Long Distance Wireless Mesh Network Planning: Problem Formulation and Solution

Sayandeep Sen
Dept. of CSE, IIT Kanpur
Kanpur INDIA 208016
sdsen@iitk.ac.in

Bhaskaran Raman
Dept. of CSE, IIT Kanpur
Kanpur INDIA 208016
braman@iitk.ac.in

ABSTRACT

Several research efforts as well as deployments have chosen IEEE 802.11 as a low-cost, long-distance access technology to bridge the digital divide. In this paper, we consider the important issue of planning such networks to the minimize system cost. This is a non-trivial task since it involves several sets of variables: the network topology, tower heights, antenna types to be used and their orientations, and radio transmit powers. The task is further complicated due to the presence of network performance constraints, and the inter-dependence among the variables. Our first contribution in this paper is the formulation of this problem in terms of the variables, constraints and the optimization criterion. Our second contribution is in identifying the dependencies among the variables and breaking-down the problem into four tractable sub-parts. In this process, we extensively use domain knowledge to strike a balance between tractability and practicality.

We have evaluated the proposed algorithms using random input sets as well as real-life instances with success. We have been able to show detailed planning of network topology, required tower heights, antenna types, and transmit powers for the Ashwini project, a long distance WiFi network under deployment in Andhra Pradesh, India, In this case, we are able to achieve within 2% additional cost of a lower bound estimate.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Wireless communication, Network Topology

General Terms

Algorithms

Keywords

802.11 mesh networks, Long-distance WiFi, Low cost networking, Rural Connectivity

1. INTRODUCTION

802.11-based long-distance networks have been proposed as a cost-effective option to provide Internet connectivity to rural areas in developing regions, to enable Information and Communication Technology (ICT) services [1, 2]. Fig. 1 depicts one such network: the current Ashwini project deployment [3] in Andhra Pradesh, India. Other such networks include the Digital Gangetic Plains

(DGP) [1], Aravind [4], Nepal Wireless [5], DjurslandS.Net [6], etc. In such networks, long-distance wireless links, up to several tens of kilometers are used to connect neighboring villages to a *landline* node *i.e.* a node with wired Internet connectivity (say, optical fiber). Each network node has a WiFi radio mounted atop an antenna tower. The towers are necessary to achieve Line-of-Sight (LOS) for the long distance links. Further details on how such long-distance operation may be achieved using 802.11 are in [1, 7].

The deployments thus far have been ad-hoc in nature and planning, if any, was mostly done manually. The lack of a systematic approach has several shortcomings. First, an ad-hoc approach or manual planning does not pay attention to the overall system cost. As indicated in [1, 2], cost minimization is the primary criterion for any technology deployment for rural areas in developing regions. In 802.11-based long-distance networks, the antenna tower costs form a significant component of the total cost. The antenna towers can be quite expensive: a 30m antenna tower can cost as much as U.S.\$1,000, while the WiFi radio itself may cost \$50 or less [7]. Network infrastructure planning for such long-distance networks thus becomes essential.
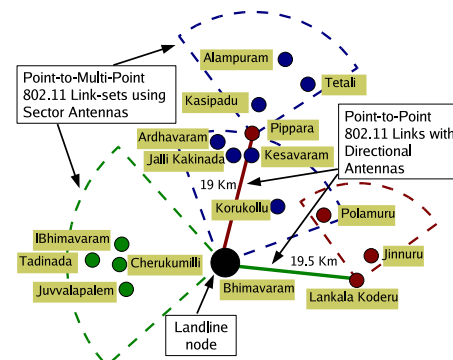

Figure 1: The Ashwini network, Andhra Pradesh, India

Second, lack of planning implies that network performance cannot be guaranteed. This is undesirable since a network operator may wish to provide, bandwidth guarantees to each village. Finally, manual planning cannot scale to large networks. For instance, even for the current 16-node deployment of Ashwini (Fig. 1), determining tower heights optimally and assigning antennas and transmit powers to guarantee network performance can be a daunting task without a methodical approach. During our studies at the Ashwini project site, we have found cases of inter-link interference and unpredictable throughput performance due to inadequate planning.

Thus, the *planning* of such long-distance wireless mesh networks *prior* to deployment is very essential. To the best of our knowledge,

we are the first to articulate and study this problem. Our first contribution in this paper is the formulation of the problem in terms of the variables, constraints and cost optimization objective.

We quantify the desired network performance in terms of three main sets of constraints (1) The application *throughput* constraint: this simply states that a particular download (or upload) capacity is required per village node. (2) *Power* constraints: these consist of an upper bound on the Effective Isotropic Radiated Power (EIRP) at each radio, and a lower bound on the received power at the receiving end of each link (receiver sensitivity requirements). (3) Inter-link *interference* constraints: at each radio, the received signal strength should exceed the total interference at that radio by a certain threshold.

We identify four main sets of variables which have to be determined in the course of network planning. The first two sets of variables: (1) the multi-hop network topology, and (2) antenna tower heights, primarily affect the system cost. The next two sets of variables are (3) the antenna types to be used atop the towers as well as their orientations, and (4) the transmit powers of the radios in the system. The latter two sets of variables primarily affect the overall network performance. We identify the various inter-dependencies in determining these variables for a deployment.

The problem of network planning can thus be stated as:

Given a set of villages to be provided with network connectivity from a given landline node, determine the minimum cost network topology, the tower heights, antenna types and orientations, and radio transmit powers, such that the three constraints: throughput, power, and interference, are satisfied.

A subset of the above problem is the aspect of determining the topology and tower height variables to minimize cost. This subset is akin to the well known hard problem of minimum energy broadcasting in wireless networks, where the complexity arises due to the "wireless multicast advantage" [8]. The presence of the three constraints as well as the inter-dependence between the four sets of variables further complicates our overall problem.

Our second contribution in this paper is the articulation of the various design considerations, and the approach we have taken to solve this problem. For this purpose we have drawn upon experiences and lessons from various deployments. We identify the dependencies and break down the problem into four tractable parts: (a) *Topology-Search (TS):* exploring the search space to find a topology, i.e., deciding which links to form, (b) *Height-Assignment (HA):* assigning optimum heights to tower locations once a topology has been formed, (c) *Antenna-Assignment (AA):* assigning appropriate antennas at each node and deciding their orientations, and (d) *Power-Assignment (PA):* assigning transmit powers to each of the radios in the system. We have applied extensive domain knowledge to strike a balance between problem tractability and solution practicality, while breaking down the problem into its subparts.

