening nodes 2, 5, and 6, correspondingly. Thus, a pipeline is accepted from frame 2, and the sinks (base station) continue to collect aggregated packets from all the nodes every 6 time slots. Thus, the minimum schedule length is 6.

4.1.1. Lower Bound on Schedule Length

First consider aggregated convergecast when the entire interfering links are eliminated by use transmission power control or several frequencies. Although the problem of minimizing the schedule length is NP complete on general graphs, shows in the following that once interference is eliminated, the problem reduces to one on a tree, and can be solved in polynomial time. To this end, first give a lower bound on the schedule length, and then recommend a time slot assignment scheme that achieves the band.

Discription1: If every interfering links is eliminated, the schedule length for aggregated convergecast is lower bounded by Δ (T), where Δ (T) is the maximum node degree in the routing tree T.

Proof: If every interfering links is eliminated, the scheduling problem reduces to one on a tree. Now since each of the tree edges desires to be scheduled just once inside each frame, it is correspondent to edge coloring on a graph, which requirements of a number of colors at least equal to the maximum node degree. Once all the interfering links are eliminated, concurrency is still inadequate by the adjacency constraint due to the half-duplex transceivers, which prevent a parent from transmitting when it is already getting from its children, or whenever its parent is transmitting.

4.1.2. Assignment of Timeslots

Given the lower bound Δ (T) on the schedule length in the absence of interfering links, now present a time slot assignment scheme in algorithm 1, called BFS TIME SLOT ASSIGNMENT or SHORTEST PATH ALORITHM; that achieves this band. In each iteration of BFS-TIME SLOT ASSIGNMENT (line 2-6), an edge e is choose in the Breadth First Search (BFS) starting from one node, and is assign the lowest time slot that is different from all its closest edges respecting interfering constraint. Note that, since evaluate the performance of this algorithm also for the case when the interfering links are present, check for the corresponding constraint in line 4; whenever, interference is eliminated this check is redundant. The algorithm runs in O (|ET|2) time and minimizes the schedule length when there are no interfering links, as prove in Description 1. To illustrate, Fig. 6 in which all the interfering links are removed, and thus the network is scheduled in 3 time slots. Although BFS-TIME SLOT ASSIGNMENT may not be an estimate to ideal scheduling under the physical interference model, it is a heuristic that can get the lower bound if all the interfering links are eliminated. Therefore, together with a method to remove interference the algorithm can optimally schedule the network.

4.2. Raw Data Converge-cast

In this section, we consider the one-shot data collection where every sensor reading is evenly important, and aggregation may not be desirable or still possible. Thus, every packet has to be individually scheduled at every hop en route to the sink or base station. Since before, we focus on minimizing the schedule length. In other case like periodic aggregated convergecast somewhere a pipelining get place and each tree edges is programmed only one time within every frame, now the edges could be scheduled several times and there is no pipelining.

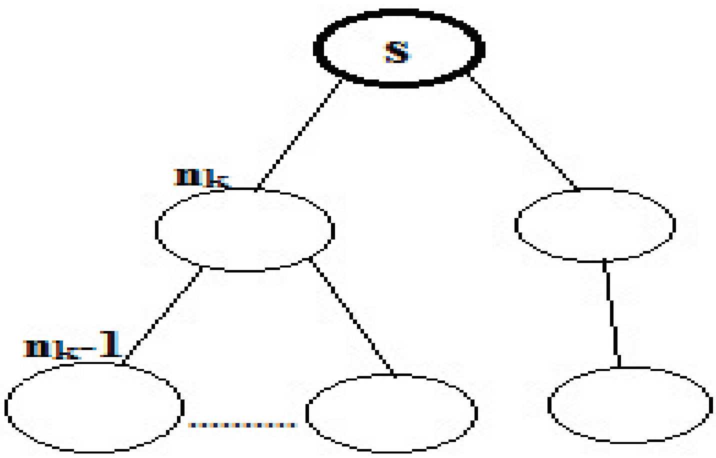

Figure 7: Raw-data Converge-cast: Largest top-sub tree with no nodes

The problem of minimize the scheduling length for raw-data converge-cast is prove to be NP-complete even below the protocol interference model by a decrease from the well known Partition Problem. Before getting into the details, we first explain the following terms: a branch is definite as a sub tree containing the sink as an endpoint; a top-sub-tree is definite as sub trees that have a child of the sink as its root. For example, in Fig. 7, the branches are {s, 1, 4}, {s, 2, 5, 6}, and {s, 3, 7}, while the top-sub trees are {1, 4}, {2, 5, 6}, and {3, 7}.

