Topology Planning for Long Distance Wireless Mesh Networks

Abstract

Cost optimization is an important criterion in technology deployment for developing regions. While IEEE 802.11-based long-distance networks may provide a cost-effective option to connect remote villages, planning such a network to cover the villages in a given region can be a non-trivial task. This is especially so since the antenna tower cost is the dominant cost in such scenarios. In this thesis, we consider the problem of topology planning in the context of 802.11-based long-distance rural networks. Our first contribution is the formulation of this problem in terms of its constraints and the optimization metric. We find the problem to be of combinatorial nature, and hence not scalable for large input sets. Our next contribution is in breaking down the problem into its constituent sub-parts some of which are independent, this makes the solution procedure tractable. Next we propose several searching and pruning strategies in generation of an optimal solution.

We then present a set of results generated by the algorithm. We have been able to find topologies for sets of nodes of size 31 (so far). For a sub-set of the Ashwini (31 out of 34 nodes) project, a long distance rural network currently being deployed in the West Godavari district of Andhra Pradesh we have been able to show a cost benefit of the order of 15 times on their cost savings, by improving the process of assigning tower heights. Also careful topology planning led to a further 22 % savings over their current deployment plan. Another fact to be borne in mind is the current deployment utilizes three independent WiFi channel while our solution only uses one channel. A final contribution of this thesis is the discussion of further work necessary in this domain.

Chapter 1

Introduction

There has been a huge proliferation of Internet and other communication based services in the last two decades. However, this spread is confined to developed countries, and metropolitan pockets of developing countries. This is really unfortunate for developing countries like India, where around 74% of the population is rural and are on the wrong side of the digital divide. Bridging this divide necessitates, providing internet connectivity to each and every village. Providing the same by expanding the current telephone network to rural areas is infeasible because of the huge initial infrastructure costs. Also, deployment of cellular wireless would not be sustainable because of its business model, which demands more high-paying consumer density. Emerging technologies like 802.16 WMAN, have not yet reached the scale of competitive mass production, hence the equipments are expensive. In this regard, the 802.11 WiFi has shown tremendous growth and acceptance as a last hop access solution, because of their low price. Although 802.11 was primarily designed for indoor operation, but [13] has established the possibility of using 802.11 in long-distance networking.

802.11-based long-distance networks has been proposed [13, 19, 15] as a cost-effective option to provide Internet connectivity to rural areas in developing regions, to enable ICT (Information and Communication Technology) services. Fig. 1.1 depicts such a network schematically. It has a *land-line* node with wired Internet connectivity (say, optical fiber). The several villages around are connected with this land-line node, possibly through multiple hops, using long-distance wireless links. Each village has an antenna tower atop which the radio is mounted. The towers are necessary to achieve Line of Sight (LOS) for the long distance links.

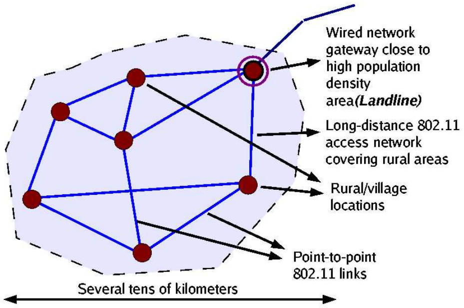


Figure 1.1: Schematic of an 802.11-based long-distance network

From our experience in the Digital Gangetic Plains project, we have learned that the primary cost in 802.11-based long-distance networks is the set of antenna towers [22]. While the 802.11 radio itself costs about U.S.$50 or less, a 30m antenna tower can cost as much as U.S.$1,000. Topology planning for such long-distance networks with emphasis on cost minimization, hence becomes an important problem, and is the focus of this thesis. The network topology is among the first decisions to make while planning a deployment, but it has cyclic dependences with several crucial aspects such as the inter-node interference, the Medium Access Control (MAC) protocol in use, and the resultant application performance seen.

Our first contribution in this thesis is the formulation of the topology construction problem. We stipulate three main constraints: (a) the application *throughput* constraint, (b) inter-link *interference* constraint, and (c) *power* constraint. And cost as the primary optimization objective.

For simplicity, we capture the target application requirements solely in terms of the throughput achieved. The application throughput constraint simply states that a particular throughput is required per village node. Note that the throughput is dependent on the MAC protocol in use. The inter-link interference constraint is closely tied to the MAC protocol in use. In this thesis, we focus on the 2P MAC protocol, as proposed in [19], although other possibilities can also be accommodated in the same problem formulation and solution. The power constraints consist of an upper bound at the transmission end of each link Effective Isotropic Radiated Power (EIRP) regulations, and a lower bound at the reception end of each link (receive sensitivity requirements). As mentioned above the optimization criterion we consider is the *cost* of the resultant deployment. This cost is dominated by the cost of the towers installed [22], and we focus on minimizing the same.

The problem of topology construction can thus be stated as:

Given, (a) the locations ($< x, y, z >$ *co−ordinates*) *of a set of villages to be provided with network connectivity and that of the land-line providing it and (b), the specific bandwidth requirement per village node, what is the minimum cost topology which satisfies the three constraints: throughput, interference, and power?*

Several aspects of the above formulation were abstracted out of a situation where we had to design the network topology for a planned deployment of a 34-node network. This network is being deployed as part of the Ashwini project, by Byrraju foundation, in West Godavari district, Andhra Pradesh, India [7]. At that time (April 2005), for lack of any further knowledge, we designed the network topology based on intuition, without really optimizing the cost in any serious manner. We did however pay attention to the throughput requirement of 384Kbps, which was specified as the minimum requirement per village, to run high-quality video-conferencing between the central location (Bhimavaram) and each of the villages (to be used for educational content, tele-medicine, etc.).

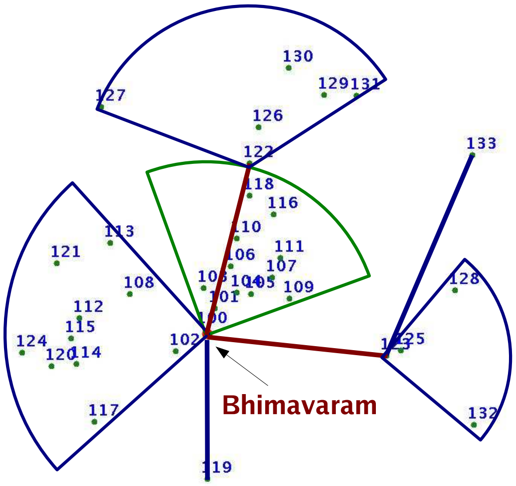


Figure 1.2: The Ashwini Deployment

The locations of the 34 villages in the Ashwini project, as well as the topology we had designed then, is shown in Fig. 1.2. This design assumed a proprietary MAC protocol as supported by a particular vendor for outdoor WiFi products, this MAC allows the user to specify the distribution of available bandwidth (both upstream and downstream), between the competing receivers. We are assuming the 2P MAC protocol in our thesis.

Our next contribution in this thesis is the approach we have taken to solve this problem. For simplicity, we focus on the construction of a *tree* topology. This simply provides connectivity and does not address any fault-tolerance issues. We first break down the problem into four parts: (a) exploring the search space to find a tree (i.e., deciding which links to form) which satisfies all the constraints, (b) assigning optimum heights to tower locations once a tree has been formed, (c) assigning appropriate antennas, and (d) assigning transmit power so that all the links formed can function simultaneously. We are using a Branch-and-bound algorithm for the first problem, a set of Linear Programming (LP) equations for the second problem, after introducing appropriate domain based relaxations. For the third problem which is the antenna assignment problem, we first propose an optimum combinatorial algorithm, and then replace it with a heuristic algorithm of polynomial time complexity. The fourth problem of power assignment is solved by solving an LP formulation, while there have been other proposed power assignment LP formulation, ours is novel and dependent on the MAC considered, (we specify the set of equations for 2P only).

We introduce several search strategies and pruning strategies to reduce the search time, for the first problem. These are based on our domain knowledge. The two constraints are, (a) constraints on the depth of the generated trees, and (b) dynamic cost bounding of the generated sub-trees.

As soon as a tree is formed the two LP problems are solved, the first one to assign optimal height values for the formed tree, and the second one to assign power values for the formed links. Before assigning power values however, we also assign antennas for each link, using the heuristic algorithm.

We have been able to find topologies for sets of nodes of size 31 (so far). For a sub-set of Ashwini (31 out of 34 nodes) project a 15 times reduction in cost on the estimated cost of the overall deployment, by careful tower height assignment. Also careful topology planning led to a further 22 % reduction in cost over their current deployment plan.

In the rest of the thesis we elaborate the above mentioned procedures. Chapter 2 introduces the problem and presents a breakdown of the same in terms of its constituent sub-problems. Chapter 3 details the solution procedure of the first sub-problem of tree enumeration, while Chapter 4 describes the same for the second sub-problem of height assignment, and Chapter 5, Chapter 6 describe the antenna assignment and power assignment sub-problems respectively. Subsequently, Chapter 7 presents a detailed evaluation of our approach. Chapter 8 describes related work. Finally, Chapter 9 presents the conclusion of our work and also discusses the possible future work.

Chapter 2

Problem Formulation And Solution Procedure

In this chapter we give a high level overview of the problem and the solution procedure. In Section 2.1 we describe the equipments required to carry out a deployment. We then introduce a few definitions in Section 2.2, and explain the dependences between them in Section 2.3. In Section 2.4 we first state our assumptions and then give a high level view of the solution procedure devised.

2.1 A Typical Deployment

Long distance rural networks based on 802.11 technology, have been deployed at a few places around the world, examples include [22], [7]. While the previous two projects are in a developing country, even developed countries have been utilizing long distance Wi-Fi links to extend network in remote areas [3]. The idea is to link up distances which are separated by tens of kilometers, by a multi hop wireless network. Each link setup has two 802.11 radios, which is bought off-the-shelf, the radios are connected to high-gain directional antennas mounted on tower tops. The typical deployment of a link can be understood by Figure 2.1. The equipments required to establish a link as well as their costs are summarized in Table 2.1 (partially reproduced from [21]). As can be seen from Table 2.1 the dominant cost in a deployment is that of setting up a tower. Hence, the focus for designing the networks is on tower cost minimization.

WiFi link	**Cost**
Antennae	120
RF cable	120
Connectors & adapters	50
2 wireless bridges	2,000
Antenna tower	6,600
Total	8,890

Table 2.1: Cap-Ex of a typical deployment(in U.S.$)

2.2 Definitions

In this section we introduce the key concepts which will be utilized in problem formulation.

Link, Parent, Child and Tree topology

A link is said to be formed when there is a two way communication between two geographically separated points using, the setup depicted in Figure 2.1. In a link the node that is nearer to the landline node is known as the parent, while the other node is known as the child. Also, we call a generated topology a tree topology if the root node (landline node) has no parent and all other nodes have exactly one parent.

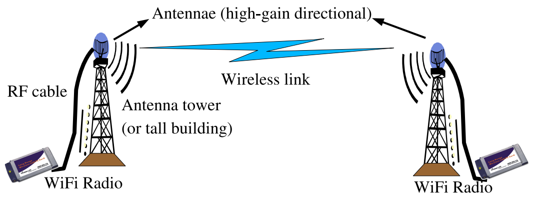


Figure 2.1: Long-distance link setup

Point-to-Point Links

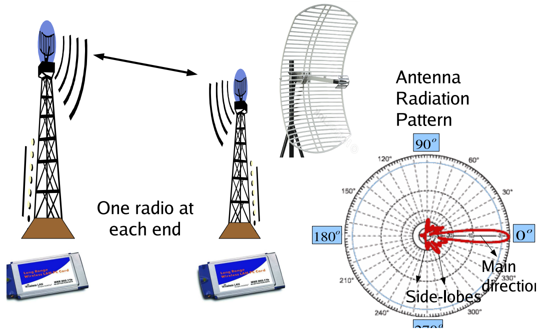


Figure 2.2: *P-to-P* link

A Point-to-Point (*P-to-P*) link is a link involving only two radios (one radio at each end). A *P-to-P* link is shown in Figure 2.2. These links generally have high gain directional antennas mounted on tower tops. The directional antennas have a very small half angle beam-width, and hence can be thought of focusing their energy to a narrow region.