We use a branch-and-bound algorithm for the *TS* problem and for simplicity, focus only on the construction of a *tree* topology. The challenge here is in the application of various domain-knowledge based techniques to prune the search space. Next, we formulate *HA* and *PA* as Linear Programming (LP) problems. For *AA*, we propose a heuristic algorithm of polynomial time complexity.

Due to the inter-dependence between these problems, the stage at which we solve them is significant. We perform *AA* locally for each node, while we do *PA* and *HA* for an entire tree, whenever we have a spanning tree during the branch-and-bound search. Such an approach leads to an efficient algorithm which produces topologies satisfying the various constraints.

In our evaluation so far, we have been able to plan networks of sizes up to about 30 nodes. We have tested our algorithm with a real-life instance too: for a 31-node input set corresponding to the original goal of the Ashwini deployment (a super-set of the current 16-node deployment). The generated plan includes details of all four sets of variables, and is within 2% cost of a lower bound estimate. We believe that this is significant given the complexity of the problem. In the future however, it is not unimaginable to have deployments of O(100) nodes, to cover the set of villages in a 20-30km radius [9]. Further algorithmic and methodological enhancements are needed to achieve this. We view our work as the first step in this direction.

Although network topology planning has been studied in different contexts in prior work, several aspects make our setting unique. Unlike cellular network planning, we do not require carpet coverage, but only at the specific village locations. The consideration of cost minimization and the different sets of variables are also unique to our problem setting (unlike for e.g. in wired networks or in ad-hoc networks).

While our work is primarily for 802.11 mesh networks, in retrospect, our approach can also readily accommodate other wireless technologies in similar usage settings, such as 802.16 mesh networks. The rest of the paper is organized as follows. In Sec. 2, we articulate the problem in terms of its objective criterion, the variables, and the constraints. Subsequently, Sec. 3 describes the overall solution approach, and the problem breakdown. We also provide a detailed description of the formulation and solutions to the four sub-problems. We evaluate our algorithms and heuristics in Sec. 4. We present related work in Sec. 5 and conclude the paper in Sec. 6.

2. PROBLEM FORMULATION

In this section, we articulate the problem of network planning. In the course of our discussion, we present the optimization objective, the problem constraints, and the variables involved.

Our primary goal is to minimize the deployment cost. This is because, system cost is a central consideration in any technology deployment for rural regions [1, 2]. The long-distance links are typically formed using high gain antennas placed atop antenna towers. The main system components are the WiFi radios, the set of antennas, and the antenna towers. Among these, the dominant cost is that of the antenna towers [7]. Table 1 shows some tower costs from the current Ashwini & DGP deployments. (The cost for the towers taller than 15m also includes the cost of providing grounding for lightning protection: this cost component is about $100-150 per tower.) In comparison, the cost of a WiFi radio or an antenna is about one or two orders of magnitude lesser (~$50).

Tower ht. (m)	10	15	21	24	27	30	45
Cost ($)	100	150	800	950	1,100	1,850	5,000

Table 1: Antenna tower costs

Our optimization *objective* can thus be stated as: *minimize the cost of the towers* in the system.

In Table 1, the tower cost increases, piecewise-linearly, with tower height. The required tower height in turn depends on the link length since we need to achieve line-of-sight (LOS) clearance. Further, a single tall tower may be able to support links to several neighboring nodes. The tower heights are thus dependent on the network topology as well. Hence we have two main sets of *variables* here: the *mesh-network topology* and the *tower heights*.

While minimizing cost, we also wish to achieve a certain desired level of network performance. We formulate this as a *constraint* in the *TS* problem. In this paper, we quantify performance simply

by saying that each village node requires a certain download/upload capacity from/to the landline. For instance, in the Ashwini network, the original stated performance constraint was a download/upload capacity of $384Kbps$ per village node, to support high-quality interactive video. In the problem formulation, we denote this throughput requirement by R and term the constraint as the *throughput constraint*. We assume R to be the same for each village node, and in the download/upload directions.

The dependencies of the throughput constraint are as follows. Each link has a certain capacity, which is determined by the PHY layer. For instance, operating at 11Mbps using 802.11b allows a maximum application-level throughput of about 7Mbps [10]. The operation of a link at a particular PHY rate (e.g. 11Mbps) depends on (a) the received signal strength (RSSI) being above a certain threshold, termed P_{min}, and (b) the signal-to-interference ratio (SIR) being above a threshold, termed SIR_{reqd} (in the network, each link acts as interference to the others). For typical commercial 802.11b equipment, $P_{min} = -85dBm$ and $SIR_{reqd} = 10dB$ for 11Mbps operation [10]. In our problem, we formulate the above two requirements as the *power constraint* and *interference constraint* respectively. It is worth noting that for a given transmit power, the RSSI and SIR are fairly stable for the long-distance links [10].

The RSSI and SIR depend on several variables: the link length, transmit power, and the antenna types in use. Thus we now have two additional sets of *variables* the *antenna types and their orientations*, and the radio *transmit powers*.

Apart from the power and interference constraints, the throughput constraint also has a dependence on the antenna type, as we explain now. We distinguish between *point-to-point (p2p)* links and *point-to-multipoint (p2mp)* link-sets (see Fig. 1). In a p2p link, we have a single radio at either end. It is typically formed using antennas of small beam-width at either end of a link (e.g. parabolic-grid antennas with 8° beam-width). A p2mp link-set consists of multiple logical links. Each of the logical links is between a central radio and a radio at the far-end. A p2mp link-set is usually formed by using a sector antenna at the central location. Sector antennas with beam-widths of 30°, 60°, 90°, and 120° are available commercially [11]. In Fig. 1, p2mp link-sets are shown as sectors.

The link capacity in a p2mp link-set is shared between its constituent links. And whether we have a p2p link or a p2mp link-set is clearly dependent on the antenna type. Further, in a p2mp link-set, the number of nodes sharing the capacity is dependent on the antenna orientation. Thus the antenna types and orientations also affect the throughput constraint.
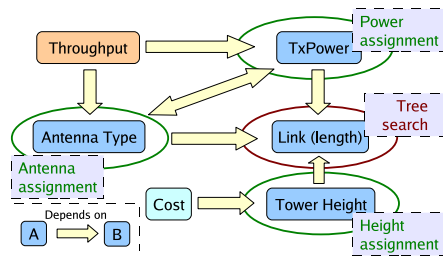
Figure 2: Dependencies in network planning

To summarize the above discussion, we state:

Objective: tower cost minimization
Constraints: throughput, power, interference
Variables to determine: network topology, tower heights, antenna

types & orientations, transmit powers
Fig. 2 summarizes the dependencies among these.

