Data Aggregation with Fast Convergecasting in Wireless Sensor networks

By

Vinit Ambwani
[10MCEC01]

NIRMA UNIVERSITY
INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
AHMEDABAD-382481

March 2012

Data Aggregation with Fast Convergecasting in Wireless Sensor Networks

Major Project II

Submitted in partial fulfillment of the requirements

For the degree of

Master of Technology in Computer Science and Engineering

Vinit Ambwani

10MCEC01

GUIDED BY :

Prof. Gaurang Raval

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD

DECLARATION

I, **Vinit Anil Ambwani**, **10MCEC01**, give undertaking that the Major Project entitled **"Data Aggregation with Fast Convergecasting in Wireless Sensor Networks"** submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in Institute of Technology of Nirma University, Ahmedabad, is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Vinit Ambwani

DEDICATION

This thesis is dedicated to my Grandfather, who has been an inspiration to me. He has taught me to never give up and believe in oneself. The most important lesson that I have learnt from him is to work hard irrespective of situations and surroundings and victory will be yours.

I would also like to dedicate this thesis to my family, who have encouraged me in the tenure of 2 years of pursuing my post graduation. They have taken care of all my needs and provided me with constant support and encouragement.

CERTIFICATE

This is to certify that the Major Project, entitled "Data Aggregation with Fast Convergecasting in Wireless Sensor Networks", submitted by Mr. Vinit Anil Ambwani [10MCEC01], towards the partial fulfillment of the requirements for the degree of Master of Technology in Computer Science and Engineering of Nirma University of Science and Technology, Ahmedabad is the record of work carried out by him under my supervision and guidance. In many opinion, submitted work has reached a level required for being accepted for examination. The result embodied in this major project, to the best of my knowledge,haven't been submitted to any other university or institution for award of any master degree.

Prof. Gaurang Raval
Guide, Associate Professor,
Department of Computer Engineering,
Institute of Technology,
Nirma University, Ahmedabad

Dr. S.N. Pradhan
Professor,P.G. Coordinator,
Department of Computer Engineering,
Institute of Technology,
Nirma University, Ahmedabad

Prof. D. J. Patel
Professor and Head,
Department of Computer Engineering,
Institute of Technology,
Nirma University, Ahmedabad

Dr. Ketan Kotecha
Director,
Institute of Technology,
Nirma University, Ahmedabad

ABSTRACT

Wireless Sensor Networks (WSN) is collection of small sized nodes. These networks have gained popularity lately because of their self organizing nature and ease of deployment. Data collection from multiple nodes known as convergecast is one of the main functions of Wireless Sensor Networks. Aggregation is performed to conserve energy, which is very dear since the size of nodes is very small and so is their battery capacity. Lots of applications demand fast data delivery, which can be obtained by parallelizing communication among nodes. But, there are factors like interfering links, half duplex transceivers, topology of the network that play as spoil sport. It is shown that with the use of multiple frequencies the parallelism in communication can be effectively employed. That does not give the freedom to use unlimited frequencies. The numbers of frequencies which can be used in WSN are limited because of limitations of communication hardware so the nodes which use the same frequency are further assigned time slots to efficiently use the channel assigned to them. We have used different frequency allocation schemes to show their influence on latency. When multiple frequencies and Time Division Multiple Access (TDMA) techniques are used the latency becomes a function of topology. For different topology with same number of nodes the time required for the convergecast process also varies. Schedule length is obtained by applying different frequency allocation schemes on different tree topologies. Later, the obtained schedule length is compared.

ACKNOWLEDGEMENT

It gives me immense pleasure in expressing thanks to **Prof. Gaurang Raval**, Associate Professor, Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad for his valuable guidance and continual encouragement throughout the Major Project. I am heartily thankful to him for his time to time suggestions on various topics that helped me a lot during my project work.

I like to give my special thanks to **Dr. S. N. Pradhan**, P.G. Coordinator, Department of Computer Science and Engineering,Institute of Technology, Nirma University, Ahmedabad for his continual kind words of encouragement and motivation throughout the project.

I would like to thank my family who have encouraged and motivated me through out my tenure of post graduation. Almighty has been kind to me and guided me in tough times.

I would like to thank all my colleagues who have over the time given me lot of suggestions whenever I used to get stuck. Discussions within lab were good means of sharing knowledge. Also candid moments with them have helped me in taking a break from work and get back to it with furore.

Vinit Ambwani
[10MCEC01]

Abbreviation Notation and Nomenclature

WSN .... Wireless Sensor Network
TDMA .... Time Division Multiple Access
FDMA .... Frequency Division Multiple Access
MAC .... Medium Access Control
SPT .... Shortest Path Tree
MIT .... Minimum Interference Tree
DCT .... Degree-Constrained Tree
TPC .... Transmission Power Control
MST .... Minimum Spanning Tree
BDMRST .... Bounded-Degree Minimum Radius Spanning Tree
CSMA .... Carrier Sense Multiple Access
MMSN .... Multi-Frequency Media Access Control for Wireless Sensor Networks
TMCP .... Tree-based Multi Channel Protocol
CTCCAA Convergecasting Tree Construction and Channel Allocation Algorithm
HYMAC .... Hybrid TDMA/FDMA Medum Access Protocol
BFS .... Breadth First Search
MATLAB .... MATrix LABoratory

Contents

List of Figures

Chapter 1

Introduction

Wireless Sensor Network (WSN) has developed in recent years. WSN because of its flexibility in arrangement as well as the less effort demanded for maintenance, have exhibited promising applications in many fields like military, healthcare, environmental applications, etc. WSN comprises of large number of tiny sensor nodes. Because of their small size and use of wireless medium for communication these nodes can be deployed in the phenomenon or close to it. These sensor nodes because of its size have some limitations. They have limited computation power, memory, communication capabilities and energy. WSNs have been extensively studied with the objective of energy efficiency whereas throughput, bandwidth utilization, fairness and latency were considered as the secondary objectives.
One of the most prominent operations of WSN is convergecasting. 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 wireless sensor networks (WSN). Two types of data collection: (i) aggregated convergecast where packets are aggregated at each hop (ii) raw-data convergecast where packets are individually relayed toward the sink[14]. Aggregated convergecast is applicable when a strong spatial correlation exists in the data, or the goal is to collect summarized information such as the maximum sensor reading. Raw-data convergecast, on the other hand, is applicable within each sensor reading, is equally important, or the correlation is minimal. Aggregated convergecast also results in energy conservation and one of the most popular techniques used. In this it reduces the number of packets to be transmitted from source to sink which saves the energy of transmitting each packet individually. This process does increase the latency of communication. In traditional wireless sensor network (WSN) applications, energy efficiency may be considered to be the most important concern

whereas utilizing the bandwidth and maximizing the throughput are of secondary importance. A lot of applications have recently come up such as structural health monitoring, which require high amounts of data to be collected at a faster rate. Thus a bound on convergecast latency is highly desirable in such applications and mission critical applications, e.g. surveillance and security.

For convergecast, latency of communication is defined as time taken from start of transmission from leaf nodes until all the data is received by base station[14]. In convergecast, latency depends on the number of parallel transmissions. Higher the number, lower the latency. In wireless sensor networks, each device is typically equipped with a single radio transceiver. The broad range of emerging applications requires more complex operations like detection of events in real-time or responsive querying of the network by collecting streams of data in a timely manner. During bursty traffic, the large number of packets generated within a short period leads to a high degree of channel contention and thus a high probability of packet collision. Limited channel capacity and the influence of interference among the sensor radios or the interference due to external networks or electronic devices, that share the same parts of the spectrum, result in a competitive communication environment.

Use of multiple frequencies properly allocated to the interfering links and scheduling the access of these channels will definitely give respite from interference and collisions, which will directly affect the collection rate and latency. The applications of WSN adopt much smaller packet sizes compared to those in general wireless ad hoc networks. Hence, the multi-frequency MAC protocols proposed for general wireless ad hoc networks is not suitable for wireless sensor network applications. Also there is and energy tradeoff between latency and energy. Since the methods generally used for energy conservation make the transceiver switch to sleep mode to conserve energy and the nodes in that case have to wait to send data or send data for the longer period till the receiver node wakes up which eventually delays the packet. In the next chapters, describes the problems in performing fast convergecasting, followed by the previous work that has been in this respect also the protocols that have been developed with the perspective of reducing the latency. Finally, we will discuss the approaches that can be taken to achieve good results in fast convergecasting.

1.1 Objective & Scope

1.1.1 Objective

Wireless medium is inherently of broadcast nature. Thus when a node send some data it is indeed listened by all the nodes within its reach. But when parallely data needs to be sent to different destinations by different nodes which are close to each other then there is a problem. This scenario results into interference also known as collisions due to which none of the nodes completes their communication successfully. In WSN nodes communication reach is less and they are mostly densely populated and have limited energy. So these collisions are responsible for energy wastage and delay which is undesirable for WSNs.

1.1.2 Scope Of The Project

As discussed in previous section, collisions occur in a wireless sensor network when multiple nodes simultaneously transmit to the same node over the same channel or a receiver is in the transmission range of another communication taking place over the same channel. Such collisions waste resources (e.g. bandwidth and energy) as well as increase data latency and hence they are undesirable. For convergecast to work in a collision-free manner, tree topology needs to be constructed and allocate the schedule that specifies for each node in the network the time-slots in which it will communicate. The latency will be further reduced by assigning different time-slot and different frequencies to interfering communications.

Chapter 2

Literature Survey

In typical WSN applications it is of interest to extend the network lifetime due to the battery limitations of the sensor devices. Communication in WSN is the major operation responsible for energy consumption. Extensive study has been performed with the objective of energy conservation whereas throughput, bandwidth utilization, latency have been considered secondary objectives[3]. This chapter contains in brief the details of literature reviewed before starting the work on project and during the execution of the project. It contains information about convergasting, frequency allocation schemes, existing protocols, etc.

2.1 Convergecast

Data gathering is a basic capability expected of any wireless sensor network. The usual means of performing data gathering is to have all nodes send their measurements to a particular node, the sink. The corresponding many-to-one type of communication is called convergecast. Convergecast usually operates by building a logical tree on top of the physical topology with the sink located at the root, and subsequently by routing packets along the tree. There are two types of convergecast methods (i) Raw-data convergecast (ii) Aggregate convergecast.

Aggregation convergecast is defined as the routing and the en-route aggregation of data as they travel to the sink. Aggregation is a means to achieve energy efficiency by reducing the transmitted traffic volume. Aggregation operates by ensuring that a node receives a specific number of incoming (fan-in) messages from a correspondingly specific number of nodes from its neighbors, then combines the received data along with its own, and it generates a single output message that describes collectively the received and its own data together. Typical examples of aggregation function are min, max, sum, average, etc.

A regular convergecast requires more slots than an aggregation convergecast, since there is no aggregation and therefore the volume of traffic is not reduced en-route to the sink.

2.2 Problems in fast convergasting

Communication latency is defined as the time required to complete the process of sending the data from source to the sink. There are applications which have strict latency requirements where data needs to be delivered to the sink almost immediately. If not done there can be serious consequences. Primary limiting factors of fast data collection are: (i) interference in the wireless medium, (ii) half-duplex transceivers on the sensor nodes, and (iii) topology of the network.

2.2.1 Half-duplex transceivers

Since nodes of wireless sensor networks are very small in size they cannot have contain complex hardware of high end transceivers also because of the power constraints the transmission power of these transceivers is very limited. Thus single transceiver is used in these nodes. And since single transceiver is used the nodes can either receive or send at a particular moment. Due to this these transceivers are also said to be the half duplex transceivers. There are transceivers which supports multiple frequencies but that also doesn't improve the situation since the nodes can listen or send data on any one particular frequency at a particular at a time. Due to constant improvement in the semiconductor technology we can expect to have multiple transceivers in a node with only small change in energy consumption but until then all the applications use the general approach of half duplex transceivers.

2.2.2 Interference in the wireless medium

Intersecting links, which are defined as the links with a common destination 2.1(a), cannot transmit on the same time slot since they have to wait for each other's transmission. Assigning non-conflicting frequencies to these nodes does not improve the situation, either. Then the receiver should be assigned a frequency and the senders should use this frequency to transmit to the parent. Interfering or interference links are the links which causes excessive levels of interference if they are scheduled simultaneously. 2.1(b) shows an example where the dotted line represents interference. Interfering links should not get the same time slot and frequency. Since our aim is to minimize the number of time slots, the best option then is to assign the same time slot on non-conflicting frequencies.

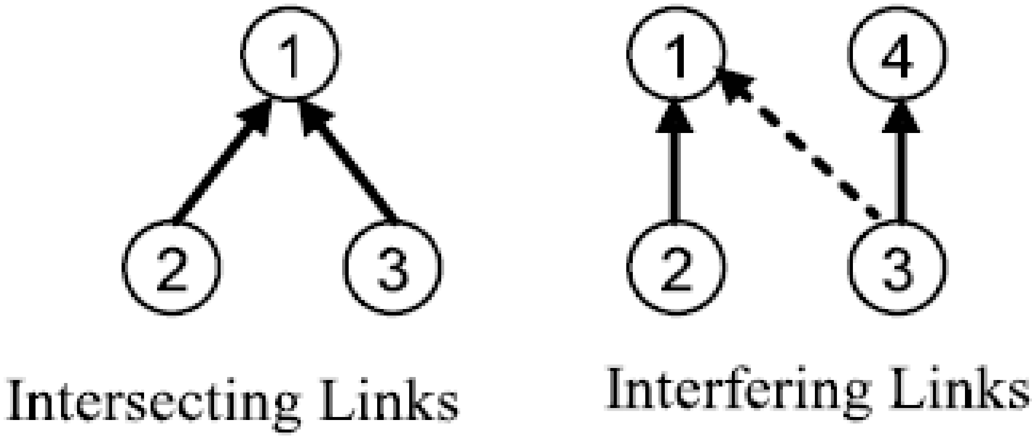

Figure 2.1: Intersecting and Interfering links

Interfering Parents: Interfering parents are pair of parent nodes p and q, such that a transmission by any child of p causes excessive levels of interference with a simultaneous transmission by any child of q on a tree topology. As illustrated in 2.1(b), nodes 1 and 4 are interfering parents when assigned the same frequency because simultaneous transmissions by their respective children 2 and 3 cause interference on parent 1.

