14 OPTIMIZATION ISSUES IN DISTRIBUTION NETWORK DESIGN

Geraldo R. Mateus¹ and Zenilton K. G. Patrocínio Jr.¹

¹Department of Computer Science Federal University of Minas Gerais 30123-970 Belo Horizonte, MG, Brazil mateus@dcc. ufmg. br zeniltonOdcc. ufmg.br

Abstract: A distribution network design problem arises in a lower level of an hierarchical modeling approach for telecommunication network planning. Improvements of technologies used to deploy distribution networks have contributed to make distribution network planning more similar to other levels of access network. The major points that differentiate distribution network design problems are its huge dimensions and the several technological options that could be used to connect customers. Major technological trends to deploy distribution networks are discussed here. As an extension of the capacitated network design problem, it is a NP-hard combinatorial optimization problem. The need to install facilities and capacities in discrete levels and the incorporation of addition technology-related cost terms and constraints makes the exact solution of the mixed integer programming model even harder. There are several models and strategies that might be devised for solving those models, we present some of them.

Keywords: Distribution network, telecommunication system, capacitated network design.

14.1 INTRODUCTION

The urban telecommunication system is composed by intricate networks that enable the communication of hundreds of thousands of customers. The hierarchical organization of this network plays a major role, in as much as optimized levels of customers concentration enables the substantial economies of scale of increasing transmission bandwidth. The design of such networks with an hierarchical modeling approach is also a consequence of the fact that people solve complex problems by solving higher level problems first and then solving resulting lower level problems.

Figure 14.1 Spatial demands of a distribution network

A distribution network problem arises in a lower level of the hierarchical organization of the telecommunication system. The problem can be viewed as an extension of a capacitated single commodity network design problem, as illustrated in Figure 14.1. The commodity flow over an arc can be interpreted as the equivalent number of telecommunication channels that links a switching center or a concentrator to the customers served across that arc. The problem assumes demand points (sinks) as individual customer, a corporate customer or customers clustered and connected to terminal boxes, normally installed in electricity poles. The problem also assumes given the location and capacity of concentrators (sources) that can be optical remote units, which are supposed to be connected to the switching centers, distribution boxes or radio stations. We need to install adequate capacities on the arcs to route the required flow to each sink.

A method in which a large problem is broken up into several smaller problems is normally implicit in the literature concerning specific models of telecommunication network design (Gavish, 1982; 1983; 1991; Minoux, 1989). In order to provide a better description of the distribution network context we prefer to follow here the practice of working explicitly with an hierarchical design organization that can have different hierarchical levels depend on the technologies, services and customers (Balakrishnan et al., 1994; Mateus et al., 1994). Figure 14.2 depicts the classical hierarchical structure of an urban telecommunication network. The symbol $\bigcirc \Rightarrow$ means that the element at left is a generic node of the network described at right, and the symbol $\odot \Longrightarrow$ means that the element at left is the *root* node of a *tree* network referred at right. The first level of the figure states that the *urban space* is partitioned in *local areas.* Each local area is served by a *switching center.* The communication among local areas is performed by a *backbone network.* The *switching center or concentrator* is the element of linkage between the *backbone network* and the *local access network.* Remark that a *switching center is* a node representing a local area in the *backbone network* and that it is also the *root* node of the correspondent *local access network.*

Our main focus here is on the second phase of the local access network design process. Figure 14.2 shows that, in order to provide access for the customers assigned to each switching center, we need to perform three levels of network design :

- 1. At the first level, each local area is partitioned in *service sections,* and the *primary network* provides optical access of the service sections to the assigned switching center.
- 2. At the second level, each service section is partitioned in *terminal sections,* and the *secondary network* provides the access of these terminal sections. An *optical remote unit,* a *distribution box* or a *radio station* is the element of linkage between the *primary* and the *secondary* networks.
- 3. At last, the *domestic network* encompasses the *tertiary networks* that assure the links of the customers to assigned terminal boxes.

The distribution network studied here concerns a particular local access network design problem, that in Figure 14.2 is associated with the secondary network. The literature on local access network design problems covers a variety of settings, which raises issues of dimensioning, topological design and routing. Balakrishnan et al. (1991) discuss several local access network design formulations, see also (Magnanti et al., 1993; Bienstock and Gtinluk, 1995). The local access design can be made either according to a Steiner tree (Luna et al., 1987; Balakrishnan et al., 1989) or else be based in an extension of the minimum spanning tree problem (Gavish, 1983; 1991; Hochbaum and Segev, 1989; Mirzaian, 1985). In any case the problem is NP-hard. In the first case we express the real existence of intermediate or transshipment nodes, but we face the difficulty of a NP-hard subproblem, thus looking for approximate solutions of the Steiner problem in graphs (Hakimi, 1971; Aneja, 1980; Wong, 1984; Maculan, 1987; Beasley, 1989; Duin and Volgenant, 1989; Agrawal et al., 1994). In the second case we deal with a restricted modeling approach, but we can take advantage of the fact that a greedy algorithm is able to find optimal solutions of the minimum spanning tree subproblem (Kruskal Jr., 1956).

A complementary problem arising in local access design of computer and communication systems concerns the location of facilities. The facilities may be switching centers, remote units, concentrators, radio stations, distribution or terminal boxes (Minoux, 1989; Gavish, 1991). Normally a facility is the *root* (supply) node of a local access *tree* network, and it is one of many demand nodes of a higher level network, that may be a tree, a *ring* or a *multiconnected* network. We can be inspired by classical methods to solve facility location problems, either in capacitated versions (Sa,

Figure 14.2 Hierarchical structure of an urban telecommunication network

1969; Beasley, 1988; Mateus and Luna, 1992) or in uncapacitated versions (Balakr-

Figure 14.3 Technologies for distribution networks

ishnan et al., 1989). Distribution network planning consists of both locating optical remote units, concentrators or radio stations and selecting adequate cables or wireless channels to route the required flow to each demand point.

14.2 TECHNOLOGY FOR DISTRIBUTION NETWORK

The linkage of the urban communication customers is nowadays performed by fiberoptics for the backbone and for the primary networks. For the secondary (distribution) and the tertiary networks we can have wireless channels, fiber and/or copper cables (see Figure 14.3). For wireline networks, models should accommodate any possible configuration: *fiber-to-the-home (FTTH), fiber-to-the-curb* (FTTC) or traditional copper links. In this case the distribution network is composed by cables with standard capacity levels, normally given by a number of pairs of fiberoptics or copper wires. In the wireless network we need to cover all the urban area, assign frequencies and optimize the power control.