There are two other factors, not mentioned above, on which the throughput depends. The first is the MAC protocol in use. In this work, we assume the use of the 2P MAC protocol proposed in [12]. 2P is much more efficient than CSMA/CA for the long-distance links, since it allows the simultaneous operation of the multiple radios at a node. Further, it uses a single channel throughout the wireless *backbone* network. This allows us to use the remaining two (of the three total) independent channels of 802.11b for local usage within each village node. 2P effectively reduces each half-duplex wireless link to look like a wired link of half the capacity in each direction (see Fig. 4 in [12]). We take this into account while checking for the throughput constraint. Changing the multi-hop protocol to CSMA/CA or TDMA would result in different formulation of the throughput constraint, as has been shown in [13].

While the work in [12] considers a network of only p2p links, we consider p2mp link-sets too. A minor modification is required in 2P to accommodate p2mp link-sets. The original 2P protocol has two phases (hence its name): SynTx and SynRx. These two phases correspond to synchronous transmission on all the links from a node, and synchronous reception on the same links. For p2mp link-sets, we need to note that the links in the set share the radio at the central node. Thus only one of them will remain active during SynTx or SynRx.

The throughput at a node depends on the routing too. But as mentioned earlier, for simplicity, we only focus on tree topologies. This only provides connectivity and does not address any fault-tolerance. For a given tree topology, the routing is fixed. Hence the dependence of routing on throughput translates to dependence on the network topology, which we have already considered as a variable in the problem statement.

3. SOLUTION TO TOPOLOGY PLANNING

In this section, we present our solution to the above problem statement. We first present our overall strategy and the breakdown of the problem in Sec. 3.1. Subsequently, in Sec. 3.2-Sec. 3.5, we present formulations of the individual sub-problems and their solutions. Our main contribution here is in terms of the partitioning of the problem as well as the application of various domain knowledge based simplifications.

3.1 Overall Solution Strategy

As depicted in Fig. 2, there is significant inter-dependence in determining the four sets of variables. Given this, even the formulation of the problem in its entirety is complex. We initially tried this approach, but soon realized that the problem was better addressed by breaking it up into smaller parts.

We do not seek to design a single algorithm which determines all the variables and satisfies all the constraints at one go. As mentioned earlier, even just the sub-problem of height assignment seems akin to the well known difficult problem of minimum energy broadcasting in wireless networks [8].

The four sets of variables suggests a natural division of the problem along these lines. However, due to the inter-dependence, it is still challenging to determine the *order* in which a set of variables are to be solved for. That is, the order in which the dependencies are resolved is non-trivial.

We begin by making several observations. (O1) We first observe from Fig. 2 that all three of tower heights, antenna types, and transmit powers are dependent on the network topology (set of links). (O2) Next, we observe that the tower heights are independent of

the antenna types or the transmit powers. This can be easily understood: the tower heights merely ensure that there is LOS. (O3) The antenna types and the transmit powers are directly dependent on one another.

The first observation (O1) suggests that we should first arrive at a network topology, and then determine the (dependent) tower heights, antenna types, and transmit powers. We term this subproblem of finding a tree topology as *Topology-Search (TS)*. We use a branch-and-bound based tree-enumeration approach for this. In this approach, it turns out that we can easily check the *throughput constraint* at each stage of the tree search. Whenever a spanning tree is enumerated, we solve the *Height Assignment (HA)*, *Antenna Assignment (AA)*, and *Power Assignment (PA)* problems, as defined below.

The *Height Assignment (HA)* problem is defined as: for a given topology, determine the tower heights at each of the nodes such that LOS is achieved, and such that the sum total of the tower costs is minimized. Observation (O2) implies that the order of solving the *HA* problem versus *AA* or *PA* are mutually independent. In the following paragraphs, we shall show an LP formulation of *HA*.

Observation (O3) raises significant questions. Should antenna types to be placed at a node and orientations be decided first, or should transmit powers be determined first? In either case, how can we ensure that the power and interference constraints are addressed? A result from prior work helps us resolve this problem. In [12], the power assignment problem is defined as: for a given topology and set of antennas, what should the transmit powers be such that the power and interference constraints are satisfied? The problem is formulated as an LP.

Given that we have a *Power Assignment (PA)* formulation as above, we define the *Antenna Assignment (AA)* problem simply as: given a node and its set of children, what should be the types of antennas used for link formation, and in what directions should they be oriented? Note that *AA* is defined local to each node, and there is no global view across the network. In this approach, we solve various local instances of *AA* (one instance for each network node). We then finally solve the *PA* problem, which has a global view of the network, and also incorporates the power and interference constraints.

Thus, in each stage of the tree-search, we get a spanning tree topology which satisfies the three constraints. For each such topology, we perform *HA*, and keep track of the topology which has the minimum cost. We now elaborate on the four sub-problems.

3.2   Topology Search (TS)

The *topology search (TS)* algorithm is the all-encompassing component, which invokes the other three sub-problems at the appropriate junctures. We view the tree topology as being rooted at the landline, with the natural notions of *parent* and *child* relations. We use a Breadth First Search (BFS) based spanning tree enumeration approach, with several pruning mechanisms. *TS* also incorporates the *throughput constraint* by ensuring that the trees it produces satisfy that constraint. We explain this below.

Checking the throughput constraint

Recall from Sec. 2 that each link has a particular throughput capacity. This is about 7Mbps at the application level, while operating at 11Mbps using 802.11b [10]. Note that in our design, the power assignment step ensures that each link can operate at this capacity. Further, the 2P MAC protocol is going to ensure that this capacity is divided equally in the download/upload directions. That is, we have $7/2 = 3.5Mbps$ in each direction. Denote this as B_{max}.