2.2.3 Topology of the network

The topology of the network is one of the factors where the number nodes in an area i.e. its density do affect the latency. Connectivity may also limit the performance of scheduling. Consider the nodes that select the same parent. So, in case of tree topology there is no bound in the number of children to a node. They have to wait for each others' transmission which simply increases the length of a schedule. The number of required time slots per frame in a TDMA based schedule depends on the connectivity of the network topology.

2.3 Existing protocols

2.3.1 Multi-Frequency Media Access Control for Wireless Sensor Networks (MMSN)[10]

MMSN protocol consists of two aspects: frequency assignment and medium access. The frequency assignment is used to assign different frequencies if enough frequencies exist, or evenly allocate available frequencies if there are more neighbors than available frequencies, to nodes that have potential communication conflicts. MMSN allows users to choose 1 of 4 frequency assignment strategies. The four frequency assignment strategies offered are: (i)Exclusive frequency assignment (ii) Even selection (iii)Eavesdropping (iv)Implicit-consensus. In media access design, nodes that have potential conflicts coordinate to access the shared physical frequencies, in a distributed way. It is a slotted CSMA protocol and at the beginning of each timeslot nodes need to contend for the medium before they can transmit.

MMSN assigns channels to receivers. When a node intends to transmit a packet it has to listen for the incoming packet both on its own frequency and the destination's frequency. A snooping mechanism is used to detect the packets on different frequencies which makes the nodes to switch between channels frequently. MMSN uses a special broadcast channel for broadcast traffic and the beginning of each time slot is reserved for broadcasts.

2.3.2 Multi-Channel Lightweight Medium Access Control (MC-LMAC)[3]

MC-LMAC is a schedule-based multi-channel MAC protocol that takes the advantage of interference and collision free parallel transmissions on different channels. It is designed to provide higher throughput over multiple channels. A node selects a timeslot and a channel on which it is allowed to transmit. Timeslot and channel selection is fully distributed and guarantees the same slot/channel pair not to be used with the 2-hop neighborhood.

MC-LMAC protocol has a scheduled access where each node is granted a time slot beforehand. A timeslot consists of a control period and a data transmission period. During the control period, all the nodes switch their interfaces to a common channel. The control period is used for notifying the destination about the incoming packet and the channel on which the data transmission will take place such that the receiver should switch its interface. It does not require a dedicated broadcast channel. All nodes at the start of each timeslot are required to listen to a common channel in order to exchange control information.

2.3.3 Tree-based Multi Channel Protocol (TMCP)[4]

This protocol partitions the network into multiple sub-trees. The goal of this protocol is to minimize the intra-tree interference. The protocol partitions the network into subtrees and assigns different channels to the nodes residing on different trees. TMCP is designed to support convergecast traffic and it is difficult to have successful broadcasts due to the partitions. Contention inside the branches is not resolved since the nodes communicate on the same channel.

TMCP has three components, (i) Channel Detection (CD), (ii) Channel Assignment and (iii) Data Communication (DC). Given k-orthogonal channels, the CA module partitions the whole network into k subtrees and assigns one unique channel to each sub tree. The goal of partitioning is to decrease potential interference as much as possible. After partitioning the interference in the original network can be divided into two categories, one is the interference among trees, called inter-tree interference, which is eliminated by assigning different orthogonal channels to each subtree and the other is the potential interference among nodes within a tree, called the intra-tree interference. Since intra-network interference cannot be avoided in this scheme it becomes the main performance bottleneck.

After assigning channels, the DC component manages the data collection through each subtree. When a nodes wants to send information to the sink, it just uploads packets along the subtree it belongs to.

2.3.4 Hybrid TDMA/FDMA Medium Access Control (HyMAC)[2]

It is multi-channel MAC protocol for WSN. It is a combination of TDMA and FDMA. However, assignment of timeslots and frequencies is done according to the Breadth First Search (BFS) algorithm on a tree topology.

It performs a Breadth First Search (BFS) constructing a tree having the base node as its root. As each node is traversed by BFS, it is assigned a default time slot and a frequency Then the possibility of having an interference with any of its same-height previously visited one-hop And two hop neighbors is checked. If a conflicting neighbor n2 is found for a particular node n1, the algorithm checks whether the two nodes are siblings. If so, the node n1 is assigned a different frequency than that of n2. If they are not siblings then n1 is a different frequency than that of n2, allowing both n1 and n2 to send messages to their parents at the same timeslot but in different channels. When BFS is about to start a new level (height) of nodes the default time slot number will be increased by one.
Once all the nodes are processed according to the above heuristic, all of the time slot assignments will be inverted such that the slot number assigned to every node is smaller than that of its parent.

2.3.5 Convergecasting Tree Construction and Channel Allocation Algorithm (CTCCAA)[1]

The CTCCAA algorithm utilizes a greedy approach for tree construction and channel allocation. The formation of tree takes place one level at a time and simultaneously allocates channel for communication.

The algorithm has some constraints which it follows: (i)Co-channel Criterion: If children of one group are in transmission range of another group's parent then the groups' children must have channel for communicating with their parent if available or they should use different slot. (ii)Proximity criterion: A node is assigned a child to the closes possible parent node. (iii) No two children of a group share the same slot for communication with their parent. (iv)If distance of node 1 is less than the distance of node 2 from A then the slot assigned to node 1 is greater than node 2, i.e. The child nodes of each group are allocated a slot that is greater than the slot used by their parent.

The algorithm starts channel allocation from the root node (basestation) and proceeds in a BFS manner. The root node runs the algorithm and after termination the root node transmits a broadcast packet with the schedules allocated for each node in the network. The algorithm takes the neighbor list of each node at the current level and starts assigning a channel for each node in the list such that it adhers to the constraints and these nodes become part of the next layer of the tree.

The algorithm maintains a list called 'currentlist' and a list called 'nextlist'. Initially currentlist consists of all nodes in level 1 (i.e. the root node) and the nextlist will be empty. For each node(P) in level 1 it chooses the children and allocates a channel for each of the child nodes to communicate with P. It uses the proximity criterion while choosing children for each node. All these nodes are inserted in the nextlist. Once allocation for all nodes in the currentlist is done, all nodes in the nextlist are copied onto the currentlist. Now the algorithm chooses each node from the currentlist (i.e. all nodes in the current level (level 2)) and chooses children for them and also allocates channel for the child nodes. For convergecasting the slot allocated to a parent node should be greater than the slot allocated to the child node. But in this case after allocation, the timeslot for a parent node will be lesser compared to the slot allocated to a child node. So the order in which time slots allocated to a nodes in the network is reversed by finding the slot with the highest number and from that subtract the slot number allocated to each node to get actual slot number.

2.4 Suggested Approaches

2.4.1 Transmission power control

It improves the performance of the network in several aspects. First, power control technique improves the reliability of a link. Upon detecting that link reliability is below a certain threshold, the MAC protocol increases the transmission power, improving the probability of successful data transmission. Second, only nodes which really must share the same space will contend to access the medium, decreasing the amount of collisions in the network. This enhances network utilization, lowers latency times and reduces the probability of hidden and exposed terminals. It has been found that for moderate size networks of 100 nodes power control can reduce the schedule length by 15 - 20%[7].

2.4.2 Contention free protocols

It is well known that collisions present a major challenge in covergecast when contention-based MAC protocols like CSMA are employed. Collisions result in loss of packets and recovery methods such as retransmission increase the latency. Moreover, retransmissions might lead to further collisions in high data rate scenarios. In addition, retransmissions drain the scarce energy reserves of sensor nodes.

Contention-free MAC protocols like TDMA can be used to eliminate collisions and obtain a bound on the time required to complete convergecast. A TDMA schedule can be determined such that convergecast is completed in minimal number of timeslots. In TDMA, time domain is sliced in timeslots. Multiple, spatially separated, non-interfering transmissions can be scheduled in each timeslot.

2.4.3 Use of multiple frequencies

In general wireless networks, multiple channels have been provisioned to mitigate the effects of interference by assigning different frequency channels to interfering links.

A typical sensor device is usually equipped with a single half-duplex radio transceiver, which can not perform simultaneous transmission and reception, but can work on different channels separately. On the other hand, traditional wireless ad-hoc networks usually assume more powerful radio hardware and multiple transceivers per node[12]. If the channels are orthogonal, simultaneous transmissions can take place on multiple channels without interference

An important fact to be observed is that the current WSN hardware such as MICAZ, Telos and CMU FireFly use CC2420 radio which provides multiple channels. Although the typical WSN radios operate on a limited bandwidth, the operating frequency of the radios can be adjusted over different channels. Once different channels are assigned to interfering or contending links, more simultaneous transmissions can take place and more data can be delivered to the sink node in limited time. In the next section different frequency allocation schemes are discussed.

Figure 2.2 shows a scheduling example on a tree toplogy. In 2.2(a), the solid lines between the nodes show the transmission links whereas the dotted lines show the interfering links. The numbers inside the circles represent the node ids. Figure 2.2(b) shows the tree after time slot assignment with a single frequency channel. The numbers on the links show the assigned time slots. In this case, it takes 6 time slots to schedule the network. In 2.2(c) it can be seen how the scheduling is performed with 2 frequencies. First, the frequencies are assigned to the parents (represented inside the boxes next to each parent, F1 is the initial frequency).

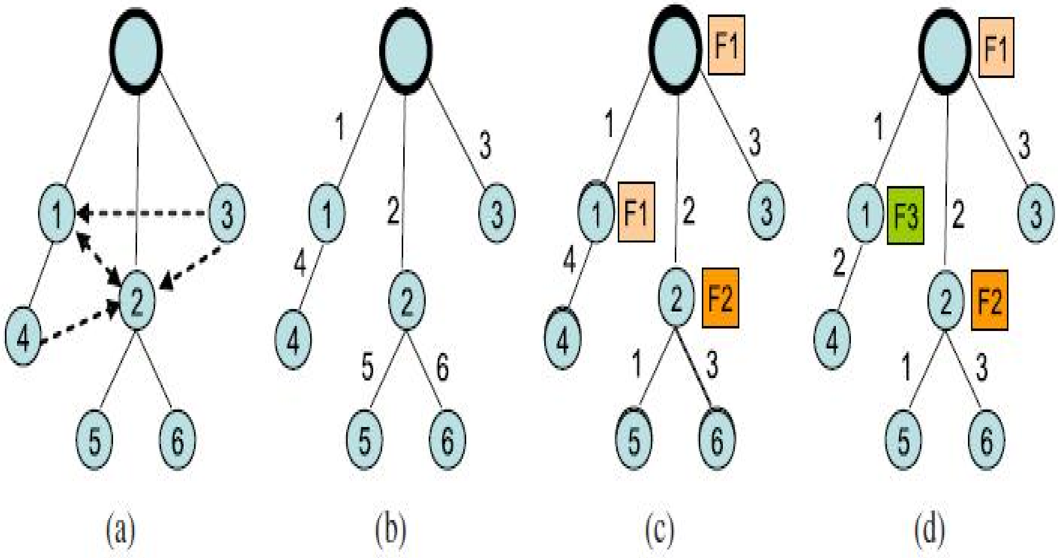

Figure 2.2: (a)Communication links and interfering links (b) Schedule with 6 time slots on a single channel (c) Schedule with 4 time slots in 2 channel (d) Schedule with 3 time slots in 3 channels

Then, the time slots are assigned to the senders. With 2 frequencies, the network is scheduled in 4 slots.

Figure 2.2(d) shows the case with 3 frequencies. The network is then scheduled in 3 time slot. 50% reduction is achieved on the schedule length thus the data collection rate at the sink is doubled with the sufficient number of frequencies. Once multiple frequencies are employed along with spatial-reuse TDMA, the data collection rate often no longer remains limited by interference but by the toplogy of the network.

2.4.4 Topology control

The main goal of topology control is to reduce node power consumption in order to extend network lifetime. Since, the energy required to transmit a message increases with distance, it makes sense to replace a long link by a sequence of short links. On one hand, energy can be conserved by abandoning energy-expensive long-range connections, thereby allowing the nodes to reduce their transmission power levels. On the other hand, reducing transmission power also confines interference, which in turn lowers node energy consumption by reducing the number of collisions and consequently packet retransmissions.

Consider the nodes that select the same parent. They have to wait for each others' transmission which simply increases the length of the schedule. One option would be to construct balanced trees. But, with balanced trees sink remains the high-degree bottleneck. To avoid the bottlenecks, there should be a limitation on the number of children per parent.

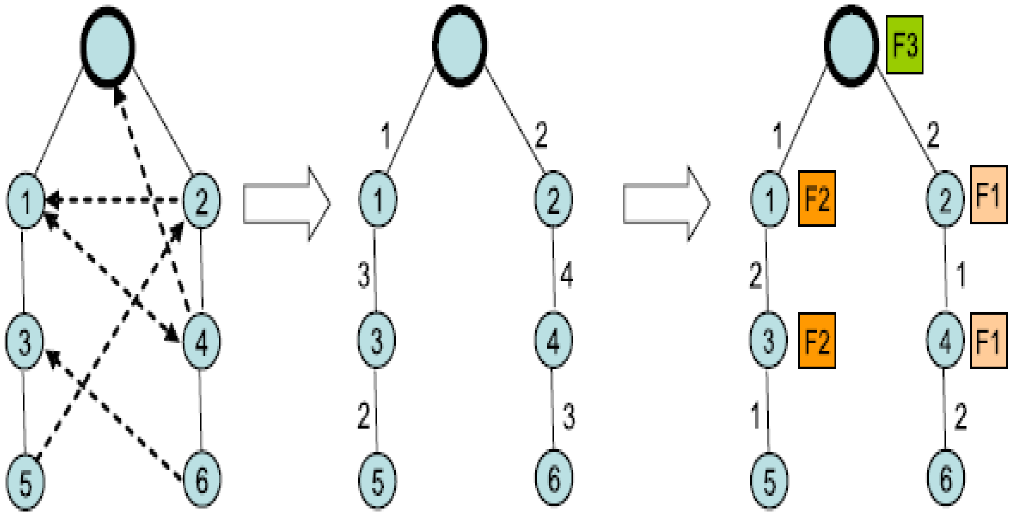

Figure 2.3: Scheduling in a degree constrained tree

Consider a case when all n nodes are in range of each other and the sink. If nodes select their parents accordiing 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 subtrees rooted at the sink. If there are enough number of frequencies to eliminate all the interference then the network can be scheduled in 2 time slots. Figure 2.3 shows the same network as in in Figure 2.2 with a different routing tree. The second part of the figure shows the time slot scheduling which takes 4 time slots on a single channel frequency. The last part shows that the time slots are scheduled over different frequency channels. This takes 2 time slots to schedule all the links which is 3 times better than the baseline with single frequency over a non-degree constrained tree.