Point-to-Multipoint Links

A Point-to-multiPoint (*P-to-mP*) link set involves one antenna being mounted with a single radio which communicates with more than one radio (at different geographic locations) to form multiple links. This concept can be better understood by Figure 2.3. The common radio point is generally connected to a sector antenna with a half angle beam-width of 30 to 180 degrees. The common radio communicates with all the radios at the other end in a time shared fashion.

Transmit Power and Interference

Transmit power is the power that is being fed into the radio. When more than one communication links are operating in physical vicinity of one another in the same frequency channel, they cause disruption for each other, this phenomenon is known as interference. The concept of interference can be better understood from Figure 2.4. In the figure there are two links, both of which have a common end point at P1. Due to physical proximity (antennas for both the links are on the same tower), the antenna for the lower link receives signal

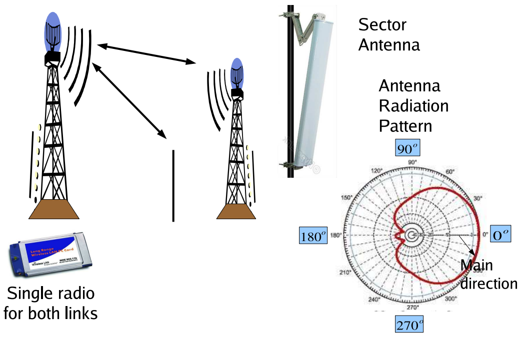


Figure 2.3: *P-to-mP* link set

meant for the upper link as well (and vice-versa). This signal is unwanted and it obstructs proper reception of the designated signal, i.e., it *interferes* in proper reception. For proper link performance the amount of interference should be minimized.

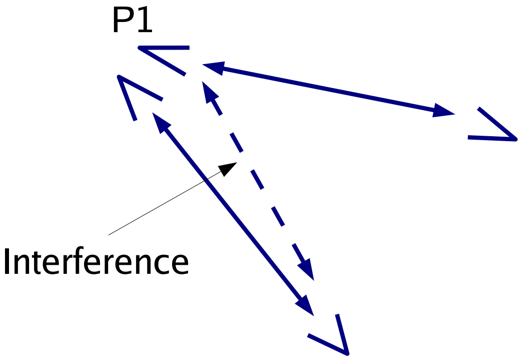


Figure 2.4: Interference between two links

Medium Access Control (MAC) Protocol

The MAC layer decides which radio has the control of the shared transmission media. We consider 2p [19], Time Division Multiple Access (TDMA) based protocols and Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) for long distance wireless networks. In TDMA based systems the links operate in predetermined turns, while in CSMA/CA the links have no predetermined ordering of turns and hence they contend for a small duration to select the link which will operate in the given turn. In 2p multiple links operate simultaneously. The MAC protocol used determines the maximum throughput achievable by a node in the topology (formulated in Section 3.2), and also the transmit power assigned to each node (formulated in Section 6.4).

2.3 Dependences

We now state the various dependences between the above mentioned entities, that have to be taken care of, in order to arrive at a feasible topology for a given set of villages. We end the section by representing the problem of topology construction as a collection of sub-problems, few of which are interdependent.

Throughput depends on MAC

The throughput that can be achieved on a single link depends on the MAC protocol to be used. This can be understood if we consider the case of 2p, TDMA and CSMA/CA MACs. In 2p the links can operate simultaneously, while in the other two MACs the channel is essentially time shared. This assuming the same physical encoding for all three MACs implies that the capacity of 2p would be higher than that of the other two MACs. We further elaborate this aspect in Section 3.2.

Throughput depends on Antenna/Link type

The throughput depends on whether the link is *P-to-P* or a *P-to-mP* link set, in case of the former the link can use the entire throughput allowed by the MAC while in case of the latter the maximum allowed throughput by the MAC is always divided between the constituent links. This is because in a *P-to-mP* link set one of the ends has only one radio, which can only service the set of multiple antennas at the other end in turns. We illustrate this in Figure 2.5.

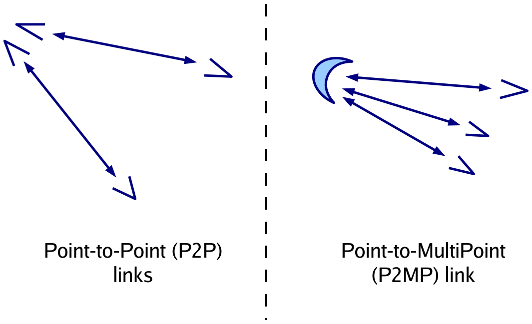


Figure 2.5: Top view diagram of *P-to-P* and *P-to-mP* link set

Transmit power depends on Link length and Antenna type

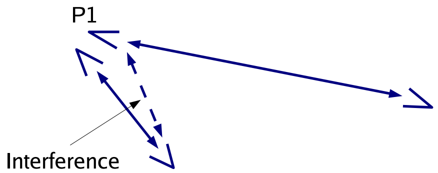
The radio signal transmitted suffers degradation in its strength which is dependent on the distance. Also the antennas have a specific gain (capacity to enhance the signal power) in a given direction. The value of the transmit power is calculated while taking care of the previous points and ensuring that high enough signal is received at the other end of the link. We mathematically formulate these in Chapters 5 and 6.

MAC (feasibility) depends on Transmit Powers

This is specific to the 2P MAC protocol, in which multiple links can operate simultaneously. Simultaneous operation of multiple links necessitates ensuring that the ratio of strengths of actual signal and of interference (SIR) is greater than a system specified margin. This is explained by Figure 2.6. We mathematically formulate this concept in Section 6.1.

Tower Height depends on Link Length

Formation of a long distance WiFi link between two distant nodes is based on the presence of an unhindered line-of-sight (LOS) and clearance of at least 60% of the first Fresnel curve at all the intermediary points. The dimensions of the Fresnel curve depend on the link length. The mathematical formulation of this is provided in Section 4.2.



Signal to Interference Ratio should be above threshold

$$P_R - I_R \geqslant SIR_{reqd}$$

Figure 2.6: Diagram showing relationship between signal and noise strength

Cost of deployment depends on Tower height

The cost of the deployment depends mainly on the heights of the towers used. A point to be noted in this regard, is that the tower costs don't grow linearly with tower heights this is because with height the design of the tower also changes. We state this in more details and then present a mathematical formulation in Chapter 4.

All these dependencies are summarized in Figure 2.7.

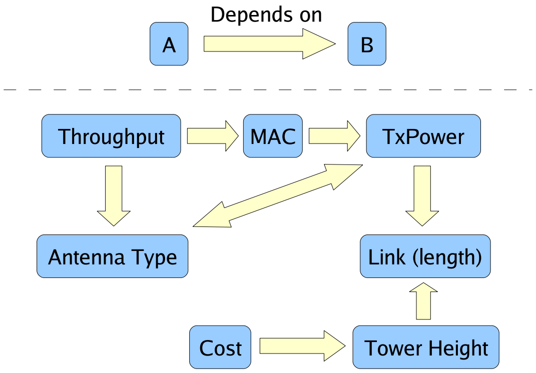


Figure 2.7: The Dependence Diagram of Topology Construction Problem

2.4 Overview of Solution Approach

Before stating our sequence of solving the sub problems we state the simplifying assumptions.

Assumptions

1. Only tree topologies to be generated. Making a tree topology would imply that in case any of the links fail, one or more nodes would become disconnected. Hence, the generated topology would not be fault tolerant. Generation of topologies without this restriction is something we leave open for future research.
2. Towers and masts to be placed only in villages. While they could ideally be placed in any point on the convex hull we consider only this setting and leave the more ideal solution open for future research.
3. 2P MAC to be used, as its more efficient in terms of available bandwidth utilization. We also point out the changes to be made to the formulation in case another MAC is used.

4. The throughput requirement of each node (R) to be fixed (say to 384 Kbps in case video applications are to be supported). This assumption is justified as rural networks are generally to be used for specific applications, which generally involve video and audio applications such as tele medicine, distance education etc.

Solution Approach

Under the above stated assumptions, there remain four sub-problems that need to be solved, they are shown in Figure 2.8.

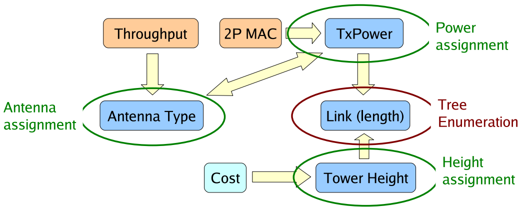


Figure 2.8: The sub problems to be solved

1. **The tree enumeration problem:** This step involves finding all the *feasible* tree topologies. A tree is said to be *feasible* if it satisfies all the constraints imposed on it. We use a exhaustive enumeration algorithm for this. An enumerated tree give the information regarding the parent child relationship that must exist between the various nodes.
2. **The height assignment problem:** This involves assigning optimum heights to the towers to be placed at villages. This step assumes that the parent child relationship between the various nodes is known.
3. **Antenna assignment problem:** This involves assigning appropriate antennas to the radios to placed at villages. This step should have all the parent child relationship information available. The aim of the step is to ensure coverage within half beamwidth region of each antenna of the opposing point of the link, while reducing interference.
4. **The Power Assignment problem:** This involves assigning appropriate power to the radios to placed at villages. Power assignment step should take care of the following: (a) too much power would result in high interference between the various links decreasing the performance, while (b) too less power might result in failure in formation of links.

The exact sequence of solving these sub-problems is as follows.

1. We enumerate all the trees, and for each tree
2. we try to assign appropriate
   (a) antennas
   (b) power values
3. If we are successful
   (a) We assign the heights to the towers at all the nodes
   (b) We then compute the cost of the generated tree and compare it with the previously generated best cost tree, the lower costing of the two is saved.

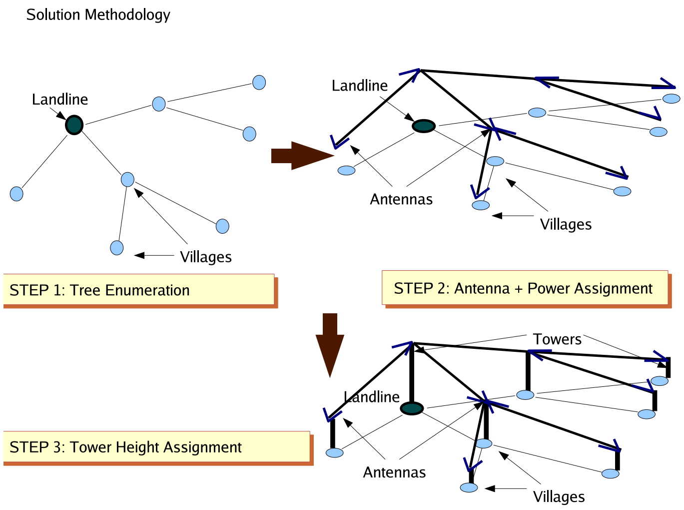


Figure 2.9: Sequence of steps in topology generation algorithm

The sequencing of steps is shown in Figure 2.9.

The first step of determining the parent child relationship between the various nodes and the third step of assigning heights have a dependency on each other. We can either form out the trees (decide parent child relationship) and then assign heights to the towers to be placed so that the links actually form, or we could assign arbitrary heights first and then decide the feasible parent child relationships. We chose the former sequencing, as given the parent child relationship between the various nodes, we could formulate the optimal height assignment problem as a LP.

Also, transmit power and antennas can not be assigned before deciding the parent child relationship. The assignment of transmit power and antenna are dependent on each other as well this is because, the amount of transmit power needed to form the links and the interference caused to other links is determined by the by the gain of the antennas put up at various locations. Similarly, we can also argue that if the powers to be assigned between various links are known, we can assign the antennas as well. We have chosen the first methodology. Also, we are assigning the antennas while considering only a node and its child set instead of considering the entire topology tree, the reason for doing so is that the problem of antenna assignment seems to be equivalent to minimum set-cover problem which is a NP complete problem, hence a more generalized version of the problem would be even harder to solve. Also, we assign the power values to links globally, which ensure that only topologies in which the links can operate together are selected.