How B_{max} is shared across the various nodes depends on the topology. In the BFS-based enumeration, at each stage, we have a

connected sub-tree rooted at the landline. In this setting, we wish to check if a given sub-tree satisfies the throughput constraint. To achieve this, we make the following design decision. We stipulate that all the links with the landline are p2p links. This ensures that we have a download/upload capacity of B_{max} on each of these links with the landline. This is illustrated in Fig. 3.
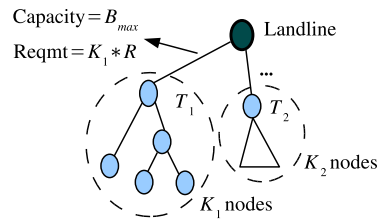

Capacity $= B_{max}$
Reqmt $= K_1 * R$

Figure 3: Throughput constraint: $K_i < \frac{B_{max}}{R}$

Clearly, the links with the landline are going to be the throughput bottlenecks of the network, as the entire traffic of their corresponding sub-trees has to be routed by them. We we denote the sub-trees hanging from these links as T_i, and the number of nodes in T_i as K_i, as depicted in Fig. 3. Assuming a fair allocation of the bottleneck capacity among the nodes in T_i, we have the throughput constraints represented as $K_i < \frac{B_{max}}{R}$. Adherence to this criterion can be easily checked for incrementally during the tree-search.

Search pruning strategies

An important aspect of *TS* is the set of domain-knowledge based strategies that we use to cut the time taken by the search procedure. While the search is still combinatorial, the pruning strategies are very effective.

Eliminating long links: The first heuristic we apply is to ignore "long" links. The intuition for this comes from several factors. First, long links require more transmit power, thus increasing the chances of interference with nearby links. More importantly, longer links require taller towers. The significance of this can be understood from Table 1. We can see that the cost of sub-15m towers is much lower than that of taller towers. We term such short towers as *masts*. Such masts are inexpensive since they can simply be constructed by using a couple of water pipes.

From experience in CorDECT deployment efforts, it is known that typically links up to 10-15km long can be formed with only a mast at one of the ends [9]. The threshold link length we use in our algorithm thus corresponds to this. In a given input data set spanning a diameter of 40-60km, it is easy to see how the elimination of long links will significantly cut down the search space.

Tree depth restriction: Another effective pruning strategy is to restrict the depth of the tree topology to be at most two hops from the landline node. This would also significantly reduce the search space. The depth restriction is justified since in most practical cases, we would be looking to plan a network of radius about 20-30km from the landline [9].

Dynamic Cost Bounding (DCB): While the previous two pruning strategies are heuristics, DCB is a pruning strategy based on lower bounding the cost of the sub-tree during the BFS-based *TS* process. Now, such lower bounding of cost is tricky for the following reason. Cost is determined by the tower heights, but during the *TS*, we have not yet assigned any tower heights. And, invoking *HA* for each sub-tree during the tree search process is not an option since it would prove to be time consuming (*HA* itself is an LP, which may not be solvable in polynomial time).

Tower height lower bounding forms the basis for cost lower bound-

ing for a sub-tree. It builds on the following insights. (I1) The landline is typically in a town or a city, with reasonably tall buildings around. (I2) About 15km long links can be formed between a 40m tower at one end of the link and a sub-15m mast at the other [9]. (I3) Due to land undulations (even in the plains) and the presence of 10-12m trees, it is unlikely that we can achieve LOS even over distances as short as 1km using just sub-15m masts at either end.

(I1) above implies that we require a tall tower at the landline node. In practice this turns out to be about 50-60m high. (I2) implies that for a link, it is always more economical to have one end as a sub-15m mast. Now, combining these with (I3) above; in two-hop configurations, we can have only one of two possibilities: (a) a mast at the second-hop (leaf-node), or (b) a mast at the first hop (non-leaf). This is shown in Fig. 4.
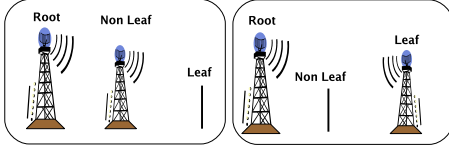

Figure 4: The two choices for mast location

Among the two choices, locating the mast at the leaves is clearly more economical. This is so; because, for every non-leaf one-hop from the landline, there is at least one (and possibly many more) leaf node(s). This means that choice (a) above is always less expensive than (or as expensive as) choice (b).

Now we can use the above discussion for calculating a lower bound on the tower heights (cost) of a sub-tree as follows. We assume a tall tower (50 to 60 m) having a fixed height at the landline node. We have already argued above that masts (negligible cost) are sufficient for leaf nodes. And for the non-leaf nodes which are one-hop from the landline, the minimum required height (cost) can be easily determined by its current child set.

In Sec. 4 we shall show that this DCB technique is indeed very useful in pruning the search space. We shall now turn to the *HA* sub-problem.

3.3   Height Assignment (HA)

In *HA*, we seek to determine the heights of the towers at the various nodes, such that the overall cost is minimized. In doing this, we are given the network topology.

For long-distance WiFi links, line-of-sight (LOS) is required. Prior experience in long-distance WiFi [7] as well as CorDECT links [9] has indicated that the obstructions (such as trees, buildings etc.) play a significant role in determining the height required for achieving LOS.
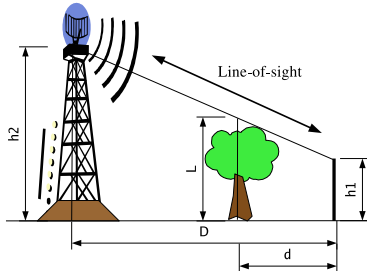

Figure 5: The LOS constraint on the tower height

Fig. 5 shows how the tower heights at either end of a link are

related to the height and location an obstruction in-between. We denote by D the link length, by L the obstruction height, and by d the distance of the obstruction from one end of the link. The heights of the towers are denoted h_1 and h_2 respectively. It is easy to see that for LOS clearance, we can write an inequality using basic geometry:

$$h_1 * (D - d) + h_2 * d \geq L * D \qquad (1)$$

Generally, in rural areas (which is where such long-distance networks are meant to be primarily used), we typically do not have many tall buildings, and the only obstructions are the trees in-between. It is possible to have a upper-bound on the height of trees in a region. This gives us the value of L. Also for a given link, we can estimate d based on a site survey. Thus in Eq. 1 only h_1 and h_2 are the variables. It should be noted that presence of obstruction sources other than trees does not change the formulation, it only necessitates determination of L from site survey.

Although Fig. 5 shows the two tower locations at the same level, this need not be the case. The same approach can readily accommodate tower locations at different Mean Sea Levels (MSL). The MSL data can also be gathered from a site survey (as is the geographical location of the towers). Although our notation does not indicate it (for simplicity), the parameters L and d can be different for each link.