2.5 Frequency Allocation Schemes

2.5.1 Receiver based frequency assignment[8]

Receiver-based frequency assignment strategy is the scheme in which nodes are assigned the frequency on which they will receive data from the nodes. So when other nodes want to send data to a particular node at that time they have to tune in to the frequency of that particular node. So in that case any node will operate on atmost two frequencies. One on which it receives data and the other on which it needs to send data. Receiver-based channel assignment is a widely used approach in sensor networks as a convenient way to organize multi-channel protocols, because it simplifies synchronization issues as all receptions take place on the same channel at each node.

2.5.2 Exclusive frequency assignment[10]

In exclusive frequency assignment, nodes first exchange their IDs among two communication hops, so that each node knows its two-hop neighbors' IDs. A simple way to implement this is for each node to broadcast twice. In the first broadcast, each node beacons its node ID, so that each node knows its neighbors' IDs within one communication hop. In the second broadcast, each node beacons all neighbors' IDs it has collected during the first broadcast period. Hence, after the second beacon period, each node gets its neighbors' IDs within two communication hops.

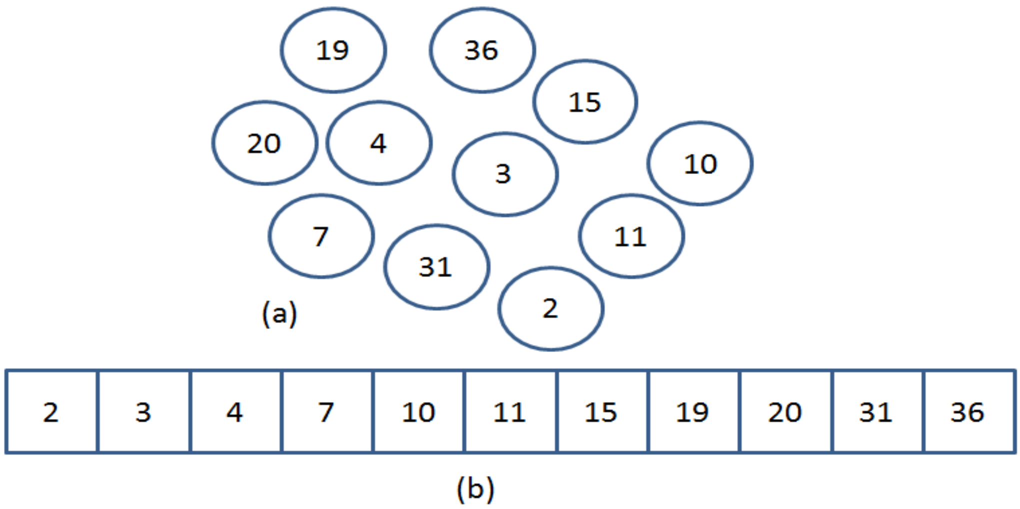

Figure 2.4: Example of a topology with node with two hop neigbor node IDs

After nodes collect ID information of all neighbors within two hops, they make frequency decisions in the increasing order of their ID values. If a node has the smallest ID among its two communication hops, it chooses the smallest frequency among available ones, and then beacons the frequency choice within two hops. If a node's ID is not the smallest one among two hops, it waits for frequency decisions from other nodes within two hops that have smaller IDs. After decisions from all those nodes are received, the node chooses the smallest available (not chosen by any of its two-hop neighbors) frequency and broadcasts this choice among two hops. This scheme guarantees to assign different frequencies to different nodes within any two-hop neighborhood, when the number of frequencies is at least as large as the two-hop node number.

In the figure 2.4(a) an example of topology is given. After exchange of node IDs on two hop communication distance the node with the smallest ID is allocated the frequency first. (In figure 2.4(b) the nodes are arranged in ascending order of node IDs which is info of two hop neighborhood for node 31).Thus, node 2 will be assigned the frequency first. All nodes will be assigned exclusive frequencies.

2.5.3 Even Selection[10]

In exclusive frequency assignment, when there are not enough frequencies, it is possible that when a node makes its frequency decision, all physical frequencies have already been chosen by at least one node within two hops. In this case, the exclusive frequency assignment is extended by randomly choosing one of the least chosen frequencies. For convenience, we call this extension even selection, which makes an even allocation of available frequencies to all nodes within any two communication hops.

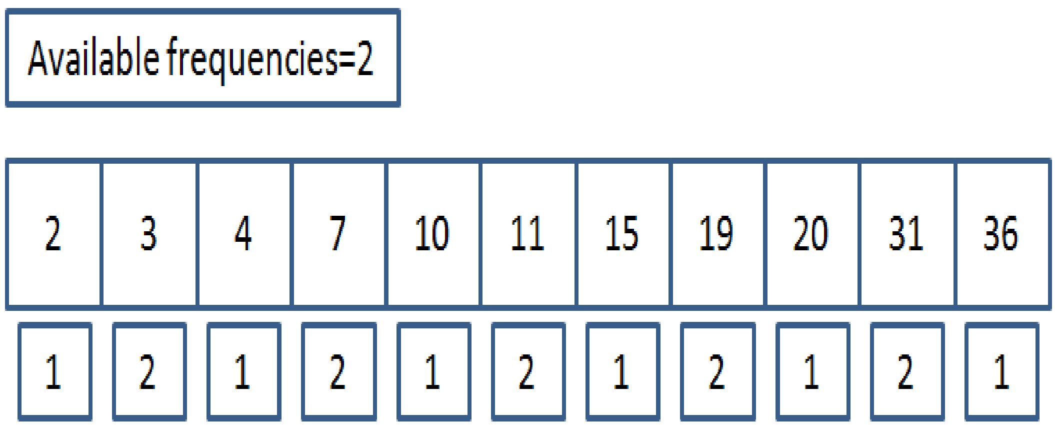

Figure 2.5: Even selection (Available frequencies=2)

Considering the same topology given in figure 2.4 if the number of frequencies available are two then the nodes are evenly allocated the two frequencies and when

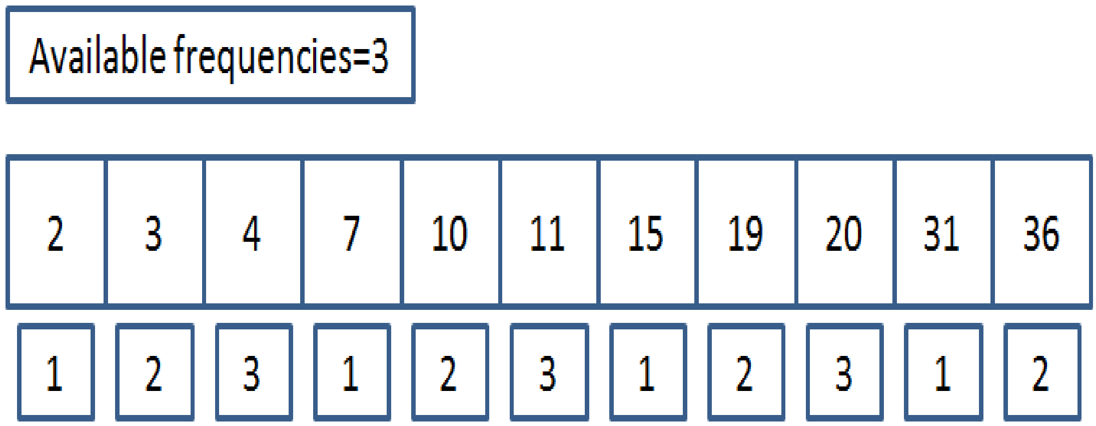

Figure 2.6: Even selection (Available frequencies=3)

the number of available frequencies are three then nodes are evenly allocated the three frequencies. This is shown in figure 2.5 and figure 2.6 respectively.

2.5.4 Eavesdropping[10]

Even though the even selection scheme leads to even sharing of available frequencies among any two hop neighborhood, it involves a number of two-hop broadcasts. To reduce the communication cost, there is a lightweight eavesdropping scheme. In eavesdropping, each node takes a random backoff before it broadcasts its physical frequency decision. During the backoff period, each node records any physical frequency decision overheard. When a node's backoff timer fires, it randomly chooses one of the least chosen frequencies for data reception. Compared with even selection, eavesdropping has less communication overhead, but it also results in more potential conflicts, because it only collects information within one hop for frequency decisions.

2.5.5 Implicit-consensus[10]

When physical frequencies are abundant, the communication overhead in exclusive frequency assignment can be further reduced, while all nodes within any two-hop neighborhood can still be guaranteed to get assigned different frequencies. To achieve this performance, the implicit-consensus scheme can be used, which is inspired by the pseudo random number generator algorithms. It is a distributed frequency assignment algorithm for multi-frequency MAC designs.

In implicit-consensus, nodes' IDs need to be collected within two hops, in the

same way as what is done in exclusive frequency assignment. Then, each node calculates its frequency number with a local computation.
In the system, all nodes share the same pseudo random number generator, which is able to generate a unique random number sequence for each specified seed, the node ID here. For each frequency number, each node calculates a random number for itself and a random number for each of its two-hop neighbors, with the same pseudo random number generator. A node wins the current frequency number if and only if its current random number is the highest among all current random numbers generated by all nodes within two hops. When two random numbers tie, the one with the larger node ID wins. In this way, each node explores all frequency numbers from zero to positive infinity until it finds the frequency that it has the highest priority for. By using the same pseudo random number generator, it is guaranteed that when a node decides that it wins frequency number FreNum i, all nodes within two hops automatically agree with that decision and consensus is implicitly achieved, without any communication.

2.5.6 Tree based frequency allocation[4]

In tree based allocation schemes, all the children of the sink nodes are considered to be roots of the subtrees. Exclusive frequency is assigned to every subtree. Thus all the nodes that belong to a particular subtree are assigned the same frequency.

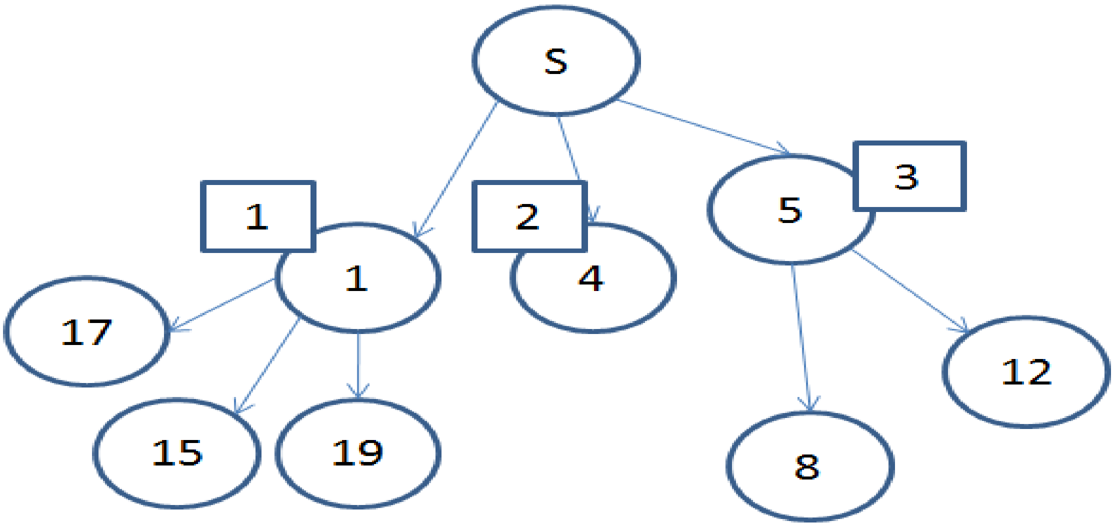

Figure 2.7: Tree based frequency allocation

The sink node S has 3 children (2.7). These 3 children are considered to be the roots of 3 subtrees and exclusive frequencies are allocated to these trees. Thus nodes 1, 17, 15, 19 are allocated with frequency 1, node 4 is allocated frequency 2 and nodes 5, 8, 12 is allocated frequency 3.

2.5.7 Load balanced frequency assignment

In the load based frequency allocation schemes. The objective is to assign frequencies to the nodes in such a manner that the communication finishes evenly. Thus for that reason the nodes which are heavily loaded i.e. the nodes which have more children will definitely need more time to acquire data from them. If the frequencies are allocated at receiver based then the children of the node will tune in that frequency to send data. Thus nodes with more children are assigned the frequencies which are least loaded. In this manner the frequencies are assigned in such a way that the load is distributed evenly on frequencies.

2.6 Challenges and Requirements

2.6.1 Syncronization

If the channel assignment is done and the radios are switching between channels instead of being fixed on one channel, co-ordination is required between the senders and receivers in order to be on the same channel at the same time.

2.6.2 Broadcast support

If the nodes are switching between channels dynamically, it might be problematic to support local broadcasts. Local broadcasts are important for WSN traffic since sensor nodes may require in-network processing before they transmit the data towards the sink node.

2.6.3 Channel switching

The radio cannot switch between the channels immediately but takes some time, for instance it is around 650 micro sec for Nordic Nrf905 radio. The timeslot size should be large enough to accommodate the switching time.

2.6.4 Joining the network

A new node joining the network may disrupt the channel organization, it may be required to scan all channels to find the suitable channel to transmit.