The use of copper-based and/or fiber-based systems is specially interesting for telecommunication operators that already have an installed infrastructure (switching centers, ducts, cables, access equipment, etc.), which can evolve in order to provide other services such as broadband access. Recently, new operators have joined the telecommunications market and usually they are not disposed to build a brand-new entire network and look for less expensive alternatives (such as, infrastructure renting). To these new operators, wireless technologies, either fixed or mobile, seem to represent an interesting way of creating an access/distribution network.

Wireless technologies are in their road to maturity and are supposed to be able to provide broadband services in the short-term. At the mobile side, the planning of *third generation* (3G) networks based on *wideband code division multiple access* (WCDMA) air interface, poses several differences and singularities when compared to the planning of *second generation* (2G) networks. The well known systems are the *universal mobile telecommunications system* (UMTS) or the *code division multiple access* (CDMA). In the case of fixed systems, the *local multipoint distribution service* (LMDS) is one of the most studied technologies.

Type	Upstream Data Rate	Downstream Data Rate	Max Local Loop Reach (Km)	Life Line Voice
Analog Modems	14.4-33.6 Kbps	14.4-33.6 Kbps	N/A	Yes
56 Kbps Modems	33.6 Kbps	56 Kbps	N/A	Yes
Cellular Modems	14.4 Kbps	14.4 Kbps	N/A	N/A
ISDN	128 Kbps	128 Kbps	N/A	Yes
ASDL	176 Kbps 224-640 Kbps	1.54 Mbps 6-8 Mbps	5.4 3.6	Yes Yes
VSDL	640 Kbps $1.6 - 2.3$ Mbps	13 Mbps 52 Mbps	1.4 0.3	Yes Yes
Cable Modems	0-768 Kbps	30 Mbps	N/A	No
LMDS	64 Kbps	34-40 Mbps	5.0	N/A
APON [®] FITH/B	155 Mbps	622 Mbps	N/A	N/A
$EPON^b$ FTTH/B	10-1000 Mbps	10-1000 Mbps	N/A	N/A

Table 14.1 Technologies characteristics (^aATM-based passive optical network; ^bEthernetbased passive optical network)

In the following subsections, each of these technologies are briefly described and its major features are pointed out (see also Table 14.1). The focus here is not on the technical aspects of the systems, but rather on their socio and techno-economic dimensions.

14.2.1 Wired

14.2.1.1 Copper-based. In spite of antiquated technology and slow transmission rates, the *public switched telephone network* (PSTN) system is, on a world basis, the most reliable and inexpensive alternative for voice and low-rate data transmission. Moreover, it is this system through which the current Internet developed via broad access to modems of various speeds. It is also worth to mention that it is often quicker to carry out a variety of services by simply calling (via the PSTN system) rather than going through the procedure of starting a PC and logging onto a higher speed network.

However, the future of PSTN is not bright when it comes to services requiring high bandwidth. It is also doubtful that any compression technologies will be able to reduce the transmission requirements enough in order to guarantee transmission of even basic services over PSTN in the near future. So, in the context of the copper-based telephone network, on should consider other technologies for providing broadband access, such as *integrated services digital network* (ISDN) and the various *digital subscriber line* (xDSL) technologies.

ISDN. Employing the traditional telephone copper infrastructure and allowing for both voice and data transmission, ISDN is now widely available on a world basis.

Demand for this technology was quite high at the end of the last decade. The important factor to consider is its ability to provide faster Internet access as well as a support system for some business and LAN applications. However, ISDN must be defined as a medium range technology and, from the perspective of non-telecom operators, ISDN is rarely a solution to be consider in the development of future systems. Finally, ISDN represents a movement toward higher bandwidth. In the short-term, ISDN will continue being used, but in the future its limitations will likely mean that it will be superseded.

xDSL. The various xDSL technologies increase bandwidth to the subscriber by utilizing the copper wiring already installed in the PSTN infrastructure. xDSL comes in a seemingly endless number of versions that vary in the capacity carried and reached. In addition, while some version on the technology allows for symmetrical traffic others transmit two separate data streams with much more bandwidth devoted to the downstream leg to the customer than returning. One of the most popular versions is the *asymmetric* DSL (ADSL). It is effective because symmetric signals in many pairs within a cable (as occurs in cables coming out of the switching center) significantly limit the data rate and possible lime length. This variant increases the capacity of the copper system by 30 to 200 times when compared to PSTN modems. One of xDSL's drawbacks is that there is an inverse relationship between capacity and range. Versions of xDSL, such as *very high-bit-rate* DSL (VDSL), can operate at speeds over 50 Mbps but have a range of only a few hundred meters. The major advantage for xDSL is that it allows the further use of the copper-based telecom network. A second advantage is that the approach will allow for incremental development. Therefore, xDSL allows the operators of copper-based networks to offer broadband services at reasonable prices. For markets such as the *small office home office* (SOHO) users, xDSL offers advantages in those areas near to switching centers. As one moves farther into less densely settled areas the use of wireless technologies, such as cellular and LMDS, becomes a more realistic alternative.

14.2.1.2 Fiber-based. Fiber to home has always been an attractive option. It has all the benefits of the fiber. It provides a future-proof network in that we do not have to go through the hassles of upgrading from ISDN to ADSL (and then, to xDSL, etc.). It does not have to contend with Electro Magnetic Interference (EMI) problems. No outside plant component implies highest reliability. It does not need electric powering and is immune to lightning and other transients. These properties of the fiber lead to lowest powering and operational costs (such as maintenance, provisioning and facilities planning).

The use of fiberoptics for the backbone and for the primary networks is not new, but only recently its use for the secondary (distribution) and the tertiary networks has started to be explored as a result of several new technologies that are leading to cost reduction such as *Passive Optical Networks* (PON).

Moreover, many protection schemes designed for fiber-based topologies rely on availability of another path between each remote unit and the associated switching

center. So, establishment of local access rings has become more attractive as many operators have started to use fiberoptics.