4.2.1. Lower Bound on Schedule Length

As mentioned in Section 4.1.1, if all the interfering links are eliminated using many frequencies, the only limiting factor in minimizes the schedule length is the half duplex transceivers. In the following, we provide a lower bound on the schedule length under this scenario.

Discription2: If every interfering links is eliminated, the schedule length for one-shot raw-data convergecast is lower bounded with max $(2n_k - 1, N)$, where n_k is the maximum number of nodes in some top-subtree of the routing tree, and N is the number of sources in the network.

Proof: If all the interfering links are eliminated, the scheduling problems reduce to one on a tree. Now since all tree edges desires to be scheduled only one time within each frame, it is correspondent to edge coloring on graph, which desires number of colors at least equivalent to the maximum node degree. Once all the interfering links are eliminate, concurrency is still restricted by the adjacency control due to the half-duplex transceiver, which prevent from transmitting when it is already receiving from its children, or as its parent is transmitting.

4.2.2. Assignment of Timeslots

We currently explain a time slot assignment method in Algorithm 2, called DATA AGGREGATION ALGORITHM, which is run nearby through each node at each time slot. The key idea is to: (i) schedule transmissions in similar along several branches of the tree, and (ii) keep the sink active in getting packets for as a lot of time slots as possible. We imagine that the sink is aware of the amount of nodes in all top-sub trees. All source nodes maintain a buffer and its related state, which may be either full or empty depending on whether it contain a packet or not. Our algorithm does not involve any of the nodes to store more than one packet in their buffer at any time. We initialize each buffer as full, and imagine that the sink's buffer is forever full for the easiness of explanation.

Algorithm 1: Shortest Path Algorithm

The shortest path algorithm is used to calculate the path between the sink and the source. Given a network G = (N,E), with a positive cost Dij for every edges (i, jϵN), starting node S and a set P of permanently label nodes, the shortest path from starting node S to each other node j .

1. Initially P = {S}, DS = 0, and Dj = DSj for j ε N
2. Find i ε P such that Di = Min Dji ε P
3. Set P = P U {i}
4. If P contains all nodes then
5. Stop the algorithm is complete.
6. For all j ε P
7. Set Dj = min[Dj,Di + Dij]
8. Go to step 1.

Algorithm 2: Data aggregation algorithm

This algorithm has no transmission overhead and is computationally simple. Because both the sender and the receiver have the first dictionary and every new dictionary entries are created based on presented dictionary entries, the recipient can recreate the dictionary on the fly as data is established.

1. Send start code
2. For each character {
3. If new entry appended with character is not
4. In dictionary
5. {
6. Send code for new entry

7. Add new entry appended with character as
8. New dictionary entry
9. Set new entry blank
10. }
11. Appended character to new entry
12. Send code for new entry
13. Send stop code.

- Input

A = {a1, a2,…., an} - symbol of alphabet size n
W = {w1, w2,….wn} - set of symbol weights.
i.e wi = weight (ai), 1 < i < n

- Output

C (A, W) = {c1, c2,…., cn} - set of binary codeword where ci is the codeword; for ai, 1 < i < n
Let L(C) = i=1Σn wi x length (ci) be the weighted pathway length of code C.
The condition is L (C) < L (T) for any code T (A, W)
A node can be there a leaf or an inner node. Primarily, each node is leaf nodes; internal nodes have symbol weight, links toward two child nodes and the probable link to a parent node. As an ordinary convention, bits '0' represent follow the left child and bit '1' represent follow the right child. A completed tree has N leaf nodes and N−1 internal nodes. The resulted compressed data is again compacted using compression technique and then the data are transmitting to the sink node where in the data are decompressed.

Algorithm 3: Scheduling algorithm for multi-line network

Suppose N represent the total number of nodes in the network and n_k represent the greatest number of nodes in a branch, then the number of timeslots required by our convergecast scheduling algorithm for multi-line networks is given by maxð3n_k -1; $N_Þ$. The pseudo code for scheduling in tree network is the same as the one for multi-line networks, excluding that the dynamic condition used for a node has to be modified to: t < t_I and D_{nI} P W v.