We can write equations similar to Eq. 1 for each of the links in the given topology. This gives a set of linear inequalities. In addition, we also impose an upper bound H_{upper} (60m in our evaluations) on the tower heights: $h_i \leq H_{upper}$. This is done to ensure that we do not end up with impractical, arbitrarily high towers. The minimization criterion for *HA* is:

$$minimize(\sum C(h_i)) \qquad (2)$$

where $C(h)$ represents the tower cost as a function of its height. As indicated in Table 1, the cost function $C(h)$ is super-linear. Hence the above formulation is linear but for the optimization criterion.
Piece-wise linearity

The cost function $C(h)$ can be approximated by a piece-wise linear function (as can be any function). This is further facilitated by the notion of *families* of tower structure designs. A particular tower design can accommodate towers within a particular range of heights. The cost typically takes quantum jumps only when the tower design itself changes. If $C(h)$ can be approximated as a piece-wise linear function, we can use the standard technique of separable programming to solve the formulation much like an LP [14].
The special case of trees with depth at most two

The above formulation of the *HA* problem makes no assumptions about the nature of the topology. However, if we take into account the fact that we restrict the depth of the tree to be two at the most, a simplification is possible.

We first observe in Table 1 that the main jump in cost comes as we go from masts to taller towers. While the tower cost still remains a super linear function of tower heights beyond 15m , the cost increments across different tower designs is not that substantial. So, in engineering terms, it would not be too bad of an approximation if we assume it to be linear in the range beyond 15m (so long as we do not exceed much higher than 50m). We thus have two linear sections in the cost function: one for masts, and another for taller towers.

Recall our argument from Sec. 3.2 that (a) the landline node will

have a tall tower (50-60m in height), (b) the leaf nodes will have masts, and (c) the non-leaf nodes one hop from the landline will have tall towers. Given these, for each node, we can select the (linear) region of the cost function which is applicable for that node (mast or taller tower). Thus the minimization objective in Eq. 2 now becomes linear too and the entire *HA* becomes an LP. The tree depth restriction, thus not only helps in tree search pruning, but also in meaningfully approximating the *HA* to an LP. We now discuss our solution to the third sub-problem: the *AA* problem.

3.4 Antenna Assignment (AA)

The *AA* problem is addressed individually for each node, in a given tree topology. The problem involves deciding the set of antennas to be used at the node as well as their orientations. We assume that we are given a set of antenna types to choose from. We assume that we know the cost associated with each antenna type. For instance, we could be given antennas of beam-widths 8°, 22°, 30°, and 60°. Before we present the details of the problem and the solution, it is worth noting two points.

First, as mentioned already, we assume that all the links from the landline node to its children are p2p links. So, we do not need to perform *AA* for the for the landline node. A more significant aspect to note, is the following. While assigning antennas to a node, we only consider its child set. Specifically, we do not consider the link to its parent. We simply use a p2p link, with a low-beam-width antenna directed towards the parent. The reason for this is that we do not want a p2mp link-set to the parent: the throughput of the link to the parent is the bottleneck, and we do not want it to be shared with any other (child) node.

We have two loose goals in *AA*: cost reduction, and interference reduction. We do not have these as strict optimization goals. One reason for this is that *AA* resembles the minimum set cover problem: where we need to cover a set of children using a set of antenna types. The minimum set cover problem is a well known NP-complete problem. Hence instead of formulating and seeking an optimal solution, we simply use a heuristic approach for *AA*.

Another reason for not seeking even a formulation of strict optimality is that it is not important. First, cost optimization is not significant here since the cost of the antennas are small anyway ($50 to $150), compared to that of the towers. And interference consideration too need not be strict, since in our overall approach, we push the careful consideration of the interference constraint (and power constraint too) to the *PA* sub-problem.

Our heuristic algorithm for *AA* first pre-sorts the child set radially. We then use the following recursive procedure.

1. If the given child set S can be covered within the beam-width of a particular antenna type, return that antenna type.

2. Else, identify the radially adjacent pair of children which have the maximum angular separation. Split S at this point into S_1 and S_2.

3. Call the above procedure recursively for S_1 and S_2 and merge the antenna sets returned by the recursive calls.
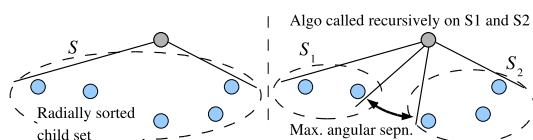

Figure 6: Recursive antenna assignment: an illustration

It is easy to see that the worst-case complexity of the above algorithm is $O(|S|^2)$. An illustration of the recursion is given in Fig. 6. The reasoning behind step-1 above is the cost reduction objective. And the reasoning behind finding out sets with the maximum angular separation in step-2 is to achieve interference reduction. The farther away the main lobes of two antennas, lesser is the interference effect of their side lobes on one another.

3.5 Power Assignment (PA)

The formulation of the *PA* problem is akin to that in our prior work [12]. In [12], we have an LP formulation for *PA*, incorporating the power and interference constraints. One significant additional complication which arises in this work is the consideration of p2mp link-sets. In this case, not all the links in a p2mp link-set are operational at the same time since they share the same radio. Hence while considering the interference of a p2mp link-set on other radios, each link of the link-set must be considered separately. The formulation however still remains a LP.

In our setting, we also incorporate another insight in *PA*. Recall that no two nodes with masts (i.e. leaf nodes in our tree topology) will interfere with one another (I3 from Sec. 3.2). We thus leave out these cases in the LP formulation of *PA*.

4. EVALUATION

In this section, we present an evaluation of the main aspects of our approach and our algorithms. We seek to answer the question of whether our approach is practical. Also, the *AA* step in our approach is based on heuristics. This opens up the question of whether our solution is indeed optimal.

We have implemented our algorithm in C, and have used the QSOPT [15] library for the solving LP problems. All of the simulations were run on a 3GHz desktop.