Chapter 3

Implementation

As per our study, we have found out that the collisions while transmission results into energy wastage and delay. Since our objective is to reduce the latency quotient of the whole convergecast process. Since it has been found that most of the convergecast processes works in two phases:

- Tree creation
- Scheduling of nodes such that interference is avoided.

In this project modules are created for creating trees and modules for frequency and timeslot allocation to avoid interference. Since it is modular it gives us opportunity to get results for using different frequency allocation schemes with different tree topologies and frequency allocation schemes. The frequency allocation and timeslot allocation are further decoupled into different phases. As shown in 3.1 first step consists of tree creation. In this step different tree creation module can be called. Similarly, in the second step of frequency allocation we can call different allocation schemes. Which at the end will give us results according to those selections.

3.1 Platform Used: MATLAB

MATLAB (matrix laboratory) is a numerical computing environment and fourth-generation programming language. Developed by MathWorks, MATLAB allows matrix multiplications, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interacting with programs written in other languages including C, C++, Java, and Fortran. MATLAB has been extensively been used in this project for implementing the algorithms for creating tree topologies, frequency allocation schemes and timeslot assignment.

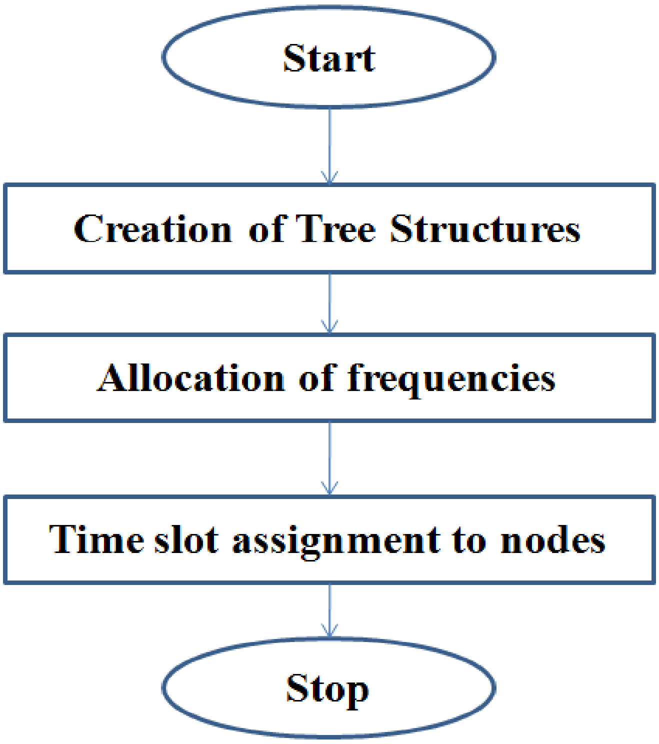

Figure 3.1: Flowchart

3.2 Assumptions

I have considered an undirected graph G=(V,E), where V is the set of nodes and E is the set of edges representing communication links. Network is a connected and all the nodes have a uniform transmission range R. Thus any two nodes can communicate if the distance between them is atmost R. The nodes are deployed over the area of 200 x 200. Each node is equipped with a single half-duplex transceiver. We consider that two nodes interfere to each other when (i) the nodes are adjacent to each other, or (ii) both the nodes are transmitting on the same frequency, and atleast one of the receivers of a node is within the interference range of the non-intended transmitter. We assume that transmissions on different frequencies are orthogonal and non-interfering to each other. All the frequency allocation schemes are receiver-based, in which frequencies are allocated to the nodes statically. So, when children have to transmit data to the parent they have to tune into the frequency assigned to the parent. Due to this static assignment, each node operates on at most two frequencies. Typically, the number of frequencies on which a transceiver of a WSN node can operate is limited.

3.3 Shortest Path Tree

After giving positions of the node the SPT is created. The tree is first constructed by putting only 125 nodes as shown in figure 3.5. Later in the same area we increase the number of nodes. Figure 3.6, 3.7, 3.8, 3.9 shows the SPT with 200, 400, 600 and 800 nodes in the same area.

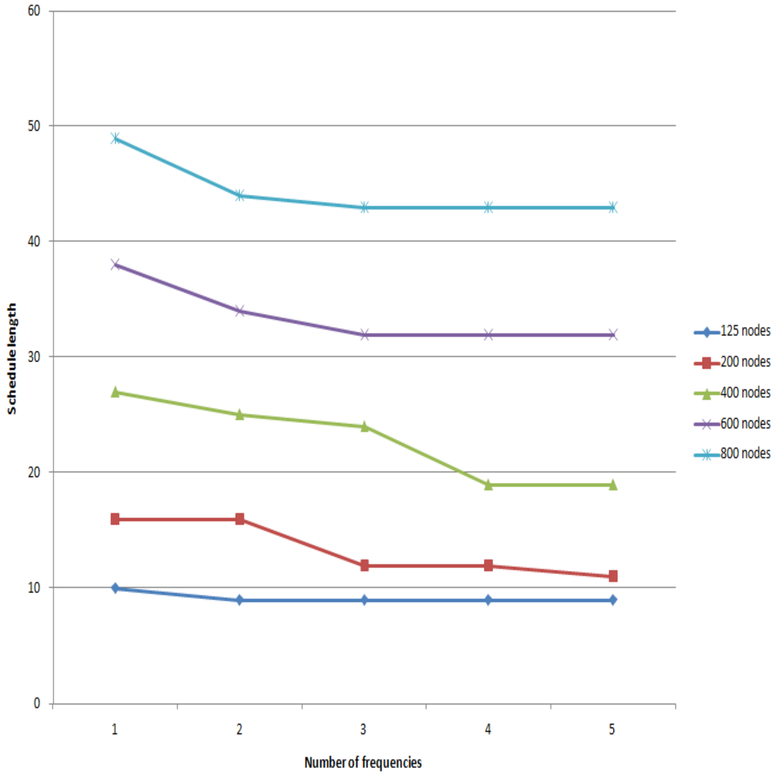

Figure 3.2: Schedule length for SPT with Even Selection allocation

Figure 3.2 shows the graph of the schedule length for SPT tree. After assigning frequency using Even Selection strategy and increasing the nodes in the area gradually. We have found that as the number of nodes are increased within a unit area definitely the schedule length will increase to accomodate all the nodes. What is also observed that when the nodes are scarcely placed in an area the interference levels are very low so assigning more number of frequencies is of very less use.

Figure 3.3 shows the graph of the schedule length for SPT tree. Schedule length is obtained after assigning frequencies using load balanced frequency assignment scheme.

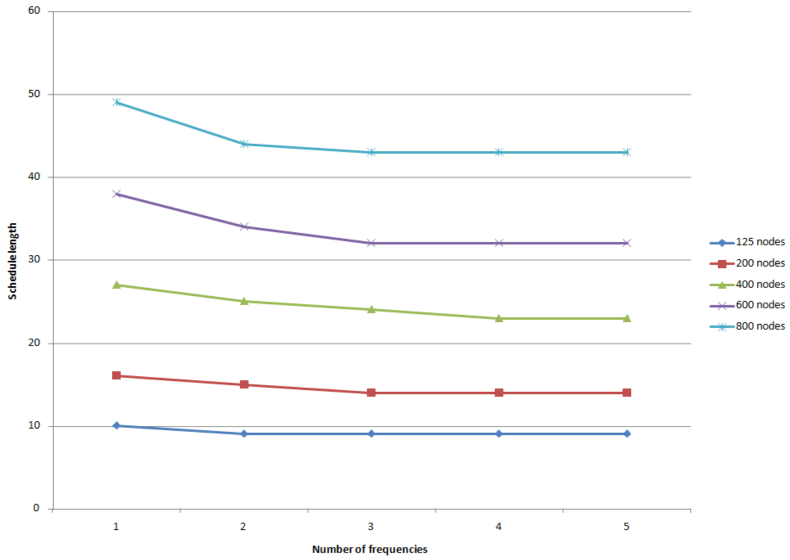

Figure 3.3: Schedule length for SPT with Load Balanced allocation

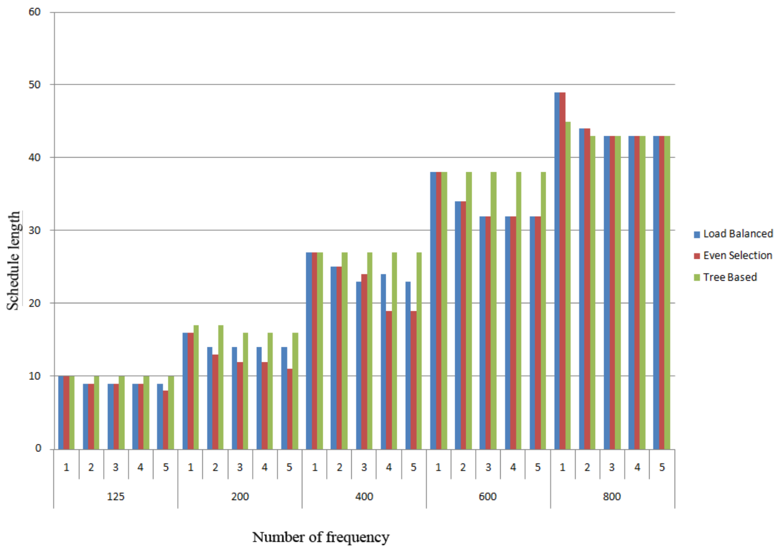

Figure 3.4: Comparison of different frequency allocation used in SPT

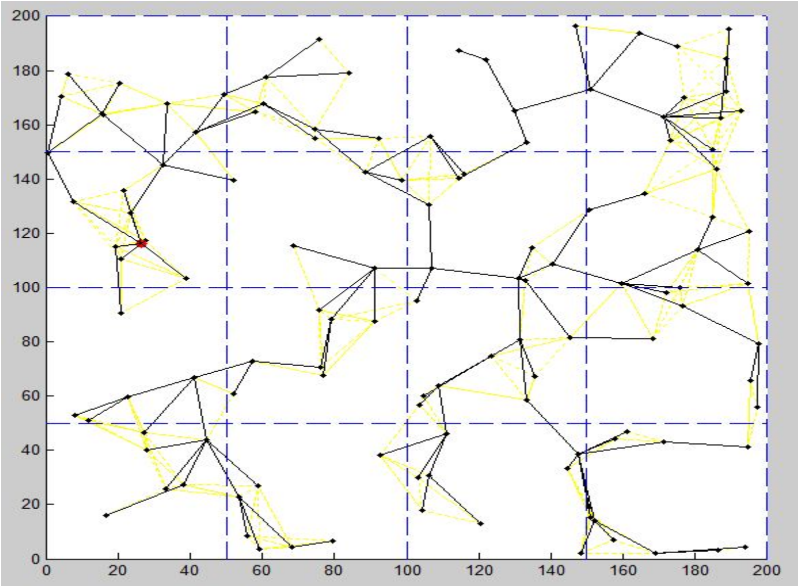

Figure 3.5: Shortest Path Tree (125 nodes)

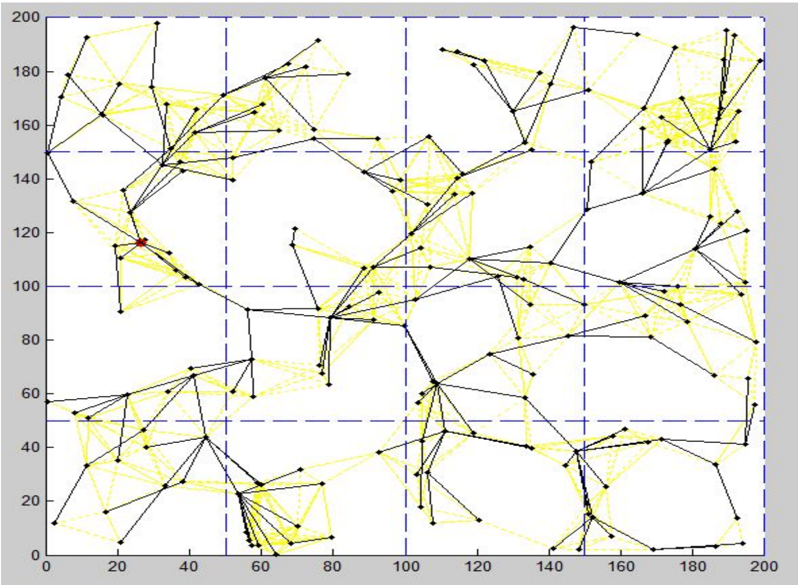

Figure 3.6: Shortest Path Tree (200 nodes)

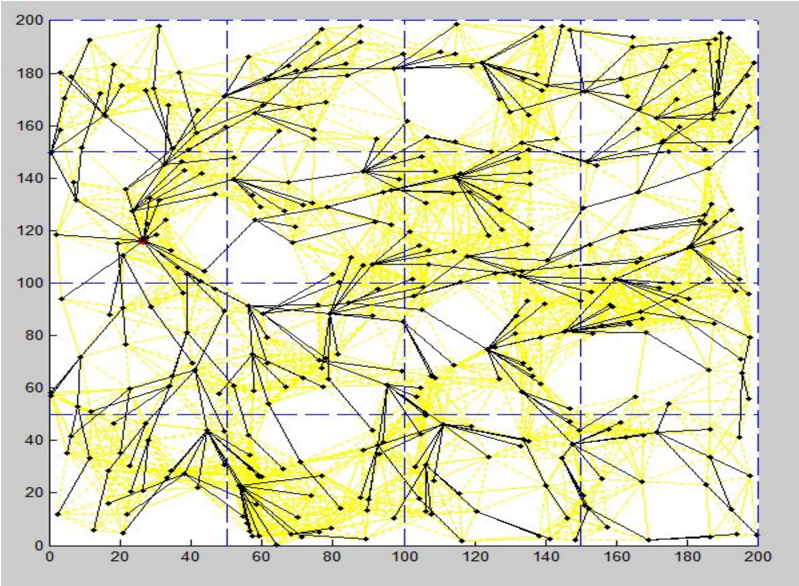

Figure 3.7: Shortest Path Tree (400 nodes)