Notation
1. Q: transmission queue
2. I 2 f1; . . . ; kg: branch ID of this node
3. S 2 fT ; I; Rg: current state
4. N: ½n1; n2; . . . ; nk : number of packets in each branch
5. t: current timeslot
6. T : ½t1; t2; . . . ; tk : last active timeslot

Initialization

7. S set according to hop-count
8. N set with initial numbers given by the base station
9. t = 1
10. ti ¼ 08i 2 f1; 2; . . . ; kg
11. at timeslot t:
12. Let L ¼ fijti < tg; set of eligible branches to transmit at t
13. Let j ¼ arg maxfniji 2 Lg be the first branch in L that

has the maximum number of nodes left

14. nj maxðnj 1; 0Þ; branch j will transmit at timeslot t
15. tj tj þ 2; branch j will not be eligible for next two timeslots
16. if t 6 tI; this branch is active
17. switch S
18. case T: transmit one packet from Q
19. case I: idle

```
case R: push the received packet into Q
end switch
S nextðSÞ; according to Figure 2
end if
t t þ 1
if maxðNÞ ¼ 0 and t > maxðTÞ
initialization
end if
```

Where, D_{nI}¼n 0 I_{nI} is the number of packets forwarded by one-hop-sub tree I to the base station before timeslot t. We conclude that the number of timeslots required by our converge-cast scheduling algorithm for tree networks is at most maxð3nk1; NÞ (Algorithm 2). Here, n_k represents the number of nodes in the largest one-hop-sub tree. During multi-line networks employ sleep schedule outcome in rather additional energy savings. Consider multi-line network of N nodes and k branches. Suppose that the total number of nodes in every branch is equivalent. Note that convergecast in this network can be proficient in N timeslots (Algorithm 2).

5. Simulation and Result

5.1 Required Number of Channels

In this section, for the different channel assignment method, the dissimilar channel assignment methods, we calculate the required total amount of channels to entirely remove interference as a function of consumption density. In our simulation results as shown in Figure 8, we assume that the number of available channels is unlimited so as to show the upper bounds. TMCP require various channels as every branch is on a special channel. This is costly for deployments where a group of nodes can straightly attach to the sink, and therefore are assigned special channels because they form special branches. Thus, one needs to optimize the channel discussion.

5.2 Aggregated Converge-cast on Degree-Constrained Trees

Figure 9 shows the variant of schedule length by density when the maximum tree degree is 3 (in sparser scenario, through a maximum degree of 2, it was not forever possible to make connected topologies). The peak two lines are for nodes transmit at highest power, and nodes with power control.

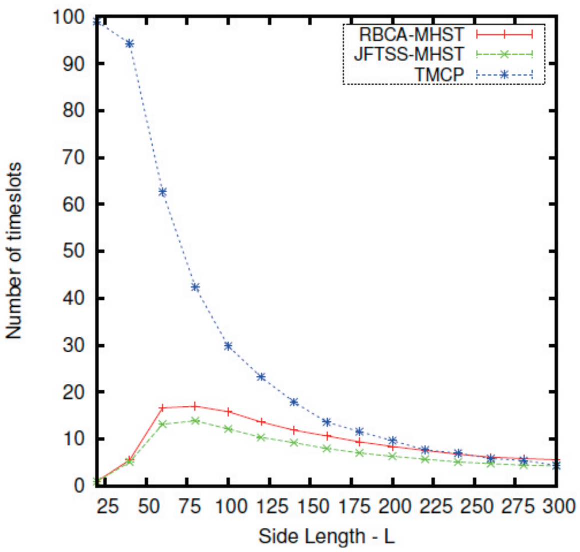

Figure 8: Bounds on the number of frequencies

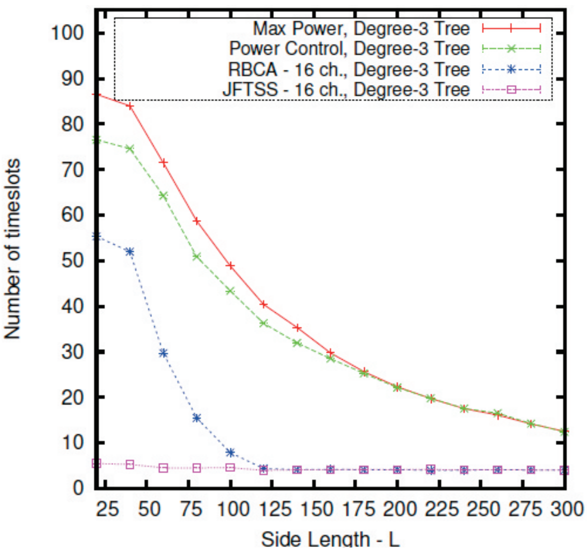

Figure 9: scheduling on degree-constrained

Sensor nodes around 100 are uniformly distributed over a 1000m×1000m area. Initially, 10 Joules of energy is assign to each node and then we add the network with 1000 randomly generate packets. The values of parameter intended for simulations are as shown in Table 1. The source and destination of each packet are randomly select and the sizes of packet are drawn from a uniform allotment between 1 and 100 units. The effective radio range is 250 meters. The long distance path loss model is used and the path loss exponent is put to 4.0. Data packets are generate at gap of 1 second.