FTTC. This scheme uses optical fibers to connect the switching center or central office (CO) to the *optical network units* (ONUs). The downstream traffic follows this path, namely, from the core to the CO and further on to the ONUs through fiber. The upstream traffic can also follow the same link. However, the ONUs are connected to the CO through a coaxial cable as well. The purpose of this cable is to send normal analog TV signal and to provide power to the ONUs. Each ONU is further connected in a star topology to a few (10-100) homes through a coaxial cable or a twisted pair. The cost of deploying FTTC must be equivalent to the cost of next generation *digital loop carrier* (DLC) — technology that makes use of digital techniques to bring a wide range of services to users via twisted-pair copper phone lines. It also provides a broadband-ready platform. And because the network is fiber to the serving pole or pedestal, maintenance savings are likely, too. The switched digital services FTTC architecture, as good as it is, fails to meet the needs and cost targets of a total service application. It costs about the same as a DLC system with a copper distribution for new build and rehabilitation projects. But it may not be cost-efficient for scattered demands of second lines and requirements for residential data, such as work at home, high-speed modems and LAN/WAN interfaces in existing neighborhoods.

FTTH/B. The optoelectronic equipment cost is a significant enough fraction of the total cost in the last mile that the future of FTTH/B is going to depend significantly on cost reduction in the optics. For example, one recent evaluation estimates 15% of the cost as lying in CO optoelectronics, 40% in the distribution network and its installation, and 45% in the customer premises optoelectronics and its installation. The same source breaks down the 40% for the infrastructure as follows: 53% in construction, 10% in engineering, 20% in couplers and splitters, 9% in splice closures, and only 8% in the cost of the fiber itself. A lot of optoelectronic cost reduction is already happening. While clever things are happening in the optical component world, equally inventive things are happening in the civil engineering of fiber installation. Machines can now cut narrow grooves in the pavements along city streets to lay in fiber bundles. Small trenching machines install plastic ducts one to two feet underground and remotely controlled directional drilling robots extend the path of the duct under highways. If the shorter ducts do not already contain the fiber, it can be later blown in by compressed air, a technique that allows upgrades and rehabs without further excavation. Thus, the costs, particularly the lifetime costs, of the all-glass solution are comparable to or less than those of any of the copper-based solutions, since the latter include along the right-of-way hot, finite-lifetime electronics with periodically spaced localized backup power sources. Some of the electronics cost is for data compression and decompression, unnecessary with fiber because of its huge offered bandwidth. Cost dominates everything in the last mile to a much greater extent than with the more traditional metro and long-haul situations.

14.2.2 Wireless

Telecom operators have long experience with copper access networks and will use them as long as possible. But, as we have mention before, new operators have recently joined the telecommunications market and usually they are not disposed to build a brand-new entire network and look for less expensive alternatives. To these new operators, wireless technologies (fixed or mobile) provide an interesting alternative for the deployment of an access/distribution network as shown in Figure 14.4.

14.2.2.1 Cellular. The deployment of a cellular system depends greatly upon the cellular network planning. Third generation mobile system for personal communication are becoming a reality faster than expected. Along with the new systems, a new set of services is intended to be offer to their user community. This new service set includes traditional voice and data access, network services, multimedia services and new services yet to be defined. All these services must be supplied with ubiquitous access, high performance and quality standard, must be able to run in a great variety of terminals and competitive price. Then, the network planning must consider radio propagation predictions, geographic and traffic parameters evaluation, radio network design optimization and network resource allocation. The radio station location at minimal cost aiming to cover the area under study. Link capacity is defined as the number of channels available in the wireless link. The capacity is limited by the reverse link rather than the forward link. Reverse link capacity depends on the interference received at the base station. Moreover, the link capacity varies with power control, cell coverage area, traffic load and radio path loss. The power control mechanisms allow each mobile unit to emit the minimum power needed to communicate with the target quality and hence generate the minimum possible interference at all others. So the network planning problem should consider the base station location and power control problems and the support for services based on higher transmission rates.

Figure 14.4 Wireless distribution network

14.2.2.2 LMDS. A fixed wireless cellular system, such as LMDS, is an alternative to the copper-based (xDSL), fiber-based (FTTx) and *hybrid fiber coax* (HFC) systems. Information broadcasting allows the high-speed downstream transmission, while service flexibility is obtained by an individual return channel whose performance is adequate to its end user (typically, residential or small/medium enterprises). LMDS has the advantage of low material and labor costs, and the ability to match construction to user demand avoiding cash-flow problems. Another advantage is the quick ability to install a system and lower maintenance because of less exposed equipment. The limitations for LMDS include the still unresolved constraints of the technology, lack of availability, the provisional nature of equipment, regulatory issues and unproven background when it comes to interactive services.

14.2.3 The cost structure

The available capacities of the distribution network come at different cost. The information about the costs of these capacities is rather accurate; for every particular type of communication link, the structure of the cost can be described by length and capacity dependent cost functions. These functions are structurally different and depend on which technologies are really used. The cost functions for wireline networks are well known, so in this section, we present typical cost functions for other two important types of links: leased lines and microwave links.

14.2.3.1 Leased lines. A new operator, such as a mobile-communication network operator, may rent part of its network from a leased line provider. As illustrative example. Figure 14.5 shows a typical cost structure for this case.

Figure 14.5 illustrates that the typical cost structure for a particular capacity, such as 30, 480, or 1920 channels, i.e., 2, 34, or 140 Mbps *of plesiochronous digital hierarchy* (PDH), is piecewise linear and monotonically increasing with the length of the link. The slope, however, decreases with the length of the link (e.g., it decreases at specified lengths such as l_1 , l_2 , and l_3 in Figure 14.5).

Changing the view from a particular capacity to available capacities on a particular link, as illustrated in Figure 14.6 for distinct digital signal levels of the PDH, it is easy to see that economies of scale are large. It depends on the distance between two endnodes of a link, but as a rule of thumb, a capacity of six to eight times 30 channels is

Figure 14.5 Typical cost structure for leased lines

Figure 14.6 Typical cost structure for leased lines on a link

more expensive than a capacity of 480 channel, and three times 480 channels are more expensive than 1920 channel.