Chapter 3

Tree Enumeration Problem

In this chapter we explain the first of the four sub-problems that we solve. The tree enumeration problem involves finding out the parent child relationship between the various nodes. While the number of trees is exponential in the number of nodes, not all trees would be viable topologies. Hence, in order to reduce the number of trees to be searched (the search space) we have formulated a set of constraints which has to be satisfied if the tree is a feasible topology tree. In Section 3.2, we describe the constraint placed by the MAC protocol to be used on the topology formed. While in Section 3.3 we describe the depth restriction criteria used, we describe a search space minimization heuristic employed by us in Section 3.4.

3.1 Tree enumeration algorithm

In order to explore the search space we use Branch-and-Bound technique. In this method each node is assigned a set of feasible children (based on whether the link length is below certain threshold). The root node in each iteration selects a combination of children (from the feasible child set), and then all the child nodes (of the root) do the same. While each of the nodes is selecting its child set various constraints are used to prune it. Pruning operations are carried out at each step (ie., after each node been assigned one possible child combination).

3.2 The capacity constraint placed by MAC

For the planned long-distance networks generated by our algorithm, all the traffic is to be routed through one root node. The link between a child and parent can support only a fixed amount of traffic which depends on the physical layer encoding (PHY) used by the MAC, this constraint is cascaded to all other nodes down the tree as well. The value of maximum traffic is determined by the PHY used, for 802.11b PHY this value is 11 Mbps, while for 802.11a/g PHY this value is 54 Mbps.

We illustrate this concept with Figure 3.1. Here, B_{max} denotes the maximum available bandwidth over a link under the current MAC, and R denotes the bandwidth requirement of each node. Hence, the maximum number of villages that can be supported by a parent which has the B_{max} link to its own parent is $\frac{B_{max}}{R}$.

In this section we describe the constraints placed by the various MAC protocols on the generated topologies. We only consider the MAC protocols discussed in Section 2.2. We first consider the 2P MAC [19], which is especially suited for planned long distance wireless networks. We then describe the constraints that need to be placed in case the MAC is changed to any Time Division Multiple Access (TDMA) based MAC and Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) respectively.

2P MAC

2P is a MAC designed specifically for long distance wireless networks. In this MAC a set of *P-to-P* links which have a common end point simultaneously transmit or receive (Synchronous operation). With careful

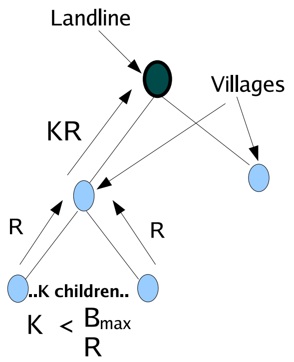
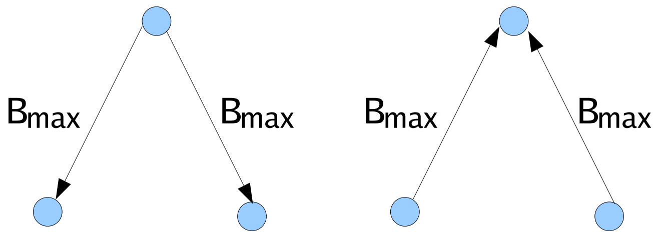


Figure 3.1: Maximum number of villages that can be present in a subtree is $\frac{B_{max}}{R}$



(a) Simultaneous transmission by two links at B_{max}

(b) Simultaneous reception by two links at B_{max}

Figure 3.2: Synchronous operation in 2P MAC

power assignment, such links have been shown to operate at B_{max} throughput individually. We illustrate using Figure 3.2, here the two links which share a common end-point are able to simultaneously transmit and receive at B_{max}. In 2P weighted allocation of time slots between the siblings is also possible this ensures that a parent now can distribute the available bandwidth between its children in a weighted manner. Also, in 2P multiple links (*P-to-P* or any one of a *P-to-mP* link set) which have a common end point can operate simultaneously provided certain power assignment constraints are satisfied, we provide the mathematical formulation of the same in Section 6.4. Hence, the only constraint on capacity is that none of the non root nodes should have more than $\frac{B_{max}}{R}$ descendants in their subtrees, provided all the links from the root node are *P-to-P*. We ensure that this is the case in our algorithm.

TDMA MAC

As explained in Section 2.2 a MAC based on TDMA, would have all the links working mutually exclusive of each other in time domain. Hence, inn case of normal TDMA if the number of nodes is greater than $\frac{B_{max}}{R}$ we cannot ensure a fixed bandwidth of R. This is possible in case of 2P as simultaneous operation of links is not allowed in normal TDMA protocols. This is illustrated in Figure 3.3. However, as mentioned previously weighted allocation of time slots is possible in TDMA based protocols. So, if the number of villages to be served exceeds $\frac{B_{max}}{R}$ then ensuring a fixed bandwidth while only one root node is available, is infeasible, in case the number of nodes is less than $\frac{B_{max}}{R}$, then none of the nodes can have more than $\frac{B_{max}}{R}$ children.

CSMA/CA MAC

In CSMA/CA, while the down-link traffic can be time scheduled (the traffic from parent to children), in case of up-link traffic this is not possible. This is beacause for each up-link time slot there exists a contention period when all the child nodes who have data to send try and gain control of the time slot. Under such circumstances, each node has an give equal priority to send data in the nest time slot. Hence whatever

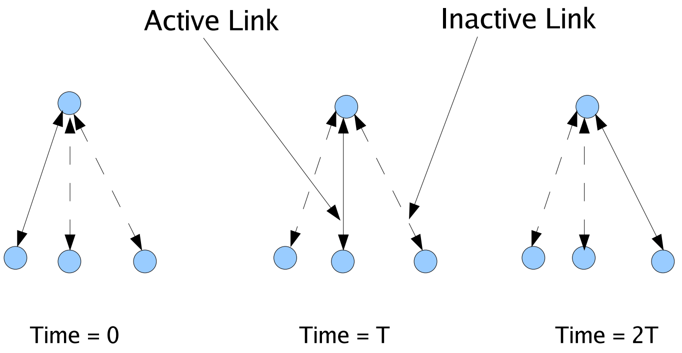


Figure 3.3: Mutually exclusive (in time) operation of links in TDMA protocols

bandwidth is received by a node is equi-distributed among all its children (*i.e.*, weighted allocation of time-slots is not possible). This implies that we have to ensure that a node which is really deep down the tree gets his promised share (*i.e.*,the depth of the tree is restricted by capacity constraint, the promised bandwidth and the number of children at each level). The functioning of the CSMA/CA protocol is explained by Figure 3.4. Here, the link that will send its data is not pre-decided, and hence over a long period of time all the nodes have equal amount of channel share.

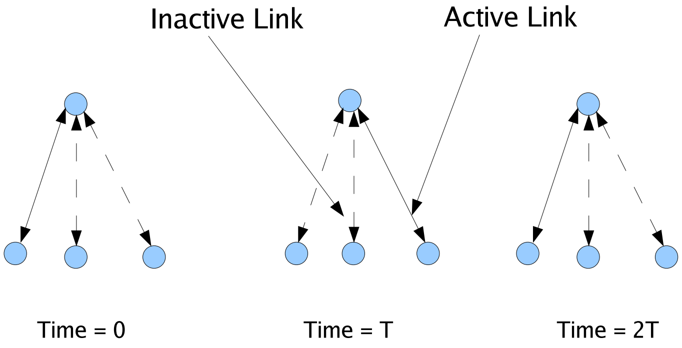


Figure 3.4: Equi-priority working of links in CSMA/CA protocol

The above mentioned constraints can be written as follows,

(a)Number of nodes (excluding the landline node) $\leq \frac{B_{max}}{R}$

(b) For all nodes (j) which are not root,

$C_j = \frac{C_i}{S_i}$
$C_j \geq S_j * R$

Where i is the parent of j, C_k denotes the maximum allowed bandwidth for a node, S_k denotes the number of descendants of a node k (i.e., all the nodes which are present in the sub-tree rooted at node k) and R denotes the bandwidth requirement of the nodes.

3.3 The Depth Restriction

In this section we describe the constraint on the depth of the valid topology trees formed by the tree enumeration algorithm. By *depth restricting* a tree we mean fixing the depth of the tree to some predefined value and not expanding a tree along any vertex in case this restriction is violated. The reason for adding this constraint was to reduce the search space. We have restricted the depth of trees to 2, choosing this value had two reasons, (a) as stated previously it reduces the search space of the tree enumeration algorithm, and hence reduces running time, (b) the restriction of the tree height to two hops makes formulation of the height assignment problem as a linear programming problem possible, this point would be explained in detail in Chapter 4.

While restricting the topologies to two hops seems like a significant constraint on the type of topologies formed, which would affect the validity of the results generated. In Indian context it is not so, as in India 85% of the villages are within 20 Km radius of a fiber (Point-Of-Presence) POP [18]. Assuming the maximum link length to be 15 Km, gives the the generated topologies a diameter of 60 Km, hence, ensuring that the solution would work for most situations.

3.4 The Search Pruning strategies

The search space for the tree enumeration problem is exponential in nature, hence to reduce the time we are using two heuristics based on our domain knowledge. We describe them next.

Eliminating long edges to begin with

Before starting the enumeration algorithm we restrict the set of feasible edges to only those whose length is below a certain threshold which is to be decided by the user. In our evaluation we take the value to be 10-15 km in our evaluations. This assumption is valid in the sense that even though links of greater length can be easily formed, most of the nodes can be linked up with smaller edges, and it leads to a substantial reduction in the search space as instead of searching a search space of the order of n^n we search in a space of size $(n-k)^{(n-k)}$, where n is the number of nodes and k is the number of nodes which are eliminated from each nodes child set due to height thresholding.

Dynamic cost bounding

At any given time addition of new nodes to a topology tree can only result in increase of cost. Hence, it would make sense if we stop expanding all those sub-trees whose current cost has already exceeded that of the current known best topology.

Now the assignment of cost (which in turn depends on height) and the design of topology are interdependent on each other as pointed out in Section 2.4. As the optimum cost of topology could only be determined when the entire topology is known. However, we can always assume a lower bound on the cost of the generated topology using the following observations,
Observation-1: The height of a tower at should be at least high enough to have LOS with its child set.

Observation-2: Given a link-length, we can lower-bound tower height necessary at one end, assuming the presence of mast at the other end (this results in minimization of cost, as we would show in Chapter 4).
These, two observations together imply that at the time of generation of a subtree, we can lower-bound the cost. Which is what we do, and once the cost of a subtree grows beyond that of the current best cost topology, we stop growing it further and backtrack. Also, the set of lower-bounds of tower heights for the various links is pre-computed for efficiency.

Chapter 4

Assigning Optimal Heights

In this chapter we formulate the problem of height assignment for the towers to be placed at various village locations once the parent child relationship between all the nodes is known. Formation of a long distance WiFi link between two distant nodes is based on the presence of an unhindered line-of-sight (LOS), (under the assumption that received power level, which we are ensuring anyway) and clearance of at least 60% of the first Fresnel curve at all the intermediary points. Now the problem of ensuring the required clearance between two distant points with some hindrances in between has more than one solution (i.e., the towers can be of different heights and still attain LOS), hence the need to find the optimum height assignment (least cost one). We describe the process of assigning the heights to the towers in this chapter. In Section 4.1 we give the Non-linear programming formulation of the height assignment problem. In Section 4.2 we describe the two issues in formulating the problem as an LP, (a) obstruction height assignment problem and (b) piece-wise nature of the tower cost function, we then describe the relaxations for the same. In Section 4.3 we suggest a solution strategy for the obstruction height assignment problem.

4.1 The Height Assignment problem formulation

As mentioned previously, once a topology tree is generated (by the Branch and Bound algorithm) the next step is to optimum heights to the towers at all the nodes. This problem can be stated formally as, *Given the parent-child relationships between a set of n nodes, what it the height of tower to be placed at each location so that (a) LOS and (b) minimum Fresnel clearance is achieved and the resultant topology is least cost possible.* This requirement is captured in the Figure 4.1.