Unless otherwise mentioned, we use the following parameter settings. We use 15km to be the link length threshold (as explained in Sec. 3.2). The obstruction height L is assumed to be 18m, which represents a situation where we may have tall coconut or palm trees in the vicinity. This is, however, a worst-case situation as apart from the coastal regions of India, we mostly encounter shorter trees. We set d to be 1km: this means that we choose the tower location within each village such that we do not have any tall trees within 1km in the exact direction of the intended link. This is practical given that most villages are surrounded by large expanses of farming fields anyway. We assume a variant of Free Space Pathloss model for radio propagation, which closely models the behavior of long distance Wi-Fi links [16]. We use $SIR_{reqd} = 15dB$, which gives a headroom of 5dB above the theoretical SIR for 802.11b 11Mbps operation. We use $R = 384Kbps$, which corresponds to the requirement for high-quality video as stated in the requirement specifications of Ashwini project. We use $H_{upper} = 60m$. The antenna types, costs, and side-lobe patterns resemble those of actual off-the-shelf antennas [11].

For the randomly generated input sets, we use 15-node cases, these sets have *density* of the nodes resembling that of the real deployment scenario (Ashwini). For the cost function $C(h)$, we use the two-line approximation mentioned in Sec. 3.3. The actual function used is approximated from Table 1, and is shown in Fig. 7.

We present several aspects of our evaluation below. We first evaluate the algorithm on some real input sets (Sec. 4.1). Next we present the usefulness of the pruning strategies, specifically that of DCB (Sec. 4.2). Finally, we present the effect of varying the input parameters on the performance of algorithm, using random input sets (Sec. 4.3).
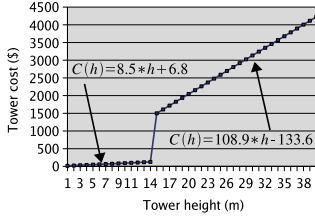
Figure 7: Tower cost versus tower height

4.1 Evaluation on Real Deployment Cases

The original goal of the Ashwini project was to deploy a long-distance WiFi network to provide connectivity to 33 villages surrounding the Bhimavaram town in, Andhra Pradesh, India (33+1=34 including the landline). The current deployment, shown in Fig. 1, is consists of 15 of the proposed 33 nodes. This deployment was planned manually, and we are already experiencing several performance and interference issues in the network.

Fig. 8 shows the topology generated by our algorithm for the input set. However, we should note that only the X and Y coordinates were taken as input, not the MSLs. Also, we assumed $L = 18m$ and $d = 2km$ to be uniform for all links. Three of the nodes 30, 31, and 32 could not be connected due to our restriction of link length to 15km, and the 2-hop restriction. We were able to successfully connect the remaining 31 nodes.
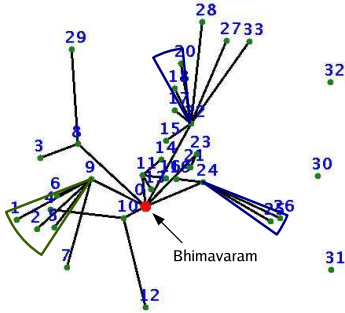

Figure 8: Topology generated for Ashwini

In the generated topology, node-0 is the landline node (Bhimavaram). There are 7 other non-leaf nodes, with tower heights in the range of 20-36m. The AA step assigns 8° directional antennas to all but three of the nodes: two of 22° beam-width and one of 30°, as shown in Fig. 8. The topology satisfies all the three constraints: throughput, power, and interference. Its overall cost was about \$55K, and the algorithm run-time was about 10 hours.

Cost lower bound: We now compare the above topology cost with a lower bound, computed as follows. We construct a lower bound tree topology by assigning as many children as possible to the root, so long as this is allowed by our link length constraint. After this, there may be some nodes remaining which are unreachable from the root directly. Let these nodes be $m_i, \{i = 1, 2, ...k\}$. Each such node has a set of possible parents P_i. One choice of parent for each m_i produces a tree topology. There are a total of $\Pi_{i=1}^{k}|P_i|$ such trees. We enumerate all these trees, and for each tree, we find the optimal height assignment using the LP formulation given in Sec. 3.3. Call the set of these trees to be F, and let the topology which has the minimum cost be $\hat{T}$ and its cost $\hat{C}$. In this process, we completely ignore PA and AA.

Let T_A be the tree output by our overall algorithm, and its cost

be C_A. We claim that $\hat{C}$ represents a lower bound on C_A. First observe that both $\hat{T}$ and T_A use the same height for the root node. Now, denote the set of children of the root in $\hat{T}$ as S. Let T_{A1} (with cost C_{A1}) be the tree obtained from T_A by assigning the root as parent, for all nodes in S. Clearly, $C_{A1} \le C_A$ (since the root's height is fixed). And $T_{A1} \in F$, and hence $\hat{C} \le C_{A1}$.

Using the above method we found the lower bound for our 31-node Ashwini input, to be about \$54K. This is only slightly less than the cost of the topology generated by our algorithm. Further, we observed that the lower bound topology does not have a feasible power assignment, while the output of our algorithm clearly does. We depict the topology with lowest cost in Fig 9. By looking at the links in the region marked by a circle in the figure and comparing the same with Fig 8, the utility of our algorithm becomes evident. As can be seen, in our topology, node 11 is selected as a non-leaf node (as a parent of nodes 13, 14 and 16), whereas in the topology with least cost nodes 11,13,14 and 16 are all children of root, which makes AA and PA difficult, due to low angular separation amongst them. In our experience, we have found that in general
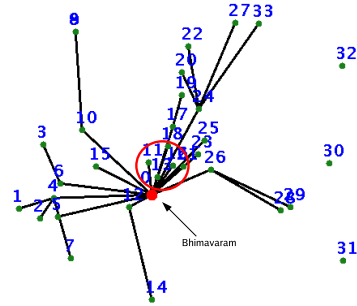

Figure 9: Ashwini topology with least cost (infeasible)

the power assignment is not easy to achieve using naive schemes. To further stress this point, we note the following. In our original (manual) plan for the Ashwini network, we had sought to use all the three non-overlapping channels available in 802.11b. Even in such a scenario, using naive power values of 20dBm (the default setting in many outdoor WiFi products) creates inter-link interference in several cases.

There are a number of links in the current deployment which cannot function simultaneously due to violation of the above criteria. A few of these are depicted in Fig 10. In the current setup, there are sectoral antennas of 90 degree beamwidth at nodes 100 and 123, a sector antenna of 120 degree is present at node 122 for the $p2mp$ links, while all other $p2p$ links have directional antennae. As can be seen from the figure, links from node 100 to nodes 114 or 120 or 124 cannot function when the link from node 125 to node 123 is active (SIR at 114 and 120 of 9 db and of 10 db at 124). Similarly, the link from node 100 to 119 cannot function when the node 126 is transmitting traffic to node 122, (SIR at 119 of 8 db). This example further underlines that our algorithm is essential to ensure an operational network.