14.2.3.2 Microwave. The cost structure for microwaves is different since the maximum transmission distance through the air is limited. To guarantee a specified quality of the signal, it is necessary to periodically amplify the digital signal. So-called repeaters are needed after every interval of constant distance. For instance, if the maximum distance without amplification of the signal is 50 kilometers, two repeaters are necessary on a link of length 130 kilometers. The necessary amplification of the digital signals dominates the structure of the cost function for microwaves. Figure 14.7 illustrates a typical cost structure of microwaves for a particular capacity. The cost in dependence of the length is a staircase function with equal width intervals of constant cost. The width l_1 is the distance at which a repeater becomes necessary, and $l_i = i \cdot l_1$.

Figure 14.8 illustrates the cost structure of a microwave connection for a particular link. Similar to leased lines, it is a staircase function with considerable economies of scale.

The illustrated capacity and cost structures add significant complexity to the design of a distribution network. In the sense of complexity theory, the problem becomes difficult because of the discrete structure of the available capacities. It is not possible, for instance, to install 30.5 channels. Even if this is a required value, the network designer must choose between two discrete levels, e.g., 30 and 60 channels. The illustrated economies of scale cause further difficulties since it is not clear at which point it is appropriate to choose a 480 channel link instead of several 30 channel links. Of course, as shown in Figure 14.6, there exists a break-even point from which on it is

Figure 14.7 Typical cost structure for microwave connections

Figure 14.8 Typical cost structure for microwaves on a link

cheaper to use the higher capacity link, but it might pay to choose this higher capacity even below the break-even point because of the additional capacity. Using the larger capacity of 480 channels instead of six to eight times 30 channels, additional 240- 300 channels are available at relatively small extra cost. Because of this additional capacity on one link it might be possible to decrease capacities on other links, and thus the overall network cost might decrease.

New local loop architectures (such as HFC, FTTC, FTTH/B, and LMDS) tend to have reduced economies of scale because of the new role played by variable costs associated with electronics and Optronics (Pupillo and Conte, 1998). This means that it is very unlikely that a single uniform architecture will prevail in the way that twisted copper pair dominated the telephone network and coax cable dominated cable television in the past. Rather, the local loop of the future (and specially the distribution network) is more likely to be characterized by heterogeneous technologies. As a result, the inevitable competition will lead the successful operators to adopt most, if not all, of the forthcoming access/distribution technologies. Moreover, the competition in the local loop will increase due to a low probability of effective entry preemption. On the one hand, the decrease in economies of scale facilitates entry; on the other hand, the increase in economies of scope incentives to offer a new range of services through which entrants can differentiate themselves from dominant operators. Economies of scope mean that customers will be increasingly able to satisfy multimedia needs through any of a number of suppliers from formerly distinct industries (Pupillo and Conte, 1998).

14.3 OPTIMIZATION MODELS

In regard to network design problems, prior to the 1980s, location-allocation models were natural choice when designing telecommunication and computer networks (Boorstyn and Frank, 1977; Chou et al., 1978; Gerla and Kleinrock, 1977; McGregor and Shen, 1977). These models sought to determine the location of concentrators (or network access facilities) between a switching center and terminal units having known demands. The resulting network constituted an hierarchy of nodes between terminals and a switching center (or a gateway node) having a binary-tree-like structure such that whenever the flow passed to an upper level hierarchy, the flow would be combined or consolidated. The major cost savings from this design accrued from the fact that direct network connections from demand points to a switching center were obviated, and the installation permitted a more efficient use of link capacity from a concentrator to a switching center.

Given a capacitated network and point-to-point traffic demand, the objective function of more recently considered capacity expansion problems is to add capacity to the edges, in integral multiples of various modularities, and route traffic, so that the overall cost is minimized (Bienstock, 1993a;b; Bienstock and Günlük, 1996; Brockmüller et al., 1996; Cook, 1990; Dahl et al., 1995; Gunluk, 1996; Mateus and Franqueira, 2000; Shaw, 1993; 1995). The underlying network has a directed tree rooted at a switching center or central office (CO). The arcs correspond to feeder sections along which signals can be sent, and the nodes represent geographic areas where demand for signals can occur. Each node, other than the CO, is entered by exactly one arc, and the CO is not entered by any arc. We are also given the demand at each node for equivalent DSO signals *(Digital Signal Level* 0), which transmit at the rate of 64 Kbps. Furthermore, certain nodes are designated as remote terminals, where it is possible to install multiplexing equipment to convert DS3 signals *(Digital Signal Level 3)* to 28 DSl signals *(Digital Signal Level* 7), and DSl signals to 24 DSO signals. For each arc, we are given its length, and the number of spare spaces available in the corresponding feeder section, that is, the number of cables that can be installed along the feeder section without undertaking the costly operation of constructing a new channel. For different types of signals, the space used by fiber cables is much less than that for copper cables. Hence, for large demand nodes, when we consider a capacity expansion to meet with such an increase in demand, the use of fiber optic cables is a more appropriate alternative. Naturally, there is a tradeoff between using cheaper copper cable installations or more expensive fiber optic cables, depending on the required feeder section capacity and the demand of customers. The objective function is to minimize the total cost of the copper and fiber cables and the multiplexers that must be installed in order to satisfy the demand.

As mention before, distribution networks are closely related to other levels of hierarchical organization of the telecommunication system especially to local access networks. Balakrishnan et al. (1991) demonstrated how to cast the general layered network framework for local access network planning problem as a fixed charge network design model. In order to do that the following assumption concerning the cost structure has to be made. The network design model ignores joint costs between various transmission media, and assumes a one-to-one correspondence between transmission rates and media, i.e., the planner preselects a preferred transmission medium for each transmission rate (Balakrishnan et al., 1991). Installation/expansion costs are also assumed to be piecewise linear, consisting of fixed and variable components.

In spite of that, a network design model can incorporate all features of the general layered network framework. In particular, it can handle general topologies, multiple service types, sectional and point-to-point media types, economies of scale in processor and transmission cost functions, and existing transmission and processing capacities. It also permits backfeed and bifurcated routing. Comprehensive surveys on the applications of network design models and their resolution by mathematical programming techniques can be found in (Magnanti and Wong, 1984; Minoux, 1989). If the cost functions are piecewise linear and concave, and if the network does not

contain any existing capacities, the model reduces to an uncapacitated network design problem.

In general words, an uncapacitated network design problem is given by:

minimize *Design cost* = *Fixed cost* + *Variable cost*

subject to

Flow balance constraints Linkage constraints

Fixed costs are related to installation/expansion of the facilities associated with arcs and/or nodes, while variable (or flow) costs are associated with routing customers demands through network topology. It is clear that, in order to minimize fixed costs, one should look for a solution that installs the minimum amount of facilities as possible, but this will probably rise variable costs as routing becomes harder. The same reasoning can be applied backwards. In order to minimize variable costs, more facilities have to be installed (to make routing easier and cheaper), but fixed costs will get higher.