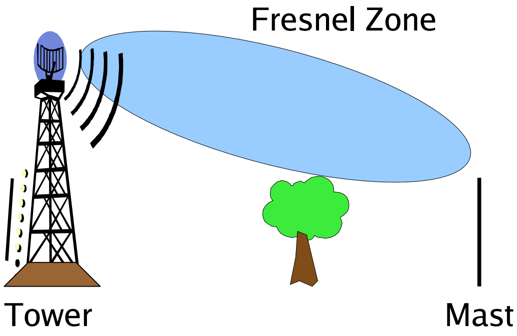


Figure 4.1: The Line-of-sight requirement of long distance Wi-Fi links

We now formulate the optimal height assignment problem. Let us assume that we have two nodes 1 and 2, where node 2 is the non-leaf node and 1 the leaf node. This is shown in Figure 4.2.

where,

h1,h2 = heights of the two towers,

D = distance of separation between the two points
L = height of the obstruction
d = distance of separation of the obstruction from location 1

Given that we need to ensure LOS between the two nodes, the heights of two towers placed at points 1 and 2 is related as follows,

$$h1 * (D - d) + h2 * d \geq L * D$$

This is derived by trigonometry involving triangles ABC and AEF, in Figure 4.2.

So, for a set of n nodes with n-1 edges between them the formulation becomes,
We can derive similar equations for all the n-1 edges in a n node tree. Let,

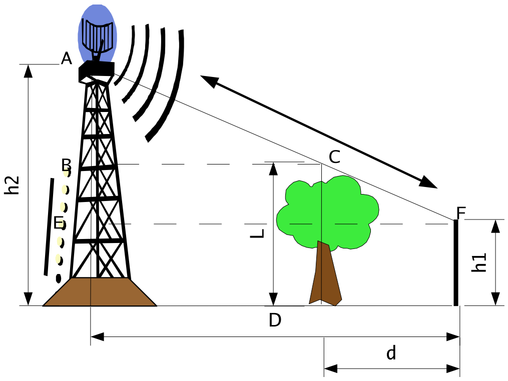


Figure 4.2: The minimum clearance constraint on the tower height

h_{thres} =Upper bound of mast heights.

A_i is the cost coefficient of the tower at point i, which for root and non-leaf nodes, would be that of a tower and for leaf nodes would be that of a mast.
Also, Assume the obstruction height and the distance of separation is the same for all the points. Hence, the formulation becomes,

$minimize(\sum A_i * h_i)$,
Subject to,
for all nodes (i,j) which have an edge between them,
1. $h_i * (D - d) + h_j * d \geq L * D$, where h_i is the child of h_j
2.(a) $h_k \leq h_{thres}$, k is a leaf node
2.(b) $h_{thres} \geq h_k$, k is a non leaf node

4.2 The non-linearity of tower cost function and unknown obstruction heights

However, there are two major issues that need to considered, before formulating an algorithm to assign cost optimal heights to the towers. They are as follows.

The Obstruction height assignment

As the links are long-distance, potentially tens of kilometers apart, we need to know the exact height of the maximum obstruction, and its exact location, before we can go on and assign tower heights to clear it.

This is necessary because if significant obstructions are present in between the two antennas at two ends the signal would get substantially reduced in power due to diffraction caused by the obstruction. This can be understood by seeing Figure 4.3.

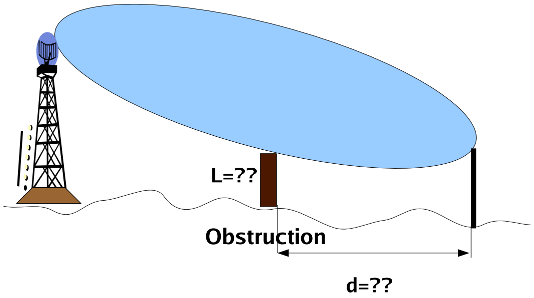


Figure 4.3: The Obstruction Height Location Problem

The piece-wise linear cost function

The cost of the towers is not linear in height, this is so because the basic build (technology) of tower changes with height. Up to a height of 12-15 meters we can simply use metallic water-pipes (mast) as a substitute for regular towers, while for heights above that threshold would necessitate using proper towers. Figure 7.1, shows the nature of the curve that needs to be optimized for getting optimum cost topology, clearly this is piecewise linear and hence can't be modeled as a LP problem. **will put figures of mast and tower**

Domain Based Relaxations

Relaxing Obstruction Height Assignment

As stated in Section 4.2, its necessary to know the exact height and location of the maximum obstruction. To relax this we make the assumption that the maximum obstruction height is a constant and its location from one of the tower co-ordinates is known. We will remove this assumption in Section 4.3.

Relaxing the piece-wise linearity of tower heights

In order to force the linearity on the tower cost function, we make two assumptions (a), the land line node will always be a tower. and (b),we constrain the tree generated to be two hops. That two hop topologies are still realistic for Indian scenario is already justified in Section 3.3.

Given the above two assumptions, on the generated topologies there are only two feasible placement choices, for masts and towers, either (a), the towers are to be placed at the leaf nodes and the towers at non-leaf nodes or (b), the masts are to be placed at all the non-leaf lodes except root and the towers are to be placed at the leaf nodes. These choice are depicted in Figure 4.4, under the assumption that masts cannot be used at both ends for link formation.

We can now prove that the topologies of first kind would be at least as cheap (if not cheaper) by considering the following observation.
Observation: The number of non-leaf nodes in a tree is always greater than or equal to the number of leaf nodes.

This observation coupled with the fact that masts are order of magnitude cheaper than proper towers, proves that the topologies in the first arrangements would be at least as cheap (if not cheaper) than the topologies with the second kind of arrangement.

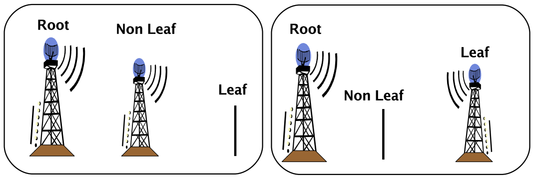


Figure 4.4: The two choices for tower type assignment

4.3 Methodology to calculate obstruction heights

One of the main issues in determining the height of towers erected to achieve LOS is the knowledge of the obstructions present along the straight line. Determining the same by manual survey would be a tedious process as this would involve covering tens of kilometers. In this section we describe an automated mechanism to calculate the same.

The Shuttle Radar Topography Mission (SRTM), an international project spearheaded by the National Geo-Spatial Agency (NGA) and National Aeronautical and Space Administration (NASA), have made terrain elevation data of the entire world freely available on their website [10]. This data which has an resolution of 3 Arc seconds (which translates to height measurements on a grid of dimensions 90 meters), can be used to approximately infer the terrain height values between any two end points.

Also given the aerial separation between two points, we can calculate the radius of the Fresnel curve as well by using the following formula,
The radius of the Fresnel zone can be calculated from the following formula:
$R(n) = M\sqrt{\frac{N}{f}\frac{D1D2}{D1+D2}}$ Where R (n) is the radius of the nth Fresnel zone
M = 17.3, R (n) in meter and D1, D2 in km
= 72.1, R (n) in feet and D1, D2 in statute mile
as mentioned earlier wireless link formation clearance of (60% for 2.4GHZ) first Fresnel zone is required.
We make an simplifying assumption that, the angle of elevation of the line segment joining the two antennas with the horizontal axis, would be small as the difference between the tower heights at both ends is in order of meters while the link length would be in order of kilometers. This essentially means that we can assume the Fresnel curve to be horizontal, i.e., the radius of the curve can be assumed to vertical.

Given the previously calculated height values of the intermediary points and the Fresnel radius at those points we can easily calculate the point of maximum obstruction. This is the point which if cleared would result in clearance of all the other points.

The main issue with this method would be the (lack of) accuracy of the SRTM3 data. Hence, an independent study was undertaken to find out the accuracy of the theoretical readings [11]. The main aim of the study was to find if prediction of tower heights can be carried out based on the available satellite data. We found the actual feasible tower height values are within a 2 meters of their predicted values, hence leading lending credence to the method.

The detailed methodology of extraction of height values form the SRTM3 format and calculation of tower heights as well as the evaluation can be found in [11].

Chapter 5

Antenna Assignment

In this chapter we present the formulation of the antenna assignment problem and an algorithm to solve it. As discussed in 2, there is an inherent dependency between the selection of antennas to be put up at a tower and the assignment of the power valued to them, in Section 5.1 we formulate the antenna assignment problem. In Section 5.2 we present an heuristic algorithm for the same problem.

5.1 Problem Formulation

As mentioned in Section 2.4 after finding out a topology tree we assign antennas to the various nodes. The antenna assignment is carried out for a parent and its set of children. We make the domain knowledge based assumption that the child node always has a high gain directional antenna focused towards its parent. This assumption is justified because a child node has to link to its parent node only. As pointed out in Section 2.4 the antenna assignment is solved in a local context (i.e., only for a node and its children). Whereas is should had been solved globally i.e., the assignment of antenna to a node should be dependent on the entire tree formed and not only on its child set. This does not result in any violation as we take care of the ultimate aim of ensuring proper working of all the links by assigning power globally. The reason for choosing this arrangement is that we can formulate the power assignment problem as an LP, while even the local version of the antenna assignment problem seems equivalent to set cover problem which is NP complete.

During antenna placement also we focus on ensuring the lowest cost possible. Also we say that a link is formed only when the end points are covered by the half beam width region of the antennas at the other end. Given this the problem of antenna assignment can be formulated as follows,

Given a the co-ordinates of a parent node and that of a set of child nodes, find the (a) type, (b) count (c) orientation of different kinds of antennas to be put up at the parent node so that (a) every child node falls within the half beam-width of one of the antennas placed at the parent node and (b) the cost of deployment is minimized.

The algorithm is combinatorial in nature. Even though the input size would be small, but as the antenna assignment algorithm is to be run for all the non-leaf nodes of each tree generated, the running time becomes an issue.

Hence, we look for an algorithm which is polynomial time and ensures the coverage of all the points within the half beam width region. The fact that algorithm needn't result in optimal cost assignment is not an issue because the antenna are far less costlier than an tower as shown in Table 2.1. Also, this problem seems to be equivalent to some variant of set cover problem, though we have not been able to prove the same. Hence, instead of concentrating on formulation of an optimum algorithm we designed an heuristic algorithm for the problem. In the next section we state an heuristic algorithm which has the above mentioned desirable properties.

5.2 Heuristic algorithm for antenna assignment

In this section we mention an heuristic algorithm with $\O(n^2)$ time complexity, where n is number of child nodes. In the heuristic algorithm we focus on minimizing the interference between the various nodes instead of minimizing the cost which is not a dominant factor as pointed out already.

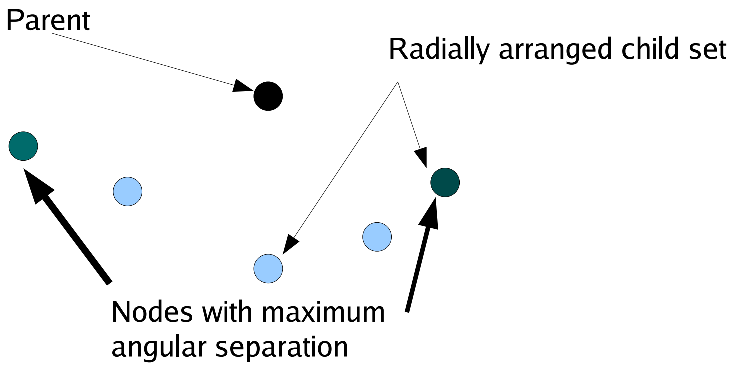
The heuristic algorithm
INPUT:

1. The co-ordinates of the parent node and the child set.
2. Radially arranged child set.
3. The cost price the various kinds of antennas.
4. Tables which given the angle with the main axis of the antenna gives the gain of the antenna in that direction (for all the antennas).