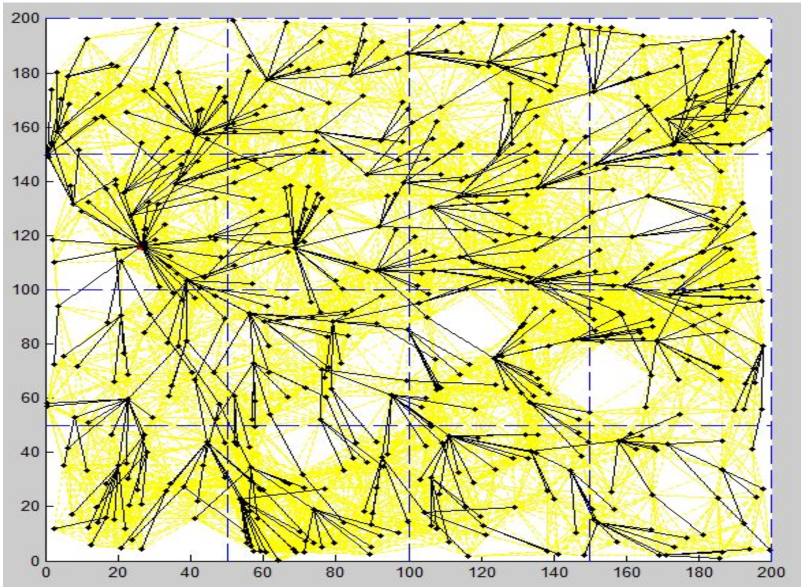

Figure 3.8: Shortest Path Tree (600 nodes)

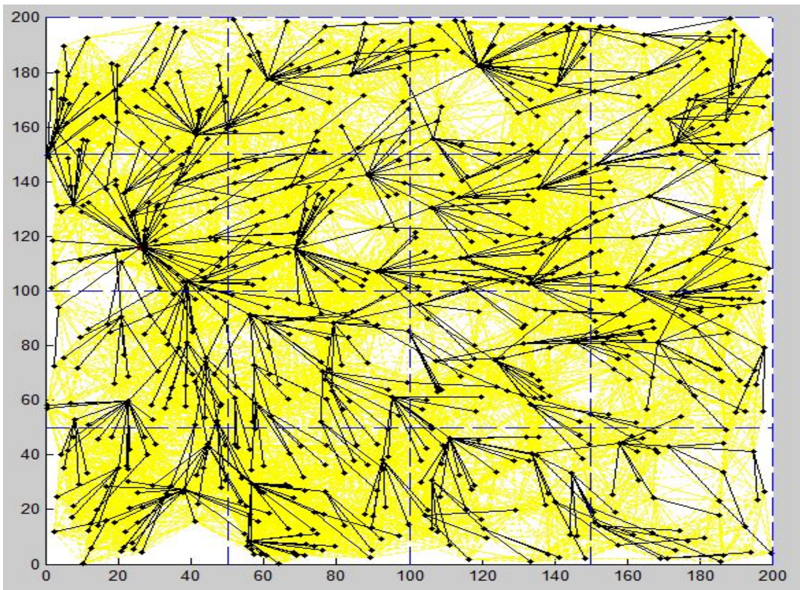

Figure 3.9: Shortest Path Tree (800 nodes)

3.4 Minimum Interference Tree

This tree is nothing but a Minimum Spanning Tree (MST) where the cost function for selecting a link is the number of nodes covered by the union of the two disks centered at nodes forming the link. This cost function gives the measure of the interference by counting the number of nodes affected by the two nodes intended to form a link. The MST thus constructed is called the Minimum Interference Tree (MIT).

After giving positions of the node the MIT is created. The tree is first constructed by putting only 125 nodes as shown in figure 3.13. Later in the same area we increase the number of nodes. Figure 3.14, 3.15, 3.16, 3.17 shows the MIT with 200, 400, 600 and 800 nodes in the same area. Figure 3.10 and 3.11 shows the graph of the schedule length for MIT tree using even selection and load balanced frequency allocation scheme respectively. The graph in figure 3.12 gives the comparison of the frequency schemes applied to the nodes.

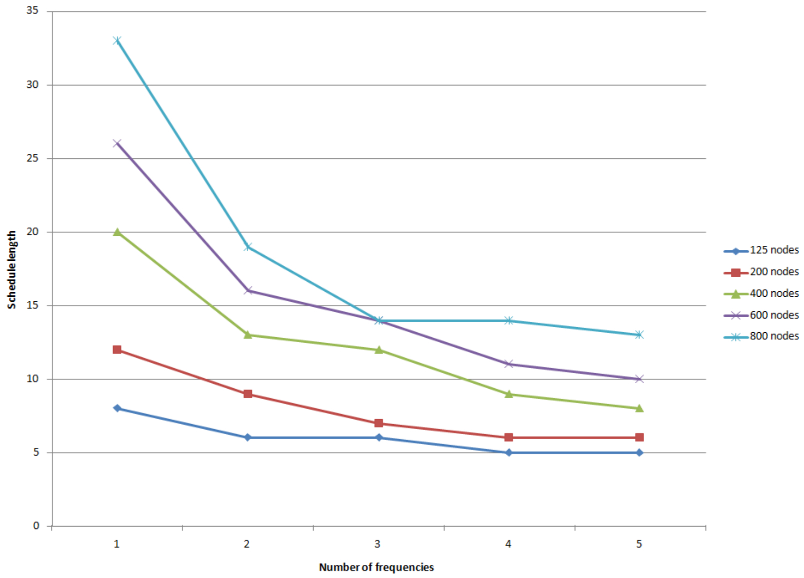

Figure 3.10: Schedule length for MIT with Even Selection allocation

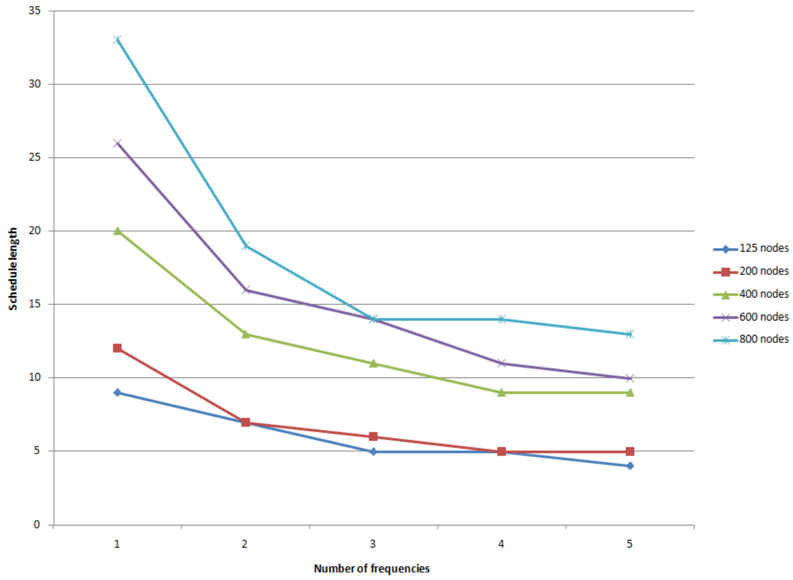

Figure 3.11: Schedule length for MIT with Load balanced allocation

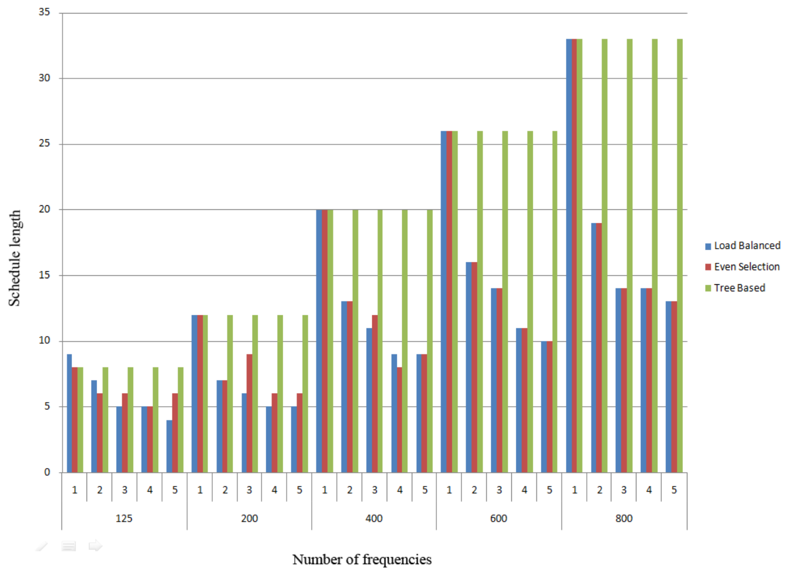

Figure 3.12: Comparison of different frequency allocation used in MIT

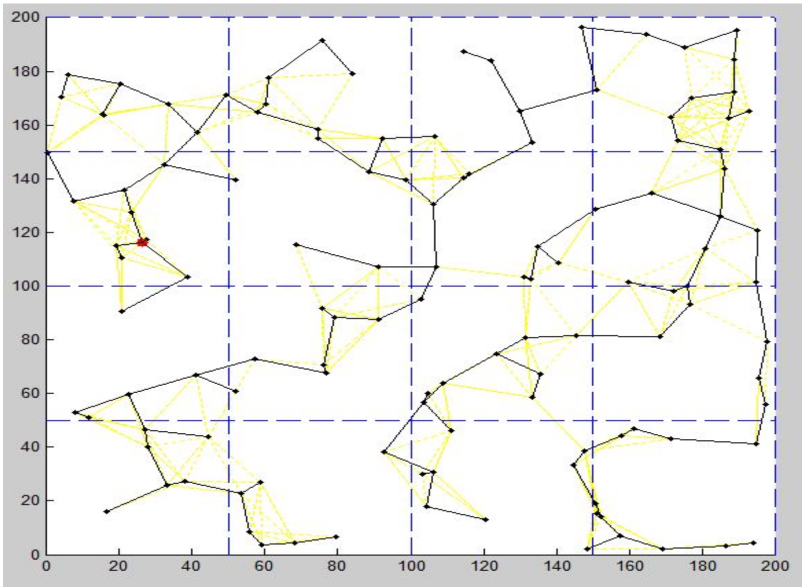

Figure 3.13: Minimum Interference Tree (125 nodes)

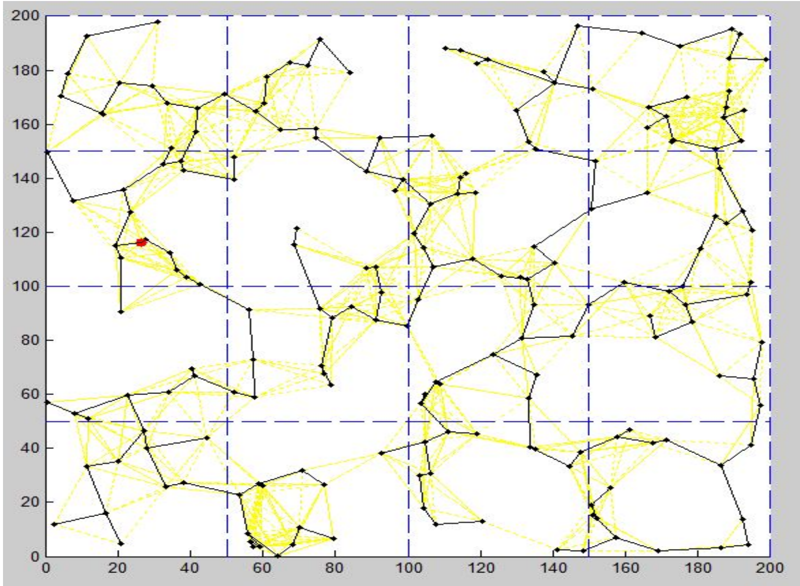

Figure 3.14: Minimum Interference Tree (200 nodes)

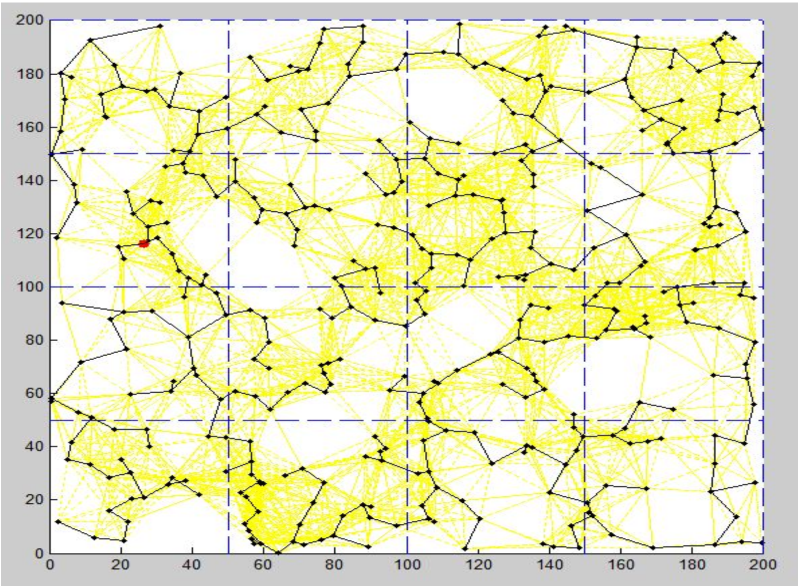

Figure 3.15: Minimum Interference Tree (400 nodes)