So, the objective function above tries to minimize the summation of fixed costs and variable (or flow) costs. Flow balance constraints assure that installed (node/arc) facilities will meet all customers demands. Finally, linkage constraints establish a connection between decision and flow variables, i.e., if there is a flow unit passing through a facility (associated with an arc and/or a node), the existence of that facility should indicate by the corresponding decision variable, and vice-versa.

An uncapacitated network design solution that conserves flow at each node, and satisfies all demands at minimum total fixed plus flow costs corresponds to an optimum distribution network plan. It is possible to enrich this network design model in several ways. For instance, we can model economies of scale in processor and transmission costs if these economies can be approximated by piecewise linear concave cost functions.

Moreover, additional constraints can be used to set a limitation on the selected technology, e.g., one could limit the number of splices in a copper-based distribution network as described in (Mateus et al., 2000; Mateus and Franqueira, 2000); or the existence of overlapping cells in a LMDS system could be forbidden as in (Carlson and Authie, 2001). Similarly, additional terms (cost values) could also be incorporated to the objective function. Doing so, one could represent technology-related costs, e.g., splicing costs associated with copper-based infrastructure (Mateus et al., 2000), modularity (Mateus and Franqueira, 2000), or interference costs and the cost to connect the located base stations to the fixed networks related to wireless scenarios (Mazzini etal.,2003).

Existing resources and non-concave cost functions introduce arc capacities. Capacitated network design problem is much harder to solve than the uncapacitated version. One possible solution strategy is to dualize arc capacity constraints (i.e., multiply capacity constraints by Lagrange multipliers, and add these to the objective function). The resulting subproblem is an uncapacitated network design problem which is easier to solve. By iteratively modifying the Lagrange multipliers, e.g., using subgradient optimization, we can possibly originated good heuristic solutions and lower bounds for the original capacitated problem.

However, experiments with this approach for other related models (such as capacitated plant location and capacitated minimum spanning tree) suggest that the addition of arc capacities significantly increases the gaps between the upper and lower bounds. So, one should expect distribution network design problems with existing capacities and non-concave cost functions to be computationally more difficult to solve.

To cope with that, one could minimize the computational burden by simplifying the capacitated network design model. This can be done by ignoring some model aspects or by focusing on more relevant ones. For instance, if the location sites of node facilities (such as remote units, distribution boxes and radio stations) are predefined (or ignored), only arc facilities have to be selected along with the routes of customers traffic. In this case, Steiner tree problem (Luna et al., 1987; Balakrishnan et al., 1989; Agrawal et al., 1994; Mateus and Franqueira, 2000) is obtained, if there is intermediate or transshipment nodes, i.e., network elements which do not represent customers premises (e.g. concentrators). However, if every node has an associated demand value, which means that they all should be connect to some remote unit, distribution box or radio station, the generated problem is a minimum spanning tree problem (Gavish, 1983; 1991; Hochbaum and Segev, 1989; Mirzaian, 1985).

Considering the following aspects: node facilities, arc facilities and traffic routing, different simplified models can be obtained depending on which aspects are predefined or ignored. Table 14.2 summarizes some models that will be generated by this way. Although these problems are easier than the original one, many of them are still NPhard. So, one might not except to obtain an exact procedure to some of them, and approximated solutions should be generated using good heuristics.

Many works that address those problems could be found in the literature, and it is not feasible to describe all of them here. In order to illustrated how those problems could be solved, some of those algorithms are revised in the next section.

14.4 ALGORITHMS

Because the underlying graph to be designed is a directed rooted spanning tree, the capacitated network model with point-to-point traffic demand described at the beginning of Section 14.3 can be formulated as a well structured integer linear program. Cook (1990) formulated and solved this problem using an algorithm based on the basis reduction technique of Lovász and Scarf (1990). Bienstock (1993a) formulated several subproblems related with Cook's problem, and developed various exact and e-approximation algorithms.

After the success of cutting plane algorithms based on the facial structure of the convex hull of feasible solutions (so called branch-and-cut methods), recent research on telecommunication network design focuses on exploiting polyhedral descriptions to obtain tighter reformulations of the problem (Bienstock and Giinluk, 1996; Brockmiiller et al., 1996; Dahl et al., 1995; Grotschel and Monma, 1990; Grotschel et al., 1992a;b; 1994; Giinluk, 1996; Stoer and Dahl, 1994). The capacitated network design problem described before generates a series of research problems related to expanding capacity on the communication network. Bienstock and Günlük (1996), Günlük (1996), and Dahl et al. (1995) consider installing more capacity on the edges and routing the traf-

Table 14.2 Distribution network optimization models

fic on the capacitated network, given a set of traffic demands between certain nodes of the network.

A second category of related problems is called the general network design problem, and is defined as follows. Given an undirected or directed graph $G = (V, E)$ where V and *E* are respectively the sets of vertices and edges of the graph G, and given the cost c_e , for corresponding each edge $e \in E$ (or for each arc in the directed case), the network design problem is to find a minimum cost subset $E' \subseteq E$ that meets some design criteria. Examples of the design criteria are Steiner tree and node- and edgeconnectivity or reliability restrictions. In the node- and edge-connectivity problem, each node $s \in V$ has an associated nonnegative integer r_s , known as the type of s. We say that the graph *G* to be designed satisfies the node connectivity requirements, if, for each pair $s, t \in V$ of distinct nodes, G contains at least

$$
r(s,t)=\min\{r_s,r_t\}
$$

node-disjoint $[s,t]$ -paths. Similarly, we say that G satisfies the edge connectivity requirements, if, for each pair s, t of distinct nodes, G contains at least $r(s,t)$ edgedisjoint $[s,t]$ -paths. These conditions ensure that some communication path between *s* and *t* will survive a prespecified level of edge (or node) failures. Cutting plane algorithms are developed using the information on the facial structure of the convex hull of feasible solutions (Grotschel and Monma, 1990; Grotschel et al., 1992a;b; 1994). For $r_s \in \{0, 1, 2\}$, $\forall s \in V$, this problem is applied to Intra-LATA fiber network design problems faced by the regional Bell Operating Companies (BOC) and the reported computations show that their cutting plane algorithm with preprocessing solves problems having up to 116 nodes and 173 edges (Grötschel et al., 1991).