Parameters	Value	Parameters	Value
Bandwidth	2Mb/s	Transmission range	250 m
Transmit power	660mW	Topology Size	1000m x1000m
Receive power	395mW	Number of sensors	100
Idle power	35mW	Packet rate	5 packets/s
Initial energy in batteries	10 Joules	Packet size	512 bytes

Table 1: Assumed parameters

The simulation is run for 750 seconds therefore each protocol has enough time to discover the route from the sink to the source and produce substantial amount of data traffic. The power consumption for transmitting one unit of data is 660mW, the power consumption for data reception is 35mW and the power consumption in the idle mode i.e., when the sensor node is not in the sending or in the receiving and data is 35mW.

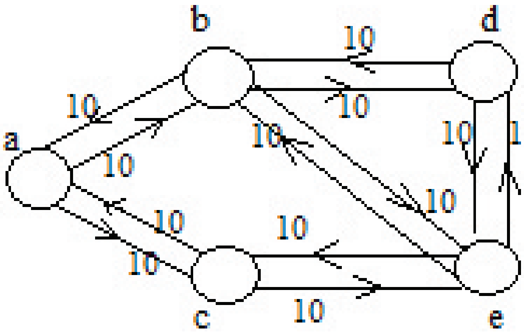

Figure 10: Routing Tree (a)

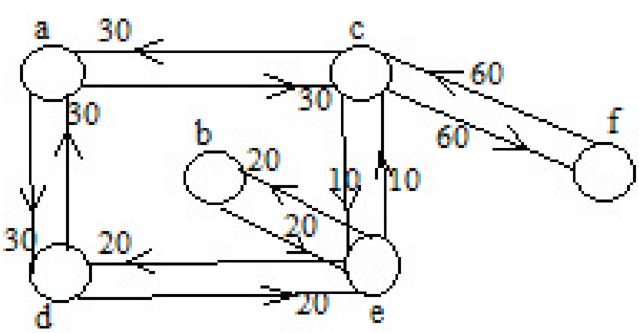

Figure 11: Routing Tree (b)

In this section, we consider two different routing trees, and also calculate the shortest path with the help of BFS-Time Slot Assignment algorithm. In Figure9 and Figure10 are represent the two routing trees on which applying BFS-Time Slot

	A	B	C	D	E	F	G
Routing1	-	10,B	10,C	30,E	20,C	N/A	N/A
Routing2	-	60,E	30,C	30,D	40C	90,C	N/A

Table 2: Node ids from which (aggregated) packets are established by their equivalent parents in each time slot over dissimilar frames

Assignment scheme to calculate the shortest path from source node to destination node with the help of measured distance for intermediate node. In Table 2 we represent the distance between various nodes.

6. Future Work

Although we evaluated our proposed algorithms using real implementation on sensor nodes where schedules are computed locally and are adaptive to network dynamics are necessary to enhance the operation of sensor networks and to meet application requirements. For instance, we observe a trend in using WSNs to support additional difficult operations range from developed control to health care, which necessitate difficult operations similar to exposure of actions in real-time or responsive query of the network by collect streams of data in a suitable approach. Thus, supporting QOS metrics such as delay and reliability become more important. Therefore, distributed implementation and performance testing of the proposed algorithms on tested or real deployments becomes essential. Additionally, real implementation and consumption will help in addressing the problems of alternating connectivity and channel errors with untrustworthy links and handling asymmetric links

7. Conclusions

In this work, it is considered that the fast converge-cast in WSN where node communicate via a TDMA protocol to reduce the schedule length. The basic boundaries due to interference and half-duplex transceivers on the nodes and investigate techniques to overcome the similar problem. It is originate that while transmission power control helps in reducing the schedule length, a variety of channels are more effective. We also observed that node-based (RBCA) and link-based (JFTSS) channel assignment schemes are extra efficient in terms of eliminate interference as compared to assigning dissimilar channels on dissimilar branches of the tree (TMCP).Once interference is totally eliminated, it is proved that with half-duplex radios the reachable schedule length is lower-bounded by the maximum degree in the routing tree for aggregated converge-cast, and by max $(2n_k - 1, N)$ for raw-data converge-cast. Here, we are using best possible converge-cast scheduling algorithms; it is revealed that the lower bounds are feasible once a suitable routing scheme is used. Through general simulations, we verified up to an order of magnitude decrease in the schedule length for aggregated, and a 50% reduction for raw-data converge-cast. In the future, we will investigate scenarios with variable amounts of data and implement and calculate the combination of the schemes considered.