Topology for East Godavari: We present the proposed solution of our algorithm for another real life input. In an extension of the Ashwini project, a deployment for 18 nodes is being planned in the East Godavari district. We have run our algorithm successfully for this input set too. The generated topology is shown in Fig. 11. The cost in this case was about \$40,000, and the algorithm run-time was less than 5 seconds.
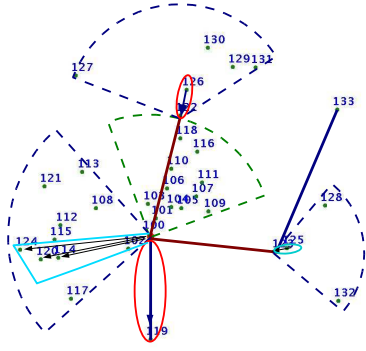
Figure 10: Links incapable of simultaneous operation in proposed Ashwini deployment
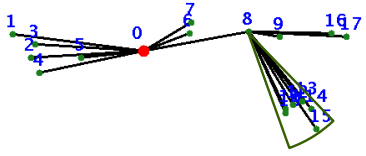

Figure 11: Topology generated East Godavari

4.2 Effectiveness of pruning

To evaluate the effectiveness of our pruning strategies, Fig. 12 shows the run-time of the algorithm for random input sets of different node counts. This is compared with the case where we do not have any pruning based on DCB or the elimination of long edges. We can clearly conclude that pruning based on DCB and elimination of long edges makes the algorithm much more scalable.
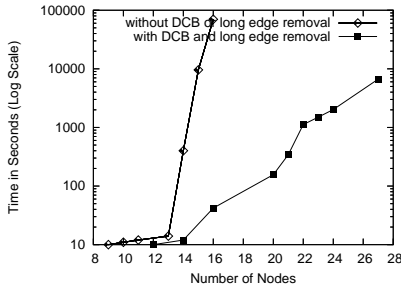

Figure 12: Run-time with and without pruning

Table 2 specifically shows the usefulness of the DCB pruning strategy for 5 different random input sets of 10 nodes each. We can see that DCB is very effective in pruning sub-trees early on to cut the run-time significantly.

4.3 Algorithm Characteristics

While carrying out network planning, the run-time of an algorithm is not a very serious concern, however it matters in certain contexts. For instance, while a run-time difference between a few seconds and a few hours hardly matters, a difference between a few hours and a few weeks certainly does. To this end, it is useful to know the performance of the algorithm as the various input parameters are varied. We specifically look at the variation of d, L, and SIR_{reqd}. The different parts of Table 3 shows the variation of the run-time for a specific random input set of 15 nodes, with increas-

S.No.	% subtrees rejected by DCB	Runtime with DCB (sec)	Runtime without DCB (sec)
1	97	0.029	382
2	74	0.025	274
3	53	0.009	217
4	94	0.010	130
5	88	0.015	199

Table 2: Effectiveness of DCB

ing d, L, and SIR_{reqd}. While the data corresponds to a specific topology, we observed similar trends in other random input sets too.

First, we observe in Table 3 that the run-time is high for small values of d, and also for large d. The reason for this is that when d is small, too many of the generated tree topologies are adjudged infeasible in the *HA* step, due to tower height requirement beyond the imposed threshold H_{upper}. If too many topologies are infeasible, then the running cost bound gets updated only infrequently. This in turn means that DCB is not very effective. On the other hand, when d is too high, then too many tree topologies become low-cost, and then as well DCB filtering becomes less effective.

d (km)	Runtime (sec)	L (m)	Runtime (sec)	SIR_reqd (dB)	Runtime (sec)
0.75	280	16	0	11	1
1	155	18	0	13	38
1.5	97	20	0	15	38
2	95	22	14	17	54
2.5	95	24	101	19	92
3	160			21	2160

Table 3: Run-time variation with d, L, SIR_{reqd}

Next, we see in Table 3 that the run-time increases sharply with the obstruction height L. The reason for this is the same as the higher run-time for low d, i.e. we have several topologies rejected in the *HA* step. Similarly, the run-time increases sharply for large SIR_{reqd} too. The reason for this is that several topologies are now being rejected in the *PA* step.

We also wish to point out that the overall system cost increases with increasing L or with decreasing d. This is as expected. On the contrary, it is noteworthy that the system cost does not increase significantly if the SIR_{reqd} is incremented. This is because the cost is decided by the tower heights, which is independent of the interference constraint (see Fig. 2).

Performance of Antenna Assignment: We now show the usefulness of having p2mp link-sets, and of our *AA* approach in particular. Fig. 13 shows the run-time of the algorithm for two cases: (a) where we simply assign 8° beam-width (directional) antennas for all links (i.e., all links are p2p links), and (b) where we consider the a choice of antennas with 8°, 22°, and 30° beam-widths to select from and use our *AA* algorithm.

We use SIR_{reqd} as the x-axis in Fig. 13 since it is a measure of how stringent the interference constraint is, and how effective our AA approach is in alleviating that. Although Fig. 13 is for a specific random input of 15 points, this is representative: we have run for other random inputs too with the similar results. We find in Fig. 13 that the run-time is lesser with the use of p2mp link-sets and with our antenna assignment approach. More important than the reduced run-time itself is the reason for this: the *PA* step becomes less stringent with the use of p2mp link-sets (recall our discussion in Sec. 3.5). This means that we are more likely to find feasible topologies with our *AA* approach than without it.

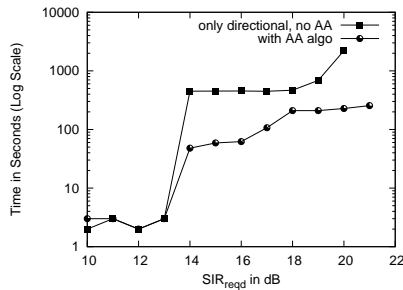It is also worthy noting that the inclusion of antennas of larger

Figure 13: Effectiveness of *AA*

beam-widths does not seem to help. For instance, we tried including antennas with 60° beam-width in our set of antenna types. With this, the *AA* step ends up choosing to use a 60° antenna at one or more nodes. However, the side-lobe pattern of these antennas are not good enough, and this makes the LP in *PA* infeasible.