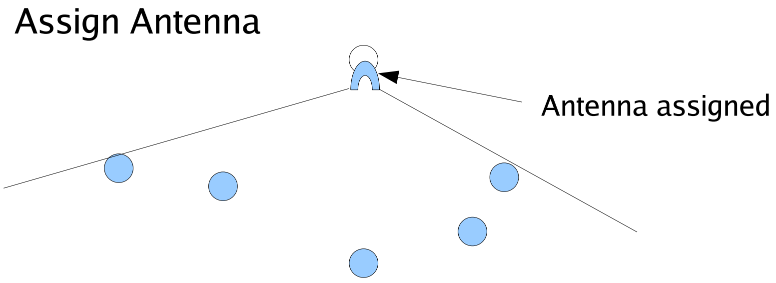
ALGORITHM:

1. Rearrange the child set so that the points with maximum angular separation are the ones which farthest apart.
2. Recursively do the following
   (a) Assign an antenna which can cover all the points within this set and set cost equal to cost of this antenna
   (b) Find the neighboring points which have the maximum angular separation, and along this point divide into two sets and call the algorithm on these two points, return the configuration from the previous run or this run based on which one is the lower costing one.
3. **TERMINATION (of recursive function):** If the set of points have an angular spread (the angular gap between the farthest two child points), less than that of a directional antenna return the cost of the same.

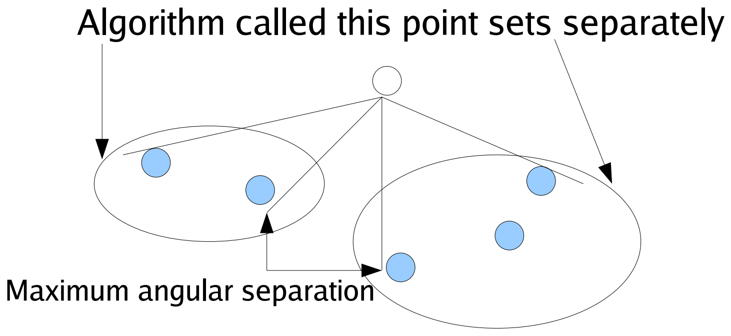
The functioning of the algorithm can be better understood from Figure 5.1. Radial ordering of the points is required to ensure that we can determine the point of maximum angular separation by comparing the points with their neighbors only. The reason for finding the maximum angular separation point between the points is interference minimization, as incrementing angular separation ensures the the main lobes of the two antennas are also separated.



(a) The radially arranged point set is provided as input



(b) Maximum bounding antenna is found and its cost is stored



(c) The points with maximum angular separation is found and the algorithm is called recursively on the two subsets

Figure 5.1: Diagram explaining the working of the antenna assignment algorithm

Chapter 6

Power Assignment

One critical issue to be addressed while designing the topologies of long distance Wi-Fi networks is proper power planning. This essentially means deciding how much power is to be fed into the radios present at either end of each link. The power assignment is dependent on the antennas being assigned and the link lengths. In Section 6.1, we briefly introduce the factors to be kept in mind while assigning powers. The next section describes the importance of satisfying the power equations, especially in case of planned long distance links form the field studies conducted by us. In Section 6.3 we formulate the power equations, and in Section 6.4 we demonstrate the additional power requirements to be satisfied for the various MAC's being used.

6.1 Factors in assigning power

As discussed in Chapter 4, clearing the LOS is critical in formation of a long distance Wi-Fi link. Apart from the tower height, (a) the power transmitted, (b) the pathloss, (c) the power received and (d) interference from other sources determine the feasibility of link formation. We next elaborate the above mentioned four factors briefly. There is a restriction on how much power can be transmitted by a transmitting radio, which is in terms of Effective Isotropically Radiated Power (EIRP) . EIRP is defined as *is the amount of power that would have to be emitted by an isotropic antenna (that evenly distributes power in all directions) to produce the peak power density observed in the direction of maximum antenna gain. EIRP takes into account the losses in transmission line and connectors and the gain of the antenna* [4]. The constraint on power transmitted is that it should not cross the maximum of 30dBm, for any single link [1].
Pathloss refers to the loss of transmitted RF signals as they propagate through the media, the pathloss model used by us is
PL $=log\frac{c}{4\pi fd}^2$, which is known as the Frii's free space pathloss model. Where f in MHz is the frequency of the radio waves, d in km is the distance of separation between the points, and c is the speed of light.

The equation for recieved power in db scale then becomes,
$P(r) = P(t) + G(t) + G(r) - log(\frac{c}{4\pi fd}^2)$

Since for the IEEE 802.11b the f is 2.4 GHz, so the above equation reduces to

$P(r) = P(t) + G(t) + G(r)(32.5 + 67.6 + 20logd)$
$i.e., P(r) = P(t) + G(t) + G(r)(100.1 + 20logd)$

Where, d is the link distance measured in km.
Also each receiver has a sensitivity level which signifies the minimum expected strength of the received signal, which in general for 802.11b radio equipment working at 11Mbps is -85dBm [19]. This signifies that the received power should be above this threshold level which is calculated as,
$P(r) \geq -85$

Another factor is the interference which is defined as the noise signal received from other links working in the same channel,for meaningful link formation.This is the signal to interference ratio which should be of the order of 10 to 15 dbm [19].
ie, $P - I \geq SIR_{reqd}$.

6.2 Importance of satisfying equations

The links in long distance networks are of the order of tens of kilometers, and hence the necessity to ensure proper power assignment is not that clear. Hence, we decided to study the effects of interference on the performance of the long distance links. In order to study the effects on the throughput of interference between long distance links, we carried out an experimental study. We present the results of the same and in this section. We define inter-link interference to be those from non-WiFi sources, or from WiFi sources in neighboring deployments and those from our own deployment. typically its expected the interference in rural settings would be mostly from other links in the same deployment.

6.2.1 Experimental Study

For the experimental study we selected a 23 km link, between Indian Institute of Technology Kanpur (IITK), and Banthar. In IITK a parabolic grid antenna was mounted atop the Faculty Building at an height if 33 m. While at Banthar a similar kind of antenna was mounted on a 25 m tower. Both ends had Single Board Computer from [9], along with Senao 2511CD plus ext2 PCMCIA cards based on Prism2 chipset as the WiFi radio.

The Banthar-IITK link has high error rate due to significant presence of other WiFi sources in the vicinity at site Banthar. These WiFi sources form the source of inter-link interference. We confirmed this by performing an active scan at Banthar. The active scan showed as many as a dozen different WiFi sources.

The experiment consisted of sending 1400 byte UDP packets at an 100 ms interval from one end at a fixed transmit power. We had a series 30min runs for each experiment and this ran for a duration of 2 days. The error rate averaged over the 30min durations is shown in Fig. 6.1. The SNR was about 18-20dB for the various runs. We see that there are lengthy durations of high packet error rates. These error rates are as high as 50%. Expt 20 (about 1am) shows the first large spike in the error rate. We found it strange that other interference sources were operating at this time, but the pattern repeated the next day too. It is likely that the interfering WiFi links are used for late-night data backups in addition to regular Internet access. The other WiFi sources remain active through most of the working day, up to about 4pm in the evening.

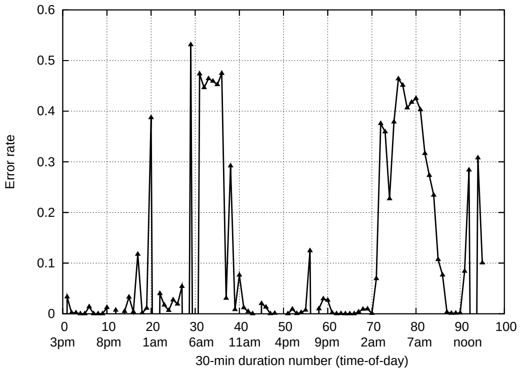


Figure 6.1: Error rates over various 30min runs

Implications

For a long-distance link with directional antennae, interference from any other WiFi source is a classic case of the *hidden node* problem. For instance, in our IITK-Banthar link, WiFi sources near Banthar are hidden from the transmitter at IITK. Of course this can be addressed partially by the use of RTS/CTS. However, RTS/CTS has two issues. The first is that it would add to the overhead in the protocol. The second is that it may not necessarily work if the two links are operating in adjacent, overlapping channels (e.g. channels 1 and 3). For example, suppose another WiFi source W were operating in channel-3 near site Banthar, and IITK-Banthar were operating in channel-1. Then W may not hear the RTS/CTS exchange between IITK and Banthar, and may still end-up interfering with Banthar. It is common for many deployments to be configured in channels other than the three non-overlapping channels 1, 6, and 11. In such a scenario, the RTS/CTS would not necessarily solve the issue.

The extent of the effect seen due to interference on the long-distance links is rather startling. Fig. 6.1 indicates that error rates can be as high as 50%. Thus, unlike in [12], we see a direct cause-effect relation between interference and packet error rate. Further, unlike in urban community networks, for a long-distance link setup with significant infrastructure investment (antenna towers) to operate with 50% error rate seems rather wasteful.

Naturally, at such high error rates, application performance (TCP/video) will be affected critically. Networks with long-distance links are already being laid out, e.g. the Ashwini project [7], and intended to be used for applications such as high quality two-way video. Video performance is affected significantly even with error rates as low as 1%.

This experiment proves that extra care has to be taken of proper power assignment during topology planning.

6.3 The Power equation formulation

In this section we formulate the power equations. The power equations are formulated globally, and solved as an LP. This ensure that all the links (as determined by MAC) can function simultaneously. We use the notation $A \Rightarrow B$ to denote the fact that node A is transmitting towards node B. Hence, we are concerned about the interference at the antenna at node B.

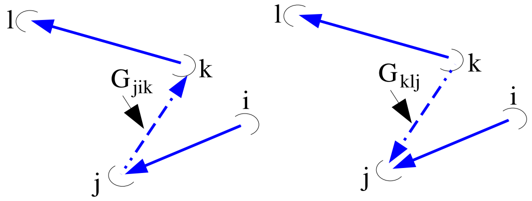


Figure 6.2: Figure showing G_{jik} and G_{klj}

Let, P_{ij} denote the power transmitted by the antenna placed at i designated for j towards j.
G_{ijk} denote the gain of an antenna placed at i designated for j in the direction of k.
and, PL_{ij} denote the pathloss suffered by the signal while traversing from node i to j. In case two links $i \Rightarrow j$ and $k \Rightarrow l$ interfere with each other, the strength of interference at node j is calculated as,
$Interference(I) = P_{kl} * G_{jik} * G_{klj}$
This is illustrated by Figure 6.2. **NOTE:** All the power values are in absolute scale (instead of db scale).

Hence, the net interference at a node j, is the summation of all the interference at the node from all the interfering links.

We now formulate the set of linear equations for power assignment problem. We do not consider the set of fixed losses due to RF cables and connectors in our formulation, but these are negligible and also they can easily be replaced by a constant upper bound (so that the losses are never greater than this value). For a given link $i \Rightarrow j$, the following three criteria should be satisfied,

1. The transmitted power by i, $P_{ij} * G_{ijj}$ should be lower than a maximum value P_{max}, this is the restriction placed by spectrum regulatory authority of the country (like FCC in USA and WPC in India) in EIRP [1]. This can be written as,

$$P_{ij} * G_{ijj} \leq P_{max}$$

2. The recieved signal strength by j, $\frac{P_{ij}*G_{ijj}*G_{jii}}{PL_{ij}}$ should be higher than the receiver sensitivity P_{min}.
   $$\frac{P_{ij}*G_{ijj}*G_{jii}}{PL_{ij}} \leq P_{min}$$

3. The ratio between the signal strength and noise strength at node j should be greater than a bound SIR_{reqd}.
   $$\frac{\frac{P_{ij}*G_{ijj}*G_{jii}}{PL_{ij}}}{\sum_{(<k,l>|<k,l>\epsilon R)} \frac{P_{kl}*G_{jik}*G_{klj}}{PL_{kj}}} \geq SIR_{reqd}$$
   Where, R is the set of all links which interfere with the link $i \Rightarrow j$.