As a new alternative standardized network architecture, Synchronous Optical Network (SONET) technology with high-speed add/drop multiplexing technology is a recent innovation for protecting cable cuts or hub failures in the network. The Self-Healing Ring (SHR) structure has a ring structure such that any distinct pair of nodes *s, t* on this ring have two node-disjoint paths. For SONET ring architectures, the initial design problem is to determine the capacity of the ring and the clustering of the central offices into each ring (Wasem, 1991; Laguna, 1992). Another related problem is concerned with the load balancing issue (Shulman et al., 1991; Myung et al., 1997). This problem seeks to minimize the maximum aggregated link flow resulting from bidirectional link flows between all pairs of demand nodes on a ring.

The general location-allocation problem (LAP) is a class of mathematical programming problems that seeks the least cost method for simultaneously locating a set of service facilities and satisfying the demands of a given set of customers. A variant of this class of location-allocation problems is the p -median problem which is concerned with the location of *p* new facilities, called medians, on a network, in order to minimize the sum of weighted distances from each node to its nearest new facility (Handler and Mirchandani, 1979). If $p > 2$, this problem can be viewed as a locationallocation problem, because the location of the new facilities control the allocation of their service in order to best satisfy the demands. Hakimi (1964) has shown that for the p -median problem defined on a network, a set of optimal locations will always coincide with the vertices. Cavalier and Sherali (1986) have designed exact algorithms to solve the stochastic p -median problem on a chain graph and the 2-median problem on a tree graph, when the demand density functions are assumed to be piecewise uniform. Sherali and Nordai (1988) have developed certain localization results and algorithms for solving the capacitated p -median problem on a chain graph and the 2-median problem on a tree graph. Sherali (1991) later analyzed capacitated locationallocation problems on chain and tree graphs where the locational decisions need to be made sequentially, and Sherali and Rizzo (1991) addressed the case of unbalanced capacitated p-median problems having probabilistic link demands.

The fixed-charge or discrete location-allocation problem is a variant of the class of location-allocation problems that restricts the potential locations of the supply centers that might be constructed to certain preselected sites (Davis and Ray, 1969). Rardin and Choe (1979); Rardin (1982); Rardin and Wolsey (1993) have developed enhanced formulations and valid inequalities for uncapacitated fixed-charge network flow problems. Rardin and Wolsey (1993) introduced the class of dicut valid inequalities for an aggregated formulation, generating a tighter linear programming relaxation equivalent to the disaggregated formulation presented in Rardin and Choe (1979). Note that Steiner tree problems and uncapacitated facility location problems (UFL) are special cases of the class of uncapacitated fixed-charge problems. For the class of fixedcharge capacitated network flow problems, Herrmann et al. (1996) have generalized the dual ascent algorithm of Erlenkotter (1978). The related capacitated facility location problem (CFL) has been studied by Padberg et al. (1985); Leung and Magnanti (1989); Aardal et al. (1995); Aardal (1998), and several valid inequalities have been derived based on the facial structure of the problem. This problem seeks to construct any subset of a given number of fixed capacity facilities, so as to satisfy the specified demand at a minimal total construction plus distribution cost. Leung and Magnanti (1989) consider the case where at most one such facility having a capacity of *U* units can be constructed at each potential site, with no limit on the total number. Marathe (1992) considers a problem that permits any number of such facilities to be constructed at each site (hence, the net capacity is a multiple of U), with the total number of facilities located being *p* and presents a dynamic programming algorithm on a chain graph, and then solves a Lagrangian relaxation of a formulation on a general network via a conjugate subgradient optimization procedure.

The test problems considered in Marathe (1992) are solved exactly by a branchand-bound algorithm that includes Lagrangian relaxation and Benders' decomposition strategies for computing lower bounds. The size of problems solved ranges from (number of network nodes, p) = (7, 5) to (20, 10) within about 1 CPU minute on an IBM 3090-300E supercomputer. For the capacitated facility location problem, Jacobsen (1983) has empirically compared various heuristics using a set of test problems derived from data published in Kuehn and Hamburger (1963). The size of problems solved involves 50 customers and 25 locations. More recently, Aardal (1998) has solved problems ranging up to 100 customers and 75 facilities using various strategies for generating valid inequalities. Although the initial linear programming (LP) relaxation generates tight lower bounds within about 1% of optimality, the effort to solve such problems to optimality ranges from 5.9 to 32.5 CPU hours on a SUN Sparc ELC computer. The LP relaxations of these problems produces significantly larger gaps, ranging up to about 70% of optimality, hence rendering them considerably more challenging to solve. For the fixed-charge capacitated network flow problems considered in Herrmann et al. (1996), test problems having the number of arcs ranging from 20 to 60 and the number of commodities ranging from 10 to 35 have been solved on a SUN Spark IPX Workstation. Problems having less than 30 arcs were solved to optimality via a branch-and-bound method and were used to evaluate the proposed dual ascent approach. The results indicated that the dual ascent procedure produced lower bounds lying typically within 12% of the optimal solution value.

14.5 CONCLUSIONS

Decision support systems for engineering design in telecommunication require an hierarchical organization of optimization models. The rising complexity of integrated computer and communication systems imposes for the emerging problems a coherent divide-and-conquer solving strategy. Optimal and heuristic algorithms must be adapted to the different classes of models, with consistent transference of information and decision among the models. We have showed a typical hierarchy of decision levels, and we have focused here the lower levels of local access network design.

Improvements of technologies used to deploy distribution networks have contributed to make distribution network planning more and more similar to other levels of access network, e.g., feeder network design. The major points that differentiate distribution network design problems are its huge dimensions (maybe involving hundreds of customers) and the several technological options that could be used to connect customers

to the switching centers, distribution boxes or radio stations (maybe originating a distinct model for each distinct technology).

Capacitated network design problem is a referential model to design the lower levels of the local access networks. This makes distribution network design problem an extremely difficult one. As an extension of the capacitated network design problem, it is a NP-hard combinatorial optimization problem. The need to install facilities and capacities in discrete levels and the incorporation of additional technology-related cost terms and constraints makes the exact solution of the mixed integer programming model even harder. The computational results show that, from the theoretical point of view, it is very difficult to prove the optimality of the solutions. Heuristics is perhaps the only practical solution methodology for such a complex problem faced by telecommunications network providers. On the other hand, the computational results shows fast solution times, thus suggesting that heuristics can solve the large scale problems met in practice.