7. REFRENCES

[1]Shivendra Dubey, and Chetan Agrawal, et al, “A survey of data collection techniques in wireless sensor network,” *IJAET Journal of Advances in Engineering and Technology*, vol. 6, issue 4, 2013.

[2]DurmazIncel, Ozlem, et al, "Fast data collection in tree-based wireless sensor networks," *Mobile Computing, IEEE Transactions*, pp. 86-99., 2012.

[3]Incel, O. Durmaz, and BhaskarKrishnamachari, et al, "Enhancing the data collection rate of tree-based aggregation in wireless sensor networks," *Sensor, Mesh and Ad Hoc Communications and Networks, 2008. SECON'08.5th Annual IEEE Communications Society Conference on*. IEEE, 2008.

[4]M. C. Rodríguez-Sánchez, S. Borromeo, *Member, IEEE,* and J. A. Hernández-Tamames, ”Wireless Sensor Networks for Conservation and Monitoring Cultural Assets,” *IEEE Sensors Journal,* vol. 11, no. 6, 2011.

[5]F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: A survey,” *Computer Networks,* vol. 38, no. 4, pp. 393– 422, 2002.

[6]Jonathan Henderson .”A Study of the LEACH Protocol for Wireless Sensor Networks.”Mississippi Valley State University, IttaBena, Mississippi 38941.

[7]Wendi RabinerHeinzelman, AnanthaChandrakasan, and HariBalakrishnan, ”Energy-Efficient Communication Protocol for Wireless Microsensor Networks,” 33rd Hawaii International Conference on System Sciences, 2000.

[8]Hu¨seyin O¨ zgu¨r Tan and IÚbrahimKo¨rpeogßlu “Power Efficient Data Gathering and Aggregation in Wireless Sensor Networks”, Department of Computer Engineering, Bilkent University.

[9]W. Heinzelman, A. Chandrakasan and H. Balakarishnan, “Energy-Efficient Communication Protocols for Wireless Microsensor Networks,” *Proceedings of the Hawaaian International Conference on Systems Science*, 2000.

[10]S. Lindsey and C. S. Raghavendra, “Pegasis: Power-efficient gathering in sensor information systems,” in *IEEE Aerospace Conference,* March 2002.

[11]John A. Stankavic, Tarek F. Abdelzaher, Chenynng Lu, LuiSha and Jennifer C. Hou, “Real-Time Communication and Coordination in Embedded Sensor Networks”, Proceedirigs of the IEEE*,* V01.91.No.7. pp. 1002-lM2, 2003.

[12]C. Siva Ram Murthy, B. S. Manoj, “Ad Hoc Wireless Neyworks”, Pearson Education, Inc. and Dorling Kindersley Publishing Inc., pp.647-666, First impression 2007.

[13]Madden, S., et al., “TAG: A Tiny Aggregation Service for Ad-Hoc Sensor Networks”, 2002: OSDI.

[14]Karvonen H., Shelby Z., and Pomalaza-Raez C., “Coding for Energy Efficient Wireless Embedded Networks,” in Proceedings of International Workshop on Wireless Ad Hoc Networks, Finland, pp. 300-304, 2004.

[15]W. Heinzelman, A. Chandrakasan and H. Balakarishnan, “Energy-Efficient Communication Protocols for Wireless Microsensor Networks,” Proceedings of the Hawaaian International Conference on Systems Science, January 2000.

[16] Choe, Hyun Jung, Preetam Ghosh, and Sajal K. Das, “QoS-aware data reporting control in cluster-based wireless sensor networks,” Computer Communications, pp.1244-1254, 2010.

[17] Jang, Ungjin, Sunggu Lee, and Sungjoo Yoo, "Optimal wake-up scheduling of data gathering trees for wireless sensor networks", Journal of Parallel and Distributed Computing, pp. 536-546, 2012.