Where, the noise strength is calculated as the summation of the noise signal strength of all the interfering links at the receiver. The nodes that are interfering are selected based on the MAC protocol. The selection procedure of interfering nodes in case of 2p, TDMA and CSMA/CA MACS is described below.

6.4 The dependence of power equations on MAC

2P MAC

In 2p protocol a set topology needs to be bipartite, and the two partitions of nodes, switch between simultaneous transmission and simultaneous reception phases, such that if one end of an edge is in reception phase the other is in transmission. This are shown in Figure 3.2. While using 2p protocol,for a link $A \Rightarrow B$ the various links which interfere is summarized below.

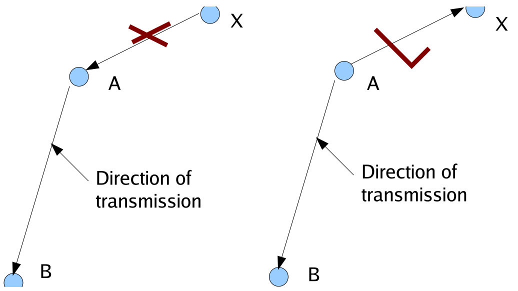


Figure 6.3: Figure showing node A in transmit mode

1. If the link is $X \Rightarrow A$ it **doesn't interfere**. This is because according to 2P MAC, the direction of transmission would be from $A \Rightarrow X$. We depict this in Figure 6.3.

2. If the link is $A \Rightarrow X$ and the antenna is not the same as that of $A \Rightarrow B$ it **interferes**. This is because both the links would be in transmit mode and as the antennas are not common i.e., the links are not part of a *P-to-mP* link set which would mean that they can not operate simultaneously.

3. If the link is $X \Rightarrow B$ it **interferes**. This is because the node B is in reception phase and, hence would be receiving traffic from its child as well. We show this in Figure 6.4.

4. If the link is $B \Rightarrow X$ it **doesn't interfere**. This is because, as mentioned above node B is in receive phase and hence, can not transmit.

5. For link $C \Rightarrow D$ which does not satisfy any of the above mentioned conditions, it **interferes** $A \Rightarrow B$.

Here X, C and D denotes any arbitrary node.

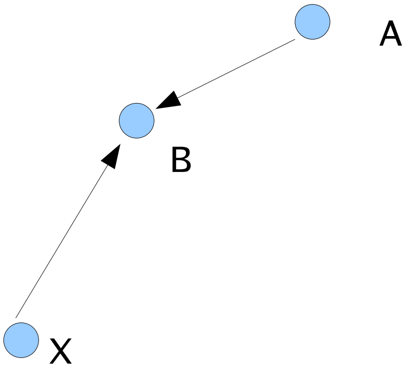


Figure 6.4: Showing node B in receive mode

TDMA MAC

For TDMA protocol, the selection of interfering nodes again depends on what kind of TDMA is taking place, in case the TDMA is local to a node (ie, all the parent nodes are running TDMA between the child nodes without any synchronization between themselves), for a node $A \Rightarrow B$ all the links which don't have either of the vertices common (ie., either A or B as vertices) would potentially interfere. While if the TDMA is a centralized one none of the nodes interfere.

CSMA/CA MAC

In CSMA/CA protocol all the nodes contend simultaneously for the media, this implies that for a link $i \Rightarrow j$, all other links would potentially interfere with it.

6.4.1 Modification for *P-to-mP* link set

The above formulation holds true if the only links are of type *P-to-P*, however we consider topologies with *P-mP* links as well. In this case, the not all the links which are supposed to interfere with a link $A \Rightarrow B$ do not actually interfere. This is because as pointed out in 2.2 the constituent links of a *P-to-mP* link set are not active simultaneously. This is illustrated in Figure 6.5, where the link $A \Rightarrow B$ interfers, with links $C \Rightarrow D$, $C \Rightarrow E$ and $C \Rightarrow F$. However these three links can not be active simultaneously because they share the same radio. Hence, the actual criteria to be satisfied in such case is that for *P-to-mP* link set interfering links consider only the constituent link which causes the maximum interference for the current link. However this would make the interference constraint non-linear (due to inclusion of max() function).

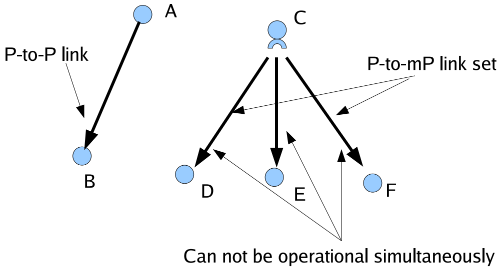


Figure 6.5: Links of an interfering *P-to-mP* link set are not active simultaneously

Hence, the other alternative is to write an interference constraint equation for each of the constituent links of the *P-to-mP* link set. While this solution is optimal it greatly increments the number of constraints to be satisfied. We solve this problem by replacing the interference constraint for a *P-to-mP* link set by an average case constraint. In average case constraint ensure that the interference contributed by a *P-to-mP* link set is the average of its constituent links. This method however suffers from the drawback that the maximum intereference of a link would be higher than the average case interefence, and hence, the resultant links might not function. A point to be noted, is that we assign power on a per link basis. This assumption in turn implies that the radio equipment placed at the common node (parent) of the *P-to-mP* link set have the ability to change their transmission power on a per packet basis. Such functionality is available in radios developed by Atheros [**?**].

Chapter 7

Results

In this chapter we seek to evaluate whether our algorithm generates topologies which are feasible, and least cost. Also, we wish to document the efficiency gained by applying various constraints to the tree enumeration problem as described in Chapter 3. We also present a few unique observations made by us regarding usage if antennas for topology planning that we made during simulations.

7.1 Details of Coding and Libraries

Coding and simulation was done in C on a Linux based Intel PC. We used QSOPT [8] linear programming library to solve the two linear programming problems (one for power assignment and the other for tower height assignment).

The cost function curves for the tower utilized is plotted in Figure 7.1. The tower cost figure were taken from the site of [2], while the mast costs were enquired from local market. The details of the antenna radiation patterns were taken form [5]. We used Breadth First search to enumerate out all the feasible trees. Also we assume the following values for the various user input constraints,

(a) Maximum link length varies from 10 to 15 km, (b) height of obstruction (l in Figure 4.2) is 18 meters, (c) maximum height of a mast (h_{thres}) is 15 meters, (d) the minimum SIR (SIR_{reqd}) is 15 db, and (e) the value of minimum recieved signal strength is -85 dbm.

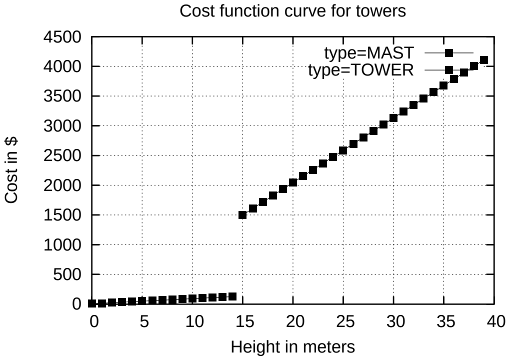


Figure 7.1: The Height Vs Cost plot used for simulations

A point to be noted is the setting of SIR_reqd to 15 db gives us a head room of around 5 db from the minimum required SIR [19].

7.2 Runtime

The runtime of the tree enumeration algorithm is exponential in the number of the nodes considered, this can be seen from Figure 7.2. This figure shows the time required to enumerate out all the trees of depth 2 without using any of the additional pruning constraints mentioned in Section 3.4. It also plots the time required to enumerate the number of all the trees in presence of the constraints. The runtimes are for a specific set of nodes, and hence might vary with another set. The set of points have been generated randomly using the rand() function present in stdlib.h library of standard C. The generated points are within a square of sides 40 km.

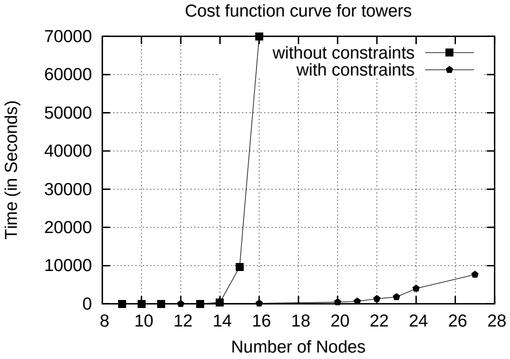


Figure 7.2: Comparative plot of runtimes with and without constraints

7.3 Evaluation of topology for Ashwini project

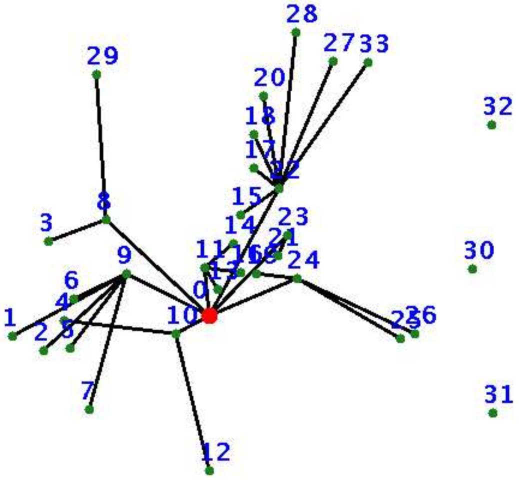


Figure 7.3: Topology generated by the algorithm for Bhimavaram (31 out of 34 villages)

We have run our algorithm for a set of 31 villages (out of 34) considered for connectivity in the Ashwini Project. The current deployment is being carried out following the deployment plan shown in Figure 1.2. As mentioned previously the current plan uses all three available independent channels. The topology generated by our algorithm is shown in Figure 7.3. The cost of the topology generated by our algorithm is $ 55771.39, while the cost of the hand crafted topology shown earlier in, Figure 1.2 is not known totally, we have the information of a few links that have been already deployed by them, we find that a gross over estimation in height assignment is being carried out. We present the present height values and the theoretically required height values in Table 7.1 (partially reproduced from [11]), the procedure for calculating these values is presented in [11]. From this table the average over estimation is 15 times the theoretically required value, suggesting a potential 15 times savings in tower deployment cost. Also, even after assigning heights properly, the cost of current topology is $ 71251.40. This translates into a savings of the order of 22%. This savings is achieved by virtue of careful topology planning.

The reason for not being able to find topologies for the entire set of 34 nodes is probably the assumption of 15 km of link distance and the fact that we have made the topologies two hop.

One key observation made by us is as follows.
While experimenting we found that usage of sector antennas with beamwidth greater than 30° is not feasible for power assignment, as their half beamwidth covers a huge area which makes it difficult to assign power. Hence we use only directional antennas of 8° half beamwidth (Model No. HG2424G) and sectoral antennas of 22° (Model No. HG2418P) and 30 degree (Model No. HG2414P) for assignment. The radiation patterns for these antennas have been taken from [5].

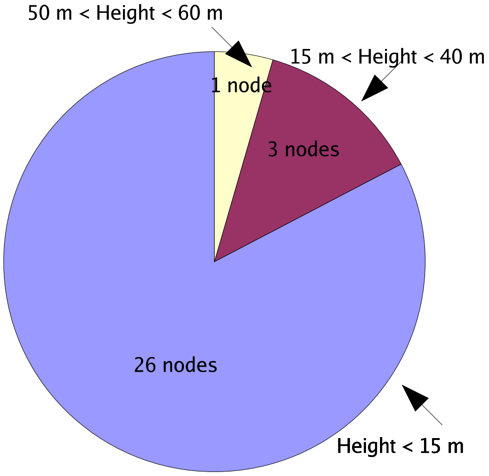


Figure 7.4: Height distribution by the algorithm for Bhimavaram (31 out of 34 villages)

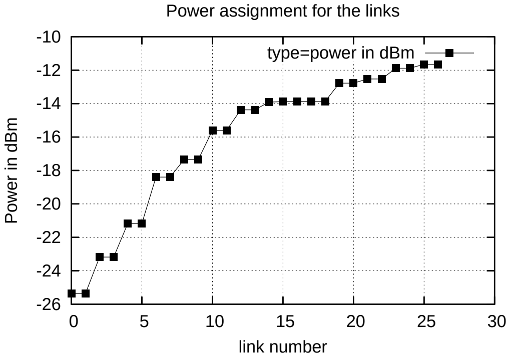


Figure 7.5: Power distribution by the algorithm for Bhimavaram (31 out of 34 villages)

We next present a plot showing the height distribution and power distribution generated by the algorithm for the Bhimavaram topology. The height assigned by the algorithm is shown in Figure 7.4. As can be seen from the pie-chart, the most of the links are assigned masts, while the non-root nodes are assigned towers, the root node is assigned a tower whose height is greater than 45 meters (around 50 meters). The assignment of masts to the nodes results in lowering of cost as pointed out earlier. The power assigned is shown in Figure 7.5, we have sorted the power values in increasing order. All the power values are within bounds of maximum and minimum allowed power values.