For this class of problems, duality gaps are substantial; however, when a more accurate lower bound procedure is used, the bounds found by many researchers were shown to be good. There are several strategies that might be devised for those models, but most of the approaches are not so easy to implement. Finally, researchers have long recognized that, for capacitated problems, optimal solution methods that use only the standard flow-based problem formulation are ineffective. The literature suggests that strong formulations, based on results from polyhedral combinatorics, appear to provide the best means to solve these problems.

Bibliography

- K. Aardal. Capacitated facility location: Separation algorithms and computational experience. *Mathematical Programming,* 81:149-175, 1998.
- K. Aardal, Y. Pochet, and L. A. Wolsey. Capacitated facility location: Valid inequalities and facets. *Mathematics of Operations Research,* 20:562-582, 1995.
- A. Agrawal, P.N. Klein, and R. Ravi. When trees collide: An approximation algorithm for the generalized steiner problem on networks. Technical Report CS-90-32, Department of Computer Science, Brown University, Providence, Rhode Island 02912, 1994.
- Y.P. Aneja. An integer linear programming approach to Steiner problem in graphs. *Networks,* 10:167-178, 1980.
- A. Balakrishnan, T.L. Magnanti, and P. Mirchandani. Modeling and heuristic worstcase performance analysis of two-level network design problem. *Management Science,* 40:846-867, 1994.
- A. Balakrishnan, T.L. Magnanti, A. Shulman, and R.T. Wong. Models for planning capacity expansion in local access telecommunication networks. *Annals of Operations Research,* 33:239-284, 1991.
- A. Balakrishnan, T.L. Magnanti, and R.T. Wong. A dual-ascent procedure for largescale uncapacitated network design. *Operations Research,* 37:716-740, 1989.
- J.E. Beasley. An algorithm for solving large capacitated warehouse location problems. *Journal of the Operational Research Society,* 33:314-325, 1988.
- J.E. Beasley. An sst-based algorithm for the Steiner problem in graphs. *Networks,* 19: 1-16, 1989.
- D. Bienstock. Computational experience with an effective heuristic for some capacity expansion problems in local access networks. *Telecommunication Systems,* 1:379- 400, 1993a.

- D. Bienstock. A lot-sizing problem on trees, related to network design. *Mathematics of Operations Research,* 18:402-422, 1993b.
- D. Bienstock and O. Giinluk. Computational experience with a difficult mixed-integer multicommodity flow problem. *Mathematical Programming,* 68:213-237, 1995.
- D. Bienstock and O. Günlük. Capacitated network design polyhedral structure and computation. *INFORMS Journal on Computing,* 8:243-259, 1996.
- R.R. Boorstyn and H. Frank. Large-scale network topological optimization. *IEEE Transactions Communications,* 25(l):29-47, 1977.
- B. Brockmiiller, O. Giinluk, and L. A. Wolsey. Designing private line networks polyhedral analysis and computation. Technical report, CORE Discussion Paper, 1996.
- C.M.F. Carlson and Gerard Authie. Optimized design of LMDS cells: site location, equipment sizing and users homing. In *Proceeding 3rd IEEE International Conference on Mobile and Wireless Communications Networks (MWCN'2001),* pages 111-118, Recife, Brazil, 2001.
- T.M. Cavalier and H.D. Sherali. Network location problems with continuous link demands: p-medians on a chain and 2-medians on a tree. *European Journal of Operational Research,* 23:246-255, 1986.
- W. Chou, F. Ferrante, and L. Gerke. An algorithm for optimally locating network access facilities. In *Proceedings of ICC 78,* pages 24.5.1-24.5.8, 1978.
- W. Cook. Integer programming solution for capacity expansion of the local access network. Technical Report TM-ARH-017914, Bell Communications Research, 1990.
- G. Dahl, A. Martin, and M. Stoer. Routing through virtual paths in layered telecommunication networks. Technical report, Telenor Research and Development, 1995.
- RS. Davis and T.L. Ray. A branch and bound algorithm for the capacitated facilities location problem. *Naval Research Logistic Quarterly,* 16(3):331-334, 1969.
- C.W. Duin and A. Volgenant. Reduction tests for the Steiner problem in graphs. *Networks,* 19:549-567, 1989.
- D. Erlenkotter. A dual-based procedure for uncapacitated facility location. *Operations Research,* 26(6):992-1009, 1978.
- B. Gavish. Topological design of centralized computer networks formulations and algorithms. *Networks,* 12:355-377, 1982.
- B. Gavish. Formulations and algorithms for the capacitated minimal directed tree. *Journal of the ACM,* 30:118-132, 1983.
- B. Gavish. Topological design of telecommunication networks Local access design methods. *Annals of Operations Research,* 33:17-71, 1991.
- M. Gerla and L. Kleinrock. On the topological design of distributed computer networks. *IEEE Transactions on Communications,* 25(l):48-60, 1977.
- M. Grötschel and C. Monma. Integer polyhedra arising from certain network design problems with connectivity constraints. *SIAM Journal on Discrete Mathematics*, 3 (4):502-523, 1990.
- M. Grötschel, C. Monma, and M. Stoer. Polyhedral approaches to network survivability. In *DIMACS Series in Discete Mathematics and Theoretical Computer Science,* volume 5. American Mathematical Society, Providence, Rhode Island, 1991.
- M. Grötschel, C. Monma, and M. Stoer. Computational results with a cutting plane algorithm for designing communication networks with low-connectivity constraints. *Operations Research,* 40(2):309-330, 1992a.
- M. Grötschel, C. Monma, and M. Stoer. Facets for polyhedra arising in the design of communication networks with low-connectivity constraints. *SIAM Journal on Optimization,* 2(3):474-504, 1992b.
- M. Grötschel, C. Monma, and M. Stoer. Design of survivable networks. In M. O. Ball, T. L. Magnanti, C. L. Monma, and G. L. Nemhauser, editors. *Handbooks in Operations Research and Management Science 7'.Network Models.* North-Holland, 1994.
- O. Giinliik. A branch-and-cut algorithm for capacitated network design problem. Technical report. School of Operations Research and Industrial Engineering, Cornell University, 1996.
- S.L. Hakimi. Optimal locations of switching centers and the absolute centers and median of a graph. *Operations Research,* 12:450-459, 1964.
- S.L. Hakimi. Steiner's problem in graphs and its implications. *Networks,* 1:113-133, 1971.
- G.Y. Handler and P.B. Mirchandani. *Location on Networks Theory and Algorithms.* The MIT Press, Cambridge, Massachusetts, 1979.
- J.W. Herrmann, G. loannou, I. Minis, and J. M. Proth. Fixed-charge capacitated network design problem. *European Journal of Operational Research,* 95:476-490, 1996.
- D.S. Hochbaum and A. Segev. Analysis of a flow problem with fixed charges. *Networks,* 19:291-312, 1989.
- S.K. Jacobsen. Heuristics for the capacitated plant location model. *European Journal of Operational Research,* 12:253-261, 1983.
- J.B. Kruskal Jr. On the shortest spanning tree of a graph and the travelling salesman problem. *Proc. Amen Math. Society,* 7:48-50, 1956.