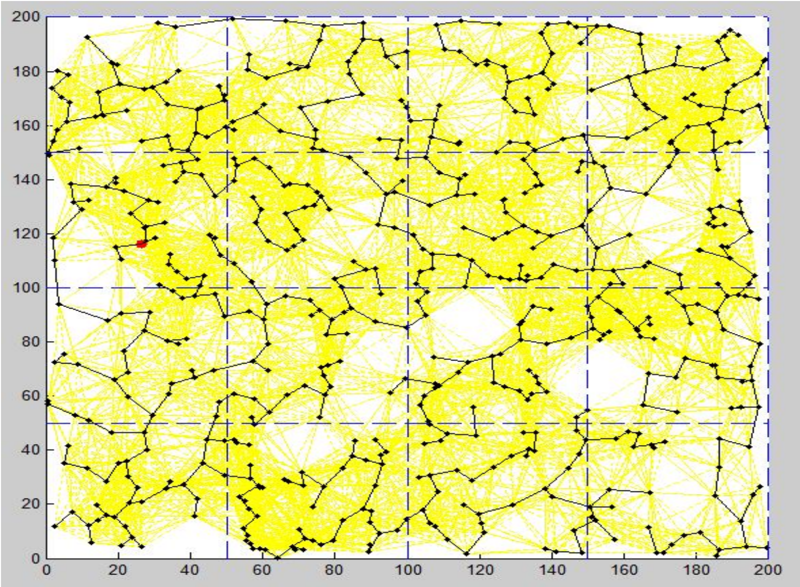

Figure 3.16: Minimum Interference Tree (600 nodes)

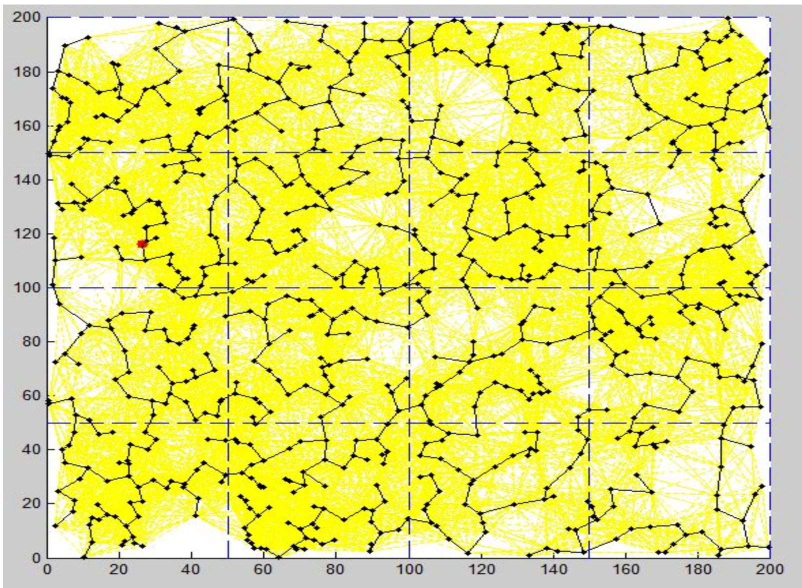

Figure 3.17: Minimum Interference Tree (800 nodes)

3.5 Bounded Degree Minimum Radius Spanning Tree (BDMRST)

In this tree, first of all backbone tree is created. One representative is chosen called a local root arbitrarily from each non-empty cell, and connecting them in Breadth-First-Search (BFS) order starting from sink. After this a local spanning tree is created within each cell from the remaining nodes lying in that cell while respecting the bounded degree constraint. Finally the union of the backbone tree and all the spanning trees is taken.

After giving positions of the node the BDMRST is created. The tree is first constructed by putting only 125 nodes as shown in figure 3.18. Later in the same area we increase the number of nodes. Figure 3.19, 3.20, 3.21, 3.22 shows the BDMRST with 200, 400, 600 and 800 nodes in the same area. Figure 3.23 and 3.24 shows the graph of the schedule length for BDMRST tree using even selection and load balanced frequency allocation scheme respectively.

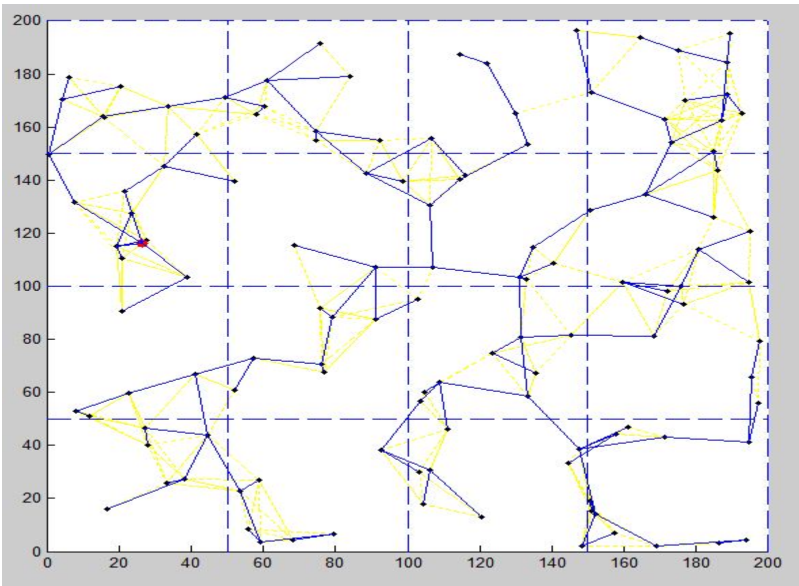

Figure 3.18: BDMRST (125 nodes)

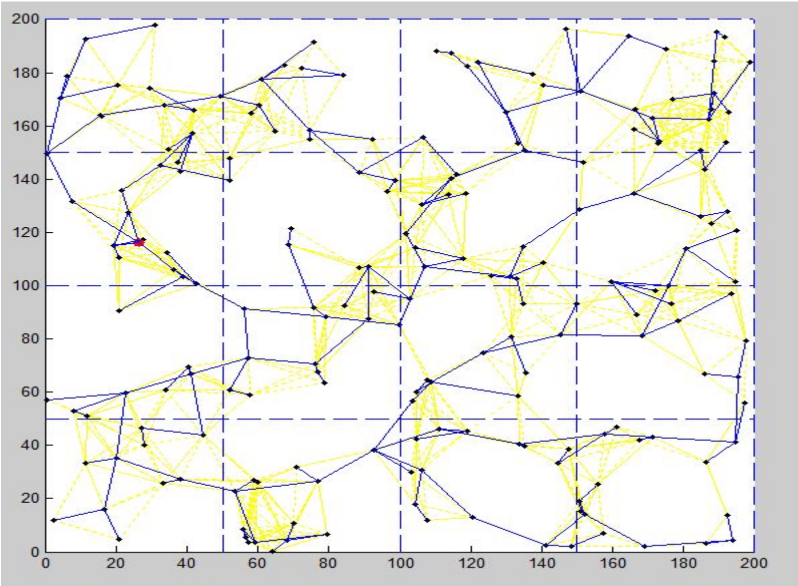

Figure 3.19: BDMRST (200 nodes)

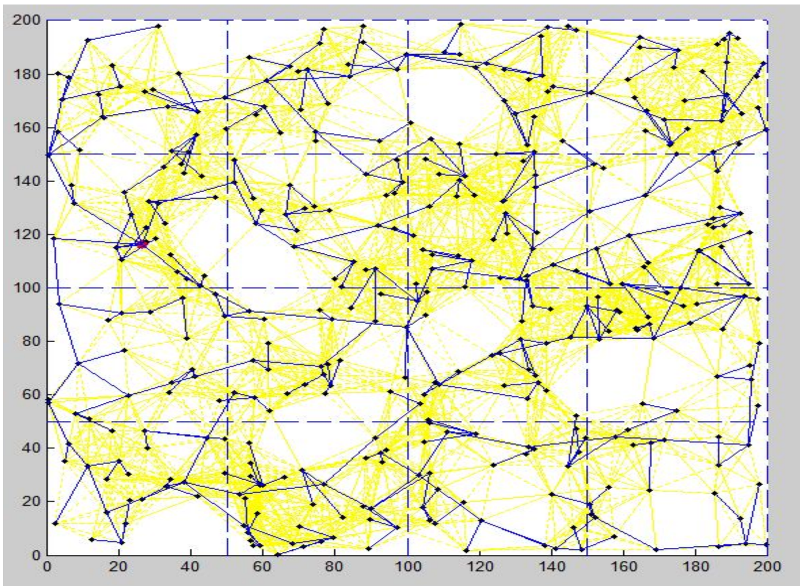

Figure 3.20: BDMRST (400 nodes)

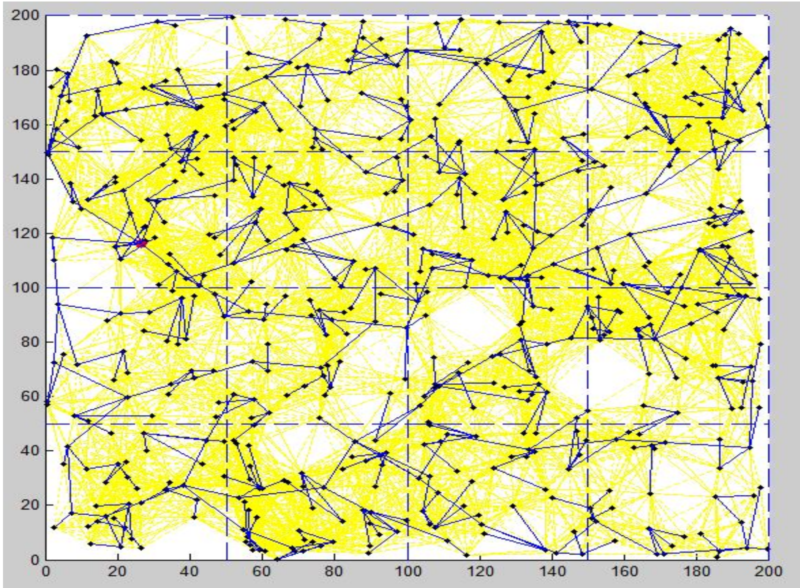

Figure 3.21: BDMRST (600 nodes)

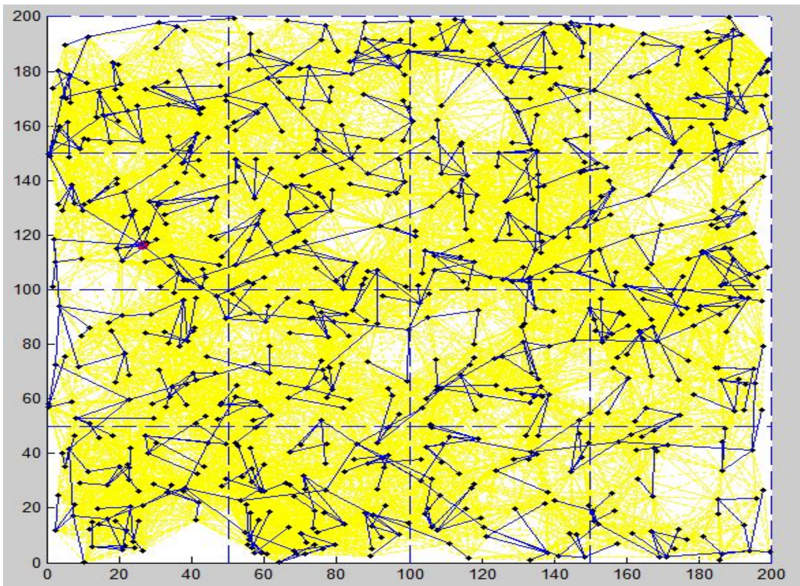

Figure 3.22: BDMRST (800 nodes)

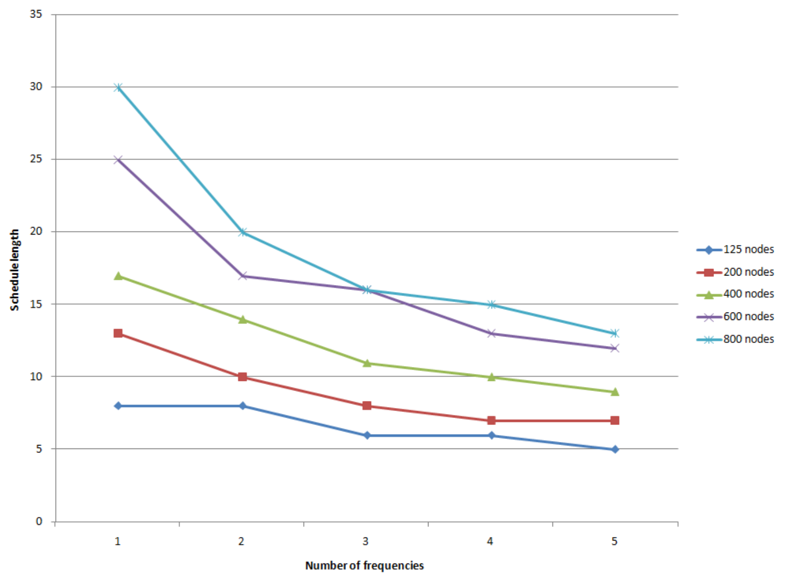

Figure 3.23: Schedule length for BDMRST using even selection

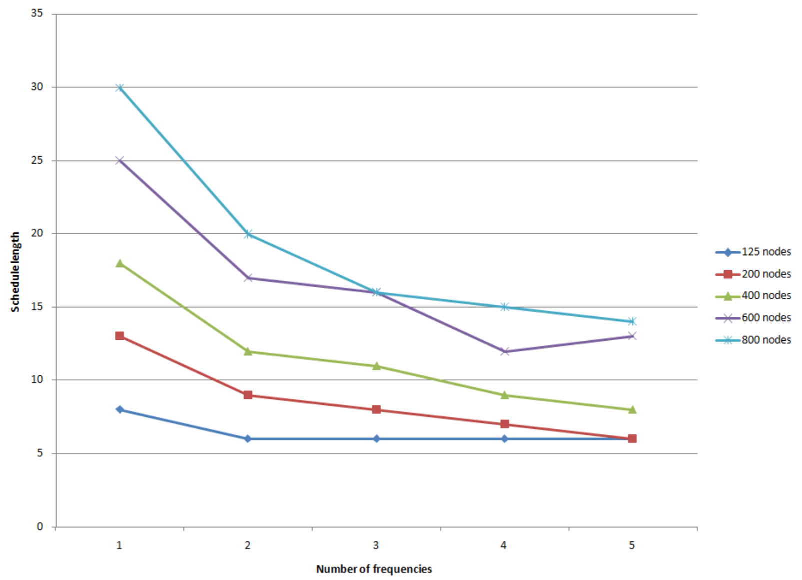

Figure 3.24: Schedule length for BDMRST using load balanced

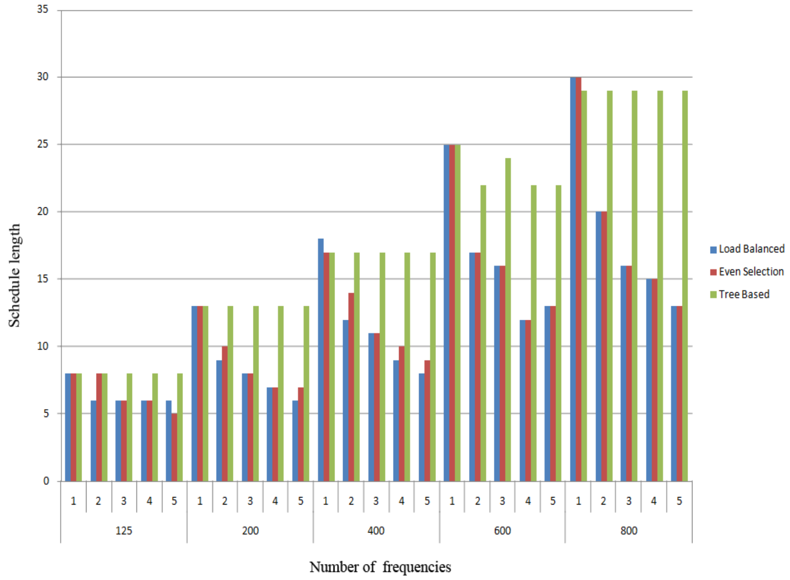

Figure 3.25: Comparison of different frequency allocation schemes used in BDMRST

3.6 Degree Constrained Tree (DCT)

This is the tree in which the degree of the tree is bounded. Thus no node can have children more that the upper bound. The trees in different topologies are varied from upper bound 2-5. The upper bound is usually kept low to increase the data collection rate. But, this also results in increasing the hop count which in turn increases the radius of the network. Which will also bring in some latency.