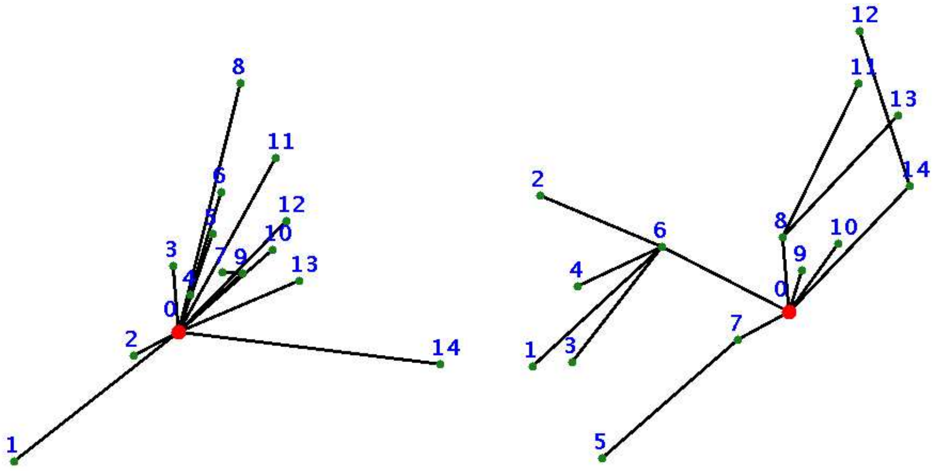
Parent Node	**Child Node**	**Present Child height (in m)**	**Optimum Child height (in m)**	**Estimated Savings (in $)**	**Estimated Savings Factor**
Pippara	Kasipadu	18	4	1744.90	22.6
Pippara	Alampuram	30	8	2977.10	20.2
Bhimavaram	Kesavaram	30	14	2866.10	11.8
Bhimavaram	Korukollu	27	8	2650.55	18.1
Bhimavaram	Charukumilli	24	15	2194.50	8.7
Bhimavaram	Juvvalapalem	30	14	2866.09	11.8

Table 7.1: Estimated cost improvements for the various links of Bhimavaram

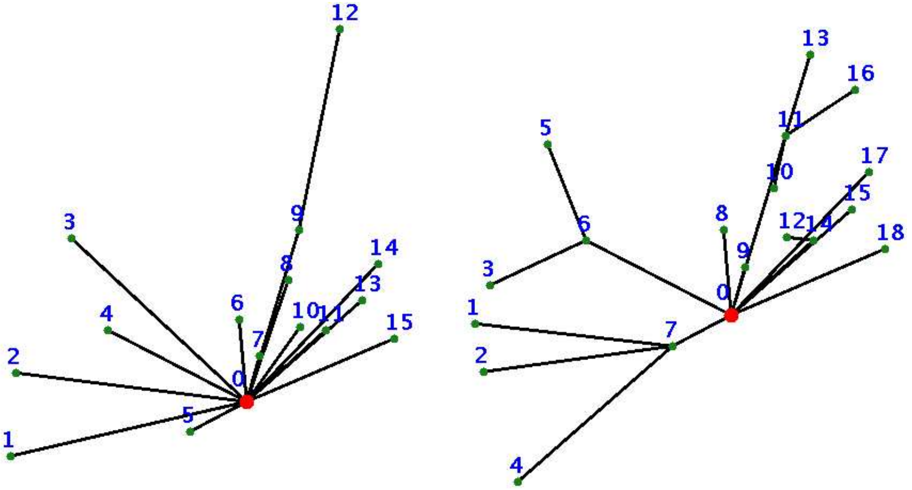
7.4 Evaluation of Random topologies

We have also evaluated our algorithm for randomly generated point sets, to verify if it works out for various scenarios. Topologies for a few randomly generated point set is shown in Figure 7.6. Point sets were generated using methodology pointed out in Section 7.2. The following general comments can be made about the topologies generated, they are bushy in nature. The majority of the nodes are assigned with masts. The height distribution pie-charts for the same set of topologies is plotted next in Figure 7.7.

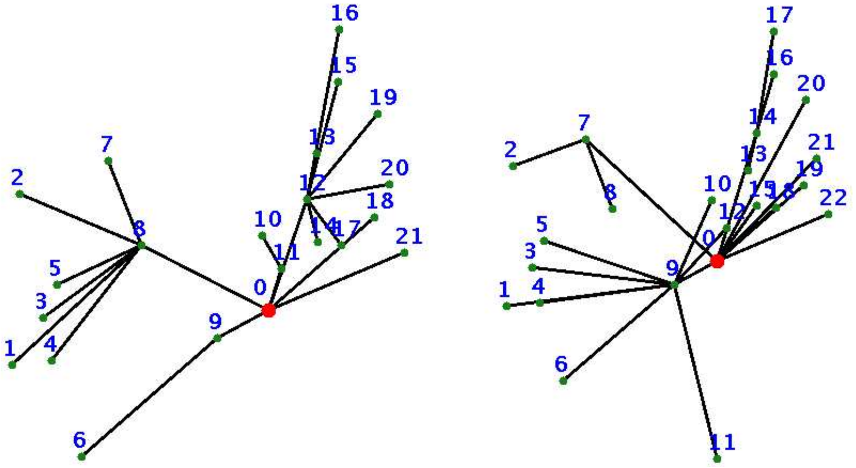
We also present the power assigned to various links by our algorithm in Figure 7.8.



(a) Topology for randomly generated point set 1 (b) Topology for randomly generated point set 2

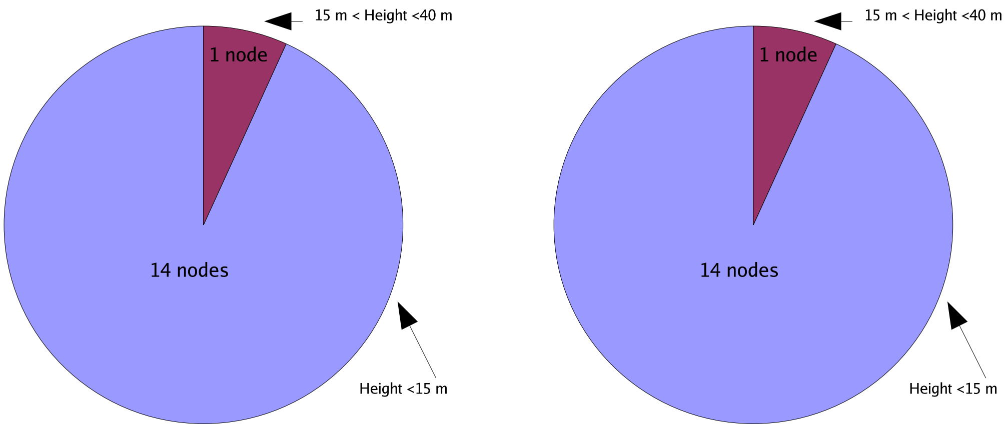


(c) Topology for randomly generated point set 3 (d) Topology for randomly generated point set 4



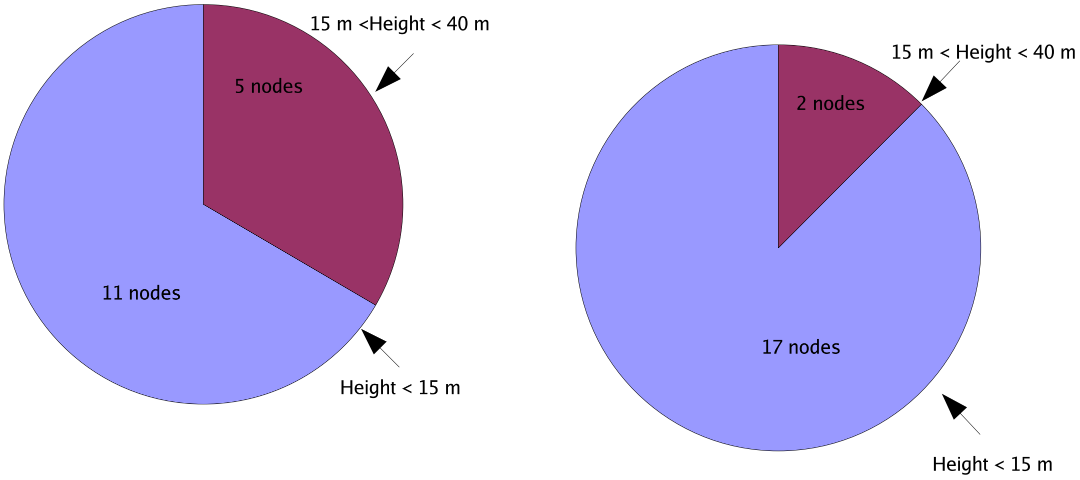
(e) Topology for randomly generated point set 5 (f) Topology for randomly generated point set 6

Figure 7.6: Topologies for randomly generated point sets



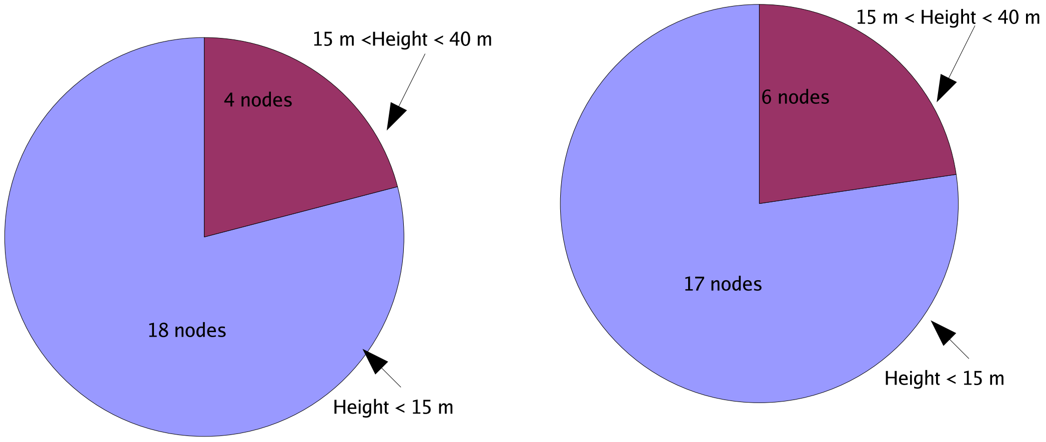
(a) Height distribution for randomly generated point set 1 (in meters)

(b) Height distribution for randomly generated point set 2 (in meters)



(c) Height distribution for randomly generated point set 3 (in meters)

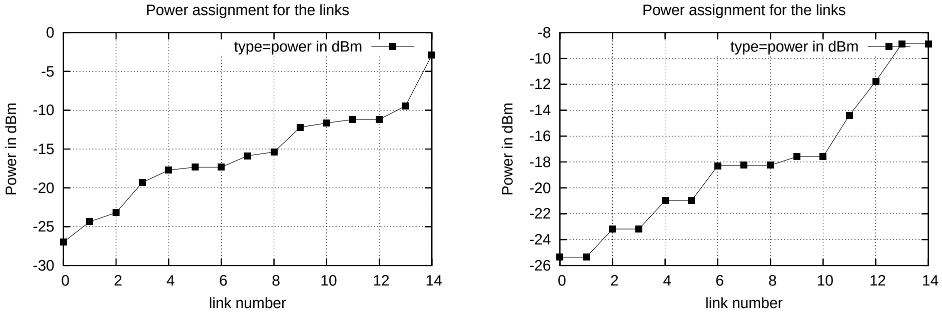
(d) Height distribution for randomly generated point set 4 (in meters)