- A. A. Kuehn and M. J. Hamburger. A heuristic program for locating warehouses. *Management Science,* 9(4):643-666, 1963.
- M. Laguna. Optimal design of sonet rings for interoffice telecommunication. Technical report, Graduate School of Business and Administration, University of Colorado, 1992.
- J.M.Y. Leung and T.L. Magnanti. Valid inequalities and facets of the capacitated plant location problem. *Mathematical Programming, 44:211-291,* 1989.
- L. Lovasz and H. Scarf. The generalized basis reduction algorithm. Technical Report No. 946, Cowles Foundation Discussion Paper, Yale University, 1990.
- H.P.L. Luna, N. Ziviani, and R.M.B. Cabral. The telephonic switching centre network problem: Formalization and computational experience. *Discrete Applied Mathematics,* 18:199-210, 1987.
- N. Maculan. The Steiner problem in graphs. *Annals of Discrete Mathematics,* 31: 185-212, 1987.
- T.L. Magnanti, P. Mirchandani, and R. Vachani. The convex hull of two core capacitated network design problems. *Mathematical Programming,* 60:233-250, 1993.
- T.L. Magnanti and R.T. Wong. Network design and transportation planning: Models and algorithms. *Transactions Science,* 18(1): 1-55, 1984.
- V. Marathe. A discrete equal-capacity p-median problem. Master's thesis, Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1992.
- G.R. Mateus, F.R.B. Cruz, and H.P.L. Luna. Algorithm for hierarchical network design. *Location Science,* 2:149-164, 1994.
- G.R. Mateus and R.V.L. Franqueira. Model and heuristic for a generalized access network design problem. *Telecommunication Systems,* 15:257-271, 2000.
- G.R. Mateus and H.P.L. Luna. Decentralized decision-making and capacitated facility location. *The Annals of Regional Science, 26:361-377,* 1992.
- G.R. Mateus, H.P.L. Luna, and A.B. Sirihal. Heuristics for distribution network design in telecommunication. *Journal of Heuristics,* 6:131-148, 2000.
- F.F. Mazzini, G.R. Mateus, and J.M. Smith. Lagrangean based methods for solving large-scale cellular network design problems. *Wireless Networks,* 9:659-672, 2003.
- P.V. McGregor and D. Shen. Network design: an algorithm for the access facility location problem. *IEEE Transactions on Communications,* 25(l):61-73, 1977.
- M. Minoux. Network synthesis and optimum network design problems: Models, solution methods and applications. *Networks,* 19:313-360, 1989.
- A. Mirzaian. Lagrangian relaxation for the star-star concentrator location problem: Approximation algorithm and bounds. *Networks,* 15:1-20, 1985.
- Y. Myung, H. Kim,, and D. Tcha. Optimal load balancing on sonet bidirectional rings. *Operations Research,* 45(1):148-152, 1997.
- M.W. Padberg, T.J. Van Roy, and L.A. Wolsey. Valid linear inequalities for fixed charge problems. *Operations Research*, 33:842-861, 1985.
- L. Pupillo and A. Conte. The economics of local loop architecture for multimedia services. *Information Economics and Policy,* 10:107-126, 1998.
- R.L. Rardin. Tight relaxations of fixed charge network flow problems. Technical report. School of Industrial Engineering, Purdue University, West Lafayette, Indiana, 1982.
- R.L. Rardin and U. Choe. Tighter relaxations of fixed charge network flow problems. Technical Report J-79-18, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, 1979.
- R.L. Rardin and L.A. Wolsey. Valid inequalities and projecting the multicommodity extended formulation for uncapacitated fixed charge network flow problems. *European Journal of Operational Research,* 71:95-109, 1993.
- G. Sa. Branch-and-bound and approximate solutions to the capacitated plant location problem. *Operations Research,* 17:1005-1016, 1969.
- D.X. Shaw. Limited column generation technique for several telecommunication network design problems. Technical report. School of Industrial Engineering, Purdue University, West Lafayette, Indiana, 1993.
- D.X. Shaw. Local access network design and its extensions. Technical report. School of Industrial Engineering, Purdue University, West Lafayette, Indiana, 1995.
- H.D. Sherali. Capacitated, balanced, sequential location-allocation problems on chains and trees. *Mathematical Programming,* pages 381-396, 1991,
- H.D. Sherali and F.L. Nordai. A capacitated balanced 2-median problem on a tree network with a continuum of link demands. *Transportation Science,* 22(l):70-73, 1988.
- H.D. Sherali and T.P. Rizzo. Unbalanced, capacitated p-median problems on a chain graph with a continuum of link demands. *Networks,* 21:133-163, 1991.
- A. Shulman, R. Vachani, J. Ward, and P. Kubat. Multicommodity flows in ring networks. Technical report, GTE Laboratories Inc., 1991.
- M. Stoer and G. Dahl. A polyhedral approach to multicommodity survivable network design. *Numerische Mathematik,* 68:149-167, 1994.

- OJ. Wasem. An algorithm for designing rings for survivable fiber network. *IEEE Transactions on Reliability,* 40:428-432, 1991.
- R.T. Wong. A dual ascent algorithm for the steiner problem in directed graphs. *Mathematical Programming,* 28:271-287, 1984.