The combination of degree constrained tree and tree based frequency allocation scheme emulates none other than the TMCP protocol. It has been observed that when the number of frequencies available is greater than the upper bound it does not affect the schedule length when using tree based frequency allocation.

Figure 3.26 and 3.27 shows the graph of the schedule length when using even selection and load balanced frequency allocation respectively.

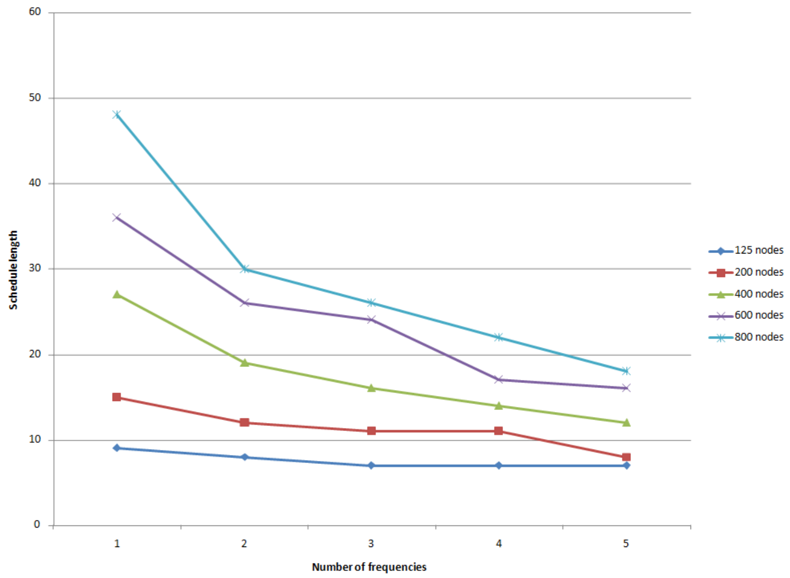

Figure 3.26: Schedule length for DCT (degree=5) using even selection

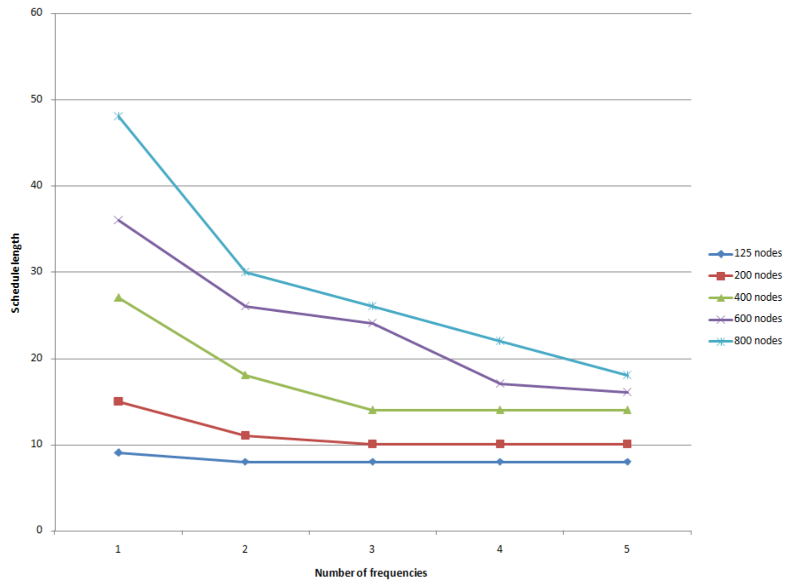

Figure 3.27: Schedule length for DCT (degree=5) using load balanced

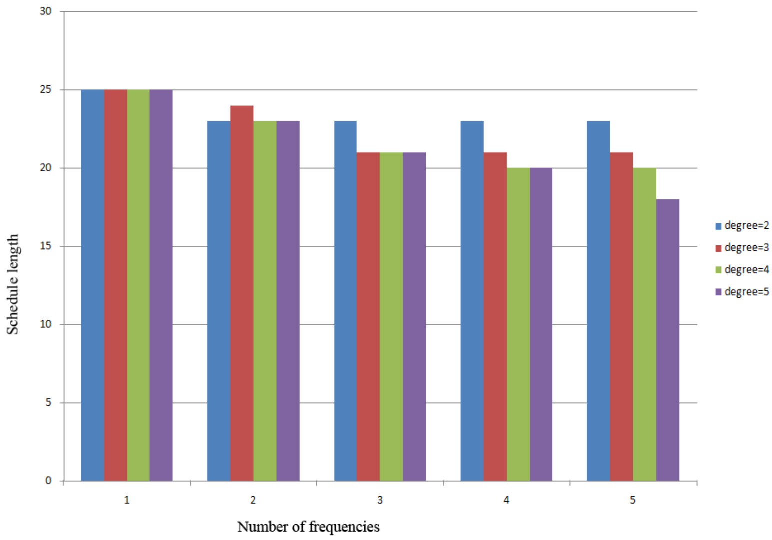

Figure 3.28: Comparison of schedule length varying upper bound of degree (No. of nodes :400, load based frequency allocation)

It can be seen in 3.28 that there is no improvement in the schedule length once the number of frequencies made available are more than the upper bound of the degree of nodes when using tree-based frequency allocation. But as the upper bound is increased the schedule length improves. Since the nodes get distributed among the subtrees to which the frequencies are allocated.

3.7 Comparison of schemes on the basis of density of nodes

As we have seen that the schedule length usually increases as we increase the number of nodes to accommodate them. But since different tree topologies are created with same number of nodes their results are compared in this section. As it has been mentioned in the start itself that number of nodes are increased in a particular area by placing more number of nodes with the existing nodes. So the results have been shown for the different frequency allocation schemes. Graphs in figure 3.29, 3.31, 3.33, 3.35 and 3.37 are for 125, 200, 400, 600 and 800 nodes respectively when load balanced allocation scheme is used. Whereas, Graphs in figure 3.30, 3.32, 3.34, 3.36 and 3.38 are for 125, 200, 400, 600 and 800 nodes respectively when even selection allocation scheme is used.

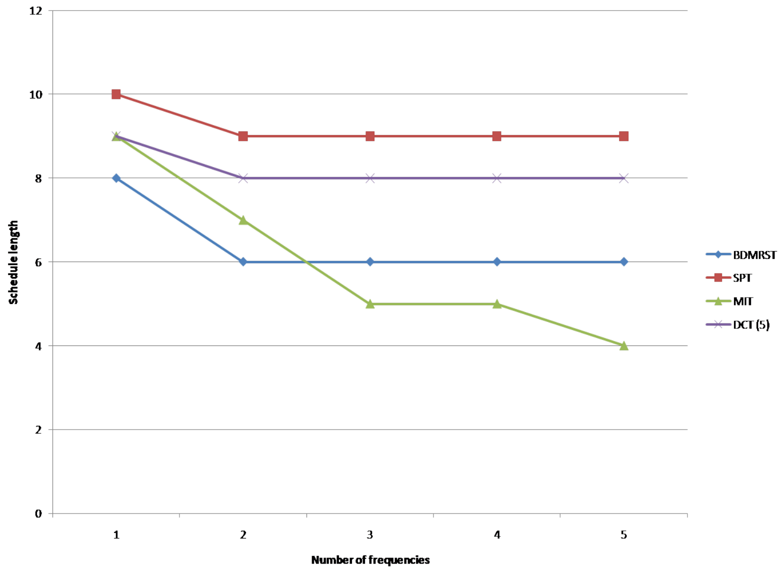

Figure 3.29: Schedule Length v/s Number of frequencies (125 nodes) Load Balanced

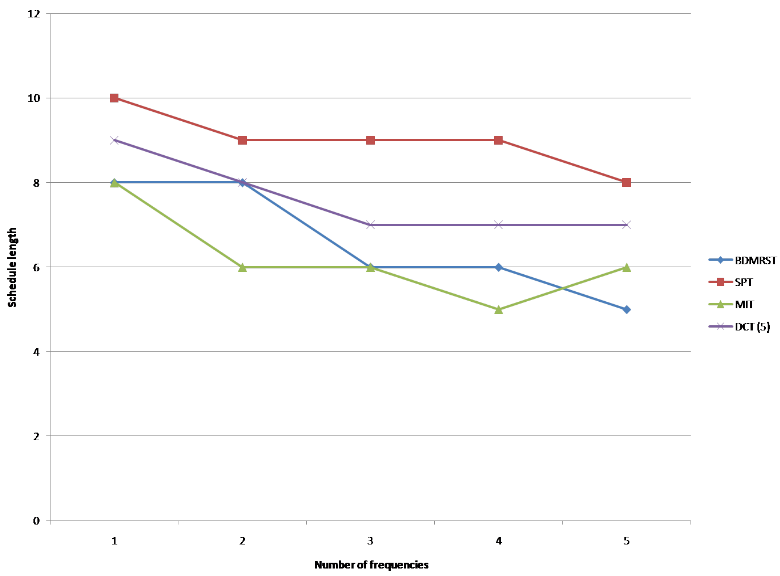

Figure 3.30: Schedule Length v/s Number of frequencies (125 nodes) Even Selection

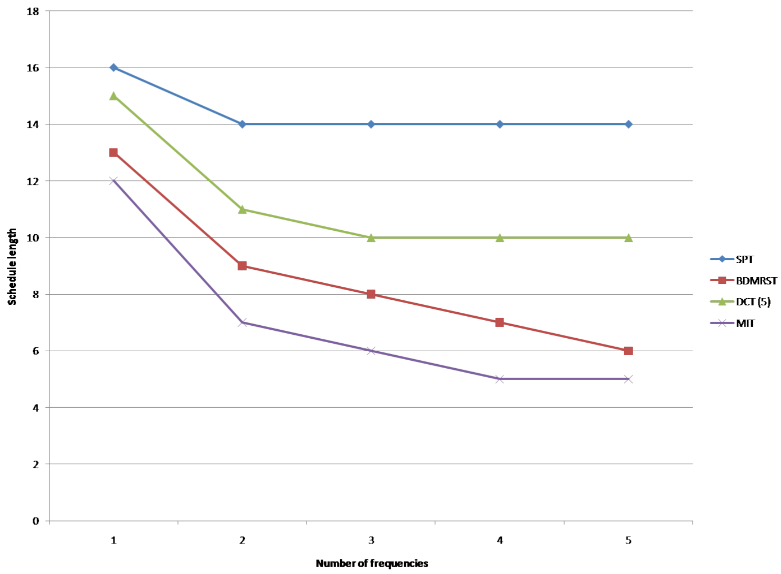

Figure 3.31: Schedule Length v/s Number of frequencies (200 nodes) Load Balanced

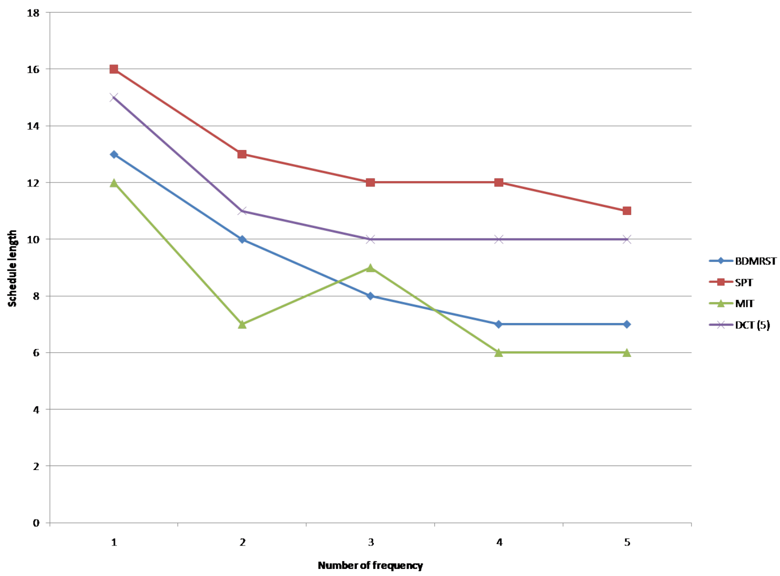

Figure 3.32: Schedule Length v/s Number of frequencies (200 nodes) Even Selection

It has been observed that in case of MIT sometimes the schedule length marginally increases even after increase of number of frequency. It depends on the way that frequency is allocated to the nodes and how the nodes are distributed. This can be observed in the graphs in figure 3.32 and 3.34.