(e) Height distribution for randomly generated point set 5 (in meters)

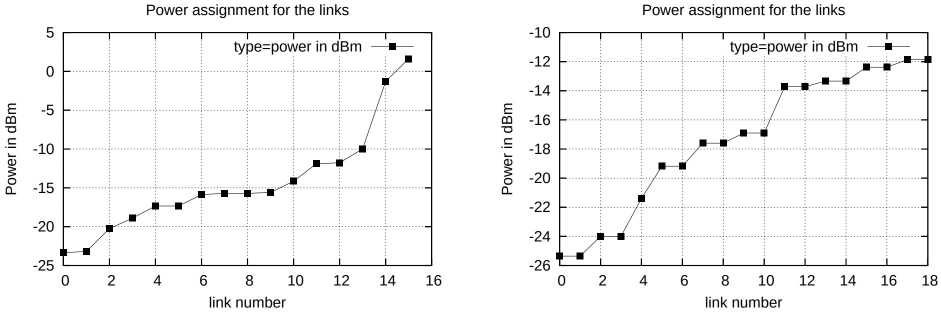
(f) Height distribution for randomly generated point set 6 (in meters)

Figure 7.7: Height distribution for randomly generated point sets



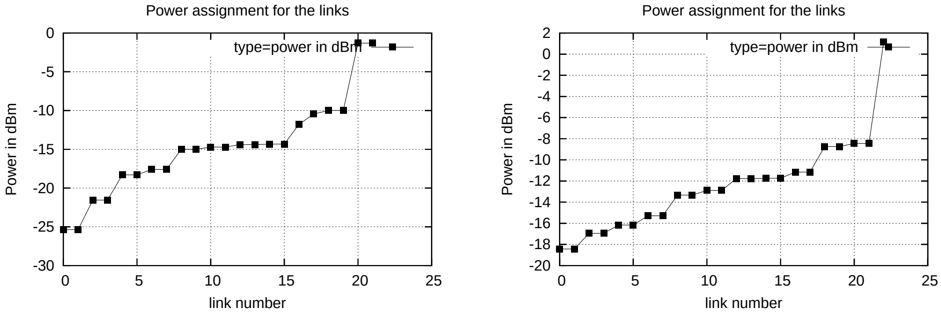
(a) Power distribution for randomly generated point set 1 (in dBm)

(b) Power distribution for randomly generated point set 2 (in dBm)



(c) Power distribution for randomly generated point set 3 (in dBm)

(d) Power distribution for randomly generated point set 4 (in dBm)



(e) Power distribution for randomly generated point set 5 (in dBm)

(f) Power distribution for randomly generated point set 6 (in dBm)

Figure 7.8: Power distribution for randomly generated point sets

Chapter 8

Related Work

Determining the topology in multi-hop wireless mesh network has been a topic of significant research. However, this research has been focused mainly on the problem of Topology Control, which is the determination of a connectivity pattern while minimizing the amount of energy transmitted by the radios placed at each node.

[20] did the seminal work for topology control problem, other works in this field include [23], [16] and [14]. [20] showed that construction of network topology which minimizes the maximum transmitter power allocated to any node is polynomial time solvable. [23] proposed a cone based distributed algorithm for topology control, [16] describes a distributed protocol which is designed for sectored antennas,while [14] uses the concept of range neighborhood graphs (RNG) for topology control.

However, our work differs significantly in its aim, as our focus is on minimization of cost of initial topology deployment while maintaining the received power values above a given threshold. Proper allocation of power values for the radios placed at various nodes given the nodes they are going to form links with has been studied in [17]. To our knowledge cost minimization of antenna towers while generating a topology is unique in our setting. No work concentrating on cost minimization while generating a topology has come to our knowledge.

While there has been a significant amount of work in topology planning for cellular networks, their focus is different from ours as their networks are designed to provide carpet connectivity, whereas our goal is to provide connectivity at each village location, at the minimum cost possible.

Chapter 9

Conclusions and Future Work

Conclusions

100% Rural connectivity is one of the stated goals of Government of India. Currently rural teledensity is only of the order of 3 per thousand square kilometers [18]. The problem is essentially that of providing last mile connectivity, as 85% of Indian talukas already has fibre optic connectivity. Using conventional means of wired or cellular technology for this purpose is not possible due to the high costs involved (India has around 6 lakh villages) . Long distance wireless networks using the popular 802.11 radio equipment along with corDECT [6] are a prime contender for being the enabling technology, because of their cheap cost and ease of setting up.

Research has been going on at several places [19], [15], [6] to fine-tune the 802.11 equipment for long distance functioning. But to the best of knowledge no one has worked on automating the topology planning procedure for this kind of networks. The importance of automating the planning process cannot be overemphasized as (a) there is a huge investment involved and (b) doing it on a trial and error basis with multiple iterations for even a district (let alone state or country) is simply impossible.

In this thesis we for the first time formulate the problem of topology construction in terms of the various factors involved. Also our work is significantly different from current planning mechanisms of campus wide WiFi networks and cellular as they require carpet connectivity, and focus there is on ensuring minimization of interference, while we need point to point connectivity with the focus being cost minimization. In current scenario the towers (the main cost source) are commissioned only by taking the elevation at the two end-points into consideration, which generally leads to the keeping of a huge safe margin and hence inflates cost. We have devised and field tested mechanisms which prove that there can be substantial reductions in this costs, by utilizing the freely available satellite data. Our results for this are currently being tested in Ashwini deployment, after going through successful field tests in our Kanpur testbed. We have been able to reduce the costs by factors of tens. We also consider the relationship between the Medium Access Layer (MAC) technology, the throughput requirements of the nodes into account while planning the networks. To the best of our knowledge this is also unique in our formulation as all previous work assumed unlimited supply of bandwidth. Long distance wireless communication depends on placement of proper antennas and transmitting right power through them. As to low power would lead to failure to establish the link, and too much power would result in high interference in the network and hence degradation in performance. We had undertaken a thorough field experimentation to understand this phenomenon on our testbed. Based on the results of these experiments we have formulated a more realistic power assignment equation set for the links in our network (solved as a Linear Programming problem). Also to the best of our knowledge we are the first to formulate the antenna planning problem.

We have been able to find topologies for sets of nodes of size 31 (so far). For a sub-set of Ashwini (31 out of 34 nodes) project, (a long distance rural network currently being deployed in the West Godavari district of Andhra Pradesh) we have been able to show an expected cost benefit of the order of 15 times over their current deployment plan, by improving the process of assigning tower heights. Also careful topology planning led to a further 22 % savings over their current deployment plan.Another fact to be borne in mind is the current deployment utilizes three independent WiFi channel while our solution only uses one channel.

Also, given the current state-of-art this translates into a much bigger cost savings in reality. Another fact to be borne in mind is the current deployment utilizes three independent WiFi channel while our solution only uses one channel.

Although we have focused on 802.11 based technologies the algorithms work for any networks utilizing Line of Sight (LOS) based communication like corDECT.

Future Work

Whereas, in this thesis we have designed algorithms for optimal topology construction, for the case when the intermediary hops are to placed in the villages only, the most immediate work in this regard, would be to extend the model to consider the entire convex hull of the region when finding feasible tower construction locations. Another drawback in the topologies generated by this model is that they are tree topologies, hence removal of one link can disconnect a node form the rest of the network. Making the network fault tolerant, even in the case of two or more node failures, is also important.

Pruning of the search trees, is currently done by generating the trees in Breadth First Search starting from the root node and checking the constraints for violation, making this procedure more intelligent using algorithmic techniques like hill climbing, constraint programming etc., is also a research focus.

Still another issue that remains to be explored properly is that of the height assignment to the the various nodes. Instead of restricting the topologies, to two hops, a model of piece-wise linear optimization should be tried.

Also as mentioned in Chapter 1, we capture the application performance required only in terms of the application throughput requirement. Time delay guarantees are also required proper functioning of real time applications. A formulation which takes this issue also into consideration is a task we leave open for future research.

Bibliography

[1] 47cfr15.247. `http://frwebgate.access.gpo.gov/cgi-bin/get-cfr.cgi?TITLE=47\&PART=15\&%SECTION=247\&YEAR=2001\&TYPE=TEXT`.

[2] An wireless tower company, llc - self supporting wireless communications towers. `http://www.anwireless.com`.

[3] DjurslandS.net: The story of a project to support the weak IT infrastructure in an low populated area of Denmark. `http://djurslands.net/biblioteket/international/djurslands_net_english_presentation.ppt`.

[4] Effective isotropically-radiated power-wikipedia, the free encyclopedia. `http://en.wikipedia.org/wiki/Effective_isotropically-radiated_power`.

[5] HyperGain 2.4 GHz Outdoor Directional 802.11b and 802.11g WiFi Antennas for Wireless LAN Systems. `http://www.hyperlinktech.com/web/antennas_2400_out_directional.php`.

[6] Media labs asia, iit madras. `http://www.tenet.res.in`.

[7] Project ashwini. `http://www.byrrajufoundation.org/ashwini_overview.htm`.

[8] Qsopt linear programming slover. `http://www2.isye.gatech.edu/~wcook/qsopt/`.

[9] Soekris Engineering. `http://www.soekris.com`.

[10] Srtm repository. `ftp://e0srp01u.ecs.nasa.gov`.

[11] Madhuresh Agrawal. Optimum tower height assignments for long distance 802.11 links. Technical report, cs397 Report,Indian Institue of Technology, Kanpur, 2006.

[12] Daniel Aguayo, John Bicket, Sanjit Biswas, Glenn Judd, and Robert Morris. Link-level Measurements from an 802.11b Mesh Network. In *SIGCOMM*, Aug 2004.

[13] Pravin Bhagwat, Bhaskaran Raman, and Dheeraj Sanghi. Turning 802.11 Inside-Out. In *HotNets-II*, Nov 2003.

[14] S.A. Borbash and E.H. Jennings. Distributed topology control algorithm for multihop wireless networks. In *IJCNN*, 2002.

[15] Eric Brewer, Michael Demmer, Bowei Du, Kevin Fall, Melissa Ho, Matthew Kam, Sergiu Nedevschi, Joyojeet Pal, Rabin Patra, and Sonesh Surana. The Case for Technology for Developing Regions. *IEEE Computer*, 38(6):25–38, June 2005.

[16] Zhuochuan Huang, Chien-Chung Shen, C. Srisathapornphat, and C.Jaikaeo. Topology control for ad hoc networks with directional antennas. In *ICCN*, 2002.

[17] Paul Ipe. Power allocation issues in a wireless mesh network. Technical report, BTech Project Report,Indian Institue of Technology, Kanpur, 2004.

[18] Ashok Jhunjhunwala and Sangamitra Ramchander. Role of wireless technologies in connecting rural india. In *Indian Journal of Radio & Space Physics*, pages 363–372, 2005.

[19] Bhaskaran Raman and Kameswari Chebrolu. Design and evaluation of a new mac protocol for long-distance 802.11 mesh networks. In *MOBICOM*, 2005.

[20] Ram Ramanathan and Regina Hain. Topology control of multihop wireless networks using transmit power adjustment. In *INFOCOM (2)*, pages 404–413, 2000.

[21] Sayandeep Sen, Sukant Kole, and Bhaskaran Raman. Rural telephony: A socio-economic case study. In *International Conference on Information and Communication Technologies and Development (ICTD-2006)*, may 2006.

[22] Edited by Bhaskaran Raman The DGP Media Labs Asia Team. Digital gangetic plains: 802.11-based low-cost networking for rural areas, 2001-2004: A report. `http://www.iitk.ac.in/mladgp`, 2004.

[23] R. Wattenhofer, L. Li, P. Bahl, and Y.M. Wang. Distributed topology control for power efficient operation in multihop wireless ad hoc networks. In *INFOCOM*, 2001.