5. RELATED WORK

Determining the topology in multi-hop wireless mesh network has been a topic of significant research [17, 18, 19, 20]. The work in [17] showed that the problem of constructing network topology which minimizes the maximum transmitter power allocated to any node is polynomial time solvable. In [18] the authors propose a cone based distributed algorithm for topology control, while [19] describes a distributed protocol which is designed for sector antennas. The work in [20] uses the concept of range neighborhood graphs for topology control.

However, this research has been focused mainly on the problem of topology control, where the aim is to obtain a connectivity pattern which minimizes the energy consumption by the radios in the network. 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. Power consumption of the radio itself is not a major concern in our setting [21].

There has been a significant amount of work in topology planning for cellular networks too. However, 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. To our knowledge cost minimization of antenna towers while generating a topology is unique in our setting.

The work in [22] considers an cost analysis of various technologies for rural connectivity. However, they do not consider antenna tower cost, and in this sense our work is complementary to [22].

6. CONCLUSION AND FUTURE WORK

Usage of long distance wireless networks is inevitable to connect the rural areas of the world and bridge the digital divide. Although several deployments have already sprung up, the issue of network planning has not been considered thus far. In this paper, we first formulate this problem in terms of the relevant objective function, constraints, and variables. The problem is unique and quite different from other network planning or topology construction problems considered thus far, the primary difference being our consideration of system cost.

We introduce a variety of design considerations, with experiences and lessons drawn from different deployment efforts. Based on these, we present a solution approach which breaks down the

problem into four tractable sub-parts. The four sub-problems are inter-dependent, and hence the order in which they are solved is significant. We then present solutions to the four sub-problems.

Subsequently, we undertake an evaluation of the algorithms. Our evaluation shows that we are able to generate topologies of up to 31 points in practical settings. The cost of the generated topology is within about 2% of a lower bound. We also show our search pruning strategies to be effective. Further experience is required in the future in terms of putting our algorithms to practical test.

Scalability much beyond what we have achieved is desirable, but there are two bottlenecks. One is the exhaustive search approach. While this can possibly be addressed by intelligently partitioning the set of points into geographical groups, there is another issue too. We have found that power assignment for topologies involving over about 35 nodes has a high chance being infeasible. The presence of too many radios in the system causes significant inter-link interference. We believe that addressing this would require the consideration of multiple channels, and channel assignment in parallel with the four variables that we have considered. However, we believe that even formulating this problem meaningfully, without unduly complicating it would be very challenging.

We view our work as a significant first attempt at formulating and solving the problem of long distance rural wireless network planning. Apart from the considerations mentioned above, there are several avenues of improvements possible. In this work, we have restricted the towers to placed only at the given village locations. A more generic approach could consider the placement of towers at any point in the plane. In terms of methodology, we believe that the usage of computational geometry techniques can be quite useful in our setting, and is an interesting direction for future work. While we have considered only tree topologies, an important and significant challenge would be the consideration of fault tolerance in the network topology design.

Finally we would like to point out that while our work was primarily driven by long-distance WiFi deployments, much of problem formulation and solution is equally applicable for other similar wireless deployments too (e.g. 802.11a/g, 802.16 mesh networks, or CorDECT).

7. ACKNOWLEDGMENTS

8. REFERENCES

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

[2] 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.

[3] Project Ashwini. `http://www.byrrajufoundation.org/ashwini_overview.htm`.

[4] New wireless networking system brings eye care to thousands in India. `http://www.berkeley.edu/news/media/releases/2006/06/06_telemedicine.sht%ml`, June 2006. U.C.Berkeley News.

[5] Nepal Wireless. `http://www.nepalwireless.net/`.

[6] 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.

[7] Bhaskaran Raman. Digital Gangetic Plains: 802.11-based Low-Cost Networking for Rural Areas, 2001-2004: A Report. http://www.cse.iitk.ac.in/users/braman/dgp.html, 2004.

[8] Jeffrey E. Wieselthier, Gam D. Nguyen, and Anthony Ephremides. On the Construction of Energy-Efficient Broadcast and Multicast Trees in Wireless Networks. In *INFOCOM*, 2000.

[9] Bhaskar Ramamurthi. Broadband Wireless Technology for Rural India. In *XXVIII URSI General Assembly*, Oct 2005.

[10] Kameswari Chebrolu, Bhaskaran Raman, and Sayandeep Sen. Long-Distance 802.11b Links: Performance Measurements and Experience. In *MOBICOM*, 2006. To appear.

[11] 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.

[12] Bhaskaran Raman and Kameswari Chebrolu. Design and Evaluation of a new MAC Protocol for Long-Distance 802.11 Mesh Networks. In *MOBICOM*, Aug/Sep 2005.

[13] Sayandeep Sen. Topology planning for long distance wireless mesh networks. Master's thesis, Indian Institute of Technology Kanpur, 2006.

[14] A Ravindran, Don T. Phillips, and James J. Solberg. *Operation Research Principles and Practice*, chapter 11.2. John Wiley & Sons, 1986.

[15] QSOPT Linear Programming Solver. http://www2.isye.gatech.edu/~wcook/qsopt/.

[16] Rajesh Gandhi. Empirical path loss models for 802.11b links. Master's thesis, Indian Institute of Technology Kanpur, 2003.

[17] Ram Ramanathan and Regina Hain. Topology Control of Multihop Wireless Networks Using Transmit Power Adjustment. In *INFOCOM*, pages 404–413, 2000.

[18] 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.

[19] Zhuochuan Huang, Chien-Chung Shen, C. Srisathapornphat, and C. Jaikaeo. Topology Control for Ad Hoc Networks with Directional Antennas. In *ICCN*, 2002.

[20] S. A. Borbash and E. H. Jennings. Distributed Topology Control Algorithm for Multihop Wireless Networks. In *IJCNN*, 2002.

[21] Nilesh Mishra, Kameswari Chebrolu, Bhaskaran Raman, and Abhinav Pathak. Wake-on-WLAN. In *The 15th Annual Interntional World Wide Web Conference (WWW 2006)*, May 2006.

[22] Shridhar Mubaraq Mishra, John Hwang, Dick Filippini, Tom Du, Reza Moazzami, and Lakshminarayanan Subramanian. Economic Analysis of Networking Technologies for Rural Developing Regions. In *1st Workshop on Internet and Network Economics*, Sep 2005.