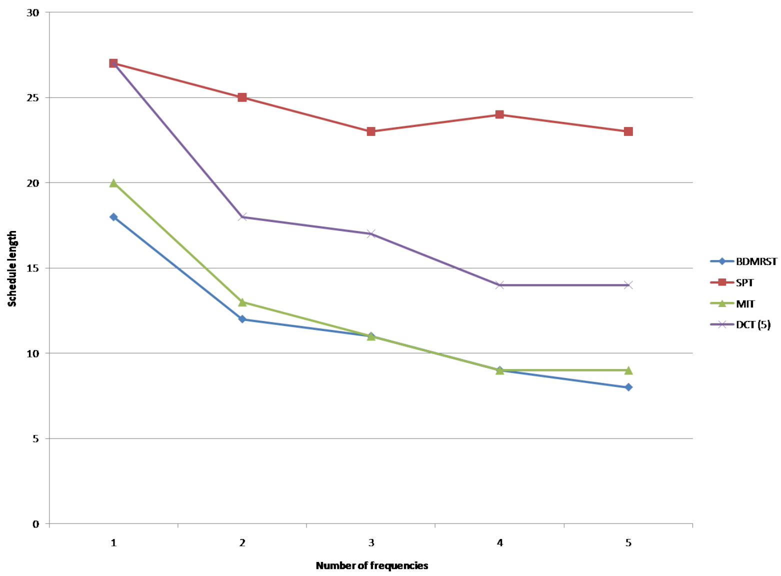

Figure 3.33: Schedule Length v/s Number of frequencies (400 nodes) Load Balanced

From the graphs in figure 3.35 and 3.36 it can be seen that there is hardly any difference in the schedule length when using both the frequency allocation schemes with different tree topologies. The same can be seen in the graphs in figure 3.37 and 3.38. It can be inferred from this that in some dense scenarios it can happen that these two frequency allocation schemes can yield same results.

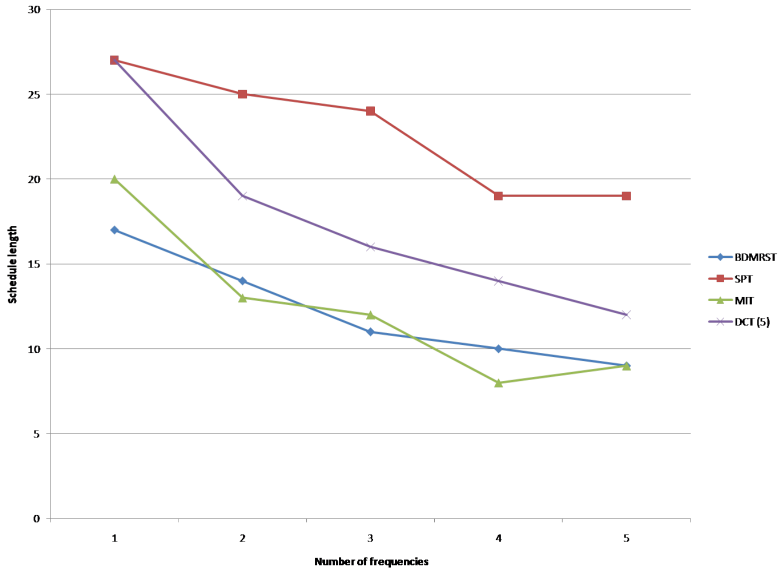

Figure 3.34: Schedule Length v/s Number of frequencies (400 nodes) Even Selection

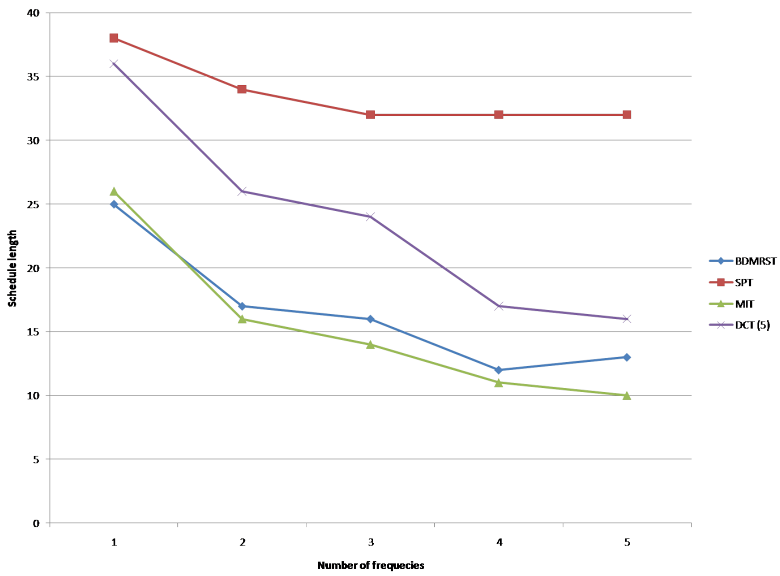

Figure 3.35: Schedule Length v/s Number of frequencies (600 nodes) Load Balanced

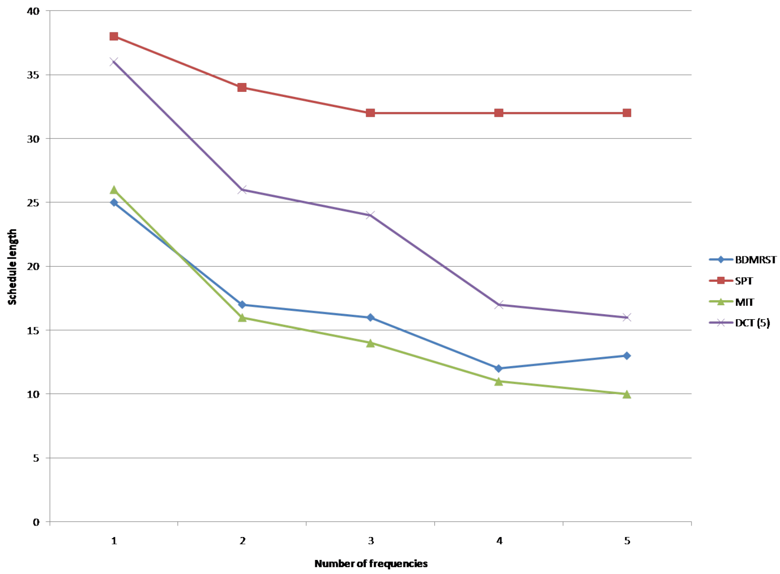

Figure 3.36: Schedule Length v/s Number of frequencies (600 nodes) Even Selection

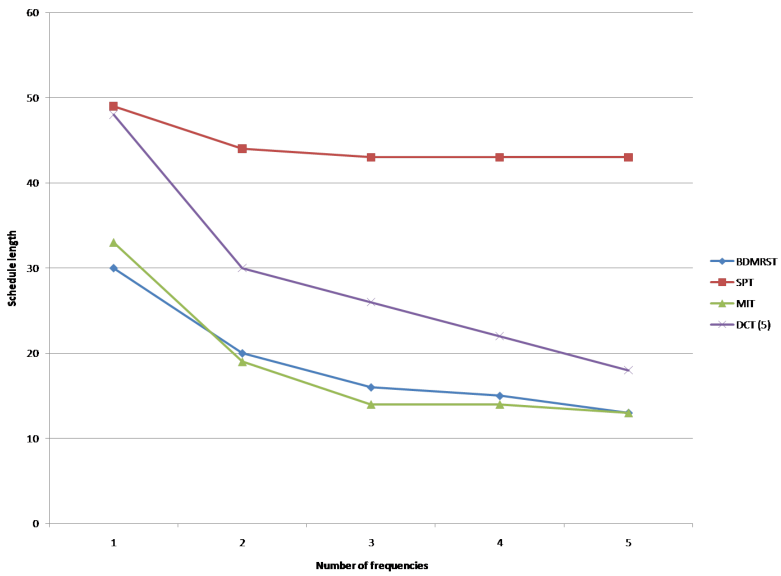

Figure 3.37: Schedule Length v/s Number of frequencies (800 nodes) Load Balanced

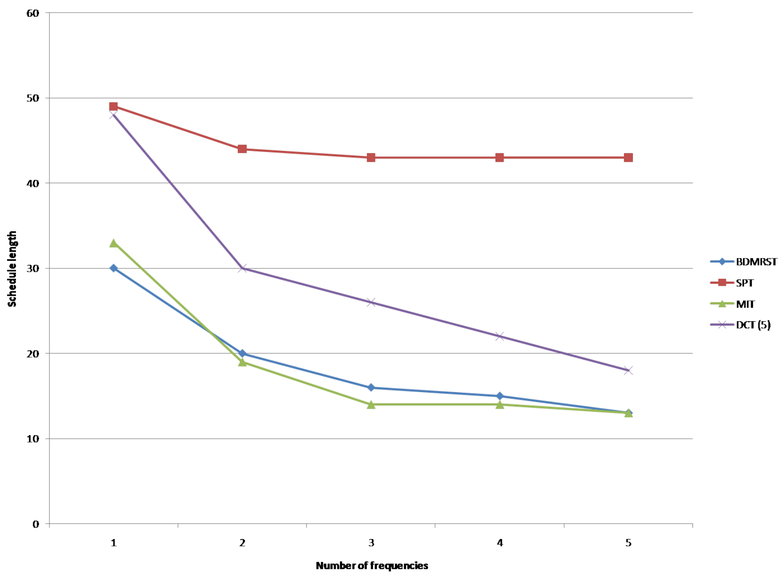

Figure 3.38: Schedule Length v/s Number of frequencies (800 nodes) Even Selection

Chapter 4

Conclusion

By observing the results it can be clearly concluded that use of multiple frequencies if supported by WSN hardware is one of the best solutions for getting rid of interference. It helps in employing simultaneous transmissions which drastically reduces delay. More simultaneous transmissions can be performed with more number of frequencies. Since WSN hardware can only provide limited frequencies, exclusive frequencies cannot be allocated to nodes. So to effectively use the limited bandwidth further TDMA scheduling can be employed. Wherein the nodes which operate on the same frequency are scheduled to use different time slots.

But, what we have observed that once multiple frequencies are employed along with the spatial-reuse TDMA, the data collection rate often no longer remains limited by interference but by the topology of the network. The density of nodes when increased the schedule length also increases to accommodate the new nodes. MIT and BDMRST trees perform better in most of the cases.

Chapter 5

Future Work

Through this work it has been proved that use of multiple frequencies definitely reduce the delay. It has been found that the schedule length reduces as more frequencies are used. But how much is the overhead of tree construction, frequency and timeslot allocation schemes is still unknown. How much time and energy is consumed for this process is definitely one of the major factors in selection of these schemes along with their performance during convergecasting. Transmission power control when employed can further help in reducing the interference by some percentage.

References

[1] V. Annamalai, S.K.S. Gupta, L. Schwiebert, "On Tree-Based Convergecasting in Wireless Sensor Networks",Wireless Communications and Networking, March 2003, VOL 3 pp 1942-1947.

[2] M. Salajegheh, H. Soroush, A. Kalis, "HyMAC: Hybrid TDMA/FDMA Medium Access Control Protocol for Wireless Sensor Networks, in the 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07), Athens, Greece, September 2007.

[3] Ozlem Durmaz Incel, Lodewijk van Hoesel, Pierre Jansen and Paul Havinga, "MC-LMAC: A Multi-Channel MAC Protocol for Wireless Sensor Networks, Elsevier Ad Hoc Networks Journal, Volume 9, Issue 1, January 2011, pp. 73-94.

[4] Y. Wu, J. Stankovic, T. He, and S. Lin, "Realistic and efficient multichannel communications in wireless sensor networks, in Proceedings of IEEE INFOCOM, 2008, pp. 11931201.

[5] Ozlem Durmaz Incel, Bhaskar Krishnamachari, "Enhancing the Data Collection of Tree-Based Aggregation in Wireless Sensor Networks ,

[6] M. Burkhart, P. von Rickenbach, R. Wattenhofer, and A. Zollinger, "Does topology Control Reduce Interference , in MOBIHOC 04, pp. 919.

[7] Amitabha Ghosh, Ozlem Durmaz Incel, V.S. Anil Kumar, and Bhaskar Krishnamachari, "Bounded-Degree Minimum-Radius Spanning Trees for Fast Data Collection in Sensor Networks, INFOCOM IEEE Conference on Computer Communications Workshops , March 2010.

[8] Ozlem Durmaz Incel, Amitabha Ghosh, Bhaskar Krishnamachari, Krishna Chintalapudi, "Fast Data Collection in Tree-Based Wireless Sensor Networks, IEEE Trans. Mob. Comput., 2012, 11(1): 86-99.

[9] 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,

[10] G. Zhou, C. Huang, T. Yan, T. He, J. A. Stankovic, and T. F. Abdelzaher, "MMSN: Multi-frequency media access control for wireless sensor networks, in Proceedings of IEEE Infocom, 2006.

[11] S. Khuller, K. Lee, and M. Shayman, "On degree constrained shortest paths, Lecture Notes in Computer Science, vol. 3669/2005, pp. 259 270, 2005.

[12] O. D. Incel, S. Dulman, P. Jansen, and S. Mullender, "Multi-channel interference measurements for wireless sensor networks, in Proceedings of LCN 2006, 2006, pp. 694701

[13] Baljeet Malhotra, Ioanis Nikolaidis, Mario A. Nascimento, "Aggregation convergecast scheduling in wireless sensor networks, Wireless Networks, 2011, Vol. 17(2), pp. 319-335.

[14] Ozlem Durmaz Incel, Amitabha Ghosh, and Bhaskar Krishnamachari Scheduling Algorithms for Tree-Based Data Collection in Wireless Sensor Networks, Theoretical Aspects of Distributed Computing in Sensor Networks, Eds. S. Nicoletseas, J. Rolim, Springer Verlag, 2010.

[15] MATLAB's Product Help.